

Contributions to Political Science

Rafal Olszowski

Collective Intelligence in Open Policymaking

 Springer

Contributions to Political Science

The series Contributions to Political Science contains publications in all areas of political science, such as public policy and administration, political economy, comparative politics, European politics and European integration, electoral systems and voting behavior, international relations and others. Publications are primarily monographs and multiple author works containing new research results, but conference and congress reports are also considered. The series covers both theoretical and empirical aspects and is addressed to researchers and policy makers. All titles in this series are peer-reviewed. This book series is indexed in Scopus.

Rafał Olszowski

Collective Intelligence in Open Policymaking



Springer

Rafał Olszowski 
Faculty of Humanities
AGH University of Krakow
Kraków, Poland

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2024

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Paper in this product is recyclable.

*For Natalia, who has walked every step of
this long journey with me.*

Preface

Human intelligence is unique. It appears that we are the most intelligent species living on Earth. To grasp the essence of our existence, it is crucial to understand what intelligence actually is, encompassing everything from individual minds to large collectives. This represents one of the most significant scientific challenges humanity faces. The question “What are we,” as well as “What are we doing together,” is posed by Hannah Arendt in the prologue to her famous book *The Human Condition*, where she regards civic activity in the public forum as one of humanity’s most significant collective duties. The inquiry “What are we, and what are we doing together?” echoes an existential curiosity about human nature, origin, and destiny, as well as about the collectives in which people unite to pursue common goals.

Individuals do not operate in a vacuum. Their identity is shaped by a network of connections with their surroundings, family, communities, interest groups, nations, and civilizations. As Edmund Burke posited, human collectivity is our “second nature.” Although it has an artificial character (being a theoretical construct), it simultaneously embodies and personifies our primal nature. Our social relationships are of paramount importance to understanding who humans truly are. This “second nature” results from human activity within collectives, which is associated with principles assimilated through cooperation with others and transmitted with customs and prejudices, existing social relations, and commitments. Since, as Aristotle believed, man is “by nature a social animal,” and the collective has a crucial impact on his existence.

In this book, we will explore collective intelligence (CI) and its crucial role in shaping public policy. This work is not a mere academic exercise but a deep dive into the essence of how collective intelligence can be harnessed to foster more inclusive and effective policymaking processes. We challenge the traditional dichotomy between the individual and the collective, often misrepresented and maligned by radical ideologies and totalitarian regimes that seek to subsume individual identity within a monolithic collective. Contrary to this, I contend that true collective intelligence emerges not from the erasure of individuality but from the synergistic interaction of independent minds.

Drawing on interdisciplinary insights from cognitive science, social sciences, computer science, and political theory, we will explore how humans, inherently social beings, thrive in complex networks of relationships that extend from the personal to the global scale. Human beings belong to numerous social groups, ranging from small friend circles to city communities or citizens of a country. The inherent need for belongingness drives us to seek acceptance and connection with others. Whether it is family, co-workers, interest groups, or political associations, we desire to be part of something greater than ourselves. As we interact, our brains sync, creating a beautiful dance of connection.

This book argues that understanding the mechanisms of collective intelligence is tantamount to understanding the very nature of humanity itself, from the operation of a single mind through the dynamics of communities to the workings of vast social structures. From my perspective, our inclination toward collaboration and group belonging is not just a cultural phenomenon but an integral part of our biological and metaphysical constitution. Just as cognitive scientists and neurobiologists inquire about the nature of an individual mind and its functions, attempting to deconstruct the mechanisms of individual intelligence through reverse engineering, so too do social scientists by observing human behavior within collectives and the factors that facilitate successful cooperation. In matters pertaining to the public sphere and policymaking, the foundation of collaboration is a debate. Again, let us heed Arendt: “Men in the plural (...) insofar as they live and move and act in this world, can experience meaningfulness only because they can talk with and make sense to each other and themselves.” Arendt perceives the realm of this mutual sensemaking, or the “public space,” as a world we all share and distinct from our private domains, imparting a dimension to our existence that transcends ephemeral and fleeting endeavors. People enter the public domain, aspiring for their shared experiences and creations to outlive mortal existence.

Collective intelligence is an enduring phenomenon interwoven with human civilization’s development. However, the introduction of the Internet and the subsequent rise of web-enabled collectives mark a significant shift in the contemporary landscape of CI. While the fundamental mechanisms that drive CI remain constant, their application and utilization have witnessed transformative changes in the digital era. The newfound ability to harness collective intelligence on a global scale, transcending geographical and temporal constraints, has opened up new prospects for online collaborative problem-solving and innovation. Open policymaking is a phenomenon that is part of this transformation, offering hope for improved public policy formulation in the guise of diverse collectives and communities that shape public policies. Although the idea of community involvement in the co-creation of public policies was proposed by Elinor Ostrom long before the internet age, the expectancy for the transition from a model of the citizen as a *customer* to the expected *citizen as a partner* model has accelerated significantly in recent years. Public institutions and grassroots NGOs alike are undertaking efforts to implement open policymaking through digital technologies. The most notable successes are observed at the local community level, where private individuals often possess more relevant, practical knowledge than officials—which is particularly visible in participatory budgeting

projects. Conversely, at the national level, adopting open policymaking is more a postulate than actual practice.

Why do we need different policymaking procedures than we used to have? The world is experiencing swift and significant changes. These changes manifest in financial crises, the spread of armed conflicts, and the recent global pandemic. We are also observing a transformation from a unipolar to a multipolar world, a crisis of democracy giving way to authoritarian political models, and controversies around energy crisis and climate policy. These shifts align with macro-trends that shape the development of our technical civilization: well-known forces like globalization and digitization are linked to what Scott Galloway has termed “The Great Dispersion.” This trend is most apparent in the digital economy, where companies founded in the last two decades are replacing traditional suppliers with innovative work models. Amazon dispersed retail, Netflix dispersed DVDs and cinemas to every screen, and the pandemic was causing dispersion by spreading remote work, telemedicine, and distance learning. Social media plays a crucial role in the dispersion of communities, fostering new forms of connection, competition, and debate across physical divides. Yet, it also serves as a breeding ground for fake news, misinformation, and conspiracy theories. These developments significantly impact the public sphere and political institutions, with state administrations no longer being the sole actors; internet companies now wield considerable influence over public debate and opinion-shaping; government functions are increasingly digitalized through e-government; and public institutions are experimenting with new management and decision-making methods.

The drive to open up policymaking stems from the widely held expert view that government bodies are not reaching their potential due to their operations being isolated from the rest of society (Beth Noveck, among others, expresses such an opinion). Internet technologies bring a significant opportunity for governmental organizations to shift from a closed-off model to one that capitalizes on the knowledge and experience inherent in our digitally interconnected society. Public policy formation undoubtedly requires this knowledge and expertise. Policymaking is an inherently complex process, taking place within ever-changing environments and influenced by varying interests. Governments and grassroots organizations worldwide are experimenting with e-participation platforms to enhance citizen engagement in developing policy, improving service delivery, opening public organizations, and gathering the distributed wisdom of diverse participants. We can also observe a number of bottom-up initiatives designed to significantly impact policy, change the way it is shaped, or even replace the old fossilized political structures—governments, parliaments, and parties—with new ones based on online direct democracy. It is increasingly clear that the most precious asset in our societies is the engaged and informed citizen.

Intelligence ought to be a measure of success in open policymaking. As a trait of both individuals and collectives, intelligence is a gradable, qualitative criterion. It is possible to reason and make decisions in ways that are more or less intelligent. We recognize that openness in public policy creation offers the opportunity to involve a significantly larger number of people in the process than before, with technology providing the means for this involvement. However, mere participation is not enough

to act intelligently: there are projects that are more and less successful, as well as complete failures. There are also empirical observations across various sectors of public life that offer examples of what can make our collective actions intelligent. If internet-based collective intelligence is to be useful, its outcomes must lead to improved decisions, more accurate conclusions, and effective solutions to problems, in other words, intelligent outcomes by some standard. One of the most significant challenges is identifying public policy outcomes that are somehow better or worse, meaning more or less intelligent. In areas such as open-source software development or corporate innovation jams, it is much easier to assess the usefulness of the solutions developed. In the field of public policy, however, it is difficult to define a uniform standard for assessing results. Nevertheless, there are some methods to achieve this. One such example is the evaluation of the maturity of a debate through the diversity and richness of the arguments presented within it.

In the following chapters, we will delve into case studies, theoretical frameworks, and practical observations. We will explore grand, utopian visions of planetary intelligence, grassroots internet projects rooted in the engagement of local communities, and examples of the “wisdom of crowds” as described in James Surowiecki’s famous book. We will examine cognitive research concerning the operations of minds, both in collective and individual contexts; we will review studies on the phenomenon of swarm behavior and reflect on how generative AI is triumphing, all of which enrich our comprehension of natural intelligence. We will also observe that despite the fact that our knowledge in this area seems extensive, its application within public policy remains surprisingly limited, and public opinion tends to highlight instances of collective stupidity rather than intelligence. Addressing these problems, this book not only compiles instances of successful internet projects, research endeavors, laboratory experiments, and social media debates but also presents an attempt to assess them through the lens of intelligence-related processes and identify the factors that may enhance them.

“Collective Intelligence in Open Policymaking” is a call to action for policymakers, scholars, activists, and citizens alike to rethink the foundations of governance and public decision-making. By combining the principles of transparency, inclusiveness, and collaboration, policymakers can tap into the collective knowledge of a wide range of stakeholders and make better-informed decisions. This model recognizes the value of each individual’s contribution while viewing a community’s collective wisdom as a critical asset in navigating the complexities of modern life. It offers a forward-looking vision in which policy formulation transcends traditional top-down approaches, advocating instead for a collaborative process where policies emerge from the collective input of independent citizens. In this vision, the artificial divide between experts and engaged citizens is bridged, allowing for a truly informed public policy of the future.

Kraków, Poland
December 2023

Rafał Olszowski

Acknowledgments

I extend my deepest gratitude to Mark Klein, whose support and insightful feedback were very helpful in shaping the direction of this work. Furthermore, my studies on collective intelligence would not have been possible without the support of my other research collaborators: Marcin Chmielowski, Piotr Pięta, Sebastian Baran, Wasim Ahmed, Marcin Pietroń, and Tomasz Piróg. I truly appreciate their tireless work, innovative ideas, and unwavering commitment to research.

This book owes a great deal to the invaluable insights contributed by a distinguished group of experts who generously participated in interviews to evaluate collective cognitive processes in online policymaking projects. I am profoundly grateful for their time, patience, and thoughtfulness. Particular gratitude is extended to Robert Bjarnason, Robert Guthrie, Miguel Arana-Catania, Yago Bermejo Abati, and Paolo Spada.

I am deeply grateful to the Faculty Authorities and my colleagues at the Faculty of Humanities, AGH University of Krakow, for fostering an environment of intellectual curiosity and cooperation. Special mention goes to dr hab. Barbara Gąciarz, Anna Kulpa, dr hab. Jacek Gądecki, dr hab. Jowita Guja, dr hab. Jakub Gomółka, dr hab. Grzegorz Ptaszek, dr hab. Leszek Porębski, dr hab. Dariusz Wojakowski, Grzegorz Świerk, and Kamil Jaszczynski. Additional thanks to the Faculty Authorities for providing funding for proofreading services.

My deepest appreciation goes to my family and friends for their understanding, patience, and endless support. Their belief in my work and encouragement during the most challenging times have been the source of my strength and perseverance. Heartful thanks go to my friends associated with the Aurea Libertas Institute, particularly Janusz Batko and Rafał Lis.

Special thanks to the editorial team at Springer Nature for their professionalism, patience, and guidance throughout the publication process.

I am profoundly grateful to the **Narodowe Centrum Nauki** (National Science Centre, Republic of Poland) for their generous funding of my research project, “Collective Intelligence on the Internet: Applications in the Public Sphere, Research Methods and Civic Participation Models” (UMO-2018/28/C/HS5/00543). Their support has been crucial in enabling the investigations that form the backbone of this book.

Contents

1	Opening Policymaking	1
1.1	Policy is Different from Politics	1
1.2	Defining the Policy and Policy Sciences	5
1.3	The Era of Great Projects and Great Disappointments	9
1.4	Policymaking as a Positivist, Rational Project	11
1.5	The Policy Cycle: Disaggregating the Complexity	17
1.6	Post-positivist Approaches to Policymaking	27
1.7	The Question of Values and Group Identity	31
1.8	Argumentative Turn and the Role of Deliberation	39
1.9	The Question of Civic Activity and Self-organization	43
1.10	Opening Policymaking in Practice	46
1.11	How to Evaluate Open Policymaking? In Search of Collective Intelligence	53
	References	56
2	Beyond the Individual: Understanding the Evolution of Collective Intelligence	63
2.1	Introduction: The Concept of Collective Intelligence	63
2.2	The Computer Mouse, Online Communities, and “Boosting Our Collective IQ”	67
2.3	Linux and the Open-Source Revolution	70
2.4	Global Brain: A Utopian Vision of the Collective Mind	76
2.4.1	Pierre Lévy	76
2.4.2	Around Lévy’s Thought: Origins and Development of the Global Brain Concept	78
2.5	From the Wisdom of Crowds to Crowdsourcing Projects	80
2.5.1	The Difference Between Crowdsourcing and Collective Intelligence	80
2.5.2	From Madness to Wisdom of Crowds	81
2.5.3	The Rise of Crowdsourcing	84

2.6	Crowdsourcing Beyond Business: Science and Policymaking	91
2.6.1	Citizen Science Projects	91
2.6.2	Crowdsourcing in Public Policies	92
2.7	The Current State of Collective Intelligence Research	105
2.8	Conclusions	119
	References	119
3	A Review of the Projects Using Collective Intelligence in Policymaking	127
3.1	Researching Collective Intelligence in Policymaking: An Overview of Past Approaches	127
3.1.1	Looking for an Appropriate Analysis Method	127
3.1.2	Review of Methods of Studying Collective Intelligence in Policymaking	130
3.2	Collective Cognitive Processes: Foundations for an Analytical Method	141
3.2.1	Theoretical Basis for the Evaluation Framework	141
3.2.2	Developing the Evaluation Framework	145
3.3	Evaluation of the Case Studies	158
3.3.1	Civic Budget of the City of Kraków: The Pandemic Edition	159
3.3.2	Better Reykjavík	170
3.3.3	Loomio	180
3.3.4	Decide Madrid	190
3.3.5	Deliberatorium	199
3.4	Insights from the Evaluated Projects	210
	References	218
4	Online Public Debate. How Can We Make It More Intelligent?	225
4.1	The Online Public Sphere and its Problems	225
4.1.1	Erosion of the Traditional Public Sphere	225
4.1.2	The Online Public Sphere as an Alternative to the Traditional Public Sphere	228
4.1.3	The Problems of the Online Public Sphere	235
4.2	Empirical Research on Online Debates	249
4.2.1	How Different is a Policy Debate from Other Debates? Results of Our Experiment	249
4.2.2	Online Civic Debate in Polarizing Topics: The Vaccination Mandate	259
4.2.3	Predicting Stock Prices by Collective Intelligence on Twitter: McDonald's Stock Performance During Russia's Invasion of Ukraine	267

4.3	Models of Intelligent Public Debate in an Online Environment	272
4.3.1	Engaged and Independent Online Citizens	272
4.3.2	Deliberation, Antagonism, or Agonism: In Search of the Model of Online Debate	276
4.4	Conclusions. Collective Intelligence in Future Policymaking	284
	References	293

About the Author

Rafał Olszowski is a social scientist specializing in research on e-participation, the public sphere, collective intelligence, social media in politics, AI ethics, e-government, and the history of political doctrines. He is affiliated with AGH University of Krakow and the MIT Center for Collective Intelligence. He is the author of a monograph on permanence and change in the political thought of Edmund Burke (2021). He is also a seasoned business consultant in the field of new technologies, and a manager of educational projects. Apart from professional work, he is actively involved in the non-governmental sector with the Aurea Libertas Institute Foundation.

Chapter 1

Opening Policymaking



One should not practically aspire to perfect government, whatever that means; “perfect” defined by one group might be “disaster” for another.

Peter deLeon (1988, p. 118)

1.1 Policy is Different from Politics

Let's begin our exploration of open policymaking by defining what policy is and how it differs from politics. Generally, policy can be defined as a deliberate system of principles designed to guide decisions and achieve rational outcomes (Água & Correia, 2021). At a glance, this broad definition fails to capture the essence of the subject. If we were to stop here, we would struggle to differentiate policymaking from, for instance, business management methods. We might refer to a company's policy, a golf club's policy, or a philatelic association's policy. In this book, however, we will primarily focus on public policy, which encompasses governmental activities related to aspects of social life that are of significant importance to citizens of a particular country. An overly broad definition would also be the one proposed by Thomas Dye, who argues that “policy is anything government chooses to do or not to do” (Howlett et al., 1995, p. 4). This, however, could result in the loss of the distinction between *policy* and *politics* altogether.

To grasp the most relevant definition, let us examine how the concept of *policy* emerged in the social sciences. How did the term *policy* come into general use and become the name of a new field of study? The so-called *policy orientation* in the social sciences emerged in the 1950s and 1960s. Harold Lasswell (1902–1978), an outstanding specialist in political communication, declared the need to develop an interdisciplinary science to produce practical, empirically-tested, systemically-developed, and technologically-supported knowledge that can be continuously applied to solve public problems. Lasswell was a careful observer of reality; he saw the growing importance of public institutions and the new, compelling means at their disposal, both technical (telecommunications, computers, mass media) and financial

(economic prosperity). This reality paved the way for the planning of large public projects. Lasswell's proposal to design public policies systematically, based on the analysis of needs, resources, actors, and public value measurement, was inspiring. At the same time, Laswell's proposal was well rooted in the pragmatism of John Dewey and the economic views of Kenneth Arrow and public choice theory. According to Lasswell, policymaking should cease to be treated as common knowledge, becoming a true science to clarify the process of policymaking in society and/or supply data needed to make rational judgments on policy questions (Lasswell, 2017, p. 121). Like many intellectual trends of the mid-twentieth century, this school originated as a response by Western societies to the threat to democracy posed by the spread of communist (and earlier National Socialist) ideology. The typical approach of this policy school called for separating political decisions from ideology and incorporating a systemic analysis that went beyond the narrow circle of power. Rather than leaving the discretionary decision-making to a single leader, Lasswell emphasized the collective development of justifications, evidence, and programs as a basis for implementing political actions. In the dissemination of the term *policy*, he saw a chance to appreciate the pure art of governing the state, separating it from unproductive power struggles. The term *policy* did not carry as much emotion as *politics*. It was separated from factional power struggles, election campaigns, charismatic leaders, fiery speeches, and gripping ideologies, as well as from those darker sides of political life, including corruption, populism, and vested interests. Lasswell (1970) argues:

We can speak of either government policy, business policy, or my own policy regarding investments and other matters hence policy is free of the many undesirable connotations clustered about the word political, which is often believed to imply partisanship or corruption).

Thus, the term *policy* was significantly different from the term *political*. The science of policy was to be practical, based both on the scientific study of problems and policymaking around these problems, cutting across the existing specializations. Lasswell sees the policy sciences as providing insights, “challenging and informing ongoing processes and decisions, and foretelling of future scenarios, all with the intent of steering government and society toward greater human dignity for all” (Weible et al., 2020).

In turn, *politics*, as defined by Max Weber, revolves around the pursuit of gaining and maintaining power, leadership, and authority. Weber's renowned definition posits that politics “is striving for a share of power or for influence on the distribution of power” (Weber, 2002, p. 311). A *political action* in that sense means that interests in the distribution, maintenance, or transfer of power determine the decision-making process. Weber asserts that those involved in politics seek power either to serve other aims (ideal or egoistic) or for the sake of power itself, to experience the prestige it bestows.

As understood in this manner, politics constitutes a critical battle for influence, occurring both between overt and covert centers, and as a balance of power between parties. For instance, when examining international affairs, a distinction can be drawn

between *international politics* and *foreign policy*. The former refers to the relationships between individual states, shaped by factors such as location, resources, strengths and weaknesses, coalition potential, history, and the temperament of their leaders. The explanation for states' behavior is found at the international level, focusing on their strategic goals, efforts to expand their influence, and the impact of external forces on their conduct. In contrast, *foreign policy* encompasses a government's specific actions that affect its relations with other countries as well as cooperation or competition in particular areas (Waltz, 1996). Politics not only involves executing a political party's agenda but also the process of acquiring and maintaining power through policies that resonate with the electorate. While ideology plays a crucial role, practical considerations also influence every decision made by a government.

Thus, *public policy*, as discussed in this book, pertains to the planning, implementation, and evaluation of activities undertaken by public institutions and their non-public partners in distinct areas of societal life. The art of policymaking may employ specialized tools from project management to advanced data analysis, including decision trees, cost-benefit analysis, and complex systems modeling. Additionally, policymaking can utilize the concept of *collective intelligence*, which will be examined in subsequent chapters. Above all, policy should address the needs of both decision-makers and citizens by introducing changes or maintaining certain practices while striving to achieve specific, verifiable outcomes. Individual efforts involved in shaping public policies concentrate on devising practical recommendations in the form of projects and programs while, as Martin Bulmer noted, "blending theoretical insights with empirical inquiry" (deLeon, 1988, p. 96). A policy must offer sufficient specificity to guide the development of applicable programs. According to Lasswell, the entire policy process should emphasize planning and enhancing the quality of information upon which decisions are based.

Let us now look at the spheres in which public policies are shaped. Figure 1.1 illustrates the interrelationships among various aspects of government, including policy, politics, and administration. Effective governance necessitates coordination across these domains. The head of the government, typically the prime minister or president, is the epicenter of power. While some actors function in multiple domains, they all seek coherence from the center. A crucial responsibility of a leader is to ensure synchronization across different domains, fostering collaboration among all government elements to achieve shared objectives. By assessing various perspectives, the prime minister or president can evaluate the alignment of a specific proposal with the government's overall agenda concerning strategic matters. The *politics* sphere encompasses political parties and factions, as well as the individuals affiliated with them. This applies to both party leadership and grassroots members. Members of parliament constitute a group of politicians who wield tangible influence on the implementation of their programs through both government support and constructive opposition. Ministers associated with political parties also form a significant group. Additionally, lobbyists and various covert pressure groups play distinct roles.

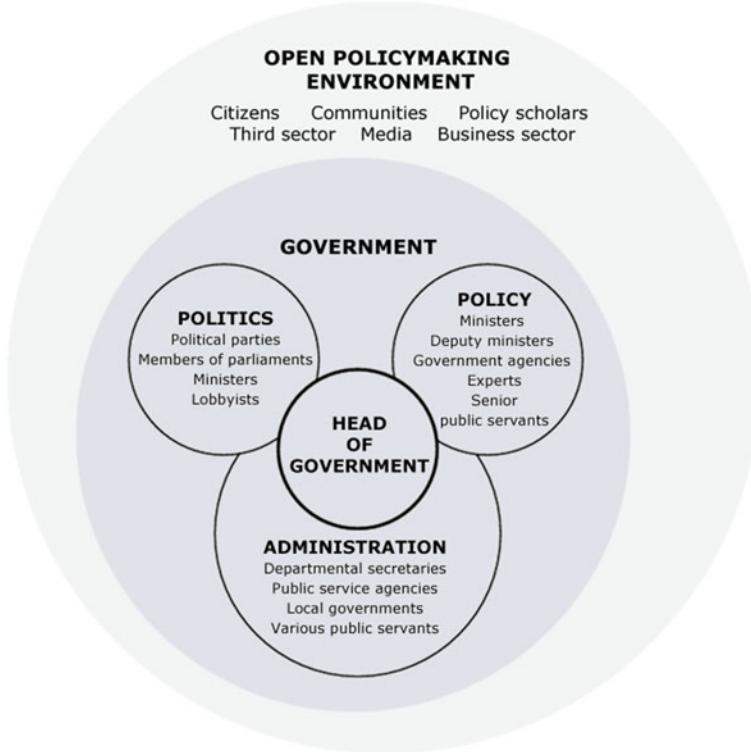


Fig. 1.1 The spheres where public policies are shaped

The *policy* sphere also encompasses the activities of ministers; however, in addition to them, it involves a wide array of full-time employees across various government agencies. When politicians assume ministerial roles, they become accountable for public outcomes, which entails setting goals for the government. Their policy responsibilities also include engaging with the public service, exposing them to a domain beyond mere *politics*. Public servants provide policy advice to governments without delving directly into political issues. Their duties span administrative tasks, policy implementation, and policy enforcement. Public servants operate in diverse branches of government, including the civil service, military, law enforcement, and education. As heads of government agencies and members of expert groups, they perform crucial tasks in this sphere. Naturally, political actors and agendas often exert a considerable influence on policy development. While some policies are assigned to the public service or public enterprises for implementation, others ultimately fall under the minister's purview, who may seek the prime minister's final approval. Consequently, even thoroughly researched and broadly consulted policy positions can be overturned at the discretion of the governing administration.

The *political realm* predominantly depends on parliamentarians, their advisors, political parties, and organized interest groups. In contrast, the *policy realm* encompasses a more extensive range of the public sector. For policies to effect change, however, they necessitate the resources and authority of the administration. Once decisions are finalized, the government focus must shift to directing efforts toward achieving the desired outcomes. This is the domain of the administration where services are provided, taxes are collected, and laws are enforced. The majority of policy implementations occur in the policy sphere, removed from the scrutiny of political actors (Althaus et al., 2007, p. 23).

Finally, an area that some authors have excluded from the direct shaping of policies but will be particularly important for us in this book is the *open policymaking environment*. This domain includes: the citizens in general, especially those with a sense of common good and civic virtue, various communities, independent policy scholars, the third sector, and the media and business sectors. We will revisit the concept of *openness* and discuss the potential impact of this sphere on policymaking at the end of this chapter.

1.2 Defining the Policy and Policy Sciences

Numerous authors have attempted to provide their own definitions that tell us a bit more about what policy is and is not. For example, Heclo (1972) points out that for policymakers it is not always necessary to act; sometimes it is enough to refrain from acting: “A policy may usefully be considered as a course of action or inaction rather than specific decisions or actions.” In the same vein, Howlett, Ramesh, and Perls (1995) write: “Decisions by governments to retain the status quo are just as much policy as are decisions to alter it.” This brings to mind the famous idea of *salutary neglect*, Britain’s unofficial policy initiated by Prime Minister Robert Walpole to lessen the enforcement of strict regulations, particularly trade laws, imposed on the American colonies at the turn of the seventeenth and eighteenth centuries. Refraining from regulating and imposing a fixed social order on the colonies contributed to their amazing development, which later gave rise to the United States. Schaffer (1977), in turn, points out that “Policy is a committed structure of important resources.” Therefore, as in well-run project management, we should carefully estimate the expenditure of resources available to us in relation to the goals planned to be achieved. This was supplemented by Easton, (1953) who wrote that a policy “consists of a web of decisions and actions that allocates values;” it follows that values are not only material resources, but also social goods, culture, and other things important to people. In turn, Jenkins (1978), writing that policy is “a set of inter-related decisions (...) concerning the selection of goals and the means of achieving them within a specified situation,” shows us that individual public policy decisions are not isolated from each other, but form a complex, interdependent system. The long-term significance of implemented policies and their historical context is indicated in the definition formulated by Friend, Power, and Yewlett (1974): “Policy is essentially a stance which, once

articulated, contributes to the context within which a succession of future decisions will be made.” By writing that “policy process is complex, messy and often appears unpredictable,” Cairney (2020) draws attention to the exceptional complexity of these processes, the analysis of which always distorts real phenomena, using useful—but limited—simplifications. The diverse environment in which policies are created is pointed out by Kal H. Colebath (1998): “Diverse activities by different bodies are drawn together into stable and predictable patterns of action which (as often as not) come to be labeled ‘policy,’” while Schneider and Ingram (1997) describe policy designs as enshrined in “statutes, administrative guidelines, court decrees, programs, and even practices and procedures of street-level bureaucrats.” Finally, Wildavsky (1979) notes that in policy all actions are closely related to the result and should not be considered in isolation from it: “Policy is a process as well as a product. It is used to refer to a process of decision-making and also to the product of that process.”

Let us note the distinction between public policy as governmental action and public policy as science. The term *public policy* is used to describe both the sphere of practical political activities and a certain type of scientific inquiry. Policy analysis and policymaking are primarily practical activities. On the other hand, theoretical (research) approaches to public policies constitute the discipline of teleological (purpose-oriented) sciences, which should justify, explain, and improve practical activity. First, policymaking refers to the actual policies and programs of governments that affect people’s lives and well-being. It is in this sense that we will primarily use the term in this book. Second, it denotes the area of scientific research concerning the description, explanation, and evaluation of these policies (Midgley et al., 2000).

Policy analysis as an organized process has a long tradition. Since ancient times, when structures of power emerged from the darkness of history, there was a high demand for competent advisors, sages, and trusted counselors of monarchs. It is difficult to consider them as open and transparent, however. As Peter deLeon put it, there was “only scarce and episodic regard for recording and routinizing the activity. These were, in fact, zealously protected prerogatives; their practitioners had little incentive to share their closely guarded secrets of access” (deLeon, 1988, p. 15). If we look at the famous oracle of Delphi, we see that mystical ceremonies had a significant impact on the politics and laws of ancient Greece, but only *after* the proper analytical work, i.e., translation into hexameter by priests who were interpreting the words of Pythia. The Athenian strategist Themistocles, guided by the words of the oracle, prepared a defense strategy against the Persian invasion based upon the revelation that the Greeks should trust their wooden walls, which was read as an indication to rely on their fleet. Philip of Macedon also sought advice at Delphi and heard: *With silver spears, you may conquer the world.* So, the king took control of the silver mines. This affected the condition of the state treasury, and thus the ability to win the favors of the rulers of the surrounding city-states and to shape their policy under his own dictate (Orrell & Chlupatý, 2016).

Like so many accomplishments of European civilization, systematic and continuous policy counsel probably grew out of relationships between the Church and the state in the Middle Ages, as these two centers of power sought to balance their influence (deLeon, 1988, p. 15). Taking place in Europe between the end of the

seventeenth century and the beginning of the nineteenth century, the Enlightenment brought forth the development of systematic inquiry in the form of empirical research and the increased prominence of applied science in European society. As a result, we observe the professionalization of political action. An example is the work of the eighteenth-century British politician Edmund Burke, who, over 15 years of his involvement in the supervision of the East India Company, amassed an impressive dossier on the company's activities in India, including seventeen volumes of financial statements, eyewitness accounts from various social groups, reports on abuses in the exploitation of Bengal, studies on Indian culture, power relations and traditions, and travelogues and business documents. This allowed him to prepare a plan for reforming the country's political system, which Burke defined as *Magna Charta Libertatum of Hindustan*. Although these reforms were not implemented during the author's lifetime, the accumulated knowledge allowed the British monarchy to put an end to the Company's worst abuses (Whelan, 1996).

In the twentieth century, another very important circumstance appeared, which, apart from the threat of totalitarian ideologies, had a major impact on the development of policy sciences. Governments around the world were undergoing transformations related to the tremendous growth of public sector activities. These policies had their first apogee in the era of Roosevelt's New Deal, and then again in the two decades after World War II. Factors such as recurring economic crises, growing tension in international politics, new technical tools allowing for greater control of social processes, and economic development bringing new financial resources increased the appetites of governments to bring further areas of social life under its control. The era of *laissez-faire* was becoming a thing of the past, and a new model of the relationship between governments and citizens emerged where large social programs implemented by active governments came to the fore. Meticulous government policies began to appear when the state administration became interested in areas of life that had previously belonged to the private sphere. Strategies, programs of resource allocation, specific measures of a regulatory nature; goals, means of action, and deadlines have become the bread and butter of public administration.

Our current knowledge of the complexity and interaction of various spheres of social life leads us to the conclusion that there is almost no area that does not have a public dimension. This perception of the world distinguishes us from people living in the past. The dichotomous division into the private and public spheres, which, according to Arendt (1998), characterized European cultures until the industrial revolution, originated in ancient Greece and was illustrated by the opposition of the household (*oikos*) and the political community of free citizens (*polis*). In antiquity, there was a fundamental division between the private and political spheres. The domain of politics was the public space: the space of *action, disclosure to others, and perception of others*, while the domain of the private sphere was secrecy. Home and family, as the center of the private sphere, were related to the maintenance of life (both individual and as a species). The private sphere also included matters related to economics and trade, which were traditionally associated with *oikos*. Everything that had an economic character, and therefore was related to the life and survival of an individual, was inherently non-political and non-public. On the other hand,

everything related to the participation of a free individual in making policy decisions belonged to the public sphere. Participation in policymaking meant activity in the public arena, expressed through argumentation and persuasion, and the implementation of the results of these debates in the activities of the state. In modern times, there has been a significant expansion of the field of public affairs, i.e., the field of interest to state officials.

The radical population growth associated with the Industrial Revolution and the accompanying urbanization processes resulted in the emergence of a new type of society, which Jose Ortega y Gasset called the *mass society*. It was preceded by the weakening of old communities centered around traditional norms and values, which in turn led to loosened social control mechanisms. Mass society changed the perception of the public and private spheres. The capital economy transformed formerly social relations, based on personal dependence, into relations based on monetary valuation and expected rate of return. Instead of permanent social relations, short-term and superficial contacts began to dominate. When economic problems rose to the rank of general social issues, the relationship between the private and the public sphere underwent a fundamental change. Since then, the concern for ensuring prosperity or general means of subsistence has ceased to be an individual matter, instead becoming the central problem of politics. Breaking the border previously separating these spheres opened the way to further transformation. The place of personalized governments was taken by professional administration, which supervised larger and larger swaths of society. This can be seen in the number of people employed in public administration, as an example. In 1792, at the very beginning of the United States, only 780 people worked in the federal services (if you exclude deputy postmasters). In 2020, the number of federal government (civilian) employees reached 2.93 million people (Noveck, 2015). If you count everyone working in the federal, state, and local governments, you get a staggering 19.77 million people.

We can divide the policies shaped by these officials into two general types. The first one covers areas traditionally belonging to the state, such as defense, international affairs, taxes, the legal system, or internal security. The second is related to broadly understand economic and social issues. Excluding the traditional areas of activity, states began to pursue a permanent policy related to the economy, i.e., the labor market, education, social security for citizens and residents, and health in the mid-nineteenth to the twentieth centuries. Problems of increased interest in public institutions included employment and unemployment, economic crises, poverty, and wealth distribution, family matters, the natural environment, social conflicts, and inequalities. The inefficiency of tasks undertaken by the state caused additional problems including bureaucratization, inflation of law, and maloperation of public institutions. In the last 100 years, health policy, which had been the domain of private physicians and charities for centuries, has been covered by a state-supervised compulsory insurance system in most countries. Considering the scale of intervention of public institutions in this sphere of life, it was a real revolution. One of the most famous institutions in this field, the British National Health Service, was recognized as an “experiment in health care (...), a lone example in Western Europe of ‘socialized medicine,’” at the time of its establishment in 1948. Aneurin Bevan, architect of

this new service, believed the NHS had become “the most civilized achievement of modern Government” (Webster, 2002).

Not every country, however, has gone through such a significant transformation of the health care system. In the United States, health care is still suspended in the old model based on private medical providers. The proposal to introduce public health care, similar to the British one, is the subject of a fiery dispute. Supporters of such reform believe that medical services should be guaranteed by the state and available to a wide range of citizens. On the other hand, a postulate to maintain the voluntary nature of insurance is put forward by critics of the model in which state officials act as administrators of doctors. Supporters of private health care believe that the high quality of medical services in the US is the result of the fact that they develop outside the area of direct government intervention.

1.3 The Era of Great Projects and Great Disappointments

Large public projects became a determinant of the politics of many countries in the post-World War II period. International economic recovery programs, the most famous of which being the Marshall Plan, were followed by extensive domestic projects. Initially, natural scientists, such as Vannevar Bush and Karl Compton of MIT, criticized the social sciences for their inability to produce objective, practical knowledge (Cherrier & Fleury, 2017). The National Science Foundation was even established in the US in 1950 without a division for social sciences. Funding for policymaking research, however, was primarily sourced from military agencies and large philanthropic foundations like the Ford Foundation. These entities placed a large emphasis on problem-focused research, quantification, mathematical foundations, impartiality, and, most importantly, scientific neutrality and value-free analysis (Cherrier & Fleury, 2017). The emergence of new concepts, such as public choice theory, and the growing interest of governments in running large social programs led to a greater use of the achievements of policy sciences in public administrations. President Lyndon B. Johnson declared in 1964:

An ‘unconditional war on poverty in America’ that aimed ‘not only to relieve the symptom of poverty, but to cure it and, above all, to prevent it.’ The ‘chief weapons’ in this war he pledged to use were ‘better schools, and better health, and better homes, and better training, and better job opportunities.’ He also promised more libraries, public transportation, and food relief for the poor. (Council 2014)

The rhetoric of the War on Poverty quickly found its way into legislation, establishing new federal programs and agencies. The Economic Opportunity Act of 1964 was passed by Congress and became law in August of the same year. This act created the Office of Economic Opportunity, which provided funds for vocational training, Job Corps to train youth in conservation camps and urban centers, and Head Start, an early-education program for children of poor families, among other programs (Cooley, 2020).

The War on Poverty is remembered as a monumental policy experiment. What's more—it still arouses controversy in American society, among some it is remembered as the implementation of a great, wonderful vision (Bailey & Danziger, 2013), among others as a complete failure, which not unsuccessfully solves social problems, but also worsens them. This program directly involved the government in schools, hospitals, neighborhoods, and labor markets. Programs and policies originating in this era continued into the 1970s when two major social programs, Supplemental Security Income (SSI) and the Earned Income Tax Credit, were introduced. The full legislative agenda laid out in Johnson's 1964 State of the Union speech and extended in the Economic Report of the President included: maintaining high employment, fighting discrimination, improving regional economies, rehabilitating urban and rural communities, improving labor markets, expanding educational opportunities, extending access to health service, promoting adult education, reducing food insecurity and improving child health, as well as assisting the aged and disabled.

Head Start, for example, provided early education and health services to low-income preschoolers. Critics claim that Head Start's impacts on test scores are middling, but the program also affects noncognitive skills, such as motivation and attitudes, that contribute to longer-term educational achievement and earnings growth. In their research, Jens Ludwig and Douglas Miller stated that a 50–100% increase in Head Start funding was associated with an increase in schooling attainment of about one-half year and a 15% increase in the likelihood that participants attend college (Bailey & Danziger, 2013). Workforce development programs, another important field of state intervention, were focused on increasing the employment and earnings of the disadvantaged. In the early 1980s, the Manpower Development Research Corporation ran a number of *supported work* programs for disadvantaged youths financed by the federal government with the aim to instill positive work habits and attitudes. These programs were frequently criticized by opponents of excessive state intervention. Reporter Ken Auletta described in his book *Underclass* attending one of these courses in New York City and discovering that students typically showed no signs of attention or respect during class, even sleeping or reading newspapers. The trainer in charge explained that if the rules were applied, "we'd lose just about everyone in the class." The overall effect—as Auletta claimed—has been to *train* participants in irresponsibility: they learn that the world will keep rewarding them even when they do not live up to their obligations (Auletta, 1983, p. 23).

One of the most acclaimed legacies of War on Poverty was a significant decline in the official poverty rate of the elderly. This was a result of the enactment of the Older Americans Act of 1965, several congressional actions that increased Social Security benefit levels and then indexed them to inflation, and the passage in 1974 of the federal Supplemental Security Income program that provides the elderly poor and disabled with a guaranteed annual income. SSI replaced Old Age Assistance programs and expanded access to public benefits for the poor elderly, particularly in southern states. There was little doubt of the validity of the reforms and raised little controversy from either commentators or political opponents.

It was different with health care reforms. Medical care has been consuming an ever-increasing share of the federal budget over the past 50 years. Supporters of medical care programs see their positive impact in the reduction in infant mortality and increased life expectancy. Medicare, as an example, provides health insurance coverage to all individuals aged 65 and older, but—what was an unintended effect—increased medical costs for all Americans, not just the elderly (Bailey & Danziger, 2013). In recent years, the Affordable Care Act, introduced by President Barack Obama's administration and popularly known as Obamacare, has been the subject of political controversy. On the one hand, the program introduced compulsory insurance; on the other hand, it enabled the purchase of insurance policies for 22 million of the poorest Americans. Despite attempts by his opponents to remove it, Obamacare has proven difficult to change due to the popularity it garnered. The major project to transform American healthcare, initiated under Johnson, is still a work in progress.

1.4 Policymaking as a Positivist, Rational Project

As we can see, public policy programs in the 1960s and 1970s were primarily treated as state intervention tools to address market failures, appearing in complex systems at the interface between the economy and society. The problems that public policies had to challenge, according to Howlett, Ramesh, and Perl (1995, p. 22), were primarily seen as: *natural monopolies*, *imperfect information*, and *the tragedy of the commons*. The first of them, *natural monopolies*, appeared in certain industries with large capital requirements and disproportionate profits where there is room for a small number of large companies operating at the expense of others, which could lead to higher prices and lower quality at the expense of consumers. This situation is particularly common in industries such as telecommunications, railways, electricity, and heat production. In turn, the problem with *imperfect information* occurs when consumers or stakeholders have insufficient information to make rational decisions in the market. An example is the activity of large pharmaceutical companies, which “have no incentive to reveal adverse side effects of their products, nor the consumers have the expertise required to evaluate such products prior their use” (Howlett et al., 1995, p. 23). Another example of observed problems, *the tragedy of commons*, occurs when common-pool resources, such as forests, fisheries, clean air, natural beauty, etc., are exploited without a requirement to maintain the resource for future generations. In the absence of clear rules, farmers using a common pasture may seek to increase their short-term benefits, regardless of the fact that in the long-term, they may deplete what was once a common resource (Ostrom, 2015).

For years, the belief that social science had sufficient resources to breakdown this complex reality and design changes that would enable governments to correct social processes and control their operation was the dominant approach in policy development. Analytical work carried out by competent experts rose to the rank of *most important activity* in the field of policymaking. The ambition of this school of thought was to train a class of apolitical professionals who are able to judge society

coolly and without prejudice. The systematic collecting, analyzing, and evaluating of data was to allow for reasonable, objective analysis of policy goals and outcomes. We call this approach positivist, or rational-comprehensive. As Weiss (1977, p. 4) summarized this belief in the late 70s:

There is much hoopla about the rationality that social science would bring to the untidy world of government. It would provide hard data for planning, evidence of need, and resources. It would give cause-and-effect theories for policymaking, so that statesmen would know which variables to alter in order to effect the desired outcomes. It would bring to the assessment of alternative policies a knowledge of relative costs and benefits, so that decision-makers could select the options with the highest payoff. And once policies were in operation, it would provide an objective evaluation of their effectiveness so that necessary modifications could be made to improve performance.

The logic of inquiry employed by positivists involves analyzing problems with the aim of calculating solutions that are specific to the issue at hand. When using modern management techniques, administrative state operations are seen as controllable, efficient, and subject to calculation. This approach prioritizes the calculation of solutions for policy problems, with analysis becoming a central element of policy formulation and implementation that aims to resemble an automated mechanism. The emphasis is on routine and formalized procedures, with analytic efforts focused on treating policy issues seen as programmed decisions. This approach is rooted in a technocratic ideal, where analysis becomes closely associated with management science (Torgerson, 1985).

At the core of this concept, we see the epistemological belief that a perceptible reality truly exists, and that observations have temporal and contextual independence from the observer; therefore reality can be objectively known (Morcöl, 2001). Accurate, objective descriptions of reality, based on formalized analytical work, are to be formed as they are in the natural sciences. The study of physics, for example, directly observes and measures natural phenomena. In the social sciences, according to the positivist approach, we discover laws and describe them regardless of our personal, moral views much as a physicist observes the universe. We use quantitative analysis, the objective separation of facts and values, generalizable findings independent of cultural customs, and local traditions or moral imagination to describe the political reality in detail. Objectivism is the most fundamental epistemological assumption of the positivist approach (Lincoln & Guba, 1985; Hawkesworth, 1988).

The consequence of adopted objectivism is the separation of facts from values. In this view, policy analysis, as based on empirically verified knowledge and scientific objectivity, assumes value neutrality. According to positivist belief, an objective reality exists and is reflected in our minds, so *facts* which pertain to it can be distinguished from the emotional states of the mind (*values*) (Morcöl, 2001). According to this assumption, the policy analyst does not need to consider how the economy and the political system *ought* to function. Rather, they focus on the best possible description of how a system works, how it is designed, and how it is effectively managed. Lasswell, especially at the beginning of his career, followed the Weberian tradition which refused to prioritize values and perceived individual preferences in terms of goals (Zittoun, 2019). As Philipe Zittoun states:

Lasswell in his famous book ‘Politics: Who gets what, when and how’, proposes that, distinct from political philosophy which seeks to ‘justify preferences’, political science should be considered as the scientific ‘study of influence’. Following his previous work on propaganda, his book focused on how elites used symbols, violence, values, and goods to maintain their dominant positions. (Zittoun, 2019)

According to Lasswell, bringing ethical values to the foreground of political analysis explained nothing because, ultimately, they served to increase the influence of decision-makers. This way of thinking focuses on technical problems and *management*, rather than posing problems which require knowledge of ethics. This mindset also focuses on the ability to learn from practice rather than the application of *ideologies* or moral principles. As summarized by Anderson (1979), “in most theories of policy rationality, derived from economic theories of the utility-maximizing individual and a positivist conception of valuation, values are to be regarded as the ‘preferences’ of the policymaker.”

This is a departure from the idea of the common good in favor of identifying particular goals and interests relating to specific problems. As Peter deLeon summed it up, “the prevalence of single-interest politics has overwhelmed the concept of the common good” (DeLeon, 1988), which, since ancient times, has been the subject of reflection by thinkers and statesmen. The common good refers to what is beneficial for all members of a community or society; it is a collective outcome that is valued and sought after because it contributes positively to the well-being of everyone. In policymaking, various terms can be used as synonyms for the common good, including general benefit, social good, common interest, common benefit, general happiness, general interest, public good, community spirit, and social harmony (Kaul et al., 1999).

The classical political philosophy related common good to the virtue of justice. In Plato’s view, justice, perceived as the distribution of natural equality in society, should be treated as the primary goal of the state’s activities. In turn, this would lead to a just government and societal happiness. He expressed it with the words:

And this is justice, and is ever the true principle of states, at which we ought to aim, and according to this rule order the new city which is now being founded, and any other city which may be hereafter founded. To this, the legislator should look,—not to the interests of tyrants one or more, or to the power of the people, but to justice always; which, as I was saying, is the distribution of natural equality among unequals in each case. (Plato, 360 B.C./2008)

This idea was further developed by Aristotle, who emphasized that:

It is evident, then, that all those governments which have a common good in view are rightly established and strictly just, but those who have in view only the good of the rulers are all founded on wrong principles, and are widely different from what a government ought to be, for they are tyranny over slaves, whereas a city is a community of freemen. (Aristotle, 350 B.C./2004)

The school of policy science, however, has narrowed the field of interest in the study of good governance. Jenkins (1978) defines policymaking as a goal-oriented behavioral science. In this definition, shaping public policies is making governmental

decisions that define a goal and setting out a means to achieve it without bringing to the foreground such general ideas as justice or the common good.

Capturing reality as an objectively accessible space, independent of the world of ideas, requires specific tools. One of the positivist's favorite tools is the use of quantitative analysis, which focuses on aggregate data sets, metrics, measurable goals, mathematical methods, and computational techniques. Application of such a methodology to a policy problem involves the use of a mathematical model as a simplified representation of that problem or of some phenomenon important to the problem. Quantificationism as a methodological orientation assumes that quantification is a value in and of itself and that quantitative answers are, by definition, better than qualitative ones: “discrete entities and events can be enumerated and measured precisely and that the inherent quantitative characteristics of entities and events define their behavior in a wholly definite and predictable way” (Morcöl, 2001).

The desire to treat policymaking as a fully rational process emerges from the aforementioned factors: objectivism, separating facts from values, and the wide use of countability. As Stone (2002) points out, *the rationality project* in policymaking has shared the “common mission of rescuing public policy from the irrationalities and indignities of politics, hoping to make policy instead with rational, analytical, and scientific methods”(p. 6). Enthusiasts of this approach believe that rationality in shaping public policies should not only include analyses but also the entire decision-making process. The first stage of this process is the accumulation of rational knowledge. This knowledge is acquired methodically, formulated in a language that is intersubjectively understandable (i.e., communicative, clear, precise), and—most importantly—fulfills an informative (not persuasive) function. This knowledge is logically systematized (non-contradictory and consistent) and free from emotional and volitional states. The issue of rationality (or lack thereof) should also be decidable, i.e., its notion should be formulated in such a way that rational objects can be distinguished from others.

The theoretical foundation for the science of rational action is praxeology, the main goal of which, according to Tadeusz Kotarbiński, is to explain “how to act in order to act as efficiently as possible,” and how efficiency also means effectiveness (Kotarbiński, 1965). According to this theory, rational action requires us to direct our actions toward a certain goal (realizing something that we evaluate positively and avoiding something that we evaluate negatively). Preparation for action takes place through prior consideration (developing a goal and a plan) and the selection of the best means in the given circumstances. These means should both lead to the goal and be appropriate relative to the capabilities of a given entity. The efficiency and effectiveness of the action depends on the knowledge possessed, the knowledge available to the participants of the action, whether it was true, complete, how it was obtained, and how this knowledge was used when formulating the justification for the action. The rationality of action is defined by Kotarbiński (1965) as the justification of a practical directive (action plan) on the basis of the knowledge of the justifier (the justifier may be the actor, but also the observer).

Whether an action is more or less rational depends on whether the answer to the problem of how to act as efficiently as possible, adopted as a directive for action, is more or less justified on the basis of the knowledge available to the justifier (...). The comparative scale for choosing a particular course of action is expected utility, and the maxim of rationality would be the rule ordering to choose in such a way as to maximize expected utility.

It is worth noting that optimization, the choice of the best means to achieve a given goal, goes hand in hand with the exclusion of evaluations (moralizing) from the sphere of rationality, leaving only certain ordering activities there. Therefore, it becomes rational to achieve the assumed goal, regardless of its moral value.

The desire to make policy sciences practical and based on methodological assumptions of objective scientific laws, individualism, and rational self-interest led to the development of *rational choice theory* or *public choice theory* in the post-World War II period. These theories were formed based on the works of three economists—Duncan Black, James Buchanan, and Kenneth Arrow. Their idea was to use the theory of neo-classical economics to explain all social processes, asserting that “the first principle of economics is that every agent is actuated only by self-interest” (as expressed by Francis Edgeworth in his *Mathematical Psychics* published in 1881). This theory holds that individuals do not cooperate to achieve common goals unless coerced, always acting as *self-interested, strategic, and rational actors* and constantly calculating the utilitarian value of their participation in social life. The concept of rational choice is widely attributed to utility theory, initially proposed by John von Neumann and Oskar Morgenstern in 1944. According to their theory, individuals tend to follow pre-determined patterns of behavior with the underlying assumption that they always prefer having more wealth than less. The rational agent, in this case, is expected to have knowledge of the *probability distribution*, which is a list of possible events that are associated with a probability of occurrence. Utility theory assumes that individuals have consistent and stable preferences and make decisions based on the maximization of subjective, expected utility. Essentially, given a set of options and beliefs expressed in probabilistic terms, the individual is expected to maximize the expected value of their utility in all the activities they conduct, whether in their professional or private life. Probability estimates and utility values are both essential elements of calculation for maximizing the expected utility function. The individual assesses relevant probabilities and utilities based on personal opinion and all available relevant information. When faced with two possibilities that have the same consequences, an individual is expected to prefer the opportunity that offers the highest probability of the best result (Schilirò, 2012). As Buchanan, the first among public choice theorists to win the Nobel Prize, put it: “In one sense, all public choice or economic theory politics may be summarized as the ‘discovering’ or ‘re-discovering’ that people should be treated as rational utility maximizers, in all of their behavioral capacities” (Howlett et al., 1995, p. 32).

Arrow's 1951 *Social Choice and Individual Values* was one of the earliest texts that applied rational choice theory to explain public decisions. Author presented political decisions as a collection of choices of individuals pursuing private preferences. Arrow's (1950) *impossibility theorem* stated that no collective choice could truly satisfy the rationality conditions: “if we exclude the possibility of interpersonal

comparisons of utility, then the only methods of passing from individual tastes to social preferences which will be satisfactory and which will be defined for a wide range of sets of individual orderings are either imposed or dictatorial.”

In conclusion, until we accept comparing the differing levels of utility understood by different individuals, it is impossible to formulate a social-preference order to make any kind of general public decisions. The individual agent is taken as the fundamental building block for all analyses. Agents are assumed to have certain objectives or preferences and interact in particular institutional settings. Aggregate outcomes are discovered and characterized by examining the consequences of the assumed behavior of individual agents and the institutional constraints on that behavior. Therefore, we can explain a pattern of social behavior, or any enduring social arrangement, as the aggregate outcome of the goal-directed choices of large numbers of rational agents. Little (1991) cites a few examples of how any social event can be explained in this way by referring to events in Chinese history:

Why did the Nian rebellion occur? It was the result of the individual-level survival strategies of north China peasants. Why did the central places of late imperial Sichuan conform to the hexagonal arrays predicted by central-place theory? Because participants-consumers, merchants, and officials-made rational decisions based on considerations of transport cost. Why was late imperial Chinese agriculture stagnant? Because none of the actors within the agricultural system had both the incentive and the capacity to invest in agricultural innovation (Little, 1991).

As we can see, in the *public choice theory* a concept of common interests has been sidelined for the individual rationally seeking their own gains. Arrow argues that if we think about *collective* choice we should always have in mind the sum of individual preferences and called for rejecting “mysticism” and “the organism approach to social problems” (Arrow, 1950). In his appreciated work concerning public expenditures, Samuelson (1954) later uses the same words to dismiss the existence of a “mystical collective mind” and Buchanan (1949) claims that “State decisions are, in the final analysis, the collective decisions of individuals,” adding that the state had “no ends other than those of its individual members.”

To sum up, the so-called scientific approach to public policy has allowed for the discovery that policy creation is both a process and a craft which does not depend on the vision of an *outstanding individual*, but can be organized and institutionalized. However, such a view also carries risks; e.g., the professionalization of policymaking can lead to detachment from debate and ideas, thereby limiting openness. As a result, policymakers may become overly reliant on technical expertise and bureaucratic procedures, neglecting the importance of public input, deliberation, and the diversity of perspectives that contribute to a vibrant policy discourse.

1.5 The Policy Cycle: Disaggregating the Complexity

One of the most important applications of the positivist approach to public policy-making analysis is based on a model called *the policy cycle*. Like practitioners in the physical sciences, policy cycle advocates believed that scientific truth would result from the application of the rigorous scientific method and that good policy is likely to result from the application of a similarly rigorous process (Althaus et al., 2007; Edwards, 2001). Public policy specialists saw themselves as empirical researchers of objective reality, gaining evidence through observation, experimentation, and scientific investigation as the basis for understanding the world: “An empiricism similar to that used in physics was considered the best device for learning political truths” (Goodin & Klingemann, 1996, p. 559).

The policy cycle was seen as a mechanism for comprehensive and systematic policymaking where the focus is on the process. This prescriptive model—as Australian researcher Sophia Everett called it—comprised of:

A number of logically ordered sequential steps comprehensively canvassed the policy operations. The rigorous application of this mechanism ensured a ‘rational’ outcome by selecting the most effective means of achieving an end, breaking down decision-making into phases, ensuring a comparison of options, and providing a single most appropriate answer (Everett, 2003).

Despite the fact that, in various forms, this framework was present in the work of policy researchers from the very beginning, it gained particular popularity in Australia, where subsequent editions of the book *The Australian Policy Handbook* are one of the most important reading materials for public administration employees. The policy cycle—as authors of this handbook suggest—“brings a system and a rhythm to a world that might otherwise appear chaotic and unordered” (Althaus et al., 2007).

To imagine what a policy cycle is in the simplest terms, we can assume that there is a *system* that has inputs (*problems*) and outputs (*policies*). But what happens in this *magic box* along the way, before the problem turns into policy? In the real world, it is a whole variety of things. The model is supposed to show a simplified picture of that variety by breaking the process down into a number of discrete stages, although it fails to explain everything of course. For example, it will not give us the answer to why issues arise on the policy agenda or why a particular decision is made. However, it can be useful for us in two ways:

- (a) As prescriptive model, showing on the basis of accumulated experience how policymakers *should* operate to make sure that their decisions are made in a systematic way;
- (b) As descriptive model, explaining how policymakers actually *do* operate, to simplify the study of how they make decisions (Cairney, 2020, p. 26).

The idea to model policymaking processes in this way appeared at the very beginning of public policy studies. Lasswell proposed to divide the policy process

into seven stages (1970). They were: (1) intelligence, (2) promotion, (3) prescription, (4) invocation, (5) application, (6) termination, and (7) appraisal. This model has become a very influential achievement in policy studies and an inspiration for many subsequent attempts, including Brewer (1974), and Jenkins (1978). Brewer and Jenkins both extended the staged theoretical approach, proposing a modified decision sequence. More recently, Althaus, Bridgeman, and Davis (2007) developed the Australian Policy Cycle as a heuristic to articulate the policy process and guide public policy practitioners in Australia.

As Howlett, Ramesh, and Perl have noted, the policy cycle model had little to say about external influences on the state: “it simply assumed that policymaking was limited to a small number of officials” (Howlett et al., 1995, p. 11). Public policies in this paradigm were created in a rather hermetic world of bureaucratic institutions, hierarchies, expert groups, and strict reporting. There was not much room for doubt. Moreover, in these early models, policy creation was treated quite linearly. The individual phases were to follow one another. This was intended to reduce the complexity of studying public policy by isolating each phase as a separate process and examining it before the whole process was put back together again. It was undoubtedly a valuable undertaking, intended to be a way out of the chaos of ad hoc decisions made by visionary leaders, not professionals. By “disaggregating the complexity of the process” (Howlett et al., 1995, p. 13) into several stages and sub-stages, this model facilitates understanding of policymaking by allowing the processes to be investigated alone or in terms of their relationship to any or all the other phases in the cycle. But taken literally—in a *technocratic* way, as critics (like John S. Dryzek) called it (Dryzek, 2002)—this approach could lead to wrong decisions while leaving decision-makers blissfully unaware that they were wrong. The linear nature of the modeled processes also resulted in policy evaluation being envisaged only after the end of the entire cycle. There was no room for real-time feedback or two-way interactions with stakeholders or individuals during the cycle.

An intriguing aspect found in Lasswell’s policy cycle, which later vanishes from subsequent proposals, is the initial stage, referred to as *intelligence*. While this stage can be seen as merely gathering information, the term *intelligence* suggests a more collective thinking process, which is pertinent to our discussion in this book. Authors who later expanded upon the policy cycle concept took a different path, however, primarily drawing from management sciences. They focused on making specialized calculations of risks, costs, and benefits associated with each stage, developing lists and rankings of options, and narrowing down possible choices through feasibility studies and other methods.

In the works of Brewer (1974), Jenkins (1978), and Jones (1994), the guiding principle of this model is the logic of applied problem-solving. The individual stages within the policy cycle find parallels in the managerial approach to addressing issues (Howlett et al., 1995, p. 12) (Table 1.1).

What actually happens at each stage? First, *agenda-setting* refers to the process by which problems come to governments’ attention. This phase is regarded as the most crucial in the policy cycle; it is focused on how problems arise, or do not arise, as potential issues that the government should address. A policy agenda is the

Table 1.1 Five stages of the policy cycle and their relationship to applied problem-solving (Howlett et al., 1995, p. 12)

Applied problem-solving	Stages in policy cycle
1. Problem recognition	1. Agenda-setting
2. Proposal of solution	2. Policy formulation
3. Choice of solution	3. Decision-making
4. Putting solution into effect	4. Policy implementation
5. Monitoring results	5. Policy evaluation

process of reducing potential policy issues to a select few that the government finds significant. John Kingdon in his work *Agendas, Alternatives, and Public Policies* notes: “The agenda, as I conceive of it, is the list of subjects or problems to which governmental officials, and people outside the government, closely associated with those officials, are paying some serious attention at any given time” (2010). This process appears to be critical in the positivist model of policymaking as it determines which issues receive attention and resources from the government and which ones are left unaddressed. Independent experts were to recognize subjects that would become problems requiring further government attention. However, decision-makers could only focus on a limited subset of all available options. Howlett, Ramesh, and Perl claim that, in line with the original, deterministic understanding of the policy cycle, “agenda-setting [was] a virtually automatic process occurring as a result of the stresses and strains placed on governments by industrialization and economic modernization” (Howlett et al., 1995, p. 94). Creating a public policy agenda was supposed to be the task of professionalized structures and institutions focused on impartial assessment of reality and capturing problems, regardless of the pressure of interest groups, public opinion, or the media. Kingdon used a biological metaphor to describe this process, calling issue identification some sort of natural selection in which factors, such as experts’ agreement on a problem or technical feasibility, identify a few issues for the next phase of the policy cycle (Althaus et al., 2007, p. 51). Additionally, Herbert Simon remarked that the inclusion of an issue on the agenda is determined by its proper structuring. In order to distinguish between *ill-structured* problems (too general, unsolvable, i.e., the prevalence of poverty in society) and addressable *well-structured* problems (e.g., poverty in a particular region resulting from the collapse of a specific industry), Simon suggests that problems should be broken down into smaller *well-structured* issues, which will allow policy solutions to be found for them (Simon, 1973).

The second stage, *policy formulation*, refers to the process of generating options on what to do about a problem identified in the agenda-setting phase, namely: set objectives, identify the costs, create a list of solutions, estimate the effects of a solution, and choose the appropriate policy instruments (Cairney, 2020, p. 26). During this phase of the cycle, potential policies to address the problems and issues are identified, improved upon, and put into a formal plan or transformed into government programs. The distribution of power among various interests is reflected in the creation of policy alternatives, which impacts both the implementation of the policy and the results it produces. During this stage, the selection of available options is

limited to those that are acceptable to policymakers. The policy cycle theory says that doubtful scenarios, inconsistent with the applicable canons of conduct or unacceptable due to the existing political system, should not be developed (Howlett et al., 1995, p. 110). The core issue in this phase is to clearly define and evaluate the advantages and disadvantages of the possible options. During this phase, public servants identify and examine data and evidence in various forms, including research reports, expert opinions, and feedback from stakeholders. Public officials then evaluate the evidence regarding the different options and formulate a proposal that identifies the option(s) that can proceed to the next stage, if any. To search for solutions effectively, it's important to identify not just the actions that are technically capable of resolving a policy problem, but also those that are both politically acceptable and administratively feasible. The selection can be quite discretionary; as Carlsson puts it: "If significant actors in the policy subsystem believe that something is unworkable or unacceptable, this is sufficient from its exclusion in further consideration in the policy process" (Howlett et al., 1995, pp. 112–113).

The formulation stage includes not only the consideration of what needs to be done, but also the alignment of potential policy tools with policy problems and the selection of instruments which allow the plan to be executed. The options available to the government are, after all, limited by the available instruments. The policy instruments include:

- (1) regulatory instruments: legislation, economic and social regulations, deregulation and allowing self-regulation in selected areas of social life (Weingast, 1981), standards imposed by formal standards bodies, administrative decisions, creation of new institutions, and the launching of public service programs. Legal instruments are the traditional way of writing government policy; they guarantee that the proposed policy can be translated into action.
- (2) economic instruments: taxes, tax incentives, fees, and user charges (Brunori, 1997), certificate trading, procurement policies, subsidies, and grants. State-owned enterprises, direct provision of the public services, and/or public–private partnerships may play a significant role here as well (Howlett et al., 1995, pp. 113–134).
- (3) cooperation instruments: voluntary agreements, producer and consumer associations, and cooperation with interest groups or think tanks (Fraussen & Halpin, 2017). These instruments are often funded by the private sector or grassroots financing, but the government can play an active role in their development by granting them tax breaks or subsidies and inviting them to negotiate public policies as partners.
- (4) information instruments: public information campaigns (Weiss & Tschirhart, 1994), labeling schemes, reporting requirements, advice services by commissions and ad hoc expert councils (Crombez & Hug, 2000), and technology transfer (Fig. 1.2).

The *decision-making* phase of the policy process takes place when any of the options that have been proposed and refined during the previous stages are approved as the official plan of action. The choice is made from a relatively small number of



Fig. 1.2 Five-stage policy cycle model

well-developed options and tools refined in the previous stages while considering all possible effects for solving the problem in accordance with the rationalist paradigm. It is also a choice between taking action and not taking action as “decisions can be ‘positive’ in the sense that they are intended, once implemented, to alter the status quo in some way, or they can be ‘negative’ in the sense that the government declares that it will do nothing new about a public problem but will retain the status quo” (Howlett et al., 1995, p. 139). During the decision-making stage of the policy cycle, the number of policy actors involved decreases significantly. This is because, except in cases such as referendums, only those with political power are included in the group of policy actors entitled to make a binding public decision. In other words, most often the policy actors are people holding formal positions in the government or chairmen of parliamentary factions, and in some cases representatives of the judiciary. We omit here situations when political systems are facades. An example of such systems was in communist countries where the government and parliament only formally approved decisions that were made by the leadership of the ruling party. In parliamentary systems, however, decision-makers, despite their position, cannot make whatever decisions they wish.

From the very beginning of policy sciences, the establishment of public decision-making based on rational reasoning was one of the most important goals. For this reason, early analysts tried to transfer decision-making models developed in the field of business management to policymaking. Business management, in a simplified way, showed the possibility of maximizing expected outcomes based on rational analysis and choices made. The marketplace behavior of buyers and sellers, who were seeking the best utility from their limited resources by minimizing costs and maximizing benefits, was considered a process that could be translated into a procedure to select the most efficient means possible to achieve clearly defined policy goals. (Howlett et al., 1995, p. 144). This model assumes that maximal outcomes can be achieved through the ordered gathering of relevant information which in turn allows the objectively best alternative to be identified and selected (Weiss, 1977). Howlett described this process, claiming that:

Decision-makers are assumed to operate as technicians or business managers, who collect and analyze information that allows them to adopt the most effective or efficient way of solving any problem they confront. It is for its ‘neutral’ technical application to problem-solving that this mode is also known as the ‘scientific’, ‘engineering’, or ‘managerialist’ approach (Howlett et al., 1995, p. 144).

Despite its weaknesses, when confronted with real political life and the emergence of competing models (e.g., Incremental, described by Dahl and Lindblom (1953), Mixed-Scanning proposed by Etzioni (1967) or the Garbage-Can Model (Cohen et al, 1972)), the rationalist model of decision-making has been the base point of reference for policy scientists for years. The analysis of this rational-comprehensive model was undertaken by, among others, John Forester, who described the conditions that must be met for this model to work:

- (a) the number of *agents* who are decision-makers has to be limited, possibly to as few as one person: a utility-maximizing, economically rational actor—the decision-maker;
- (b) the organizational *setting* for the decision must be simple and insulated from the influences of other policy actors (e.g., the decision-maker’s office);
- (c) the *problem* has to be well defined, especially its scope, time horizon, value dimensions, and chains of consequences;
- (d) the available *information* must be as close to ideal as possible, which means it must be complete, accessible, and comprehensible;
- (e) decisions must be made carefully and without haste: *time* must be infinitely available to the decision-makers to consider all eventualities, incidents, and their expected consequences (Forester, 1984; Howlett et al., 1995, p. 156).

Once a public issue has gained enough public attention and been placed on the policy agenda, different alternatives are considered. The government can then establish policy objectives and choose a path to achieve them. Next comes the phase known as *policy implementation*, where the decision is put into effect. Typically, policy implementation involves a list of actions to accomplish the desired objectives, including but not limited to allocating funds, assigning staff, and developing operational guidelines to ensure the policy’s success. According to the positivist model,

Table 1.2 Relationship of policy goals and means in policy implementation
(Source Howlett et al., 1995, p. 173)

		Policy Tools	
		Consistent	Inconsistent
Policy goals	Coherent	Optimal	Ineffective
	Incoherent	Misdirected	Failed or suboptimal

a dedicated group of administrative employees is involved in the implementation of the policy. This stage generally involves the participation of multiple bureaucratic agencies at various levels of government, including national, regional, and local levels. Under the rational paradigm, implementation was regarded as a purely technical problem to select the means best suited for the objectives pursued. This initial assumption was supplemented with two approaches: a top-down perspective, where effectiveness was defined as keeping to the original intent of the ratifiers, and a bottom-up approach, where experience was monitored in detail, data from the implementation of the plan was collected, and effectiveness was viewed as the bureaucracy's adaptive behavior to make the best use of the activities carried out to improve the environment (Howlett et al., 1995, pp. 163–164).

Implementation also requires the selection of an appropriate instrument from those analyzed in the *policy formulation* phase. Rationalist policymaking assumes that tool selection is a technical exercise, based on evaluating the advantages and disadvantages of the available tools in relation to the assumed goals and choosing the least costly and most effective one. According to Howlett:

Instrument choices, to be effective, must be closely and carefully related to policy goals, and (...) any new goals and tools must also be carefully integrated with existing policies if implementation is to succeed. New and old goals must be coherent, in the sense of being logically related, while new and old instrument choices must be consistent, in the sense of not operating at cross-purposes (Howlett et al., 1995, p. 172).

Relations between policy goals and tools in the implementation phase are shown in Table 1.2.

The last stage of the cycle, which also paves the way to finding new issues for the public agenda, concerns *policy evaluation*. At this stage, the government assesses how the policy is working, tracks the reactions of the public and influential organizations to the policy, and identifies support or opposition to the program(s). The aim of this phase is to verify if the policy was appropriate, implemented correctly, had the desired effect in terms of perceived intentions and results, and to generally estimate the extent to which the policy was successful (Cairney, 2020, p. 27). The positivist approach to policymaking viewed this kind of assessment as a neutral, technical process, being described by David Nachmias as “the objective, systematic, empirical examination of the effects of ongoing policies and public programs have on their targets in terms of the goals there are meant to achieve” (Howlett et al., 1995, p. 178). Developing neutral standards to evaluate success or failure and using quantitative methods based on adequate and acceptable measures turns out to be an extremely complex task, however. What is a measure of success for some does not meet these criteria for others. The most important evaluation tasks are usually carried out by government

departments, specialized executive agencies, or committees dedicated exclusively to evaluation activities. In the US, for example, these works are performed by staff in the General Accounting Office, the Congressional Budget Office, the House and Senate Budget Committees, and the legislative and appropriations committees. On the other side of the globe, the Australian National Audit Office plays an important role in evaluating implemented policies, with its work complemented by the Office of Evaluation and Audit in the Department of Treasury and Finance. Reports, periodic audits, recommendations, and policy adjustment plans are the expected outputs of these institutions.

Types of evaluation conducted can be matched with different resources and various steps in the policy cycle. Take the example of the Australian Department of Finance, which distinguished four types of evaluation in its documents (Althaus et al., 2007, pp. 182–183):

- (a) *Appropriateness Evaluation*, which helps decision-makers to determine whether a given program really meets social needs and falls within government priorities. The subject of the investigation is whether the objectives align with the government program and people's needs.
- (b) *Efficiency Evaluation*, which examines how well inputs are used to obtain a given output and whether funds are spent sparingly. The goal is to see if the same quantity and quality of outputs can be produced in a different way at a lower cost. The extent to which program inputs are minimized for a given level of program outputs, or to which outputs are maximized for a given level of inputs, is examined.
- (c) *Effectiveness Evaluation*, which asks whether the program is producing valuable results and whether the outcomes justify the expenses. The extent to which program outcomes are achieving their stated objectives is examined.
- (d) *Meta-Evaluations*, which assess the evaluation process itself. It examines whether the evaluation practices used by the agencies are professional, whether these practices are sensitive to the social and physical environment of the program, and whether the resulting reports are used by public institutions in a practical manner.

The process of policy evaluation often goes hand in hand with the temptation to predict future events based on current trends. Forecasts are seen as a necessary precursor, and often as a determinant, of policymaking (Wieland & Volters, 2013). Predictive theories became a sort of Holy Grail in policy research. Analysts use their forecasts to project the consequences of particular policy decisions on certain policy targets. Some researchers claim they can predict future events with great precision. In order to make such precise predictions, variables are operationalized. Their categories and values are subjected to quantitative analyses to reveal the relations between them. Predictions relevant for monetary and fiscal policymaking are regularly published by policy institutions such as the American Board of Governors of the Federal Reserve System and the European Commission. In the US, the Congressional Budget Office publishes its Budget and Economic Outlook. Annual predictions are made public in January and presented to the House and Senate Budget

Committees. Furthermore, the Congressional Budget Office releases an evaluation of the President's budget plans for the upcoming fiscal year (Wieland & Volters, 2013). In the European Union, the European Commission provides forecasts in a publication called the European Economic Forecast. Twice annually, projections are released along with additional updates for major member states. Predictions are calculated for specific countries, the eurozone, the European Union, prospective members, and some non-European Union nations. The projections encompass 180 variables. Country-specific predictions are generated by country desks utilizing various economic models and econometric methods (Wieland and Volters 2013). In recent years, various quantitative models have been made to predict threats such as pandemics, coup d'états, ecological disasters, and the probability of winning an election by political leaders. The government of the United Kingdom is no different from the US or EU institutions, making its forecasts with the help of Office for Budget Responsibility (OBR). These forecasts include predictions of the GDP impact of fiscal stimulus legislation, inflation forecasts, predictions of the impact of interest rate reduction on the economy, and many others. To cite an example dating back to 1958, which remains pertinent even now:

In the case of employment policy, it is necessary to obtain some idea as to the factors controlling next year's or next quarter's employment level. If this leads to a predicted employment, which is considered too low, the policymaker will try to find measures in order to increase the level; and then it is necessary to predict the effectiveness of alternative measures, i.e., of the reaction of variables, which are not controlled to changes in controlled variables. This leads to a plan of actions, which may be followed or abandoned in the light of subsequent events. (Theil, 1958)

This rational-comprehensive model was, as mentioned above, widely applied in social policy development by several federal governments in Australia since the 1980s. The policy cycle framework has become an important component of official policymaking, used by Parliament of Australia (Marsh, 2023), Australian Public Service Commission (Introduction to Delivering Great Policy 2023), the State Pollution Control Commission (Everett, 2003), state governments of Queensland, Victoria, Western Australia (Goswami & Gerritsen, 2021), and many others. First published in 1998, *The Australian Policy Handbook* (Althaus et al., 2007) has become one of the most important readings for administration employees, and Australia is perceived as an example of the success of rational policy studies, or—as critics say—the technocratic approach to policymaking. Examples of success in this field include: a higher education contribution scheme whereby students pay back tuition fees as a percentage of post-graduation income (Edwards, 1997), a child support program that utilizes the income taxation system to transfer payments from non-custodial parents who earn income to sole parents (Edwards, 1998), a policy response to HIV/AIDS based on partnership, community engagement, and bipartisan support (Fitzgerald et al., 2019), the water markets in Murray-Darling Basin—a cap-and-trade system where a specified volume of water can be extracted or diverted on an annual basis (Horne & Grafton, 2019), and the National Competition Policy implemented in 1995 through which cooperation of central, state and territory governments introduced competitive

market forces to sectors protected by regulation or government ownership, especially gas and electricity infrastructure, as well as water and transport (Fenna, 2019).

The widespread adoption of the rational approach has brought many benefits. It allowed for the systematization of administrative work and, at least partially, the independence of public policies from party ideologies, populism, and short-term popularity-seeking. The rational approach also allowed political analysis to be embedded with measurable effects, gains and costs, formal decision-making models, and strong empiricism focused on objective analysis. However, this approach has raised some doubts. Despite the undeniable successes of rational, analysis-driven policymaking, many authors have pointed out that it is impossible to encapsulate all social factors within one model, and that this type of approach is only suitable for analysis under certain conditions: political consensus, minimal ideological polarization, well-defined problems, social stability, and low international tensions (Dryzek, 2002). Furthermore, in its purest form, this model relies on trust in expert knowledge among public servants, as they were responsible for most of the work. Australia's unique geographic location and economic development made it an ideal place to apply the rational-comprehensive model. As an island nation, Australia has a distinct political and economic landscape, which may have allowed for a more controlled and stable environment to apply this approach. Additionally, Australia's high level of economic development, particularly in the latter half of the twentieth century, provided a solid foundation for evidence-based policymaking, as data and expertise were readily available. Australians also have a relatively high degree of social trust in the government, which may have allowed for a greater level of cooperation and consensus-building among different stakeholders. This level of trust may have facilitated the implementation of policies based on rational analysis and expert knowledge, as the public was more willing to accept and support policies that were evidence-based.

It is important to note, however, that even in Australia the rational-comprehensive model was not without its challenges and limitations. The complex nature of policy issues, competing interests and priorities, and changing social and economic circumstances all posed significant challenges to the effective implementation of this approach. While this model can provide a systematic and structured framework for decision-making, it may not take into account the unique cultural, social, and economic factors that shape each policy issue. Additionally, the extended use of expert knowledge may lead to a certain level of elitism and exclusivity in the policymaking process, resulting in the perspectives and experiences of marginalized communities or the *silent majority* of citizens being overlooked. Furthermore, the rational approach may struggle to adapt to rapidly changing circumstances or emerging issues that do not fit neatly into predefined categories. In such cases, a more flexible and adaptive approach may be required, which considers the evolving nature of policy issues and the need for continuous reevaluation and adjustment.

In summary, while the rational approach to policymaking offers significant benefits, acknowledging its limitations is equally important. Using the policy cycle model for the examination of policy processes allows us to organize the analysis, identify key variables, and see the strengths and weaknesses of the examined processes. Using

routine and formalized procedures for policy decision-making helps to streamline the process, reduce errors, and improve the overall effectiveness of policies. This approach also allows policymakers to make decisions based on data-driven analysis rather than intuition or personal biases. While this comprehensive model proves to be a useful tool, it is not entirely able to address the issues of conflict and dynamics that have recently gained attention in the social sciences. According to this model, problems are considered to have an *objective* existence and are simply waiting for governments to recognize them. But policy issues are complex and multifaceted, so treating them as programmed decisions can oversimplify the situation. Therefore, policymakers must also consider the nuances and complexities of each policy issue, as well as the potential impact on various stakeholders, to ensure that they make well-informed decisions. Due to this, there has been a notable increase in the popularity of an approach known as post-positivism, in which the principal components are twofold: the re-emphasis on the *human* facets of policy and the inclusion of a broader range of participants. Together, these components contribute to a more adaptive and responsive policymaking process that better addresses the needs of society at large (deLeon, 1988, pp. 112–113). The policy cycle analysis method is therefore not to be rejected but supplemented and treated non-linearly.

1.6 Post-positivist Approaches to Policymaking

In recent years, global politics has been unlike the ideal, stable environment that could be subjected to systematic analysis. It can be more accurately characterized as a state of *radical uncertainty*, as described by King and Kay (2020). The 2020 COVID-19 pandemic is a perfect example of this, being the most disruptive global event since the Great Depression and World War II. The pandemic has brought various types of uncertainty, including: organizational shortcomings in health service, lack of consensus among experts and variations in expertise, the possibility of over- and under-reacting, inconsistent and possibly ill-considered decisions restricting civil liberties, questioning trust in public administration, insufficient technical know-how, the temptation for governments to experiment with different kinds of interventions, and many others (Brik et al., 2021, p. 1). The simultaneous increase in international tensions resulting in the spread of conflicts, including Russia's military aggression against Ukraine in 2022, further deepened the global crisis. Both the pandemic and global conflicts have caused a profound global economic downturn, leading to unemployment, food insecurity, and increased global poverty. The economic consequences are vast, with most of the world falling into recession, public debt levels soaring, and future growth prospects looking bleak. This prompted inquiries into why mathematical algorithms, based on the assumption of rational economic agents and efficient markets, have failed to predict the crises. These events have also accelerated the spread of radical ideologies, aimed at remodeling the lives of entire societies, and trends toward populism or totalitarianism.

These unexpected events have raised questions about the effectiveness of using behavioral or natural scientific methods to make policy decisions, particularly in assigning probabilities to alternative scenarios. The uncritical use of models drawn from physics in the analysis of social phenomena has become the object of widespread criticism because it reduces human behavior to purely individual actions, lacking wider social context. Adrian Pabst, a critic of this method, recently described the potential misuse of probabilistic models in shaping public policy during the coronavirus pandemic:

Problems with probabilistic models and their misapplication are not new. A key factor in the 2008-2009 global financial crisis was an excessive and uncritical use in financial services of models drawn from physics (...) In the past ten years, we have come a long way in understanding the limits of modeling financial and economic processes based on human behavior that is supposedly determined by general physical laws or individual psychological dispositions. Modeling economic and financial decisions either ignored behavioral aspects altogether or reduced them to purely individual actions that can be 'nudged'. Missing from such models is the wider social embedding of individual and group choice, as well as a recognition of the intertwining of economic with social scarring in the event of severe shocks.

Some scientists (especially those who seek to popularize it, such as Richard Dawkins and Steven Pinker) assert that scientific inquiry leads to incontrovertible truths when in reality scientific findings remain conjectural and fallible even as our knowledge and understanding continue to improve. Here it is worth remembering two rival scenarios based on different models: one by the team at Imperial College predicting more than 500,000 deaths in the absence of a lockdown and the other by a team at Oxford University claiming significant levels of herd immunity. Key to genuine progress is the capacity to question established facts and entrenched methods in an attempt to correct errors – for example, modeling assumptions linked to projections of death rates and the spread of the virus as more data become available. That, in turn, is vital in adapting both lockdown rules and public policy in relation to the economic and social impact of Covid-19. (Pabst, 2021)

It seems that comprehensive rationality, in the face of social turmoil and uncertainty about tomorrow, has failed epically. The most critical policy challenges are plagued by extreme incertitude, making it difficult to establish clear objectives. Trends in national policies are often varied and in direct conflict with one another, further adding to the complexity of the issues at hand. Due to the interconnected nature of policy components, predicting outcomes is a daunting task since there are too many possible scenarios to account for. Additionally, policy processes are not straightforward and operate simultaneously, making it difficult for policymakers to have a complete understanding of the policy environment. With an overwhelming amount of information to sift through, policymakers are unable to give adequate attention to all of the issues, potential solutions, and decision-making criteria. Ultimately, the volume of information and complexity of the problems at hand exceed the capacity of policymakers to process effectively. In the face of political reality, comprehensive policy analysis often brings results that policymakers cannot use. Furthermore, the technocratic image is inaccurate to what analysts actually do. In the words of John S. Dryzek, a well-known critic of positivist approach,

Technocratic policy analysis proceeds in the image of an omniscient benevolent decision-maker, a situation in which there is no politics (Majone calls this “decisionism”) (...). In a complex political system, instrumentally rationalistic policymaking is possible at best only in rare moments of consensus amid crisis, or in the occasional area insulated from more pluralistic control (such as diplomacy or national security policy). (Dryzek, 2002)

Douglas Torgerson recognizes that positivist policymakers believe human interactions can be examined, analyzed, and manipulated using principles of scientific investigation and technology, ultimately allowing “knowledge to replace politics.” Moreover,

The place of the analyst in society as a social being dealing with others of his kind tends not to be raised as a point of discussion. Removed from society, social science produces the knowledge from which to fashion an effective social technology, and the analyst – as both scientist and technician – becomes one who performs remote operations on an essentially alien object. (Torgerson, 1986)

But should rationality as an attachment to reasonable judgments be completely rejected? Absolutely not. Simon (1955) was among the first to challenge the rigidity of the rational/scientific model, doing so not long after its inception. Recognizing that politicians and administrators may not possess the ability to thoroughly evaluate all potential options and arrive at the *correct* decision, he proposed a behavioral alternative model based on the concept of bounded rationality (Everett, 2003). According to this model, decision-makers would prioritize the need for a quick and practical solution over pursuing all possible objectives and exploring every option. The bounded rationality model recognizes that decision-makers may not have the time, resources, or knowledge to make fully informed decisions and encourages them to use their intuition, experience, and heuristics to arrive at a workable solution. The model acknowledges the importance of balancing practicality with objectivity in decision-making and highlights the need to make reasonable and rational decisions based on the available information and resources. Simon argued that the constraints of decision-makers’ information-processing abilities did not lead them to the ideally optimized decision, but rather the *satisfactory* decision with lowered expectations, having understood that access to perfect information is impossible (Simon, 1955). Expectations could rather be satisfied, than optimized (Forester, 1993, pp. 6–7).

The post-positivist approach, apart from Simon, has its roots in the ideas of pluralists like Charles E. Lindblom and Robert A. Dahl. They recognized that the process of problem recognition is primarily a socially constructed process as it involves the creation of agreed-upon definitions of normalcy and what constitutes an undesirable deviation from this norm. Therefore, problem recognition is not a simple, mechanical process of identifying challenges and opportunities but is instead a process where ideas, sentiments, and values influence the behavior of citizens, governments, and non-governmental actors. Lindblom suggested that policymaking, rather than being a scientific process, was akin to the science of “muddling through,” arguing that policymaking was an inherently political process (Lindblom, 1959). The outcome of a “power play,” Lindblom claims that policy content and the decision-making process were proceeded by a series of negotiating steps between groups using a variety of resources and techniques to reach a solution.

Post-positivism's main objective is therefore to eliminate the illusion of certainty and acknowledge the existence of divergent opinions or viewpoints (Dryzek, 2002). By doing so, post-positivism aims to prevent the adoption of simplistic and one-dimensional approaches to policymaking, which can lead to ineffective outcomes. This approach recognizes that policymaking is a complex process that involves multiple stakeholders with differing interests and values. Therefore, post-positivism advocates for a more nuanced approach to policymaking, where different perspectives are considered and conflicting views are reconciled through open dialogue and deliberation. In this way, policies can be developed that reflect the diverse needs and aspirations of society, rather than be based on a narrow and rigid understanding of the world.

The concept of bounded rationality refers to the practical impossibility rather than the logical impossibility of exercising perfect rationality, where we observe the constraints on an actor's information-processing abilities. Simon argues that most people are only partially rational and that the remaining part of their actions is either *emotional* or *irrational*. Additionally, he asserts that an individual who wishes to behave rationally must consider not only the objective environment, like the amount of information they have access to and the limited time they have to make decisions, but also the subjective environment, including their cognitive limitations, perceptual processes, abilities to discover alternative options, as well as intuitions and emotions (Simon, 1998, p. 266).

Amos Tversky and Daniel Kahneman describe the observed deviations from perfect rationality, claiming that in real-life decision-making people rely on "heuristic principles which reduce the complex tasks of assessing probabilities and predicting values to simpler judgmental operations that are quite useful, but sometimes they lead to severe and systematic errors" (Tversky & Kahneman, 1974). Instead of logical calculations, people often use *decision heuristics* (general and simple rules) that allow for a quick reaction based on accumulated experience. For example, the availability heuristic refers to the fact that people prioritize data they have easier access to; meaning if you recently had a traffic accident, you are more likely to overestimate the probability of being injured while driving. As a result, as a policymaker, you are more likely to advocate for stricter speed limits or punishments for speeding, even if your case was not representative. The behaviors described in this way, often referred to as cognitive shortcuts or cognitive biases, are sometimes equated with subjecting policymaking to herd thinking or populism. However, as Gigerenzer points out, according to modern psychological research, this is how the standard decision-making process of actual humans works, in contrast to "heavenly beings equipped with practically unlimited time, knowledge, memory, and other unlimited resources" who would be needed to fulfill the vision of fully rational policy (Gigerenzer, 2001). From a policymaking perspective, it is particularly important to include processes in which decision-making heuristics are integrated with thinking styles specific to the relevant communities. Cairney lists a few of these heuristics, namely: *framing effects* based on *pre-reasonable* emotional and moral judgments, a *need for coherence* to identify patterns and causal relationships in social life, groupthink, collective judgments, and status quo bias (Cairney, 2020, p. 60).

Another important element of the post-positivist approach is the concept of social learning. This viewpoint emphasizes that through the process of learning, both individuals and collective groups can deepen their comprehension of the values, desires, and priorities of their counterparts (Innes & Booher, 2010). Social learning can be understood as a collective, public process characterized by innovation, communication, shared understanding, and a value-driven recognition of the learning that has taken place. Evolved as a counter to the rational policy model, social learning theory emphasizes social norms, customs, institutions, and ideas as explanatory factors and potential remedies for the shortcomings of numerous policy programs in achieving sustained, noticeable change. By examining the context and continuity of knowledge and its holders, social learning facilitates the transfer of ideas, information, and innovation (Holden, 2008).

Examining social learning within planning and policy offers a framework to distinguish learning from other forms of change. The difference between a *simple change* and *learning that can endure* can be judged as an improvement which proves its value in social practice by leading to the common good. Heclo (1974, p. 306) defines the policy learning process as a lasting change in behavior stemming from experience. Typically, this change is seen as a reaction to a recognized stimulus. Adapting practices collaboratively to address unforeseen conditions and learning through experience are viewed as pathways to effectiveness. While Heclo perceived policy learning as primarily an implicit, often unintentional activity, other theorists describe it as a conscious process of adjustment or formalizing knowledge. For instance, Hirschman, when evaluating innovation and implementation in development, argues that policies cannot be pre-determined. Instead, they should evolve incrementally, with policymakers adapting, reassessing, and responding to emerging challenges and opportunities as they present themselves (Holden, 2008).

1.7 The Question of Values and Group Identity

Positivism, as we remember, assumes the separation of facts from values when creating public policies, uses standardized and structured processes as a basis, and focuses analytical efforts on managing policy issues as programmed decisions. The consequence of this adopted objectivism was the division of facts and values. According to this view, policy analysis based on empirical knowledge and scientific objectivity should be value neutral. Positivists identify emotional states with *values* and distinguish them from *facts* related to objective reality as reflected in the policymakers' minds. This approach aligns with the Weberian tradition that prioritizes goals over values and views individual preferences of policymakers in terms of goals. Lasswell, especially in his early research, rejected the idea of prioritizing values in policy analysis, as he believed that it only increases the influence of decision-makers. He advocated for *management* over ethical knowledge and learning *from practice* rather than moral principles (Zittoun, 2019). The disappointment of this situation was described by Peter deLeon in the late 1980s: "For years, policy scholars have argued

that technical, value-free information should drive policy decisions. In retrospect, this has been conceived widely to be naive, perhaps even wrong" (deLeon, 1988).

Post-positivism moves away from this radical separation. At the turn of the 1970s and 1980s, policy scientists became increasingly convinced that "policy is inherently a normative practice, which is why ethics is central to the study of public policy (...). Simply put, values matter" (deLeon, 1988). Inevitably, both elected and appointed public officials must make normative decisions. As David Easton famously stated, "Politics involves the authoritative distribution of values" (Sullivan & Segers, 2017). The evolution of thinking about values was associated with the emergence of new economic concepts that sought motivation for participation in economic processes beyond the simple maximization of individual well-being. It became clear that purely instrumental rationality, which does not ask questions about ethical values relevant to society, becomes a caricature of itself. Amartya Sen believed that "it is difficult to accept that any kind of goals—no matter how bizarre—must be accepted, i.e., it does not compromise the rationality of those who pursue them" (Sen, 1998, p. 71). Sen is very strong about this, writing that the *Instrumental Rationalist* is an expert in decision-making who, seeing a man cutting off his fingers with a blunt knife, rushes to advise that it would be worth choosing a sharper tool as it is better suited to his purpose. Even this advice could be useless if the man wanted to cause himself as much pain as possible by cutting off his fingers (Szarfenberg, 2002). It follows that a man who is guided only by economic rationality is a social fool, because as humans we are guided not by just one utility ranking, but by several—including an ethical one. We can assess a particular situation considering both our personal benefits and the greater social good. Actual decisions do not strictly adhere to the maximizing logic found in rational choice theory, and our actions are not motivated entirely by selfish considerations: "Choice may reflect a compromise among a variety of considerations of which personal welfare may be just one" (Sen, 1977). Sen suggests that decision-making encompasses a spectrum of elements including ethical values, individual identities, solidarity with others, notions of fairness and equity, and self-interest. All these elements are weighed and balanced to reach a decision suitable for the given situation. In essence, Sen highlights that the decision-making process is intricate, intertwined with both personal and societal considerations, and that this interplay profoundly influences the decision's outcome. Talcott Parsons echoes this sentiment, positing that "the rationality of instrumental action depends to a large extent on non-instrumental actions consistent with social norms or roles" (Turkel, 2005).

Thus, the post-positivist movement in policy studies contends that "systematic investigation of the normative dimensions of policy issues should have an important place in the practice of policy evaluation" (Amy, 1984). Some of the principle works on this subject include Frank Fischer's *Politics, Values, and Public Policy*, Duncan MacRae's *The Social Function of Social Science*, Charles W. Anderson's *The Place of Principles in Policy Analysis*, and Douglas J. Amy's *Why Policy Analysis and Ethics are Incompatible*. The common thread of these works is that policymakers should engage in ethical analysis because it can serve as a safeguard against the undetected influence of personal values on policy decisions. By taking an explicitly

normative approach, policymakers can ensure that all normative considerations are made transparent and visible to everyone involved in the decision-making process. This means that normative arguments and assumptions can be identified and labeled as such, which alerts policymakers to their presence and helps prevent them from being inadvertently incorporated into policy deliberations. On the other hand, so-called *value-free* analyses tend to conceal normative decisions within the underlying assumptions of the analysis, such as in the choice of options selected for analysis: “The normative decisions present in so-called ‘value-free’ analyses tend to be **hidden** deep in the assumptions of the analysis” (Amy, 1984). As a result, these normative assumptions could go unnoticed and be more likely to influence the policy debate without proper consideration. Policymaking, just like contemporary science, is consistently imbued with values: individuals participating employ their professional and personal principles and perspectives to select the most effective methods, formulate research questions, establish concepts and indicators, and evaluate the consequences and policy implications of the outcomes (Cairney, 2020). Therefore, if it is true that policy research inevitably involves a number of normative assumptions, then it is better to acknowledge and analyze these assumptions openly rather than ignore their existence. Engaging in ethical analysis and taking a normative approach can help policymakers make more informed and transparent decisions and ensure that normative assumptions do not influence policy debates without proper consideration.

Yet, one must inquire: which ethical values and norms lie at the heart of the post-positivist debate? A prevailing perspective, particularly in newer publications, suggests that values are not static or universal but are context-specific and evolve over time. Bevir and Rhodes (2003) argue that social, political, and cultural factors can plastically shape values and beliefs, making them contingent upon the historical and societal context in which they arise. According to this approach, ethical issues are relative, and often conflicting, as they depend on the particular cultural group. Consequently, policymakers should be aware of the diverse values and beliefs held by different groups and acknowledge that their perspectives may evolve over time. To foster a more inclusive policy process, researchers like Stone (2002) recognize that a universally applicable ethical standard may not be feasible—and can even be dangerous—given the diverse values and beliefs present in society. Innes and Booher (2010) advocate for context-sensitive and adaptive policymaking that is responsive to the specific needs and values of the communities involved and may change as needed. Wayne Parsons, criticizing this approach, summed it up: “You show me your ‘truth’ and I’ll show you mine. You respect my ‘values’ and I’ll respect yours” (Parsons, 2020).

This approach, which integrates epistemologies like constructivism, interpretivism, and critical theory, dismisses the notion of a singular objective truth or moral value that guides policy decisions. It posits that such decisions are shaped by a complex interplay of cultural values, interests, power dynamics, and social contexts. In adopting constructivist and interpretivist viewpoints, this perspective recognizes that policy issues are primarily molded by social constructs, collective understandings, and cultural identities, rather than objective ethical values. A notable

representative of these views, Fischer (2003) emphasizes the importance of pluralism especially the inclusion of marginalized or minority groups, and acknowledges that multiple viewpoints and ways of knowing can contribute to a richer understanding of policy issues.

This can be seen in the context of urban planning and development. Gentrification, the process of neighborhood transformation due to an influx of higher-income residents and businesses, can lead to displacement and social inequality, disproportionately affecting low-income and minority communities. Traditional top-down urban planning approaches often prioritize economic development without adequately addressing the needs and concerns of these communities. Frank Fischer highlights the importance of including these marginalized voices in the policymaking process to develop more equitable and just urban planning policies. One example could be the implementation of participatory planning processes, wherein residents are actively involved in shaping the development of their neighborhoods. In such processes, community workshops, public meetings, and other engagement activities can be organized to gather input and insights from residents who are directly affected by the changes in their neighborhood. By involving these communities in decision-making, policymakers can gain a better understanding of their concerns, needs, and aspirations, resulting in more inclusive and equitable urban development policies.

Hajer (1995), on the other hand, examines the role of language, discourse, and power relations in shaping environmental policy. He argues that policy decisions are influenced by the social context, political interests, and cultural values in a decisive way. In his view, environmental policy is not only based on scientific and technical knowledge but also on discursive struggles among various actors, each with different interests, values, and power. An example from Hajer's work is the concept of *ecological modernization*. This discourse emerged in the late 1980s and early 1990s, primarily in Western Europe, and suggests that economic growth and environmental protection can be compatible. This idea was promoted by various actors, including large industry representatives and environmentalists. Ecological modernization gradually became a dominant discourse, sort of a *new culture* of public policy, shaping policy agendas and informing the development of new regulations.

Yanow (2000) explores similar ideas, stressing the importance of interpretive policy analysis, which focuses on understanding the meanings, values, and beliefs that underpin policy decisions. She highlights the role of social context and power dynamics in shaping policy outcomes. An illustration of this can be seen in public housing policy. When policymakers decide to implement a new public housing program, they might create specific language and images to convey certain ideas and values about the program's purpose and target population. For instance, they might emphasize the need to provide *affordable housing* or *decent living conditions* for *low-income families* or *vulnerable populations*. Through interpretive policy analysis, researchers can examine how these terms and phrases are understood and used by different stakeholders (e.g., policymakers, housing authorities, residents, and the public).

Nonetheless, a different perspective in modern policy research suggests that not all values can be considered mutable cultural constructs. Authors such as Peter deLeon,

Adrian Pabst, Wayne Parsons, and Douglas J. Amy believe that there is a certain set of ethical principles that keeps society together, regardless of cultural variations and contesting particular values. Here “classic principles as authority, justice, and efficiency can be understood as necessary considerations in any rationally defensible policy appraisal” (Anderson, 1979). This approach, which can be called Universalist, assumes that there are fundamental ethical principles or moral values shared by all human beings. This in turn can serve as a basis for policymaking and promote social cohesion and harmony. According to Amy (1984), “policy analysis that avoids examining the inevitable clashes between values (...) runs the risk of being largely irrelevant to contemporary policy choices.”

This perspective, which is characteristic of classical political philosophy, saw a substantial decrease in prominence within public policy discourse during the twentieth century. Malbin observes that in a “post-Weberian world, where ‘facts’ and ‘values’ are thought to have separate cognitive foundations, politicians were embarrassed about basing political choices on principles of justice or ‘basic values’” (Malbin, 1980). However, we now see a resurgence of this approach, though not without some resistance. The renewal of interest in this topic is observed, among other things, by research in evolutionary psychology, an interesting example of which is the Moral Foundations Theory (MFT) proposed by Haidt (2013). In brief, the theory suggests that humans have innate, and universally accessible, psychological foundations for morality. According to MFT theory, morality is based on five (later extended to six) primary foundations which are:

- (1) *Care*; related to our ability to feel empathy and compassion for others, especially those who are vulnerable or in need. It drives the desire to protect and care for others, while avoiding causing harm.
- (2) *Fairness*; concerning the concepts of justice, equality, and reciprocity. It is based on our innate desire for fair treatment and cooperation, as well as our aversion to cheating or exploiting others.
- (3) *Loyalty*; centered around identity, allegiance, and loyalty. It reflects our inclination to support and defend our in-group (family, friends, or community) while opposing betrayal or disloyalty.
- (4) *Authority*; relates to our understanding of social hierarchies, respect for authority, and adherence to tradition. It highlights the importance of maintaining social order and respecting the roles and rules that govern our interactions.
- (5) *Sanctity*; involves the concepts of purity, sanctity, and the sacred. It is rooted in our desire to maintain physical and moral cleanliness, as well as our aversion to contamination, degradation, or anything considered taboo.
- (6) *Liberty*; accounts for moral concerns around individual autonomy and resistance to tyranny or oppressive control. Liberty is grounded in the human desire for self-determination and the aversion to being dominated or controlled by others.

In line with MFT theory, the ethical values pertinent to a particular community are not derived from systematic reasoning but emerge from a blend of innate factors, collective learning over time, social interactions, and group dynamics, all of which are well-established practices developed through evolution. Consequently, it is evident

that the norms guiding a specific community and the cultural identity founded on them are intimately linked to collective intelligence, a subject we will delve into in later chapters.

These moral foundations provide a comprehensive framework for understanding the ethical aspects of human cognition, conduct, and grasp of social norms. According to Haidt, these moral foundations are present in varying degrees in all human cultures and serve as the basis for the diverse moral systems we see around the world. Each culture constructs its own virtues, narratives, and institutions on top of these foundations, resulting in unique moralities that may emphasize different aspects of them (Haidt, 2013). By understanding the universal moral foundations, we can gain a deeper insight into the values and social norms that govern human behavior and social interactions across cultures. While the primary moral foundations themselves are considered stable and universal, the way they manifest and are prioritized within specific moral rules, customs, and values can change and evolve over time. Understanding the manifestation of these values across diverse societies is grounded in the empirical observations of their traditions and long-lasting institutions. It can be said that MFT is a theory operating at a higher level of generality than classical ethical theories. These kinds of ethical guidelines were expected by several policy researchers as “standards at the metapolitical level which any system of policy evaluation must meet and criteria of ‘metapolitics’—standards for judging the adequacy of any system of political judgment” (Anderson, 1979).

Let us consider an example. The significance of moral foundations in policymaking can be illustrated by the diverse policies and public responses to the COVID-19 pandemic worldwide. Pandemic created a situation where the governments had to rapidly develop policies to manage the spread of the virus and mitigate its impact on society. Successfully implementing health policies depends on understanding the moral foundations guiding specific social groups. Policies like stay-at-home orders, mask mandates, and later vaccination rollouts were primarily framed as measures to prevent harm and care for the health of the public. In the US, messaging from the Centers for Disease Control and Prevention (CDC) and public health officials often emphasized the prevention actions for protecting not only yourself, but saving other's lives (Centers for Disease Control and Prevention, 2023). Similarly, the psychological study conducted in Turkey showed that participants endorsed care, fairness, and purity foundations more than the remaining foundations in the face of a disease threat (Ekici et al., 2023).

In New Zealand, the government's campaign for citizens to ‘Unite against COVID-19’ tapped into the loyalty foundation. Policies were often presented as a way for communities to show solidarity and protect their own. As evidenced, it has been highly effective in garnering community support in New Zealand (Wilson, 2020). The campaign helped develop a collective sense of purpose, and public opinion polls consistently showed more than 80% support for the government's actions (Jamieson, 2020).

Particularly noteworthy are the variations in public attitudes toward the moral dimensions of pandemic policies across different regions of the United States. In the states dominated by people with liberal views, for whom—according to Haidt's

theory—the most important moral values are care and fairness, strict compliance with all state recommendations related to pandemic restrictions, including strict lockdowns, wearing masks and vaccinations, prevailed. In California, Governor Gavin Newsom implemented stringent measures, with a strong urging Californians to receive vaccines, wear masks, and stay home to protect themselves and others, which met with general public support (Reston & Krieg, 2021). It was different in the states where the majority of people have conservative views, for whom the basic values are liberty and sanctity (both in the religious and secular sense). In Florida, governor Ron DeSantis resisted lockdowns and mask mandates, promoting personal choice and keeping the economy open. This stance appealed to a sense of individual liberty, which resonated with many in the state, leading to less compliance with federal health recommendations (Calvan, 2020). Florida became known for its open-for-business stance, and many Floridians, especially in more conservative areas, resisted mask mandates and other measures, viewing them as government overreach.

Also, when it comes to another feature specific to conservative morality, i.e., sanctity, research has confirmed the resistance of people expressing this feature to federal health policy. A study on the impact of American citizens' values on compliance with staying-at-home and wearing masks policy found that those who attached greater importance to sanctity were more resistant to the federal government's recommendations (Chan, 2021). For some religious people, opposition to Covid-19 vaccines was also motivated by vaccine ingredients (like pork) or concerns about guidelines for everyday conduct (such as not inflicting pain on cows) in their production. For certain religious groups, these ingredients are defined as unclean or harmful in essence and therefore can destroy the bodily purity and sanctity that a person is religiously bound to preserve (Trangerud, 2023).

The varied responses to health policies enacted to combat Covid-19 pandemic underscore the significance of considering the moral foundations that resonate with different social groups when formulating public policies. The success of policy implementation is in large part contingent upon honoring these ethical values. While some groups prioritize fairness and care, others may hold loyalty, authority, liberty, or sanctity in higher regard. Researchers in social morality, such as Haidt, propose that these values, although to varying degrees, are vital elements of social life across all cultures and geographical regions.

The issue of group identity is closely related to the ethical values shared in a community. As Elinor Ostrom argues, people have a natural inclination to form groups based on a common understanding of basic values that then become an integral part of their identity. Citizens generally possess a strong motivation to participate in substantial groups beyond their individual existence, which is characterized by robust structural coherence (Ostrom, 2009). As we saw in the example above, people who share common values strengthen their identity in the face of challenges, which in recent years included the Covid-19 pandemic and related health policy. A fundamental human propensity is the desire to foster connections with others, developing communication frameworks, norms, and traditions. Strong identification with a group often makes people susceptible to collective emotions, feeling pride in the group's accomplishments and resentment when the group faces criticism or attack,

regardless of their personal involvement. This inclination leads them to build lasting social organizations and assure the persistence of the groups to which they belong (Goldstone & Theiner, 2017). Over time, these groups, diverse in size and objectives, can develop systems of rules and norms, like laws, to orchestrate group interactions. They can also establish monitoring bodies to oversee compliance with these rules, adjudicators to ascertain rule violations, and punitive measures to discipline rule-breakers (Ostrom, 2009). A functional perspective reveals, for instance, the role of gossip within communities as a conduit for disseminating crucial information, especially concerning deviations from societal norms, making it accessible to a broader audience who wouldn't have direct knowledge otherwise (Levine & Smith, 2013). Similarly, profound emotional ties to a group, manifesting as feelings of pride, guilt, or anger, motivate individuals to undertake actions that bolster, mend, or safeguard their group's integrity and standing.

The belief in the existence of basic ethical principles that shape the attitudes of major social groups is often associated with the restoration of the concept of the common good, which was dismissed as impractical by positivist researchers. Peter deLeon argued that the goal of policymaking should be to provide "common-sense solutions for the common good" (deLeon, 1988), while Wayne Parsons added that a sense of the common good in public policy requires an agreement on the principles of what a good society ought to look like (Parsons, 2020). In contemporary policy studies, the common good is most closely associated with communitarian philosophers Michael Sandel, Charles Taylor, and Michael Walzer. Michael Sandel, in particular, argues that public policies should focus on promoting the well-being of society as a whole, rather than merely serving the interests of individuals (Sandel, 2009). The common good is defined by Sandel as a set of shared values, institutions, and conditions that benefit all members of society, fostering social cohesion and solidarity. According to Sandel, the pursuit of the common good requires active civic engagement and deliberation, where citizens participate in political discourse and decision-making processes (Sandel, 2012). This approach emphasizes the importance of community, shared responsibility, and mutual support, as opposed to individualism. The pursuit of common good urges policymakers to address societal issues holistically, taking into account the interdependencies between different sectors and the potential long-term consequences of their decisions, rather than focusing solely on the interests of specific individuals or groups.

A notable attempt to translate the idea of the common good into practice as public policy is the concept of public value. Public value refers to the contribution an organization or activity makes to broader society. This perspective reframes traditional notions of *public interest*, *common good*, and *common welfare* through a more managerial and entrepreneurial lens. The public value concept, first introduced by Harvard professor Mark H. Moore, serves as an analog to the business term *shareholder value*, but in the realm of public management. Moore envisioned public value as a measure of success in the public sector, gauging the effectiveness of an entity's ability to augment its worth to the public, both immediately and over time. This idea, rooted in his work *Creating Public Value* (Moore, 1995), aligns managerial success in governmental endeavors with enhancing their societal benefits. A comprehensive

review on empirical research into public value conducted by Faulkner and Kaufman (2018) highlights four primary dimensions that define this concept:

- (1) *Outcome Achievement*: This examines the degree to which a public entity enhances outcomes that the public deems valuable—spanning various sectors.
- (2) *Trust and Legitimacy*: This dimension gauges how much an organization and its actions are trusted by the public and key stakeholders, and whether these actions are perceived as legitimate.
- (3) *Service Delivery Quality*: Focusing on user experience, this element evaluates whether services meet high standards and are tailored to the needs of their recipients.
- (4) *Efficiency*: This considers whether an entity is optimizing its resources to deliver maximum benefits.

Moore's portrayal of a *public manager* working entrepreneurially for the common good garnered significant attention worldwide and aided in preserving and reinforcing the professional identities of public servants during challenging periods. It is worth noting that the challenges related to twenty-first-century social problems, such as the coronavirus pandemic, have redirected the interest of both researchers and politicians to issues of common good. They seek to acknowledge the role of collective well-being and foster a sense of community. A compelling indicator for this shift is a statement by the British Prime Minister Johnson: "One thing the coronavirus crisis has already proved is that there really is such a thing as society" (Meynhardt, 2021).

1.8 Argumentative Turn and the Role of Deliberation

Lasswell's exploration of the policymaking process primarily concentrates on the role of decision-making within government institutions, largely overlooking the impact of external forces on state actions and the participation of the general public. This perspective assumed that the responsibility of policy formulation rested with a limited number of government officials, neglecting the potential influence of stakeholders and other outside actors. This model failed to account for the dynamic nature of the policymaking process, where feedback loops and continuous adjustments are commonplace (Howlett et al., 1995, p. 11). Proponents of *objectivity* and scientific rigor in policy studies aimed to develop a class of apolitical, *scientific* specialists. These advocates believed that the cultivation of such experts would ensure a more informed and effective approach to addressing complex policy issues. The ultimate goal of this training was to generate a cohort of *professionals*—individuals who possess essential knowledge and skills that are crucial to society and who willingly assume the responsibility of utilizing these abilities in the best interest of the public (deLeon, 1988; Lasswell, 1971, pp. 4–9). By nurturing a class of scientifically trained policy experts, the supporters of this perspective hoped to instill a sense of neutrality and impartiality in the policymaking process, thereby reducing the influence of partisan politics and personal biases. Particularly in the policy formulation

and decision-making stages of the policy cycle, the involvement of policy actors was expected to be limited to individuals holding formal positions within the government, leaders of parliamentary factions, or, in some instances, representatives of the judiciary. To streamline the decision-making process, the number of decision-makers was expected to be limited to a small group of persons who embodied utility-maximizing, economically rational actors. As deLeon noted, “early problem-oriented research was, for a number of reasons, limited to a very small community (in many cases, just the analyst and the policymaker), a closed set at the preference of most of the participants” (DeLeon, 1988, p. 76).

This vision of professional policy specialists, however, raised concerns about the potential for elitism and detachment from the broader public. Post-positivist critics challenged the technocratic view of policymaking, which assumed that rational and objective analysis by experts is sufficient for effective policy decisions. Critics argued that an overreliance on technical expertise might inadvertently exclude diverse perspectives and voices, thus limiting the inclusivity and democratic nature of the policymaking process. The observed influence of interest groups and lobbyists further contributed to the elitist nature of policymaking. These groups often possessed significant financial resources and the ability to shape public discourse, allowing them to exert substantial influence over policy decisions. Critics have expressed concern that the insular nature of twentieth-century policymaking could result in prioritizing the interests of bureaucratic elites over those of the general public. This concentration of power among a select few may contribute to a democratic deficit, as most of the population has minimal influence on the policies that directly impact their lives. A lack of transparency and inclusiveness directly contributed to public disillusionment and the erosion of trust in political institutions, ultimately undermining the legitimacy of democratic governance. Furthermore, excluding diverse perspectives and experiences from the policymaking process led to policies that inadequately address the needs and concerns of various societal groups.

The *argumentative turn in policy analysis* emerged in the late 1980s and early 1990s as a response to the concerns mentioned above, advocating for a more inclusive and participative approach to policymaking. By emphasizing the importance of open dialogue, deliberation, and the exchange of diverse perspectives, the argumentative turn sought to break down the barriers of the closed policymaking process that had been dominated by bureaucratic elites. The term *argumentative turn* was first coined by Frank Fischer and John F. Forester in their 1993 edited volume *The Argumentative Turn in Policy Analysis and Planning*. This book brought together various approaches to policy analysis, all of which highlighted the significance of argumentation, debate, rhetoric, language, and meaning as essential elements in the examination of policy-making and the planning processes (Fischer and Forester, 1993). The concept was inspired by the idea of the *linguistic turn*, popularized by Rorty (1979), and provided a methodological framework oriented toward the analysis of policy deliberation.

The desire to broaden participation in policymaking directed the research interests of policy scientists toward defining the standards of interpersonal communication, i.e., eliminating from the language or minimizing the occurrence of syntactically uncoordinated (nonsense), unstable, vague, fuzzy, unclear, or incomprehensible expressions. Here, Jürgen Habermas's concept of communicative rationality is useful, as it posits that effective human action naturally arises from successful communication and the potential for specific forms of reason is inherent within communication itself (Habermas, 1984). Habermas' theory suggests that through open and honest discourse, participants can resolve disagreements and coordinate their actions based on shared understanding and mutual agreement. The deliberative process, grounded in the principles of equality, sincerity, comprehensibility, and truthfulness, allows participants to critically examine the validity of claims and arguments. Forester emphasizes that the rationality proposed by Habermas does not belong to the technocratic school, because this researcher "has sought to develop a model of practical communicative action, action understood not as the result of a formal means-ends calculation, but as a pragmatic socially meaningful performance" (Forester, 1993, p. 68).

Habermas divided the actions performed by people into several types. *Instrumental actions*, typical for the pursuit of a technical interest, do not require social interaction. All kinds of actions of an instrumental or strategic nature, calculated for success, are devoid of voluntary acceptance of consent by both parties in the interaction. In the case of *strategic actions*, communication is just a means of achieving a goal that is superior to other elements. This is a type of interaction in which the achievement of goals is based only on complementing preferences or balancing the interests of both parties. In the case of *communicative actions*, however, the main goal is to reach a mutually satisfactory and fair agreement. The kind of reasoned communication advocated by Habermas does not refer to emotions at all, but states that we can act according to a consensus-oriented communication style. Participants in rational communication coordinate action plans and can therefore align said plans on the basis of mutually negotiated definitions of the situation. The process of reaching an agreement is expressed through three moments: comprehensibility, the possibility of criticism, and the requirement for justification. Justifiability and the possibility to criticize apply even to the type of expressions that are not accompanied by a clearly profiled claim to validity, namely expressions that articulate evaluations (Table 1.3).

According to Habermas, the desired result of such communication is a rationally motivated consensus. Consensus must be reached in accordance with an appropriate

Table 1.3 Features of communicative action in relation to other types of action according to Habermas

	Orientation	
	Oriented to success	Oriented to understanding
SCOPE: Non-social	Instrumental action	
SCOPE: Social	Strategic action	Communicative action

procedure. It has nothing to do with accidental consent, i.e., approval obtained de facto, based either on the threat of sanctions, coercion achieved by rhetorical means, calculation, or an act of desperation. The Habermasian consensus is endowed with intersubjective validity and backed by reasons ready to be articulated on demand, the legitimacy of which can be reconstructed every time, even when the participants of the interaction change. To reach such a consensus, there must be a so-called *ideal communication situation*, which is expressed by satisfying the following conditions:

- (a) *public sphere and inclusion*: no one who could make a significant contribution can be excluded, even for controversial validity claims;
- (b) *communication equality*: everyone has an equal chance to be heard;
- (c) *exclusion of delusions and illusions*: participants must be sincere and straightforward;
- (d) *no coercion*: communication must not be restricted so that a better argument can come to the fore and have a decisive influence on the outcome of the discussion (Habermas, 1984).

The argumentative turn has significantly impacted the field of policymaking and increased the role of deliberation. The concept of deliberative democracy, based on the works of Habermas (1987) and Rawls (1997), became increasingly influential, growing into the dominant paradigm for political theory in the English-speaking world. Deliberation and dialogue among stakeholders, with the aim of finding common ground or a compromise that reflects a shared understanding of the issue at hand, became one of the most important reference points in public policies in the twenty-first century (Dryzek, 2000; Fischer, 2009). By engaging in meaningful discussions, stakeholders can explore various perspectives, challenge their assumptions, and potentially identify areas of agreement or consensus (Mansbridge et al., 2010). Deliberation can be broadly described as “communication that induces reflection on preferences, values, and interests in a non-coercive fashion” (Dryzek, 2000, p. 76). In the context of deliberative democracy, this process leads to decisions that are binding for the participants or those they represent. According to Mansbridge, the deliberation should be accessible to all affected parties, with (1) equal opportunities for everyone to impact the process, (2) equal resources, and (3) fundamental rights protection. The act of providing reasons, or *reason-giving*, is a crucial aspect of this approach. Participants are expected to treat each other with mutual respect and equal consideration, attentively listening and offering reasons that others can understand and accept. The goal is to establish fair terms of cooperation among free and equal individuals while maintaining honesty in communication. Importantly, one of the most important functions of deliberation in the political process is to *justify* political decisions, and, by participating in the debate, participants are to feel bound by the obligation to respect its results: “non-deliberative mechanisms, involving coercive power in their mechanisms of decision (...) can and must be justified (...) by deliberative procedures” (Mansbridge et al., 2010). Similarly, Dryzek (2000) argues that deliberative democracy insists that the only way we can justify our collective decisions is through open and inclusive dialogue, where all affected by a decision can participate in it. Implementing deliberation in policymaking can thus be seen as a

process of sharing ideas and debating solutions crafted by experts, public employees, or politicians. In this context, the deliberative process enables a diverse range of stakeholders to engage in meaningful discussions and contribute to the refinement of policy proposals. By allowing various perspectives to be considered, policymakers can make more informed decisions, leading to the approval, modification, or even rejection of potential solutions. This inclusive approach to policymaking ensures that policy outcomes enjoy greater legitimacy and support from the wider public.

Finally, many researchers notice the epistemic value that civic deliberation can bring to policymaking. Aitamurto and Chen (2017) review many successful online participation initiatives, noting that deliberation strengthens civic values in society by fostering a more informed and active citizenry. Through joint deliberation, individuals gain deeper insights into societal issues and enhanced learning experiences, which, in turn, bolsters the legitimacy of the decision-making process. Gutmann and Thompson (2004, p. 102) notice that.

The epistemic value of deliberation is especially great when the justification for a decision must combine factual and evaluative matters (...). Although experts may be the best judges of scientific evidence, they have no special claim to finding the right answers about priorities when degrees of risk and trade-offs of costs and benefits are involved.

Going beyond the circle of experts and embedding policy decisions in civic deliberation may therefore make it possible to make the costs of public projects more realistic and reduce the accompanying risks. The exchange of diverse perspectives, data, and approaches can enable participants to choose the most potent or pertinent ones, generating synergies between individuals. These synergies can facilitate the creation of “new solutions out of the arguments, information, and solutions brought to the table” (Estlund & Landemore, 2018), reshaping viewpoints on the matter at hand and possibly leading to more informed and substantiated proposals. Consequently, this allows decision-makers to implement these solutions or choose from a variety of arguments and information to propose more effective and efficient public policies.

Civic deliberation is undoubtedly a valuable approach that can render policymaking more grassroots-oriented and inclusive. However, it is not the sole method for achieving these objectives in public debate. Indeed, some alternative strategies may be better suited to certain conditions of the debate. This topic will be revisited in Chap. 4.

1.9 The Question of Civic Activity and Self-organization

As we have noted, deliberation can enhance civic values within society by cultivating a more knowledgeable and engaged citizenry. Empowered, informed, and engaged citizens are fundamental assets of a thriving society. Conversely, the inverse is true: only active and engaged citizens can guarantee effective deliberation. Therefore, it is essential to explore the meaning of civic activity and identify the necessary

conditions for it to reach a high standard. In this section, we will address these issues by referring to two prominent twentieth-century thinkers. One of the most important contemporary theorists of civic activity was Hannah Arendt (1906–1975), a German-born historian and political philosopher. Arendt is one of the most prominent and influential modern theorists who stands for civic engagement as spontaneous and competitive human action, with an emphasis on active citizenship and civic duty (Berger, 2012). As described in her famous work *The Human Condition*, citizens engaging in the public sphere can go beyond private interests and act together in favor of the common good. They express their citizenship by being part of *vita activa* (active life) and through involvement in deliberations about what is best for their society.

In this approach, the public-spirited civic life of *polis* in ancient Greece allows the citizen to achieve full human potential. However, the author of *The Human Condition* does not limit the concept of the *polis* to ancient states, like Greek cities, but rather includes any assembly in which human words and deeds are held in high regard:

The polis, properly speaking, is not the city-state in its physical location; it is the organization of the people as it arises out of acting and speaking together, and its true space lies between people living together for this purpose, no matter where they happen to be. “Wherever you go, you will be a polis” (Arendt, 1998, p. 198).

Arendt analyzes the *vita activa* through three categories, corresponding to the three fundamental activities inherent in our existence: labor, work, and action (d’Entreves, 2022). Labor and work are related to meeting basic life needs necessary for survival, but the *life of action* celebrates human uniqueness, freedom, and individual identity. Political passivity, which—according to her—is a flaw of a society focused on consumption, is a departure from common-sense, satisfying the lowest needs of mankind leading to aimless careerism and consumerism. In the limitation of civic engagement, which is a problem for modern societies, political participation, and even a human agency itself, are perceived by passive consumers as onerous tasks best entrusted to the expertise of professionals and officials. By limiting civic activity, they condemn themselves to the threat of totalitarianism.

Arendt understands the peak human activity to be engaged in citizenship, immersed in activity, and public deliberation among equals. In *The Human Condition*, Arendt emphasizes that action in the network of human relationships is sustained by communicative interaction (Arendt, 1998, pp. 178–9, 184–6, 199–200). Action, to the extent that it requires appearing in public, making oneself known through words and deeds, and eliciting the consent of others, can only exist in a context defined by plurality—groups, assemblies, and collectives. Action entails debate in these assemblies: by means of communication, we are able to articulate the meaning of our actions and to coordinate the actions of a plurality of agents. This plurality is “specifically ‘the’ condition—not only the ‘conditio sine qua non’, but the ‘conditio per quam’—of all political life” (Arendt, 1998, p. 7). To act means to take the initiative, to introduce the novum and the unexpected into the world. It is a process that cannot occur in isolation but rather depends on the presence of a multitude of participants, each capable of presenting their distinct perspectives (d’Entreves, 2022).

According to Arendt, consensus is not a prerequisite for civic activity; instead, public debate can result in the presentation of diverse viewpoints which are heard and regarded with mutual respect. Civic engagement is valued not because it may necessarily lead to agreement, but for its role in empowering every citizen to exercise agency, cultivate the ability to make impartial judgments, and gain a degree of public influence. Political action, in her view, involves individuals expressing their freedom and engaging in activities that can lead to change. This freedom to act can result in conflicts and struggles over different visions of the common good. In pluralistic debate, where citizens can engage in public discourse and political action, competition, conflicts, and disagreements are expected as individuals with diverse perspectives come together to participate in the political process. The debate, in Arendt's view, becomes "a competitive space, in which one competes for recognition, precedence, and acclaim" (Benhabib, 1992).

Arendt distinguishes between power and strength. Power, for her, is the capacity of individuals to act together in concert to achieve common goals. Strength, on the other hand, is the capacity of the authoritative center to make decisions and enforce them. A truly civil society is based on a concept of power, whose essence arises out of a process of free, undistorted communication, discussion and debate, and concerted activities of a plurality of agents. Strength is quite the opposite: it hinges on the command-obedience dynamic emanating from a single decision-making authority. In its unwavering faith in rational infallibility, it often overlooks the potential of collective efforts (Arendt, 1972). The belief that a self-organized community of people is a much more adequate source of knowledge about public matters than an authoritative decision-making center was further developed by the outstanding economist and Nobel Prize winner Friedrich August von Hayek (1899–1992) in his concept of spontaneous order.

First, according to Hayek, knowledge is heterogeneously distributed among members of society. In the essay *Economics and Knowledge* he considers "how the spontaneous interaction of a number of people, each possessing only bits of knowledge, brings about a state of affairs in which prices correspond to costs, etc., and which could be brought about by deliberate direction only by somebody who possessed the combined knowledge of all those individuals" (Hayek, 1937). This knowledge that appears to originate in one single mind "is in fact dispersed among all the people involved in the process" (Hayek, 1945).

At the same time, Hayek emphasizes the limits of a single human's rationality. He argues that no central authority, no matter how knowledgeable, could possess the information required to plan and control complex social systems and institutions effectively. Instead, these systems can emerge and function efficiently without one authoritative decision center and without deliberate central planning or design. They evolve organically from the interactions of individuals pursuing their own self-interest and making decentralized decisions. He identifies the greatest achievements of social life as "the results of human action, but not of human design" (Hayek, 1967), calling these self-organizing processes, based primarily on the free exchange of goods, a *spontaneous order*. Spontaneous order suggests that order and coordination can arise naturally from the actions of countless individuals who are guided by

their own knowledge, preferences, and incentives. This order emerges as a result of unrestricted exchange of information. Rather than relying on top-down, hierarchical approaches where central authorities dictate solutions to complex social problems, it is based on voluntary exchange and decentralized decision-making.

Describing this process, which he believes forms the foundation of Western civilization, as the *extended order of human cooperation*, he emphasizes that this order is not coincidental but is based on cooperation grounded in shared moral principles (Hayek, 1988). Empirical study of these process can answer “the central question of all social sciences, how the combination of fragments of knowledge existing in different minds can bring about results which, if they were to be brought about deliberately, would require a knowledge on the part of the directing mind which no single person can possess” (Hayek, 1937).

This *spontaneous order*, a unique manifestation of civilizational self-organization, has, however, become vulnerable to numerous threats. Centrally managed social projects, which may appear to be fully rational and more responsive to collective needs, are, in reality, highly imperfect and disruptive bottom-up economic processes. This is because human societies, based on the spontaneous actions of individuals, operate on a different type of rationality, namely, collective wisdom, sometimes “metaphorically described as that of the ‘social mind’” (Hayek, 1937). Free communication, information exchange, and dialogue, as well as adherence to fair rules, are essential conditions for the existence of such societies. Furthermore, Hayek believes that the spontaneous order enabling dynamic and decentralized civilizational development was primarily possible due to the observance of shared principles of morality, with particular importance placed on individual freedom and justice, in conjunction with the institution of private property (Hayek, 1988). We can see here an alignment with the psychological foundations for morality described by Haidt, which emerged from collective wisdom during intergenerational interactions.

1.10 Opening Policymaking in Practice

In the early twenty-first century, the shift toward deliberation, dialogue, and civic engagement in public policy has led to the development and implementation of various innovative participatory mechanisms, such as citizen assemblies, participatory budgeting, and online consultation platforms. According to Bevir (2010), we can observe “the consequent rise in the public sector of self-organizing policy networks. The rise of such networks means that the state has to concern itself less with direct action and more with the tasks of managing and steering networks.” These practices have been shown to improve the quality of policy outcomes by fostering collective learning, building trust, and promoting social cohesion (Innes & Booher, 2003).

The concept of open policymaking entered the political agenda with the movement of open government, which began to gain momentum in the late twentieth century. The rise of the internet and digital technologies in the 1990s and early 2000s provided new opportunities for governments to be more transparent, share information with

citizens, and solicit their input in decision-making processes. As Beth Noveck, the former United States' Deputy Chief Technology Officer, explains, the fundamental concept of *open government* suggests that governmental bodies might be falling short in effectiveness and legitimacy due to their secluded operations. Therefore, the significant prospect for these institutions lies in moving from a closed-off approach toward one that harnesses the expertise and capacity found within a networked society. It is becoming increasingly clear that the most precious resource we have in our society is the smart citizen (Noveck, 2015, p. 2).

A possible way to increase openness in policymaking is through broadening social participation using Internet technologies. Technologically-supported participation, or e-participation, can be defined as “the extension and transformation of participation in societal democratic and consultative processes mediated by information and communication technologies, primarily the Internet” (Saebø et al., 2008) or as “the use of information technologies to engage in discourse among citizens and between citizens and elected or appointed officials over public policy issues” (White, 2007). This technological shift, combined with a growing recognition of the importance of civic engagement, led to the development of various related concepts relevant to open government like publicly available internet platforms, transparency, collaboration and co-creation, informed decision-making, social learning, transparency, conflict resolution (see Table 1.4). Transparency in online policymaking enables the public to take a role as active citizens and provides access to information about the policy and its reform (Aitamurto & Chen, 2017).

E-participation not only acts as a channel for communication between citizens and public administration but also transforms the relationship between the state and citizens into a more collaborative partnership. Governments around the world are experimenting with e-participation tools to enhance citizen engagement in developing policy, improve service delivery, and open public organizations, as well as gather distributed wisdom and know-how of diverse participants (Dixon, 2010; Osimo, 2008). As an example, governments can use participatory sensing on social networks using opinion mining, crowdsourcing, and immersive simulations to identify topics that should be included in the policymaking agenda (see: Mureddu et al., 2014). The trend of implementing ICT tools is visible in the support of public decision-making and optimized policy planning (Gavanelli et al., 2013; Valle-Cruz et al., 2020), analyzing large amounts of social data to detect patterns and abnormalities (Greenemeier, 2014; McKelvey & MacDonald, 2019), using dynamic models for learning, adaptation, and forecasting of policy formulation (Valle-Cruz et al., 2020; Joyner-Roberson 2019), real-time continuous policy monitoring (Grothaus, 2019; Sun & Medaglia, 2019), participating in online consultations (Milano et al., 2014; Vicente & Novo, 2014), legislative reforms (Aitamurto, 2012; Landemore, 2020), as well as urban strategy planning (Madero & Morris, 2016).

Understood in this way, openness is, on the one hand, a reflection of a changing political reality. Internet communication has enabled a dynamic flow of information and the freedom to form online groups, and as society becomes more diverse, there are more active stakeholders and more decision-making centers. On the other hand, it is a postulate of programmatically *opening* various areas of public policy, which in

Table 1.4 Summary of key terms related to opening policymaking with the use of ICT tools

Online digital platforms: tools for gathering grassroots proposals, opinions, and initiatives while also being employed by various governmental institutions. These platforms facilitate a two-way communication channel that enables citizens to actively participate in the policymaking processes and allows government institutions to better understand and address public concerns. By fostering collaboration and dialogue between citizens and authorities, these digital platforms contribute to a more inclusive, transparent, and responsive governance system
Transparency: Accessibility of information related to government decision-making, policies, and activities is a key component of Open Government Initiative and is considered essential for promoting accountability, trust, and citizen participation. Different dimensions of transparency, including proactive and reactive transparency, and the potential consequences of increased transparency for public organizations are studied (Fox, 2007; Meijer, 2013)
Informed decision-making: Participative processes can (in some cases) lead to better-informed decisions, as they encourage participants to engage with diverse viewpoints, learn from each other, and consider the potential consequences of various policy options (Fishkin, 2018)
Conflict resolution: Deliberation can help to resolve conflicts and build consensus by providing a forum for stakeholders to engage in constructive dialogue, negotiate differences, and collaboratively develop solutions to shared problems (Mansbridge et al., 2010)
Collaboration and co-creation: Open policymaking is associated with collaborative efforts and co-creation of knowledge, emphasizing the importance of engaging with various stakeholders, including marginalized and underrepresented groups, in the policy process
Social learning: A social-learning-based approach to public policies is based on the post-positivist notion that individuals and groups can gain a better understanding of the values, needs, and preferences of others via learning processes, ultimately promoting social cohesion and the development of more inclusive policies (Innes & Booher, 2010). The learning process can be supported by such ICT tools as databases open to all citizens (open data)

some domains is successfully implemented (e.g., open data, participatory budgets, provision of public services by competitive private operators, online voting), while in others it fails. An example of failure in direct democracy and its ability to shape the political agenda online is the case of the new Icelandic constitution. Despite significant political engagement and effort, it was disregarded by the political authorities, a situation I will elaborate on in the following chapter.

Let us look at openness in policymaking in three main aspects, taking into account: (1) the open policymaking environment, especially multi-level governance and multiple stream analysis, (2) open digital platforms, and (3) open data.

The *open policymaking environment* signifies a transition from corporatism, where group-government interactions are centralized and exclusive, to a more decentralized system with a larger number of participants. As government responsibilities have grown, more interest groups have emerged. In turn, policy resources have become stretched, leading to an increased reliance on external advice. The surge of involvement from various sources has transformed issues that were once privately managed by a small group of insiders into matters of public discussion. As a result, the focus has shifted from a single center of authority to multiple centers of power and influence, distributing authority among various organizations and entities.

Cairney (2020, pp. 101–102) illustrated this intricate environment where policy decisions are made as a space in which several actors and processes interact (see

Fig. 1.3). In this environment, various actors, including individuals and organizations, influence policy across numerous levels and types of governance. Numerous policymakers and influencers operate within various policymaking arenas. By concentrating solely on a select group of elites at the heart of government, we overlook the distribution of policymaking responsibilities throughout the political system. This complex landscape is characterized by an abundance of rules and norms adhered to by different institutions operating at various policy layers. Every policymaking setting has its distinct formal and informal guidelines that regulate conduct. Furthermore, networks play a crucial role in shaping policy as they encompass the relationships between policymakers and those who influence them. Ideas, particularly dominant beliefs or paradigms, often shape the discussions surrounding policy formulation. Lastly, the context in which policy choices are made is influenced by economic, social, demographic, and technological factors. Both routine and unpredictable events can abruptly shift policymakers' attention, requiring them to adapt and respond to new circumstances. This multifaceted environment illustrates the intricate nature of policymaking and the numerous elements that contribute to the decision-making process.

From the perspective of openness, the multicentric approach to policymaking, also known as multilevel governance, is particularly significant, as it describes the distribution of power from central governments to other *centers*. This approach recognizes the unclear boundaries between formal and informal sources of authority (Cairney, 2020, p. 131). The traditional model of a strong, centralized state stands in contrast to the multi-centric approach, where a fragmented, disaggregated state must share power and negotiate with various political actors, such as businesses, business associations, civic organizations, local governments, international organizations like the United Nations or European Union, trade unions, religious associations, and more. As Elinor Ostrom argues “the emergence of new and complex problems requires governments to increasingly collaborate with non-governmental actors in the understanding and addressing of these challenges” (Mureddu et al., 2014).

The United Kingdom is an example of transformation toward such a model. This process was often associated with the *new public management* (NPM) approach, meaning the application of private sector ideas to the public sector. Since the 1980s, British governments have tackled policy challenges by altering the equilibrium between the government and society, moving away from an emphasis on the public sector and increasingly focusing on the private sector. The NPM model emphasized market-based approaches, efficiency, and cost reduction resulting in increased privatization, deregulation, the introduction of quasi-markets in the public sector, and public-private partnerships (Hood, 1995). Multi-level governance in the UK also involved civil service reform, which aimed to enhance accountability among civil servants by granting them greater control over their budgets. The reform sought to distinguish policy formulation from service delivery within government departments by separating the two functions and establishing executive agencies to handle service delivery (Cairney, 2020, p. 136). The increased use of *quangos* (quasi non-governmental organizations) sponsored, but not directly controlled, by public administration is another example. One instance of a *quango* is the Environment Agency.

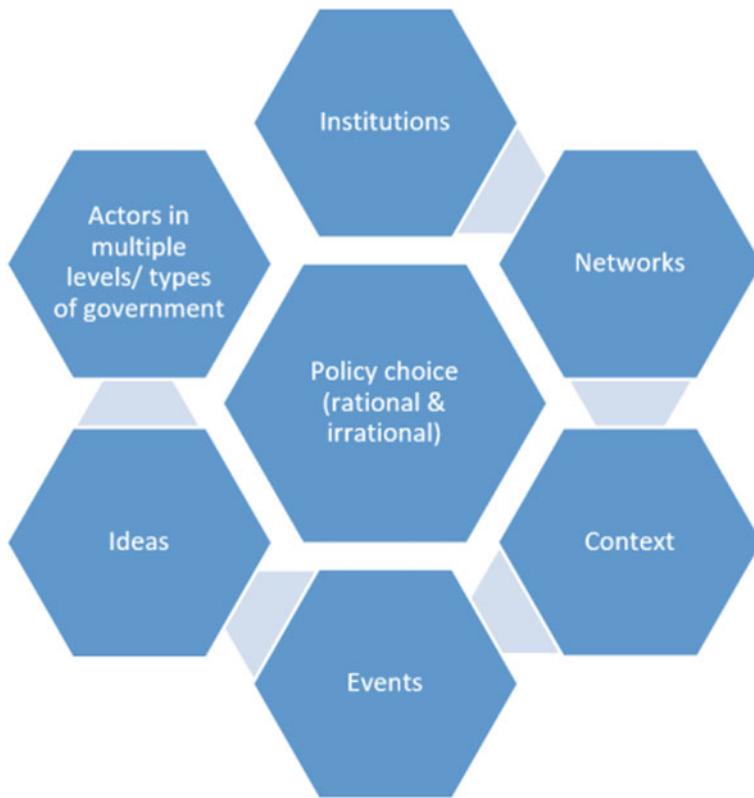


Fig. 1.3 Key elements of policy choice in a complex policymaking environment. *Source* Cairney (2020)

Established in 1996, the Environment Agency is responsible for environmental regulations, flood management, and pollution control in England. Although it receives funding from the government and operates under the UK's Department for Environment, Food, and Rural Affairs (DEFRA), it functions independently of direct government control. The agency's independent status allows it to combine expertise and resources from various sectors while still being accountable to the public and the government (Hood & Margetts, 2007).

The Multiple Streams Framework, first introduced by John Kingdon in 1984 and later updated in 2010, is a well-regarded approach to non-linear, multi-threaded analysis of policymaking. Kingdon posits that numerous alternative solutions can coexist and be developed independently in response to any policy issue (Kingdon, 2010). The central question, therefore, is why one solution is chosen over the others. The underlying assumption is that policymakers have limited time and resources at their disposal. Individual decision-makers possess finite time and capacity to address

problems. Consequently, they cannot undertake analyses on their own, but instead rely on solutions developed by various collectives.

Fluid participation in the decision-making process enables the introduction of new ideas as diverse institutions and individuals become involved. Ideas refer to shared beliefs or ways of thinking, encompassing knowledge, norms, worldviews, and ideologies. Alternative ideas materialize within three independent streams:

- (a) *Problem stream*, where perceptions of problems evolve and become redefined;
- (b) *Policy stream*, where alternatives are developed within policy communities. Debates and definitions occur within these communities, typically consisting of interest groups, academics, think tanks, experts, and members of the bureaucracy working in the relevant field;
- (c) *Political stream*, where proposals pass through initial feasibility filters and subsequently seek a *window of opportunity*—an occasion for implementation (Hoefer, 2022).

By comprehending the dynamics within and between these streams, the Multiple Streams Framework offers valuable insights into the complex, non-linear nature of policymaking, while enabling the inclusion of diverse solutions developed by competing policy communities in the process.

The development of *open digital platforms in policymaking* can be traced back to the rise of the internet and the increasing recognition of the potential for technology to transform governance and democratic processes. The development and adoption of these platforms have evolved over the past few decades. In the late 1990s and early 2000s, the internet became more widespread leading to the emergence of early internet-based initiatives. Governments and civil society organizations started experimenting with digital tools for public consultation and participation. Examples from this era include online discussion forums, e-petition systems, and rudimentary e-government services. One of the pioneering projects, Minnesota E-Democracy, originated as a grassroots initiative with the goal of promoting public discourse and involvement in political and policy matters using online forums and digital resources. Launched during the 1994 Minnesota gubernatorial race, the project eventually broadened its focus to encompass local, state, and national issues. By offering a forum for citizens to share ideas, engage in debate, and acquire knowledge about diverse policy topics, the platform enhanced civic participation and fostered a more informed electorate (Jensen, 2006). Another example comes from the United Kingdom, where the Bristol City Council emerged as one of the early local government adopters of e-petitions. The innovative online platform enabled citizens to create, sign, and submit petitions directly to the council, simplifying the process and facilitating citizen engagement with local government. The initiative's primary objective was to increase the council's responsiveness to the concerns and needs of its constituents by offering a more accessible and convenient channel for public participation in local governance (Macintosh & Whyte, 2006).

The turn of the century brings with it complex e-government systems. An interesting example comes from Estonia. The foundation was laid in the early 1990s, following the country's independence from the Soviet Union. Aiming to rebuild its

economy and infrastructure, the government embarked on a series of reforms that prioritized digitalization and modernization. One of the key milestones during this period was the creation of the Tiger Leap Program in 1996, which aimed to develop Estonia's IT infrastructure and promote computer literacy. Estonia has become a global leader in e-government, allowing citizens to access more than 99% of public services online. Its X-Road platform, launched in 2001, enables secure data exchange between various government institutions, businesses, and citizens. In 2002, Estonia introduced a national ID card system that serves as a secure digital identity for citizens. This digital identity enables Estonians to access e-services, sign documents electronically, and participate in secure online voting. Estonia also became the first country to implement online voting for national elections in 2005. The i-Voting system allows citizens to vote from anywhere in the world using their digital identity, increasing voter turnout and strengthening democratic processes (Solvak & Vassil, 2016).

In the late 2000s, the *open data movement* gained significant momentum, fueled by a growing recognition of the value of government data as a resource for innovation, transparency, and civic engagement. Open data is data that is freely available, accessible, and shareable by anyone without restrictions on its use, reuse, or redistribution. Typically, open data refers to government-generated data of public significance that is readily accessible and free from any limitations, making it easy to locate and obtain. This allows individuals and organizations to build upon existing data to develop new products, services, or insights. Open data is supplied in formats that can be easily processed by computers, facilitating efficient analysis and integration with other datasets. This data may encompass transportation data, geospatial information, weather data, reports, images, and other content relevant to the public interest (Veljković et al., 2014). The launch of Data.gov in the United States in 2009 is a prominent example of this phenomenon. Data.gov is an online platform established by the U.S. federal government to provide public access to a wide variety of datasets from numerous government agencies. The launch of Data.gov was part of a broader open government movement that emphasized the importance of transparency, participation, and collaboration. In 2009, then President Barack Obama issued the *Transparency and Open Government Memorandum*, which laid the groundwork for Data.gov and similar open data initiatives across the country (Janssen et al., 2012). The success of Data.gov inspired other countries to develop their own open data platforms, leading to a global open data movement. The United Kingdom launched data.gov.uk in 2010 and the European Union established the European Data Portal in 2015. These platforms aimed to increase government transparency, empower citizens, and drive innovation through data-driven insights.

Recent trends in the development of open digital platforms are associated with the rise of social media and mobile technologies, which have transformed the way people communicate, access information, and interact with governments. Policy-makers leveraged these platforms to engage with citizens, solicit input, and disseminate information about policy initiatives. At the same time, the rise of civic tech and GovTech movements led to the development of new tools and platforms specifically designed for policy collaboration and participation. The emergence of advanced

tools for policy co-creation, crowdsourcing, and real-time monitoring and evaluation illustrates the diversity of these phenomena. Recent trends also include the use of artificial intelligence, machine learning, and blockchain technologies to enhance the capabilities of digital platforms and improve the policy process. Data-powered collaborative governance employs technologies of opinion mining and sentiment analysis, simulations and serious gaming, participatory sensing with geo-tagging, and much more (Mureddu et al., 2014). There are many examples of such projects. Brazil's *Mudamos* platform is a mobile app that allows Brazilian citizens to create and sign legislative proposals, with the potential for successful proposals to be submitted to the National Congress. In turn, *vTaiwan* is an online platform that combines social media, AI-driven discussions, and face-to-face meetings to involve citizens in the policy process, particularly on digital and technology-related issues.

The open government phenomenon has gained significant prominence in official policy agendas over the past few years, with several milestones marking its evolution. One such key milestone was the launch of the Open Government Partnership (OGP) in 2011 by eight founding countries, including the United States, Brazil, Indonesia, Mexico, Norway, the Philippines, South Africa, and the United Kingdom. The OGP is a multilateral initiative that promotes open government reforms, increasing transparency and citizen participation in public policymaking through the development and implementation of National Action Plans. This initiative fosters collaboration among governments, civil society organizations, and citizens to create more accountable, transparent, and responsive governance systems (Veljković et al., 2014). Another pivotal moment in the open government movement came with the approval of the 2030 Agenda for Sustainable Development by all United Nations Member States in 2015. This agenda includes 17 Sustainable Development Goals that address a range of global challenges, such as poverty, inequality, climate change, and peace and justice. One of the adopted strategic goals within the 2030 Agenda emphasizes the need for responsive, inclusive, participatory, and representative decision-making at all levels. This goal underscores the importance of open government practices in achieving sustainable development and empowering citizens to take an active role in shaping their communities' futures (Transforming our world..., 2023).

1.11 How to Evaluate Open Policymaking? In Search of Collective Intelligence

The emphasis on openness is a significant trend in both the public and private sectors. Institutions, companies, and NGOs are implementing participatory management methods and encouraging free communication, knowledge sharing, flexible work arrangements, and dynamic task force creation. The development of interest in these forms of collaboration is, of course, closely related to the emergence of new technological possibilities brought about by internet communication. The virtual agora

enabling free debate and remote-working teams has become the norm. In many businesses, especially those based on expert knowledge, agile management has become the standard approach to project management. This approach prioritizes flexibility, collaboration, and responsiveness to change. One of the key features of agile management is its openness to feedback and adaptation. According to the agile model, teams are expected to be transparent about their progress, communicate frequently, and be open to changes in requirements or priorities. Agile teams work to create a culture of trust and collaboration, where team members are encouraged to share their ideas and feedback openly and honestly, without fear of judgment or retribution (Martin, 2003). The open and collaborative style of teamwork is certainly highly valued in the present moment. At the same time, however, the multitude of projects and initiatives often gives the impression of chaotic implementation, and online communication regarding public affairs is sometimes described as filled to the brim with misinformation or simply *garbage* (Pennycook & Rand, 2019). In 2018, for example, it was reported that the online platform for public consultation in Brazil was flooded with false information and malicious content during the presidential election campaign (Magenta et al., 2018).

The metaphor of policymaking as a *garbage can*, first stated by Cohen et al. (1972) paper has gained particular relevance in recent years. Cohen argues that public policymaking is a chaotic process that resembles rummaging through a wastebasket, where good ideas are mixed up with bad ones in a haphazard way. The garbage can is a container into which a volatile mix of problems and solutions is dumped. Public matters are discussed in independent streams that have a complicated and unpredictable relationship with each other. Although problems are identified, solutions proposed, and choices made, it is not necessarily guaranteed that the process will be chronological or that these streams will come together in any meaningful way (Cairney, 2020). Cohen describes a policymaking process where problems are poorly understood, goals are ambiguous, and conflict is common. Decision-makers may also have other concerns that distract them from the policy issues at hand (Cohen et al., 1972, p. 16). As a result, outcomes can be a matter of chance. This vision has become even more relevant in the age of online communication and social media, where information flows quickly and chaotically, and the public's attention is often drawn to sensationalist or misleading content. The online world has created new channels for ideas, opinions, and solutions to be dumped into the policymaking process, making the *garbage can* metaphor even more applicable today.

The current landscape of policymaking, characterized by disarray and an excess of garbage information, makes it particularly important to adopt an effective methodology to evaluate the processes through which public policies are formulated. Policy-making evaluation, a distinct concept from policy evaluation, is a meta-level assessment that focuses on the processes and methods used to create policy, rather than the outcomes or impacts of the policy itself. It examines how policies are formulated, the systems and structures that facilitate policy development, and the effectiveness of these processes in producing well-informed and relevant policies. This kind of evaluation is a systematic process that aims to determine the merit, worth, or value of policymaking by applying certain criteria and standards.

Policymaking evaluation is crucial because the quality of the policymaking process can significantly affect the quality of the outcomes. A well-evaluated policymaking process is more likely to be effective in addressing the issues it aims to solve. To conduct policymaking evaluation, researchers might look at case studies, engage in comparative analyses, use process tracing methods, and employ other qualitative research techniques. For this purpose, methods such as cost–benefit analysis (CBA) can be used. A CBA is an assessment of the economic efficiency of a policy by comparing its costs to its benefits (Boardman et al., 2017). Another possibility is a cost-effectiveness analysis (CEA), comparing the relative costs of policies based on a specific outcome which, unlike CBA, does not attempt to assign a monetary value to the benefit (Drummond et al., 2015). You can also try impact evaluation, which assesses the causal effect of policies employing experimental or quasi-experimental designs such as randomized controlled trials or natural experiments (Gertler et al., 2016).

Nevertheless, I have no doubt that open policymaking requires a different approach to evaluation in comparison to conventional policymaking methods. Open policymaking operates within the dynamic, interactive, and information-rich environment of the internet, and can occur much faster than traditional policymaking. Methods of evaluation must be agile enough to keep pace with the rapid development and iteration of policies in digital spaces. The policies developed in online communities are subject to network effects, where ideas can spread widely and spontaneously, sometimes regardless of their merit. The internet allows a vast number of participants to engage in the policymaking process, which was not an advantage for the classic policymaking, which focused on using experts. Nowadays, however, the potential multiplicity and diversity of people involved can be a great advantage. The question remains whether any evaluative methods exist that can analyze how the number and diversity of participants in online projects can translate into the quality of their outcomes. What types of analysis can be employed to assess the factors that transform the openness and widespread participation of stakeholders into effective policies with a positive impact on the common good?

In this volume, I will propose an approach that entails an analysis of open policymaking initiatives through the scrutiny of collective intelligence manifestations evident in these projects. Collective intelligence research is an attempt to capture those features of online communities that may contribute to their success in problem-solving. It is an attempt to understand how distributed groups of people communicating over the internet can achieve unexpectedly good results that exceed the capabilities of even the smartest of them acting alone. It is an attempt to capture the phenomenon of open source software, Wikipedia, crowdsourcing, and citizen science. It is an analysis of intelligence that emerges from the mutual inspiration, collaboration, collective efforts, and competition of many individuals that appear in online networks. Nevertheless, before we attempt this analysis, let's get acquainted with the evolution of the concept of collective intelligence, the examples of projects, and the state of research on this topic.

References

Águia, P. B., & Correia, A. (2021). Organizing for innovation in the armed forces: A logical thinking process approach. In S. Ojo (Ed.), *Global perspectives on military entrepreneurship and innovation* (pp. 283–305). IGI Global. <https://doi.org/10.4018/978-1-7998-6655-8.ch014>.

Aitamurto, T., & Chen, K. (2017). The value of crowdsourcing in public policymaking: Epistemic, democratic, and economic value. *The Theory and Practice of Legislation*, 5(1), 55–72.

Aitamurto, T. (2012). Crowdsourcing for democracy: New era in policy-making. *Publications of the Committee for the Future, Parliament of Finland*. 1/2012. Helsinki, Finland.

Althaus, C., Bridgman, P., & Davis, G. (2007). *Australian policy handbook*. Allen & Unwin.

Amy, D. J. (1984). Why policy analysis and ethics are incompatible. *Journal of Policy Analysis and Management*, 13(4), 573–591.

Anderson, C. W. (1979). The place of principles in policy analysis. *The American Political Science Review*, 73(3), 711–723. <https://doi.org/10.2307/1955399>

Arendt, H. (1972). *Crises of the republic*. A Harvest Book.

Arendt, H. (1998). *The human condition*. The University of Chicago Press.

Aristotle. (350 B.C./2004). Politics: A treatise on government (W. Ellis, Trans.). The Project Gutenberg EBook. <https://www.gutenberg.org/cache/epub/6762/pg6762-images.html>.

Arrow, K. J. (1950). A difficulty in the concept of social welfare. *Journal of Political Economy*, 58(4), 328–346. <https://www.jstor.org/stable/1828886>.

Auletta, K. (1983). *The underclass*. Random House.

Bailey, M. J., & Danziger, S. (2013). *Legacies of the war on poverty*. Russell Sage Foundation.

Benhabib, S. (1992). Models of public space: Hannah arendt, the liberal tradition and jürgen habermas. In S. Benhabib (Ed.), *Situating the self: Gender, community and postmodernism in contemporary ethics* (pp. 89–120). Polity Press.

Berger, B. (2012). Political engagement as intrinsic good: Arendt and Company. In *Attention deficit democracy: The paradox of civic engagement* (pp. 52–82). Princeton University Press. <https://doi.org/10.1515/9781400840311-004>.

Bevir, M. (2010). *Democratic governance*. Princeton University Press.

Bevir, M., & Rhodes, R. A. W. (2003). *Interpreting British governance*. Routledge.

Boardman, A. E., Greenberg, D. H., Vining, A. R., & Weimer, D. L. (2017). *Cost-benefit analysis: Concepts and practice*. Cambridge University Press.

Brewer, G. D. (1974). The policy sciences emerge: To nurture and structure a discipline. *Policy Sciences*, 5(3), 239–244.

Brik, A. B., & Pal, L. A. (Eds.). (2021). *The future of the policy sciences*. Edward Elgar Monographs.

Brunori, D. (1997). Principles of tax policy and targeted tax incentives. *State and Local Government Review*, 29(1), 50–61. <https://doi.org/10.1177/0160323X9702900106>

Buchanan, J. M. (1949). The pure theory of government finance: A suggested approach. *The Journal of Political Economy*, 57, 496–505.

Cairney, P. (2020). *Understanding public policy*. Red Globe Press.

Calvan, B. C. (2020). *Florida reopens state's economy despite ongoing pandemic*. Associated Press. <https://apnews.com/article/virus-outbreak-florida-business-ron-desantis-donald-trump-e64376aba8306681b53d52956d15bcd7>.

Centers for Disease Control and Prevention. (2023). How to protect yourself and others. <https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/prevention.html>.

Chan, E. Y. (2021). Moral foundations underlying behavioral compliance during the COVID-19 pandemic. *Personality and Individual Differences*, 171, 110463. <https://doi.org/10.1016/j.paid.2020.110463>

Cherrier, B., & Fleury, J. B. (2017). Economists' interest in collective decision after World War II: A history. *Public Choice*, 172, 23–44. <https://doi.org/10.1007/s11127-017-0410-7>

Cohen, M. D., March, J. G., & Olsen, J. P. (1972). A garbage can model of organizational choice. *Administrative Science Quarterly*, 17(1), 1–25.

Colebatch, H. (1998). *Policy*. Open University Press.

Cooley, A. (2020). War on poverty. *Encyclopedia Britannica*. <https://www.britannica.com/topic/War-on-Poverty>.

Council on Contemporary Families. (2014). *Was the war on poverty a failure? Or are anti-poverty efforts simply swimming against a stronger tide?* <https://sites.utexas.edu/contemporaryfamilies/2014/01/06/was-war-on-poverty-a-failure-report/>.

Crombez, C., & Hug, S. (2000). Policy making and commission appointment in the European Union. In P. Moser, G. Schneider, & G. Kirchgässner (Eds.), *Decision rules in the European Union*. Palgrave Macmillan. https://doi.org/10.1007/978-1-349-62792-9_3.

d'Entreves, M. P. (2022). Hannah arendt. In E. N. Zalta & U. Nodelman (Eds.), *The stanford encyclopedia of philosophy*. <https://plato.stanford.edu/archives/fall2022/entries/arendt/>.

Dahl, R. A., & Lindblom, C. E. (1953). *Politics, economics, and welfare*. Harper & Bros.

DeLeon, P. (1988). *Advice and consent: The development of the policy sciences*. Russell Sage Foundation.

Dixon, B. E. (2010). Towards e-government 2.0: An assessment of where e-Government 2.0 is and where it is headed. *Public Administration and Management*, 15(2), 418–454.

Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, G. L., & Torrance, G. W. (2015). *Methods for the economic evaluation of health care programmes*. Oxford University Press.

Dryzek, J. S. (2000). *Deliberative democracy and beyond: Liberals, critics*. Oxford University Press.

Dryzek, J. S. (2002). A post-positivist policy-analytic travelogue. *The Good Society*, 11(1), 32–36. <https://doi.org/10.1353/gso.2002.0004>

Easton, D. (1953). *The political system: An inquiry into the state of political science*. Knopf.

Edwards, M. (1998). The demise of the policy management review: The device for evaluating policy process. *AJPA*, 57(4).

Edwards, M. (2001). *Social policy, public policy—from problem to practice*. Allen and Unwin.

Ekici, H., Yücel, E., & Cesur, S. (2023). Deciding between moral priorities and COVID-19 avoiding behaviors: A moral foundations vignette study. *Current Psychology*, 42, 5922–5938. <https://doi.org/10.1007/s12144-021-01941-y>

Estlund, D., & Landemore, H. (2018). The epistemic value of democratic deliberation. In A. Bächtiger, J. S. Dryzek, J. Mansbridge, & M. Warren (Eds.), *The oxford handbook of deliberative democracy* (pp. 1–20). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198747369.013.26>.

Etzioni, A. (1967). Mixed-scanning: A “third” approach to decision-making. *Public Administration Review*, 27(5), 385–392. <https://doi.org/10.2307/973394>

Everett, S. (2003). The policy cycle: Democratic process or rational paradigm revisited?. *Australian Journal of Public Administration*, 62, 65–70. <https://doi.org/10.1111/1467-8497.00325>

Faulkner, N., & Kaufman, S. (2018). Avoiding theoretical stagnation: A systematic review and framework for measuring public value. *Australian Journal of Public Administration*, 77(1), 69–86.

Fenna, A. (2019). National competition policy: Effective stewardship of markets. In J. Luetjens, M. Mintrom, & P. Hart (Eds.), *Successful public policy: Lessons from Australia and New Zealand* (pp. 191–206). ANU Press. <https://doi.org/10.22459/SPPANZ.09.2019.10>.

Fischer, F. (2003). *Reframing Public Policy: Discursive Politics and Deliberative Practices*. Oxford University Press. <https://doi.org/10.1093/019924264X.001.0001>.

Fischer, F. (2009). Participatory governance as deliberative empowerment: The cultural politics of discursive space. *American Review of Public Administration*, 39(4), 385–410.

Fischer, F., Forester, J. F. (Eds.). (1993). *The argumentative turn in policy analysis and planning*. Duke University Press.

Fishkin, J. S. (2018). *Democracy when the people are thinking: Revitalizing our politics through public deliberation*. Oxford University Press.

Fitzgerald, L., Mutch, A., & Herron, L. (2019). Responding to HIV/AIDS: Mobilisation through partnerships in a public health crisis. In J. Luetjens, M. Mintrom, & P. Hart (Eds.), *Successful*

public policy: Lessons from Australia and New Zealand (pp. 29–58). ANU Press. <https://doi.org/10.22459/SPPANZ.05.2019.03>.

Forester, J. (1984). Bounded rationality and the politics of muddling through. *Public Administration Review*, 44(1), 23–31. <https://doi.org/10.2307/975658>

Forester, J. (1993). *Critical theory, public policy, and planning practice*. State University of New York Press.

Fox, J. (2007). The uncertain relationship between transparency and accountability. *Development in Practice*, 17(4–5), 663–671.

Fraussen, B., & Halpin, D. (2017). Think tanks and strategic policy-making: The contribution of think tanks to policy advisory systems. *Policy Sciences*, 50, 105–124. <https://doi.org/10.1007/s11077-016-9246-0>

Friend, J. K., Power, J. M., & Yewlett, C. J. L. (1974). *Public planning: the inter-corporate dimension*. Tavistock.

Gavanelli, M., Riguzzi, F., Milano, M., & Cagnoli, P. (2013). Constraint and optimization techniques for supporting policy making. In T. Yu, N. Chawla, & S. Simoff (Eds.), *Computational intelligent data analysis for sustainable development*. Chapman & Hall/CRC.

Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. J. (2016). *Impact evaluation in practice*, second edition. World Bank Publications.

Gigerenzer, G. (2001). The adaptive toolbox. In G. Gigerenzer & R. Selten (Eds.), *Bounded rationality: The adaptive toolbox*. MIT Press.

Goldstone, R. L., & Theiner, G. (2017). The multiple, interacting levels of cognitive systems (MILCS) perspective on group cognition. *Philosophical Psychology*, 30(3), 334–368. <https://doi.org/10.1080/09515089.2017.1295635>

Goodin, R. E., & Klingemann, H. D. (Eds.). (1996). *A new handbook of political science*. Oxford University Press.

Goswami, K., & Gerritsen, R. (2021). Policy life cycle analysis of three Australian state-level public policies: Exploring the political dimension of sustainable development. *Journal of Development Policy and Practice*, 6(1), 9–35. <https://doi.org/10.1177/2455133321998805>

Greenemeier, L. (2014). Smart machines join humans in tracking Africa Ebola Outbreak. *Scientific American*. <https://www.scientificamerican.com/article/smart-machines-join-humans-in-tracking-africaebola-outbreak/>.

Grothaus, M. (2019). China's airport facial recognition kiosks should make us fear for our privacy. *Fast company*. <https://www.fastcompany.com/90324512/chinas-airport-facial-recognition-kiosks-should-make-us-fear-for-ourprivacy>.

Gutmann, A., & Thompson, D. (2004). *Why deliberative democracy?* Princeton University Press.

Habermas, J. (1984). *The theory of communicative action*. Beacon.

Habermas, J. (1987). *The theory of communicative action: Vol. 2, Lifeworld and system: A critique of functionalist reason*. Beacon Press.

Haidt, J. (2013). *The righteous mind: Why good people are divided by politics and religion*. Penguin Books.

Hajer, M. A. (1995). *The politics of environmental discourse: Ecological modernization and the policy process*. Oxford University Press.

Hawkesworth, M. E. (1988). *Theoretical issues in policy analysis*. State University of New York Press.

Hayek, F. A. (1937). Economics and knowledge. *Economica*, 4(13), 33–54.

Hayek, F. A. (1945). The use of knowledge in society. *The American Economic Review*, 35(4), 519–530.

Hayek, F. A. (1967). The results of human action but not of human design. *Studies in philosophy, politics and economics*, 96.

Hayek, F. A. (1988). *The fatal conceit, volume 1: The errors of socialism*. University of Chicago Press.

Heclo, H. (1972). Review article: Policy analysis. *British Journal of Political Science*, 2, 83–108.

Heclö, H. (1974). *Modern social politics in Britain and Sweden: From relief to income maintenance*. Yale University Press.

Hoefer, R. (2022). The multiple streams framework: Understanding and applying the problems, policies, and politics approach. *Journal of Political Practice and Research*, 3, 1–5. <https://doi.org/10.1007/s42972-022-00049-2>

Holden, M. (2008). Social learning in planning: Seattle's sustainable development codebooks. *Progress in Planning*, 69, 1–40.

Hood, C. (1995). The “new public management” in the 1980s: Variations on a theme. *Accounting, Organizations and Society*, 20(2–3), 93–109.

Hood, C., & Margetts, H. (2007). *The tools of government in the digital age*. Palgrave Macmillan.

Horne, J., & Grafton, R. Q. (2019). The Australian water markets story: Incremental transformation. In J. Luetjens, M. Mintrom, & P. Hart (Eds.), *Successful public policy: Lessons from Australia and New Zealand* (pp. 165–190). ANU Press. <https://doi.org/10.22459/SPPANZ.09.2019.09>.

Howlett, M., Ramesh, M., & Perl, A. (1995). *Studying public policy: Policy cycles and policy subsystems*. Oxford University Press.

Innes, J. E., & Booher, D. E. (2010). *Planning with complexity: An introduction to collaborative rationality for public policy*. Routledge.

Innes, J. E., & Booher, D. E. (2003). Collaborative policymaking: Governance through dialogue. In M. A. Hager & H. Wagenaar (Eds.), *Deliberative policy analysis: Understanding governance in the network society* (pp. 33–59). Cambridge University Press.

Introduction to Delivering Great Policy. (2023). *Policy hub*. <https://www.policyhub.gov.au/model>.

Jamieson, T. (2020). “Go Hard, go early”: Preliminary lessons from New Zealand’s response to COVID-19. *The American Review of Public Administration*, 50(6–7), 598–605. <https://doi.org/10.1177/0275074020941721>

Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. *Information Systems Management*, 29(4), 258–268.

Jenkins, W. I. (1978). *Policy analysis: A political and organizational perspective*. Martin Robertson.

Jensen, J. L. (2006). The Minnesota E-democracy project: Mobilising the mobilised? In S. Oates, D. Owen, & R. Gibson (Eds.), *The internet and politics: Citizens, voters and activists* (pp. 39–58). Frank Cass Publishers.

Jones, B. D. (1994). *Reconceiving decision-making in democratic politics: Attention, choice, and public policy*. University of Chicago Press.

Joyer-Roberson, E. (2019). What do drones, AI and proactive policing have in common? SAS. https://www.sas.com/en/_us/insights/articles/riskfraud/drones-ai-proactive-policing.html.

Kaul, I., Grubberg, I., & Stern, M. (1999). Defining global public goods. In I. Kaul, I. Grubberg, & M. Stern (Eds.), *Global public goods: International cooperation in the 21st century* (pp. 2–19). Oxford University Press.

King, M., & Kay, J. (2020). *Radical uncertainty: Decision-making for an Unknowable Future*. Little, Brown.

Kingdon, J. W. (2010). *Agendas, alternatives, and public policy* (2nd ed.). Pearson.

Kotarbiński, T. (1965). *Traktat o dobrej robocie*. Zakład Narodowy im. Ossolińskich, Wydział Nauk Społecznych PAN.

Landemore, H. (2020). When public participation matters: The 2010–2013 Icelandic constitutional process. *International Journal of Constitutional Law*, 18, 179–205.

Lasswell, H. D. (1970). The emerging conception of the policy sciences. *Policy Sciences*, 1, 3–14. <https://doi.org/10.1007/BF00145189>

Lasswell, H. D. (1971). *A pre-view of policy sciences*. American Elsevier.

Lasswell, H. D. (2017). *Power and personality*. Routledge.

Levine, J. M., & Smith, E. R. (2013). Group cognition: Collective information search and distribution. In D. E. Carlston (Ed.), *The Oxford handbook of social cognition* (pp. 616–633). Oxford University Press.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.

Lindblom, C. E. (1959). The science of muddling through. *Public Administration Review*, 19, 79–88.

Little, D. (1991). Rational-choice models and Asian studies. *The Journal of Asian Studies*, 50(1), 35–52.

Macintosh, A., & Whyte, A. (2006). Evaluating how eparticipation changes local democracy. *Egov-
ernment Workshop '06 (Egov06)*. https://www.researchgate.net/publication/228370359_Evaluating_how_eParticipation_changes_local_democracy.

Madero, V., & Morris, N. (2016). Public participation mechanisms and sustainable policy-making: A case study analysis of Mexico City's Plan Verde. *Journal of Environmental Planning and Management*, 59(10), 1728–1750. <https://doi.org/10.1080/09640568.2015.1085841>

Magenta, M., Gragnani, J., & Souza, F. (2018). How WhatsApp is being abused in Brazil's elections. <https://www.bbc.com/news/technology-45956557>.

Malbin, M. J. (1980). *Unelected representatives*. Basic Books.

Mansbridge, J., Bohman, J., Chambers, S., Estlund, D., Føllesdal, A., Fung, A., Lafont, C., Manin, B., & Martí, J. L. (2010). The place of self-interest and the role of power in deliberative democracy. *Journal of Political Philosophy*, 18(1), 64–100. <https://doi.org/10.1111/j.1467-9760.2009.00344.x>

Marsh, I. (2023). *Opening up the policy process*. https://www.aph.gov.au/About_Parliament/Senate/Powers_practice_n_procedures/pops/pop34/c17.

Martin, R. (2003). *Agile software development: Principles, patterns, and practices*. Pearson Education.

McKelvey, F., & MacDonald, M. (2019). Artificial intelligence policy innovations at the Canadian Federal Government. *Canadian Journal of Communication*, 44(2), 43–50.

Meijer, A. (2013). Understanding the complex dynamics of transparency. *Public Administration Review*, 73(3), 429–439.

Meynhardt, T. (2021). Public value is knowable, public value creation is not. *Administration & Society*, 53(10), 1631–1642. <https://doi.org.libproxy.mit.edu/https://doi.org/10.1177/00953997211053500>.

Midgley, J., Tracy, M. B., & Livermore, M. (Eds.). (2000). *The handbook of social policy*. SAGE Publications.

Milano, M., O'Sullivan, B., & Gavanelli, M. (2014). Sustainable policy making: A strategic challenge for artificial intelligence. *AI Magazine*, 35, 22–35.

Moore, M. (1995). *Creating public value: Strategic management in government*. Harvard University Press.

Morcöl, G. (2001). Positivist beliefs among policy professionals: An empirical investigation. *Policy Sciences*, 34, 381–401. <https://doi.org/10.1023/A:1012749120909>

Mureddu, F., Misuraca, G., Osimo, D., Onori, R., & Armenia, S. (2014). A living roadmap for policymaking 2.0. In P. Sonntagbauer, K. Nazemi, S. Sonntagbauer, G. Prister, & D. Burkhardt (Eds.), *Handbook of research on advanced ICT integration for governance and policy modeling* (1st ed.). IGI Global: Hershey, PA, United States. <https://doi.org/10.4018/978-1-4666-6236-0>.

Noveck, B. S. (2015). *Smart citizens, smarter state: The technologies of expertise and the future of governing*. Harvard University Press.

Orrell, D., & Chlupatý, R. (2016). *Origins. In the evolution of money* (pp. 7–30). Columbia University Press. <https://doi.org/10.7312/orre17372-003>.

Osimo, D. (2008). *Web 2.0 in government: Why and how?* Office for Official Publications of the European Communities: Luxembourg.

Ostrom, E. (2009). A general framework for analyzing sustainability of social-ecological systems. *Science*, 325, 419–422.

Ostrom, E. (2015). *Governing the commons: The evolution of institutions for collective action*. Cambridge University Press.

Pabst, A. (2021). Rethinking evidence-based policy. *National Institute Economic Review*, 255, 85–91.

Parsons, W. (2020). *On the road to common good*. <https://togetherforthecommongood.co.uk/leading-thinkers/on-the-road-to-the-common-good>.

Pennycook, G., & Rand, D. G. (2019). Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. *Journal of Personality*, 88(2), 185–200.

Plato. (360 B.C./2008). The laws (B. Jowett, Trans.). The Project Gutenberg EBook. <https://www.gutenberg.org/files/1750/1750-h/1750-h.htm>.

Rawls, J. (1997). The idea of public reason revisited. *University of Chicago Law Review*, 64(3), 765–807.

Reston, M., & Krieg, G. (2021). Gavin Newsom Holds Onto His Job as California Governor, CNN. <https://edition.cnn.com/2021/09/14/politics/gavin-newsom-california-recall/index.html>.

Rorty, R. (1979). *Philosophy and the mirror of nature*. Princeton University Press.

Saebø, Ø., Rose, J., & Flak, L. S. (2008). The shape of eParticipation: Characterizing an emerging research area. *Government Information Quarterly*, 25, 400–428.

Samuelson, P. (1954). The pure theory of public expenditures. *The Review of Economics and Statistics*, 36(4), 387–389.

Sandel, M. J. (2009). *Justice: What's the right thing to do?* Farrar.

Sandel, M. J. (2012). *What money can't buy: The moral limits of markets*. Farrar.

Schaffer, B. (1977). On the politics of policy. *Australian Journal of Politics and History*, 23(1), 146–155.

Schilirò, D. (2012). Bounded rationality and perfect rationality: Psychology into economics. *Theoretical and Practical Research in Economic Fields*, 3(2), 101–111.

Schneider, A., & Ingram, H. (1997). *Policy design for democracy*. University of Kansas Press.

Sen, A. K. (1977). Rational fools: A critique of the behavioral foundations of economic theory. *Philosophy and Public Affairs*, 6(4), 317–344. <https://doi.org/10.2307/2264946>

Sen, A. K. (1998). Rational behavior. In J. Eatwell, M. Milgate, & P. Newman (Eds.), *The new Palgrave: A dictionary of economics* (Vol. 4). MacMillan Press.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69, 99–118.

Simon, H. A. (1973). The structure of ill-structured problems. *Artificial Intelligence*, 4, 181–201.

Simon, H. (1998). Bounded rationality. In J. Eatwell, M. Milgate, & P. Newman (Eds.), *The new Palgrave: A dictionary of economics* (Vol. 1). MacMillan Press.

Solvak, M., & Vassil, K. (2016). *E-voting in Estonia: Technological diffusion and other developments over ten years*. University of Tartu and Estonian National Electoral Committee.

Stone, D. (2002). *Policy paradox: The art of political decision making*. W. W. Norton & Company.

Sullivan, E., & Segers, M. (2017). Ethical issues and public policy. In *Handbook of public policy analysis: Theory, politics, and methods* (pp. 309–327). Taylor and Francis. <https://doi.org/10.4324/9781315093192-33>.

Sun, T. Q., & Medaglia, R. (2019). Mapping the challenges of artificial intelligence in the public sector: Evidence from public healthcare. *Government Information Quarterly*, 36(2), 368–383.

Szarfenberg, R. (2002). *Podstawy i granice racjonalizacji polityki społecznej*. Uniwersytet Warszawski.

Theil, H. (1958). *Economic forecasts and policy*. North-Holland.

Torgerson, D. (1985). Contextual orientation in policy analysis: The contribution of Harold D Lasswell. *Policy Sciences*, 18(3), 241–261.

Torgerson, D. (1986). Between knowledge and politics: Three faces of policy analysis. *Policy Sciences*, 19(1), 33–59. <https://doi.org/10.1007/BF00138709>

Trangerud, H. A. (2023). What is the problem with vaccines?" A typology of religious vaccine skepticism. *Vaccine X*, 14, 100349. <https://doi.org/10.1016/j.jvacx.2023.100349>

Transforming our world: the 2030 Agenda for Sustainable Development. (2023). <https://sdgs.un.org/2030agenda>.

Turkel, G. (2005). Talcott Parsons: The social relations of individualism and economy. *Sociological Focus*, 38(1), 65–82. <https://doi.org/10.1080/00380237.2005.10571204>

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124–1131.

Valle-Cruz, D., Criado, J. I., Sandoval-Almazán, R., & Ruvalcaba-Gomez, E. A. (2020). Assessing the public policy-cycle framework in the age of artificial intelligence: From agenda-setting to policy evaluation. *Government Information Quarterly*, 37(4), 101509. <https://doi.org/10.1016/j.giq.2020.101509>

Veljković, N., Bogdanović-Dinić, S., & Stoimenov, L. (2014). Benchmarking open government: An open data perspective. *Government Information Quarterly*, 31(2), 278–290.

Vicente, M. R., & Novo, A. (2014). An empirical analysis of e-participation: The role of social networks and e-government over citizens' online engagement. *Government Information Quarterly*, 31, 379–387.

Waltz, K. N. (1996). International politics is not foreign policy. *Security Studies*, 6(1), 54–57. <https://doi.org/10.1080/09636419608429298>

Weber, M. (2002). *Political writings*. P. Lassman & R. Speirs (Eds.), Cambridge University Press.

Webster, C. (2002). *The national health service: A political history*. Oxford University Press.

Weible, C. M., Nohrstedt, D., Cairney, P., et al. (2020). COVID-19 and the policy sciences: Initial reactions and perspectives. *Policy Sciences*, 53(2), 225–241. <https://doi.org/10.1007/s11077-020-09381-4>

Weingast, B. R. (1981). Regulation, deregulation, and deregulation: The political foundations of agency clientele relationships. *Law and Contemporary Problems*, 44, 147–177.

Weiss, J. A., & Tschirhart, M. (1994). Public information campaigns as policy instruments. *Journal of Policy Analysis and Management*, 13(1), 82–119. <https://doi.org/10.2307/3325092>

Weiss, C. H. (1977). Introduction. In C. H. Weiss (Ed.), *Using social research and public policy making*. D.C. Heath.

Whelan, F. G. (1996). *Edmund Burke and India*. University of Pittsburgh Press.

White, J. (2007). *Managing information in the public sector*. M.E. Sharpe.

Wieland, V., & Wolters, M. (2013). Forecasting and policy making. In G. Elliott & A. Timmermann (Eds.), *Handbook of economic forecasting* (Vol. 2, Part A, pp. 239–325). Elsevier. <https://doi.org/10.1016/B978-0-444-53683-9.00005-0>.

Wildavsky, A. (1979). *Speaking truth to power: The art and craft of policy analysis*. Little, Brown.

Wilson, S. (2020). Pandemic leadership: Lessons from New Zealand's approach to COVID-19. *Leadership*, 16(3), 279–293. <https://doi.org/10.1177/1742715020929151>

Yanow, D. (2000). Conducting interpretive policy analysis. *SAGE Publications*. <https://doi.org/10.4135/9781412983747>

Zittoun, P. (2019). The two Lasswells: Implications for critical policy studies. *Critical Policy Studies*, 13(2), 211–215. <https://doi.org/10.1080/19460171.2019.1620622>

Chapter 2

Beyond the Individual: Understanding the Evolution of Collective Intelligence



The institutions and professions threatened by disintermediation and the growth of transparency will only be able to survive and prosper in cyberspace by migrating their skills toward the development of collective intelligence and navigational aids. (...) It is said that we are becoming the neurons of a planetary hypercortex.

Pierre Lévy (1998)

2.1 Introduction: The Concept of Collective Intelligence

Collective intelligence (CI) is as old as civilization. Intelligent collaboration within communities and social groups, resulting in the collective creation of goods, and their accumulation within a shared culture, is widely considered to be humanity's defining characteristic (see e.g. Tomasello, 1999). Our advanced cognitive functions are a unique feature of humans that lie at the heart of our capacity for cooperation. While our hearing, vision, and manual skills may be similar to those of primates (e.g., chimpanzees), it is our exclusive use of complex language, our capacity to create sophisticated tools, and our ability to reason about intricate concepts—such as quantum physics, the common good, or transcendence—that sets humans apart. Highly developed products of civilization, significant scientific discoveries, exceptional cultural works, and the prudent governance of nations are achievements typically attributed to the high intelligence of the people responsible for them.

Intelligent behavior can be broadly defined as the ability to learn from experience, adapt to new situations, understand complex ideas, and employ various forms of reasoning to overcome challenges. It involves a range of cognitive processes, including but not limited to creative problem-solving, strategic decision-making, and applying knowledge to manipulate one's environment. Measures of these cognitive abilities are often operationalized through standardized assessments (like IQ tests) that aim to provide objective criteria for intelligence (Wechsler, 2008).

Robert Sternberg's triarchic theory of intelligence (1985) argues that intelligent thinking transcends mere IQ scores, encompassing a synergy of analytical, creative, and practical abilities that enable individuals to effectively navigate their environment. Similarly, Howard Gardner's theory of multiple intelligences suggests that intelligence is not a single, unitary construct but rather a compilation of distinct modalities of cognitive aptitude (Gardner, 1983). According to Gardner, these modalities range from logical-mathematical intelligence, which enables one to discern patterns and reason deductively, to linguistic intelligence, which pertains to the skill in using language to express oneself rhetorically or poetically. Further expanding upon these ideas, Goleman (1995) introduced the concept of emotional intelligence as the ability to recognize, understand, and manage our own emotions as well as to recognize, understand, and influence the emotions of others. This form of intelligence has been linked to success in social and professional domains, as much as or even more than traditional cognitive intelligence (Goleman, 1995).

Cooperation in groups, showing features of intelligent thinking that go beyond the capabilities of a single person, is one of the most important processes that has arisen since the dawn of humankind. The earliest forms of CI appeared about two million years ago and were most likely centered on the control of fire. Subsequently, CI concerned the emergence of language, tools, and intricate institutions around 200,000 years ago. The evolution of language enabled the human species to create increasingly sophisticated symbolic systems, which expanded the cognitive abilities of our species with each new medium (Peters, 2015). In ancient Greece, the Athenian Assembly was a gathering of all male citizens who met to make important decisions about the city-state. They would discuss and vote on issues such as war, trade, and taxation, and their decisions were considered as the will of the people. The Mayan civilization, on the other hand, collectively developed a complex and accurate calendar system that was based on a sophisticated understanding of astronomy and mathematics. This system required collaboration between astronomers, mathematicians, and religious leaders and allowed the Mayans to make predictions about celestial events with great accuracy. The collective intelligence of scholars and librarians who worked at the Library of Alexandria led to the preservation and dissemination of knowledge, which in turn had a profound impact on the development of science, philosophy, and literature in the ancient world. During the Middle Ages, the First Crusade stood out as a remarkable example of collective cooperation. Despite its controversies, this initiative showcased collective leadership and involved hundreds of thousands in both military and logistical collaboration, leading to significant civilizational impacts.

However, technological changes, especially the communication revolution of the late twentieth century, raised the interest in CI to a whole new level. The size of groups that can collaborate in real time has changed: thanks to internet communication we can track the immediate cooperation of thousands or even millions of people. The extended use of algorithms, computing, mass media, and online communication introduced new symbolic systems, decreasing the distance barrier and allowing for flexible, dynamic creation of communities consisting of dispersed participants. The digitization of communication has also allowed for more reliable measurements of

the parameters affecting group collaboration. The advent of advanced communication technologies influenced the development and application of multidisciplinary research on CI in many areas of social and political life.

Nowadays, collective intelligence is studied both as a feature of small groups, where ties and interactions between participants are strong and the deliberation process leads to the informed intellectual outputs (Bonabeau, 2009; Malone, 2018; Woolley et al., 2010), and as a statistical phenomenon resulting from the aggregation of a vast number of dispersed opinions from incoherent crowds (Howe, 2008; Surowiecki, 2005), or even the entire global internet community (Mayer-Kress & Barczys 1995; Engel & Malone, 2018). Inspired by both the growth of communication networks and the study of analogous patterns of behavior of intelligent communities in disciplines such as biology, economics, management, neurocognitive science, and artificial intelligence, studies on CI developed.

There are numerous definitions of CI in scientific literature. Pierre Lévy, a Tunisian-born French philosopher and one of the pioneers in the analysis of cyberspace, stated that CI is “a form of universally distributed intelligence, constantly enhanced, coordinated in real time, and resulting in the effective mobilization of the skills” (Lévy, 1997). Francis Heylighen, cyberneticist studying intelligent organizations, claimed that CI is the ability of a group to “find more or better solutions than the whole of all solutions that would be found by its members working individually” (Heylighen, 1999). According to Heylighen, if there is no collective cooperation in a social structure, the structure has a limited capability for solving a certain pool of problems: each person seeks solutions independently, and thus, there is neither positive nor negative interaction. He also notes, recalling an analogy with the life of insects, that the result of CI’s work is often not a single solution, but a kind of shared memory, which he calls the collective mental map (Heylighen, 1999). Tom Atlee and George Por argue that CI is older than humankind itself and primal forms of CI manifest in the synergies and resilience of ecosystems. They describe this process as “the wisdom of nature,” which “learns from its experience” through the evolutionary process of iterative creation and testing. In social animal groups such as ants, bees, specific fish and birds, and mammals like wolves and primates, this phenomenon is also observed (Atlee & Por, 2000).

In the field of management, collective intelligence surrounding group decision-making, knowledge management, and organizational learning is studied. Thomas W. Malone, the founding director of the MIT Center for Collective Intelligence, proposed a simplifying definition, describing CI as “groups of individuals acting collectively in ways that seem intelligent,” at the same time highlighting goals, motivation, and incentives, as well as cognitive processes related to this kind of work (Malone, 2018). A definition by Barry Wellman, a prominent sociologist and network scientist, emphasizes the importance of social networks and their influence on collective intelligence, claiming that CI emerges from the collaboration and competition of many individuals connected through social networks (Wellman, 2001). An alternative term is “crowdsourcing,” a business strategy with roots in the open-source movement. The term, a combination of the words *crowd* and *outsourcing*, highlights the idea of leveraging the wisdom of the crowd to achieve specific goals or outcomes. The genesis of

the term can be traced back to 2006 when Jeff Howe, a contributing editor at *Wired Magazine*, coined it in an article titled “The Rise of Crowdsourcing” (Howe, 2006). Biological sciences evoke the term *swarm intelligence* as self-organized behavior emerging from swarms of social insects in which “accumulated interactions lead to an irreversible outcome” (Gordon, 2015). The term *swarm intelligence* is also widely used in the domain of artificial intelligence, standing for decentralized, self-organized systems, inspired by social insects. The definition provided by Bonabeau et al. (1999) highlights the importance of agent interactions: “Swarm intelligence is a property of systems of unintelligent agents exhibiting collectively intelligent behavior by means of local interactions and simple behavior rule.” Moreover, while Anita Woolley, professor of organizational behavior, defined CI as “the general ability of a group to perform a wide variety of tasks” and proposed a general, single CI factor highly correlated with the level of cognitive diversity within a group (Woolley et al., 2010), several psychologists challenge this thesis, connecting CI with the average individual intelligence level of the group and describing CI level as a multi-dimensional vector rather than a single factor (Bates & Gupta, 2017; Crede & Howardson, 2017).

From the above review, we can observe a significant increase in interest in the subject of CI in the last 20 years. Simple analysis with the use of Google Ngram Viewer, an online search engine that charts the frequencies of search strings using a yearly count of n-grams in printed sources, shows that in the early 1980s the term CI was practically unheard of, while the term *artificial intelligence* was beginning to gain popularity (Fig. 2.1). Beginning in 2005, interest in CI began to gradually increase, which was further supported by the new term *crowdsourcing*, which became extremely popular after 2010.

Thus, one may ask: what factors drove the surge in interest in computer network-powered collective collaboration and its potential impacts during that period? How did this phenomenon begin? What do we currently understand about CI? To address these questions, we must journey back over half a century.

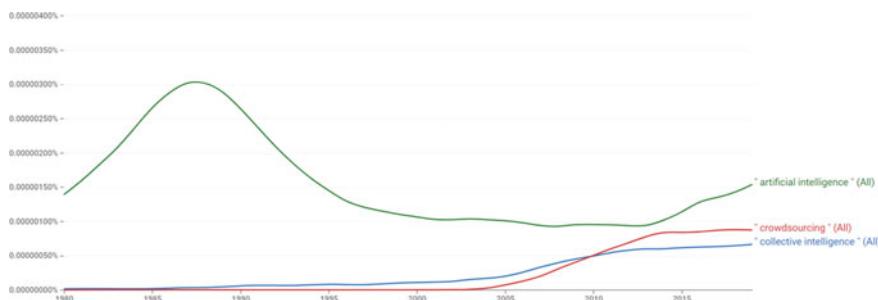


Fig. 2.1 Popularity of the concepts of AI, CI, and crowdsourcing in english-language literature, based on google N-gram query

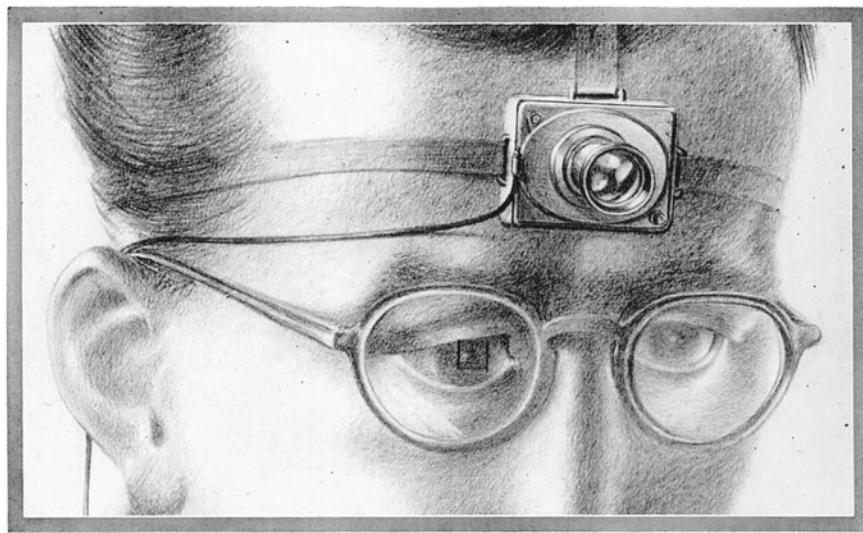
2.2 The Computer Mouse, Online Communities, and “Boosting Our Collective IQ”

In a dimly lit conference hall in San Francisco, an audience of over a thousand computer professionals gathered with anticipation. The year was 1968, and they were attending the Fall Joint Computer Conference at the Brooks Hall Auditorium. Douglas Engelbart, a then little-known engineer working for the Stanford Research Institute and funded by the ARPA agency, took the stage with a confident stride. The group he led was nicknamed “a research center for augmenting human intellect,” (Engelbart & English, 1968) and this presentation, later famous as “The Mother of All Demos,” was intended to expose their innovations to a wider audience.

The stage was set with a large projection screen, rows of seats brimming with attendees, and a workstation where Engelbart would showcase his projects. He began his presentation, describing his vision of augmenting human intelligence through the use of technology. Then, he unveiled the first of his innovations: a small, wooden device with two metal wheels and buttons on top. It was the computer mouse, a humble but groundbreaking invention that would redefine the way people interacted with computers. Engelbart connected the mouse to a large computer, the oN-Line System (NLS). With a flick of his wrist, he moved the device across a flat surface, and the on-screen cursor glided smoothly in response. The audience was in awe, realizing they were witnessing something truly extraordinary (Bardini, 2000).

Engelbart’s inspiration for the computer mouse came from his broader vision of augmenting human intellect and improving human-computer interaction. He believed that computers could be used as powerful tools for enhancing our problem-solving and collaborative abilities (Engelbart, 1962). His interest can be traced back to his 1945 encounter with Vannevar Bush’s article “As We May Think,” published in *The Atlantic Monthly*. In the article, Bush envisioned a device called the *Memex*, which he described as a desk-like system that would allow users to store, retrieve, and annotate information through microfilm technology. The main conclusion was that a machine could be used anywhere that a logical thought process occurs. Bush proposed the *Memex* as a solution to the issues of knowledge accessibility, hoping to create a sort of collective memory machine. Bush was very concerned with information overload inhibiting the research efforts of scientists, arising from the unprecedented demands on scientific production and technological application during World War II. He argued that scientists, operating under conditions of *information explosion*, require respite from the tide of scientific documents in an information saturated society. He hoped to transform an *information explosion* into a *knowledge explosion*. Bush envisioned the *knowledge explosion* as an expansion of human understanding and intellectual capacity. While the term *information explosion* implies a rapid increase in the amount of available data, the concept of a *knowledge explosion* highlights the importance of efficiently organizing, accessing, and interpreting this information to enhance human comprehension (Bush, 1945) (Fig. 2.2).

In the late 1950s and early 1960s, when Engelbart was developing his ideas for devices augmenting human intellect, computers were difficult to access and required



A SCIENTIST OF THE FUTURE RECORDS EXPERIMENTS WITH A TINY CAMERA FITTED WITH UNIVERSAL-FOCUS LENS. THE SMALL SQUARE IN THE EYEGLASS AT THE LEFT SIGHTS THE OBJECT

AS WE MAY THINK

A TOP U.S. SCIENTIST FORESEES A POSSIBLE FUTURE WORLD
IN WHICH MAN-MADE MACHINES WILL START TO THINK

Fig. 2.2 The illustration for Vannevar Bush's essay "As We May Think" (Bush, 1945). *Source* [https://commons.wikimedia.org/wiki/File:The_Memex_\(3002477109\).jpg](https://commons.wikimedia.org/wiki/File:The_Memex_(3002477109).jpg)

specialized knowledge to operate. Interaction with computers was limited to text-based command-line interfaces and input devices like punch cards or light pens. Engelbart recognized the need for a more intuitive and efficient way to interact with computers, particularly for tasks that involved the manipulation of on-screen objects and text. The idea for the mouse came to Engelbart during a conference presentation on computer graphics in 1961. While sitting in the audience, he began to imagine a more user-friendly input device that could move a cursor on the screen. He sketched a simple design on a notepad, which resembled a small box with two perpendicular wheels mounted on the bottom, allowing the device to track movement in both the X and Y directions. Engelbart's concept evolved into the first computer mouse prototype, developed with the help of his colleague Bill English at the Stanford Research Institute.

The computer mouse was just the beginning; the Augmentation Research Center had much more in store. As "The Mother of All Demos" demonstration continued, Engelbart showcased other designs like hypertext and collaborative real-time editing. One of the innovations presented was video conferencing, a technology that seemed like a futuristic dream at the time. Engelbart and his colleague in Menlo Park, connected via a video link, demonstrated the possibility of having face-to-face

conversations over long distances through a computer screen. Another revolutionary concept introduced was hypertext, the foundation upon which the World Wide Web would eventually be built. Engelbart showcased a new way to organize and navigate information by connecting pieces of text through clickable links, allowing users to easily traverse from one document or topic to another. This innovative approach to information management left the audience awestruck as they began to imagine the potential for an interconnected network of information. Perhaps one of the most captivating demonstrations was that of real-time collaborative editing (English et al., 1967). With the aid of the oN-Line System, Engelbart and his team showed how multiple users could work on the same document simultaneously, making edits and changes that were visible to all participants in real-time. The implications were profound, as the audience envisioned a future where barriers of time and distance would no longer impede collaboration and teamwork. In the days, months, and years following “The Mother of All Demos,” the impact of Engelbart’s inventions reverberated throughout the world of computing. The computer mouse and other tools to *augment intellect* became integral parts of modern computing, making it both more accessible and user-friendly.

It was the same for online group collaboration, which Engelbart envisioned as a new type of interaction designed to unleash a new type of intelligence. In the cooperation of communities connected by a computer network, Engelbart saw the best implementation of Bush’s idea of organizing, accessing, and interpreting information to achieve a *knowledge explosion*, i.e., to produce high-quality knowledge. Engelbart noticed that it was not the parameters of devices, but rather unlocking the intellectual potential of people using them, that was key to technological change. He later summarized his approach: “I’m not ‘numerically oriented;’ my vision has always facilitated discursive thinking and collaboration” (Engelbart, 1995).

These ideas were further developed in a 1972 article, “Coordinated Information Services for a Discipline- or Mission-Oriented Community.” In this paper, Engelbart proposed working in distributed communities of users connected through resource-sharing computer networks, which would consist of both technical core devices and diverse collectives (referred to as: *vendors, brokers, and customers*). He believed that this interconnected system would stimulate the growth of an information market, ultimately leading to a *knowledge society*. By combining the abilities of both computers and human intelligence, Engelbart proposed the emergence of the way of life “in an integrated domain, where hunches, cut-and-try, intangibles, and the human ‘feel for situation’ would coexist with powerful concepts, streamlined terminology and notation, sophisticated methods and high-powered electronic aids” (Provenzo, 1997). He later called this method “boosting our collective IQ” (Engelbart, 1995), with tools for augmenting not just individual worker’s knowledge, but also the knowledge of teams of people both coresident and distributed over the world who were interacting through a networked environment (Engelbart & Lehtman, 1988).

Engelbart’s vision foresaw a cultural shift that “will produce a discontinuity in our cultural evolution of a scale commensurate with that of the industrial revolution.” He recognized the potential for such *network-coupled, distributed community* when connected to the ARPAnet, the precursor to the modern-day Internet, and engaged in

collaborative dialogue, document development, and research intelligence (Engelbart, 1972). ARPAnet, the first high-speed computer network spanning across continents, was developed by the U.S. Department of Defense as a digital communications experiment. It expanded to connect numerous universities, defense contractors, and research institutions, significantly enhancing collaboration and accelerating technological progress through fast and flexible information exchange. For us, however, the most important is the additional effect of ARPAnet: it brought together hackers (referring to the original meaning of skilled programmers, not cybercriminals) from all over the nation who had previously been part of isolated groups with transient local communities. They created a critical mass. Instead of remaining in isolated small groups each developing their own ephemeral local cultures, they discovered themselves as a networked tribe (Raymond, 2001). It took a few years for online group collaboration to produce results that amazed the world, however.

2.3 Linux and the Open-Source Revolution

The history of Linux serves as a remarkable example of the power of online group collaboration, as its development is an unexpected result of collective efforts among enthusiasts and developers from around the world. In 1991, a Finnish computer science student named Linus Torvalds began working on a personal project to create a free and open-source operating system as an alternative to the existing environments, starting with a free Unix kernel for 386 PC machines (Fig. 2.3). The name Linux is a combination of his first name, Linus, and Unix, the operating system that inspired his projects. At the time, most operating systems were proprietary and expensive, but Torvalds wanted to create one that would be freely available to anyone who wanted to use it. In its early stages, Linux's development synergized with another phenomenon: the public discovery of the Internet. In the early 1990s, the internet-provider industry began to thrive, offering affordable connectivity to the general public. With the introduction of the World Wide Web, the Internet's growth rate increased dramatically, allowing groups of unorganized individuals who rarely met in person to collaboratively develop complex projects. Many of the early internet services, such as web hosting, email, and file sharing, relied on Linux-based servers to ensure continuous uptime and smooth operations. This has in turn promoted the development of many other open-source projects and technologies that have been instrumental in the popularization of the Internet, such as the Apache web server, the MySQL database system, and the PHP programming language. By late 1993, Linux could compete on stability and reliability with many commercial systems. With a strong focus on security and a large and active community of developers working together to identify and fix vulnerabilities, Linux-based servers were highly resistant to cyberattacks. The surprise of this success is reflected in the words of Eric S. Raymond, an experienced hacker of the older generation: "Who would have thought that a world-class operating system could coalesce as if by magic out of part-time

hacking by several thousand developers scattered all over the planet, connected only by the tenuous strands of the Internet?" (Raymond, 2001, p. 21).

Torvalds originally released Linux as free software under the GNU General Public License. This meant that anyone could use, modify, and redistribute his source code. Upon sharing the initial kernel source code on the internet, he allowed developers and enthusiasts worldwide to collaborate and contribute to the project, turning it into a collective effort where all willing people could collaborate, contribute, and provide feedback. This marked the beginning of the Linux project as a collaborative endeavor, which has reached an entirely unexpected scale. Thousands of developers worked to integrate GNU components with the Linux kernel, creating a fully functional and free operating system.

The underlying idea of Linux was inspired by the concept of free software, which centers on making the source code of a software program freely available for anyone to access, modify, and redistribute. The free software movement, initiated by hackers, aimed to preserve the freedom and collaborative spirit of software development that was prevalent in the early days of computing. Back then, software was often shared among researchers and developers without legal or financial restrictions. This sharing culture, however, was gradually replaced by the commercialization of software and the imposition of proprietary licenses. In response to the rise of proprietary systems, the free software movement, led by Richard Stallman, emerged in the early 1980s.

```
From: torvalds@klaava.Helsinki.FI (Linus Benedict Torvalds)
Newsgroups: comp.os.minix
Subject: What would you like to see most in minix?
Summary: small poll for my new operating system
Message-ID:
Date: 25 Aug 91 20:57:08 GMT
Organization: University of Helsinki

Hello everybody out there using minix -

I'm doing a (free) operating system (just a hobby, won't be big and
professional like gnu) for 386(486) AT clones. This has been brewing
since april, and is starting to get ready. I'd like any feedback on
things people like/dislike in minix, as my OS resembles it somewhat
(same physical layout of the file-system (due to practical reasons)
among other things).

I've currently ported bash(1.08) and gcc(1.40), and things seem to work.
This implies that I'll get something practical within a few months, and
I'd like to know what features most people would want. Any suggestions
are welcome, but I won't promise I'll implement them :-)

Linus (torvalds@kruuna.helsinki.fi)

PS. Yes - it's free of any minix code, and it has a multi-threaded fs.
It is NOT portable (uses 386 task switching etc), and it probably never
will support anything other than AT-harddisks, as that's all I have :-.
```

Fig. 2.3 This is how Linus Torvalds' original post, inviting collaboration on the development of a new system, might have appeared on computer screens of that era. The figure retains the text of the original email, preserving its authentic spelling

Stallman founded the Free Software Foundation (FSF) and created the GNU General Public License (GPL) to ensure that software could be freely used, modified, and redistributed while protecting it from becoming proprietary (Williams, 2010, p. 128). The most significant achievements of this movement include the GNU Project, which, launched in 1983, aimed to develop a complete Unix-compatible operating system, and Emacs, which was introduced in 1984 as a powerful, extensible, and customizable text editor (short for “Editor MACroS”). The GNU Project, in particular, played a pivotal role in the development of the Linux operating system. The Linux kernel was combined with the GNU tools and utilities to create a fully functional operating system. The scope of the Linux project, which is still ongoing, is impressive. According to the Linux Foundation’s annual kernel development report, each kernel release involves contributions from over 1,000 developers representing more than 200 different corporations. Over time, this number has grown with more developers joining the project and contributing to its ongoing evolution. According to the Linux Foundation, approximately 777,000 developers have contributed to the projects realized in this community since its inception (The Linux Foundation, 2023).

Many observers, however, asked why it was Linux, which emerged from a spontaneous movement of young enthusiasts, and not a much older GNU project meticulously built by experienced hackers, that became a worldwide phenomenon. In fact, the project initiated by Linus Torvalds was initially more of a conglomerate than a separate operating system; although, unlike its predecessors, it was not written for demanding Unix machines, but for the increasingly popular PC microcomputers. Linux contained all the best products and free software created over the years: BIND (introducing verbal domain names instead of numerical IP addresses), the PPP communication protocol operating in the TCP/IP model, the GNOME graphical overlay, and others (Williams, 2010, p. 155). Soon it was Linux, and not GNU, that became the benchmark for open projects created by a dispersed community, and, at the same time, a serious competitor to commercial systems. What could have been the reasons?

Both projects were distinguished by the transparency and lack of financial motivations of the participants. Linux, as well as GNU, was driven by passion and prestige rather than monetary gains. A diverse group of people participated in its creation, including hobbyists, computer science students, professional programmers, and researchers. Motivated by the desire to learn, solve technical challenges, and contribute to an open and free alternative to proprietary operating systems, these individuals were drawn to Linux and GNU projects by their shared passion for innovation and the prestige associated with being part of a groundbreaking effort in the field of computing. The sharing of knowledge and ideas among developers meant publishing the full programming code so that any interested person with technical competence could understand its operation (Williams, 2010, p. 111). Stallman, an ideological opponent of using any commercial IT products, believed that software *by definition* should be free and that the culture of sharing is one of the most important development levers: “sharing knowledge is the most fundamental act of friendship; because it is a way you can give something without losing something” (Royse, 2020). The motivation to contribute to these projects came primarily from passion, a sense of

mission, and a sense of shared identity within the hacker community. As proponents of software “liberation” believed, fascination with the problem itself, both in software as in other kinds of creative work, is a far more effective motivator than money alone (Raymond, 2001, p. 59). Raymond, one of the prominent representatives of the open-source movement, argues:

Reporters often ask me these days if I think the open-source community will be corrupted by the influx of big money. I tell them what I believe, which is this: commercial demand for programmers has been so intense for so long that anyone who can be seriously distracted by money is already gone. Our community has been self-selected for caring about other things —**accomplishment, pride, artistic passion, and each other** (Raymond, 1999).

What’s more, the motivation to maintain a positive reputation persists regardless of a participant’s conscious efforts or awareness, what can be heard in the words of an anonymous hacker: “you may not work to get reputation, but the reputation is a real payment with consequences if you do the job well” (Raymond, 2001, p. 83). Prestige also has an important social dimension, as it is a good way to attract attention and cooperation from others: “if one is well known for generosity, intelligence, fair dealing, leadership ability, or other good qualities, it becomes much easier to persuade other people that they will gain by association with you” (Raymond, 2001, p. 84).

The characteristics of Linux that distinguish it from GNU were outlined by Raymond in his famous essay, “The Cathedral and the Bazaar” (Raymond, 2001). In this work, two distinct software development approaches are presented. The first of them, called *building cathedrals*, consists of systematic work on the implementation of a project based on the vision of a remarkable *architect*. Stallman, according to this opinion, had the features of a classic cathedral architect: a programming *wizard* who could disappear for several months and return with an extraordinary design (e.g., the GNU C Compiler), leaving a distinct mark of his personality on his creations. The source code in this kind of projects was transparent with each software release, but code development between releases was restricted to an exclusive group of insiders. GNU, Emacs, and GCC were examples of *cathedrals*, or monuments of hacker ethics (Williams, 2010, p. 164). Despite the fact that a large community could benefit from the results of the work, the most important part of the creative process was essentially individual.

The prevailing paradigm in software development, influencing both commercial programmers and hackers, has thus far been Brooks’ Law, which states that introducing additional workforce to a project can make it less effective. This concept was first formulated by Fred Brooks in his 1975 book, *The Mythical Man-Month* (Brooks, 1975). The author contended that, in specific situations, incorporating extra individuals into a team results in a longer, rather than shorter, completion time. Raymond recalls that the success of Linux, which was built in a completely different way, took him by surprise:

Linux overturned much of what I thought I knew (...). I believed there was a certain critical complexity above which a more centralized, *a priori* approach was required. I believed that the most important software (operating systems and really large tools like the Emacs programming editor) needed to be built like cathedrals, carefully crafted by individual wizards or

small bands of mages working in splendid isolation, with no beta to be released before its time (Raymond, 2001, p. 21).

Raymond credits Linus Torvalds as the inventor of the second model, called *the Bazaar*, a community-driven approach where the code is developed publicly over the Internet. The Linux community appeared like a bustling, seemingly chaotic, diverse marketplace, where various groups had their own ideas and methods for creating software. This open exchange of ideas and collaboration allowed for rapid development and a continuous flow of fresh perspectives. There was no clear leader and everyone could freely propose their solutions, getting feedback quickly, thanks to frequent publishing and testing of the updated code. Torvalds did not impose his vision because he did not consider himself an authority. Instead, he valued his community; he believed that delegating both creative work and verification of results and co-deciding on the future of the project with them was the best solution. Raymond recalls how unexpected the success of the project was for him and the other developers:

Linus Torvalds's style of development—release early and often, delegate everything you can, be open to the point of promiscuity — came as a surprise. No quiet, reverent cathedral-building here — rather, the Linux community seemed to resemble a great babbling bazaar of differing agendas and approaches (...). The fact that this bazaar style seemed to work, and work well, came as a distinct shock. I worked hard (...) trying to understand why the Linux world not only didn't fly apart in confusion but seemed to go from strength to strength at a speed barely imaginable to cathedral-builders (Raymond, 2001, pp. 21–22).

The essential feature of Linux was not its technical aspects, but rather its social dynamics. The core of this project was not the construction of the Linux kernel itself but the innovation of a novel development model. Torvalds summarizes this idea with self-reflection: “I’m basically a very lazy person who likes to get credit for things other people actually do” (Raymond, 2001, p. 27). Before Linux, the conventional view was that a small, tight-knit group was required for the careful and coordinated development of any complex software like an operating system. Linux, however, was developed in a unique way, with a large number of volunteers working together solely through the Internet, maintaining quality not through strict rules or authoritative control but by frequently releasing new versions and gathering feedback from hundreds of users within days, which proved to be surprisingly successful. Compared to cathedral-builders, Torvalds was more like a genial party host: “In letting others lead the Linux design discussion and stepping in only when the entire table needed a referee, Torvalds had created a development model where the most important managerial task was not imposing control but keeping the ideas flowing” (Williams, 2010, p. 164).

Linux remains one of the most inspiring examples of collaborative work to this day. The scale of the project and, most importantly, the value of the results produced are measurable and empirically verifiable, unlike many past and present social phenomena in which harnessing the attention and brainpower of entire communities takes place. The opinions of people involved in the open-source movement indicate that the strength of this project primarily lies in common bugfixes and sharing the results of this work. It is not the creation that is most important, because elements of

Linux had already been created in the GNU project and many other smaller initiatives. It is public testing, scrutiny, and experimentation that allowed for the evolutionary development of a coherent whole and rapid detection of errors (bug-spotting). In contrast, in the *cathedral model*, working versions of the code are available only to a few developers, so hunting for bugs is labor intensive and imperfect. Torvalds claims that detecting and correcting bugs is a highly collaborative effort, and the person who understands and fixes the problem is usually not the person who first characterizes it: “Somebody finds the problem, and somebody else understands it. And I’ll go on record as saying that finding it is the bigger challenge.” Speaking less formally, “given enough eyeballs, all bugs are shallow” (Raymond, 2001, p. 30).

Early and frequent releases are *the* critical part of the Linux development model that keep community members stimulated and rewarded so that they could engage in a continuous testing and debugging process. Instantly introducing the results of work to the public forum and subjecting them to discussion, according to Shaun Gallagher’s theory, affects the *sense of agency* that can be defined as “the sense that I am the one who is causing or generating an action” (Gallagher, 2000). Further, the hackers were stimulated by the idea of obtaining a sense of self-importance from their actions and rewarded by the sight of constant (even daily) improvement in their work. They felt a deep sense of involvement in an extraordinary and crucial project, believing that their efforts significantly contributed to its success.

Although the community was open to virtually anyone interested, in practice, the technical skills required to understand the programming code served as a barrier to entry. Potential members had to possess a certain level of proficiency and demonstrate an interest in putting in the necessary work. Project participants had to *speak a common language* in order to understand each other and recognize basic concepts. In this case, the common speech was the C programming language. They had also to follow certain rules. For example, even though equality was valued in the hacker community, the principle that enabled resolving conflicts was seniority (Raymond, 2001, p. 103). If two contributors or groups had a dispute, and the dispute could not be resolved objectively, the side that had put the most work into the project as a whole (or the side with the most property rights in the whole project) decided. In fact, the Linux community was a very strict meritocracy, where “one’s work is one’s statement, [and] the best craftsmanship wins” (Raymond, 2001, p. 89). The participants shared a strong ethos that quality should be left to speak for itself.

Without the ability to communicate and without a common hacker ethos, this project would certainly not have achieved such success. The community’s strong sense of mission and distinct identity effectively connected individual selfishness to challenging objectives that required persistent cooperation. Community members retained their individuality and fueled their egos by attaining prestige, all while dedicating substantial commitment to the common good—the system they were developing. The aspiration to earn a reputation within the community and a shared passion for crafting *beautiful software* brought individual contributions together into a cohesive, collaborative endeavor.

2.4 Global Brain: A Utopian Vision of the Collective Mind

2.4.1 *Pierre Lévy*

The emergence of the Internet as a new space for human activity and the spread of online collaboration in the form of projects like Linux has captured the imagination of many observers of social processes. One of them was Pierre Lévy, a French philosopher and sociologist who successfully introduced into public discourse the concept of collective intelligence. He imagined CI as a description for the new dimension of human activity using IT networks, and at the same time a paradigm of thinking, communication, and cooperation of the whole humankind. The main difference between Lévy and Engelbart or the hacker community was that the latter were coming from engineering and business environment, while Lévy was deeply immersed in the European philosophical and sociological tradition. In his 1994 book *L'intelligence collective: Pour une anthropologie du cyberspace (Collective Intelligence: Mankind's Emerging World in Cyberspace)* Lévy presents a vision of CI centered around the idea that human beings, when connected through digital networks and technology, could pool their knowledge and skills to create a more powerful and intelligent entity. He believed that the Internet could function as a catalyst for knowledge sharing, collaboration, and problem-solving on a massive scale, transcending geographical, cultural, and linguistic barriers. Lévy suggests that the most important purpose of new forms of communication is to create the tools necessary for sharing our mental capabilities in the formation of a collective intellect and imagination. The technical infrastructure of the Internet can provide a foundation for the collective minds of intelligent communities, which can then mutually develop and enhance our social and cognitive potential. According to Lévy, the primary goal for the future is to envision, construct, and improve an interactive, constantly evolving cyberspace based on this principle. These novel forms of communication have the potential to significantly alter the framework of social bonds and address the challenges that humanity currently confronts. The emergence of the Internet allowed knowledge to be shared and recognized by others, creating a necessary foundation for collective thought and envisioning a future where humanity could address complex challenges and create a more equitable, democratic, and innovative society.

Lévy's book can be seen as a utopian manifesto, as it presents a comprehensive system of utopian ideals that are founded on emerging models of computer-assisted communication. Collective intelligence is here understood not only as a cognitive process, but also as a global project joining people in a new way. The word *intelligence* itself comes from the Latin *inter legere* (joining together)—uniting knowledge and people to construct a new form of society. This new kind of society will be deterritorialized and based on establishing social bonds as the foundation of our relationship to knowledge. Therefore, Lévy defines CI as a form of “universally distributed intelligence, constantly enhanced, coordinated in real time, and resulting in the effective mobilization of skills” (Lévy, 1997, p. 13). The notion of the universal distribution of intelligence is summarized by phrase: “No one knows everything, everyone knows

something, all knowledge resides in humanity" (Lévy, 1997, p. 14). Lévy believes that the concept of CI implies the technical, economic, legal, and human enhancement of a universally distributed intelligence, leading to a general skill development. In order to coordinate intelligence in real-time, communication methods must be based on digital transformation technologies. The development of new communication systems should enable community members to coordinate their interactions in the virtual world of knowledge. Dynamic maps of shared context would represent events, decisions, actions, and individuals and would constantly change the universe. In this way, cyberspace would transform into a constantly changing space where knowledge and people interact within intelligent communities that exist beyond physical borders. CI is seen here as a great scientific, technical, and political project that aims to "make people smarter with computers, instead of trying to make computers smarter than people" (Peters, 2015). Lévy views CI not as the opposite of *collective stupidity* or *individual intelligence*, but as the necessary extension of *artificial intelligence*, which allowed for the development of a renewed cultural cognitive system that could leverage the problem-solving ability and ubiquitous memory provided by networked computers (Peters, 2015).

Lévy, the author of *L'intelligence collective*, had a vision that humanity will reach a new stage of evolution by creating a new human quality that is as fundamental as language but operates at a much higher level than traditional media or bureaucratic institutions. Reflecting upon the history of civilization, including the breakthroughs caused by the invention of writing, the spread of commerce, etc., Lévy argues that computer networks are another breakthrough, the beginning of a new era. The Internet has made the entire corpus of cultural artefacts ubiquitous, interconnected, and capable of being automatically transformed using algorithms that fetch, translate, and modify information. This feature of the Internet sets it apart from mass media, which automates copying, and static writing, which is self-preserving. Most of the previous symbolic systems, however, came from the era of static writing and mass media. According to Lévy, online communication, interconnected using hyperlinks, reflects the impending symbolic revolution. Because knowledge is becoming the key factor of the world, a new anthropological space, called *knowledge space* is being formed. It is presented as the new stage of civilizational development, following ecological space (land), territorial space (states and landed goods), and commodity space (economy and trade): "Knowledge space" that is being formed today, will almost certainly "take precedence over the spaces of earth, territory and commerce that preceded it" (Lévy, 1997, p. 5). "It is a part of cosmopolitan and borderless space of relations and qualities, (...) a space in which the processes of individual and collective subjunctionalization come together. (...) This space will harbor forms of self-organization and sociability" (Lévy, 1997, p. 141). Within the knowledge space, humanity is supposed to become a *nomadic* tribe once again and pluralize its identity. According to Lévy, the inhabitants of the knowledge space will have a distributed and nomadic identity, different from identities based on belonging (to a territory or a commodity space). It is supposed to be a return to a time before the division into territories, trade, and property rights. In this space, there will be enough room for everyone and cost will not be a barrier. Everyone will be able to use the full resources

of the space and travel as a virtual nomad. Instead of a *symbolic* life (as in literature and the media), it is supposed to be a *real* life consisting of constant participation in public life via the Internet (Lévy, 1997, p. 168).

In Lévy's view, the spread of internet technologies in the knowledge space has the potential to "promote the construction of intelligent communities in which our social and cognitive potential can be mutually developed and enhanced" (Lévy, 1997, p. 17). The self-organized communities (*molecular* in Lévy's terms) that he expects to appear in cyberspace are supposed to realize the ideal of direct democracy. They will make use of every human act, enhance individual qualities, integrate people with each other and with technologies and create synergy between creativity, initiative, and diversity of skills. (Lévy, 1997, pp. 51–54). The members of molecular communities (*cyberspace cities*) will communicate "laterally, reciprocally, outside categories and hierarchies, folding and refolding, weaving and reweaving, complicating their metaphoric fabric of their peaceful cities" (Lévy, 1997, p. 55). Lévy claims that in this way we are entering a period of accelerated deterritorialization—creating new forms of political and social regulations (Lévy, 1997, p. 57). Within a community, cyberspace has the potential to become the most seamlessly integrated medium for analyzing problems, engaging in group discussions, developing awareness of complex processes, making collective decisions, and evaluating outcomes. But Lévy believes that current forms of government are not suitable for this new model. To address these challenges, digital simulation technologies, real-time information access, interactive communication methods, intelligent tools for filtering data, navigation through information streams, and mutual recognition of individuals and groups based on their activities and skills are necessary (Lévy, 1997, pp. 60–61). In this idealistic vision, real-time democracy is the opposite of the demagoguery seen in live-action broadcasts and the impulsive behavior of crowds. Instead, it involves the slow, but ongoing, development of a collective and interactive debate, in which everyone can contribute by posing questions, taking positions, proposing and weighing arguments, and making and evaluating decisions.

2.4.2 Around Lévy's Thought: Origins and Development of the Global Brain Concept

Lévy presented his inspiring, though strictly theoretical, utopian vision of a global mind in *L'intelligence collective*, a book that became a publishing success and popularized the concept of CI beyond a narrow circle of enthusiasts. The ideas cited in his book have extensive sources and a longer history. Some of them, derived from the engineering approach to computer networks, have been presented in previous sections of this chapter. But neuroscience-inspired and futurological visions of the planetary *global brain* have also been explored by several thinkers, including Vladimir Vernadsky (1863–1945) and Pierre Teilhard de Chardin (1881–1955), who coined the term *noosphere*, as well as H.G. Wells (1866–1946), who envisioned a *world brain*

(Wells, 1938). Building on these antecedents, contemporary authors such Barbara Marx Hubbard popularized the term *conscious evolution*, motivated by the prospect of a *noosphere*, imagined as a rapidly evolving, globally interconnected network of human minds, communicating and collaborating to solve problems and create new knowledge (Hubbard, 1998). In this view, the *noosphere* would function like an informational cortex of the planet.

Coined by Peter Russell in 1983, the term *global brain* imagines as a complex system, comprised of the emerging intelligent network formed by all people on this planet and the computers and communication links that connect them together. The metaphor of the information network as global brain was extended to the whole of society as a global organism. According to Rusell, and many *complex systems* researchers, like G. Mayer-Kress, F. Heylighen, J. de Rosnay, V. Turchin, J. Bollen and B. Goertzel, if the information processes in the network constitute the *mind* of this system, all people together with their artefacts form its *body*. Since individual people are organisms themselves, this encompassing system is supposed to be an organism consisting of organisms, that is, a super-organism or *super-being* (Turchin, 1977). The authors use the achievements of philosophy of mind and cognitive psychology, describing processes in the real human brain (e.g. learning, associations, making decisions), to attempt to describe what the imagined *global brain* would be and how thinking processes would take place in it.

The development of internet technologies made this theoretical concept applicable to reality. Additionally, concepts derived from artificial intelligence, such as software agents, were used to metaphorically describe these collective thinking processes. The Principia Cybernetica Project (see <http://pespmc1.vub.ac.be/>), followed by The Global Brain Institute (see <https://globalbraininstitute.org/>) emerged from this community of researchers in 1993. One of them, Francis Heylighen, argues that within the analogy of society as a macro-organism, the online communication pathways function like nerves, relaying messages between various organs and muscles (Heylighen et al., 1996). Unlike prior understandings, there is no center of leadership in this group organization. Data on the Internet is not centrally controlled, but rather distributed over millions of individuals and documents, interlinked by billions of connections. Heylighen and other Principia Cybernetica researchers assume that cognitive processes at the planetary level can force all this chaotic, heterogeneous information to interact, facilitating the emergence of collective patterns. Therefore, the global superorganism, superior to both individual people and existing social groups and nations in intelligence, was supposed to be the end result of an evolution towards ever-stronger interconnections between humans, software and machines across the planet. This would allow the capacity of humankind to evolve towards higher order complexity and harmony.

Reading these words today, many of the utopian expectations of the global brain presented by Lévy and other authors seem exaggerated or illegitimate, but it is impossible to deny their acuteness in capturing the importance of the phenomenon. We must remember that most of these texts were written before the era of Web 2.0 and social media, when it was not so obvious that the Internet would enable widespread participation on a large scale. In this sense, the presented visions can be considered, in a

way, prophetic. But only in a way. After all, today it would be difficult to say that, thanks to the development of the Internet, the world has risen to a higher level of meta-consciousness. Utopia has not been realized, there has not been a thorough revolution of the existing order as the utopians had expected. Instead, there has been a transfer of many human activities to the Internet without the expected coordination of their activities into one global mind. Equally important to us, these works—especially Lévy’s immensely popular book—introduced the term *collective intelligence* into the language of internet sociology and research, resulting in a great increase in interest in problems outlined in this vision and the emergence of a new generation of researchers addressing these issues in great detail.

Finally, the visionary concept of a new epoch in human development, one where individuals engage in a worldwide information network that elevates collective intelligence, has spurred enthusiasm for the development of a new kind of Internet project. These projects, drawing on the contributions of vast numbers of participants to address commercial, social, and scientific challenges, gained momentum in the 2000s. This emerging trend came to be recognized as “crowdsourcing.”

2.5 From the Wisdom of Crowds to Crowdsourcing Projects

2.5.1 *The Difference Between Crowdsourcing and Collective Intelligence*

How is *crowdsourcing* different from *collective intelligence*? Crowdsourcing is an open call for anyone to participate in an online task by submitting information, knowledge, or talent. As we will see in this section, there are several specific features of crowdsourcing, e.g., the huge size of the population that participates in it (*crowds*) and the mode of operation that is less focused on interactions between people and more on aggregating the creations of individuals. However, there is a close affinity between the concepts of crowdsourcing and CI which can be seen in their interchangeable application to the same phenomena and projects. A prime example is the experimental off-road traffic law-making initiative in Finland. Launched in 2013 by the Finnish Ministry of the Environment and the Committee for the Future of the Finnish Parliament, this project addressed concerns associated with motor-powered vehicles traversing natural landscapes beyond the confines of established roads, primarily involving snowmobiles and ATVs. In this project, the public was involved in the lawmaking process, first by being asked to voice their concerns, experiences, and problems regarding off-road traffic, and second by being asked to propose solutions. This comprehensive project has been extensively documented in a collection of papers by Tanja Aitamurto, a researcher at Stanford University. Intriguingly, some of these publications employ terminology specific to collective intelligence research (Aitamurto, 2016; Aitamurto et al., 2014), while others refer exclusively to

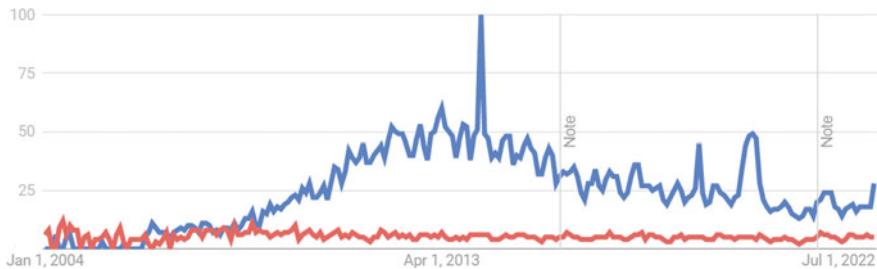


Fig. 2.4 Comparison of searches for the terms “*Crowdsourcing*” (blue) and “*Collective Intelligence*” (red) in Google Search in the years 2004–2023. Data retrieved from Google Trends 2023-05-13

the concept of crowdsourcing (Aitamurto & Chen, 2017; Aitamurto & Landemore, 2016).

Since these terms can be used interchangeably in certain cases, why do some authors choose to use the former and others the latter? Crowdsourcing as a concept primarily originated from business initiatives aimed at creating value for the market through the end product or service. The term *crowdsourcing* as a neologism is catchy and unique, although its use can also cause conceptual confusion because many average internet users confuse it with crowdfunding, which is indeed a subtype of crowdsourcing projects, but not the only one. Since its appearance in the early 2000s, the term *crowdsourcing* has gained immense popularity around the world, while CI has always remained a niche term. A comparative summary of the uses of these terms in the English-language literature is presented at the beginning of this chapter in Fig. 2.1. Another comparison can be seen in Fig. 2.4, showing the popularity of both terms in the Google search engine (based on Google Trends). You can see here a huge leap in popularity in the spring of 2014, when the crowdfunding platform Kickstarter reached an important milestone: more than \$1 billion had been pledged to projects on the platform.

2.5.2 *From Madness to Wisdom of Crowds*

We ought to remember that the term *crowd* has for a long time been given a negative connotation both in social sciences and in popular language, particularly in the context of crowd psychology or crowd behavior. The roots of this perception can be traced back to the nineteenth and early twentieth centuries when some of the first theories of crowd behavior were developed. In his 1841 work, *Extraordinary Popular Delusions and the Madness of Crowds*, Scottish journalist Charles Mackay presents a comprehensive study of crowd psychology, underscoring the irrationality and impulsiveness of collective human behavior. Mackay explores how individuals,

when part of a crowd, could be swayed into adopting beliefs and participating in actions that they might not engage in individually.

Mackay provides numerous historical examples to illustrate his points. For instance, he delves into the hysteria and financial ruin surrounding the South Sea Bubble, a speculative frenzy in the early eighteenth-century Britain. He also discusses the Dutch Tulip Mania of the seventeenth century, in which people invested extraordinarily high amounts into tulip bulbs, creating an economic bubble that eventually burst (Mackay, 1980). Mackay's work illustrates the potential dangers of herd mentality and the susceptibility of individuals to suggestion and manipulation in crowd situations. While the view of medieval history in this work is tainted by a nineteenth-century sense of superiority and may raise concerns about the accuracy of facts, it is an important example of the perception of crowds as a threat at the time.

Gustave Le Bon, another pioneer in this field, viewed crowds as irrational and susceptible to manipulation. His 1895 work, *The Crowd: A Study of the Popular Mind*, suggests that individuals in a crowd lose their sense of personal responsibility and rationality and that a sort of *group mind* takes over, leading to phenomena such as mob violence and riots. Le Bon believes that crowds lend a sense of anonymity, making individuals feel invincible and less accountable. This anonymity, coupled with the rapid spread of emotions and behaviors—a phenomenon he terms *contagion*—can lead to herd mentality, even when actions contradict personal beliefs. Le Bon also warns of crowds' suggestibility, asserting they can be easily manipulated, often by skillful leaders or orators, towards potentially harmful outcomes. Above all, he portrays crowds as largely irrational, driven more by emotion than by reason, leading to potentially destructive behavior (Le Bon, 1896).

James Surowiecki's book *The Wisdom of Crowds*, published in 2004, represents a significant shift in how we perceive crowds. Whereas earlier authors often viewed crowds as irrational and prone to hysteria, Surowiecki posits that under the right circumstances, crowds could be incredibly wise and that their aggregated judgments could often be more accurate than those of any individual member, even experts. He proposes that when the individual judgments of a diverse group are aggregated properly, the resulting collective prediction or decision is often remarkably accurate. This insight stems from a simple but profound realization: the aggregation of a multitude of independent estimates, even if many are wildly inaccurate, often converges towards the true value. Surowiecki argues that large groups of people, when their individual judgments are aggregated properly, can produce accurate predictions and make sound decisions.

It's important to differentiate between the type of crowd that Surowiecki refers to and the one that Le Bon describes. The latter is a group of people gathered for some reason in one place, and its actions are unpredictable. The reason may be one of common interest, as in the case of trade unionists protesting in front of the Prime Minister's office. It can be conflicting interests, as in the case of people storming supermarkets on Black Friday. Or a mere twist of fate, as in the case of the onlookers watching the scene of the accident. Each of these crowds can, under certain circumstances, become an unpredictable, impulsive group, and gathering in one place triggers psychological mechanisms that, according to Le Bon, can lead to a distortion

of the perception of reality. In turn, the crowd in Surowiecki's works is understood very broadly: as a collection of individuals who have an opinion on a given issue or undertake certain activities in a given field. They do not have to physically gather in one place because modern communication can be the platform connecting them. In this sense, a crowd is a community of Wikipedia editors, consumers, or stock market investors, any group consisting of people who do not know each other but are united by a common cause.

As Surowiecki claims, collectively given answers to questions related to spatial reasoning, estimation of size, or general knowledge of reality are no worse (and sometimes even better) than the answers given by individual members of the group who are experts. To illustrate his point, he drew upon an array of examples, from attendees at a country fair accurately guessing the weight of an ox to the aggregated predictions of sports enthusiasts outperforming those of professional sports analysts. The example of attendees at a country fair comes from a famous experiment conducted by the scientist Sir Francis Galton in 1906. Galton was interested in understanding how accurate the *average voter* was when it came to estimating values. He found his answer at a country fair in Plymouth, where attendees were invited to guess the weight of an ox after it had been slaughtered and dressed. Around 800 people, both experts and non-experts, participated in the contest. The guesses varied widely, ranging from under 1,000 pounds to over 1,400 pounds. Yet, when Galton calculated the mean of these guesses, he found that it was 1,197 pounds. Astonishingly, the actual weight of the ox was 1,198 pounds. This result showed that while individual guesses varied widely, the collective judgment of the crowd, represented by the mean of their guesses, was remarkably accurate—only one pound off from the actual weight. This example is used by Surowiecki to illustrate the surprising accuracy of collective judgment under the right conditions. In the described case of ox weight estimation, the conditions were that the guesses were independent, diverse, decentralized, and could be aggregated (Surowiecki, 2005).

Surowiecki identified four conditions as essential for a crowd to be *wise*. The first of them is *diversity of opinion*, which requires that each person should possess private information, even if it only constitutes an eccentric interpretation of commonly known facts. The second condition is *independence*, meaning that individuals' opinions should not be overly influenced by those around them. The third, *decentralization*, suggests that individuals should have the ability to specialize and leverage local knowledge. Lastly, *aggregation* posits that a mechanism must be in place to transform individual judgments into a collective decision. Surowiecki argues that if the conditions of diversity, independence, and decentralization are maintained, the average opinion of the people forming the community will be close to the truth and the final result of their actions will be more effective than in the case of individual actions. In addition, members of a given group must be motivated to give the right answer and have at least basic knowledge of the topic. This thesis is supported by many contemporary examples. One of them is the story of the Space Shuttle Challenger disaster in 1986. Using investment mechanisms, an independent and anonymous group of stock market investors accurately predicted that Morton Thiokol was to blame for the disaster. The share price of this company fell much lower on the day

of the disaster than three other companies involved in the production of components and their findings were proved by a special investigative commission six months later (Surowiecki, 2005, p. 7). According to Surowiecki, stock market investors met the criteria of differentiation and decentralization at that time and were independent in making decisions. Unfortunately, this independence was found to be lacking in a similar case a few years later, when the Space Shuttle Columbia disaster occurred. Then, guided by the experience of the Challenger case, investors assumed a similar scenario. Therefore, they did not show independence of thinking, but took mental shortcuts which turned out to be wrong (Surowiecki, 2005, pp. 173–191).

The inspiring stories given in Surowiecki's book can be countered by a whole host of examples of completely opposite phenomena, when the wisdom of the crowds did not work. Modern financial markets remind us of situations where too many consumers or investors replicate herd behavior, the perspicacity of the crowd loses its power, and a market of emotions and vanity begins, best exemplified by stock market bubbles and financial pyramids just before their collapse. How exactly to recognize and ensure favorable conditions for the emergence of crowd wisdom remains an open question. The value of Surowiecki's observations, however, is the recognition of the impact of diversity, independence, decentralization, and aggregation on the group's performance. In the following years these observations became the subject of many research projects, both with the use of empirical science as well as modeling and simulations. An even more significant influence of this work is the interest in consciously harnessing the *power of crowds* in business, which was presented by Jeff Howe in 2006.

2.5.3 *The Rise of Crowdsourcing*

Howe, a journalist and contributing editor at Wired Magazine, first coined the term *crowdsourcing* in the article “The Rise of Crowdsourcing” (Howe, 2006). He later expanded on this concept in a book titled *Crowdsourcing: Why the Power of the Crowd Is Driving the Future of Business* (Howe, 2008). In this initial conception, crowdsourcing is described as an online, distributed, problem-solving, and production model. In this model, businesses tap into the collective intelligence of a large group of people—a *crowd*—to perform specific tasks, solve problems, or generate ideas. The main incentive for business organizations, called here a *crowdsourcers*, to engage in these types of initiatives is the assumption that the tasks can be done more efficiently or at a lower cost than using traditional methods. According to Howe, the crowd typically consists of outsiders, amateurs, volunteers, or part-time workers who willingly participate. The types of tasks that could be crowdsourced ranged widely. They could include everything from creating digital content to solving complex problems or funding business startups. Howe's core argument is that technological advances have made it possible to harness the intelligence and skills of a global crowd in a way to transform various industries. He also notices a growing trend among businesses who were increasingly outsourcing tasks to the public, or

the *crowd*, rather than to specific, specialized employees or contractors. This trend is made possible thanks to the connectivity provided by the Internet, which allows for mass collaboration and participation from people all around the world.

In his article and further publications, Howe compares traditional outsourcing—where tasks or processes are given to a third-party provider, usually in a different location or country where labor is cheaper—to this new model:

Simply defined, crowdsourcing represents the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call (...). The crucial prerequisite is the use of the open call format and the large network of potential laborers (Brabham, 2008).

The term *outsourcing* was used to denote the practice of engaging external entities to carry out one or multiple functions of an organization, depicting a scenario where a business delegates services to other firms that could otherwise be executed by its internal staff. Crowdsourcing was conceived as a specialized variant of outsourcing, accentuating the significance of internet platforms and interactive technologies. Both crowdsourcing and outsourcing engage outside capabilities to work on company or organizational tasks (Ikediego et al., 2018).

There remain substantial distinctions between outsourcing and crowdsourcing, however. One important difference is the interpretation of the term *contract*. In outsourcing, the client organization looks for one specific provider, outlines requirements, and then the chosen provider furnishes the client with products or services based on a contract. In contrast, in crowdsourcing, forming a contract becomes more intricate as it would mean establishing an agreement with several participants who are frequently unidentified. Tasks or problems are broadcast to a large, often anonymous, crowd which can contribute its ideas, solutions, or work. These contributions are then sifted and evaluated, with the best ones being utilized. The power of crowdsourcing lies in its ability to tap into the diverse knowledge, skills, and the creativity of participants. Nevertheless, from the business point of view, there remains a kind of contract in which the contractor has strictly defined tasks to perform, there is a specific budget, and there are clear success factors. It's less about the deliberation, more about the tangible results. As usual in business!

Afuah and Tucci (2012) claim that crowdsourcing holds the potential to outperform both internal production and traditional outsourcing in terms of efficiency and effectiveness. This is particularly true when issues can be partitioned into simpler, smaller tasks, and when the necessary expertise is accessible beyond the organization's boundaries. This approach tends to be especially successful when those involved in crowdsourcing exhibit a high degree of motivation to resolve problems, and when the resulting solution is both rooted in user experiences and subject to user evaluation.

Howe provides some interesting examples of the phenomenon he was trying to define. One of them was Threadless.com, a trailblazing online clothing company. Threadless' approach to crowdsourcing was unique and innovative in many ways, particularly in how it democratized the process of design and production in the clothing industry. The firm has built a vibrant online community wherein members

were invited to generate and submit their graphic designs for t-shirts. In an open contest, these designs are voted upon by the community members themselves. So, unlike traditional clothing companies that rely on in-house or contracted designers to create new ideas, Threadless empowered its community members to co-create the products. This not only fosters a strong sense of community engagement, but also leads to a diverse range of creative, crowd-approved designs that may not have been considered in a more traditional design process. The top-rated designs are eventually brought to life, printed on t-shirts, and sold through the company's website. Designers of the winning entries received cash prizes as recognition for their creative contribution. This model has led to considerable benefits for Threadless. It allows the company to embed market research directly into its production process, resulting in products that closely align with the tastes of its audience. Moreover, the company substantially reduced overhead costs, given that product ideas are sourced from the consumer base itself. Perhaps the most profound advantage of this model was its ability to engage and retain a dedicated audience, fostering a strong sense of community and brand loyalty (Howe, 2008).

An example of a long-lasting crowdsourcing campaign is Doritos' *Crash the Super Bowl*, a unique contest that started in 2006. It is an annual competition inviting fans and aspiring filmmakers from around the world to create their own commercials of Doritos—cheese-flavored triangle-shaped chips. The number of participants increased each year, with thousands of individuals and teams submitting entries from across the globe. In 2014, the competition reportedly received around 5,400 entries from 30 countries (Orange, 2014), marking the highest participation rate in its history. The creativity and humor expressed in the submissions was an integral part of the competition's appeal. For instance, *Doritos Time Machine* (2014 winner) featured a clever little boy who tricks a man into entering a makeshift “time machine” (a cardboard box) to get his hands on his bag of Doritos. In turn, *His dying wish was for Doritos* (2010 winner) is set at a funeral, where the deceased's final wish was to be buried in “a jumbo casket of Doritos.” The camera then cuts to the inside of the casket, in which a very-much-alive guy, covered in Doritos, bites into a chip while watching football on television (Furdyk, 2023).

One of the most compelling elements of the Doritos project was the public voting process. Participants were encouraged to promote their submissions via social media to gather votes, effectively making the competition a viral marketing campaign. This not only boosted the contest's reach and engagement but also gave the participants valuable exposure. Another intriguing aspect was the prize. The creators of the winning commercials were rewarded with cash prizes up to \$1 million, depending on how well the commercial ranked on the USA Today Super Bowl Ad Meter. Winners were also offered opportunities to work on set with renowned directors and producers in the advertising industry. Finally, the winning commercial would air during the Super Bowl, one of the most-watched television events in the United States. Given the high cost of Super Bowl ad slots, this was a prestigious and coveted opportunity.

InnoCentive.com (now called Wazoku Crowd), another of Howe's examples, is of a different kind. We are not dealing with a project for one specific corporation, but a

platform that broadcasts scientific research or development challenges from several companies to an online community, on which individuals can attempt to solve the challenges in exchange for substantial cash prizes (Brabham, 2012; Howe, 2008). The platform unites a global network of problem solvers (known as *solvers*) and organizations (known as *seekers*) that need solutions to complex research and development challenges. Seekers, often businesses or nonprofits, post challenges anonymously on the platform. These challenges can range from scientific research problems and engineering difficulties to innovation or process-improvement needs. The challenges are made visible to the global community of solvers, who include scientists, engineers, professionals, students, and other interested individuals who have registered with InnoCentive. Solvers, working alone or in teams, tackle a challenge and submit their solutions within the deadline. Seeker organizations evaluate the solutions, and if a solution meets the criteria, the solver (or team of solvers) is awarded a cash prize. For example, Ed Melcarek, a 57-year-old physicist from Canada, helped Colgate-Palmolive find the answer to the question of how to fill a toothpaste tube with fluoride powder without it escaping into the environment. Italian Giorgia Sgargetta invented a blue dye for Procter and Gamble that enabled perfect dosing of dishwashing liquid, visible through the water turning blue when it is saturated with the right amount of liquid. And Maria Vikomon, a 63-year-old retiree from Budapest, found valuable information about so-called antioxidant connections which could potentially prevent the formation of sick cells (Henhappel & Niedzielski, 2015). Throughout its years of activity, InnoCentive and its successor Wazoku gathered over 380,000 problem solvers from nearly 200 countries. It has also run over 2,500 external challenges, capturing app. 200,000 innovations and awarding over \$60 million in the process (Wazoku Crowd, 2023). This project stands as one of the most renowned instances of crowdsourcing, and an examination of its structure reveals characteristics unique to this form of collective intelligence.

So, let's take a look at the features of crowdsourcing projects that establish a specific framework for the CI emerging within them. The extensive definition, drawn by Estellés-Arolas and González-Ladrón-de-Guevara (2012) from a survey of thousands of papers on the topic, describes it as something for which:

- (a) there is a clearly defined crowd;
- (b) there exists a task with a clear goal;
- (c) the recompense received by the crowd is clear;
- (d) the crowdsourcer is clearly identified;
- (e) the compensation to be received by the crowdsourcer is clearly defined;
- (f) it is an online assigned process of participative type;
- (g) it uses an open call of variable extent;
- (h) it uses the Internet.

Since these projects must meet the well-defined expectations of a particular enterprise, which is most often also the sponsor of the project, the management of the entire process usually rests on the shoulders of one coordinating team (Brabham, 2012). It could be a dedicated branch of the enterprise or an external company specializing in organizing crowdsourcing projects. This process manager must define what results

are expected and the success criteria, manage the process of collecting proposals and their evaluation, and settle the crowdsourcing campaign.

According to Afuah, this process is organized in four major steps (Afuah, 2018). First, the managers articulate and outline the problem and disseminate it to an undefined set of people via an open call. In crowdsourcing projects, the companies are often modularizing the problem (Viscusi & Tucci, 2018), i.e., breaking down a complex problem into smaller, manageable tasks or *modules*. Each of these modules can then be solved independently, often by different individuals or teams. Modularization allows contributors with different expertise, perspectives, and approaches to focus on aspects of the problem that they are best suited to solve.

In the next stage of the process, individuals interested in addressing the problem self-select, without any specific assignment to the task by the process managers. These problem-solvers bear their own costs associated with resolving the issue, with no guarantees of successful outcomes or desired results. We have to remember that creating crowds, unlike dispersed online communities, begins as a result of a *top-down decision*: it is not a spontaneous process, although participation is fully voluntary and open to anyone. However, it is naturally limited by participants' competence. People in the crowd are often anonymous to each other because of its size, so in this sense it is also different than a typical community. The practice of crowdsourcing proves that the idea of an *open crowd*, having no limits on its growth, thanks to the process of self-selection, leads to greater engagement of talented individuals (or *crowd crystals*; Viscusi & Tucci, 2018) and not to the *representativeness* and social equity that theories of participative democracy expect (see e.g. Best & Krueger, 2005). So, we can see that in crowdsourcing diversity is important, but competence is probably even more so. By drawing from a wide pool of individuals with varying competencies, companies increase the likelihood of receiving innovative and effective solutions.

The third stage can be conducted in two ways. In the context of contest-based projects, submitted solutions are assessed to select a winner. It's worth mentioning that the crowd can be involved not just in problem-solving, but also in problem formulation or evaluation of the proposed solutions. Conversely, in a collaborative crowdsourcing model, component solutions are compiled together. The final phase involves the integration and implementation of the chosen solution (Fig. 2.5).



Fig. 2.5 The crowdsourcing process

Following the publication of Howe's article and book, the concept of crowdsourcing gained huge traction and popularity. Various businesses began adopting crowdsourcing as a strategic method to solve problems, generate new ideas, and harness the collective intelligence of the masses. Companies in many sectors, such as technology, design, and retail, have utilized crowdsourcing for diverse functions. Examples include Amazon's Mechanical Turk for micro-tasks, Netflix's algorithm improvement contest, and Lego's Ideas platform for new product designs. The serial editions of product or process innovation contests organized by technological corporations are particularly impressive. Notable examples include Dell's Ideastorm, which has received over 23,000 ideas and implemented around 500 of them (Lam, 2016), and IBM's Innovation Jam. The 2006 edition of Innovation Jam alone engaged 150,000 IBM employees, family members, business partners, clients from 67 different companies, and university researchers, yielding more than 46,000 ideas (Bjelland & Wood, 2008). The organization of this type of project became very fashionable in traditional sectors as well, including fast-moving consumer goods. Figure 2.6 shows the number of completed crowdsourcing projects through 2013. As you can see, the food industry enthusiastically embraced this model.

We can also observe a substantial increase in academic research on crowdsourcing, demonstrating its growing recognition as a legitimate field of study—including, but not limited to, business management. As of June 2023, the renowned Web of Science database contains over 13,000 scientific articles on the subject. Since 2015, the annual number of new articles has consistently surpassed 1,000, with the peak year being 2018, which saw 1,553 publications. Business applications are the primary catalyst for interest in this research, including product creativity, the direct involvement of customers in corporate development, and cost savings. Numerous online platforms

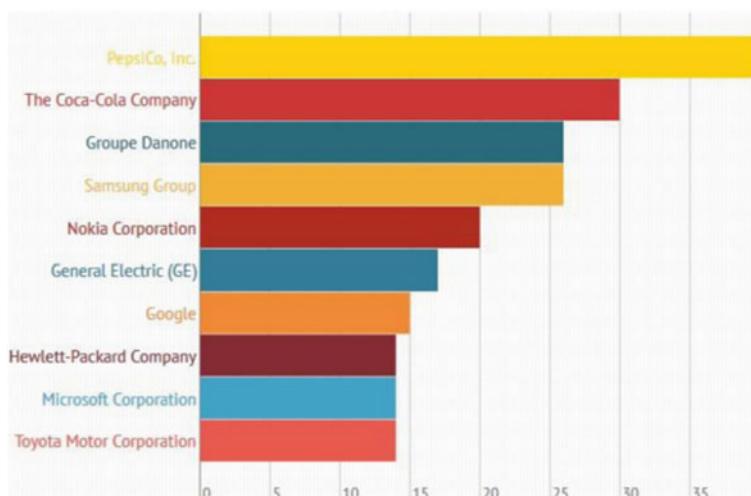


Fig. 2.6 The number of completed commercial crowdsourcing projects, from 2006 to 2013 (Ikediego et al., 2018)

serve as intermediaries connecting problem owners (those who need tasks to be completed) and solvers. Besides above mentioned InnoCentive, examples include Kaggle or Topcoder.

In summary, we can say that crowdsourcing appears in business projects with collective intelligence implemented in a specific model, is recognized as a way to access external knowledge, and enables companies to draw on “the power of the crowds” (Michelucci & Dickinson, 2016). Not every type of online CI project can be categorized as crowdsourcing, as many of them value personal interactions within smaller groups over large participant volumes. On the other hand, nearly every crowdsourcing project can be treated as a form of CI implementation on the Internet, albeit with certain operational boundaries imposed by market practices.

Owing to its business origins, the crowdsourcing model has elevated interest in CI to a new level. Numerous projects emerged swiftly, akin to mushrooms after the rain, all aimed at achieving commercially significant values, such as the introduction of new products and services, workflow improvement, an increase in profits, and the enhancement of innovation. The starting point was here the enterprise (*crowdsourcer*) and its needs. Characteristics highlighted by Surowiecki, such as diversity, independence of judgment, and decentralization, are still deemed significant success factors. However, more frequently than not, the feature of these projects is not the cooperation of the entire group, but the competitive selection of highly competent individuals: the search is on for “solvers with the incentives and valuable, immobile, difficult to imitate or substitute resources to self-select to solve problems” (Afuah, 2018). The belief in the critical role of a project leader or manager is also prevalent, a concept already highlighted by Howe (2008): “people might be enthusiastic and capable at some level of self-organization according to their interests and abilities, but they also require direction and guidance and someone to answer their questions.” However, looking from the perspective of the crowdsourcer, what really matters is the development of “very high-value solutions to some problems at a very low cost” (Afuah, 2018).

Both the crowdsourcing company (seeker), and the participants (*solvers*), derive value when the compensation they receive for their contributions exceeds the cost of providing these benefits (Lepak et al., 2007). Thus, it is a simple economic calculation in which both parties should be satisfied. The rewards garnered from solving a problem should exceed the costs of its resolution on both sides. This includes organizational processes, technical costs, as well as the time and commitment dedicated to participating in the project. Rewards for the solvers encompass both monetary and non-monetary benefits, as well as extrinsic and intrinsic rewards, which the seeker should ensure (Belenzon & Schankerman, 2015). For example, this could take the form of micropayments distributed to numerous contributors or a substantial reward given to the individual who provides the most compelling solution. Not everyone is motivated by money, however. While many solvers are incentivized by immediate monetary payments, others are motivated by factors such as skill development, expertise improvement, showcasing their skills to potential employers, reputation building, enjoyment, and the desire to make a difference in the world (Afuah, 2018). As solvers may have non-monetary incentives for resolving certain problems, and

those whose solutions do not win most often bear their own costs, the cost of solutions to seekers can be extremely low.

2.6 Crowdsourcing Beyond Business: Science and Policymaking

Although the term crowdsourcing originated in the business sector with terminology rooted in business concepts, its adaptation to non-commercial areas, such as scientific research and public administration, began swiftly. It is important to note that in these non-commercial applications, the fundamental values and motivations differ from those in business. Instead of focusing on the profit of a single organization, the emphasis is on social benefits, and participants are often driven by motivations other than financial gain. In this section, we will review the most interesting projects in which a wide audience is involved in scientific research or public policy development.

2.6.1 *Citizen Science Projects*

First, let's examine citizen science, in which public volunteers contribute to the collection and classification of data, thereby enhancing the capabilities of the scientific community. As Oxford English Dictionary defines it, citizen science is “scientific work undertaken by members of the general public, often in collaboration with or under the direction of professional scientists and scientific institutions” (*‘Citizen Science’ Added to Oxford English Dictionary, 2014*). The notion of involving a broad internet audience in volunteer science work dates back to the close of the twentieth century, as highlighted by Strasser et al. (2019). However, these early initiatives did not call for active intellectual participation from non-professionals. SETI@home, a project initiated in 1999, is a prime example. This project leveraged the Internet for distributed computing, enlisting the participation of individuals to lend their machines' computing power. The objective was to perform computation tasks aimed at finding potential evidence of radio transmissions from extraterrestrial intelligence. The observational data used in this endeavor came from the NRAO Green Bank Telescope and the Arecibo radio telescope (Korpela et al., 2011). SETI@home capitalized on the unused resources of volunteer internet-connected computers worldwide. The system disseminated millions of data segments to be processed by personal computers at various locations. These computers then relayed the processed results back to the SETI@home system. Participation in this project was generally passive, however. Beyond the initial configuration, the computational tasks demanded very little engagement.

Drawing inspiration from the model described by Howe, new types of projects were created in which substantive work is actually delegated to the crowds. The first

endeavor of this new type was the 2006 Galaxy Zoo, a crowdsourced astronomy project that invited people to assist in the morphological classification of large numbers of galaxies. The Zooniverse.org web portal grew around this original concept, becoming the hub for the Internet's most popular citizen science initiatives. Some recent projects launched on this platform include: *Cloudspotting on Mars* (2022), where volunteers search and make annotations for mesospheric clouds on Mars with data from the Mars Climate Sounder on the Mars Reconnaissance Orbiter, *Snapshot Serengeti* (2012), where people classify animals at the Serengeti National Park in Tanzania using images gathered from 225 camera traps to study how species interact with each other, or *Old Weather* (Terras, 2015), where participants use a special interface to digitally transcribe weather and sea ice information from the log books of US Arctic exploration into digital form, data which was recorded while these ships were navigating between 1850 and 1950.

Citizen science projects offer significant value across several dimensions. One of the most direct forms is the large volumes of data that volunteers can gather and, in some instances, analyze, thereby significantly advancing scientific understanding in various fields, from astronomy to zoology. Beyond data collection, these projects hold immense educational value. By engaging the public in research activities, citizen science projects foster improved scientific literacy and promote greater public interest in science. Furthermore, the insights derived from these projects often inform policy decisions and conservation strategies, particularly in the realm of environmental monitoring and biodiversity. As for the motivations driving participants, many are drawn to the educational benefits these projects afford, offering opportunities to deepen their understanding of specific scientific topics. Personal interest in a scientific field or the subject matter of the project is another common motivator. For some, the prospect of contributing to scientific knowledge, even outside the realm of professional science, is a significant draw. Lastly, the strong community elements in many of these projects serve as a social incentive, providing an opportunity for participants to engage with like-minded individuals and contribute to a shared goal.

2.6.2 *Crowdsourcing in Public Policies*

2.6.2.1 **Promising Beginnings and Worldwide Development**

Attempts to use the crowdsourcing model in various sectors of public administration appeared very quickly after its popularization. The *Peer-to-Patent* initiative, managed by Beth Simone Noveck, originated in late 2006 as a response to the growing backlog and quality issues facing the United States Patent and Trademark Office (USPTO) (Noveck, 2006). The project aimed to improve the patent examination process by opening it up to public participation and addressing some of the systemic problems faced by the USPTO. This not only reduced the backlog of unexamined patents but also increased the quality of the patents being issued. Noveck, a New York University law professor, started the Peer-to-Patent project as a pilot program in collaboration

with the USPTO. The goal of this initiative was to revamp the patent examination process. The participation of crowds may be able to be harnessed to find *prior art*—publications and evidence that might challenge the claims of a patent application for being non-novel or obvious. The traditional patent review process faced challenges in this area because patent examiners often did not have sufficient time or resources to find all relevant prior art. This led to the issuance of patents that possibly should not have been granted, known as *low-quality patents*. Under the Peer-to-Patent system, patent applications were posted online, and the public, particularly those with expertise in the relevant fields, was invited to submit prior art and commentary relevant to the patent’s claims. This information was then made available to the patent examiners, providing them with more data and perspectives to make informed decisions. The work carried out under Peer-to-Patent originally concerned applications in the area of computer architecture, software, and information security (Loiselle et al., 2009). This pioneering online system facilitated public involvement in the patent examination process, effectively demonstrating how crowdsourcing can augment the efficacy of government procedures.

In 2009, as the Peer-to-Patent project was gaining notoriety thanks to its innovative approach, Noveck became the United States deputy chief technology officer for open government and led President Obama’s Open Government Initiative (OGI). She was based at the White House Office of Science and Technology Policy and served as an expert on governance, technology, and institutional innovation. Under the auspices of OGI, several projects were launched that used various forms of crowdsourcing. One of them is a platform called Challenge.gov, managed by the General Services Administration (GSA) to support a new policy instrument called “Prizes and Contests.” The guiding principles for this platform include mechanisms that support the potential to “bring new ideas to the table from unlikely sources” in order to support major breakthroughs on enduring social and technological challenges, and to “help address social [needs] in addition to science and technology challenges” (Mergel & Desouza, 2013). Launched in 2010, Challenge.gov has become the dominant platform for federal agencies to host tournament-based crowdsourcing. Since its genesis, the Challenge.gov platform has seen visitors from every country in the world and has issued more than 400 awards by 69 Federal agencies, totaling more than \$72 million in prizes (Hameduddin et al., 2020).

Challenges on this platform are designed to enable citizens to not only provide solutions but also review and evaluate them, cast votes, and even participate in the implementation of solutions and the subsequent assessment of new policies or other public sector innovations. Solution providers can range from individual citizens and teams of citizens to private or nonprofit organizations, and even industry consortia. Challenge.gov also serves as a tool for agencies to amplify the recognition of their purpose and societal significance (Mergel & Desouza, 2013). To this day the platform remains an active field for federal agencies to operate in. In July 2023, challenge topics included: community-led solutions to advance mental health of children and youth (announced by Department of Health and Human Services), universal

design for more accessible federal buildings (Public Buildings Service), innovations in measuring community perceptions (National Institute of Justice), incentivizing photovoltaic system owners to share information-rich datasets from their assets (Department of Energy), or developing a prototype sensor to detect water toxicity (Environmental Protection Agency).

Within a few years, open innovation established itself as the norm across numerous American government agencies. In 2015, it became mandatory for all U.S. federal agencies to designate at least one officer with the responsibility of coordinating crowdsourcing and citizen science initiatives (Liu, 2021). Projects created as a result of the Open Government Initiative were an inspiration for many similar platforms implemented at the state or municipal level, as well as in other countries (e.g. as a part of the UK Open Government Network). These examples inspired a new way for governments to integrate user-generated geographical data into their policy frameworks (Parker et al., 2014), execute translation tasks during emergencies (Sutherlin, 2013), or volunteer to organize documents. For instance, the US National Archives Records Administration initiated the Citizen Archivist program, where citizens were trained to categorize documents and records that are safeguarded by the government on a dedicated online platform. The primary responsibilities of the volunteers included tagging archival records, indexing, and so on. According to Bowser and Shanley (2013), over 170,000 participants contributed to indexing 132 million names within a span of 5 months. Another great example is Fix My Street, a popular project initiated in the United Kingdom that was designed to simplify the process of reporting non-emergency issues like potholes and broken streetlights to local authorities. Launched by the non-profit organization mySociety in 2007, it operates as a website where citizens can mark issues on a map and these reports are automatically sent to the responsible council. Additional crowdsourcing functions, such as crowdsensing, situated crowdsourcing, spatial crowdsourcing, and wearable crowdsourcing, were designed to further enhance public policies and services (Taeihagh, 2017). Studies documenting the outcomes of these federal activities (e.g. Mergel & Desouza, 2013; Noveck, 2009), as well as local initiatives (e.g. Brabham, 2008), have highlighted the outcomes achieved by implementing crowdsourcing within governmental entities.

Citizens' involvement in the practical implementation of public entities' objectives led to many successful projects, but the ambitions of government crowdsourcing enthusiasts went much further. The success of projects in which citizens are involved in the implementation of *top-down* public tasks seemed to many researchers an important, but insufficient, step in the field of reconstruction of the public sphere (e.g. Aitamurto & Landemore, 2015; Noveck, 2009). Crowdsourcing in government was expected to transcend traditional purposes of policy legitimization and reconceptualize participation based on proportional samples. Experimental projects intended to harness the *wisdom of crowds* using collective public opinion as a potential reservoir of unique and valuable policy insights began. The role of the public was expected to be elevated from mere validators to active contributors, underscoring the potential of crowdsourcing as an innovative approach to democratic governance and decision-making.

As shown in the previous chapter, phases such as policy formulation and policy implementation are particularly important in the policy cycle process. One might ask whether it would be possible to identify crowdsourcing projects dedicated to these phases. This task was undertaken by Linders (2012) who conducted an analysis of the White House's Open Government Innovation Gallery, and distinguished between *design* (related to the policy design phase) and *delivery* (related to the policy implementation phase) in crowdsourcing projects. Using the ChicagoShovels.org platform as an example, which involves residents in city snow removal, Linders shows that as part of implementing this public service, "citizens can 'claim' streets for cleaning and volunteer for a 'Snow Corps' to help the disadvantaged." However, they can also participate in the formulation of the snow removal policy through "consultation and ideation in which citizens can provide suggestions and ideas via the snow portal's Facebook and Twitter accounts" (Linders, 2012). But what happens when you try to use crowdsourcing to generate policy ideas on a scale larger than the city and in relation to problems of a larger dimension than snow removal?

Crowdsourcing that is centered on policy formulation, rather than solely on collaboration in the execution of public services, poses greater challenges. This complexity arises both from the reluctance of traditional political institutions to relinquish control over public affairs and the concern that by lowering entry barriers and allowing, often anonymous, online communication, the system may be flooded with irresponsible, radical, manipulated, or nonsense proposals. A common problem is the lack of a clear link between participation in such projects and the emergence of a specific solution, i.e., something that can be compared to *winning* in commercial projects, the lack of what we can define as a *prize* and the resulting satisfaction.

The "CrowdLaw for Congress" project, which is managed by The Governance Lab at New York University, sought to systematically organize efforts made in different countries to utilize crowdsourcing in the creation of high-level policy. This initiative aims to examine how lawmaking bodies and other public institutions around the world are employing crowdsourcing technologies to cultivate an effective dialogue with the citizens, thereby enhancing the quality and legitimacy of lawmaking. The team at GovLab has done an outstanding work reviewing initiatives from all over the world. These projects include both top-down strategies executed by a diverse array of public institutions such as the Province of British Columbia, Ministry of Electronics and Information Technology in India, and the European Commission, as well as bottom-up efforts started by civic movements. Additionally, the reviews feature local projects managed by municipalities worldwide, as well as *private* ventures established by individuals, which may include politicians or technology startups. As of 2023, the CrowdLaw Catalog encompasses 105 initiatives. These are categorized based on the types of participatory tasks they involve, such as *ideation*, *expertise*, *opinions*, and *actions*. Additionally, the classification takes into consideration the stages of the lawmaking process, which may include problem identification, generation of solutions, drafting, decision-making, implementation of decisions, or assessment. However, if we take a closer look at these projects, the problems and controversies associated with them become apparent.

2.6.2.2 Challenges and Limitations of Crowdsourcing in Policymaking

First of all, let us note that initiatives proposed by government institutions often have limited causative power. In some cases, their potential impact was minimized during the course of their development, e.g. by allowing only some proposals, selected according to an unclear key, to be processed. As an example, consider “We the People,” a system launched by the White House in 2011 which served as a platform for petitioning the administration’s policy experts and was accessible to all interested citizens. The justification for the project was that “the right to petition the Government for a redress of grievances” is guaranteed by the United States Constitution’s First Amendment (U.S. Const. amend. I). Users who wished to create a petition were required to register a free whitehouse.gov account. To sign a petition, users only needed to supply their name and email address. Petitions attaining a specified threshold of signatures underwent review by administration officials, who were tasked with drafting an official response. The original threshold for a response was set at 5,000 signatures but was soon raised to 25,000 and then raised again to 100,000 in 2013. Regardless of statements otherwise, many petitions remained unanswered despite exceeding the required threshold. This led to doubts about the reasonableness of this project. In August 2013, the Washington Post website published an article about 30 petitions that had been left unanswered for an average of 240 days despite having met their signature goals (Peterson, 2013). More importantly, even if they did get answers, they were often enigmatic and vague. The administration was in no way obliged to give the submitted cases a chance.

We will understand why if we look at some popular petitions. One known case is a November 2011 petition urging the government to build a Death Star like the space station featured in the Star Wars franchise, which was presented as a potential economic stimulus and job creation measure. Another well-known initiative called for “the pardon of Edward Snowden,” and gained a total of 167,954 signatures. It stated that “Edward Snowden is a national hero and should be immediately issued a full, free, and absolute pardon for any crimes he has committed or may have committed related to blowing the whistle on secret NSA surveillance programs.” Both petitions, due to the large number of signatures, were answered, but it is unlikely any of these answers satisfied the signatories (Rhodan, 2015). What was the follow-up? The “We the People” project was curtailed under Donald Trump, and on Joe Biden’s inauguration day it disappeared completely from the Internet.

Another example representative of many government-initiated projects was *Rahvakogu*, or *People’s Assembly*, introduced in 2013 by the Estonian government and aimed at involving citizens in the policymaking process by enabling them to propose changes to the electoral laws, political party laws, and other issues. The *Rahvakogu* platform allowed citizens to submit proposals and ideas online, as well as discuss and refine them. In a multi-stage selection, the submitted proposals were grouped by analysts into 59 bundles of different possible scenarios and supplemented with impact analyses. Then, a stratified random sample from different subsets of the population was selected to participate in the deliberation day. On this day, seminars were held on the main topics, their timeliness and effectiveness were discussed, and

the selected topics were presented for the Estonian parliamentary to discuss. By summer 2014, three of the proposals out of the 15 that were sent to the Parliament had become laws. Critics, however, claimed that the process, originally presented as a simple and direct channel of communication between citizens and the government, was distorted by far-reaching filtering of submitted ideas and cherry-picking by selecting only those matching the previously adopted program of the ruling party. For example, Alar Klip, lecturer in comparative politics at the University of Tartu, pointed out that, in practice, too much responsibility for the selection of proposals rested on the moderators of the debate (Klip, 2013). The well-known Estonian journalist Priit Hõbemägi saw in these actions foremost a desire to improve the government's PR, not a real crowdsourcing of new ideas to improve public policy (Hõbemägi, 2013).

On the other hand, projects that were created from the bottom up, on the basis of citizen initiative, often took the form of a *protest movement* against certain events in public life. Very often, despite attracting a large number of people, they did not result in practical implementation, nor even serious public debate. One of many such examples is the Argentine project DemocracyOS. In Buenos Aires in April 2012, a group of activists, students, and hackers started conceptual work to design tools to adapt the outdated political system to the *internet era* by enabling direct participation in setting the political agenda and decision-making. The motto that guided this group was "upgrade our democracies to the internet era." As one of the leaders of the movement, Pia Mancini, put it: "While the Internet allows us unprecedented access to information, low costs for collaborating and participating, and the ability to express our desires, demands, and concerns, our input in policymaking is limited to voting once every two to five years" (Mancini, 2015). Members of the movement founded the DemocracyOS Foundation and created an online platform enabling large-scale participation by proposing, deliberating, and voting on policy issues. They imagined that it would serve as a bridge between members of parliament or political decision-makers and citizens. Policy programs, new regulations and laws, and other proposals could be proposed and mutually evaluated by a large online community and then selected in an online voting process. Politicians and policymakers should have been able to gain a channel enabling constant and direct contact with citizens, collecting ideas and gaining feedback; at the same time, they were supposed to be obliged to respect the decisions of the DemocracyOS community. And here the first serious obstacle appeared: none of the existing Argentine political parties or government institutions were interested in using this tool because they saw it as a potential threat to their independence. Arguments about the substantive potential of solutions based on crowdsourcing were not taken seriously!

The DemocracyOS community decided to start the project despite these difficulties. They themselves formed a political party: *Partido de la Red* (Eng. The Net Party). Its task was to win seats in parliament with the stipulation that the candidates were obliged to vote in line with the decisions of the citizens that engaged with DemocracyOS. This action was called "hacking the system." Unfortunately, the party only won 1.3% of the support in 2013, and in subsequent years the support was even smaller. The DemocracyOS platform gained some popularity and was occasionally used for social consultations by The Congress of the City of Buenos Aires, and it

also saw implementations in Tunisia and Mexico (Mancini, 2015). However, the activists' ambitious plan to completely rebuild their political system using this tool never came true.

Another well-known example of a grassroots initiative was the 2010–2013 project focused on Icelandic constitutional reform. The reform process was instigated by the 2008 financial crisis, which led to a profound economic and political crisis in Iceland. In the fall of 2008, all four of the Icelandic national banks collapsed and Iceland entered into a deep economic crisis. Public trust in the existing political institutions was severely undermined. To restore it, the idea of a new constitution, directly involving the citizenry, was proposed. This new process of constitutional reform began with the National Assembly of 2009, also known as *Pjóðfundur*. The National Assembly was a gathering of 1,500 Icelandic citizens selected randomly from the national registry, along with 300 participants chosen from various institutions and organizations. This one-day initiative was focused on defining the principles and preferences that should steer the revitalization of government and public administration and establishing “the principal viewpoints and points of emphasis of the public concerning the organization of the country’s government and its constitution” (Kok, 2010).

Following the National Assembly, a decision was made to draft a new constitution. The Constitutional Assembly was elected in 2010, which consisted of 25 citizens who were not professional politicians. When the elections to the Constitutional Assembly were nullified due to some irregularities in the voting process, the same elected members were appointed to a Constitutional Council by a parliamentary act. The most original and directly participatory part of the drafting stage was the *crowdsourcing* phase, which included a series of twelve crowdsourcing moments. What made the Icelandic constitutional reform process unique, was the unprecedented use of the *wisdom of crowds* techniques and digital technology. Helene Landemore summarized this process, claiming that:

the twenty-five members of the Council, far from isolating themselves from popular input, regularly posted online the version of the draft they were working on. All in all, they posted twelve drafts, at various stages of completion. The Constitutional Council used online platforms and social media to share the draft constitution and get public feedback. This represented a radical new model of participatory democracy and constitutional development, as it enabled the general public to contribute directly to the drafting of their constitution. (...) The participants' online (and offline) contributions were actually read by the constitution-drafters and made a measurable difference to the final text (Landemore, 2020).

All draft clauses were published on the Constitutional Council's website, as well as on its Facebook page, where citizens could comment on the proposed text. The Council also had a Twitter account and a YouTube channel. Anyone interested in the process could send feedback by posting comments on social media platforms, on the Council's own webpage, or in regular email. The Icelandic procedure made every effort to maintain transparency at each vital phase, particularly during the critical stage of drafting the constitutional text. The objective was to encourage public involvement not only in the ratification of the constitution but also in its development.

This procedure was intended to be direct, rather than represented through intermediaries. In comparison, the constitutional work in elder, pre-Internet era cases were typically seen as processes best left to experts, at least when it comes to agenda-setting and actual writing of the constitutional proposal. Ordinary citizens were usually involved only at the very end of the process (Landemore, 2020).

In spite of the innovative approach, the outcomes of the constitutional reform did not meet expectations. The event organizers anticipated substantial engagement, yet only about 3600 comments were posted online. Each comment initiated discussions of varying lengths, with multiple response chains springing from the initial remarks. In the end, only approximately 360 suggestions were put forward, which, considering Iceland's population of around 320,000, is not a high figure. Furthermore, some participation theorists were hoping for the involvement of a varied demographic, but it appeared that the crowdsourcing process primarily attracted the attention of middle-aged and older white males (Landemore, 2020). The Council's draft constitution was completed in July 2011 and was later approved in a non-binding referendum in October 2012 by two-thirds of voters (Ginsburg & Elkins, 2014). However, the draft constitution was not formally adopted. Although the draft was submitted to the Parliament, the Parliament did not complete the ratification process before its term ended, and subsequent legislatures did not pick up the process.

Various reasons have been proposed by observers as to why the draft was not ultimately codified into law. Some attribute it to an inconsistent Icelandic public who lost urgency once the harshest phase of the economic crisis had passed. Notably, the 2013 election yielded a clear victory for parties that had staunchly resisted the constitutional process. Opposition emerged almost immediately from parties, politicians, and even scholars who felt threatened by the prospect of change or were frustrated by their lack of influence over it. Additionally, significant economic entities that were opposed to certain elements of the new draft also contributed to the resistance (Landemore, 2020).

2.6.2.3 Successes of Crowdsourcing in Policymaking

Nonetheless, among the projects that engage internet users in co-creating public policies there are a number of successful initiatives, many of them described in the CrowdLaw Catalog. Projects focused on local issues of the urban community, combined with submitting and voting projects financed from the participatory budget, noticeably thrive. Participatory budgeting (PB) is as a form of civil participation in which citizens are involved in the process of allocating public budgets that has gained great popularity in recent years. Originating in Porto Alegre, Brazil, in 1988, PB has evolved in numerous ways, featuring diverse methodologies, formats, and technological applications. The Participatory Budgeting Project, an independent NGO dedicated to supporting the implementation of PB, indicates that over the past 35 years PB has been adopted by more than 7,000 cities worldwide. It has been used for budget decisions in various contexts, including states, counties, cities, housing authorities, schools, and other institutions.

The implementation of such projects, most often carried out on a cyclical, annual basis, is an opportunity for many municipalities to launch online platforms supplementing the voting process with crowdsourcing about local problems, civic debate, and co-creation of ideas to improve the lives of city residents. The success of this type of project is influenced by the combination of crowdsourcing and online deliberation with the causative factor that—as we wrote earlier—is so important in the case of commercial projects. The possibility of influencing the choice of the best solution, and then observing its implementation, significantly affects the involvement of citizens. In addition, PB projects are characterized by a high degree of *depoliticization*, i.e., detachment from ideologically marked party programs, and by their status as a practical solution that improves the life of the local community.

In the next chapter we will analyze in greater detail the most interesting cases belonging to this group, but first we will present an example to better understand how PB projects work. One of the most famous projects of this type is the citizen participation platform *Decidim* (Catalan for “we decide”). It is an online platform for participatory democracy launched by the City Council of Barcelona. This digital tool enables citizens to propose, debate, and prioritize policy proposals and projects through an interface that facilitates threaded discussions and aligns comments with proposals. Initially developed to support participatory processes beyond just participatory budgeting, *Decidim* Barcelona played a pivotal role in drafting the city’s strategic plan for 2016–2019. The platform’s launch catalyzed the rewriting of the municipal strategic plan, attracting significant engagement: within two months of its introduction, over 25,000 users had registered, contributing 10,860 proposals and participating in 410 meetings (Peña-López, 2019).

Decidim originated from a broader movement seeking to enhance political representation and transparency, particularly following Ada Colau’s election as Barcelona’s mayor in 2015. Subsequently, an initiative emerged within the city’s municipal government aimed at bolstering democratic legitimacy through the strategic use of technology. Colau, an activist and member of the platform *Barcelona en Comú* (‘Barcelona in Common’ in Catalan), has been a staunch advocate for social justice and anti-corruption, especially in response to the 2008 Eurocrisis and its aftermath (*Decidim: Participatory Budgeting in Barcelona*, 2023). The 2016 launch of *Decidim* followed the introduction of Madrid’s online participatory platform, *Decide Madrid*, a recipient of the United Nations Public Service Award. In their initial development both platforms utilized the same open-source software, CONSUL (Royo et al., 2020). *Decidim* was developed through the collaboration of both private and public funders. Key contributors included the Open University of Catalonia, alongside Barcelona-based developers Codegram and aLabs (Previle, 2019).

The platform features four main participatory spaces: *processes*, *assemblies*, *consultations*, and *initiatives*. *Processes* enable the creation and management of various participation tools, such as elections or participatory budgeting. Citizens can decide where up to 75 million euros of the municipal budget are spent, that is, 5% of the overall budget. Residents can submit proposals for how the budget should be spent. These proposals can range from neighborhood improvements to city-wide initiatives. *Assemblies* allow users to form groups or collectives, organizing

and listing their meetings and locations. *Consultations* facilitate the organization of referendums, fostering discussion, debate, and voting on specific topics. *Initiatives* empower users to launch and support citizen proposals collaboratively, including gathering support and collating discussion results (Ennadif, 2020). Additionally, the platform's components section offers a range of actions like proposal submission, voting, supporting, amending, and survey participation. Users, who can register as individuals, collectives, or user groups, are able to navigate, create, endorse, and share content, engaging actively in the platform's participation processes (Decidim: Participatory Budgeting in Barcelona, 2023).

Decidim Barcelona is just one of numerous examples of projects tied to local politics that employ participatory budget mechanisms. It offers citizens the opportunity to propose ideas on city services and operations and vote (up or down) on suggestions that are either appealing or unappealing. Other similar initiatives include Better Reykjavik in Iceland, Decide Madrid in the Spanish capital, Bogotá Abierta in Colombia, Lisboa Participa in Portugal, and Future Melbourne in Australia.

Another notable category of successful crowdsourcing projects in the field of policymaking are singular, specialized issue initiatives that engage an audience with in-depth knowledge on the subject. This audience might encompass groups of stakeholders with conflicting interests. The crowdsourced law reform in Finland, mentioned at the beginning of this subchapter, can serve as an example here. In this initiative, launched in 2013 by the Finnish Ministry of the Environment and the Committee for the Future of the Finnish Parliament, citizens were invited to contribute knowledge to law reform online (Aitamurto, 2014). Laws pertaining to offroad traffic regulate transportation that extends beyond established roads, such as motor-powered movement in natural environments, predominantly involving snowmobiles in winter and all-terrain vehicles (ATVs) in the summer. Participants had the ability to propose ideas on the platform, vote ideas up or down, and comment on others' suggestions. Users accumulated points through their participation, submission of ideas, voting, and commenting. The information on the platform, along with the input generated by the crowd, was highly accessible and visible online for anyone interested. The crowdsourced parts of the process varied depending on the stage of the project, and included three distinct phases:

- (1) Ministry of Environment officials, in conjunction with researchers, delineated ten principal sectors in which issues could be detected via crowdsourcing. This stage produced around 340 ideas and conversation starters, elicited 2,600 comments, and attracted 19,000 votes from roughly 700 users. The researchers synthesized the participants' contributions by summarizing the ideas and issues contained within the submissions. In collaboration with civil servants, they then categorized the ideas and comments into broader *challenge areas*, which were subsequently fine-tuned and utilized to shape the second phase. The interpretation of these ideas was conducted by the researchers' teams and then cross-verified within the group.

- (2) In the second phase, participants were encouraged to suggest solutions to the issues identified during the initial phase. For this purpose, the broad challenge areas were segmented into more specific topics. This stage resulted in approximately 170 ideas, 1,300 comments, and 6,000 votes.
- (3) In the third phase, the participants' ideas were evaluated by both the crowd and experts. Crowd evaluation took place on a dedicated online platform where participants could assess the ideas conceived in the earlier stages.

It should be noted that participation in this process did not mean balance between users. During the idea crowdsourcing stages, only 700 out of 7,000 visitors registered on the site, which was necessary to contribute. Most of the ideas proposed in the Finnish crowdsourcing process were generated by a minority of registered users: 46.44% of the ideas were produced by the ten most active participants (Aitamurto & Landemore, 2015).

2.6.2.4 Organization, Representativeness and the Scale of Participation

As we can see in the presented overview, experiments using crowdsourcing for policy creation have become an interesting alternative in recent years, aligning with the idea of *open government*. At the same time, however, these projects are more problematic than initiatives where citizens get involved in the implementation of already-planned public tasks. The kind of controversy largely depends upon whether crowdsourcing is organized *top-down* (like *Rahvakogu* or *We the People*), or *bottom-up* (like DemocracyOS or the Icelandic constitutional initiative). If the project is initiated by the broadly understood public administration, there is a risk that the government is making mock processes just for PR gains, in practice *validating* what they already planned to implement. On the other hand, if the project arises *from the bottom*, it is possible that it will have the character of a *protest movement* and for that reason may not be taken seriously, or even be thwarted by the establishment as it is perceived to be a threat.

The issue of representativeness presents a distinct challenge. Given that representativeness is identified as one of the pillars of social participation in both participatory democracy theories (see, for example, Pateman, 1970; Roussopoulos & Benello, 2005) and deliberative theories (e.g., Habermas, 1991; Gutmann & Thompson, 1996; Gastil et al., 2005), many observers have questioned the extent to which self-selected, online crowdsourcing initiatives represent the desires and priorities of the broader population. Moreover, they ask: do these types of projects meet the legitimacy criteria if they do not involve a statistically representative sample that forms a mini-public in relation to the whole society (Dryzek & Goodin, 2006; Grönlund et al., 2014)? For instance, there were concerns that the participants in the Icelandic constitutional process were predominantly middle-aged and older white males. Similarly, in the crowdsourcing of off-road traffic law reform in Finland, 86% of participants were educated, full-time working professional males with a strong interest in off-road traffic (Aitamurto & Chen, 2017).

In response to this, Aitamurto and Landemore contend that despite the absence of statistical representativeness among the participants, the policy crowdsourcing experiments captured a range of diverse viewpoints and opinions. This mitigates, to an extent, concerns about bias that could stem from the self-selected nature of citizen participation. “Lack of statistical representativeness thus does not necessarily mean a poverty of views, information, and arguments and low-quality deliberation” (Aitamurto & Landemore, 2016). An approach enforcing representativeness, or, for example, the random selection of participants as proposed by Fishkin (2009), could in fact lower the quality of the process. Artificial inclusivity or enforced statistical representativeness would be a denial of the concept of crowdsourcing. Presenting an analysis of participants in a crowdsourced law reform, Aitamurto, Landemore, and Galli conclude that the presence of a disproportionately represented participant sample should not necessarily be viewed as a problem: “an interesting finding in our experiment is that the more active participants were at least sometimes seen as representing the views of the more passive ones” (Aitamurto et al., 2017).

Proponents of using crowdsourcing in policymaking must consider the scale of participation. Recent experiences have made it clear that local projects, those engaging a public that is restricted to residents of a specific city, stand a better chance of succeeding. Such initiatives often utilize participatory budget mechanisms to involve residents in resolving local issues, deliberating on new investments, or overseeing social and cultural city programs. This approach tends to yield measurable results in the form of implemented projects and reinforces a sense of agency among participants. Given the reduced level of politicization around such matters, they provoke far less controversy than national programs. This holds true whether the initiatives are organized by local governments, as with Decide Madrid, or by non-governmental organizations in partnership with public institutions, as seen with Fix My Street or Better Reykjavik.

The experience gained from off-road traffic law reforms in Finland shows that the merits of the debate and the effectiveness in generating outcomes can go hand in hand with the participation of a smaller, but well-prepared, group of people particularly interested in a given issue. This, and other similar experiences, prompted Beth Noveck to propose an alternative to business-specific *crowdsourcing widely*, in which the broadest possible crowd is invited to participate in a given project, instead promoting *crowdsourcing wisely*, emphasizing targeted involvement of experts or those with relevant knowledge and experience. Noveck points out that in mass participation services such as We the People, even the most popular petitions had no real impact on policymaking. The platform hosted over 200,000 petitions that have garnered over 13 million signatures and comments. Nevertheless, only about 165 received a response from the White House, and practically none of them can be linked to a decision taken or a program implemented by the government (Noveck, 2015, p. 76). According to the author of *Smart Citizens, Smarter State*, even if the government “were willing to build Death Star, the website (...) would offer no implementable guidance around how to do so, No one is invited to share relevant knowledge, experience or expertise” (Noveck, 2015, p. 76).

Delving deeper, it appears that in many cases mass participation does not lead to efficient problem solving or decision-making, as individuals often lack the necessary time, education, and motivation to engage in a meaningful and productive manner. We are faced with an overabundance of information and challenged to extract valuable knowledge from such a vast amount of data. Thus, according to Noveck, we need *human expertise* from individuals and groups to curate and align data with decision-making processes. Reference points could be engaging expert networks, locating and matching expertise within and across organizations, or allowing project managers to target the most appropriate people for the planned tasks. Examples of such networks include CyberCompEx, where participants are sorted based on criteria such as reputation, credentials, skills, and experience (Noveck, 2015, p. 106). These expertise technologies can possibly unlock the potential for *crowdsourcing wisely*, assisting institutions in locating and developing the right expertise to make sense of the ever-growing flood of data and make more effective decisions. This type of crowdsourcing would involve the targeted creation of smaller groups consisting of employees of public institutions or volunteers, the members of which would be individuals with confirmed knowledge and/or experience for the specific topic and motivation to solve problems and exchange opinions with other specialists. This model assumes a certain exclusivity. Entry barriers will simultaneously motivate participants, facilitate the analysis of complex problems, and satisfy their sense of worth. In the field of policymaking, as the author of *Smart Citizens, Smarter State* contends, the untapped potential lies in public institution employees at various levels, whose unique knowledge and skills could interact with others to create unique, digitized knowledge networks, activated as needed when new challenges arise. An attempt to create such a network from the Obama administration era includes Experts.gov, established by the Food and Drug Administration. However, this expert network differs from their predecessors (for example, those established in the 1970s under the Federal Advisory Committee Act). They wouldn't be assembled based on schematic selection or imposed representativeness. Instead, these networks would welcome the expertise of self-organizing groups arising from the voluntary involvement of highly motivated individuals.

Crowdsourcing initiatives that focus on collaborative policy development bear an importance transcending the mere pursuit of validation of governmental actions. These projects rest on the notion that everyday citizens can contribute valuable information and knowledge, rather than simply serving as a rubber stamp for policy. Public managers' expectations have evolved beyond merely involving the public due to their right to participate (though this continues to be a justified reason) as their participation can enhance the intelligence and caliber of project outcomes, leading to improved results. On the other hand, even with all the excitement and acknowledgement of the possible advantages of crowdsourced policymaking, it can be argued that the concept of open government has made a surprisingly minimal impact on how public decisions are actually made, how governments solve problems, or how public goods are allocated (Noveck, 2015, p. 17). While technology has disrupted every other industry, its influence on policymaking remains limited and is only visible in selected areas.

The escalating tensions worldwide in the third decade of the twenty first century—wars, pandemics, and social unrest—seem to be pushing the prospect of openness even further away. Therefore, it is important to study these powerful collective-action processes—that exist within interconnected structures based on information exchange—in a scientific manner. Recent years have witnessed a dynamic progression in applied research on the factors influencing collective intelligence, largely driven by the advent of new tools and methodologies, alongside the wealth of data sourced from real-world projects.

2.7 The Current State of Collective Intelligence Research

In parallel with the burgeoning interest from both private and public stakeholders in crowdsourcing projects, there has been a vibrant and multidisciplinary progression in collective intelligence research over the last 15 years. There are a few drivers of this: the proliferation of dedicated crowdsourcing projects, a huge increase in the popularity of social media, and researchers' desire to gain a better understanding of these phenomena. However, the scholars dealing with the crowdsourcing aspect of CI mainly focus on project organization methodologies and practical applications of projects. Yet, the research on CI goes beyond the merely practical aspect of executing large-community projects. Researchers seek to understand the extent to which we can discuss collective thinking: is it merely a metaphor, or do collectives possess their own distinct modes of reasoning that differ from individual cognitive processes? If collective thinking is not a metaphor but a tangible process, then the question arises: can we parameterize it and measure a collective intelligence quotient? This research on CI is not limited to crowdsourcing—it encompasses large groups formed ad hoc, small task-oriented teams, and long-standing communities with a strong sense of identity. The development of Artificial Intelligence (AI), particularly neural networks, along with new research approaches such as cognitive science, serve as inspirations for research on CI. This has led to a fascinating synergy between human intelligence and AI, both of which are undergoing a rapid and profound evolution. Whether in designing new systems for collaboration or refining the methodologies for crowdsourcing projects, the lessons learned from CI research have far-reaching implications.

The inception of the MIT Center for Collective Intelligence (CCI) marked a novel approach in the field of CI research. As the premier research unit specifically dedicated to examining the CI phenomena, it is affiliated with the renowned MIT Sloan School of Management and the MIT Media Lab. Founded in 2006 under the leadership of Professor Thomas W. Malone, the CCI has distinguished itself as an influential academic institution dedicated to understanding how people and computers can collaborate more intelligently and how collective intelligence can be harnessed to address complex societal issues. The center is home to a multidisciplinary team of researchers specializing in fields such as computer science, economics, sociology, management, and human and artificial intelligence. Their research portfolio spans

a wide range of topics, from understanding group dynamics in problem-solving and decision-making to exploring new computational tools for promoting collective intelligence. Among its numerous significant contributions to the field, CCI is renowned for developing theoretical models, empirical studies, and designing distinctive methodologies to both measure and enhance collective intelligence.

One of the first important achievements of the scholars affiliated with this center was the investigative study delineated in the research report *Harnessing Crowds: Mapping the Genome of Collective Intelligence*, by Malone, Laubacher, and Dellarocas (2010). This seminal work introduced a relatively intricate typology and outlined distinct foundational elements, referred to as CI Genome: the *building blocks*, which could inform the structure of collective intelligence communities.

These blocks can be categorized into four main types:

- *What*: The nature of the task or problem being worked on. This could range from simple tasks, like image identification, to complex ones, such as software development or strategy planning.
- *Who*: The participants involved in the collective intelligence activity. They may be experts, amateurs, or a mixture of both. The participants could be internal to an organization, external, or a blend of internal and external participants.
- *Why*: The motivations of the participants to contribute. The incentives could be financial, such as wages or prizes, social motivations like reputation and community, intrinsic motivations such as enjoyment or a sense of purpose, or a combination of these.
- *How*: The process or method of gathering and managing the contributions. This could include collection processes like voting or averaging or integration processes such as collaborative editing or competition. It also includes the way the tasks are divided among the participants, whether in a centralized or decentralized manner, or whether the tasks are independent or interdependent.

The *genome* received its moniker because different combinations of these building blocks can lead to a wide variety of different kinds of collective intelligence systems, much like different combinations of genes lead to different organisms. But it is just a neat metaphor to describe existing web projects like Linux, Wikipedia, and InnoCentive. The *genome* acts as a prelude to an extensive program of laboratory studies.

These studies, which attracted substantial attention, involved measuring the level of collective intelligence within a specific group and determining factors that might impact CI level. This approach was inspired by the concept of measurable intelligence of individual humans, proposed by Spearman in 1904 and subsequently investigated by many psychologists. IQ researchers have repeatedly shown that a singular statistical factor, often referred to as *general intelligence* or the *g-factor*, emerges from the correlations among individuals' performance across a diverse array of cognitive tasks. In relation to this, the CCI research team led by Anita Woolley posited a counterpart hypothesis of a general collective intelligence factor, or the *c-factor*, which gauges a group's overall efficiency. They hypothesized that a group's CI, akin to individual intelligence, could be a factor affecting its performance across various

efforts. Critically, they aimed to determine whether collective intelligence is more than the sum of individual intelligences and therefore does not have a strong correlation with the average or maximum individual intelligence of group members. For this purpose, they recruited a total of 699 people to participate in a laboratory experiment in which randomly formed groups of 3–5 people working face-to-face were tasked to jointly solve visual puzzles, brainstorm, make collective moral judgments, and negotiate over limited resources. In the 2010 article, “Evidence for a Collective Intelligence Factor in the Performance of Human Groups,” published in *Science*, they describe the results of these experiments, which attracted the attention of many researchers (Woolley et al., 2010). First, they found evidence supporting the existence of a *c-factor*, discovering that groups with higher *c* performed better across various tasks in fields as disparate as logic, creativity, or negotiation. Second, the *c-factor* was observed to be less related to the average or maximum individual intelligence of group members but was significantly correlated with the average social sensitivity of group members and equality in conversational turn-taking: teams where members had higher levels of social sensitivity and where conversational turns were distributed more evenly, demonstrated higher CI. The authors also stated that a higher proportion of females (presumably because they are better at social sensitivity on average) was predictive of higher collective intelligence. But does that mean that all-female groups would be the best performers? Not necessarily. The second critical factor is cognitive diversity: the variety of skills and thinking styles. These findings suggest that enhancing collaboration within groups could improve their collective intelligence and that features related to social interaction play a crucial role in this process. Furthermore, the conclusions underscore the significance of diversity in team composition (Woolley et al., 2010).

This line of research was progressively expanded over the following years. It was demonstrated that the *c-factor* of a group is also evident in online collaboration (Engel et al., 2015). Subsequent studies explored various contexts, such as the effects of group structure (comparing short-term ad hoc groups with long-term ones), and the role of cultural environments, including the United States, Germany, and Japan (Engel et al., 2015). Further research examined the relationship between physiological synchrony and CI in computer-mediated teams (Chikersal et al., 2017) and tested whether the level of CI, as measured by the *c-factor*, could predict a group’s performance in other activities, like online gaming performance (Kim et al., 2017). Significantly, the ongoing research at CCI examined the convergent and discriminant validity of the *c-factor*, and collective intelligence in general, compared with other frequently studied group-based states and processes. This research observed a distinction between CI and measures of group climate or inter-member relationships (Woolley et al., 2018). No correlation was found between CI and group satisfaction (Chikersal et al., 2017; Engel et al., 2014), relationship quality (Woolley et al., 2018), or psychological safety (Glikson et al., 2016). By contrast, researchers *do* see a strong association between CI and transactive memory systems (Kim et al., 2016) and some forms of group learning (Aggarwal et al., 2015).

The findings developed by Woolley and her collaborators have received extensive recognition. According to Web of Science data as of July 2023, the article “Evidence

for a Collective Intelligence Factor in the Performance of Human Groups” has been cited as many as 1146 times, making it one of the most influential texts in the history of CI research. Both confirmatory findings by other researchers (Askay et al., 2019; Dickinson et al., 2012; Miller, 2010) and opposing viewpoints have emerged. For example, Credé and Howardson (2017) argue that the inference of a general collective intelligence construct is unwarranted because the *c-factor* does not correlate strongly with all tasks performed by groups. They also pointed out other *statistical artifacts* in the study design and resulting data suggest that group performance might be more of a situationally specific phenomenon. On the other hand, Bates and Gupta (2017) argue that the aggregate *g-factor* of individual group members still presents the most credible challenge to the hypothesis of situational specificity in group performance settings. In their experimental study, the mean IQ of a group’s individuals accounted for around 80% of the variance in the group’s collective IQ test scores. Hence, it appears that the *c-factor* does not equally manifest itself in every situation, and its level could be affected by additional factors contingent on the circumstances. Nevertheless, many researchers believe that the *c-factor* for CI “has been well established in the literature” (Askay et al., 2019, p. 492).

In Woolley’s research, cognitive diversity was one of the factors most strongly correlated with the group’s level of collective intelligence. As we may recall, diversity was defined by Surowiecki as one of the key factors influencing the emergence of the wisdom of the crowds, alongside independence, decentralization, and aggregation. Today, many researchers studying CI emphasize the role of diversity as a significant factor in group performance. For instance, Lo, in discussing the intelligence of financial markets, emphasizes that “CI may function best in relatively diverse and egalitarian settings in which many different opinions are represented and heard” (Lo, 2015). Landemore, in her book *Democratic Reason: Politics, Collective Intelligence, and the Rule of the Many* (2012), recognizes diversity as a pivotal factor that enables the manifestation of CI in democratic decision-making processes. She describes this diversity as the different ways of viewing the world and associated it with the law of large numbers: as the size of the group increases, the precision of individual predictive models becomes less crucial, whereas the diversity of the group models gains progressively more significance. In other words, “as groups become large, the criterion becomes less ‘is this person accurate’ than ‘is this person different’” (Landemore, 2012). Similarly, Baltzersen (2022) contends that in contemporary research on group intelligence “the predominant strategy is to scale up the size of the group and hope this can create more diversity benefits. (...) The group should be diverse, so different individuals can supplement each other with different pieces of information.”

The Diversity Prediction Theorem, developed by Hong and Page (2004), is a mathematical theory often referenced in Collective Intelligence (CI) research. This theorem delineates the relationship between collective accuracy and the diversity of predictions in relation to expected errors. In their theorem, Hong and Page assert that the error a problem-solving group generates is equal to the average squared individual error minus the predictive diversity. This may seem complex, but let’s delve deeper into the matter. Firstly, the mean square error, which is the average of individual squared errors, encompasses the errors each individual commits expressed as

a distance from the correct or true value. Secondly, predictive diversity is calculated as the average squared distance of individual predictions from the mean prediction. In statistical terms, this is analogous to variance, reflecting the dispersion of predictions. However, Page opts to use the term *diversity* to underscore the significance of variations in predictions—essentially framing this as a dimension of crowd diversity. Thirdly, the crowd error—the squared error of the collective prediction—is equivalent to the average squared error of the crowd's prediction minus predictive diversity. According to the theorem, the crowd square error is invariably equal to or less than the mean square error. Thus, the prediction made by a crowd is always superior to, or at least equal to, the average prediction made by its members. A significantly superior prediction demands a higher degree of diversity, while a crowd with no predictive diversity will perform only as well as its average member. The critical takeaway here is that when a group displays a high level of diversity, the crowd error will be minimal. This suggests that *wise crowds* are dependent not only on individual accuracy, represented by mean individual error, but also on crowd diversity (Baltzersen, 2022; Hong & Page, 2004). In broad terms, Hong and Page argue that when it comes to problem solving, diversity trumps individual ability. Landemore even goes so far as to assert that a larger collective is inherently more diverse, and therefore more effective at problem-solving, than a smaller one: “If three députés are more cognitively diverse and thus smarter than just one, then five hundred are likely even more cognitively diverse, and thus smarter” (Landemore, 2012, p. 104).

Hong and Page's theory has garnered considerable attention for asserting that homogeneous groups of experts may not be the best problem-solvers, contrary to intuitive belief. In light of growing interest in diversity issues, this theory has also been invoked in non-academic contexts, such as its citation in support of a diversity requirement at the University of California and its inclusion in a brief to the Supreme Court of the United States advocating for diversity in the Armed Forces (Singer, 2019). Interestingly, some authors use these theorems to substantiate the concepts they propose. This acts as an effort to provide a scientific endorsement to reassure their readers: the mathematical finding is presented as an irrefutable fact with the aim of bolstering their argument (Houlou-Garcia, 2017).

Nonetheless, the Diversity Prediction Theorem has also stirred controversy. For instance, Abigail Thompson, a professor of mathematics at the University of California and one of the vice presidents of the American Mathematical Society, critiqued it as “an example of the misuse of mathematics in the social sciences.” She attributed this misuse to oversimplification and the conflation of randomness with diversity (Thompson, 2014). Houlou-Garcia (2017) also raised concerns regarding poor mathematical inference. However, we should remember that collective intelligence studies are predominantly empirical rather than theoretical. Therefore, it would be pertinent to question to what extent empirical studies confirm or falsify this mathematical model.

At first glance, the series of studies conducted by Woolley and her team, described above, might seem to fully confirm the Diversity Prediction Theorem. After all, one of the most highly correlated features with team efficiency was diversity. However, upon closer inspection, significant reservations arise. In the text *The Impact of Cognitive*

Style Diversity on Implicit Learning in Teams by Aggarwal et al. (2019), the results of studies concerning empirical studies of diversity were presented. The authors initiate their study with the claim that despite prevalent theories, there is no conclusive empirical evidence that establishes a direct correlation between team diversity and team outcomes. Hence, it is crucial to undertake such a study. Now, let's delve into how the researchers carried out their investigation. The participant pool consisted of 337 individuals who were randomly assigned into 98 teams, each comprised of two to five members. Each team embarked on the collective intelligence battery test, initially devised by Woolley et al. in 2010. Following this, they participated in minimum-effort tacit coordination game, which differs from the standard prisoner's dilemma game in that teams are rewarded for coordinating rather than for competing (Van Huyck et al., 1990). Subsequently, the object-spatial imagery and verbal questionnaire, a tool developed by Blazhenkova and Kozhevnikov in 2009, was utilized to assess the range of cognitive styles within the team. The diversity of cognitive styles was quantified as the sum of the standard deviations of each cognitive style within the team. What did the findings reveal? They are presented in Fig. 2.6. The diversity measure, namely *cognitive style diversity* has a curvilinear (inverted U-shaped) relationship with collective intelligence, with high levels of diversity disrupting CI. The next phase of the study concerned team learning abilities. It was observed that high levels of diversity impede the learning of implicit coordination through the disruption of collective intelligence.

These observations led the authors to conclude that, considering the array of problems faced by teams, achieving an optimal balance between cognitive diversity and collective intelligence can be quite intricate. If there is insufficient cognitive diversity, the team may lack vital cognitive resources needed to address various tasks. Furthermore, a deficiency in diversity could mean that the team does not have clear differentiators and indicators, making it harder to recognize which member brings necessary knowledge and skills to the table, thereby hindering the effective use of team skills. On the other hand, excessive cognitive diversity may result in communication and understanding barriers among team members, bringing the possibility that members of the team may not understand each other. It could also lead to coordination costs outweighing the potential benefits of individual contributions. Consequently, teams might perform below their potential given their collective skills and inputs. So, as argued by Aggarwal et al., if diversity has a curvilinear—inverted (U-shaped) relationship with CI, it follows that too much diversity can in some cases lead to a reduced CI level (Aggarwal et al., 2019) (Fig. 2.7).

However, there is one more important aspect to discuss: whether collective intelligence researchers believe that any type of diversity impacts a group's level of CI? Do diverse groups, like deliberative mini-publics, have to represent a miniature society in which all social groups and various types of minorities are represented? A careful look at the chart above will reveal to us that it's not about arbitrary diversity, but cognitive diversity, meaning the differentiation in thinking styles or preferred ways of utilizing the abilities one possesses. (Sternberg & Grigorenko, 1997). As Landemore concludes, this cognitive diversity relies on “diversity of perspectives (the way

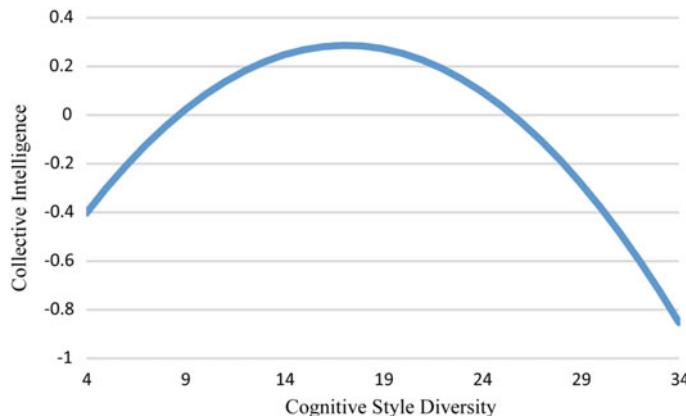


Fig. 2.7 The relationship between cognitive diversity and collective intelligence, taking into account team size and the level of cognitive style (Aggarwal et al., 2019)

of representing situations and problems), diversity of interpretations (the way of categorizing or partitioning perspectives), diversity of heuristics (the way of generating solutions to problems), and diversity of predictive models (the way of inferring cause and effect)" (2012).

The concept of cognitive style was developed by behavioral psychologists during the twentieth century. Klein and Schlesinger (1951) initially introduced this notion, focusing on the potential relationship between individual differences in perception and personality traits. Klein places strong emphasis on the controlling aspect of cognitive styles, highlighting their guiding function in an individual's activity. Witkin, another early researcher, viewed cognitive styles as patterns or modes of adjusting to the world. He believed adjusting to the world may lead to the adoption of different cognitive strategies and can result in variations in the perception of reality. He also associated cognitive styles with individual differences in the way people perceive, think, problem-solve, learn, and interact with others (Witkin et al., 1962).

Kozhevnikov (2007), in a review of the primary trends in cognitive style studies, suggests that this research has revealed that individuals adopt different methods to solve cognitive tasks. Furthermore, the preferences for these methods appear to be relatively stable over time, relating to both intelligence and personality traits. Studies carried out in applied fields have broadened the concept of cognitive styles, describing individual differences in both lower-level (primarily perceptual) cognitive functioning and more complex cognitive processing. These studies also emphasize that cognitive styles are not simply innate structures, dependent solely on an individual's internal characteristics. Instead, they are interactive constructs that develop in response to social, educational, professional, and other environmental requirements. Kozhevnikov highlights Nosal's (1990) proposal as one of the most theoretically grounded models for systematizing cognitive styles within the framework of information processing theory. In Nosal's model, a matrix is formed with four levels

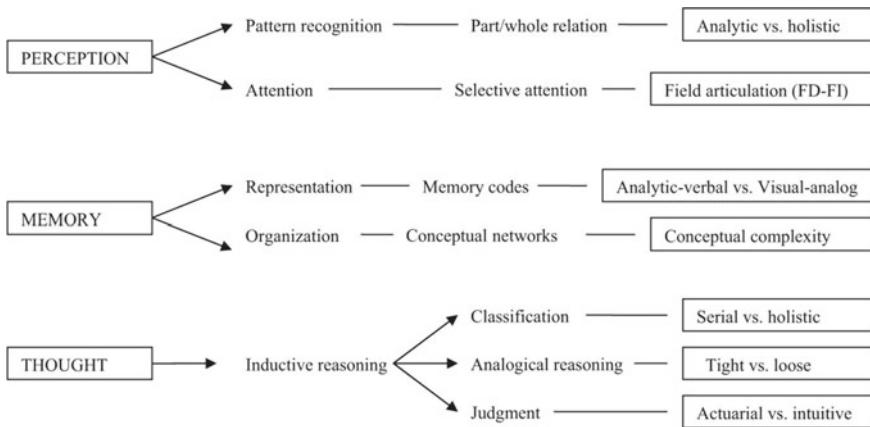


Fig. 2.8 A model of cognitive styles and the related cognitive processes proposed by Miller (1987), in “Cognitive Styles: An Integrated Model,” *Educational Psychology*, 7, p. 253 (Kozhevnikov, 2007)

(ranging from simple perception to complex decision making) and four methods (ranging from automatic data encoding to conscious allocation of mental resources). Twelve cognitive styles can then be placed at the intersections of the matrix. Nosal’s categorization of cognitive styles includes field dependence–independence, field articulation (element versus form articulation), breadth of conceptualization, range of equivalence, articulation of conceptual structure, tolerance for unrealistic experience, leveling-sharpening, range of scanning, reflectivity-impulsivity, rigidity-flexibility of control, locus of control, and time orientation. A similar concept was also presented by Miller (1987), who shows an integrated model of cognitive styles and the related cognitive processes (Fig. 2.8).

This framework offers a theoretical grounding for categorizing cognitive styles based on the level of information processing they involve and the types of information processing they regulate. That is to say, Nosal was the first to suggest a theory asserting that cognitive styles function at varying levels of cognitive complexity and can regulate different types of mental processes that could be employed at any of these levels (Kozhevnikov, 2007). Kozhevnikov concludes that cognitive styles embody the heuristics that an individual utilizes to process information about their surroundings. These heuristics are identifiable at each level of information processing, from perceptual to metacognitive. Their primary function is regulatory, controlling processes from automatic data encoding to the conscious allocation of cognitive resources. They serve an adaptive purpose, mediating the relationship between an individual and their environment. From this viewpoint, cognitive styles can be seen as unique patterns of adjustment to the world that evolve gradually through experience. This evolution is a result of the interplay between fundamental individual characteristics (such as general intelligence and personality) and enduring external requirements (like education, both formal and informal training, professional demands, and the cultural and social environment).

The terminology related to individual thinking styles and their role in interaction with the environment have been largely adopted by collective intelligence researchers from the field of behavioral psychology. However, there has been a parallel progression in the study of cognition through the emergence of a new, interdisciplinary field of study called cognitive sciences. This interdisciplinary field, incorporating linguistics, psychology, neuroscience, philosophy, computer science/artificial intelligence, and anthropology, strives for a more profound understanding of the mind and learning. The focus here is on how nervous systems represent, process, and transform information. The contemporary culture of cognitive sciences can be traced back to the early cyberneticists of the 1930s and 1940s, such as Warren McCulloch and Walter Pitts, who sought to understand the organizing principles of the mind. McCulloch and Pitts developed the first versions of what are now known as artificial neural networks, computational models inspired by the structure of biological neural networks (Dupuy, 2001). The cybernetic underpinnings of this approach are most clearly signified by the fact that the term *cognitive sciences* itself was coined by Christopher Longuet-Higgins in his 1973 commentary on the Lighthill report, which discussed the prevailing state of artificial intelligence research at that time (Longuet-Higgins, 1973). Research methods from cybernetics, neuroscience, and neuropsychology have been used by cognitive scientists to understand how intelligent behavior is implemented in a physical system, for instance, providing insights into how damage to specific areas of the brain affects cognition. Cognitive sciences has also given rise to models of human cognitive bias and risk perception and has played a significant role in the development of behavioral finance, a branch of economics (Dupuy, 2001).

What is particularly interesting for us, cognitive sciences adapt concepts known from psychology related to attention, cognition, memory, learning, development or information processing. However, cognitive scientists go a step further: they use these concepts to describe not only human individuals, but also communities. Cognitive sciences, unlike psychology, assumes that thinking is not just a one-person activity, but it can also be extended to families, communities, and nations. One of the most famous researchers in this area, evolutionary anthropologist Michael Tomasello, has attracted much attention in the academic world and beyond in recent years. Tomasello, conducting his research at the Max Planck Institute for Evolutionary Anthropology in Leipzig, examines cognitive abilities from a comparative perspective. Along with his research team, he has designed a series of experimental devices to test the spatial, instrumental, and social cognition of toddlers (ranging from 6 to 24 months) and apes. The results of these studies highlight that social, or even ultrasocial, cognition is what truly distinguishes humans from other species (Tomasello, 2014a, b).

In Tomasello's view, the thinking of each individual is interwoven into a social-species matrix and is associated with the shared intentionality inherent to a given group. Collective thinking develops when we feel like we are part of a community, thus fostering the collective cultivation of the best decision-making methods. These methods are then solidified through the cumulative gathering of group experiences as behavioral patterns, customs, and shared values. Tomasello's empirical research

suggests that both humans (including infants) and primates participate in cooperative actions within their groups. However, only humans are capable of coordinating their thought processes towards the achievement of common goals. They perceive themselves as a collective agent—a group with common interests—sharing experiences and tasks (exhibiting *we-intentionality*) (Tomasello, 2009, 2014b). Moreover, evolutionarily advantageous patterns established by the group are reinforced in the form of an evolutionarily advanced *intuitive heuristic*, which allows for the gradual refinement of cognitive processes. Empirical evidence indicates that only humans self-regulate and evaluate their own thinking in relation to the normative perspectives and standards of other group members. In contrast, primates, though social creatures, essentially lead individualistic and competition-based lives, aimed solely at achieving individual goals. Early humans were forced by natural conditions to invent ways of coordinating work with others to achieve common goals with other individuals, or even collective goals of the entire group. Tomasello argues that, over time, this led to the creation of a common cultural ground in which cooperation and communication were conventionalized, institutionalized, and became normative, leading to the creation of collective intentionality (Tomasello, 1999, 2014b).

An important component of this evolution was also the differentiation of the context of the *in-group*, the *out-group*, and the competition between them. Bennet and Sani (2008), other cognitive researchers dealing with children's subjective identification with social groups, note that the most remarkable phenomenon related to group identity is the collective sense of pride, guilt, or shame: when an individual belonging to a given group accomplishes something noteworthy, other individuals of that group feel pride, shame, or guilt in the same way as if they had done it themselves. Tomasello believes that in the course of cultural evolution, from primitive societies to the first civilizations, humans were no longer satisfied with reasoning from simple causal and intentional relations and started to create shared cognitive models of the world. These models transcended the thoughts and attitudes of individual members and were upheld by collective normative judgments. An essential component of these judgments is cooperative argumentation, in which we strive to reach a common, group decision about the direction of actions or beliefs. We do this through assertions, which are supposed to represent the truth, but also through supporting these assertions with reasons and justifications that we collectively recognize as true and credible. Importantly, individuals whose assertions contradicted each other, or who changed the meaning of the concepts they used during the discussion, were excluded from the group decision-making process (Tomasello, 2014b).

Such a complex system of collective thinking based on multifaceted interactions and social organization is, according to Tomasello, the result of thousands of years of cumulative cultural evolution aimed primarily at species survival and, subsequently, joint action and social problem-solving. The coordination of common actions is a particularly important aspect of collective intentionality. Tomasello cites as examples: collective actions of entire communities directed towards survival in primitive conditions, the division of tasks in hunting and childcare, issues of leadership and marriage, as well as protection against norm violations, e.g., tyranny within the

group or aggression from other groups that allow predation on the work of others (Tomasello, 2009).

Examples related to life in primitive societies, despite their apparent simplicity, serve as an interesting reference point for researchers of contemporary forms of collective thinking. The achievements of cognitive sciences, particularly the concepts of cognitive processes and cooperative argumentation, can be significant for empirical research on online collective intelligence, though, as we will see in the next chapter, they are not fully applied. However, these studies, to date, are often fragmented and typically focus on a specific subset of phenomena related to collective thinking, not the whole process. An interesting approach to what Tomasello refers to as *cooperative argumentation* is the subject of research by another scholar associated with MIT's Center for Collective Intelligence, Mark Klein. His work has involved experimental attempts to create an online environment in which certain organizational frameworks were proposed for group argumentation, with the aim of building consensus. For a community already in agreement on crucial matters, this strategy can deepen comprehension around a particular subject and facilitate methodical, structured decision-making. Adhering to this method, group members can contribute more than just additions to an ongoing, free-flowing online discussion. They have the option to represent their perspectives in an online deliberation map, which explicitly lays out the logical structure of the argument (Malone, 2018). Participants in the discussion can contribute different arguments in an orderly manner, find their place in the debate structure, propose arguments for and against, and rate other people's arguments on a scale. Empirical research using the Deliberatorium platform included a debate within the student community of the University of Naples concerning the future of biofuels (Gürkan et al., 2010), a discussion within the Intel community about *open computing*, and a large-scale deliberation involving approximately 600 members of Italy's Democratic Party, where potential changes to Italian election laws were discussed (Malone, 2018). The most intriguing aspect of these studies is the testing of various types of quantified *deliberation metrics*, within which measurements such as the maturity of the debate (based on the breadth and depth of issues, ideas, and arguments), controversy of arguments, group thinking, and clustering regarding polarization or balkanization are carried out (Klein, 2015).

Among other interesting research related to cognitive processes, such as group attention and information filtering, it's worth noting the work presented by Engel and Malone on measuring the level of information integration in groups. They utilized the *phi* metric, originally developed by neuroscientists as a measure of consciousness in the brain. The definition of *phi* was originally formulated to analyze the activation patterns within groups of neurons, which send out electrical signals in a specific manner when active. The mathematical concept of integrated information provides a quantitative way of measuring a combination of two properties: (1) differentiated information, and (2) the level of integration. The theoretical premise underlying *phi* is the ability of individual parts of a system to function as a whole, not just as independent, mutually duplicating subsystems, and extract a single integrated result from a plethora of information (Engel & Malone, 2018). Moreover, the neuroscientists who

proposed *phi* argued that these two properties are essential to the subjective experience of consciousness (Tononi, 2004, 2008). Thus, a high level of *phi* observed in studied groups, in conjunction with their high level of efficiency, may suggest that they have attained a sort of group consciousness.

This research is one among many examples of specialized empirical studies, simulations modeling collective intelligence phenomena, and developments in swarm AI. These are currently of great interest to researchers in the field of collective intelligence. Such endeavors include a study suggesting that introducing intermittent breaks in group interactions during task execution can enhance collective intelligence levels (Bernstein et al., 2018), measuring user-perceived values of collective intelligence in online social networks (Weng et al., 2018), and simulating the cost of cooperation in collaborative problem-solving (Guazzini et al., 2018). A series of studies conducted by scientists from the Polytechnic University of Bari was also noteworthy. They utilized a model based on statistical physics where collective dynamics are governed by a continuous-time Markov process to simulate various aspects of collective intelligence. These included investigating the influence of the scope of distrust on group performance (Massari et al., 2019), the impact of the level of conflict, criticality, and disagreement on collective intelligence compared to a situation where consensus prevails (Vincenzo et al., 2017), and the simulation of how collective intelligence is influenced by individual interest, consensus-seeking, self-confidence of individual agents, and their cognitive levels (Guazzini et al., 2018). Significantly, in most other existing studies on CI, the general ideal is consensus, a fundamental aspect of deliberative theory. Only a handful of studies suggest that dissent can be advantageous for the debate as it prevents trivialization and declining engagement.

A separate domain of investigation that operates as a subdivision of artificial intelligence research, reaching beyond the confines of CI but frequently serving as inspiration for it, is known as swarm intelligence (SI). SI is based on the investigation of collective behavior of decentralized, self-organized systems. These systems can be either natural, like ant colonies, bird flocking, animal herding, bacterial growth, and fish schooling, or artificial, typically computer simulations mimicking such behaviors for problem-solving or optimization tasks. In swarm intelligence, there is no centralized control structure dictating how individual agents should behave. Instead, each agent in the swarm acts independently, following simple rules, and changes its behaviors based on local interactions with other agents and with the environment. The basic principle is that the collective behaviors of these simple agents interacting locally with their environment cause a process of self-organization leading to the emergence of ‘intelligent’ global behavior, unknown to the individual agents (Bonabeau et al., 1999; Kennedy et al., 2001).

The best-known swarm intelligence models developed in research projects include:

- *Ant Colony Optimization (ACO)*: Inspired by the foraging behavior of ants, ACO is a technique for solving computational problems which can be reduced to finding good paths through graphs. It has been used successfully in a number

of applications, from routing in telecommunication networks to protein folding in bioinformatics (Dorigo & Caro, 1999).

- *Particle Swarm Optimization (PSO)*: Inspired by the social behavior of bird flocking or fish schooling, PSO is a method used for optimizing a problem by iteratively trying to improve a candidate solution with regard to a given measure of quality. It has been applied in many fields such as neural network training, fuzzy system control, and other areas of science and engineering (Kennedy & Eberhart, 1995).
- *Bee Algorithm (BA)*: This is an optimization algorithm inspired by the natural foraging behavior of honeybees to find the optimal solution. It has been used for optimization problems, including multi-objective tasks (Pham et al., 2005).
- *Swarm Robotics (SR)*: Inspired by social insects, swarm robotics is an approach to collective robotics that emphasizes decentralization and self-organization. SR has shown its potential in tasks that demand high levels of autonomy and cooperation, such as in search and rescue operations. They could also do pollution cleanup or environmental monitoring where individual robots might fail due to the complexity of the task or harshness of the environment (Sahin, 2005).

SI algorithms have proven highly efficient for a variety of optimization problems, including traveling salesman problems, vehicle routing, task allocation, and scheduling problems, among others. For example, in 2018, a team of researchers at MIT used swarm intelligence to design an algorithm called “Orienteering,” which improves the efficiency of taxi fleets. This algorithm takes inspiration from the collective behavior of ants and bees and applies it to optimize taxi dispatching. In nature, ants and bees collectively find the shortest path to a food source. When an ant finds food, it returns to the nest, leaving a trail of pheromones. Other ants follow this path, reinforcing the pheromone trail and allowing the colony to reach the food source more quickly. This behavior, when translated into an algorithm, helps in optimizing paths in various contexts, like vehicle routing. The “Orienteering” algorithm uses this approach to predict demand for taxis and optimally route them to areas with high demand. By doing so, the algorithm increases the productivity of the taxi fleet by reducing the time taxis are idle or roaming without passengers, leading to higher profitability and better service. The algorithm constantly learns and adjusts to changing conditions such as varying traffic patterns, weather conditions, and demand trends, making it a flexible and dynamic solution (Alonso et al., 2018).

Even though empirical research on CI has ventured into numerous compelling threads, it seems that several concepts—those which Tomasello and other cognitivists regard as cornerstones of collective thinking—continue to receive only minimal attention within CI investigations. I am primarily concerned with collective identity grounded in shared values, or the preservation of gathered knowledge in collective memory. Among the investigations that distinguish themselves for emphasizing these topics, one can include the study by Becker et al. (2019). In their experiments, they enlisted a total of eight groups, each comprising 35 individuals. They posed a series of questions related to current social issues and political events, the answers to which could be easily verified, for instance, ‘What was the U.S. unemployment

rate at the end of Barack Obama's presidential administration?" Importantly, each of these groups was homogeneous, i.e., composed of people who declared themselves either as Republicans or Democrats, in line with American political affiliations. The first experiment showed, contrary to Surowiecki's assumptions about the independence of members of crowds, that the accuracy of answers increased when participants communicated with each other (although they were not yet aware that the group was ideologically homogeneous). The authors call this the positive impact of social influence on group efficiency: "in decentralized communication networks, group estimates become reliably more accurate as a result of information exchange" (Becker et al., 2017). Moreover, despite the fact that the groups were not politically diverse, researchers found no evidence that social influence either reduces accuracy or increases polarization. Instead, they found out that social influence increased accuracy both for Republicans and Democrats, and also "decreased polarization despite the absence of between-group ties" (Becker et al., 2019). In the second experiment, the ideological identity of the groups was further strengthened by displaying symbols and colors associated with both political parties on the screen, and informing participants that the groups were politically homogeneous. And what was found? Group efficiency increased even more, and opinion polarization decreased! (Becker et al., 2019) Therefore, it seems that in this case, shared identity and group cohesion were more important for its efficiency than diversity and ideological impartiality. A somewhat similar phenomenon can be observed in Philip Tetlock's well-known Good Judgment Project: the best forecasters, whose group identity was reinforced by calling them *Superforecasters*, achieved better results than people operating in heterogeneous groups without a common identity (Mellers et al., 2015).

The concept of collective memory, which is another field not yet sufficiently explored in CI research, has garnered some attention in sociology, with Maurice Halbwachs (1877–1945) being a seminal figure in its development. Halbwachs, predominantly a theorist of macro-social processes, associated collective memory with the symbols, stories, narratives, and images that bind societies together (Eustache & Peschansky, 2022). Similar to Tomasello, Halbwachs links collective memory primarily with the ideas and customs that shape the lives of social classes, religious communities, or even entire nations and civilizations. The particular nature of a group's lived experience creates a shared memory and identity (Russell, 2006).

The mechanisms of preserving acquired knowledge in long-term memory are mentioned in some theoretical works as an element of CI processes but have not yet been raised to the foreground of CI. In the theoretical considerations presented in this chapter, especially in Lévy (1997) and Heylighen et al. (1996, 2004), collective memory is described as an important component of collective intelligence, and Baltzersen in his review of cultural-historical perspectives on CI cites many general examples of enhancing access to memory by digitizing information (2022, p. 361). However, there appears to be a dearth of empirical research examining the specific role of memory in the process of collective thinking, particularly within the context of projects and small to medium groups—a gap that should capture the interest of researchers of CI. The subsequent chapter delves further into this subject.

2.8 Conclusions

Thus, we arrive at the overarching conclusions of this chapter, wherein I have made an attempt to delineate theoretical considerations, present practical projects, and discuss ongoing research on collective intelligence. First and foremost, we observe a broad thematic range, diverse approaches, and the utilization of methods characteristic of varied disciplines, a trait typical for a young field of knowledge. Delving further, we notice the awareness of project leaders, visionaries, and researchers engaged in exploring these topics. They understand that by studying collective thought processes and supplementing them with the support of internet technologies, we can employ mechanisms as old as time itself, yet execute them on a larger scale, more effectively, and in a more conscious manner. However, a considerable amount of work still needs to be done in this area. As a scientific field, CI remains largely undeveloped and insufficiently theorized. The gulf between general considerations and highly specialized studies focusing on specific CI aspects is vast. There's a deficiency in systematic, empirical analysis of CI on a larger scale, such as in organizations, cities, nations, and networks. The existing typologies are centered on practice, often intending to categorize and synthesize distinct online practices without any use of a dedicated theoretical framework (Baltzersen, 2022, p. 6). Limited information is available regarding the high-level norms governing cognitive processes involved in collective thinking. Despite these challenges, the persistent interest in the subject among decision-makers, project managers, and scientists is encouraging. The sustained enthusiasm gives hope that research on collective intelligence is likely to become more systematized in the coming years, paving the way for a comprehensive theory rooted in empirical evidence and providing structure to the existing body of knowledge. As a scholar in the field of CI, I hope to make my modest attempt to contribute to this systematization, a goal I intend to pursue in the subsequent chapters.

References

Afuah, A., & Tucci, C. L. (2012). Crowdsourcing as a solution to distant search. *Academy of Management Review*, 37(3), 355–375.

Afuah, A. (2018). Crowdsourcing: a primer and research framework. In C. L. Tucci, A. Afuah, & G. Visconti (Eds.), *Creating and capturing value through crowdsourcing*. Oxford University Press.

Aggarwal, I., Woolley, A. W., Chabris, C. F., & Malone, T. W. (2019). The impact of cognitive style diversity on implicit learning in teams. *Frontiers in Psychology*, 10, 112. <https://doi.org/10.3389/fpsyg.2019.00112>

Aggarwal, I., Woolley, A. W., Chabris, C. F., & Malone, T. W. (2015). Cognitive diversity, collective intelligence, and learning. In *Proceedings of collective intelligence 2015*. Santa Clara, CA.

Aitamurto, T., & Chen, K. (2017). The value of crowdsourcing in public policymaking: Epistemic, democratic and economic value. *The Theory and Practice of Legislation*, 5(1), 55–72.

Aitamurto, T., & Landemore, H. (2016). Crowdsourced deliberation: The case of the crowdsourced off-road traffic law in Finland. *Policy and Internet*, 8(2), 174–196.

Aitamurto, T., Landemore, H., Lee, D., & Goel, A. (2014). Crowdsourced off-road traffic law experiment in finland. *Publications of the Committee for the Future, the Parliament of Finland*, 1, 2014.

Aitamurto, T., Landemore, H., & Galli, J. S. (2017). Unmasking the crowd: Participants' motivation factors, expectations, and profile in a crowdsourced law reform. *Information, Communication and Society*, 20(8), 1239–1260. <https://doi.org/10.1080/1369118X.2016.1228993>

Aitamurto, T., & Landemore, H. (2015). Five design principles for crowdsourced policy-making: Assessing the case of crowdsourced off-road traffic law in Finland. *Social Media for Organizations*, 2(1).

Aitamurto, T. (2014). *Crowdsourced off-road traffic law experiment in Finland*. Parliament of Finland.

Aitamurto, T. (2016) collective intelligence in law reforms: When the logic of the crowds and the logic of policymaking collide. *49th Hawaii international conference on system sciences. IEEE transactions*, pp. 2780–2789. HICSS'16.

Alonso, L., Goldbloom, A., Núñez, A., & Goldbloom, M. (2018). Orienteering algorithms for autonomous vehicles in uncertain environments: An approach based on multiobjective particle swarm optimization. *IEEE Transactions on Intelligent Transportation Systems*, 19(7), 2218–2229. <https://doi.org/10.1109/TITS.2017.2754980>

Askay, D., Metcalf, L., Rosenberg, L., & Willcox, G. (2019). Enhancing group social perceptiveness through a swarm-based decision-making platform. *Proceedings of the 52nd Hawaii international conference on system sciences*.

Atlee, T., & Por, G. (2000). Collective intelligence as a field of multi-disciplinary study and practice. https://www.academia.edu/1981875/Collective_Intelligence_as_a_Field_of_Multi-disciplinary_Stud...

Baltzersen, R. K. (2022). *Cultural-historical perspectives on collective intelligence*. Cambridge University Press.

Bardini, T. (2000). *Bootstrapping: Douglas Engelbart, coevolution, and the origins of personal computing*. Stanford University Press.

Bates, T. C., & Gupta, S. (2017). Smart groups of smart people: Evidence for IQ as the origin of collective intelligence in the performance of human groups. *Intelligence*, 60, 46–56.

Becker, J., Brackbill, D., & Centola, D. (2017). Network dynamics of social influence in the wisdom of crowds. *Proceedings of the National Academy of Sciences*, 114(26), E5070–E5076. <https://doi.org/10.1073/pnas.1615978114>

Becker, J., Porter, E., & Centola, D. (2019). The wisdom of partisan crowds. *Proceedings of the National Academy of Sciences*, 116(22), 10717–10722. <https://doi.org/10.1073/pnas.1817195116>

Belenzon, S., & Schankerman, M. (2015). Motivation and sorting of human capital in open innovation. *Strategic Management Journal*, 36(6), 795–820.

Bennett, M., & Sani, F. (2008). Children's subjective identification with social groups. In S. R. Levy & M. Killen (Eds.), *Intergroup attitudes and relations in childhood through adulthood* (pp. 19–31). Oxford University Press.

Bernstein, E., Shore, J., & Lazer, D. (2018). How intermittent breaks in interaction improve collective intelligence. *Proceedings of the National Academy of Sciences*, 115(35), 8734–8739.

Best, S. J., & Krueger, B. S. (2005). Analyzing the representativeness of internet political participation. *Political Behavior*, 27, 183–216. <https://doi.org/10.1007/s11109-005-3242-y>

Bjelland, O. M., & Wood, R. C. (2008). An inside view of IBM's 'innovation jam'. *MIT Sloan Management Review*, 50(1), 32–40. <https://www.proquest.com/scholarly-journals/inside-view-ibms-innovation-jam/docview/224960315/se-2>.

Bonabeau, E., Dorigo, M., & Theraulaz, G. (1999). *Swarm intelligence: From natural to artificial systems*. Oxford University Press.

Bonabeau, E. (2009). Decisions 2.0: The power of collective intelligence. *MIT Sloan Management Review*, 50(2), 45–52.

Bowser, A., & Shanley, L. S. (2013). *New visions in citizen science*. Commons Lab Case Study Series, Vol. 3. <http://www.scribd.com/collections/3840667/Commons-Lab-Science-and-Technology-Innovation-Program-STIP>.

Brabham, D. C. (2008). Crowdsourcing as a model for problem solving: An introduction and cases. *Convergence*, 14(1), 75–90. <https://doi.org/10.1177/1354856507084420>

Brabham, D. C. (2012). The myth of amateur crowds. *Information, Communication and Society*, 15(3), 394–410. <https://doi.org/10.1080/1369118X.2011.641991>

Brabham, D. C., & Radin, B. A. (2015). *Crowdsourcing in the Public Sector*. Georgetown University Press. <http://www.jstor.org/stable/j.ctt14bs09w>

Brooks, F. P. Jr. (1975). *The mythical man-month: Essays on software engineering*. Addison-Wesley.

Bush, V. (1945). As we may think. *The Atlantic Monthly*, 176(1), 101–108. <https://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/>.

Chikarsal, P., Tomprou, M., Kim, Y. J., Woolley, A. W., & Dabbish, L. (2017). Deep structures of collaboration: Physiological correlates of collective intelligence and group satisfaction. *Proceedings of the 20th ACM conference on computer-supported cooperative work and social computing (CSCW 2017)*.

‘Citizen Science’ Added to Oxford English Dictionary. (2014). <https://daily.zooniverse.org/2014/09/16/citizen-science-in-dictionary/>.

Credé, M., & Howardson, G. (2017). The structure of group task performance—A second look at “collective intelligence”: Comment on Woolley et al. (2010). *Journal of Applied Psychology*, 102(10), 1483–1492.

Decidim: Participatory Budgeting in Barcelona. (2023). <https://participedia.net/case/7425>.

Dickinson, J. L., Shirk, J., Bonter, D., Bonney, R., Crain, R. L., Martin, J., Phillips, T., & Purcell, K. (2012). The current state of citizen science as a tool for ecological research and public engagement. *Frontiers in Ecology and the Environment*, 10(6), 291–297. <https://doi.org/10.1890/110236>

Dorigo, M., & Caro, G. D. (1999). Ant colony optimization: A new meta-heuristic. In *Proceedings of the 1999 congress on evolutionary computation* (Vol. 2, pp. 1470–1477). IEEE Press. <https://doi.org/10.1109/CEC.1999.782657>.

Dryzek, J. S., & Goodin, R. E. (2006). Deliberative impacts: The macro-political uptake of mini-publics. *Politics and Society*, 34, 219–244.

Dupuy, J.-P. (2001). The mechanization of the mind: On the origins of cognitive science. *Princeton University Press*. <https://doi.org/10.1515/9781400823819>

Engel, D., Woolley, A. W., Aggarwal, I., Chabris, C. F., Takahashi, M., Nemoto, K., Kaiser, C., Kim, Y. J., & Malone, T. W. (2015). Collective intelligence in computer-mediated collaboration emerges in different contexts and cultures. *Proceedings of the SIGCHI conference on human factors in computing systems (CHI 2015)*, Seoul, Korea.

Engel, D., & Malone, T. W. (2018). Integrated information as a metric for group interaction. *PLoS ONE*, 13(10), e0205335. <https://doi.org/10.1371/journal.pone.0205335>

Engelbart, D. C. (1962). Augmenting Human Intellect: A Conceptual Framework. Stanford Research Institute, Summary Report AFOSR-3223.

Engelbart, D. C. (1995). Toward augmenting the human intellect and boosting our collective IQ. *Communications of the ACM*, 38(8), 30–33.

Engelbart, D. C., & English, W. K. (1968). A research center for augmenting human intellect. In *AFIPS conference proceedings of the 1968 fall joint computer conference* (Vol. 33, pp. 395–410). San Francisco, CA: Stanford Research Institute.

Engelbart, D. C., & Lehtman, H. (1988). Working together. *BYTE Magazine, December 1988*, 245–252.

Engelbart, D. C. (1972). Coordinated information services for a discipline- or mission-oriented community. *Proceedings of the second annual computer communications conference, San Jose, CA, Jan. 24 1972*. <https://www.dougengelbart.org/content/view/113/>

English, W. K., Engelbart, D. C., and Berman, M. L. (1967). Display-selection techniques for text manipulation. *IEEE Transactions on Human Factors in Electronics*, HFE-8(1), 5–15.

Ennadif, G. (2020). The city of Barcelona's participatory democracy open source platform. *Open Source Observatory (OSOR)*. <https://joinup.ec.europa.eu/collection/open-source-observatory-osor/news/participatory-democracy>.

Estellés-Arolas, E., & González-Ladrón-de-Guevara, F. (2012). Towards an integrated crowdsourcing definition. *Journal of Information Science*, 38(2), 189–200. <https://doi.org/10.1177/0165551512437638>

Eustache, F., & Peschanski, D. (2022). Toward new memory sciences: The programme 13-Novembre. In S. M. O'Mara (Ed.), *Progress in brain research* (Vol. 274). Elsevier.

Fishkin, J. (2009). *When the people speak: Deliberative democracy and public consultation*. Oxford University Press.

Furdyk, B. (2023). *14 Doritos super bowl commercials ranked worst to best*. <https://www.mashed.com/322430/doritos-super-bowl-commercials-ranked-worst-to-best/>.

Gallagher, S. (2000). Philosophical conceptions of the self: Implications for cognitive science. *Trends in Cognitive Sciences*, 4(1), 14–21. [https://doi.org/10.1016/S1364-6613\(99\)01417-5](https://doi.org/10.1016/S1364-6613(99)01417-5)

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. Basic Books.

Gastil, J., & Levine, P. (Eds.). (2005). *The deliberative democracy handbook: Strategies for effective civic engagement in the twenty-first century*. Jossey-Bass.

Ginsburg, T., and Elkins, Z. (2014). Stjórnarskrágerð á tímum gagnsæis: Ísland í samanburði [Drafting Constitutions in an Era of Transparency: Iceland in Comparative Perspective]. In J. Ólafsson (Ed.), *Tilraunir með Lýðraði—Ísland í Kreppu og Endurreisn*.

Glikson, E., Harush, R., Kim, Y. J., Woolley, A. W., & Erez, M. (2016). Psychological safety and collective intelligence in multicultural globally dispersed teams. *Interdisciplinary network for groups research (INGRoup) conference, Helsinki, Finland*.

Goleman, D. (1995). *Emotional intelligence*. Bantam Books.

Gordon, D. H. (2015). Collective behavior in animals: An ecological perspective. In: Malone, T. W.; Bernstein M. S. (ed.): *Handbook of collective intelligence*. The MIT Press, Cambridge-London.

Grönlund, K., Bächtiger, A., & Setälä, M. (Eds) (2014). *Deliberative mini-publics*. ECPR Press.

Guazzini, A., Duradoni, M., Lazzeri, A., & Gronchi, G. (2018). Simulating the cost of cooperation: A recipe for collaborative problem-solving. *Future Internet*, 10(6), 55. MDPI AG. <https://doi.org/10.3390/fi10060055>.

Gürkan, A., Iandoli, L., Klein, M., & Zollo, G. (2010). Mediating debate through on-line large-scale argumentation: Evidence from the field. *Information Sciences*, 180(19), 3686–3702. <https://doi.org/10.1016/j.ins.2010.06.011>

Gutmann, A., Thompson, D. (1996). *Democracy and Disagreement*. Harvard University Press.

Gutmann, A., & Thompson, D. (2004). *Why deliberative democracy?* Princeton University Press.

Habermas, J. (1991). *The structural transformation of the public sphere*. MIT Press.

Hameduddin, T., Fernandez, S., & Demircioglu, M. A. (2020). Conditions for open innovation in public organizations: Evidence from Challenge.gov. *Asia Pacific Journal of Public Administration*, 42(2), 111–131.

Henhappel, S., & Niedzielski, P. (2015). Crowdsourcing jako współczesne narzędzie kształtowania otwartych modeli procesów innowacyjnych. *Ekonomiczne Problemy Usług*, 121, 65–94. <https://doi.org/10.18276/epu.2015.121-04>.

Heylighen, F. (1999). Collective intelligence and its implementation on the web: Algorithms to develop a collective mental map. *Computational and Mathematical Organization Theory*, 5(3), 253–280.

Heylighen, F., Bollen, J., & Apostel, L. (1996). *The world-wide web as a super-brain: From metaphor to model*. <http://pespmc1.vub.ac.be/Papers/WWWSuperBRAIN.html>.

Heylighen, F., Heath, M., & Van Overwalle, F. (2004). The emergence of distributed cognition: A conceptual framework. *Proceedings of collective intentionality IV*, Siena, Italy.

Höbemägi, P. (2013). Eestis on vaimne vaakum ning puudus uue põlvkonna juhiomadustega tippintelktaalidest. <https://web.archive.org/web/20130313073540/http://www.delfi.ee/news/paevauidised/arvamus/priit-hobemagi-eestis-on-vaimne-vaakum-ning-puudus-uue-polvko-nna-juhiomadustega-tippintelktaalidest.d?id=65796264>.

Hong, L., & Page, S. (2004). Groups of diverse problem-solvers can outperform groups of high-ability problem-solvers. *PNAS*, 101, 16385–16389.

Houlou-Garcia, A. (2017). Collective wisdom, diversity and misuse of mathematics. *Revue Française de Science Politique*, 67(5).

Howe, J. (2006). The rise of crowdsourcing. *Wired*. <https://www.wired.com/2006/06/crowds/>.

Howe, J. (2008). *Crowdsourcing: Why the power of the crowd is driving the future of business*. Crown Business.

Hubbard, B. M. (1998). *Conscious evolution: Awakening the power of our social potential*. New World Library.

Ikediego, H. O., Ilkan, M., Abubakar, A. M., & Bekun, F. V. (2018). Crowd-sourcing (who, why and what). *International Journal of Crowd Science*, 2(1), 27–41. <https://doi.org/10.1108/IJCS-07-2017-0005>

Kennedy, J., & Eberhart, R. C. (1995). Particle swarm optimization. In *Proceedings of ICNN'95 - international conference on neural networks* (Vol. 4, pp. 1942–1948). IEEE Press. <https://doi.org/10.1109/ICNN.1995.488968>.

Kennedy, J., Eberhart, R., & Shi, Y. (2001). *Swarm intelligence*. Morgan Kaufmann Publishers.

Kim, Y. J., Aggarwal, I., & Woolley, A. W. (2016). How communication impacts team performance: Exploring collective intelligence and transactive memory system as mechanisms. *Annual Convention of the International Communication Association, Fukuoka, Japan*.

Kim, Y. J., Engel, D., Woolley, A. W., Lin, J., McArthur, N., & Malone, T. W. (2017). What makes a strong team? Using collective intelligence to predict performance of teams in league of legends. *Proceedings of the 20th ACM conference on computer-supported cooperative work and social computing (CSCW 2017)*.

Klein, G. S., & Schlesinger, H. J. (1951). Perceptual attitudes toward instability: I. Prediction of apparent movement experiences from rorschach responses. *Journal of Personality*, 19, 289–302.

Klein, M. (2015). The CATALYST deliberation analytics server. *MIT Working Paper*. <https://doi.org/10.2139/ssrn.2962524>.

Klip, A. (2013). *KOLUMN: Rahvakogu ja probleemitaju*. <https://web.archive.org/web/20130331083741/http://www.ajakiri.ut.ee/1180885>.

Kok, A. (2010). Icelandic national forum 2010. *Participedia*. <http://participedia.net/en/cases/icelandic-national-forum-2010>.

Korpela, E. J., Anderson, D. P., Bankay, R., Cobb, J., Howard, A., Lebofsky, M., Siemion, A. P. V., von Korff, J., & Werthimer, D. (2011). Status of the UC-Berkeley SETI efforts. In *Instruments, methods, and missions for astrobiology XIV* (Vol. 8152, p. 815212). SPIE. <https://doi.org/10.1117/12.894066>.

Kozhevnikov, M. (2007). Cognitive styles in the context of modern psychology: Toward an integrated framework of cognitive style. *Psychological Bulletin*, 133(3), 464–481. <https://doi.org/10.1037/0033-2909.133.3.464>

Lam, M. (2016). *Dell's ideastorm: Still co-creation?* <https://consumervaluecreation.com/2016/02/22/dells-ideastorm-still-co-creation/>.

Landemore, H. (2012). *Democratic reason: Politics, collective intelligence, and the rule of the many*. Oxford University Press.

Landemore, H. (2020). When public participation matters: The 2010–2013 Icelandic constitutional process. *International Journal of Constitutional Law*, 18(1), 179–205. <https://doi.org/10.1093/icon/moa004>

Le Bon, G. (1896). *The crowd: A study of the popular mind*. Macmillan.

Lepak, D. P., Smith, K. G., & Taylor, M. S. (2007). Value creation and value capture: A multilevel perspective. *Academy of Management Review*, 32(1), 180–194.

Levy, P. (1997). *Collective intelligence: Mankind's emerging world in cyberspace*. Plenum Press.

Levy, P. (1998). *Becoming virtual. Reality in the digital age*. Plenum Press.

Linders, D. (2012). From e-government to we-government: Defining a typology for citizen coproduction in the age of social media. *Government Information Quarterly*, 29, 446–454.

Liu, H. K. (2021). Crowdsourcing: Citizens as coproducers of public services. *Policy and Internet*, 13, 315–331. <https://doi.org/10.1002/poi3.249>

Lo, A.W. (2015). The wisdom of crowds vs. the madness of mobs. In: T. W. Malone, M. S. Bernstein (Eds), *Handbook of collective intelligence*. The MIT Press.

Loiselle, J. T., Lynch, M. P., & Sherrerd, M. R. (2009). *Evaluation of the peer-to-patent pilot program*. <https://digitalcommons.wpi.edu/iqp-all/1970>.

Longuet-Higgins, H. C. (1973). Comments on the lighthill report and the Sutherland reply. In *Artificial intelligence: Lighthill report* (pp. 35–37). Science Research Council.

Mackay, C. (1980). *Extraordinary popular delusions and the madness of crowds*. Harmony Books. (Original work published 1841).

Malone, T. W., Laubacher, R., & Dellarocas, C. (2010). The collective intelligence genome. *MIT Sloan Management Review*, 51(3), 21–31.

Malone, T. W. (2018). *Superminds*. Little, Brown.

Mancini, P. (2015). Why it is time to redesign our political system. *European View*, 14(1), 69–75. <https://doi.org/10.1007/s12290-015-0343-9>

Massari, G. F., Giannoccaro, I., & Carbone, G. (2019). Are distrust relationships beneficial for group performance? The influence of the scope of distrust on the emergence of collective intelligence. *International Journal of Production Economics*, 208, 343–355. <https://doi.org/10.1016/j.ijpe.2018.12.005>

Mayer-Kress, G., & Barczys, C. (1995). The global brain as an emergent structure from the worldwide computing network, and its implications for modeling. *The Information Society*, 11(1), 1–28.

Mellers, B., Stone, E., Murray, T., Minster, A., Rohrbaugh, N., Bishop, M., Chen, E., Baker, J., Hou, Y., Horowitz, M., Ungar, L., & Tetlock, P. (2015). Identifying and cultivating superforecasters as a method of improving probabilistic predictions. *Perspectives on Psychological Science*, 10(3), 267–281. <https://doi.org/10.1177/1745691615577794>

Mergel, I., & Desouza, K. C. (2013). Implementing open innovation in the public sector: The case of Challenge.gov. *Public Administration Review*, 73, 882–890.

Michelucci, P., & Dickinson, J. L. (2016). The power of crowds. *Science*, 351, 32–33. <https://doi.org/10.1126/science.aad6499>

Miller, A. (1987). Cognitive styles: An integrated model. *Educational Psychology*, 7, 251–268.

Miller, G. (2010). Social savvy boosts the collective intelligence of groups. *Science*, 330(6000), 22. <https://doi.org/10.1126/science.330.6000.22>

Nosal, C. S. (1990). *Psychologiczne modele umysłu* [Psychological models of mind]. PWN.

Noveck, B. S. (2006). Peer to patent": Collective intelligence, open review, and patent reform. *Harvard Journal of Law and Technology*, 20(1), 123.

Noveck, B. S. (2012). Opening the US patent process. *Science and Public Policy*, 39(1), 72–82.

Noveck, B. S. (2009). *Wiki government: How technology can make government better, democracy stronger, and citizens more powerful*. Brookings Institution Press.

Noveck, B. S. (2015). *Smart citizens, smarter state: The technologies of expertise and the future of governing*. Harvard University Press

Orange, B.A. (2014). *Universal and doritos team up for crash the super bowl contest*. <https://moveneweb.com/crash-super-bowl-2015-contest/>.

Parker, C. J., May, A., & Mitchell, V. (2014). User-centred design of neogeography: The impact of volunteered geographic information on users' perceptions of online map 'mashups.' *Ergonomics*, 57(7), 987–997.

Pateman, C. (1970). *Participation and democratic theory*. Cambridge University Press.

Peña-López, I. (2019). *Shifting participation into sovereignty: The case of decidim*. Huygens Editorial.

Peters, M. A. (2015). Interview with Pierre A. Lévy, French philosopher of collective intelligence. *Open Review of Educational Research*, 2(1), 259–266. <https://doi.org/10.1080/23265507.2015.1084477>.

Peterson, A. (2013). Here are the 30 questions the White House doesn't seem to want to answer. *Washington Post*. <https://www.washingtonpost.com/news/the-switch/wp/2013/08/13/here-are-the-30-questions-the-white-house-doesnt-want-to-answer/>.

Pham, D. T., Ghanbarzadeh, A., Koc, E., Otri, S., Rahim, S., & Zaidi, M. (2005). The bees algorithm-a novel tool for complex optimisation problems. In *Proceedings of the 2nd virtual international conference on intelligent production machines and systems* (IPROMS 2006) (pp. 454–459). Elsevier.

Preville, P. (2019) How Barcelona is leading a new era of digital democracy. *Medium*. 13 Nov 2019. <https://medium.com/sidewalk-talk/how-barcelona-is-leading-a-new-era-of-digital-democracy-4a033a98cf32>.

Provenzo, Jr., E. F. (1997). [Forward to:] Lévy, P. *Collective intelligence*. Plenum Trade.

Raymond, E. S. (1999) *Surprised by wealth, linux weekly news (1999-12-10)*, Retrieved from <https://lwn.net/1999/1216/a/esr-rich.html>.

Raymond, E. S. (2001). *The cathedral and the bazaar: Musings on linux and open source by an accidental revolutionary*. O'Reilly Media.

Rhodan, M. (2015). White house responds to petition Urging Obama to Pardon Edward Snowden. *Time*. <https://time.com/3974713/white-house-edward-snowden-petition/>.

Royo, S., Pina, V., & Garcia-Rayado, J. (2020). Decide Madrid: A critical analysis of an award-winning e-participation initiative. *Sustainability*, 12(1674), 1–19.

Roussopoulos, D., & Benello, C. G. (Eds.). (2005). *Participatory democracy: Prospects for democratizing democracy*. Black Rose Books.

Royse, M. (2020). *5 simple ways to become more useful*. <https://medium.com/illumination-curated/5-simple-ways-to-become-more-useful-13cae9870e74>.

Russell, N. (2006). Collective memory before and after halbwachs. *The French Review*, 79(4), 792–804.

Şahin, E. (2005). Swarm robotics: From sources of inspiration to domains of application. In *swarm robotics* (pp. 10–20). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-30552-1_2.

Singer, D. J. (2019). Diversity, not randomness, trumps ability. *Philosophy of Science*, 86(1), 178–191.

Sternberg, R. J., & Grigorenko, E. L. (1997). Are cognitive styles still in style. *American Psychologist*, 52, 700–712.

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. Cambridge University Press.

Strasser, B. J., Baudry, J., Mahr, D., Sanchez, G., & Tancoigne, E. (2019). 'Citizen Science'? Rethinking science and public participation. *Science and Technology Studies*, 32(2), 52–76. <https://doi.org/10.23987/sts.60425>.

Surowiecki, J. (2005). *The wisdom of crowds*. Anchor Books.

Sutherlin, G. (2013). A voice in the crowd: Broader implications for crowdsourcing translation during crisis. *Journal of Information Science*, 39(3), 397–409.

Taeihagh, A. (2017). Crowdsourcing: A new tool for policy-making? *Policy Sciences*, 50(4), 629–647. <https://doi.org/10.1007/s11077-017-9303-3>

Terras, M. (2015). Crowdsourcing in the digital humanities. In S. Schreibman, S. Siemens, & J. Unsworth (Eds), *A new companion to digital humanities*, Wiley-Blackwell, New York, NY, pp. 420–439.

The Linux Foundation. (2023). <https://www.linuxfoundation.org/>.

Thompson, A. (2014). Does diversity trump ability? An example of the misuse of mathematics in social sciences. *Notices of the AMS*, 61(9), 1024–1030.

Tomasello, M. (1999). *The cultural origins of human cognition*. Harvard University Press.

Tomasello, M. (2009). *Why we cooperate*. The MIT Press.

Tomasello, M. (2014a). The ultra-social animal. *European Journal of Social Psychology*, 44(3), 187–194. <https://doi.org/10.1002/ejsp.2015>

Tomasello, M. (2014b). *A natural history of human thinking*. Harvard University Press.

Tononi, G. (2004). An information integration theory of consciousness. *BMC Neuroscience*, 5, 42. <https://doi.org/10.1186/1471-2202-5-42>

Tononi, G. (2008). Consciousness as integrated information: A provisional manifesto. *Biological Bulletin*, 215, 216. <https://doi.org/10.2307/25470612>

Turchin, V. (1977). *The phenomenon of science: A cybernetic approach to human evolution*. Columbia University Press.

Van Huyck, J. B., Battalio, R. C., & Beil, R. O. (1990). Tacit coordination games, strategic uncertainty, and coordination failure. *American Economic Review*, 80, 234–248.

Vincenzo, I., Giannoccaro, I., Carbone, G., & Grigolini, P. (2017). Criticality triggers the emergence of collective intelligence in groups. *Physical Review E*, 96(2), 022309.

Viscusi, G., & Tucci, C. L. (2018). Three's a crowd? In C. L. Tucci, A. Afuah, & G. Viscusi (Eds.), *Creating and capturing value through crowdsourcing*. Oxford University Press.

Wazoku Crowd. (2023). <https://www.wazokucrowd.com/>.

Wechsler, D. (2008). *Wechsler adult intelligence scale–fourth edition*. NCS Pearson.

Wellman, B. (2001). Computer networks as social networks. *Science*, 293(5537), 2031–2034.

Wells, H. G. (1938). *World brain: The idea of a permanent world encyclopaedia*. Methuen and Company.

Weng, S.-S., Yang, M.-H., & Hsiao, P.-I. (2018). A factor-identifying study of the user-perceived value of collective intelligence based on online social networks. *Internet Research*, 28(3), 696–715. <https://doi.org/10.1108/IntR-03-2017-0103>

Williams, S. (2010). *Free as in freedom (2.0): Richard Stallman and the free software revolution*. GNU Press.

Witkin, H. A., Dyk, R. B., Faterson, H. F., Goodenough, D. R., & Karp, S. A. (1962). *Psychological differentiation*. Wiley.

Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., & Malone, T. W. (2010). Evidence for a collective intelligence factor in the performance of human groups. *Science*, 330(6004), 686–688.

Woolley, A. W., Kim, Y., & Malone, T. W. (2018). Measuring collective intelligence in groups: A reply to Credé and Howardson. *MIT Sloan Working Paper 5431–18*.



What is missing from the policy analyst's tool kit—and from the set of accepted, well-developed theories of human organization—is an adequately specified theory of collective action whereby a group of principals can organize themselves voluntarily to retain the residuals of their efforts.

Elinor Ostrom (2015, p. 22)

3.1 Researching Collective Intelligence in Policymaking: An Overview of Past Approaches

3.1.1 *Looking for an Appropriate Analysis Method*

If collective intelligence in policymaking is to be useful, its outcomes must lead to improved decisions, more accurate conclusions, and effective solutions to problems—in other words, intelligent outcomes by some standard. The most significant challenge has been identifying outcomes that are somehow better or worse, meaning more or less intelligent. In fields such as open-source software development or innovation jams in enterprises, it is much easier to assess the usefulness of the solutions developed. In the field of public policies, however, it is difficult to define a uniform standard for assessing results. We must also be aware that collective debates may occasionally result in more foolishness than wisdom, leading to less effective decisions and less efficient solutions (Capella et al., 2017). In this chapter, we will review the existing approaches to the study of CI in policymaking and verify the factors that can improve or hinder collective intelligence. An original project evaluation method and several evaluations we have made will also be presented.

In late 2020, within my small research team at AGH University of Krakow, we began to consider which research method would be the most appropriate for analyzing projects related to public policies that incorporate collective online intelligence. Our group, comprised of social and computer scientists, had already been established for

three years. Our primary focus was to study the various applications of information technology in the public sphere, politics, administration, and economy, including e-democracy, e-government, and the broad spectrum of society's interactions with technologies. We have been interested in collective intelligence for some time, having already conducted a preliminary laboratory experiment and published several articles and conference presentations. Thus, we felt it was time to approach the subject more rigorously.

We planned to systematically review several intriguing projects employing various CI and crowdsourcing approaches. Therefore, we compiled a preliminary list of 18 initiatives worldwide, ensuring no regions were overlooked. As such, our list featured initiatives from the United States, Western Europe, Central and Eastern Europe, Sub-Saharan Africa, New Zealand, India, and other parts of the globe. Even from this initial review, it was clear that these projects would showcase different approaches and understandings of CI. This prompted us to consider a method that would be sufficiently versatile to extract the most critical aspects from these projects while identifying their commonalities.

We began our search for the "right" method by analyzing the existing ones. We believe that in such a widely discussed domain, there must be an appropriate research approach that we could adapt to our needs. Initially, we considered using the strategy proposed by The Governance Lab (GovLab) at New York University, mentioned in the previous chapter. Two entries in their published series, *Collective Intelligence and Governance*, caught our attention. The first, *The Open Policymaking Playbook*, detailed 12 case studies and 17 online platforms or tools for planning new projects. This publication highlighted their pros and cons and considered the time and money required for their setup and project execution (Noveck et al., 2019). The second, *CrowdLaw for Congress*, described another set of 12 cases, specifying their owners, locations, implementation levels, platforms used, and the participatory tasks executed within them (Noveck et al., 2020). In both publications, the foundational analysis involved categorizing each described case into one of four stages that, according to the authors, any policymaking cycle can be divided into (1) Understanding the problem, (2) Developing solutions, (3) Drafting, and (4) Evaluation and assessment.

At first glance, this method seemed quite promising. We were dealing with an adaptation of the well-established theory of public policy cycles (see Chap. 1). Upon deeper reflection, however, I noticed certain shortcomings. In the analyses presented by GovLab, projects were assigned to specific phases of policymaking in such a way that each was viewed as executing tasks within only one phase. This could suggest that individual phases can be planned, executed, and analyzed separately. As we recall, this mechanistic approach to policy cycles was characteristic of the positivist understanding of policymaking, which we found ill-suited for addressing socially complex and multifaceted endeavors. We concurred that considering the phases of policy creation can be helpful, allowing for the systematization of administrative work. However, when studying projects prioritizing citizen participation in the public domain, we would like to see how they influence the broader social context of the policy rather than just a segment. After all, policy processes are not linear but operate concurrently. Therefore, we believed that analyzing a single process in isolation and

assigning it to one project was too narrow a view. A broader perspective was essential. We were convinced that it's impossible to encompass all social factors within one policy model. Such an approach is appropriate under specific conditions: political consensus, minimal ideological polarization, and social stability. Looking at the current state of online debate, I am inclined to believe we are witnessing a situation quite distant from ideal. Finally, in the GovLab analyses, we noted the omission of a pivotal aspect of policymaking emphasized by the post-positivist school: social learning. This perspective would involve examining the continuity and context of collective knowledge and valuing the enduring behavioral changes that arise from experience and adaptive practices.

The analyses conducted by GovLab were both useful and inspiring. However, their approach left us somewhat unsatisfied, as they only marginally referred to *collective intelligence* in the examined projects. Their reference to CI appeared to be mainly in the series title, with the analysis predominantly focusing on the policy dimension. On the other hand, our interest lies in understanding how one can genuinely assess collective intelligence in policymaking. We did not want to regard the examined projects as instruments, merely serving as useful tools for the administration to achieve predetermined goals. From our perspective, CI represented more than just *online projects outsourcing administrative tasks* designed to yield specific benefits. We were keen to understand how and why collective thinking emerges, what kind of communities can be engaged as policy actors and under what circumstances, and what processes within a group focused on public matters influence its intelligent behavior.

For this reason, we decided to broaden our search. We outlined a systematic review of academic literature on online CI and its interplay with policy, exploring how these topics have been studied to date. We were particularly interested in methods that would include cognitive processes related to intelligence in the analysis of online CI. We were fully aware of the achievements of cognitive psychology in studying these processes at the individual level, described, e.g., by Ulric Neisser (1961), Robert B. Cattell (1963), or Karin and Robert Sternberg (2012). Cognitivists sought to break down the mechanisms of intelligence into fundamental components, avoiding the study of these processes in isolation. As Sternberg, a leading figure in cognitive psychology, argues, if we are truly examining a type of *thinking* and not merely gathering discrete opinions, then information processing should occur simultaneously and coherently. Intelligence is not, after all, a sequence of disconnected activities but a flexible and parallel use of different functional areas of the brain (Eisenhardt, 1989). In the works of several scholars describing CI theory, we encountered general insights on analogous cognitive processes observed within collectives (e.g., Heylighen, 1999, Malone, 2015, pp. 207–208; Steyvers & Miller, 2015, Mulgan, 2018, pp. 35–45). However, these insights were insufficient for conducting a systematic study of the projects that piqued our interest. We were hopeful that among the recent empirical studies we would be able to identify a method that would allow us to carry out the evaluation we had in mind.

3.1.2 *Review of Methods of Studying Collective Intelligence in Policymaking*

Our inquiry centered on the specific methods and strategies employed in studies regarding CI in policymaking over the past decade (Olszowski et al., 2021). Additionally, we sought to identify trends in the annual volume of publications, prevalent concepts within these studies, and the predominant research areas they spanned. The research methods and strategies found in seminal works and those addressing topics critical to collective intelligence were of particular interest. Moreover, we aimed to discern any statistical correlations between research methodologies and other attributes of the studies under review.

This systematic literature review adhered to the PRISMA methodology—short for “Preferred Reporting Items for Reviews and Meta-analyses”—which is among the most frequently used and esteemed methods for conducting reviews (Moher et al., 2009). We selected the Web of Science database from several options due to its renowned coverage, influence on the most cited authors and articles, and precise subject classification. The timeframe for the search was set for the period from 2011 to 2020. We applied a logical search to the topic (including abstract, keywords, and indexed fields), as well as the titles of the scientific articles. In addition, we used the language filter to focus on the publications in English. When selecting keywords to identify research papers, many alternative terms of CI used in the literature were taken into account, including: “collective intelligence,” “crowdsourcing,” “swarm intelligence,” “wisdom of crowds,” and “crowdlaw.” The second set of keywords included concepts related to political science, administration, and governance: “policymaking” (variants: “policy-making” and “policy making”), “public policy,” “political science,” “public administration,” “public sector,” and “public governance.” This search led to an initial number of 169 references. After removing duplicates, checking the inclusion criteria, excluding conference proceedings, editorial materials, or reviews, and articles during the eligibility assessment because they did not concern the topic of review (e.g., their topic was tourism, citizen science, student learning environment, etc.), we received a refined list of 88 top-quality studies.

Our initial observations highlighted a burgeoning interest in the subject matter. From 2012 to 2017, there was a discernible surge in engagement with CI. As illustrated in Fig. 3.1, 2017 marked the zenith of this interest, with 18 articles published. While there was a subsequent decline, 2020 witnessed a resurgence in the number of publications, slightly outpacing the prior year.

We analyzed the content of research articles in the review, creating lists of the most frequently occurring concepts in article titles, abstracts, original keywords, and those generated by the Web of Science’s KeyWords Plus algorithm. Notably, the term *crowdsourcing* dominated in both article titles and author keywords. Since this term, in its original meaning, mainly referred to business projects, we can see that many authors remain rooted in translating patterns developed in the commercial sector into the public domain. This interpretation aligns with the analysis of research methods (as we will see below). Next, terms like *public* and *government*, often

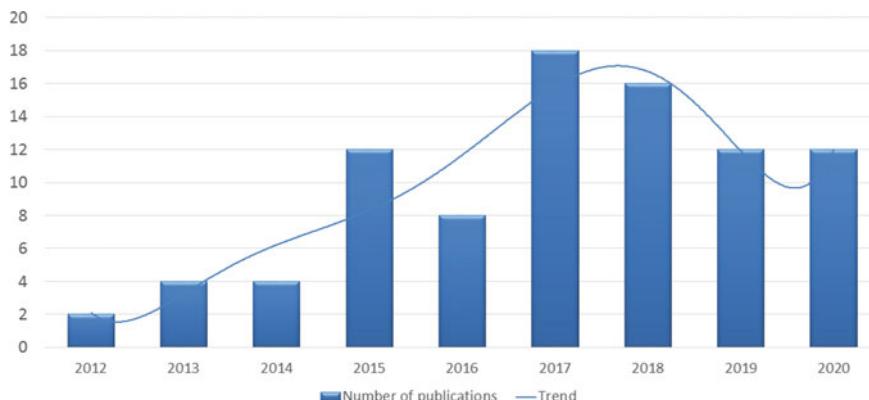


Fig. 3.1 The number of articles concerning the issues of collective intelligence and policymaking published annually and the growth trend for the period 2012–2020 (Olszowski et al., 2021)

found in article abstracts, anchor the research within political science. Furthermore, the KeyWords Plus analysis, based on the literature cited in the analyzed works, reveals that innovation and participation are the most frequently referenced concepts. Notably, *innovation*, traditionally understood in a business context as a “multi-stage process where organizations transform ideas into new or improved products, services, or processes” (Baregheh et al., 2009), is increasingly adopted in social and political sciences. In these fields, it often describes the reformative process by which public organizations open up to greater participation—an interpretation that resonates well with open policymaking (Table 3.1).

The subsequent analysis involved tracing within which research area studies on CI in policymaking were conducted. We utilized the Web of Science (WoS) Research Areas label to categorize the research domains that each journal receives upon publishing the evaluated texts. Every record within the Web of Science core collection contains a subject category, with each publication assigned to at least one category. Figure 3.2 below illustrates the WoS Research Areas where the texts were published during the review period. In the initial year of our analysis (2012), the studies were confined to only two research areas—information science and computer science—which are inherently interrelated. However, as the years progressed (with an exception in year three), we noted a systematic growth in the diversity of research areas. This diversity peaked in 2017 with 17 distinct areas, and this heightened level was almost maintained in 2018 and 2020, each boasting 16 research areas. Such expansion in research disciplines year after year underscores an increasing breadth and variety in the studies reviewed. As references to CI and policymaking emerge in increasingly specialized studies about public policy implementation, it becomes evident that discussions around CI in policymaking have evolved from broad contemplations to targeted applications in specific public policy domains.

The vast array of detailed research disciplines was quite overwhelming! We were intrigued to identify the broader academic domains where discussions on CI

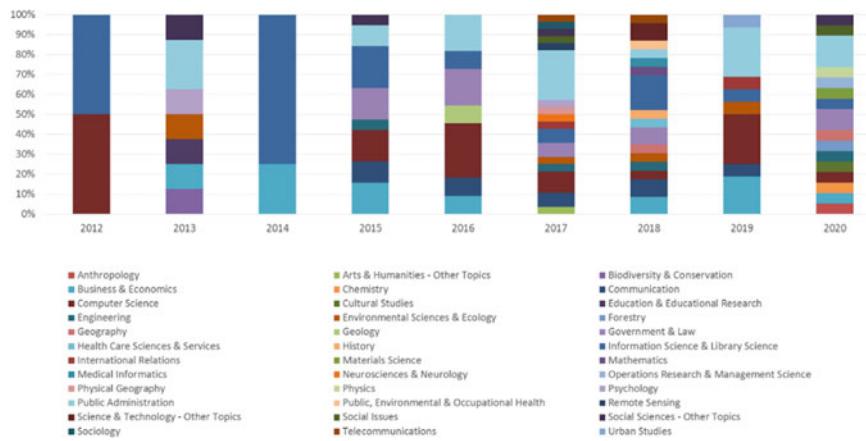
Table 3.1 Rankings of top-10 concepts in research papers on collective intelligence in policymaking, based on (a) article titles, (b) article abstracts, (c) author keywords, (d) KeyWords Plus (Olszowski et al., 2021)

Concept	Number of occurrences
<i>(a) Top 10 concepts in article titles</i>	
Crowdsourcing	24
Open	16
Public	16
Social	13
Innovation	11
Case	10
Government	9
Participation	9
Online	9
Policy	9
<i>(b) Top 10 concepts in article abstracts</i>	
Public	153
Crowdsourcing	125
Government	84
Data	82
Social	79
Open	78
Innovation	76
Research	64
Policy	63
Online	51
<i>(c) Top 10 concepts in author keywords</i>	
Crowdsourcing	50
Open	21
Public	21
Policy	19
Government	16
Innovation	16
Social	14
Participation	11
Data	10
Democracy	10
<i>(d) Top 10 concepts in keywords plus</i>	
Participation	14
Innovation	14
Media	9
Social	9

(continued)

Table 3.1 (continued)

Concept	Number of occurrences
Coproduction	8
Government	8
E-Government	7
Information	6
Democracy	6
Engagement	6

**Fig. 3.2** Research areas as per web of science in which studies on CI in policymaking were conducted, represented as a percentage per year (Olszowski et al., 2021)

in policymaking were flourishing. Was the focus mainly on political sciences, or did computer science lead the way? Or perhaps there was another dominant field? We clustered the emerging research areas into five Research Area Groups (RAGs) to simplify our analysis. We gave precedence to areas we believed were crucial, defining two broad categories: (1) Computer Science, Information Science, and related disciplines like Telecommunications, and (2) Political Sciences and affiliated areas, like Public Administration, Government and Law, etc. The remaining sectors that mentioned CI in policymaking were compiled into three categories: (3) Humanities and Social Sciences excluding Political Sciences (mainly Anthropology, Sociology, Psychology, and Cultural Studies), (4) Natural Sciences and Mathematics, and (5) Applied Sciences (e.g., Engineering, Business & Economics, Management Science). It's noteworthy that some articles spanned multiple fields and were thus classified under more than one research area, in line with the WoS categorization.

With this classification in place, we could distinctly analyze the annual number of studies within the Research Area Groups. Our findings indicated that, up until 2017, Computer Science and related fields dominated the landscape. However, from 2017 onward, political science emerged as the principal domain for research on collective

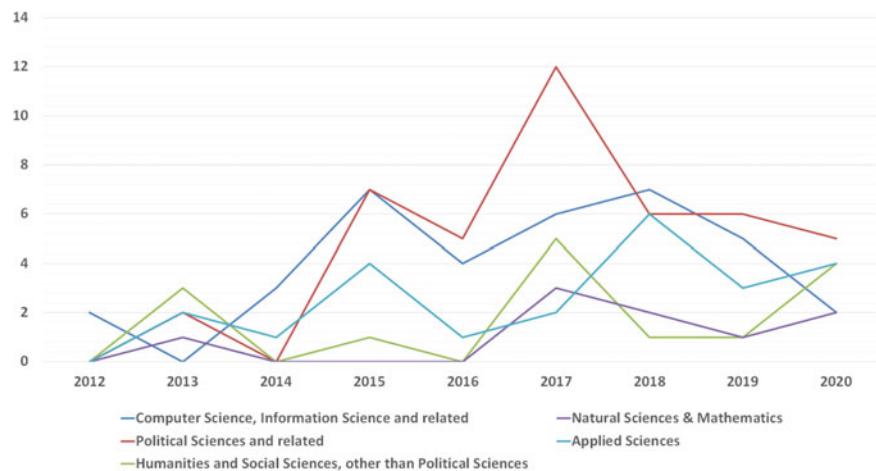


Fig. 3.3 The number of studies on collective intelligence in policymaking published yearly within the Research Area Groups (Olszowski et al., 2021)

intelligence in policymaking. In recent years, there has been a discernible decline in studies within Computer Science, making room for an uptick in diverse social research endeavors. The annual fluctuations within these grouped research areas are illustrated in Fig. 3.3. The annual distribution of studies within the RAGs suggests a declining interest in the technological facets of projects (as seen in the Computer Science and related groups). Conversely, there is a rising focus on the application of these projects within various administrative sectors and the broader public sphere, as indicated by the Political Sciences and related groups.

After these initial analyses, we turned our attention to the crux of our study: identifying methods for studying collective intelligence in policymaking. Our team of experts with experience and academic backgrounds in policymaking and information technologies evaluated the content of the articles. Our preliminary analysis compiled lists of the most frequently appearing concepts in article titles, abstracts, original keywords, and KeyWords Plus. In the subsequent qualitative research phase, we aimed to extract the methods and strategies from the analyzed texts. This phase was based on the grounded theory approach. We employed this approach to distill theoretical value from the selected studies, categorizing and presenting key concepts and articulating these concepts into distinct categories. The analysis was undertaken in stages characteristic of the grounded theory method: open coding, axial coding, and selective coding. This entailed generating high-level abstraction categories from sets of concepts, identifying key processes, and capturing the primary research outcomes described in the articles. This also involved refining or redefining their established definitions. Emerging categories were continually refined as we analyzed subsequent texts from our sample. This iterative process continued until we achieved theoretical saturation. Our analysis revealed one to five methods or strategies in each reviewed

text, leading us to compile a comprehensive list of 15 methods. The results are presented in Table 3.2.

It's immediately apparent that, despite CI's growing popularity in policymaking, its study lacks consistent rigor. The methods used to study theoretical models, successful case studies, domains for project implementation, expected results, and factors influencing CI vary widely based on the researcher's background, education, and analytical context. Furthermore, different research traditions sometimes use alternative terms to describe identical phenomena, such as the interchangeable use of *crowdsourcing* and *CI*. The limited adoption of the 'Analysis of the impact of AI algorithms' approach is indeed surprising. It appears that CI research remains largely distinct from AI research. Even though the synergy of AI and CI has been underscored as a pivotal research subject in recent times—as evidenced by the report *Identifying Citizens Needs by Combining AI and CI* (Verhulst et al., 2019) and the contributions of Mulgan (2019) and Malone (2018)—this call to action appears

Table 3.2 Methods and strategies for studying CI in policymaking identified in the reviewed literature (Olszowski et al., 2021)

No	Method of studying CI	Description	Literature	No. of assigned articles
1	Analysis of organizational structure/design	The studies were conducted from an organizational perspective. The analysis covers the structures that facilitate the coordination and implementation of rules, resources, technologies, stakeholders, and particular tasks in specific projects or initiatives of open policymaking. These studies present the systems for accomplishing and connecting the activities that occur within examined work organizations, enabling the emergence of CI	Prpić et al. (2015), Taeihagh (2017), Kerzner (2017)	31
2	Analysis of created values	The studies aim to answer the question: What kind of valuable results were produced in the analyzed projects? The analysis of outputs, given that they are more valuable than the inputs, is conducted. For example, epistemic, democratic, and economic values in increasing the quality of public service provision can be analyzed	Aitamurto and Chen (2017), Iacuzzil et al. (2020)	25

(continued)

Table 3.2 (continued)

No	Method of studying CI	Description	Literature	No. of assigned articles
3	Analysis of e-participation process	The aim of these studies is to analyze the factors that influence technologically supported participation, or e-participation, which can be defined as “participation in societal democratic and consultative processes mediated by information and communication technologies, primarily the internet” (Saebø et al., 2008) or as “the use of information technologies to engage in discourse among citizens and between citizens and elected or appointed officials over public policy issues” (Iacuzzil et al., 2020)	Saebø et al. (2008), White (2007)	17
4	Analysis of participants' behavior	These studies aim to answer the question: What sort of various activities were performed by the users of the examined policymaking platforms and initiatives, what types of operations did they engage in, and how was it related to their individual characteristics?	Aitamurto et al. (2017)	16
5	Analysis of the collaboration model	What forms of collaboration between governmental and non-governmental entities occur in the area under study and what factors influence its facilitation were investigated	Mergel (2015)	16
6	Analysis of participants' motivations	These studies focus on understanding the participants' motivations to engage in open policymaking projects	Wijnhoven et al. (2015), Aitamurto et al. (2017)	11
7	Analysis of the communication model	Analyses of the communication processes, information exchange, establishing of information channels between public and civic entities, extraction of valuable information, and the mutual understanding of the content provided are performed	Guth and Brabham (2017), Iandoli et al. (2018)	9

(continued)

Table 3.2 (continued)

No	Method of studying CI	Description	Literature	No. of assigned articles
8	Analysis of innovation process	Investigating the critical aspects of the innovation process in the studied policymaking projects and initiatives. These studies aim to answer the following questions: what influences innovation capacity, how to stimulate pro-innovative behavior, and/or what the potential positive and negative impacts of the outcomes of the innovation processes are	Leitner et al. (2016), Almirall et al. (2014), Mergel (2015)	9
9	Analysis of decision-making process	These studies aim to answer the question: How are collective intelligent policy decisions made, and what affects the quality of the decision-making process? The analysis of processes, sub-processes, and data related to collective decision-making is conducted	Epp (2017), Bose et al. (2017)	8
10	Analysis of the impact on policymaking	These studies present the observed impact of the analyzed projects on creating public policies and assess the significance of this impact and factors that influenced it	Chen and Aitamurto (2019)	7
11	Categorization of the implemented projects	Typologies of various governmental or non-governmental initiatives and projects engaging citizens in policymaking in a model that considers the emergence of collective intelligence are presented	Linders (2012)	5
12	State-of-the-art review	The state of research and practices are presented in these studies in a cross-sectional manner. The studies focus on collecting, categorizing, and situating the previously published research and practices in the field, coming from multiple disciplines	Prpić et al. (2015), Chen and Aitamurto (2019)	4

(continued)

Table 3.2 (continued)

No	Method of studying CI	Description	Literature	No. of assigned articles
13	Analysis of platform usability	These studies aim to understand the structure of policy-oriented websites, their functions, interfaces, and contents; simplicity of use; site navigation; and the ability of users to control their activities	Hogan et al. (2017), Flavián et al. (2006)	4
14	Analysis of the impact of AI algorithms	The aim of these studies is to analyze the possibilities of using AI techniques in CI processes occurring in policymaking initiatives and the possible effects of their operation	Fernández-Martínez et al. (2018)	3
15	Analysis of organizational learning	These studies focus on organizational learning as the process of creating, retaining, and transferring knowledge within a policymaking organization when an organization improves over time as it gains experience	Lenart-Gansiniec and Sułkowski (2018)	1

unheeded. The potential synergistic effect of AI and CI in policymaking seems to remain more a topic of conversation than of systematic investigation.

Our review revealed that within the entire sample, the most commonly used approaches to studying collective intelligence in policymaking were the ‘Analysis of the organizational structure’ and the ‘Analysis of the created values.’ Furthermore, when examining the two primary research areas of these studies, we found that the former method is more characteristic of political science. At the same time, the latter is prevalent in computer science. The frequent employment of the ‘Analysis of the created values’ approach reflects trends in commercial projects, where tangible results of collective efforts are often emphasized. The statistical analysis has shown significant relationships between the research methods. Notably, a negative relationship exists between ‘Analysis of created values’ and ‘Analysis of collaboration model.’ This can be explained by the fact that projects mainly oriented toward generating new values are studied in the context of the existing governance framework. In contrast, studies exploring new models of intersectoral collaboration between public and private entities—especially when a project’s scope goes beyond a single organization’s structure—necessitate a different approach.

We also compared the percentage of method usage in specific research areas to that in the reviewed literature. This helped us determine which methods and strategies were more or less prevalent in the respective research domains. In computer

science, the methods ‘Analysis of created values’ and ‘Analysis of e-participation process’ were used more frequently than in the entire sample. However, the ‘Analysis of organizational structure’ and ‘Analysis of impact on policymaking’ were underrepresented. Conversely, in political sciences, there was a heightened interest in ‘Analysis of organizational structure’ and ‘Analysis of collaboration model,’ while ‘Analysis of decision-making’ was less popular. Notably, both in this field and in the whole sample, the ‘Analysis of impact on policymaking’ method was infrequently used, which is surprising. An assessment of the most influential articles, based on usage and citation statistics from the Web of Science, highlighted their unique characteristics. Among these, the strategy focused on innovation analysis emerged as a particularly favored research approach.

In the final stage of our analysis, we delved into topics of particular interest within the context of CI in policymaking from the reviewed literature. To select these specific topics, we relied on monographs published after 1990, reflecting sources from the era when the Internet began its widespread adoption. This led us to identify seven key topics: Citizenship, Communities, Consensus, Deliberation, Diversity, Local Governance and Urban Development, and Open Data. We then scoured our literature database for keywords related to each topic. Based on the presence of these keywords, we segmented the studies into topic-oriented subgroups. The representation of these chosen topics within the analyzed research studies is detailed in Table 3.3.

Table 3.3 Saturation of the analyzed research studies with selected topics of interest (Olszowski et al., 2021)

Concept	Number of studies where the concept appeared	References in monographic publications
Citizenship	47	Landemore (2012b), Landemore (2020), Noveck et al. (2019), Ryan et al. (2020)
Local governance & urban development	30	Lévy (1997), Noveck et al. (2019, 2020), Ryan et al. (2020)
Communities	14	Lévy (1997), Landemore (2020), Ryan et al. (2020)
Deliberation	9	Landemore (2012b, 2020), Ryan et al. (2020), Noveck et al. (2020), Aitamurto (2014), Noveck (2015)
Open data	7	Ryan et al. (2020), Noveck et al. (2020)
Diversity	5	Lévy (1997), Landemore (2012b), Noveck et al. (2019), Aitamurto (2014)
Consensus	5	Landemore (2012b, 2020), Aitamurto (2014)

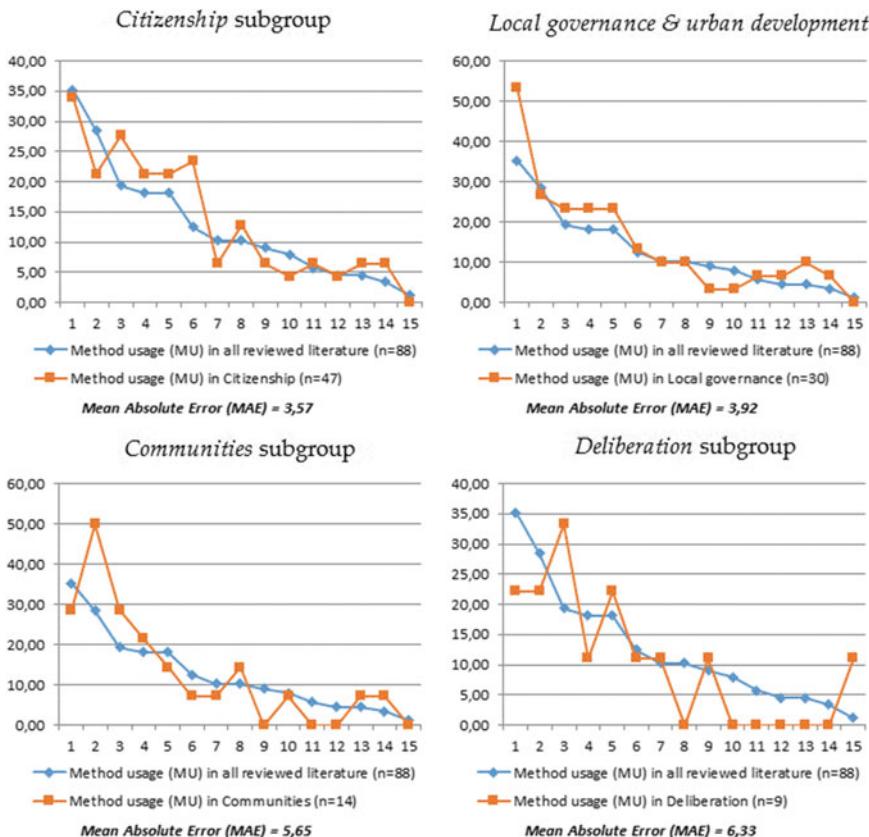


Fig. 3.4 The percentage of method usage (MU) in particular subgroups compared to the percentage of MU in all the reviewed studies. The assignment of particular methods and strategies to the labels numbered from 1 to 15 as described in Table 2.2 (Olszowski et al., 2021)

The four most popular topic-oriented subgroups were analyzed in terms of the methods and strategies adopted in the research. The aim was to verify to what extent the reviewed literature relates to the examined topics and what research methods were used in the studies focused on these topics. To achieve this, we calculated the percentage of method usage (MU) within specific subgroups and compared it to the MU percentage across all reviewed studies. We also computed the Mean Absolute Error (MAE) for each subgroup to gauge how significantly they deviate from the overall sample.

The results of the analysis are visualized in Fig. 3.4. This analysis showed that the most popular concept in our sample was citizenship, and studies using this term were very often associated with analyzing participants' motivations or broadly addressing the e-participation process. The prominence of references to citizenship issues suggests that researchers are less interested in a mechanistic approach to

enhancing policymaking. Instead, they are exploring a novel paradigm concerning the role of citizens in online initiatives. This aligns with the shifts in citizen-state relationships postulated by Noveck (2015), who envisions transforming the government from a sole authoritative problem solver to a mediator that invites citizens to collaboratively identify optimal solutions.

Another widely referenced concept in our sample was *local governance*. Over 34% of the studies we reviewed mentioned this topic. Cities, alongside local and interest-based communities, seem to have emerged as primary venues for implementing CI projects in the public domain. In the context of cities, the organizational structure of projects is often the central focus of research. Meanwhile, for communities, the emphasis lies on the values they produce. However, we observed that deeply theoretical topics such as *diversity* or *consensus* remain less popular in the analyzed studies. This could be attributed to their limited relevance to the dominant themes of citizenship and local governance.

3.2 Collective Cognitive Processes: Foundations for an Analytical Method

3.2.1 Theoretical Basis for the Evaluation Framework

Our review yielded many intriguing insights, yet it did not bring us closer to selecting a method for evaluating collective intelligence in policymaking. Although some existing research touched upon cognitive processes such as decision-making and, in rare cases, organizational learning, we did not encounter any study that holistically examined collective thought processes within policy projects. In our opinion, such studies should treat online projects not just as mere online platforms but as systems of group cognition. They should also allow for the examination of group cognitive processes within a singular project, breaking down the mechanisms of group intelligence into its essential components while avoiding the isolation of these processes. However, evidence suggested that a purely functional approach was predominant, whether approached from a social sciences or computer sciences perspective. I have thus concluded that we must devise a satisfactory method on our own.

I decided to develop a new evaluation framework tailored specifically for assessing CI in policymaking, drawing from pertinent theoretical and empirical knowledge. Following the recommendations of R. Torraco (2005), a method of creating such a framework requires outlining a preliminary theory, testing it, and setting it against the background of domain knowledge. Therefore, I needed an integrated theory of group cognition, which gathers the most important achievements of cognitive psychology thus far. Then, my goal was to supplement it with the results of empirical research in online collective intelligence and the general context of policymaking.

When choosing a theoretical basis for further development, I was influenced by its potential compatibility and complementarity with prior research on online CI.

After a qualitative analysis of approaches proposed within cognitive psychology, I initially settled on the Triarchic Theory of Intelligence proposed by Sternberg (1985), grounded in the foundational concepts of Cattell (1963) and Neisser (1967).

Let me briefly discuss the basics of these theories. Cognitivism believes that a significant portion of human behavior can be comprehended by analyzing individuals' thought processes. It opposes the idea that mental processes should be excluded from psychological studies due to their intangible nature. Cognitive psychology, as a field of psychology, focuses on how humans perceive their environment, i.e., how knowledge about the surroundings is formed and subsequently utilized in behavior. Knowledge is presented as structures (called *mental representations*), and the mechanisms of its formation are seen as processes (called *cognitive processes*). The overall issue is viewed as creating and transforming structures through processes. Hence, it can be asserted that cognitive psychology involves studying cognitive structures and processes.

Sternberg's definition of the cognitive processes says that "much of human behavior can be understood in terms of how people think—the study of how people mentally represent and process information. As such, it includes within its domain mental abilities such as perception, learning, memory, reasoning, problem-solving, and decision-making" (Sternberg & Sternberg, 2012). These main cognitive processes are derived from one of the earliest definitions presented in the first textbook on cognitive psychology published by Neisser (1967), which states that cognition is "those processes by which the sensory input is transformed, reduced, elaborated, stored, recovered, and used." Therefore, according to Neisser, the phenomenon of thinking refers to many partially overlapping notions: "increasingly sophisticated memory, attention, planning, problem-solving, search, communication, decision making." These processes concern fluid or crystallized intelligence, as described by Cattell (1963) and further developed by Horn (1969). Fluid intelligence appears when a human solves problems using deduction, logic, and inference, depending only minimally on prior learning. The *thinking* mechanism can also be based on learned procedures and gathered (crystallized) knowledge, precipitated out of experience resulting from the prior application of fluid ability.

Sternberg further developed these concepts into a theory of intelligence composed of three main parts (hence the name Triarchic Theory): (1) analytical intelligence, referring to one's ability to learn and store new information and concerned with higher-order thinking, planning, monitoring, and decision-making; (2) creative intelligence, based on the ability to deal with new and unusual tasks or situations effectively, and the skill to automate thinking and processing over time through learning; and (3) practical intelligence, taking into account the ability to adapt to existing environments and situations, the ability to modify or create new environments, and the skill to select a different environment when adaptation and shaping are not possible. Triarchic Theory emphasizes intelligence's adaptive and broad nature, going beyond traditional IQ measurements to include creativity and practical problem-solving abilities in everyday life. According to this concept, individual types of intelligence do not function independently of each other—as Gardner (1993) believes—but create a mutually complementary whole that translates into effective human functioning.

Psychological theories of cognitive processes have evolved significantly in recent years, branching off in various intriguing directions. Notably, one of the most interesting developments—in the context of CI research—has been the expansion of theory beyond the realms of individual cognition, embracing the concepts of group cognition and *thinking systems*. Initiated through empirical research in the Virtual Math Teams (VMT) project by Gerry Stahl in 2006, the concept has since been further refined and expanded by several scholars, i.e., Goldstone and Theiner in 2017. They introduced a perspective named *multiple interacting levels of cognitive systems* (MILCS), infusing cognitive psychology with a viewpoint centered on group cognition. In this perspective, various types of thinking systems collectively attain functionalities crucial to cognitive processes traditionally associated with individuals (Goldstone & Theiner, 2017).

In examining processes related to intelligence, this analysis considers not just one cognitive system but various systems operating at intersecting levels. These encompass thinking individuals and cooperative small groups driven by a common purpose and, finally, large crowds where cognitive processes occur less consciously and more chaotically. By analyzing similar processes occurring at different levels, we can gain insights beyond individual cognition to understand processes such as perception, attention, memory, problem-solving, and decision-making. We can also explore interactions among individuals within cohesive groups that share a strong identity and delve into dispersed crowds. Finally, we examine systems of interaction with artifacts and cultural resources. Understanding group cognition involves determining how systems at multiple levels can mutually reinforce, rather than undermine, each other. Embracing diverse levels of analysis allows for a better understanding of the mechanisms that enable systems to engage in flexible and adaptive behavior (Goldstone & Theiner, 2017).

I found the Triarchic Theory combined with MILCS to be a very promising approach when it comes to planned research on CI in policymaking. In my view, the necessary complement to these theoretical foundations was a distinction made by Daniel Andler, a philosopher of science at the Sorbonne University, regarding differences in collective thinking processes based on the operational level of a cognitive system. In the communities or organizations based on personal relationships, he describes *thick* cognitive processes, being a feature of communities united in a “common spirit [...], constantly coordinated and recalibrating their mutual expectations, blending into a ‘we’ capable of we-thoughts, we-intentions, we-actions,” culturally transmitting their beliefs and practices, and exhibiting similarity to the individual cognitive systems (Andler, 2012). In contrast, Andler describes the *thin* cognitive processes occurring in large moblike structures, “in which individual agents, far from deliberating or exchanging information and arguments, simply provide their conclusions, which are then fed to some aggregating algorithm or mechanism.” This possibility of getting “intelligent” results coming from incoherent crowds was also highlighted by Surowiecki (2005), who evoked many examples of surprising “order coming out of chaos.” This issue is consistent with the phenomenon of cognitive diversity that we discussed earlier. In most cases, however, we are not dealing with a pure form of any of the *thick* or *thin* models but with a combination of both: “Classical

thick procedures can be simplified to limit interactions to a more or less restricted set, and symmetrically thin procedures can be enriched to allow ‘thicker’ information to be transmitted and aggregated” (Andler, 2012).

As Goldstone and Theiner claim, individuals are often strongly driven to join robust groups, and “cognitively resourceful people will frequently form groups that effectively employ cognitive systems at higher levels than the individuals” (Goldstone & Theiner, 2017). This approach elucidates that intelligent individuals can act as the fundamental components of intelligent groups. However, the MILCS theory reveals an aspect long emphasized by CI researchers yet often overshadowed in cognitive psychology. It emphasizes that a group embodies a unique mode of thinking, not merely a simplistic aggregation of the intelligence of its constituent individuals. This leads to the rejection of the *zero-sum* perspective on group cognition as presented by Huebner (2013) and aligns fully with Woolley’s conclusion discussed in the previous chapter, asserting that a group’s intelligence is not a mere sum of individual intelligences. So, Goldstone and Theiner claim that a collective system displays an intelligence level distinct from the sum of its components. Of course, a group’s collective intelligence is expected to surpass its individual members’ aggregate intelligence. However, it is conceivable that suboptimal coordination or ineffective communication could result in CI that is, in fact, inferior. According to the MILCS theory, the level of intelligence is influenced by factors such as diffusion, preferential attachment, competitive specialization, positive feedback, negative feedback, small-world network, scale-free network, back propagation, reinforcement learning, and multi-level deep learning (Goldstone & Theiner, 2017), underscoring the irreducibility of group mentality. MILCS collects evidence that the *zero-sum* perspective loses its plausibility when we consider the dynamics of large-scale social systems that contain multiple people and an environment that facilitates their interactions (Stahl, 2006, 2009; Theiner & Goldstone, 2010). In such systems, cognition arises from the interactions between all parts, especially when its structure is based on a decentralized network.

The findings of cognitive psychologists align with the perspective presented by the researchers of CI. Consider Landemore’s opinion (2012a): “Distributed intelligence refers to cognitive processes that are stretched across individuals and their different components (mind, body, activity), as well as culturally organized settings, including groups and institutions. Distributed cognition thus refers to an emergent phenomenon that cannot be traced simply to individual minds, but rather to the interaction between those minds and between them and their constructed environment (which extends their cognitive capacities).” Further, Heylighen (1999) claims that such a system can contain several forms of interacting units: not only humans but also communication nodes, artificial agents, sub-groups, and technical infrastructure. Similarly, Malone (2018) argues that cognitive reality is a meta-system where numerous independent cognitive systems operate, often consisting of sub-systems.

According to the assumptions of cognitive theories and parallel CI research, collective cognitive processes are influenced by manageable factors. These include the group’s structure, the organization of its work, and the technical means at its

disposal (Heylighen et al., 2004). Therefore, as Malone (2018) argues, we can influence the development of the intelligence of collectives by creating optimal conditions for their operation, influencing their composition, principles, communication methods, work cycle, self-organization, and various technical measures to improve their operation. Information in these kinds of networks is propagated selectively, depending on its utility, novelty, coherence, simplicity, expressivity, authority, etc. (Heylighen et al., 2004).

In conclusion, the extension of cognitive psychology from individual to group behavior unveils distinctive patterns essential for the operation of the *thinking* system. Both individuals and groups can exhibit intelligence through decisions that integrate diverse information sources, ensuring accuracy and timeliness. Certain cognitive systems can achieve specific functionalities relevant to cognitive processes involved in perception, attention, memory, problem-solving, and decision-making (Goldstone & Theiner, 2017). These systems offer mechanisms for executing cognitive functions, such as discerning optimal choices, establishing network coherence, and developing specialized units through competitive interactions. If they work intelligently, they demonstrate flexibility in devising strategies to navigate challenges, embodying cognitive mechanisms analogous to those found in individual human cognition.

3.2.2 Developing the Evaluation Framework

The next stage was the development of a preliminary theoretical basis with the use of the grounded theory approach, following the directions of Charmaz (2006) and Corbin and Strauss (2008). I aimed to validate and refine the framework through real-life applications, integrating findings from empirical research on CI and policymaking studies.

First, I conducted several interviews to evaluate the initial framework and data collection protocol for their credibility and confirmability. Then, I selected a project for testing. The general assumption was to analyze a representative e-participation initiative, which at the same time was being affected by significant difficulties whose nature could be revealed by analyzing collective cognitive processes. An online participatory budgeting project—Civic Budget of the City of Kraków—was selected for an in-depth study based on semi-structured interviews. Data collection and analysis were conducted simultaneously; therefore, the theory was grounded in the participants' experiences. In the final stage, a framework outline was populated with the collected experiences and the extended references to relevant literature to enhance the research's internal validity and theoretical quality.

The conducted qualitative study has led to promising results. First, it was possible to identify the operation of the presumed cognitive processes in the studied project. Initially, it was not clear to me how many of these processes could be distinguished and what their scope should be. The categorization of the interviews led to the identification of the emerging concepts and sub-concepts that I decided to integrate into

the four main cognitive processes: collective sensing (perception), problem-solving, decision-making, and collective memorization (including learning and gaining feedback). Thanks to the conducted data analysis, it was possible to populate the theoretical basis with collected experiences and concepts. The last stage of the planned study was to integrate the proposed concepts and connections, develop a narrative of the theory, further populate the framework with references to the relevant literature on collective intelligence, and enhance the internal validity and quality of the framework. The results of this work were illustrated in the map of integrative findings presenting the processes, sub-processes, and data related to collective intelligent decision-making. The map of integrated findings is shown in Fig. 3.5.

3.2.2.1 The Cognitive Processes of Collective Intelligence

What happens within these successive processes, and how can they form a coherent whole? The *collective sensing* process usually begins with gathering information through open-ended questions (Clark et al., 2017), perceiving social data in the groups and communities (Chikersal et al., 2017), using a selective approach to information globally assessed as relevant, as well as “combining data from perception with top-down theories and models” (Goldstone & Theiner, 2017). This process is a particular feature of the systems performing the *thin* cognitive processes: “intelligence emerges from an assemblage of non-intelligent components, (...) by interconnecting in specific ways agents that are deliberately used as ‘sensor’ of certain narrow segments of the world (...). The ‘sensing’ they accomplish can be complex and involve inner workings that are those of a fully intelligent creature” (Andler, 2012). Gathering a large amount of diverse information and extracting from it what is important is what Goeff Mulgan calls achieving the right granularity of information (2018). The data we collect should be neither too general because then we would drown in them, nor too detailed, because then we would not know the context of the studied phenomenon. Recently, the topic of information noise reduction has been insightfully addressed by Kahneman et al. (2022). They explored how and why diminishing noise contributes to better results in various domains of individual human judgment, ranging from medicine to law. At the same time, however, the variety of information can be a valuable factor in groupthink. In many situations, some degree of unpredictability or *noise* is not just beneficial but crucial. Frequently, the phase filled with noise comes before the exploitation stage, during which such noise turns counterproductive. The importance of extracting meaningful data from a plethora of information to achieve integrated results, as measured by the *phi* metric, was also suggested by Engel and Malone (2018).

One of the key concepts affecting this process is, as we remember, cognitive diversity, i.e., the participation of people with different cognitive styles, ways of thinking, experiences, perspectives, and abilities. As we know, some researchers make a simple translation of cognitive diversity into general diversity, emphasizing “differences in demographic, educational, and cultural backgrounds” (Hong & Page, 2004). Others are more cautious in their interpretation of diversity’s scope and impact.

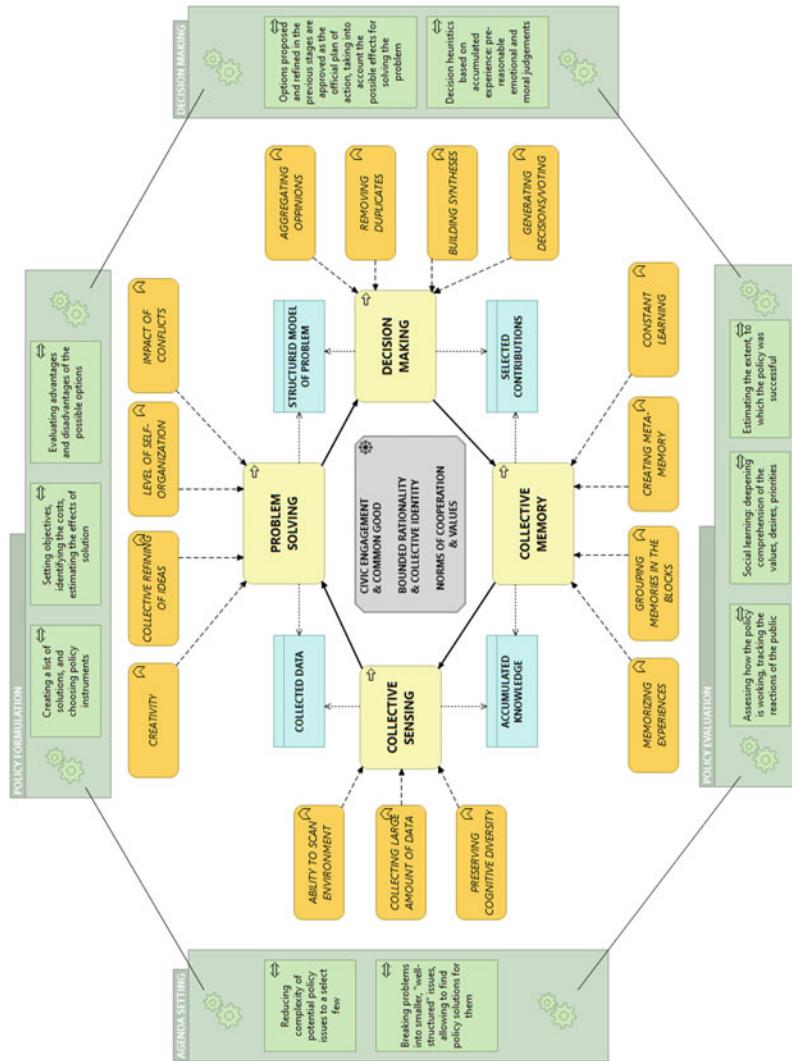


Fig. 3.5 Cognitive processes (yellow), sub-processes and impacts (orange), and data (blue) influencing the emergence of collective intelligence in policymaking. At the center of the diagram, the 'drivers' coordinating individual processes into a coherent whole are presented (gray). Around the outside, we see the phases of policymaking within which CI processes are carried out (in green)

Considering Aggarwal's warning highlighted in the previous chapter, it's crucial to consider that cognitive diversity might exhibit a curvilinear inverted (U-shaped) relationship with CI. This suggests excessive diversity could, in certain scenarios, decrease CI levels (Aggarwal et al., 2019). Moreover, the volume of data gathered is significant as it determines the potential to derive valuable insights from the raw dataset (Engel & Malone, 2018).

The *problem-solving* process is mainly related to issues of creativity and the collective refining of ideas to reach a level of maturity in the debate. Creativity is often attributed to individual actors within a system, characterized by their ability to generate novel ideas and solutions (Malone et al., 2010). Yu et al. (2012) emphasize that collective creativity extends beyond the mere generation of innovative output; it encompasses the creative nature of the tasks undertaken within a collaborative process. In such scenarios, collective creativity involves a series of non-routine tasks, each contributing to the emergence of novel ideas and solutions: it is "about both the creativity of the output from a collective process and the creativity of the tasks performed as part of this process" (Yu et al., 2012). One of the key challenges in facilitating collective creativity lies in designing conditions that foster successful creative collaborations (Bigham et al., 2015). This requires a nuanced understanding of how diverse individuals can effectively work together to leverage their unique perspectives and expertise. To achieve this, organizations and teams often need to implement strategies that encourage open communication, diverse idea generation, and a supportive environment that values contributions from all participants (Paulus & Nijstad, 2003). Under the right conditions, groups can harness the cognitive diversity of their members to achieve more creative outcomes that surpass what any single individual could achieve alone.

Many CI researchers consider refining ideas and building epistemic bases for intelligent decisions by deliberating groups as necessary elements of problem-solving. However, only a few approach it in a systematic manner. Cappella et al. (2002) propose the concept of Argument repertoire (AR), derived from the conceptual and empirical work by Kuhn (1991) on reasoning in daily life, directly eliciting and assessing counterarguments. Enabling participants in a debate to build an extensive repertoire of arguments, meaning the number of arguments related to a given issue, is considered a factor influencing the quality of reasoning during the debate and the maturity of its outcomes (Capella et al., 2017). The generation of arguments and counterarguments requires people to envision conditions that would prove or falsify their explanations, encouraging them to consider various perspectives and different solutions to the problem. These refined ideas and arguments can be organized into topically-structured tree-like frameworks in an ideal deliberative situation. In CI systems that support extensive argumentation, these structures consist of a hierarchy of questions to be answered, possible answers for these questions, and the supporting arguments for each answer (Klein, 2011a, b). This kind of organization can facilitate clarity and transparency in the deliberation process and help participants systematically explore and evaluate various facets of the problem or decision at hand.

One striking feature of collective problem-solving is the question of the group's decentralization and self-organization. In this kind of cognitive system, no single part appears as a computationally privileged locus of control (Goldstone & Theiner, 2017). If self-organization appears, it is manifested through flat relationships centered on achieving a common goal, not imposing formal structures but leaving space for the non-formal leaders to appear (Skaržauskienė, 2015). As we remember from Chap. 1, the dispersion and decentralization of knowledge, as well as the process of self-organization in society, were analyzed by Hayek. Analogous to his considerations on economics, we can see that the synergy between self-organization and decentralization in CI is evident in systems of crowdsourcing and citizen science (Silvertown, 2009). Decentralization allows participants to choose tasks that align with their skills and interests, while self-organization ensures that contributions coalesce into meaningful outcomes. As we remember, according to Surowiecki (2004), decentralization means that no single member dominates the decision, and participants are free to express their opinions independently, so the information is not concentrated in one place but distributed among diverse participants. For instance, open-source software development, exemplified by projects like Linux (Raymond, 1999), relies on decentralized problem-solving, allowing self-organized contributors from around the world to collaborate without hierarchical oversight.

The possibility of internal conflict within collective intelligence (CI) systems is a critical aspect that warrants attention, as it can significantly impact the outcomes of group problem-solving processes (Bigham et al., 2015). While conflict within CI systems may seem counterproductive, it is important to recognize that, in many cases, criticality and conflict can trigger the emergence of novel ideas and solutions within groups (Vincenzo et al., 2017). This dynamic reflects the complexity of CI processes, where the clash of diverse perspectives can lead to innovative problem-solving and decision-making. The issue of internal conflict is closely intertwined with the concept of group identity (Collins et al., 2020). Group identity is pivotal in determining how individuals interact within CI systems. When participants strongly identify with a particular group, it can lead to polarization, wherein participants self-organize into sub-groups with distinct viewpoints and ideas and have limited cross-fertilization across these groups (Becker et al., 2019). Various factors can exacerbate the phenomenon of polarization in CI. Social media tools, for instance, often contribute to the formation of filter bubbles, where individuals are exposed primarily to information and viewpoints that align with their beliefs (Klein, 2011a, b). Additionally, people's natural tendency toward homophily, or the preference for associating with others who share similar characteristics or opinions, can further reinforce polarization within CI systems.

The decision-making process usually begins with an aggregation of the generated opinions and removing repetitions, which translates to building problem syntheses (Goldstone & Theiner, 2017). One critical step in this process is the aggregation of opinions generated by participants. The system's products' epistemic quality directly depends on the implemented aggregation mechanism (Origgi, 2012). Various aggregation methods, such as voting, ranking, or weighted aggregation, can be used depending on the specific context (Hardas & Purvis, 2012).

Ideally, the possible solutions for a problem created in previous stages of work would go through systematic, well-reasoned evaluations, where “all key arguments for and against an idea have been identified” (Klein, 2011a, b). This kind of evaluation can be combined with *idea filtering*, whose role is to eliminate, as much as possible, incomplete or unsatisfactory ideas “so the community has the best possible alternatives at hand when it makes its final decisions” (Klein, 2021). Collective evaluation serves as an effective means of correcting individual biases within the decision-making process (Bonabeau, 2009). It allows the collective to recognize the presence of biases and work towards their elimination when they hinder the decision’s quality. The integration of differentiated information is another crucial phase. As we mentioned, the possible metric Engel and Malone (2018) proposed for this aim is called *integrated information* or *phi*. This metric aids in combining diverse insights and perspectives from participants, resulting in a more comprehensive decision-making process.

Finally, the natural culmination of the decision-making process is a selection of one or more results to execute. Participants’ input is assembled to generate a decision that holds for the entire group. As described by Malone et al. (2010), “In some instances, this decision determines [...] the subset of contributed items that will be included in the final output. In other instances [...], the decision relates to generating a common rank-ordering of the contributed items. In yet other instances, such as prediction markets, the decision relates to aggregating individual inputs to form a publicly visible estimate of a quantity.”

A simple majority vote is the most basic method for reaching collective decisions and avoiding conflicts: “adding all the votes together determines the relative preferences of the different alternatives for actions” (Heylighen, 1999). The famous Condorcet Jury Theorem (CJT), first expressed by the Marquis de Condorcet in 1785, is a frequent starting point for considering the majority decisions made by a group, as it calculates the relative probability of a given group of individuals arriving at a correct decision. The CJT asks, “Under what conditions does the aggregate, independent judgment of a set of individuals yield a better outcome than the most competent person alone or any random person alone?” (Capella et al., 2017). The CJT posits that the straightforward aggregation of individual judgments can often succeed. Thus, in this basic understanding, collectives often have the potential to exhibit intelligence under various scenarios, although not universally. Historically, the CJT has been examined primarily using relatively uncomplicated tasks rather than more intricate or nuanced ones, such as policy preferences or ethical decisions. Consequently, a pertinent question that arises is whether it’s feasible to assess tasks with no clear correct or incorrect outcome, i.e., those we frequently encounter in policymaking. Let’s consider collective decisions based on simply averaging the group’s judgments. Research shows that it can be effective only under certain conditions: a single, simple, distinct decision to be made (i.e., A or B), no obvious bias affects everyone in the group, and there is a rational deliberation phase preceding the decision (e.g., building an argument repertoire during the problem-solving phase).

Instead of simple majority voting or averaging the opinions of all possible participants in situations involving more complex issues, research on CI emphasizes the

potential use of decision-making approaches tailored to the problem's specific nature. Considered methods include *preference voting*, where "individuals can distribute their voting power over different alternatives, in proportion to their preference functions" (Heylighen, 1999), and *weighted voting* (Hardas & Purvis, 2012). As proposed by Heylighen, *preference voting* is a method somewhat similar to ranked-choice voting. Here, instead of casting a single vote for one alternative, participants can distribute their voting power over various alternatives based on their preferences. This allows voters to express various preferences rather than make a binary choice. Voters can show not only which option they prefer but also how much they prefer it relative to other options (Heylighen, 1999). Thanks to this, voting can be more inclusive of minority opinions and allow for a more nuanced expression of preferences, possibly leading to more broadly acceptable decisions. On the other hand, *weighted voting*, as proposed by Hardas and Purvis (2012), is based on a Bayesian mechanism for weighting the actual vote given by a user to compute an effective vote which incorporates the history of voting and also the crowd's thoughts about the value of the voter's contributions. The weights allocated to each voter will influence the overall outcome, and the voter with a higher weight will have a higher impact on the final decision. Weighted voting might be more suitable for specialized or technical decisions where expertise is crucial. It ensures that the opinions of those with more relevant knowledge or experience have a greater influence on the decision. At the same time, preference voting might be more applicable in scenarios where a broad range of preferences needs to be captured.

Another approach for regulating the decision-making process considered by CI researchers involves establishing criteria to determine eligibility for participation in the decision-making group. As mentioned in the previous chapter, Beth Noveck (2015) proposed the creation of field expert groups within which decisions could be made in their area of specialization. Using aggregate expert opinion with domain-specific factual knowledge can be a criterion for judging the success of open-ended problems. Moreover, such experts offer a greater chance of meeting the criteria set forward for success in the CJT theory, namely, that their judgments have a probability above chance of being correct so that—in the absence of other serious biases—their aggregate opinion would be more likely to constitute a wise outcome than would be the case for a random set of individuals or a randomly selected expert.

The *collective memory* process is the final stage of the presented model. It also signals a return to the beginning of the cycle. The importance of this process is essential because, if it works well, it forms the basis for increasingly efficient *thinking* in a given community gained by successive iterations of the activities described above. As Hugo Mercier and Helene Landemore (2012a, 2012b) suggested, for effective policymaking, it is crucial to analyze the actual success of a selected policy by evaluating its outcomes, taking as a reference point in the analysis the epistemic base developed earlier, i.e., diverse, competing solutions and the arguments for and against them. Also, Daniel Andler (2012) argued that just as individual intelligence is dependent on memory, collective intelligence is dependent on collective memory. However, as I mentioned in the previous chapter, within the extant research on CI, the role of memory has been relatively marginalized in favor of a focus on problem-solving or

decision-making processes. Indeed, Heylighen (1999) highlighted the importance of memory in collective thinking, yet his research remained largely theoretical. He acknowledged that accumulated knowledge improves the work of the entire cognitive system by constantly learning, gaining feedback, and making predictions. The feedback mechanism “interacts with collective memory in a non-linear way: a trail leading to a good source will be reinforced through a positive feedback loop, while a trail leading to an empty source will spontaneously decay” (Heylighen, 1999). Heylighen was aware that in the memorization process, the collected experiences and other results of problem-solving and decision-making should be grouped, archived, and stored as a shared data resource. Still, he did not present any proposals on how to do this in practice.

The concept of collective memory in the social sciences, however, has a long-standing history dating back to the early twentieth century. It draws heavily from the works of Emile Durkheim (1971/1912) on *collective consciousness* and the works of his apprentice, Maurice Halbwachs. In 1925, Halbwachs coined the term *mémoire collective*, arguing that all remembering relies on the dynamics of social groups. An individual’s social interactions with the members of his or her group determine how one remembers experiences from the past. Therefore, an individual memory is entirely socially mediated, “It is in society that people acquire their memories,” suggests Halbwachs (1992, p. 38). For Halbwachs, people remember and forget only within a social framework whose most important elements are families, social classes, and religious communities (Russell, 2006). Notably, in describing the social process of remembering, Halbwachs situated memory in the present rather than in the past and described it as a social process that looks forward instead of backward (Bachleitner, 2022). It is the process through which “the past is not preserved but is reconstructed based on the present” (Halbwachs, 1992, p. 40), therefore both individual and collective memory share the property that they are “not static but dynamic—not fixed but ever-changing. A memory is essentially re-remembered (i.e., reconsolidated) each time it is recalled” (Anastasio, 2022). Another significant observation is that truly meaningful collective memory extends beyond a mere compilation of facts or abstract information, known in psychology as semantic memory; it encompasses lived experiences, also referred to as episodic memory. This memory is specific to a particular group and forms a component of this group’s identity. Halbwachs posited that a group becomes conscious of its identity through an awareness of its accumulated experience (Russell, 2006).

For years, empirical research on memory conducted by psychologists has been dominated by the *mechanistic* approach, treating the mind as a kind of computer performing logical reasoning processes, part of which is encoding information, storage, and retrieval. In this model, the results of thought processes are saved in memory, which is a dedicated part of the mind. Encoding is understood as the transformation of information into a format that can be stored in the brain, as suggested by the multi-store model of Atkinson and Shiffrin (1968). Retrieval, the final stage, involves recalling or accessing the stored information, a process dependent on various cues and contexts, as Tulving’s Encoding Specificity Principle (1983) describes. Yet, this perspective does not adequately account for the dynamic nature of collective

memory, which, as described by Halbwachs, involves the constant reconsolidation of memory upon each retrieval. Moreover, it does not align with CI's self-organizing and decentralized characteristics. In collectives consisting of people and computers, memory is not stored in one place but scattered and diverse.

However, help in understanding memory processes comes from another field of science. Advancements in modern neuroscience, especially those concerning the neocortex, contribute significantly to our understanding of the role of memory in the process of intelligent thinking. Neuroscientists claim that we do not learn a *list of facts* about the world but a model reflecting the structure of the world and the relations between its component parts (Hawkins, 2021, p. 4). In 1978, neurophysiologist Vernon Mountcastle observed that all parts of the neocortex operate through a common principle with the cortical column—a group of vertically arranged neurons—being the unit of computation (Mountcastle, 1978). This idea was further developed by Jeff Hawkins (2021), who argued that the majority of connections between neocortex columns are not hierarchical. Hawkins emphasizes the role of the neocortex in recognizing and memorizing spatial and temporal patterns. Neocortex columns act independently to perform similar tasks, each of them being an independent thinking unit storing differentiated models of reality, and their large number and diverse perspectives have a decisive impact on thinking processes. Therefore, the knowledge accumulated in the brain is distributed (Hawkins, 2021, p. 97). Nothing we know is stored in one place; information is spread in thousands of columns, collecting thousands of complementary though diverse models in a non-hierarchical system (Hawkins, 2021, p. 110). When we memorize something, the connections between neurons are strengthened, and when we forget something, the connections are weakened, as proposed by Donald Hebb (1949). Contemporary neuroscience claims this process is constant throughout a human's life: new synapses form and old ones disappear, each occurring independently in different brain parts (Hawkins, 2021, p. 38).

According to Hawkins' theory, our brain does not permanently store all details coming from observations but *reference frames*, which act as a sort of coordinate map for different ideas. These frames are not exact descriptions of instances but their invariant representations. The process of intelligent reasoning involves finding a good *reference frame* to organize facts or observations (Hawkins, 2021, p. 88). This process does not entail a logical analysis of every situation we encounter; rather, it is about finding the most similar model in memory and using it to predict the consequences of the situation. Memory is fundamental to the entire thought process, and intelligence lies in optimally aligning a model to the specific context in which we find ourselves. This leads Hawkins to conclude that intelligent thinking is determined not only by the storage of a multitude of patterns in memory but, more crucially, by the efficient retrieval of and reference to experiences that are most pertinent to current challenges. To find an answer to a problem, neurons *compete* to provide the best-fitting models, conducting a kind of voting to reach a consensus (Hawkins, 2021, p. 110). Consequently, our brains utilize stored memories to continuously make predictions about the environment. Correct predictions result in *understanding* the situation and making the right decision; mispredictions result in confusion and prompt us to update

the model. As described by Hawkins, a memory-prediction framework presents the whole process of thinking as remembering sequences of events and their nested relationships and making predictions based on those memories, which is much more efficient than a logical analysis of each encountered problem or challenge.

It is noteworthy that Hawkins' insights regarding distributed memory complement Halbwachs' findings about the dynamic nature of collective memory. Just as the former regarded intelligence as a process based primarily on predictions based on the best-fitting models extracted from memory, the latter saw collective memory as the use of the accumulated experiences of the community to jointly prepare for future challenges.

3.2.2.2 Metacognition: Understanding Cognitive Processes in the Context of Public Policies

In line with our assumptions outlined above, in the proposed framework, the processes related to collective intelligence should be studied within the broader context of public policies. Therefore, we will consider CI processes, keeping in mind their common features with the phases of policy analysis with which we are familiar from Chap. 1. Figure 3.5 illustrates the four most relevant phases of policymaking, depicted as blocks surrounding the processes and subprocesses of collective intelligence. As one can observe, these phases have been deliberately selected and arranged such that policymaking processes bearing resemblance to cognitive processes are situated in proximity. Nevertheless, it is crucial to acknowledge that a direct correlation between these two levels of analysis is not a standard, both in the schematic representation and in practical application. This is because, according to the post-positivist paradigm I have adopted, phases of policy analysis are not linear. Furthermore, implementing actions corresponding to individual stages of public policy creation is seldom (if ever) achievable within a single project. Similarities do indeed exist, as reflected in the corresponding names and scopes of the individual phases at both the policy and CI levels. Nonetheless, when evaluating specific online projects, as intended by the framework proposed, it is impractical to consider all phases of policymaking. This is because many policy matters are addressed within the scope of long-term political or administrative decisions, and online projects have a shorter lifespan.

But let's draw our attention to several issues where these two approaches undoubtedly converge. Firstly, consider the action in the 'Agenda Setting' phase directed at reducing the complexity of potential policy issues to a select few and structuring problems such that they should be divided into smaller, *well-structured* issues, facilitating the discovery of policy solutions. This aligns with Mulgan's emphasis on achieving the right data granularity ensuring data is not distorted by informational noise while maintaining data source diversity and participants' cognitive style diversity in the debate. Furthermore, during the policy formulation phase, considerations like choosing suitable policy instruments for listed solutions, setting objectives, identifying costs, estimating solution effects, and defining and evaluating the pros and cons of possible options find their practical expression in creative collaboration. This

collaboration centers on establishing epistemic grounds for opinions or constructing an argument repertoire. Moreover, since the decision-making element present in the policy cycle involves evaluating the possible effects of adopting a chosen option and developing decision-making heuristics based on accumulated experience, it's crucial to associate this policymaking phase with CI subprocesses. These subprocesses enable the building of conclusion syntheses and opinion aggregation, as well as those belonging to collective memory, where experiences are compiled and stored in a communal data resource. It's also clear that continuous learning, a characteristic of flexible knowledge-based systems adapting to changing conditions, paired with gaining feedback and making predictions, complements the idea of social learning. This is a pivotal component of the post-positivist approach to policymaking.

When considering the integration of systemic thinking about collective intelligence with the nuances of public policies, the central part of the diagram becomes paramount. Here, the *drivers* coordinating the individual processes into a cohesive system are showcased (illustrated on the diagram against a gray background). When we inquire about the coordination of CI processes, we are essentially asking: What helps coordinate the aforementioned processes so they work coherently? It's important to realize that collective debates can sometimes lead to more foolishness than wisdom, resulting in choices that are less effective and solutions that are not optimized. Exploring the elements that can enhance or diminish group intelligence in policymaking requires a holistic view of collective thinking processes and an analysis of what coordinates them.

Using the language of cognitive psychology, we would call the process of coordinating collective intelligence *metacognition*. According to cognitivists, metacognition is the capacity to monitor and regulate our cognitive abilities (Fleming & Frith, 2014; Lund & Russell, 2022). Metacognition refers to the awareness and understanding of one's mental processes, including the ability to control these processes and understand how one thinks, learns, memorizes, and solves problems. Metacognition allows us to judge our reasoning and is often assessed with self-reported confidence measures, such as the confidence that one's task performance is correct (Fleming & Lau, 2014). Cognitive scientists have also acknowledged metacognition on a group level, which is focused on a group's shared metacognitive abilities and knowledge, as a perspective often taken within education research. As an example, the investigation of how groups of individuals coordinate group-level skills and knowledge to work together (Lund & Russell, 2022).

What would constitute metacognition in collective online policymaking? First and foremost, it is crucial to acknowledge that we are dealing with collective, not individual, metacognition. Therefore, for its existence, we need the appropriate quality of the community. So, I believe that the first essential condition for achieving coherence in the cognitive processes is the quality of participation in the public sphere, which Hannah Arendt associated with active citizenship (Arendt, 1998). Empowered, informed, and engaged citizens are fundamental assets of intelligent collective policymaking, as active citizens can ensure effective debate, which is a crucial element of CI. Citizens engaging in the public sphere can cooperate for the common good. Michael Sandel asserts that pursuing the common good requires active civic engagement

and deliberation wherein citizens engage in political discourse and decision-making processes (Sandel, 2012). For this to occur, it is essential to maintain the independence of citizens participating in the debate. Civic engagement is esteemed not merely for its potential to foster agreement but also for its capacity to empower each citizen with individual agency. Citizens' self-organization, invoking Hayek once more, requires the perseveration of the agency of individuals—not their passivity in the face of the collective—and shared rationality distributed across a large group of cooperating individuals rather than the authoritative rationality of a central decision-making body (Hayek, 1945). The independence and active engagement of citizens influence the capacity of the collective to collaborate effectively while attaining shared policy goals without succumbing to any authoritative center of political power.

In the post-positivist approach to public policies, bounded rationality is a commonly accepted guideline to coordinate policymaking, recognizing the limitations of human cognition and the practical constraints that affect decision processes. As we remember from Chap. 1, this concept challenges the notion of the *perfectly rational* decision-maker who is fully informed and capable of processing all available information to make an optimal decision. Bounded rationality suggests that decision-makers often rely on heuristics, or mental shortcuts, to drive decisions, and a decision maker's identity influences the heuristics. For instance, a policymaker's past experiences and knowledge can lead them to rely on specific heuristics that align with their identity.

Since our model involves a collective engaged in policymaking, our primary interest lies in the decision-making heuristics of the group. I believe the constraints of bounded rationality may apply to groups just as they do to individuals. Similar to individuals, collectives can use mental shortcuts and established rules to make decisions. Collective bounded rationality is not merely the sum of individual rationalities but is shaped by social interactions, discussions, and negotiations that are influenced by shared values and collective identity.

As discussed in the previous chapters, people naturally tend to form groups that become an integral part of their identity. Therefore, a significant portion of individuals' cognitive efforts is directed toward establishing reliable communication pathways, forging enduring social structures, and ensuring the continuity of the groups they associate with. A community's identity is deeply interwoven with the coherence of individual interests of the people forming a group. A strong relation between the well-being of individuals and the community, encompassing its infrastructure (area, institutions, resources), results in group members being more inclined to devote their time and commitment to it, as they benefit from the achievements of the entire community.

If the development of the community is synonymous with the well-being of its members, individuals will see value in the common good. For such a relationship between individual and community to exist, individuals should invest their resources, such as skills, material means, activity, or free time, in the development of the community and have the opportunity to participate in its development. For this to be possible, individuals must be sufficiently independent and resourced so that such an arrangement is mutually beneficial. They should possess the agency and independence to

act in their environment, guaranteeing enough time to engage in everyday affairs. This brings to mind an example from the previous chapter, namely the involvement of programmers in developing open-source projects that they can later use for their purposes. Community members should possess the potential to contribute, allowing them to play the most appropriate roles for them in public life. When individuals feel that their contributions are meaningful and acknowledged, it enhances their sense of self and belonging, which benefits both personal and collective well-being. A positive relationship between the individual and the community can lead to a more cohesive, empowered, and healthy community environment.

Cognitive flexibility in individuals often leads to the creation of communicative infrastructures, norms, and conventions that robustly support these connections. Consequently, the groups with a strong identity tend to establish norm systems that self-organize the interactions in the group, like monitoring bodies (similar to debate moderators) whose role is to look for potential violations of the rules. Finally, sanctions (like debate bans or warnings) are used to punish people found guilty of violating the rules. Goldstone and Theiner emphasize that.

Most people have an inherent drive to connect to others, and individual cognitive flexibility often finds ways to support those connections in rich and resilient ways by creating communicative infrastructures, norms, and conventions. When a person strongly identifies with a group, they experience pride when their group does well and anger when their group is attacked, even when the person is not directly involved. (...) People's strong emotions relating to group pride, guilt, and anger incite them to act in ways that will reinforce, repair, and protect the group, respectively (Goldstone & Theiner, 2017).

To find out what coordinates the cognitive processes of a collective, we must, therefore, consider its collective identity and commonly held values, which can be preserved and reinforced through collective memory. When faced with complex decisions and a lack of comprehensive information, collectives working on policy problems might rely on the prevalent values of their community to guide their choices. The influence of shared values on group collaboration, its achievements, increasing adaptive abilities, and intergenerational learning was already noted on a macroscale by Charles Darwin in his renowned work *The Descent of Man* (1998, original work published 1871). Though the idea of group selection based on shared values was somewhat contested in scientific circles, mainly by evolutionary biologists like George C. Williams (1966) and Richard Dawkins (1976), recent scholarship, including works by Tomasello (2014) and Haidt (2013), has again emphasized the link between collaboration, group identity, cohesion, and shared values. Consequently, an important query to pose when evaluating collective intelligence in the proposed model pertains to the existence of norms governing collaboration and users' sense of belonging to the community.

Policies that align with the values and identity of the collective are more likely to be accepted and considered legitimate by its members. This alignment is crucial for the successful implementation and efficacy of policies. Collective identity contains generally accepted norms and principles of cooperation that the participants follow, determining which behaviors are acceptable and unacceptable. As we remember, the positivist separation of facts from values and focusing only on facts has been

largely dismissed by contemporary policy studies. However, what remains an open question is whether online collectives reveal common identity and shared values and, if so, what impact they have on collective thinking. Another intriguing issue is the emergence of particular identities, constructed somewhat in contrast to the dominant identity and values of the community. Does the identity and goals of minorities disrupt the community, leading it to seek particular benefits instead of the common good?

The metacognitive analysis presented above aims to highlight key factors that may influence collective intelligence in the context of online policymaking projects. My objective is to explore whether and how the described cognitive processes manifest in these projects. It is crucial to discern whether the cognitive processes are complementary or if there are issues with any of these processes. With this objective, the next section of this chapter will examine a selection of case studies, endeavoring to evaluate them using the presented framework.

3.3 Evaluation of the Case Studies

Our research team based at the AGH University of Krakow selected the case studies examined in this section. The team, including myself, comprises of two social scientists and two IT specialists. Our primary criterion for selection was the recognizability of the cases. We sought examples with an international profile, those that have been the subject of previous academic discourse, and at the same time, those that were not ephemeral, i.e., they possessed enduring relevance in the field of policymaking. This consideration was crucial, as numerous intriguing projects were contingent on specific political organizations (e.g., DemocracyOS created by Partido de la Red in Argentina, or Liquid Feedback linked to the German Pirate Party) or grant funding and subsequently ceased after the funding concluded or the viability of the organization decreased. Therefore, we focused on projects that have demonstrated longevity and secured stable funding, allowing us to hope that they will function for many years to come. Additionally, we aimed to showcase a diversity of approaches to CI. We wanted to show how different approaches are implemented regarding the scale of participation (large and small collectives) and the breadth of user opportunities.

The case of the first presented project, i.e., *Budżet Obywatelski Miasta Krakowa* (eng. *Civic Budget of the City of Kraków*, BOMK), is slightly different. As I mentioned earlier, the BOMK project was the first study prepared to test and refine our method of evaluating CI processes in the public sphere. For the initial assessment, we intentionally chose a case that was not very distinctive but rather typical of many municipal participatory budgets. We selected it partly because of its geographical proximity, given that our university is in Kraków, but that was not the only reason. We wanted to verify how our framework performs in relation to a typical project that falls under the category of urban initiatives associated with local communities. What sets this project apart is not its scope but rather the particular timeframe during which the discussed edition took place, notably the period of the Covid-19 pandemic

and the accompanying strict lockdown. Of particular interest was the impact of transitioning all civic activities related to participatory budgeting to exclusively online platforms for the first time and how this shift influenced the project's characteristics.

3.3.1 *Civic Budget of the City of Kraków: The Pandemic Edition*

3.3.1.1 *Introduction and Background*

Over the past decade, the participatory budgeting (PB) model has become a standard in many countries. In this model, citizens of a given municipality gain influence over its policies related to public investments, social activities, and cultural undertakings. Participatory budgeting can be described as a policymaking process during which the citizens discuss and negotiate the issue of the *distribution of public funds* (Wampler, 2007). Participants in this process are involved in submitting proposals, deliberating, and deciding how public resources should be allocated for investment projects and social and cultural initiatives.

While Kraków distinguishes itself from other Polish cities due to its prominence as a tourist hub and an outsourcing business center, the BOMK project aligns with numerous participatory budgeting endeavors across Central-Eastern Europe. Comparative analysis indicates that the Polish model of PB parallels those in neighboring countries, with one notable distinction: From their inception, Polish projects predominantly embraced online voting, leading to heightened participation compared to other countries in the region (Džinić et al. 2016).

This online inclination was further amplified in 2020. Amid the challenges presented by the global coronavirus pandemic, almost all related activities transitioned online. This extended beyond just voting to encompass brainstorming and evaluation processes, cementing the project's status as a fully e-participatory initiative. Projects executed during the coronavirus pandemic provide a compelling insight into how civic engagement can be sustained or even amplified amidst societal challenges. Given this perspective, I believe that, despite the time that has elapsed since the pandemic, the BOMK project remains a valuable case study (Fig. 3.6 and Table 3.4).

3.3.1.2 *Project History and Key Features*

First introduced in 2013 as a pilot in four districts, the project *Budżet Obywatelski Miasta Krakowa* (eng. *Civic Budget of the City of Kraków*, BOMK) has been operating as a city-wide initiative since 2014. Each iteration of the project spans a calendar year. The 2020 edition, which we were evaluating, stood out for several reasons. For one, the participatory budget reached a record-breaking sum of over 7 million EUR, approximately 2.5 times greater than the 2018 allocation. Moreover, influenced by the

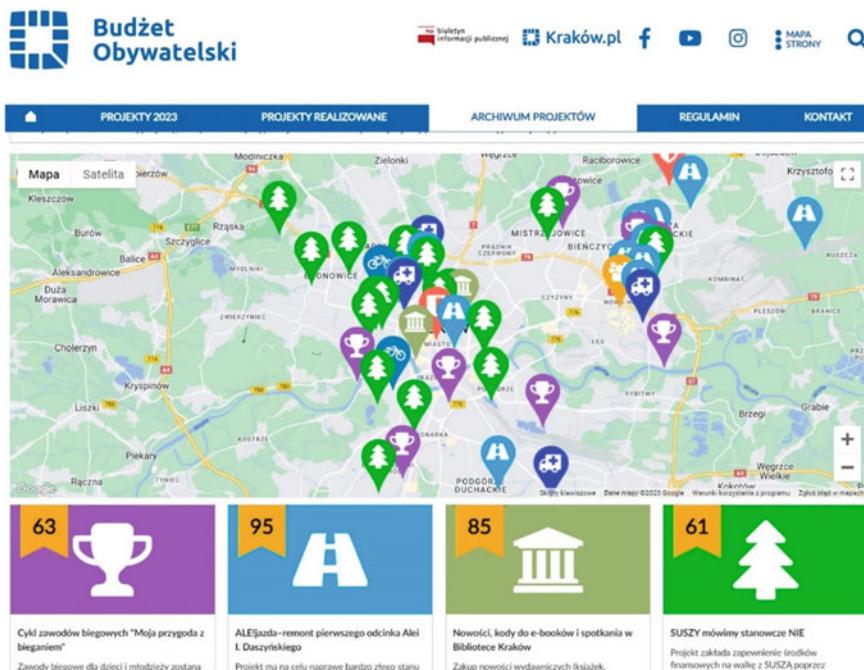


Fig. 3.6 The BOMK website presents the archive of projects from 2020. *Source* <https://budżet.krakow.pl>

Table 3.4 Basic information about the BOMK project

Project name	Budżet Obywatelski Miasta Krakowa (eng. Civic Budget of the City of Kraków, BOMK)
Owner	The City of Krakow; Department of Social Policy and Health, Participation and Dialogue Section
Web address	https://budżet.krakow.pl
Location	Kraków, Poland
Years in operation	2013–present
Implementation level	Local

coronavirus pandemic and decisions made by local authorities, the entire project transitioned to online platforms. Online engagement extended beyond mere voting—an improvement from previous years—to include data collection, evaluation, and deliberations concerning local needs. The BOMK allows citizen-proposed tasks to be executed at district and city-wide levels. Each annual cycle of the BOMK concludes with an official survey gauging citizen opinion. According to this survey, among the various facets of the BOMK's organization, respondents particularly praised the user-friendliness of the voting platform (86% positive feedback), the ease of submitting

task proposals (86% positive feedback), the simplicity of the voting process (78% positive feedback), and the quality of communication with the task verifiers (78% positive feedback). The majority of negative opinions concerned the processing of filed protests (36%), the promotion of voting in BOMK (22%), and the process of task proposal verification (14%) (Barometr Krakowski 2018). The Civic Budget Council coordinated the project activities, comprised of representatives of urban activists, NGOs, informal groups, applicants, representatives of the Kraków City Council, and representatives of district councils and administrations.

As a project, BOMK consists of several stages. The first stage is the so-called *Educational Stage*, which aims to encourage citizens to participate in consultations and prepare their proposals for public actions. It includes discussions and debates, social dialogue, cooperation between applicants, and project-writing marathons. This stage is concluded by submitting *project sheets*—preliminary proposals for actions—by the citizens. Compared to the previous years, there was a significant change in 2020, which resulted both from the coronavirus pandemic and the general trend of digitalization, i.e., virtually all processes pertaining to debate, dialogue, and cooperation between the citizens were implemented online. Preliminary actions involved the creation of *Idea Bank BO 2020* (*Bank Pomysłów BO 2020*), a collection of opinions about the citizens' most crucial needs which were obtained electronically and presented visually in the form of a publicly available e-map that grouped ideas according to locations and categories (Security, Education, Road Infrastructure, Cycling Infrastructure, Culture, Sport and Sport Infrastructure, Society, Health, Greenery and Environmental Protection). The e-map of the citizens' needs, in conjunction with an electronic database of city-owned land (land on which public projects can be implemented), was supposed to constitute an introduction to a debate and inspire project creators. Moreover, 34 inhabitants volunteered as debate coordinators (*Local Participation Ambassadors*), whose task was to clarify any queries pertaining to the formalities, as well as to moderate the debate in a given field or social environment.

The information campaign in 2020 was mostly conducted on Facebook. To support debate, feedback, interaction, and the creation of new ideas, the city officially authorized the launch of 18 Facebook discussion groups called “Joint Projects” (1 group per city district). Their purpose was to assist in the grassroots preparation of projects by local communities. These groups numbered 795 members in total, ranging from 12 to 134 per district, who were quite active—over 600 ideas and around 1,000 comments were published across all groups (Raport... 2020). Interestingly, to join the “Joint Projects” group, one must now adhere to established communication standards and behavioral norms. What happened next during the group work? Preliminary proposals for action were commented on and users created a network of cooperation, trying to find common topics. There were also many parallel groups created by grassroots on social media, where district inhabitants and local communities gathered to form new ideas (e.g., *Nowohucianie*, or the inhabitants of Nowa Huta district, with over 29 thousand members).

In the year 2020, a total of 988 proposals were submitted to BOMK (a 4.5% increase compared to the previous year), most of which belonged to the following

categories: Greenery and Environmental Protection (25%), Infrastructure (23%) and Sport & Sport Infrastructure (16%). Of the submitted task proposals, 807 concerned districts and 181 concerned the whole city (Wydział Polityki Społecznej i Zdrowia Miasta Krakowa 2020). It is striking, however, that the creative potential of the applicants from the previous editions of BOMK remains untapped. The vast majority (84%) of applicants submit their task proposals only once instead of trying to do the same in the following editions of BOMK (Raport... 2019). Many of them are daunted by the process and results of the verification stage and by limited communication with the city authorities. It would seem that the participants in previous editions should be natural partners in organizing the process of drafting new task proposals and engaging other citizens. In reality, after the first experience, the majority resigned from participating in the project. Currently, no activities aim to create lasting relations with these applicants, so the knowledge and experience they obtain are not taken advantage of in subsequent editions.

The remaining stages of the yearly project cycle include the verification of task proposals by experts to make sure that the proposals adhere to the BOMK Rules and Regulations, the filing and processing of protests against experts' opinions, voting by citizens—indicating investment priorities, and the announcement of the tasks to be implemented and their evaluation. The expert verification process constitutes a considerable obstacle: in the 2020 edition, 42% of task proposals were dismissed, mainly due to land ownership problems (Wydział Polityki Społecznej i Zdrowia Miasta Krakowa 2020). Additionally, it is not always possible to avoid project duplications. One of the applicants described the following situation in the survey:

there were several identical projects, e.g., *abreuvoirs* or green bus/tram stops. Two projects called for 'green bus stops,' which resulted in votes for one initiative being spread across two projects, so even if the majority of people want green bus stops, the fact that the votes are spread across two nearly identical projects might mean that none of them gets a sufficient number of votes to win (Raport... 2019).

The voting process was implemented mostly through an online platform. Of the total of 44,800 votes cast in 2020, 42,281 were valid. As a result, 195 tasks were selected for implementation, of which 185 were district-wide and ten were city-wide (Raport... 2020). A vast majority of voters find the online voting platform useful and transparent (91%) (Raport... 2019). However, some people expressed various complaints about the platform. The most frequently mentioned problems concerned the search engine's performance and project filtering:

The search bar on the civic budget's website didn't work for me. I tried to search for a project by its name and it never came up, but when I typed the same name into Google, it did. [...] The feature I missed was an option to filter projects by specific fields. When we submit projects, we have to define a field, which becomes the project's symbol and we should be able to filter projects by that field in the app, because some people are interested in certain fields more than others, so this type of filter would be justifiable (Raport... 2020).

High turnover among people involved in the project was also observed. So far, over two-thirds of BOMK voters participated in just one edition (68%; 113 thousand people). Only 1,660 people have voted in all editions (1%) (Raport... 2019). This

high percentage of one-time voters points to a lack of success in creating long-lasting relations and to the fact that the experience and engagement obtained by the participants do not benefit subsequent editions of the project.

3.3.1.3 Collective Cognitive Processes in the Project

Our foundational approach to evaluating CI processes pivoted around the framework detailed in Sect. 3.2.2, applying it for in-depth individual interviews. From a pool of professionals deeply involved with the BOMK's execution, we engaged 12 distinct voices: district councilors, leaders of civic movements, activists, and subject-matter experts. Yet, understanding the merit of diverse data sources, we further triangulated our findings through an ancillary method: an analysis of formal reviews of past Civic Budget editions and other pertinent local policy dossiers. Chief among these documents were the Evaluation Reports for the 2020, 2019, and 2018 BOMK Editions (Raport... 2018, Raport... 2019, Raport... 2020), a preliminary 2020 BOMK statistics overview courtesy of the city authorities (Wydział Polityki Społecznej i Zdrowia Miasta Krakowa 2020), and the *Barometr Krakowski*, a local public sentiment survey commissioned by the city (Barometr... 2018). By interweaving these varied sources, we aimed for a thorough triangulation of data, seeking to bolster the credibility and accuracy of our conclusions, all while diminishing the risk of interpretative missteps (Yin, 2009). The table below summarizes the statements of interview participants. Direct quotes from interviewees' statements are in quotation marks, with an indication of which of the 12 interview participants they refer to (coded as INT 1-INT 12) (Table 3.5).

3.3.1.4 Metacognition: Understanding Cognitive Processes in the BOMK Project

The overarching findings from the conducted analysis confirm that each of the group's cognitive processes was identified, and their performance could be evaluated, which in itself was a success and proof of the potential usefulness of our framework. But the matter becomes more complex when it comes to what binds individual processes together. Given the broad scope of the project, we must remember that many participants limit their involvement to voting without participating in consultations or collaborative project development. However, it seems that when people engage collaboratively on projects, especially in grassroots groups like the aforementioned *Nowohucianie*, it's often due to a strong emotional connection to their district. This fosters commitment and a sense of shared purpose.

Certainly, an important factor binding the operation of cognitive processes in the BOMK together was the sense of responsibility for one's surroundings (the city) and the participants' sense of agency. Participants pointed out that, especially in the socially difficult period of the coronavirus pandemic, it is important to provide a safe and comfortable space in the city and take care of its maintenance. At the same time, their involvement was helped by the awareness that the decisions they made

Table 3.5 Cognitive processes in the BOMK project

Cognitive process	How does it work in the project?
Collective sensing	<ul style="list-style-type: none"> In the initial phase of the BOMK project, an unexpectedly positive public response was observed. Newly created positions of Local Participation Ambassadors positively impacted the collection of data on local needs. “Collecting scattered opinions of the inhabitants using the Internet and translating them into proposals for action went surprisingly well (...). The response was large and overwhelmingly positive” [INT 3]. The role of Ambassadors was particularly important due to the increased isolation of people in the realities of the coronavirus pandemic Most of the respondents claimed that collecting data allowed them to raise the most important issues: “I assumed that 2–3 times a week I would post a report on my work and questions for residents on Facebook groups, and I gained feedback there. Thanks to the ongoing exchange of opinions, the residents asked directly for help in specific matters” [INT 7] When collecting opinions, previously existing Internet communities of residents enjoyed greater popularity than the official BOMK discussion groups or the interactive Idea Bank. The amount of information that has been identified on the needs and expectations of residents has been assessed as fully sufficient to prepare adequate proposals for actions: “(...) The spectrum of topics is quite wide, so there is plenty to choose from” [INT 4] The large amount of data generated had not only a direct impact on the proposed initiatives but also an indirect impact on city officials, who, having noticed an increased interest in a given topic, included it in urban programs and strategies. The impact on shaping the reality was twofold: both direct, through voting in BOMK, and indirect, through shaping the policy of the city authorities: “Even projects that are not implemented or carried out on a micro-scale draw attention to the fact that there are some topics important to the townspeople that the city should deal with. Officials can see that there is a problem, and it needs to be tackled on a larger scale” [INT 2] Most respondents believe that identifying needs made it possible to capture the diverse expectations of the city residents, and none of the topics was over- or underrepresented. However, there were some opinions that cognitive diversity was not fully preserved because of the existence of strong activist circles (e.g., from schools, sports clubs, pro-ecological NGOs) that have been lobbying for certain solutions for years, against which spontaneous grassroots initiatives are disadvantaged: “what is noticeable is the preponderance of pressure groups at the expense of the normies” [INT 4]. Others drew attention to: “abuse, which I call the pathology of the city of retirees. Retirees have more time than other social groups; they firmly believe that their problems are the most important, and they want to impose this point of view on others” [INT 9]. Sometimes, strong particular identities are also visible, e.g., groups of fans of competitive football clubs. On the other hand, in the opinion of some survey participants, the diversity is not fully preserved: people aged 35–50 are underrepresented because they are busy with their professional activities, which limits their time for social involvement. This is problematic because this group is perceived as the city’s most important taxpayers. However, the opinions collected confirm that these problems did not negatively impact the expected level of diversification of the topics discussed or the diversity of BOMK participants

(continued)

Table 3.5 (continued)

Cognitive process	How does it work in the project?
	<ul style="list-style-type: none"> The ethnic and cultural diversity of Krakow, which has been increasing in recent years as a result of immigration, helps to better identify the existing opportunities for action: “Thanks to the fact that we currently have residents who are Ukrainians, Belarusians, etc., our discussions contain examples and experiences from their previous places of residence” [INT 2]
Problem solving	<ul style="list-style-type: none"> The focus on Internet communication forced by the coronavirus pandemic resulted in a significant increase in the creative activity of people who had not been associated with any official groups before (such as social clubs, NGOs, etc.). Creativity and self-organization in establishing <i>joint</i> projects were supported by Local Participation Ambassadors, who moderated the debate, as well as by the nature of communication in social media, enabling simple and quick responses (e.g., short comments and “likes”) The response to the submitted ideas, even if small, was a great motivation for their initiators, encouraging them to work on improving the projects. However, virtually every idea that succeeded required a strong leader who refined its content, verified its feasibility (especially the question of land ownership), and was able to make realistic cost assessments. Therefore, the ideal model of cooperation was an interaction between a strong leader who ran the project and the local community who motivated the leader with their support Presenting ideas via the Internet also prompted participants to actually consider arguments instead of supporting the idea just because of neighborhood relations: “In the past, it was possible for people to run around the block and gather support from the neighbors. Now there are fewer neighborhood alliances, and we have more realistic projects” [INT 10] Although there are more and more grassroots projects, these projects still constitute a minority compared to the proposals of institutionally entrenched organizations or single individuals. It is estimated that approx. 35% of the prepared ideas were created due to deliberations within the official “Joint Projects” Facebook groups and other online communities. However, according to some respondents, the strong identity of the circles that have been lobbying for specific solutions for years (e.g., the development of bicycle paths) might have an advantage over <i>joint</i> projects; these groups have extensive experience in their narrow specializations, which results in the creation of more realistic solutions The study participants paid special attention to problems associated with cooperation with the city hall departments responsible for verifying the legal aspects of the proposals, i.e., land ownership and other regulations affecting their feasibility. The online application provided by the city authorities, which allowed for quick access to information on land ownership and the possibility of using it for investments, was found to be a positive factor in supporting group cooperation. The respondents reported that the problem was not the lack of creativity of the city’s inhabitants—this factor was rated highly—but rather the fears of city officials about the implementation of difficult and unconventional projects. Instead, they tend to discourage unusual projects to avoid extra work

(continued)

Table 3.5 (continued)

Cognitive process	How does it work in the project?
	<ul style="list-style-type: none"> The survey participants saw conflicts and disputes that arose in the course of the debates as obstacles weakening the cohesion of the online communities and discouraging activity. Therefore, they believe that the ultimate goal of group work is supposed to be consensus. Nevertheless, our respondents do not think that conflicts dominated the debates. Rather, they see a greater obstacle to collective action in the excessive passivity of group members: "Up to the stage of formal evaluation of ideas conducted by the officials, cooperation and enthusiasm was high. Only in the course of interactions with the officials did disputes arise—most often over the location of the project. However, sometimes disputes help involve people in refining the idea because they make Ambassadors choose new locations together with the initiators, encourage more people to join the community around the idea, etc. (...) When there are signs of resistance from officials, social leaders often self-mobilize to keep their proposal from being scratched" [INT 6] Refining the ideas in interaction with the group, aggregating opinions, and removing duplicates is often done during negotiations. This is where municipal officials come in, whose task is to invite the authors of duplicated proposals to meetings (usually online, e.g., using MS Teams): "We had two very similar projects submitted independently to establish a multicultural library. The applicants were invited to negotiate, they modified their assumptions a bit, and they got along, resulting in a single, even better project. It also drew the attention of officials responsible for libraries to the need for establishing such an institution, which is why foreign-language books have already started to be collected" [INT 7]. However, "there are still situations where similar projects take each other's votes away. Still, the applicants do not wish to cooperate even after meetings" [INT 9]. Sometimes, this is due to the fact that the people who have been lobbying for certain solutions for years are afraid of losing their position if a similar idea is created independently. One of the participants provided an example: "I proposed a project involving the planting of 100 new trees in each district of the city. However, the environmentalists' association felt threatened by the fact that someone else was acting in 'their' field, so they spoke out critically about my idea in the media. Still, my project was successfully implemented" [INT 12]
Decision making	<ul style="list-style-type: none"> Survey participants gave a positive opinion about the voting procedure, saying that the existing online platform is transparent and easy to use. Certain reservations were raised about the fact that voting resembles shopping in an online store (including "adding to basket" pre-selected solutions). Still, the vast majority of survey participants appreciate the simplicity and intuitiveness of the web application. The current system of preferential voting, allowing the selection of three city-wide and three district projects and prioritizing them from the most to the least supported, is also highly rated: "Preferential voting is a good solution. This prompts us to look at the various tasks and problems of the city in a broader sense, not just at the one project that the participants intended to support. Second-choice projects often win, which deepens reflection on what is generally worth doing in the city" [INT 1]

(continued)

Table 3.5 (continued)

Cognitive process	How does it work in the project?
	<ul style="list-style-type: none"> The pre-selection phase and the preliminary feasibility evaluation of the project, managed by the Civic Budget Council, were reported as more controversial. This sub-process is intended to aggregate the generated opinions, remove duplicated ones, and help build syntheses. The accusation often made against the evaluators was that they used discretionary evaluation criteria. The criteria for considering one initiative to be unfeasible and another to be feasible were not entirely clear. One respondent described a case where two similar projects were treated differently, differing only in the title. Such behavior is seen as an example of superficial assessment. However, the respondents pointed out that the negative feasibility evaluation mainly concerned the proposals submitted by participants with little experience, most often participating in the BOMK project for the first time. By contrast, several cases were described where a proposal contested by the city officials for feasibility reasons was corrected and could be put to the vote Additional factors that potentially influenced decision-making included biases that affected voting and the lack of extrapolation of the effects of the implemented projects. The study participants drew attention to the negative emotions that arose in their environment in relation to the ideological or political initiatives proposed in BOMK. Due to the increasing polarization in the country's political life, most city residents expect the local politics to be separated from ideological disputes. However, these kinds of initiatives appeared very rarely, so they were not a major issue. A more frequent problem, however, was the inability to extrapolate the long-term effects of the proposed projects, which was categorized as the creation of the structural model of the proposal. The respondents described situations in which the public was concerned about the possible negative effects of particular investments. A noticeable group of city inhabitants were afraid of any changes in their surroundings because they believed that “[these projects] will generate a lot of yet unknown problems (...) Prejudices are growing around topics that generate excessive activity in the neighborhood” [INT 8]
Collective memory	<ul style="list-style-type: none"> The study participants were generally convinced that building collective memory, aggregating experiences, learning, and gaining feedback are all issues that cause the most trouble. All of the respondents expressed the opinion that collective memory works only to a very limited extent. They agree that a lot of work is done every year, primarily collecting opinions, discussions, negotiations, designing, and presenting ideas and projects, but this whole thought process is preserved only fragmentarily, and there is very little chance of accessing and recovering knowledge

(continued)

Table 3.5 (continued)

Cognitive process	How does it work in the project?
	<ul style="list-style-type: none"> Communication is conducted using various platforms, but its results are not collected or are very difficult to access. According to the respondents, this should be changed: "We should gather information on why some things succeeded, while others failed. The feedback we receive about the projects' development is weak, while it should be the most influencing factor in developing new ideas" [INT 5]. Unfortunately, even people deeply involved in BOMK complain about the lack of feedback: "Many of the submitted projects failed for reasons that I do not know. Inhabitants often vote in favor of a promising project, but then these projects are not implemented for unknown reasons, or sometimes they come to fruition with a delay of several years" [INT 3]. City residents often feel that the actions they contributed to by voting in BOMK resulted in unforeseen consequences that could have been avoided with a well-conducted evaluation that took into account past experiences and provided the possibility to recover and access them (categorized as meta-memory) The prevailing opinion is that at the system level, no one planned the creation of collective memory in BOMK and grouping experiences into blocks: "There are many people who do not have time to make their own proposals, but who could contribute their experiences to the debate. In turn, those who have achieved success should share their experiences. However, there is a lack of such activities in an organized form, there is insufficient space and not enough incentives for it" [INT 9] This problem is associated with the high turnover of people participating in project preparation and voting in BOMK. People are discouraged because they do not notice the positive effects of the implemented activities, both on the Internet and in the urban space, even if such effects occur: "There is a lack of visual information in the city that some investments were made as a result of the civic budget. (...) There is also no information about the experiences so far and about what worked. It's about showing the residents what has gone well and why. (...) The townspeople do not associate the investments they need with the civic budget, so they don't self-mobilize and do not fully believe that something can be changed" [INT 2] Local media also contribute to this as, according to the respondents, they enjoy focusing on problems rather than presenting an objective picture of the situation: "The media are full of information about the failures of projects, but there is little information about achievements. This spreads a false, unrepresentative image. The negative aspects are disproportionately more publicized" [INT 8] Collective memory at the level of the city authorities and BOMK managers also does not work well: "The city authorities do not draw conclusions from the implemented projects. Officials rather discourage people from bold initiatives, make it difficult to collect data on relevant projects, do not support quality, but only simple solutions. Ideas that could capture the imagination and become the 'flagship' achievements of the citizens' budget are rejected, such as the proposal to build a Formula 1 racetrack, which could not even be voted for. Another obstacle is the failure to draw conclusions from previous years regarding too low rates imposed by municipalities on project pricing, which in the past resulted in the limited viability of many projects" [INT 7]

(continued)

Table 3.5 (continued)

Cognitive process	How does it work in the project?
	<ul style="list-style-type: none"> The study participants claimed that the BOMK council, which includes both city hall representatives and community organizations, is aware of the problem but does not have a plan to find a remedy for the situation: “The budget council is constantly arguing about what went wrong. To understand this better, they use evaluation studies prepared by an external company, but the company writes the same thing every year, and most often it is not implemented. We are not able to check what actually worked, why something went well and what was worse, there is no visualization of what was done” [INT 11]. According to the respondents, in subsequent editions of BOMK, the same mistakes that have already occurred are often repeated

(selection of projects in a vote) would be implemented, contributing to a change and a better environment in which they live. Furthermore, it's noticeable, particularly in the Facebook discussion groups named “Joint Projects,” that to become a member, there is a stipulation to adhere to group behavior norms and communication standards. The referenced norms include: “focusing discussions on the needs of a specific district to prepare action proposals,” “staying on topic and adhering to the group’s purpose,” “maintaining civility and politeness,” “avoiding promotions or scams,” and “prohibiting hate and bullying” (Wspólne Projekty 2020).

From the perspective of policymaking phases, the BOMK project, like most projects implementing participatory budgeting, can be placed partially in policy formulation and partially in decision-making. However, this can only be done with certain reservations. Here, there is not an opportunity to freely shape the directions of urban policy and discuss its general course, but rather a selection of methods for implementing municipal tasks that generally have consensus on the necessity of their execution. It can even be said that a significant portion of the projects align with the city’s overall policy and would likely be realized even without a participatory budget. Arguments for and against individual ideas are gathered only to a small extent during the debate. The most important method of evaluation remains the selection of projects through voting and their integration into the city’s investment plans.

Our study’s most salient findings indicate that while most cognitive processes function effectively, one significant exception is the process of collective memory, which is notably deficient. Nearly all respondents claimed that memory-related operations are the most critical problem. The editions of BOMK seems to repeat similar errors in successive iterations. New participants seldom use past experiences, and after participating in the processes once, they do not remain a part of the system. Also, learning mechanisms do not work well in practice. A cognitive system operating in this way has difficulty delivering intelligent results because, as cognitive psychology proves, all system operation depends on memory and continuous learning, which forms the basis for increasing the efficiency of *thinking* of a given community.

According to the survey participants, information about the experiences and achievements developed so far, even if available somewhere on the Internet, is very

difficult to find and not linked together in a logical structure. There is clearly no dedicated technological device that would facilitate processes related to collective memory. The observations and conclusions that emerge from debates conducted on dispersed Facebook groups and in other social media are irretrievably lost. Furthermore, the advanced *communicative actions* (as defined by Habermas, 1984), such as the negotiations held by municipal officials during which opinions are aggregated and duplicating projects are merged, could be an extraordinary source of collected practical wisdom, as proven by experience. However, this potential is not used. The respondents' answers indicate that, in the current model, knowledge gathering takes place only when *experienced* NGOs and social groups develop their organizational memory separately in their field of narrow specialization, acting independently of each other and, as it were, outside the system. In such a situation, building an integrated system that allows for the shared use of memory resources is difficult. The gathered knowledge can enhance the functioning of the entire cognitive system, forming the foundation for progressively more efficient intelligence throughout the community. However, the project consistently replicates the same mistakes in subsequent cycles, with new members seldom utilizing previous experiences. Additionally, individuals tend to disengage from the system after their initial involvement. By emphasizing the project's outcomes without fully leveraging the amassed knowledge and experience, it appears the system prioritizes identifying specific solutions over cultivating a broad knowledge foundation.

The problems related to the collective memory process in the BOMK project were confirmed not only by the semi-structured interviews but also by the second method I have adopted, i.e., examining documents obtained through desk research. The high turnover among people involved in the project means that the creative potential of the applicants from the previous editions of BOMK remained untapped. A large number of participants get discouraged upon encountering problems, which, instead of being a valuable lesson learned in the following years, is repeated by subsequent groups of people. Of course, the situation in which certain groups of participants are replaced by others is normal, but the desired course of action would be to use the organization's memory rather than starting all over again. Therefore, creating lasting relations with the participants, especially to keep them in the project, would be one of the most important tasks of the BOMK managers in the future.

3.3.2 *Better Reykjavik*

3.3.2.1 Introduction and Background

Better Reykjavik (BR) is a unique project. It's an online consultation forum where citizens have the opportunity to present their ideas on various issues concerning their city. Its inception is closely linked to the economic crisis that hit Iceland in 2008. Following the bankruptcy of the country's three largest commercial banks, most residents' savings were severely threatened, leading to a crisis of trust in the existing

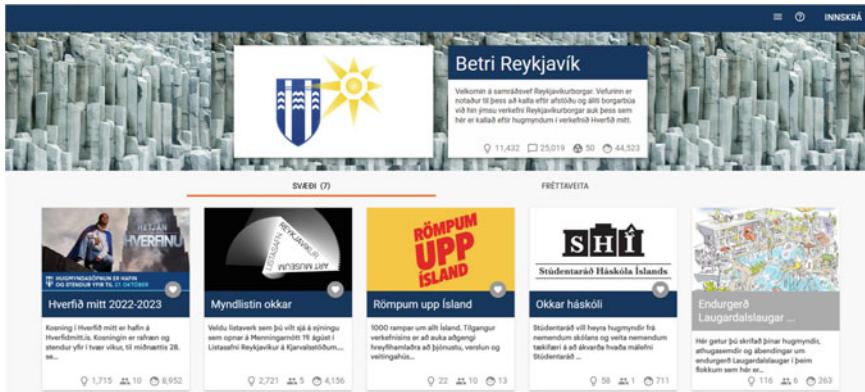


Fig. 3.7 Home page of Better Reykjavik in Icelandic. Source <https://betrireykjavik.is/>

political elites. BR began in 2010 as *Shadow City* (Skuggaborg), a participatory online platform developed by the non-profit NGO Citizens' Foundation (Adenskog, 2018). It gained significant traction riding on the popularity of the Best Party—a political group that challenged the establishment and aimed to transform the representative system into a participative democracy. This project's success led to its adoption as an official tool in Reykjavik's urban policy. BR quickly achieved considerable success, establishing itself as one of the world's pioneering projects in transitioning citizen participation to an online platform (Fig. 3.7).

Initially, the platform's focus was not on participatory budgeting but rather on fostering a free exchange of views and facilitating interaction among users. This was achieved through the presentation and mutual evaluation of arguments. It stood out as one of the first projects to innovate the presentation of policy ideas, uniquely combining them with the requirement to provide arguments both for and against. Later on, the platform evolved to include a *My Neighborhood* section, which offered users the opportunity to submit and select projects to be financed within a participatory budget. Today, the BR platform empowers citizens to propose, discuss, and prioritize ideas for city improvement, fostering open discourse between the community and the city council. It also allows participants to have a direct impact on decision-making processes, enhancing the inclusiveness and responsiveness of urban governance.

Better Reykjavik and *My Neighborhood* served as key inspirations for the Decide Madrid project, which received consultation from the Citizens' Foundation. The Norwegian Consumer Agency has also adopted the open-source version of the BR platform, known as *Your Priorities*, to engage with the Norwegian public and assist in prioritizing their initiatives. Since its inception in 2010, the *Your Priorities* solution has been officially implemented in over 20 countries. It has been utilized to crowd-source questions for the government by two majority parliamentarians in France and for various projects in Scotland, Norway, Hungary, Croatia, and the Estonian

Table 3.6 Basic information about the Better Reykjavik project

Project name	Betri Reykjavik (Icel.)/Better Reykjavik (Eng.) your priorities (open source edition)
Owner	Citizens Foundation
Web address	https://betriereykjavik.is/
Location	Reykjavik, Iceland
Years in operation	2010–present
Implementation level	Local, global

Rahvakogu in 2013, remaining one of the most famous examples of e-participation in the world (Citizens.is 2023) (Table 3.6).

3.3.2.2 Project History, Key Features, and Impact

The development of the Better Reykjavik (BR) platform commenced in 2008 and was significantly influenced by the Icelandic financial crisis. Before the crisis, Iceland experienced a remarkable economic transformation, moving from a predominantly low-value fishing industry to a high-value, finance-centered economic structure. This shift was greatly propelled by the country's embrace of economic liberalization, leading to significant growth, low unemployment, and strong development, which seemingly set Iceland on the path to a prosperous future. However, this trajectory took a dramatic turn in September 2008, when Iceland was hit by a severe recession related to the global financial crisis. Iceland's financial system was characterized by a high level of leverage, indicating that the debt compared to the GDP was extraordinarily elevated. This vulnerability left the economy particularly exposed to external shocks. When the global credit crunch struck, it hit Iceland's banking sector hard. As international financial markets struggled, Icelandic banks, which were highly dependent on foreign funding, encountered severe challenges in refinancing their short-term debt. The collapse of its three major banks not only created a deep sense of crisis and anxiety among the populace but also sparked a widespread wave of dissatisfaction and anger towards the existing political and financial systems (Carey, 2009). The launch of Better Reykjavik aligned well with the public's desire for change.

BR was established and initially financed by two private individuals, Robert Bjarnason and Gunnar Grimsson. Its success soon led to the creation of a local NGO, the Citizens Foundation. Launched in 2010, just before local elections in Reykjavik, the platform quickly gained attention from the Best Party, a group participating in the elections known for its satirical critique of local politics, which went on to win the city elections. From the beginning, BR featured a unique debate system where users contribute talking points and arguments for and against ideas, moving away from the traditional comment section. This website offers Reykjavik residents the chance to submit, discuss, and prioritize policy proposals and ideas for municipal-level issues. It enables them to express, debate, and refine various ideas they consider vital for

urban policy. A particularly interesting feature of the platform is its up/down voting system, which allows users to cast votes not only on ideas but also on the arguments presented by others. This creates a self-regulating mechanism that, without the need for moderation or administrative intervention, generates a user-prioritized list of ideas, complete with the most compelling arguments for and against each. This prioritization is exclusively driven by the citizens, providing clear evidence of their preferences. Moreover, by allowing citizens to add points for and against in separate columns, the platform encourages well-considered contributions, thereby reducing the usual online toxicity often found in comment sections. BR employs machine translation and artificial intelligence to suggest ideas and send intelligent notifications and features a toxicity sensor that alerts administrators about abusive content, although such instances are rare. Additionally, the platform incorporates an automatic classification system for ideas. (Observatory... 2023).

In 2011, the BR website was formally recognized as a collaborator by the Reykjavik City Council. As a result, ideas regarding the city's development formulated on the platform gained tangible influence on local policymaking, establishing BR as a hub for ongoing dialogue between citizens and officials on various topics. The ideas developed by the BR community are formally addressed by city officials as follows: each month, the top five rated ideas across all categories (such as tourism, operations, recreation and leisure, sports, human rights, art and culture, education, transportation, planning, administration, environment, welfare, and various others), with up to one top-rated idea in each category, are presented to the appropriate city committee (Bojic et al., 2016). Decisions to implement or discard these ideas are communicated from the department to the project manager, who then informs the citizen who initiated the idea (Adenskog, 2018). While there is no specific budget limit for each individual idea or policy proposal submitted on the BR platform, there is an annual budget allocation of approximately 300 million ISK (about 2.1 million USD) for the program overall.

Also in 2011, the participatory budgeting platform *My Neighborhood* (MN) was launched as a supplement to BR, with a focus on neighborhood improvements. MN, accessible through the Better Reykjavik platform, is tailored to manage the participatory budgeting process. It was kick-started with an initial investment of 6.2 million USD from the city of Reykjavik (Adenskog, 2018). MN enables a collaborative effort between Reykjavik residents and the city administration to decide on capital allocation for construction and maintenance projects across the city's ten main neighborhoods. It allows citizens to register in their neighborhood based on the address listed in their voter registry. Following BR's model, the annual My Neighborhood forum encourages people to propose, discuss, and assess various ideas and project proposals. A new, distinctive feature is the final voting round, where citizens select their preferred projects that can be funded within the budget constraints. This participatory budgeting initiative, with a budget of 450 million ISK (4.2 million USD), empowers the public to allocate approximately 6% of the city's capital investment budget (Observatory... 2023).

While BR is marked by its bottom-up, loosely structured approach to participation, MN follows a more formally structured process divided into four stages, spanning

approximately a year. The process begins each October with a press conference hosted by the mayor, officially initiating the collection of ideas from citizens in each of Reykjavik's neighborhoods. This idea collection phase, open to all residents of the designated neighborhoods, lasts for one month. The second stage involves a local committee evaluating these ideas and suggestions and selecting those that can be put to a vote. In the third stage, the shortlisted ideas are presented to the citizens for consideration. Through the website, they can select ideas within the budgetary constraints. Citizens then create their own list of priorities within the budgetary framework and vote on the projects they believe should be implemented (Adenskog, 2018).

The level of participation in the project is quite impressive. To date, over 70,000 residents of Reykjavik, out of a population of approximately 131,000, have engaged with the BR activities. This means that more than 53% of the city's inhabitants have interacted with the platform. There are over 44,500 registered users who have collectively published app. 11,400 ideas and over 25,000 debate points. Over 500 ideas have been sent to the city for implementation through the agenda-setting part of Better Reykjavik, and nearly 700 participatory budgeting projects have been implemented, thanks to the My Neighbourhood section (Observatory... 2023). The annual participatory budgeting voting typically sees about 10–12% of the city's population participating. The system's functionalities are constantly being expanded. For example, in 2023, an AI Chatbot was introduced, which helps users search for information from the database of projects from previous years and their details.

It is important to note that apart from its ongoing programs, the platform has hosted several special projects aimed at crowdsourcing policies in specific areas. For example, in January 2017, the Reykjavik City Council decided to crowdsource ideas for co-creating its Education Policy 2030, soliciting input from primary stakeholders, including teachers, other staff members, parents, and students. Meetings with key stakeholders were conducted from February to April 2017, identifying five essential competencies and skills in need of emphasis: social skills, self-empowerment, literacy, creativity, and health. The council then sought ideas from the public on how the education system could nurture these skills, combining offline workshops with online discussions through the BR platform. From May to June 2017, around 10,000 people participated in the crowdsourcing process, with 5,800 engaging online, leading to the generation of 56 ideas and 204 arguments. In the autumn of 2017, these ideas were consolidated into a draft proposal and action plan. Following several reviews, the city council approved the policy. The formal implementation process commenced at the end of December 2018, with a commitment from the implementation team to closely monitor, evaluate, and review the process after three years (Crowdsourcing Better... 2019).

Not everyone considers the achievements of Better Reykjavik to be a complete success. Magnus Adenskog conducted a qualitative study between 2012 and 2014, consisting of a series of interviews with civil servants who worked with BR and MN on a daily basis. They indicated a noticeable decline in engagement and participation in the debate. After an initial period of high interest in participation linked to the financial crisis, there was a significant reduction to a few percent of the most active

electorate. The conclusion drawn by the study participants was that broad-scale participation is possible in crisis situations, but on a day-to-day basis, citizens do not engage in debates on general matters. The lack of controversy significantly reduced interest in participation. As a result, ideas with only a few signatures ended up being discussed by the city council: “One of the consequences is that priorities with only a few supporters (sometimes even as few as a single supporter) also have to be treated in committee as formal request” (Adenskog, 2018). This has led to some frustration among politicians and civil servants, as the initial design of both BR and MN was predicated on the high levels of participation during the crisis, not the lower rates typical of everyday politics.

3.3.2.3 Collective Cognitive Processes in the Project

The method I adopted for evaluating the CI processes in the Better Reykjavik project was based on the framework detailed in Sect. 3.2.2. To address goals of the evaluation, the study was carried out in two stages based on a qualitative case study method. Initially, desk research was conducted to explore the fundamental factors influencing the identified CI processes to understand the context of the project. Content analyses of the Better Reykjavik website, relevant documents, official reports, statistics, and other data prepared by third parties were conducted. A particularly important source of knowledge was the aforementioned study by Adenskog, which included a total of 24 interviews with civil servants who worked with BR and BN on a daily basis, politicians in Reykjavik, researchers, and NGO representatives (Adenskog, 2018), and the study “Youth and democracy: Digital opportunities for the future of participation” by Ambrosino et al. (2023), including 12 interviews. The second stage of the study was an In-Depth Interview (IDI) with a key expert for the examined platform, Robert Bjarnason, who was president of the Citizens Foundation and one of the founders of the BR platform. The interview was conducted using a script that covered all the key issues included in my evaluation framework, which were broken down into supplementary questions to delve deeper into the expert’s viewpoints and experiences. The final stage involved extracting insights, perspectives, or areas of consensus or disagreement from the literature and interview, synthesizing the findings, and concluding the qualitative analysis. The table below summarizes the statements of the interview participant and written sources. Quotations from written sources have references to literature; direct quotes from the interviewee’s statements are in quotation marks, with the source marked [RB] (Table 3.7).

3.3.2.4 Metacognition: Understanding Cognitive Processes in Better Reykjavik

This section presents an analysis of the meta-processes influencing cognitive processes in the Better Reykjavik platform. It includes a summary of the key conclusions drawn from the analysis of cognitive processes, insights into how these

Table 3.7 Cognitive Processes in the Better Reykjavik Project

Cognitive process	How does it work in the project?
Collective sensing	<ul style="list-style-type: none"> • The diagnosis of the needs of the local community in Reykjavik primarily involves identifying what people expect in their urban environment. “That’s often the first thing you need to do, (...) to identify the problem before you start to look for solutions” [RB]. “When you’re looking at complicated problems (...) the individuals have ‘pieces of information’ that can be collected” [RB] • Challenge identification is associated with prioritizing, i.e., “breaking down problem into manageable sub problems. (...) That’s the thing that you want to know at first, before you start to do any policy, you want to know what are the most important problems” [RB]. The government has “limited resources” to decide which problems are the most important. “If you’re a government, you cannot assemble a list of 40 possible problems, connected to an issue that is being solved” [RB]. The community can help with that • The initial stage of problem identification takes place in social media “like 70% of the new visits come through from people sharing their own ideas on social media” [RB] • A natural starting point for identifying challenges is when people complain about the issue not being resolved. The benefit for the entire community is to move from complaining to looking for solutions. “For example education policy of the city of Reykjavik. People complain about the schools all the time from different perspectives. People should be criticizing, but when we did the crowdsourcing for it, it really gave a lot of people an opportunity instead of just talking about it on Facebook or Twitter to going deeper into the problem in order to contribute an idea in the next stage. When it is actually organized together with government it is a unique point that government is listening, so that’s a motivation for people” [RB] • BR managers want to increase the number of participants to attract a higher number of people (Adenskog, 2018). “The highest quality inputs are just a percent of the total. So the larger amount of inputs we have, the better” [RB] • Having a large number of participants setting the agenda decreases the risk of corruption. “...the situation where friends and family of the people working in government have more influence on policy (...)— we have a word for that. That’s called corruption” [RB] • “More diverse ideas are better. That’s what our government partners want. One of the assumptions is that it is totally inclusive.” City authorities are not satisfied if there is not much activity. However, the truth is that when it comes to selecting priorities for urban problems, if there is no social crisis, there will be no mass participation (Adenskog, 2018)
Problem solving	<ul style="list-style-type: none"> • Great emphasis is placed on ensuring that the debate is constructive and that unproductive or toxic messages are minimized (Ambrosino et al., 2023). The debate serves “a utilitarian function of bringing evidence, and bringing ideas that contribute to a solution” [RB] • The debating system on Better Reykjavik (BR) is perceived as “a sort of competitive task” [RB]. It is described as “(...) Your task is to help us find the best pros and cons for this idea, not to tell us what you think about the idea, but just to help us find the pros and cons of it” [RB]. Therefore, the fundamental mechanism for engaging participants is competing to present attention-grabbing ideas and the arguments associated with them. This method of participation, based on competing and collecting points, has its origins in the experience of the project creators in the video games industry. “We are setting people in this little debating game” [RB]

(continued)

Table 3.7 (continued)

Cognitive process	How does it work in the project?
	<ul style="list-style-type: none"> Participants are discouraged from making unconstructive or critical comments, and it is not possible to reply directly to the submitted proposal (Ambrosino et al., 2023) The substantive level of the debate is assessed as quite high: unlike the typical debate on social media, “it is not like majority is making information noise” [RB]. There is an AI mechanism that blocks toxic comments, but in practice it is used very rarely Participants’ creativity is supported by breaking the official documents into a collection of simple-language statements: “AI-supported breaking out the key parts of the policy where input is needed (...) makes it more cognitively accessible. (...) If you want to help people be creative, help people to be useful” [RB] Conflict is treated as an inherent part of debate, but the system supports the transformation of unproductive polemics into argumentative rankings of ideas. “People have different ideas, and when we talk about policies, there are often very, very different conflicting views. I mean difference of opinions is good, because that’s the only way you’re going to reach any result” [RB]. At the same time, BR is trying to maintain “as low temperature of conflict as possible” by directing debate towards arguments. This is very different from the typical situation on social media. “The type of polarizing communication on public issues as on Facebook or Twitter would be of no use to cooperation with city council (...). [If we only relied on social media] we would not be here today. (...) Differences of opinion and conflicts are good, but if it’s too hot, it totally kills the debate” [RB] Most attractive ideas are one-person proposals. However, collaborative proposals, developed entirely in the group, appear in specific situations. An example was a special project under which the Reykjavik educational policy was developed. “The policy work is a lot more collaborative. When you start, you break the policy points into several groups and sub-groups. The discussion is focused on a very specific thing in the policy, and the new ideas are based on the previous ideas. And there are ideas that fill the gap that other ideas leave out. That is a much more collaborative sort of an online work” [RB]
Decision making	<ul style="list-style-type: none"> Decision-making is the process of actually selecting the best contribution. The system is open for every citizen to vote for or against the proposal (Adenskog, 2018) “High-quality information is the basis of a good decision. (...) Citizen engagement is something that can improve the whole process of decision-making,” [RB] bringing the decision-makers actual evidence for their decisions thanks to the quality of information. The quality of information increases thanks to “displaying the arguments and options in the most accessible way” [RB] “If you’re making a policy about something that affects people’s lives” [RB], you might expect all sorts of participants to provide different types of evidence, not only the ones “considered academically the smartest people” [RB]. It increases cognitive diversity in making decisions

(continued)

Table 3.7 (continued)

Cognitive process	How does it work in the project?
	<ul style="list-style-type: none"> • The prioritization mechanism plays a key role in decision-making. “We have this prioritization mechanism on the platform. (...) We are all dealing with information overload on all levels, so prioritization helps to increase the quality of the evidence” [RB]. During the decision-making process, participants “can see not only the winning proposal but also the stages of debate and the options that appeared” [RB], even if they were not used in the final solution • Icelandic experiences indicate that decisions made in votes with high turnout are largely respected by the majority of voters (Ambrosino et al., 2023) • BR managers find it very important to secure the voting process against hacking or fraud. The security of this process is crucial for building trust in the system. For this reason, electronic identification (using ‘eIDs’) is required
Collective memory	<ul style="list-style-type: none"> • BR managers are aware that collective memory is one of the key processes, and they believe that BR stands out positively from other projects: “I think we’re doing pretty well on that. Our platform is collecting the information, but also keeping it very accessible. (...) On many other platforms, memorizing experiences is not so developed. (...) It’s so sad, you know. So much information is lost, even in EU projects, where there are requirements to keep things accessible, but often, they are not really accessible as soon as the project is finished. This may be a website that should be kept up for one or two years, but there’s no real community. No one makes sure that the historical data is something we can also learn from, (...) tracking the evolution of the ideas connected with topics” [RB] • What about the evaluation of investments that the participants have selected to be financed by participatory budgeting? “In some cases we’ve had people asking, ‘oh, so where is this? It was approved like a year ago in a vote, I don’t see it, you know what’s happening?’” [RB] • All information regarding past ideas, associated arguments, and other technical details are cataloged and made accessible through the BR portal. “All the data is publicly available and anybody can conduct searches (...). We have folders and subfolders with older projects dating back to 2011 (...). You can go back all those years. You can see all the decisions of the governments in the newsfeed to see if they are progressing in the investment process” [RB] • To facilitate better access to the collective memory resources developed in BR, an AI Chatbot based on a dedicated Large Language Model (LLM) is currently being implemented. This will allow users to communicate in natural language about the accumulated knowledge and past experiences stored in the system

processes are interconnected to form a cohesive whole, and an inquiry into what drives people’s participation in these processes. The data was collected based on an In-Depth Interview conducted with Robert Bjarnason, president of the Citizens Foundation and one of the founders of the Better Reykjavik platform, and my own analysis based on empirical research and literature review. All quotes coming from the interview are marked [RB].

Collective intelligence on platforms like Better Reykjavik “can take many forms” [RB]: it can manifest itself both in small, tightly-knit groups and in large, loosely connected communities. The goal of Better Reykjavik was to create a platform for

mass participation by a wide range of citizens who could, even with a small investment of time, exert their influence on public policies. The founders of Better Reykjavik believe that “the worst aspect of collective intelligence and how it’s being used is when the collective intelligence is essentially ignored. Government officials often initiate policy consultations after the policies have already been approved” [RB], and social consultations are sometimes conducted in a superficial manner to *legitimize* a policy that was going to be adopted anyway. It’s “not listening to the collective intelligence when it has been collected, but pretending to collect it (...). Fortunately, this doesn’t happen frequently on our platform” [RB].

In relation to the phases of the policy cycle, Better Reykjavik goes beyond the standard scope of participatory budgeting projects. Since BR empowers citizens to propose, discuss, and prioritize ideas for city improvement, fostering open discourse between the community and the city council, we can speak of a mutual partnership between citizens and the city in the *Agenda setting* area. The ideas developed by the BR community are formally addressed to the corresponding city office units, such as those dealing with infrastructure, tourism, education, etc. They can take the form of general ideas and in-depth analysis of needs, which then influence the direction of urban policy. As for the *Policy formulation* area, the emphasis on building lists of arguments for and against and collecting ranking points by individual ideas and arguments intersects with setting policy objectives and estimating the costs and effects of the solution. The policy decision-making area is not very extensive and is based on the results of work from previous phases.

Even though participation in Among Better Reykjavik is primarily of an individual nature, there is a sense of certain responsibilities towards the community: a “community motivation of being a part of something” [RB]. Nevertheless, participation is usually limited to single entries and is not very interactive. Active participation of a community nature applies to only a small percentage of the population, those who choose to dedicate their private time to it. “People are quite busy (...). A very small part of society has the luxury of dedicating hours and hours to a community project. You can’t expect them to spend hours on it, so we are mainly targeting people with free time” [RB]. Therefore, Better Reykjavik caters to those citizens who do not have much time for extensive involvement in local matters but can spare a few minutes. This is facilitated by a highly simplified presentation structure of policy ideas and investments based on gathering related pros and cons, making it easier for participants to quickly form an opinion on a specific topic.

Since the very beginning, the identity of Better Reykjavik users was built on challenging the existing form of governance: “One of the motivating drivers was to be a part of the community that organizes itself from the grassroots and takes matters into its own hands, rather than relying on inefficient governments” [RB]. Therefore, transparency is the hallmark of this new approach to policymaking. “There is definitely a public interest in understanding how policy debates are conducted, in understanding how the policymaking processes work, and in understanding the evolution of things” [RB]. The common identity of BR users also stems from the fact that the projects are “neighborhood-based, and they are supported by organically grown Facebook groups for neighborhoods” [RB].

“But then it’s also a question of interest” [RB]. In participatory budgeting, the best ideas compete with each other, and residents’ individual interests lead them to support projects in their neighborhoods. The adopted debate model places emphasis on *competition* among participants in providing pros and cons. The motivation to participate is linked to the sense of the possibility of effecting change. What unifies the community is a sense of agility, “a motivation to contribute something useful” [RB]. This attitude was particularly strong at the inception of Better Reykjavik when many residents of Iceland found themselves in difficult life situations due to the financial crisis. This stimulated people to come together to address problems and take responsibility for their country and city, regardless of the government’s perceived inefficiencies.

One could get the impression that the daily motivation to participate is primarily based on presenting competing ideas, gathering related arguments, and seeking to fulfill individual interests. The sense of community and engagement, on the other hand, tends to manifest mainly in the case of crises or exceptional events (such as creating an education strategy). It seems that people do not engage in the system as “community members” but as independent individuals, each pursuing their own goals and satisfying their own needs (including the need for agility and belonging). The collective working for the common good emerges somewhat *incidentally*, somewhat akin to Hayek’s vision of spontaneous order.

Citizen engagement is not built on unproductive disputes, as it often happens on social media platforms (as exemplified by the Facebook algorithm, which “rewards constant arguing to keep users on the site” [RB]). Better Reykjavik is different because “it is quite respectful of people’s time” [RB]. Due to its simplicity and accessibility, even a very short presence on the platform can bring valuable participation. “Your government is asking you to take part in this process, and here is an interface and content. There are pleasant images that bring you joy and nice colors. It’s like an enjoyable experience (...). Maybe they only spend ten minutes, but they do something useful. And then it is repeated one thousand times. It is like connecting people in gamified ways” [RB].

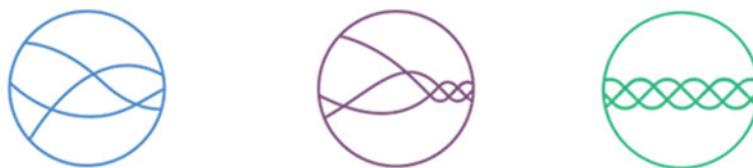
3.3.3 *Loomio*

3.3.3.1 Introduction and Background

Loomio is an online platform developed by a non-governmental initiative that emerged from anti-establishment protest movements. The name *Loomio* is derived from the English term *loom*, which in this context refers to a cloth-weaving workshop. It symbolizes *weaving* the social fabric by making decisions that incorporate numerous voices and perspectives in the debate. However, the debate here is not conducted independently but is closely linked with the decision-making process and the pursuit of consensus. This method of collective work requires that the groups using Loomio are not as large as those on participatory budgeting platforms. They

are instead more akin to closely cooperating groups working together on specific tasks. The primary goal of the project is to facilitate the management of a group working towards a common purpose. As illustrated in Fig. 3.8, Loomio comprises three parallel activities: (1) the debate, which allows for the presentation of various opinions and ideas for solutions to identified problems; (2) building agreement: each debate participant can propose a solution, which will then be commented on by others and subject to voting according to specific rules, and (3) collective decision-making.

The origins of Loomio trace back to the Occupy Movement in 2011, specifically the Occupy Wellington protest. This movement, a part of the global wave of protests, underscored the need for a new tool to facilitate project management, collaborative decision-making, and civic engagement, particularly in scenarios where physical presence in general assemblies was not possible. A group of developers and social innovators from Enspiral, a Wellington-based network and incubator for social enterprises, united to create a solution. In their search for project management software, they quickly realized that most available software catered to a hierarchical approach with a designated leader assigning tasks and a pre-established plan for the future (Bartlett & Deserriis, 2016). Contrary to this, the movement envisioned a non-hierarchical method, where all participants have equal rights and plans can organically develop from the bottom up (Table 3.8).



1. Talk things through 2. Build agreement 3. Decide together

Fig. 3.8 Three parallel Loomio activities according to early documentation. *Source* (Olszowski, 2015)

Table 3.8 Basic information about the Loomio project

Project name	Loomio
Owner	Loomio Cooperative Limited
Web address	https://www.loomio.com/
Location	Wellington, New Zealand
Years in operation	2011–present
Implementation level	Local, global

3.3.3.2 Project History and Key Features

The Occupy Movement, which served as the political backdrop for Loomio, was a global anti-establishment movement originating from a sit-in protest that commenced on September 17, 2011, in New York's Zuccotti Park. Its participants referred to themselves as *the 99 percent* to emphasize their protest against the appropriation of society by a mere one percent of citizens, which included the ruling political circles and the financial elite. Similar protests, lasting for several weeks, occurred concurrently in numerous cities around the world, such as Montreal, London, Lisbon, and Rome, among many others. Some of these cities included various locations in New Zealand, where people who were later involved in Loomio joined the protesters. Richard D. Bartlett, one of the founders of Loomio, recollects the events as follows: "We come from the Occupy movement, preceding that was the M15 Indignados movement, the Arab Spring, the Sunflower Movement, Occupy Central in Hong Kong... there's an ongoing wave of social movements" (Olszowski, 2015).

Numerous protesters aimed to transform the critical movement into a collective that could implement the principles of participatory democracy and address political issues. This led to efforts to establish a coherent system for sharing opinions, discussing them, and voting. This system was based on the work within problem-focused groups and voting during the general assembly while adhering to certain principles regarding the order of speeches and moderation. To facilitate discussions within specific groups during meetings, a set of gestures was introduced for participants to use. These gestures conveyed various intentions, such as a willingness to speak, provide a direct response, explain one's viewpoint, express agreement or disagreement, abstain from voting, or strongly object to a proposal. The primary advantage of this model was its focus on achieving consensus. "People involved in this 'consensus' process find it incredibly empowering—a refreshing contrast to the alienating remoteness and weary predictability of parliamentary representative democracy, which most people here see as totally irrelevant to their real lives" (Miller, 2011).

However, this organizational method faced a significant challenge: it became cumbersome when it came to taking the floor and moderating discussions in cases involving a larger number of participants. Technical issues soon arose, as described by Sifry (2013): "The dark side of that is that meetings last 5 h, the process breaks down, and everyone can walk away feeling disempowered." Some discussions would drag on for many hours without ever reaching a consensus, and decisions were sometimes made without input from the entire group, contrary to the initial expectations. Benjamin Knight, another of Loomio's founders, recalled how problems with organizing discussions and voting often undermined the entire effort: "A few voices could dominate the conversation, or a small group could dominate by waiting everyone else out" (Finley, 2014).

Richard D. Bartlett describes this situation as their primary inspiration for founding Loomio:

How could we learn from this amazing and inspiring methodology, this practice of direct democracy, while avoiding some of the major pitfalls? The obvious answer was to use technology, to replace the General Assembly where people have to sit together in the same place at the same time, and move that into an online space where people can participate on their own terms, when they have time, on the issues they care about (Olszowski, 2015)

To achieve this goal, a group of individuals from New Zealand, primarily from Wellington, became involved in the project. This group consisted of technology specialists, open-source developers, and social activists who decided to form a team and collaborate in the coming months. Initially, the informal group was referred to as a *social tech cooperative*. Later, it formalized its activities as a social company known as Loomio Cooperative Limited. The group quickly established a partnership with a local startup incubator, Enspiral. The popularity of the initiative extended beyond local circles at the time, and the tests involved more than 2,500 volunteers participating in approximately 400 discussion and decision groups, both from New Zealand and other countries.

After 18 months of work on the prototype and gathering feedback, it became feasible to proceed to the next stage: the development of a stable version of the application. This phase required a new source of funding, so in early 2014, a decision was made to initiate a crowdfunding campaign. The campaign proved to be successful, as it garnered over \$125,000 from 1,657 sponsors (Olszowski, 2015). This success enabled the launch of a dependable and functional software version, marking a significant milestone in Loomio's development.

In the time since its inception, the project has expanded, leading to the creation of many autonomous instances of Loomio. The software is now available in two forms: as a Software as a Service (SaaS) hosted on servers managed by Loomio Cooperative Ltd. and as an open-source solution, allowing anyone interested to download and self-install it free of charge. Various groups, including political organizations, such as the Green Party of Western Australia and Spain's Podemos, have established their own discussion and decision-making groups using Loomio. Institutions like the National Assembly of Wales have employed Loomio within their Health and Social Care Committee, while the New Zealand Government utilized it in the development of a Workforce Capability Framework (Loomio Blog: Political Parties 2023). Loomio has also found application in numerous public events. For instance, the Ministry of Economy of the Republic of China (Taiwan) chose to provide the application to participants of the National Energy Conference, using it as a tool to facilitate participation and the submission of proposals for discussion (Olszowski, 2015). Additionally, it's noteworthy that Loomio's popularity extends beyond the realm of public participation. Many enterprises that embrace non-hierarchical management principles and various types of membership organizations count among the significant user base of Loomio.

What distinguishes Loomio from other described cases in terms of functionality? It can be observed that the entire process of participation is organized around decision-making (Fig. 3.9). The most unique aspect of Loomio is that it presents deliberation and decision-making side by side. Discussion participants create a group within which they can communicate, start discussion threads, create proposals and polls,

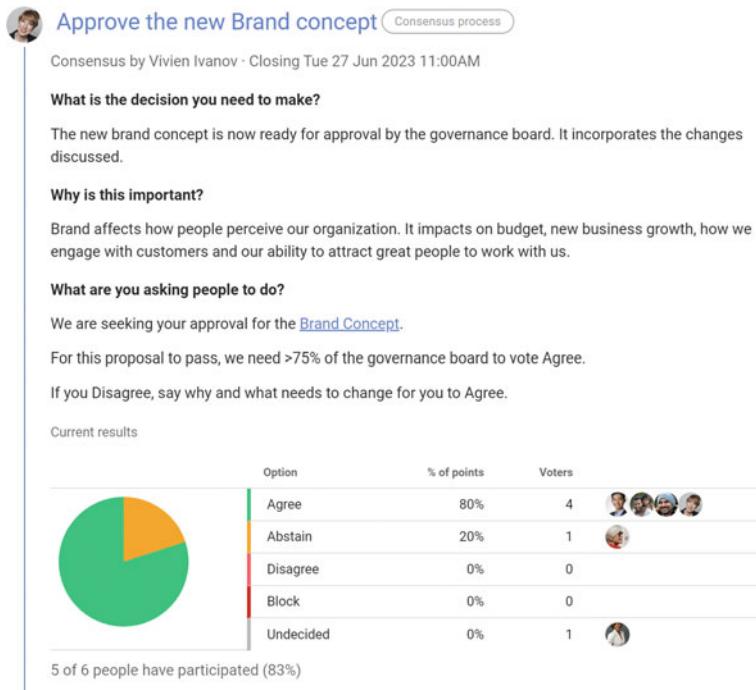


Fig. 3.9 Sample decision-making conducted within Loomio. Source <https://www.loomio.com/>

and—most importantly—make collective decisions. An initial discussion, if one occurs at all, concludes when someone transforms it into a concrete proposal. This proposal then serves as a fresh invitation for everyone in the group to engage. It is quite common at this point for those who had not participated in the initial discussion to return and express their opinions on aspects of the conversation that may have been overlooked or understated (Bartlett & Deseriis, 2016).

From this point onward, both the discussion and the decision-making process are consolidated in the same location. The screen is divided into two sections: one for discussing the issue, akin to a blog post, and the other for voting on a specific course of action to address the issue. The voting system draws inspiration from the hand signals employed by Occupy activists during General Assembly meetings. When voting, individuals have the opportunity to provide a brief summary, similar in length to a Twitter post, explaining their voting rationale. Furthermore, after engaging in the discussion, participants can modify their votes, provided they do so prior to the proposal's established deadline (Finley, 2014).

In the current version of Loomio, decision-making can be implemented in one of the four distinct methods, each designed to cater to various needs and contexts within an organization. The first method, *Advice*, is centered around gathering insights from individuals who are either directly impacted by the decision or possess specialized expertise. This approach not only ensures that the decision is well-informed but also

inclusive of different perspectives. In this method, participants are given the option to Agree or Abstain, allowing for a straightforward expression of support or a neutral stance.

The second method, *Consent*, is tailored for expedited decision-making processes. It is particularly useful when a decision needs to be made swiftly without significant objections, adhering to a *safe-to-try* principle. This method facilitates prompt and efficient resolutions, reducing the time spent in prolonged deliberations. It incorporates a two-stage voting process, offering options like Looks good, Abstain, Concerned and Undecided in the first phase, and Consent, Objection, or Undecided in the final phase, thus allowing participants to clearly express their stance on the proposed decision.

The third method, *Consensus*, is akin to Loomio's original voting system. It emphasizes a consensus-oriented approach, allowing participants to collectively reach an agreement that is acceptable to all. This method is particularly significant when the decision impacts all members and where a unanimous or near-unanimous agreement is preferred. The available voting options in this method include Looks Good, Abstain, Concerned, and Undecided in the first phase, and Agree, Abstain, Disagree, Block, or Undecided in the second phase. The inclusion of a *Block* option is crucial, as it gives participants the power to halt a process they believe should not be implemented, thereby ensuring that all voices are heard and considered.

Lastly, the fourth method, *Simple decision-making*, is an ideal starting point for those who are new to collaborative decision-making processes. This method is designed to be user-friendly and less complex, making it accessible to a wider range of users. It simplifies the decision-making process while still ensuring that all participants have an opportunity to express their views and contribute to the final decision.

In all the methods described above, decisions are made dynamically, as participants' opinions are visualized in real time. Simultaneous ongoing debate is possible, which can influence the views of those who have not yet voted or allow those who have already voted to change their stance, a process that is also visually represented. Disagreements are depicted through a pie chart, demanding attention to facilitate the resolution of concerns. This distinction sets Loomio apart from polls and other voting mechanisms: it allows for the alteration of one's position during the discussion, turning it into a collaborative game where participants collectively address concerns and seek to bring about changes (Bartlett & Deseriis, 2016).

Loomio's premise, which focuses on decision-making combined with ongoing argumentation to persuade others and approach consensus, is most effectively realized within small to medium-sized groups. According to Bartlett, this type of participation makes sense when dealing with a genuine community rather than a crowd: "I personally do not believe in large groups, I only believe in small groups (...) groups of a couple of hundred people" (Bartlett & Deseriis, 2016). This belief is rooted in Loomio's experiences, including its collaboration with the Spanish political movement Podemos. At one point, participants and supporters of this movement generated nearly 40% of the activity on the entire platform, creating an extensive network of local groups (approximately 2,500 groups). Consequently, an attempt was made to create a "Loomio of Loomios to scale direct democracy from the local to the national

(...), an index of all the discussions and the decisions that were happening within the network.” However, it became evident that the leadership of Podemos was not interested in sharing power with the grassroots structure formed in this manner. This demonstrated how large-scale digital democracy, despite being verbally championed, remains distant from political realities (Bartlett & Deseriis, 2016).

3.3.3.3 Collective Cognitive Processes in the Project

In my evaluation of the CI processes within Loomio, adhering to the methodology outlined in Sect. 3.2.2., I employed a qualitative case study approach once again. Initially, I focused on desk research to comprehend the project’s context and the key factors influencing its CI processes. This involved reviewing Loomio’s website and relevant literature, with a primary focus on R. Bartlett’s interview titled *Loomio and the Problem of Deliberation. Open Democracy* (Bartlett & Deseriis, 2016), and the *Wired* article “Out in the open: Occupy Wall Street reincarnated opensource software” (Finley, 2014). Then, I conducted an In-Depth Interview (IDI) with Robert Guthrie, the current Loomio manager. The table below summarizes the statement from IDI and from the above-mentioned press articles. Direct quotes from IDI are in quotation marks referenced as [RG]. The quotes from the press interviews and articles have their sources indicated (Table 3.9).

3.3.3.4 Metacognition: Understanding Cognitive Processes in Loomio

This section offers an analysis of the meta-processes shaping cognitive processes within Loomio. The information was gathered from an In-Depth Interview with Robert Guthrie, the current Loomio manager, supplemented by my analysis derived from empirical research and a review of the literature. All quotations sourced from the interview are marked [RG]. The central question addressed here is: What is the most critical factor influencing the cohesion of groups working in Loomio and their capacity to collaboratively create outcomes that are universally accepted? The answer can be summarized as “a toss-up between common purpose and relationships” [RG].

Contrary to large-scale social media platforms, Loomio thrives on small, high-trust groups that are successful. Groups that do not succeed simply fade away from the platform, ceasing to use its services. Therefore, in analysis, the focus can be on the characteristics of those groups that operate effectively. “A lot of groups come through and start, and 90% of them won’t be back the next year. And so, I see many groups that will come along and form and then realize that they haven’t quite got the right magic ingredients to be a group” [RG].

Loomio exemplifies how debates in small groups can be significantly more effective than those in larger ones, as smaller groups are bound by interpersonal relationships and a consensus on fundamental values, which helps to avert conflicts over crucial issues. This sense of unity creates an environment conducive to productive discussions. Loomio is effectively described as a facilitation tool suitable for groups

Table 3.9 Cognitive processes in the Loomio project

Cognitive process	How does it work in the project?
Collective sensing	<ul style="list-style-type: none"> • In Loomio, the process of gathering information can be likened to a brainstorming session. “I would call it a brainstorming situation. (...) People certainly do it with Loomio” [RG]. This dynamic approach encourages participants to freely share ideas and insights. As users engage with the platform, they collectively contribute to a rich pool of knowledge and perspectives. This collaborative approach is indeed a common practice among Loomio users, fostering a creative and inclusive space for problem-solving and decision-making • “Divergent and convergent conversation (...) is bringing together diverse perspectives” [RG]: This kind of debate involves gathering a wide array of perspectives, ideas, and insights. By initially encouraging divergent thinking, participants are free to explore and present various viewpoints, fostering an environment rich in creativity and diverse problem-solving approaches. This is crucial for capturing a comprehensive range of ideas, which ensures that multiple aspects of a problem are considered • In the case of controversial decisions to be made, the preliminary phase, known as a sense check, plays a crucial role. This can be described as a collective sensing of the problem. It’s a non-binding phase where the issue is explored more deeply and related information is gathered. Its purpose is to identify potential objections, ensure a safe environment for expressing disagreements, and refine the proposal to minimize objections later on. This includes preliminary voting as a way of gauging, “I think this is the decision we want to make. What do others think about it?” [RG]. The goal is to maintain group cohesion and the sense that consensus is achievable, essentially preparing for a decision. “It can be really upsetting to move to a decision before people are ready. So, in terms of preserving the energy of the participants and making it a nice experience to make decisions, it is a way to involve people without subjecting them to heightened pressure” [RG]
Problem solving	<ul style="list-style-type: none"> • In Loomio, the central mechanism for CI processes is a proposal put forth for collective voting. “Our main tool is the proposal. And the proposal guides people through the whole process” [RG]. The proposal serves as the focal point around which the entire problem-solving process revolves, tightly connected to the specific issue at hand. The group’s deliberation and decision-making are thus directly linked to these proposals, making them a critical aspect of the platform’s collaborative approach. “The value of Loomio is that the deliberation and the conclusion are displayed side by side” (Bartlett & Deseriis, 2016) • The process of problem-solving is directed toward discovering a method to converge on a set of statements or principles that all participants can agree upon (Finley, 2014) • In the community, members engage in “elaborating on diverse perspectives and providing clarity on them” [RG]. Concurrently, they transform these insights into a cohesive proposal. This method promotes conciseness and the addressing of a precise, well-defined question. It encourages participants to reflect and respond based on the evolution of their understanding “[of] how their views might have changed through the course of the discussion” [RG]

(continued)

Table 3.9 (continued)

Cognitive process	How does it work in the project?
	<ul style="list-style-type: none"> • Focusing attention and eliciting energy from people for engaged decision-making requires significant effort, which is why Loomio is typically not used for everyday work. The process of problem discussion must concern matters of particular importance to the collective, not every minor decision. For a discussion to be substantive, it should take place within a well-facilitated topic. The role of a facilitator is pivotal in synthesizing and integrating the various points presented by the participants • When participants formally announce moving to a proposal phase, it signifies a shift in focus, alerting others that a decision-making process is underway and their contributions are needed within a set timeframe. This transition provides a clear marker for engagement. As the proposal is put forth, the number of participants actively involved in the ongoing problem-solving process typically increases significantly (Bartlett & Deseris, 2016) • Announcing the commencement of the decision-making phase stimulates interest in the discussion and facilitates greater involvement in the ongoing problem-solving process. People often provide reasoning with their votes, effectively distilling the entire discussion into a course of action. This encourages participants to present the essence of their arguments as concise, clear statements, linking the debate to the decision-making process. “It’s much more easy to read proposal and the corresponding votes than to digest an entire discussion” [RG] • Regarding conflicts, well-facilitated discussions focused on a clear topic often make disagreements highly valuable. They aid in clarifying intentions, prompting reconsideration of viewpoints, and illuminating areas where someone’s perspective may not have been fully considered. However, when discussions shift to more principled or high-level topics, managing conflicts can become challenging. In such cases, inherent conflicts within the group may surface, revealing fundamental disagreements among participants • In the context of serious conflicts within a group, Loomio is not a panacea. However, for minor disagreements or misunderstandings, Loomio proves highly beneficial. It provides a platform where conflicts can be resolved, thanks in part to the smaller size of discussion groups and a focus on cooperation that fosters trust and relationships among participants. This contrasts with the often-impersonal nature of social media. With these foundations of trust and relationship, there is a solid base to begin resolving conflicts
Decision making	<ul style="list-style-type: none"> • Loomio’s managers assert that their initial inspiration was “hard-coded to consensus decision-making, particularly the type employed within the Occupy Movement. That’s the origin of the options to agree, abstain, disagree, and block. However, the options have evolved to be more flexible now. This is because we don’t want to dictate how people make decisions; the process is inherently fluid and dynamic” [RG] • During the decision-making process, the screen is divided into a visualization of the decision-making status and the justification for the votes cast (Finley, 2014) • Loomio currently offers three primary decision-making styles: advice, consent, and consensus. Users can adapt various decision-making approaches to suit their group’s needs, modifying and saving them for future use. In certain organizations or companies, the advice process is commonly employed. This method “gives a lot of power to the person holding the decision process, but it allows transparency and feedback through the process” [RG]. Meanwhile, in other organizations, consent or consensus methods are more favored

(continued)

Table 3.9 (continued)

Cognitive process	How does it work in the project?
	<ul style="list-style-type: none"> During the voting process, participants have the freedom to change their minds. This flexibility is crucial in establishing a shared understanding and reaching a possible consensus. The process of engaging in discussion and then iterating through multiple rounds of proposals can be extremely effective in consolidating views and uniting the group “around a common set of principles or a unified course of action” [RG] Decision-making is not merely the final stage of the thinking process; rather, it’s the focal point around which all other processes are centered. People have the opportunity to change their minds and introduce new arguments, contributing to a dynamic and evolving deliberative environment
Collective memory	<ul style="list-style-type: none"> Collective memory undoubtedly exists, yet it is significantly influenced by the individual behavior of each community member. “Do they reflect on the past to enhance their future decisions? Yes, they do. However, I don’t perceive everyone in a group as equal at engaging in that process” [RG] Utilizing collective memory in making new decisions or initiating fresh discussions within a group involves an innovative approach while referencing the past. Loomio archives all community discussions and decisions, providing a reference point. “The shape of that collective memory will vary depending on who is responsible for this process. Ultimately, it’s a highly subjective thing” [RG] The software and tools that could facilitate access to a collective’s archival discussions and decisions “are only a small part of the problem. (...) Really, if you want to make significant progress, it’s about the relationships and personal development of the group members. A purely technological approach can be disproportionate to what is required for nurturing relationships and aiding personal development within a group” [RG]. Collective work primarily revolves around relationships, effective leadership, timely participation, and individuals bringing their best selves to the situation rather than just tools, databases, and voting systems

of approximately 200 people or fewer, offering a structured yet flexible platform for collaborative decision-making: “I think that you get a lot more intelligence out of it than when you do with a very large group” [RG].

The potential application of Loomio in the policymaking cycle is also primarily associated with the decision-making phase. As has been demonstrated, all user activities in Loomio are closely related to decision-making. This direct link to decision-making makes Loomio an especially valuable tool in public policy contexts, particularly in assemblies and communities where collective decisions are crucial. In these settings, the debate is intensely focused on making binding decisions that reflect the collective will and best interests of the group. This was precisely the scenario in notable implementations like those in the National Assembly of Wales and the New Zealand Government, which utilized Loomio to develop a Workforce Capability Framework. Such instances showcase Loomio’s adaptability and effectiveness in formal governmental structures, providing a platform for structured debate and decision-making that is both inclusive and efficient. Moreover, the ability to organize

the decision-making process in various ways through Loomio—whether it's through the process of advice, consent, consensus, or simple decision-making—adds a layer of flexibility that is highly beneficial in policy contexts. This flexibility, combined with incentives to present well-thought-out arguments supporting individual positions, not only facilitates a more informed and engaging decision-making process but also positions Loomio as a useful tool in the *Policy formulation* area. This process inherently encourages a more participatory and transparent approach to policy formulation, where the collective intelligence of the group can be harnessed to craft more effective and equitable policies.

The secret to Loomio's success lies in emphasizing the dynamic decision-making process through voting as a primary feature. The proposal becomes the central element, orchestrating all the processes of collective intelligence. As such, the group's discussions and decisions are intricately tied to the proposals that are put to a vote. This makes them an integral part of the platform's collaborative ethos. By ensuring that every member's voice can contribute to the final decision, Loomio enhances the quality and acceptance of the decisions made, fostering a strong sense of commitment and unity within the group. This approach not only streamlines decision-making but also aligns it closely with the collective insights of the participants, thereby enhancing the overall effectiveness of the group.

Collective identity plays a significant role in the success of a group, but it can also be a source of conflict, sometimes unnecessarily. This issue is intertwined with the maturity of the group and its participants. For a group to exhibit collective intelligence successfully, members need to have confidence in their peers and a shared sense of value in the topics they discuss. Mutual respect and common purpose are foundational to creating a productive and cohesive collaborative environment. This dynamic tends to function well in groups where members are familiar with each other. The success and longevity of a group are often tied to common ideas or values that bind them together, fostering cohesion and unity.

3.3.4 Decide Madrid

3.3.4.1 Introduction and Background

Following a series of corruption scandals in Spain that led to a decline in public trust in governmental institutions, the Madrid City Council initiated the Decide Madrid platform in 2015. This platform, also known as Madrid Decides, functions as an official and comprehensive policymaking tool for the municipality, encompassing aspects of transparency, open data, and public participation. It utilizes Consul, an open-source civic technology, to actively involve citizens in decision-making processes. Decide Madrid is characterized by its diverse participatory features, including forums for public debates, mechanisms for citizen proposals, and avenues for participatory budgeting. These components collectively enhance civic engagement and foster a more transparent and collaborative approach to municipal governance. Today, Decide

Madrid represents one of the most successful cases of e-participation platforms in the world, having almost half a million citizens registered (Table 3.10).

3.3.4.2 Project History and Key Features

The global financial crisis, governmental austerity policies, and various corruption scandals significantly eroded public trust in politics in Spain. In response to this decline in Madrid, political movements focused on enhancing civic participation coalesced to form the coalition *Ahora Madrid*, which governed the city from May 2015 to June 2019. This coalition, recognizing the importance of public involvement, pledged in its electoral program to “implement tools for citizen participation through the Internet” (Royo et al., 2020), a commitment fulfilled through the creation of the Decide Madrid platform. Furthermore, the Madrid City Council, as a participant in the Open Government Partnership’s Subnational Government Pilot Program since 2016, has committed to developing participatory budgets and collaborative, efficient legislative mechanisms. This commitment also included expanding policies for citizen participation, as outlined by Ostrow (2017), thereby reinforcing Madrid’s dedication to fostering a more open and participatory government.

Decide Madrid, initially launched in 2015 and fully operational by 2016, is a citizen participation initiative aimed at actively involving the residents of Madrid in the local government’s decision-making processes. This platform empowers citizens to propose policy ideas, engage in debates on municipal issues, participate in budget allocation decisions, and contribute to the legislative process. It is equipped with various integrated features, including direct voting on specific topics, discussion forums, and collaborative tools for drafting regulatory texts. This allows Madrid’s residents not only to create and vote on a range of city-related proposals but also to engage in in-depth discussions about them. The platform enables citizen participation at multiple stages of the policy cycle, as described by Royo et al. (2020). These stages include setting the agenda, analyzing and preparing policy, formulating policy, and, to a certain extent, monitoring policy implementation. However, it’s important to note that the scope of topics available for discussion and decision on the platform is confined to those within the jurisdiction of the Madrid City Council. This boundary ensures that the focus remains on areas where local residents’ input is most relevant and impactful.

Table 3.10 Basic information about the Decide Madrid project

Project name	Decide Madrid
Owner	Madrid City Council
Web address	https://decide.madrid.es
Location	Madrid, Spain
Years in operation	2015–present
Implementation level	Local, global

The comment architecture of the platform is reminiscent of Reddit, where comments can not only be posted but also evaluated and further discussed, leading to the formation of branching threads. This design allows for dynamic and multi-layered conversations, mirroring the organic flow of discussions seen in popular online forums (Medialab Prado, 2019).

The Decide Madrid platform is accessible to all users, regardless of registration status, but the extent of participation varies depending on the type of activity. While associations, NGOs, and companies, alongside individual users of any age, can register on the platform, engage in creating debates or proposals, and comment in various sections, there are specific participation privileges reserved for registered individual Madrid citizens aged 16 and above. These registered individuals have the option to verify their accounts, which then grants them the ability to create proposals for participatory budgeting, as well as to support and vote on proposals. This tiered approach to participation ensures a broad base of engagement while reserving certain decision-making capabilities for verified local citizens, thereby maintaining a focused and representative input in municipal matters.

Participation on the platform is organized into four main sections: (1) debates, (2) proposals, (3) processes, and (4) participatory budgeting. The debate section is the most loosely structured. It takes the form of an e-forum, which enables users to initiate discussions, contribute to existing topics, and express their opinions through agreement or disagreement. This interactive space is accessible not only to citizens but is also employed by the city council to initiate discussions on relevant issues.

A significant part of the platform is the Proposals section. Here, users can submit proposals, often accompanied by audio-visual materials and supporting documents, to garner community support. Remarkably, proposals that receive backing from at least 1% of Madrid's residents aged over 16 (app. 27 thousand people) are brought to a vote, demonstrating a tangible impact on decision-making. The proposals can span diverse areas such as transportation, the environment, urban planning, social rights, healthcare, education, and culture. The process operates through three key stages: submission, endorsement, and voting. In the submission stage, any individual can initiate a proposal. The endorsement stage follows, where the primary objective is to highlight the most significant and relevant suggestions. Proposals that achieve the backing of 1% of eligible residents of Madrid (aged over 16, app. 27 thousand people) are approved. Prior to moving to the final stage, there is a 45-day period during which approved proposals are open for comment and discussion on the platform. During this period, proposals that fail to garner sufficient support are discarded and filed away. In the final voting stage, which lasts a week from the proposal's approval date, eligible residents vote on each approved proposal. The vote can be conducted through postal ballot, in-person voting, or digitally via the Decide Madrid platform. For the proposal to be approved, it only requires a simple majority, and there is no quorum necessary. As of now, only two out of the 26,000 citizen proposals submitted have met the necessary threshold and were subsequently approved by a significant margin, receiving support from approximately 8% of registered voters. Proposals that fail to garner the required level of support within the designated timeframe of one year are regularly archived (Medialab Prado, 2019).

Furthermore, the platform features Processes, a versatile tool employed by the city council to solicit diverse inputs for various purposes, such as developing or amending regulations or soliciting proposals for activities. The methodology for these processes is adaptive, ranging from open debates to the submission of text documents for citizen amendments (Royo et al., 2020). The most noteworthy process undertaken through the platform was the reform of Plaza de España, complemented by the subsequent renovation of eleven other squares across various districts. In addition to these major projects, the platform has also facilitated decision-making on smaller-scale issues. These include the design of public benches, the allocation of walls for street murals, and the selection of films for local archives, illustrating the platform's comprehensive scope in addressing both significant urban developments and smaller community enhancements (Medialab Prado, 2019).

Perhaps most notably, Decide Madrid includes a participatory budgeting mechanism. This innovative feature empowers citizens to have direct participation in the allocation of a portion of the city's budget for the subsequent year (app. 100 million EUR). The scope of these projects is extensive, covering city-wide initiatives or focusing on specific district-level needs encompassing current expenditures, subsidies, and public investments. Citizens are given the opportunity to vote on projects that affect the entire city or just the district of their choice, thereby integrating local perspectives into broader municipal budgeting decisions.

An intriguing aspect of the voting process is the *shopping basket* method, which operates similarly to online shopping. Participants select projects consecutively, and with each selection, the cost of the project is tallied to a running total displayed on the screen. This interactive approach allows participants to keep track of their choices, adding projects up to the limit of the allocated budget for their district or the city. Essentially, each project that fits within the total budget receives one vote. The projects garnering the most votes are selected, provided they stay within the available budget. Projects exceeding the budget are not considered. This method tends to favor cost-effective projects (Medialab Prado, 2019). Following the completion of the voting process, the proposals that receive the most votes are directly incorporated into the draft of the general budget for the city of Madrid.

The participation statistics for the platform are indeed remarkable, demonstrating a high level of civic engagement. Over 430,000 users have actively engaged with the platform, contributing to its dynamic nature. These users have collectively submitted in excess of 26,000 proposals, showcasing a wide array of ideas and suggestions for the betterment of the city. Additionally, the platform has facilitated the submission of around 125,000 comments, reflecting the depth of discussion and deliberation among its users. Perhaps most notably, these participants have cast more than 3 million votes of support, underscoring the vast scale of interaction and the strong interest of Madrid's citizens in shaping their city's policies and projects (Arana-Catania et al., 2021). In the recent years of Decide Madrid's operation, an average of 100,000 users took part in the annual edition of participatory budgeting.

The Decide Madrid project is conducted under the management of the General Directorate of Citizen Participation (GDCP), which is directly accountable to the Mayor of Madrid's Office. The development, implementation, and operational costs

of the platform are covered by the Madrid City Council's budget. In addition to this, Decide Madrid benefits from funding as part of a collaborative open government project with three other Spanish city councils. This project, supported by FEDER grants, aims to enhance the platform and develop new modules. A key collaborator in this endeavor is Medialab Prado, an enterprise owned by the city council that functions as a laboratory for citizens. It is here that several innovative projects linked to decide Madrid are being developed, pushing the boundaries of civic engagement and technological application in governance.

While Decide Madrid is dedicated to the city of Madrid, the open-source software behind it, Consul, has seen broader adoption. It is currently implemented or in the process of being implemented in over 100 organizations worldwide, predominantly in Europe (especially Spain) and Latin America. Information about these global implementations can be found at <http://consulproject.org>. The open accessibility of the Consul code online allows any organization, whether public or private, to adapt the platform for its own unique requirements. The entire Decide Madrid initiative and the broader Consul project have gained international recognition, notably the 2018 United Nations Public Service Award (Royo et al., 2020). This accolade underscores the project's significant contribution to enhancing public service through innovative citizen participation and digital governance models.

3.3.4.3 Collective Cognitive Processes in the Project

In evaluating the CI processes within the Decide Madrid project, I adopted a methodology similar to the one used in previous studies, as detailed in Sect. 3.2.2. The study was structured in two stages, utilizing a qualitative case study approach. The first stage involved comprehensive desk research to understand the context of the project and the key factors influencing its CI processes. This involved a thorough content analysis of the Decide Madrid website, relevant documents, official reports, statistics, and data compiled by third parties. An instrumental resource in this stage was the study by Royo et al. (2020), which included nine semi-structured interviews with politicians, civil servants, and platform users. Additionally, two more significant publications were reviewed: the *Future Democracies. Laboratory of Collective Intelligence for Participatory Democracy* report by Medialab Prado (2019) and the research paper “Citizen Participation and Machine Learning for a Better Democracy” by Arana-Catania et al. (2021), which discussed the application of Natural Language Processing (NLP) and machine learning in reducing informational noise in Decide Madrid. The second stage of the study entailed conducting In-Depth Interviews (IDIs) with two experts intimately familiar with the platform. The first expert was Miguel Arana-Catania, who previously served as the Director of Participation for the City Council of Madrid and the Project Director for Decide Madrid. The second expert was Yago Bermejo Abati, who worked as the project manager of the participation laboratory at Medialab Prado from 2016 to 2019. The interviews were structured around a pre-defined script aligned with the evaluation framework,

including supplementary questions to gain deeper insights into the experts' perspectives and experiences. The final phase of the study involved analyzing the gathered information, drawing out key insights and areas of agreement or disagreement from the literature and interviews, and synthesizing these findings into a comprehensive qualitative analysis. Table 3.11 summarizes the statements of interview participants. Direct quotes from interviewees' statements are in quotation marks, marked [MAC] for Miguel Arana-Catania and [YBA] for Yago Bermejo Abati.

3.3.4.4 Metacognition: Understanding Cognitive Processes in Decide Madrid

Following the previous case studies, this section analyzes the meta-processes that affect cognitive processes in the Decide Madrid platform. This encompasses a summary of the crucial findings from the cognitive process analysis, an exploration of how these processes interlink to create a whole, and an examination of the factors that motivate people to engage in these processes. The data was collected based on In-Depth Interviews conducted with two experts: Miguel Arana-Catania, previously the Director of Participation for the City Council of Madrid and the Project Director for Decide Madrid, and Yago Bermejo Abati, who worked as the project manager of the participation laboratory at Medialab Prado from 2016 to 2019, as well as my analysis based on empirical research and literature review. All quotes from the interviews are marked [MAC] for Miguel Arana-Catania and [YBA] for Yago Bermejo Abati.

The main services available on the platform are flexible and cover a broad understanding of shaping public policies. The debating section primarily pertains to *Agenda setting*, while the proposals and processes sections can have a wide application in policy formulation, with participatory budgeting joining these areas in decision-making. As with many other projects, policy evaluation is probably the most challenging to conduct directly. However, the substantial amount of interconnected data from various stages of creating local policies and their existing links can also facilitate this process.

Regarding the values that unify the community, "the norms or values that bring together the entire community in Madrid are evident when examining the winning ideas. The people feel that they should do something together with the important problem for all of them" [MAC]. For instance, when examining the projects from the last completed participatory budget (2022), it's clear that for city-wide initiatives, the majority of funding was directed towards recycling projects. In contrast, in district-level projects, the largest allocation of funds was directed towards cultural institutions and libraries. This reflects the community's prevailing priorities, which encompass both the recovery of raw materials and the promotion of culture and literature (Presupuestos Participativos 2023).

When reviewing user activity, it is observed that although the majority of users participate in a limited manner, there are also groups with strong identities and shared values. These groups typically form outside of the platform through natural

Table 3.11 Cognitive processes in the decide madrid project

Cognitive process	How does it work in the project?
Collective sensing	<ul style="list-style-type: none"> Regarding gathering information, there are very different types of users. DM generally gathers a large number of users. (1) Most do not create any ideas or proposals; they just review the existing ones and sometimes vote. (2) Some people just “come, write something and go away” (...) “They maybe doesn’t look at the other ideas. They don’t interact so much in the process.” (3) “And there is a small minority who are the ones that are very, very engaged” [MAC]. The last group contributes the most to collective sensing and shares different perspectives The large number of project participants translates into many publicly relevant topics that can be identified. “There are many people who just come with a general idea, not organized, not very well specified. This is very common. In some cases they are really experts who come to propose ideas. So they are, for example, collectives, organizations that are working around one specific topic, including architects, designers and similar” [MAC]. It often happens that non-professionals with a very general description of the problem hit the nail on the head and gain a lot of interest “The diversity is happening in practice a lot because we have a lot of participation. (...) When you have such large amount of people, diversity can essentially happen naturally, because there are so many interests presented” [MAC]. “Different interest groups provide (...) [diverse information], and have dialogues to foster common ground agreement” [YBA]
Problem solving	<ul style="list-style-type: none"> The platform was intentionally designed to facilitate collaboration among people, “but that is, of course, difficult” [MAC]. When users submit proposals, anyone can indicate if a proposal is similar to another, regardless of who authored it. This feature allows users to easily link two proposals they deem similar. Consequently, when reading a specific proposal, any user will find a list of related proposals below it. This design was implemented to enable people to effortlessly discover pertinent information without searching through each proposal “There are two levels of debate for each proposal. First, there is a general discussion at the bottom of the proposal, where everybody can talk. But besides that, in each of the proposals, there is a dedicated community space, where people are invited to collaborate around the proposal” [MAC] One of the latest innovations is the introduction of collaborative proposals powered by AI and NLP (Natural Language Processing) technology. Contributions are clustered based on their content, forming communities of patterns—groups of people focused on specific topics. Thanks to this approach, interest groups emerge organically from among unorganized citizens, fostering the creation of bottom-up proposals The role of local authority representatives in this process is not to evaluate the substance of the projects but to estimate their costs, irrespective of their substantive merit. In participatory budgeting, the government is responsible for assessing their feasibility and estimating the associated costs after the initial pre-selection of ideas. “We [as project managers] try to avoid much interference from the government because sometimes the government tends to take over the idea from the citizens” [MAC]. The aim is to prevent governments from overshadowing or taking over citizen-generated ideas

(continued)

Table 3.11 (continued)

Cognitive process	How does it work in the project?
	<ul style="list-style-type: none"> The large number of individuals naturally promotes a continuous stream of innovative ideas: “when we have so many people in the group, the creativity always happens. (...) I mean, naturally, you can’t really stop people being so creative” [MAC] “There are different perspectives on what is democracy” [YBA], so people collaborate in different ways—some engage more in debates, others less Groups of citizens with a strong identity are evident on the platform. These are frequently the most active groups, embodying the principle that “strong identity and working toward the common good are not contradictory” [YBA]
Decision making	<ul style="list-style-type: none"> People consider voting as a way to “change the society. (...) Change the way how decisions are made” [YBA] Before the voting, there is an extensive information phase lasting about two or three months. During this time, the government publishes the ideas considered for a vote and facilitates debates. Through functionalities provided by Decide Madrid, each proponent can organize an informational campaign to promote their project. This includes managing the campaign, maintaining communication with supporters, and creating a “crowdsourcing campaign around the proposal” [MAC] There are several stages in selecting the best ideas through voting. The general principle is that the first phase is designed to gauge overall support for the ideas. “You can just support any idea, or as many ideas as you want” [MAC]. There are no limits to the number of ideas that participants can endorse. The subsequent phase (in the case of citizen proposals) involves achieving a support threshold from all city residents at a minimum of 1% In participatory budgeting voting, there are ideas for the city as a whole and ideas specific to one district. Participants can vote for as many ideas in each of these categories as they wish. However, the total cost of the ideas they support must be equal to or less than the total budget allocated for the city If an idea is successful in the final voting, the government contacts its authors and discusses how to implement it. The authors are invited to collaborate with the government in making the idea a reality. They also have the opportunity to track its progress and oversee its realization According to some participants, this process differs from the theory of direct democracy because it is not entirely representative. Consequently, it may not be ideal for making major, significant decisions. As one opinion states, “I think it is not so good for large-scale direct democracy, but for small-scale distributed decisions, it is quite efficient” [YBA]
Collective memory	<ul style="list-style-type: none"> The platform offers straightforward access to all past ideas, comments, and discussions. The primary reason for this is transparency: “to make it clear to understand what is happening during the processes” [MAC]. Users can examine contributions from previous years, including each year’s statistics, ideas, resolutions, and all associated information Despite these possibilities, “the great majority of people never look at the previous participation process. Never. Because, as mentioned, most people just come to pick up one or two ideas, they don’t really care much about what happened in previous years. People are submitting proposals and even forgetting about them. Most of the engagement occurs only once” [MAC]. The exception is those who participate but are not successful, and they come back with their idea the next year. They review past records. “But, to be honest, this is a small minority” [MAC]

(continued)

Table 3.11 (continued)

Cognitive process	How does it work in the project?
	<ul style="list-style-type: none"> • Collective memory remains, therefore, one of the most difficult problems. “All the information is there, and it’s easily accessible, but I don’t think that it is making a big difference in terms of reusing the ideas” [MAC] • The prospects for developing collective memory are intriguing, particularly due to the current introduction of new functionalities based on large language models. The key question is, “How can a language model be continuously fed with people’s contributions to eventually create a model of a particular society?” [YBA]. In previous years and projects, a large amount of data was lost, so it would be wonderful to better organize this for the future

human interactions and join as already organized collectives. However, participation in online activities allows for the expansion of the engaged group and the creation of new social bonds focused on addressing societal issues, “connecting with people that feel in the same way that you, finding people on the basis of common interests, to satisfy social needs” [YBA].

Interestingly, strong opinions and biases are not seen as barriers to participation but as components of active involvement: “It is normal for people to have biases in society, so the platform is not going to resolve these biases in any way. People simply follow their established patterns of thinking” [MAC]. Groups with a strong identity appear to maintain a collective memory of their achievements outside the platform, independent of e-participation. However, for the great majority who are only interested in a single idea, submitting a proposal and then forgetting about it, the problem lies in the lack of access to this collective memory, which would facilitate the use of solutions developed in previous years.

The citizen activity on Decide Madrid is often linked to criticism of the current situation: “Because you are concerned, you want to change (...) you want to change society, how decisions are made. So, you are active (...) to offer alternatives to what the government is doing, or to bring new perspectives on doing things differently” [YBA]. Particularism is not predominant, but one can observe many examples of consideration for the entire community. Most winning, selected, and implemented ideas are intended for the whole community, not just one group of people or a specific area. An example of this approach is: “not just creating a small park in one neighborhood, but essentially creating more parks throughout the city” [MAC]. There were a significant number of ideas impacting very small collectives but considered by the majority as related to socially important problems – for example, the victims of violence, homeless people, etc. The projects deemed socially significant for the entire community have garnered substantial support.

The platform’s designers intentionally incorporate the cognitive processes of information gathering, problem-solving, and decision-making into its functionalities: “When we design platforms, we always try to include all these features, taking into account how people reason and interact, and facilitate all these positive behaviors and interactions to enable collective intelligence” [MAC]. However, “there were

many features that we introduced, designed by experts in the field of participation and collective intelligence, which were not really used.” Why? Because “if you are truly successful in engaging a large number of people, they already have their collective intelligence self-organized” [MAC]. They form groups independently, not necessarily following the intended design. More important is to observe how people actually act and facilitate that, rather than imposing solutions from above. “The ideas and proposals that were winning in Decide Madrid are amazing. But it wasn’t because we tried to help them to be more intelligent. They brought their own intelligence; our role was just to provide them a fair platform for communication and a fair process” [MAC].

According to the interviewee, “This is why many other participation platforms fail in harnessing collective intelligence, because the design of [their] platforms is very constrained” [MAC], imposing a certain way of thinking or patterns of behavior on people. In this manner, the collective intelligence of people is not utilized effectively; they are confined within something very specific and limited. A more flexible structure for responding to what is happening online would be beneficial.

3.3.5 *Deliberatorium*

3.3.5.1 Introduction and Background

The Deliberatorium is a unique project compared to other case studies discussed here, as it was created in a research center, specifically at the MIT Center for Collective Intelligence. Since 2007, it has been developed under the supervision of Mark Klein, a research scientist at this Center, and has been utilized in numerous public debate projects and commercial ventures. The platform has been used to mediate complex deliberations in various contexts, from universities, political associations, and companies like Intel and BP to public entities like the US Federal Bureau of Land Management.

The primary objective of the Deliberatorium is to facilitate online collaborative argumentation. In this context, all processes associated with collective intelligence are centered on optimally mapping arguments and maximizing the value extracted from debates through interconnected sets of arguments. The central research question guiding the development of this project is how an argument-centric approach can be effectively scaled to medium and large-scale applications (Fig. 3.10).

The Deliberatorium leverages a Collaborative Computer-Supported Argument Visualization (CCSAV) method, as outlined by Kirschner et al. (2003). CCSAV tools categorize content by topic and present it in a structured format, allowing users to view the entire discussion and engage where they see fit. This collaborative process culminates in a shared representation of collective knowledge and preferences, manifested here as an argument map. These maps visually outline the structure of an issue in informal logic, representing discussions as networks of alternative positions, along with supporting and opposing arguments for each idea. Each issue, idea,



Fig. 3.10 Sample argument map created in Deliberatorium. Source <https://deliberatorium.org/>

or argument is uniquely represented to minimize redundancy, enhancing the ease of identifying new or existing contributions. Given a holistic view of the conversation, participants can more effectively concentrate on areas of interest or relevance. The components of deliberation are depicted as trees or networks, with nodes symbolizing the fundamental elements of the argument: questions, propositions, claims, and evidence. Visual connections between these elements denote argumentative relationships, such as support, criticism, sub-issues, and alternatives. Aimed at harnessing collective intelligence, argument maps encourage participants to focus on reasoned arguments rather than personal attacks, reduce superficial, repetitive, or harmful content, and promote the integration and synergy of ideas (Klein et al., 2023).

In the Deliberatorium, we witness this concept's practical application through online debates. Simultaneously, multiple users are encouraged to create a network of posts organized into an argument map. Each post should represent a distinct issue, idea, or pro or con argument, avoid replicating points already made elsewhere in the map, and be logically connected to the relevant post. There is also a whole set of graphic symbols illustrating various aspects of the debate. This approach allows for the identification and mapping of not only the most popular opinions in the entire debate but also more *out-of-the-box*, yet potentially valuable, points (Klein, 2011b) (Table 3.12).

Table 3.12 Basic information about the Deliberatorium project

Project name	Deliberatorium
Owner	Mark Klein, MIT
Web address	https://deliberatorium.org/
Location	Cambridge, MA, United States
Years in operation	2007–present
Implementation level	Global

3.3.5.2 Project History and Key Features

The development of the Deliberatorium was influenced by communication issues observed in social media and wiki platforms. Users of social media systems often form groups with similar opinions, limiting their exposure to a subset of potentially relevant issues, ideas, and arguments. In contrast, the Deliberatorium enables users to view diverse threads of argumentation, visualized in a tree-like structure and symbols, allowing them to engage with the entirety of a debate rather than just selected threads. Social media technologies, therefore, often capture only a fraction of a community’s collective wisdom, and a significant effort is typically required to distill this wisdom from the expansive, noisy debate to inform better, more widely supported decisions. Regarding consensus, platforms like Wikipedia usually capture a *least-common-denominator* consensus among many authors, with any non-consensus elements likely being edited out by those who disagree (Klein, 2007). In the Deliberatorium, the emphasis is placed on capturing the multiple threads of the debate, and no arguments-fueled positions are excluded from the discussion under the pretext of seeking consensus. The emphasis is on the diverse arguments that emerge during the debate, aiming to capture not only the winning or dominant opinions but also the various threads of argumentation.

Within a community, members contribute to a deliberation map, a tree-structured network where each post represents a unique issue (a question to be answered), an idea (a possible answer to a question), or an argument (supporting or opposing an idea or another argument). Contributing to a deliberation map involves two primary activities: (a) breaking down the contribution into its constituent issues, ideas, and arguments, and (b) identifying the appropriate location for these elements within the map. The structure of a deliberation map resembles a tree. To correctly position a post within this tree, one must select the appropriate top-level branch, followed by the right sub-branch, and so forth, until the precise location where the post belongs is found (Klein, 2007).

Each unique point is presented only once, significantly enhancing the signal-to-noise ratio. All posts are logically positioned under the posts they refer, ensuring that content related to a specific question is grouped together. Critical thinking is fostered as users are subtly motivated to articulate the evidence and reasoning supporting their preferred ideas. Additionally, the community is empowered to evaluate each distinct

component of an argument on its merit. This organization makes it simple to determine what has and has not been discussed on any topic, promoting more comprehensive coverage. Additionally, it counteracts polarization by placing competing ideas and arguments in close proximity to one another. Users can collaboratively refine proposed solution ideas. For instance, one user might propose an idea, another might identify an issue regarding its implementation, and a third could suggest potential resolutions for that issue. The value of an argument map extends well beyond the initial problem it was created to address. It represents a comprehensive design space of possible solutions, which can be easily harvested, refined, and recombined when similar problems arise in other contexts and locations (Klein, 2011b).

Another critical challenge is the method of extracting valuable ideas from the vast corpuses containing ideas of widely varying quality. To address this, several types of idea filtering, including crowd-based filtering, were tested in which human participants were asked to select the top ideas. In the Deliberatorium, an original filtering method, known as the bag of lemons approach, was developed. Participants are given a list of candidate ideas, given a limited number of tokens, and asked to allocate their tokens to the worst ideas (the *lemons*) rather than the best ones. Empirical evaluations on the platform have shown that this type of filtering is quicker and more accurate than traditional rating systems or the standard multi-voting approach that focuses on selecting the best idea (Klein, 2017).

The first effort to evaluate this approach was the carbon offsetting experiment, which involved translating a web forum discussion on carbon offsetting (hosted on planeta.com in May 2008) into a deliberation map. The original 13-page discussion, characterized by the digressions and repetitions typical of web forums, was transformed into one argumentation map. This was soon followed by using the Deliberatorium in a real-time public affairs debate. Conducted at the University of Naples, 220 master's students in the information engineering program were asked to contribute their thoughts over three weeks on how Italy should utilize bio-fuels. The level of user participation was exceptionally high: the debate generated over 3000 issues, ideas, and arguments, in addition to 1900 comments, possibly making it the largest single argument map ever created and the most authors for a single argument map. Another significant application of the platform was its collaboration with Intel Corporation in deliberating on how *open computing* (i.e., providing users with greater access to computing tools and data) should be implemented in the company.

In 2012, the most well-known project on the platform took place. In collaboration with Italian political scientists Paolo Spada and Raffaele Calabretta, a debate was organized involving 640 volunteers from the membership of the Italian Democratic Party on the subject of planned electoral reform. The 640 volunteers were divided into four groups of 160 participants each. Out of these, 304 completed the deliberative phase and 95 filled in the post-deliberation survey. Throughout the project, 36,000 individual actions were tracked on the platform (Klein et al., 2023).

There have been fewer examples of large-scale public debates in recent years, but the platform is constantly developing and introducing new functionalities, such as decision-making criteria. The platform is primarily used for educational purposes. However, the experience accumulated over the years from various implementations

allows us to capture the advantages of an approach specific to Deliberatorium. Firstly, it excels at identifying and mapping a wide range of opinions, not just the most popular ones. This approach includes unconventional or novel ideas that may possess considerable value despite not being initially prominent. This feature contributes to a more comprehensive range of perspectives being considered, thereby facilitating the consideration of less popular and niche views. Secondly, the platform fosters synergistic creativity. As participants view and build upon the ideas of others, this collaborative environment encourages the development of new, more refined concepts through combination and extension. This dynamic interaction often leads to unexpected and inventive solutions that might not emerge in more standard settings. Lastly, the Deliberatorium significantly amplifies the diversity of ideas. It gives a voice to the opinions and insights that might otherwise go unnoticed or be overshadowed in conventional discussion forums. Here, alternative paths of argumentation are not cut off, which allows the creation of alternative solutions and action plans.

An important aspect of Deliberatorium is the measurement of various advanced indicators related to people's behavior in deliberating collectives. Over 100 metrics have been introduced and tested, utilizing the semantics provided by an argument map. Among these, an interesting one is the measure of achieved *balkanization*, a situation where a community divides itself into sub-groups. In these groups, members agree with each other but tend to reflexively ignore inputs from other groups with which they disagree. This can be seen as a deliberation dysfunction, preventing users from fully considering all available options and trade-offs. Another issue is group-think, which occurs when a community prematurely allocates a disproportionate amount of its attention to a small subset of relevant issues, ideas, and arguments, or irrational bias, which occurs when a user rates ideas or arguments inconsistently with the ratings given to the underlying arguments. On the positive side, a valuable metric is the assessment of how mature the presented topics are enriched with relevant ideas and arguments. This can be estimated in various ways, including tree topology (more mature topics tend to have both broader and deeper structures), or activity history, which shows how argument-centric deliberations tend to evolve over time, transitioning from identifying issues to proposing ideas, presenting arguments, rating posts, and ultimately reaching quiescence (Klein, 2011b).

3.3.5.3 Collective Cognitive Processes in the Project

In my analysis of the CI processes used in the Deliberatorium, I once again utilized a qualitative case study approach, following the procedure detailed in Sect. 3.2.2. My initial step involved desk research to gain an understanding of the project's context and the primary factors affecting its CI processes. This stage included an examination of the Deliberatorium's website and relevant scholarly works, particularly focusing on the research papers by Mark Klein (2007, 2011a, b, 2017) and the chapter *From shouting matches to argument maps* by Klein, Spada, and Paulson (2023) in *The Routledge Handbook of Collective Intelligence for Democracy and Governance*. Subsequently, I carried out two In-Depth Interviews (IDIs). The first interview was

with Mark Klein, the creator and manager of the platform, and the second with Paolo Spada, a political scientist involved in the platform's development, including organizing deliberations for the Italian Democratic Party. The following table consolidates insights from both the IDIs and the aforementioned research papers. Verbatim quotes from the IDIs are indicated with quotation marks and referenced as [MK] for Mark Klein and [PS] for Paolo Spada (Table 3.13).

3.3.5.4 Metacognition: Understanding Cognitive Processes in Deliberatorium

What unites the cognitive processes implemented within the Deliberatorium is the ability to capture a large amount of diverse information from the environment and to motivate users to present a substantial number of gathered arguments and alternative lines of reasoning. This can be described as building a model of reality that is useful regardless of which arguments we agree with. It seems that the Deliberatorium excels in the process of collective sensing provided there is participation from at least a moderately sized group (e.g., 300–400 people). The conviction that the core of debate should be the rational force of a pure argument, steered by its intrinsic merit rather than the author's characteristics or its contextual presentation, roots this platform in the tradition of deliberativism, inspired by Habermas's theory. The Deliberatorium primarily finds its application in the initial and final stages of the public policy shaping process. At the outset, building maps of problems and challenges is instrumental in the *agenda-setting* phase. Here, the intricate and often overwhelming complexity of policy problems is distilled down to a selective few critical issues. This simplification is achieved by deconstructing the larger problems into smaller, more manageable, and well-structured issues. These sub-problems can then be visually and logically organized into a tree of arguments, providing a clear and comprehensive overview of the various facets and considerations involved. Moreover, this structured approach to argumentation encourages a more transparent policymaking process.

As the policy moves into its evaluation phase, the logical structure of the Deliberatorium's arguments, coupled with the detailed breakdown of problems, becomes particularly valuable. It provides a framework for retrospective analysis, allowing policymakers and stakeholders to assess the real-world outcomes of the policies against the anticipated scenarios and issues outlined in the argument maps. By comparing the actual results with the predicted outcomes and alternative strategies (if developed), a comprehensive understanding of the policy's effectiveness, unintended consequences, and areas for improvement can be obtained.

Using Deliberatorium in the *policy formulation phase* also seems reasonable when it focuses on creating a list of possible solutions and evaluating their potential pros and cons. However, this platform also offers an extension beyond the theory of deliberativism in the sense that it highlights the creative conflict between alternative, diametrically opposed viewpoints and the alternative development of their argumentation paths (e.g., opponents and proponents of abortion). Argumentation can be developed by colliding with very different positions, not just by building a universally right

Table 3.13 Cognitive processes in the deliberatorium project

Cognitive process	How does it work in the project?
Collective sensing	<ul style="list-style-type: none"> How do users collect information in the Deliberatorium? It involves aggregating numerous observations made by as many people as possible. The people should have diverse perspectives or come from various regions and social backgrounds. To achieve this, it would be necessary “to get as many people as possible involved” [MK]. Many different points of view could be beneficial for grasping the general ideas about the topic. Thus, the number of people is important, and equally crucial are the diverse perspectives they provide Collective sensing can also be a more active process in the sense that “you might be revising or updating the things you’re looking for as part of the process of collecting the data” [MK]. This process can be active and adaptive when, in the course of gathering information, we learn to ask the right questions and decide what additional information is needed. In shaping the overall scope of the discussion and the questions we seek to answer, “the debate moderator plays an important role” [PS]. The role of the moderator is to pose “general questions that, in certain cases, can transform into more specific sensing goals or questions” [MK] “There were situations when the maps were becoming really large, and some people were posting bundles of ideas similar to others” [PS]. Therefore, the moderator’s task is to group threads, eliminate duplicates, etc. In the newer versions of the Deliberatorium, the co-creation of a debate structure by the users is possible. “That is somewhat innovative in terms of sensemaking” [PS] The Deliberatorium proves highly effective for gathering a large volume of diverse data and organizing it. In this context, the role of moderators and facilitators is crucial, tasked with precisely formulating initial questions, as well as structuring the debate and removing duplicate threads. The collaboration framework also aids in efficiently collecting diverse data, where participants interact with structured content rather than with other users. Group interactions, identity, and emotions take a back seat here, with the focus primarily on a large number of participants providing extensive batch information, constructing as detailed a picture of the problem as possible, seen from various alternative viewpoints What level of diversity is optimal for the collective? “When you are just observing, you probably want a diversity of sensors” [MK]. People vary in the types of patterns they can detect: “If you don’t have enough diversity, then there’s a chance you’re going to miss something important. (...) You are just not going to detect some opportunity or threat” [MK]. Thanks to the personal experiences of individuals, we can detect different information from the environment in various ways. “It is almost like everybody’s history makes them a tuning fork that will respond to a particular frequency. (...) Your tuning fork will be different from mine, so I am able to easily detect some kinds of patterns and completely miss others” [MK] “The value of diversity is asymptotic” [MK]. Initially, a little more diversity significantly boosts performance, but over time, the benefit of each additional person diminishes. Moreover, adding new people to a collective always incurs a cost: more people mean more work

(continued)

Table 3.13 (continued)

Cognitive process	How does it work in the project?
Problem solving	<ul style="list-style-type: none"> “The Deliberatorium becomes intriguing when you are formulating arguments for or against ideas” [PS]. The essence of the Deliberatorium is that participants move from a general statement to a more detailed one, fostering their creativity “I’ve noticed that other existing tools for collective ideation are essentially flat in structure.” People contribute with a “long list of individual answers. But the system doesn’t provide any structure or necessarily an incentive for people to build upon each other’s answers” [MK]. In these systems, ideas are presented at the same level. However, organizing them into more general and more detailed ideas, as well as separate families of similar ideas, aids in understanding the problem. “It seems like having that kind of structure for iterative refinement is important if you want more than just shallow single-person ideas” [MK] Participants “work on one knowledge tree. However, this tree can include zillions of branches, each representing a different family of answers” [MK]. Typically, the manual labor of a moderator is required to summarize a large deliberation. In the future, the use of AI in this role is conceivable One of the most critical design features of the Deliberatorium is its removal of authorship from ideas. “The objective is to depersonalize them” [PS]. This aims to ensure that existing biases towards individuals within the group do not overshadow the issue under consideration. “By design, the Deliberatorium reduces horizontal interaction because it really structures the discussion. [The users are] interacting not with you, but with your idea.” As a result, “the Deliberatorium is very sterilized. There is no emotion involved” [PS] The large number of people participating in collective work means we can identify an unexpected expert: “Often, the best answers for difficult problems come from unexpected places, (...) from people who do not have traditional expertise in that problem but have experience in something that turns out to be relevant” [MK] Typically, the large-scale debate and the extensive group bring benefits. However, it’s conceivable that “in some cases, adding new people can make things worse, depending on who they are” [MK]. With a large number of participants, we might receive a collection of thousands of ideas. Yet, what we likely need instead is a hundred deeply refined ideas When the topic requires not just solutions to a specific problem but more general reflections, exchanges, and clashing of views, the Deliberatorium does not provide space for extended expressions and interactions. “They [participants] hated it because they wanted to talk. They wanted to argue. (...) They wanted to write long things. The Deliberatorium was stripping out all these possibilities.” It is “not suited for those environments in which storytelling and emotions are dominating” [PS] Combining ideas and mutual inspiration are crucial for creativity. “Creativity often seems to come from combining disparate ideas. (...) It is about gathering a lot of different ideas that you can recombine” [MK]. Using an argument map supports the users’ creativity because presenting arguments in an accessible way helps avoid the blank page effect. Thus, users can begin inventing their own ideas when they have access to previously presented arguments. “That really promotes incremental creativity” [PS], meaning ideas based on others rather than complete innovations. “It is not designed for coming up with something that nobody else had ever known before, or had ever thought about” [PS] because if the participants have an overview of all the other ideas, they are framed in the thinking of the entire group

(continued)

Table 3.13 (continued)

Cognitive process	How does it work in the project?
	<ul style="list-style-type: none"> • An important aspect of group work is the recombination of ideas, i.e., using some partial answers to create a unified, more comprehensive answer. “Re-combinator is a tool in the Deliberatorium that lets people specify an answer consisting of many parts” [MK]. It allows for the selection from a list of possible answer options and the preference of specific partial solutions, determining how they combine with others. “I have not seen many systems that try to support recombination. So, I have been trying to do that [in the Deliberatorium]. It is about making a coherent whole from the partial answers. The key point is the interdependency between the parts” [MK]. The process of merging different ideas must be well-managed through analyzing their interdependencies; otherwise, the new combinations will be useless • Natural conflicts are not seen in the Deliberatorium as destructive to a group if they are transformed into something that can enhance quality: alternative solutions developed in contrast to each other. Constructively oriented conflict can serve a useful role through “interleaving critiquing in ideation” [MK]. Constructive criticism of emerging ideas can lead to improved quality. The sequence participants should follow is (1) Propose, (2) Critique, (3) Improve. Furthermore, in this model, because of the feature of depersonalizing ideas, it’s difficult for users to talk directly to each other. Removing the direct association between an idea and a person decreases the level of personal conflicts and shifts conflicts to the level of ideas • In subsequent versions of the Deliberatorium, new methods for organizing argument maps are being tested. When users deal with complex problems with interdependent sub-parts, “that means the answer space can become exponentially large, and (...) you can only explore a pretty tiny subset of the knowledge space” [MK]. In older instances, everything depended on the work of moderators; currently, methods using AI are being tested. One of these is “optimization, guiding the process to look for those combinations most likely to end up being the good ones. Optimization can be done by genetic algorithms or simulated annealing. (...) Trying to find the best answer, via a genetic algorithm, several combinations are evaluated step by step to find the best one. You are not able to test all solutions, but you can end up coming to some very good solutions, even though you only explored a tiny fraction of the full space (...). The question I was working on in the Deliberatorium is: can I use optimization techniques to suggest to people which combinations they should be paying attention to, which ideas they should try to combine together” [MK]
Decision making	<ul style="list-style-type: none"> • Knowledge gathered through the Deliberatorium aids in preparing for decision-making. This primarily requires extracting information from the large data amount, filtering it, and making it more accessible. “The advantage of the Deliberatorium is what happens before the vote; it creates a guide on how to vote. (...) Voting per se is problematic. But if we can add a good debate before the voting, then the voting gains power” [PS]

(continued)

Table 3.13 (continued)

Cognitive process	How does it work in the project?
	<ul style="list-style-type: none"> • The stage directly related to decision-making involves synthesizing threads and highlighting the most important arguments. Making syntheses is a strong suit of the Deliberatorium. “It is the core of the Deliberatorium. (...) we had facilitators who were synthesizing the discussion and argument map. And people loved it” [PS]. There are also mechanisms for group synthesis, where users act as a collective moderator. Probably in the future, it will be possible to train AI to create maps out of free conversation in natural language, making it even more accessible for users: “it can help to understand what were the highlights in the debate, in a way that it’s ready for decision-making” [PS] • Several voting methods have been tested in the Deliberatorium. However, none of them resemble those in participatory budgeting, where a single solution is selected. There were methods involving thumbs up and thumbs down, as well as prioritizing possible solutions. A long-tested method was also the elimination of the worst solutions (the so-called bag-of-lemons). There is the potential to introduce evaluation criteria and diversify votes considering different criteria (e.g., the least expensive solution vs. the safest solution). We can refer to this as advanced categorization as a part of the decision-making process • The decision-making model is not based on consensus but aims to select the most optimal alternative for people with varying viewpoints. “In many situations, it’s unrealistic to expect consensus, where everybody agrees on the best answer. People’s preferences are different” [MK]. A more realistic approach is to choose a solution that maximizes general welfare, “try to find the answer where the sum of everybody’s happiness is the greatest. Now, it’s not consensus because there might be some people who hate the idea (...). I think that consensus is a bit of a chimera in many situations” [MK] • In new implementations of the Deliberatorium within the decision-making process, there is the possibility of seeking a Pareto optimal front. If the debate brings forth two competing, distinct concepts, and we want to decide which to choose, the solution is not an averaging combination of these two concepts because then the solution won’t be satisfactory for either side. Therefore, participants should evaluate not just their best option but all available options, attempting “to limit the solutions that are worse for both sides. (...) The hard part is just making sure that you’ve eliminated all the dumb choices. If you don’t do it, then the result is people fighting to select one of the ideas, which actually is not optimal, and nobody’s going to be happy” [MK]. Eliminating options that fall below the Pareto front is intended to facilitate the choice of an optimal solution from the adversaries’ viewpoints
Collective memory	<ul style="list-style-type: none"> • Collective memory is regarded as a “super important problem (...). A few years after the debate, you realize that your current problem is similar to what was discussed several years ago” [MK]. The aim is to “efficiently organize the knowledge created by the debate, making it easy for future individuals to find what they need” [MK] • Memory is tied to the summarization aspect in the Deliberatorium: “argument maps are (...) the way to synthesize information” [PS]. They serve as an effective means to preserve debate summaries, enabling the collective to learn from its own mistakes or successes

(continued)

Table 3.13 (continued)

Cognitive process	How does it work in the project?
	<ul style="list-style-type: none"> • The argument map should serve as a record of memory, but for it to be effective, the way discussions are conducted must be orderly. Preserving insights in collective memory is a distinct process where people are expected to collectively organize content as they add it. “Every time they say something, it needs to be placed in the right spot on the knowledge graph or in the map (...). So it’s a bit more work for each individual, but the result is far more valuable for the future because it’s organized like a book instead of like a pile of post-it notes” [MK] • It is anticipated that the system will evolve to utilize AI, particularly Large Language Models (LLM), to simplify memorization. LLM will help ensure that everything added to the map is placed correctly. “But it’s still in the beginning, and it will not be so easy” [MK]. It is challenging because when people debate spontaneously, they often do not refer to the actual questions posed and mix separate threads

consensus. Participants’ engagement in deliberation seems particularly linked with seeking the most substantive arguments to support their theses and confronting alternative theses by gathering arguments against them. Activity in public issue debates using the Deliberatorium thus benefits from both a tendency towards rational justification of one’s position and a desire to confront positions deemed incorrect. However, thanks to depersonalization, users confront ideas, not people.

The assumption adopted in the Deliberatorium removes authorship from ideas and thereby depersonalizes them, leading to minimizing emotions associated with the debate and minimizing the pursuit of individual benefits by participants (such as personal acclaim for being most active). This also impacts the limitation of forming a group identity and a community of values. Interactions are vertical, with the system and the argument, not between people, reducing the possibility of influencing the debate through ideologically driven narratives. In this way, the assumption of the rationality of the debate is fulfilled. The element that binds the debate together is primarily its top-down imposed structure. Diversity is useful for gathering varied arguments, but its limit seems to be the ability to understand the structure of argumentation and make a reasonably rational assessment of other arguments.

A certain difficulty from the standpoint of debate impartiality may arise from the dominant role of the moderator, who has the power to set the agenda and influence argumentation paths. However, this seems unavoidable due to the necessity of keeping the debate structured. Perhaps a chance to make this process more objective in the future is the automatic extraction of arguments from the debate and their placement in the argument map using AI, though its effectiveness remains uncertain.

While collaboration within a collective of shared identity appears to have less significance in Deliberatorium compared to other platforms, it is possible that utilizing a common set of arguments could foster a sense of community to some extent. However, technically, due to the lack of free debate, such a community is not formed on the Deliberatorium website. On the other hand, a vast reservoir of knowledge is created, which can be preserved in collective memory. This knowledge is not

tailored to a single group with a uniform identity but is available to various groups, allowing them to present their ideas. Therefore, the risk of emotionally excluding certain views due to identity differences is mitigated here. The platform accommodates diverse concepts of citizenship and varying interpretations of the common good, as well as the potential prioritization of individual over collective benefit. Additionally, it allows for a multifaceted collective memory, which, due to its sophisticated structure for storing and visualizing information, becomes more accessible and intuitive than on other platforms. This model offers room for diverse concepts of citizenship, multiple interpretations of the common good, and even scenarios where individual interests may supersede collective ones. Additionally, it provides for a complex collective memory, which, due to its sophisticated structure for storing and visualizing information, is more accessible and intuitive than on other platforms. However, there is no certainty about if or how this memory will be utilized. For its effective use, collectives must intentionally build their identity around particular segments of collective memory.

3.4 Insights from the Evaluated Projects

The conducted review of case studies leads us to several significant observations. To start, a single platform fully encompassing all CI cognitive processes on all public policy levels seems difficult, perhaps even impossible, to implement. Collectives can respond to the challenges and needs of policymaking to varying degrees and at different levels. A platform designed to cover all processes would be neither practical nor yield satisfactory results. Although Decide Madrid aims to encompass almost the entire policymaking cycle, its driving force and attracting people is the city's participatory budget, which is where the participants' attention is largely focused. So, the success of a project depends on the set objectives, the size of the group, the method of selecting participants, and finally, the techniques used to organize collective thinking processes. The decision-making technique is also very important, as is the issue of recording the results of the work in collective memory and its subsequent utilization, which constitutes one of the most significant challenges for the future.

Our evaluation proves that civic engagement is a key social factor necessary for the success of open online policymaking. The quality of civic engagement, as highlighted by Hannah Arendt, is linked to a sense of civic duty and the pursuit of the common good. Engaged citizens' activities enhance their sense of self and belonging, which is beneficial for both personal and collective well-being. As we have observed, active citizens engaged in the debate often feel a stronger connection to the decisions made collectively. Experiences from Iceland indicate that the majority of citizens largely respect decisions made in votes with high participation.

It is also evident that the quality of civic engagement depends on the degree to which the subject of the debate is linked with the personal interests of the participant and the direct impact on their life. We notice that urban projects dominate the most

successful projects utilizing CI in the public sphere, especially those using participatory budgeting. Municipal policies attract attention because they significantly impact the daily lives and well-being of citizens, their families, neighborhood communities, and local societies. Here, we can easily identify direct interests and credibly estimate the impact of specific actions in urban spaces on our lives. The individuals comprising the successful local community are well aware that they can participate in the benefits collectively developed by investing their resources, such as skills, material assets, and free time. Beyond local communities, we see successful examples of projects among groups focused on solving narrowly defined issues and gathering individuals highly interested in the specialized topics. Examples include the debate on the program of Italy's Democratic Party in the Deliberatorium or the work on a new agricultural policy organized by the Green Party of West Australia in Loomio. Participants in the debate must be independent enough so that lack of time, lack of funds, insufficient knowledge, or the influence of the lobbying groups do not distort their engagement. Only overcoming these obstacles can ensure that the participants' agency is strong enough for their involvement to be beneficial, meaningful, and acknowledged for both them and the community.

Perhaps this is one of the reasons why, among successful CI projects related to public issues, there are few nationwide projects with a mass reach. With their complexity level, national and international policies set high demands and, due to the large number of people they cover, become abstractly distant from ordinary citizens who might engage in shaping them on an online platform. It is difficult to expect the undertaking of intellectual effort and, consequently, rational reasoning if the subject of this effort is significantly remote from our current perspective. It is irrational to have extensive knowledge about a given public issue when the low rate of return from the acquired orientation does not justify the costs incurred in terms of time and other resources.

Some may conclude that when obstacles to collective intelligence include a lack of civic engagement or knowledge deficits resulting from, for example, systematic biases, the solution is to calibrate policy projects so that their scope and user base are appropriate to the situation. This ensures that members of the collective have sufficient motivation to participate due to the alignment of their interests with the subject of the debate. It is essential not to limit this alignment to economic interests; in many cases, emotional bonds, traditions, beliefs, or social and familial connections are equally or even more strongly tied to an individual's interests and well-being.

How can we ensure that the opportunity to participate in online communities is granted to the right individuals, however? In the case of urban communities, it's relatively straightforward: aside from being of legal age, residency in the city is a common requirement. The same applies to organizations with defined membership status: being a member entitles you to participate in the collective. But what about other kinds of debates concerning specialized policy issues? Beth Noveck (2015) proposed the creation of field expert groups for handling policy issues, where decisions could be made within an expert's area of specialization. Then, the aggregated expert opinions could be utilized along with domain-specific factual knowledge. However, this idea raised legitimate concerns. It would certainly deviate from the idea

of opening policymaking and revert to the *expert* policymaking of the 1950s. Questions arise: Who will select the experts, and on what basis? How can we guarantee that political affiliations do not influence their selection?

It seems that relying on top-down, decreed expert teams for collective intelligence would be a mistake. Experiences from reviewed projects have shown that the best solutions for challenging problems often emerge from unexpected sources, involving individuals who do not possess traditional expertise. A promising direction appears to be the self-organization of responsible and aware citizens rather than imposing top-down models and strict access criteria. If the project genuinely succeeds in engaging a large number of people, their collective intelligence has the potential for self-organization. Therefore, it is more appropriate to observe how people behave and facilitate that rather than enforcing pre-made organizational solutions. The level of knowledge and engagement of participants can be rewarded within the community, as seen in cases like Better Reykjavik, where ideas and arguments gaining the most reputation rise to the top of the ranking list, increasing their visibility. We can also promote civic attitudes by rewarding them through metrics such as those available in Deliberatorium, for instance *efficiency* or *maturity*, as well as encouraging independent thinking (minimizing *groupthink*).

The issue of participant diversity is somewhat more complex. Diversity appears to be particularly effective in collective sensing, as it aids in identifying new sources of information, uncovering hidden opportunities, and detecting potential threats. A variety of sensors can enhance the quality of observations. However, diversity is not as prominently expected in the context of collective decision-making (beyond mere aggregation or ranking-based voting). For example, in platforms like Loomio, a common identity and a shared sense of the common good play a decisive role, and consensus decision-making is the expected model.

Analysis of different models of organizing online collectives, taking into account the number of platform users and the nature of their participation, confirms Andler's (2012) hypothesis regarding the existence of group cognitive processes at two fundamental levels: (1) *thin* processes occurring within large, mob-like structures where participants, instead of engaging in deliberation or exchanging information and arguments, simply provide their contributions, which are subsequently processed by some form of aggregating mechanism, and (2) *thick* cognitive processes, characteristic of closely cooperating communities, often bound by a shared ethos, and continuously coordinating or adjusting their collective expectations. These communities possess a stronger sense of a collective *we* capable of shared intentions and actions. Although we typically do not encounter a pure form of either *thin* or *thick* models in the projects evaluated in this book, we can discern the characteristics of both models. In some cases, the projects are more focused on dispersed crowds and their aggregated opinions. In contrast, in other cases, they delve into interactions among individuals within groups that share a common identity.

The extended process of collective sensing is a defining characteristic of systems engaged in *thin* cognitive processes. It involves perceiving social data through the collection of a large volume of dispersed information, with participants intentionally acting as *sensors* for specific aspects of the world. This approach is crucial

when building a broad knowledge base around a particular subject of interest. A prime example of this approach is the Deliberatorium, where the goal is to aggregate numerous observations made by as many people as possible, constructing a comprehensive picture of the problem from various alternative viewpoints. Gathering many opinions allows for the detection of diverse perspectives or *patterns of thinking* and provides greater detail. Bounded rationality acknowledges that there can be various ways of reasoning about a given problem, each with its arguments for and against, decision heuristics, and cognitive patterns. Independently, they can contribute to a better understanding of the issue, making the construction of diverse argumentation paths meaningful. A similar approach was evident in the *Idea Bank* function within the BOMK project, where ideas were additionally mapped out.

The aggregation of individual statements and their transformation into systematic arguments for and against a particular idea is also a distinctive feature of Better Reykjavik. In this model, the primary activity is individual participation, and the collective is, as with Hayek, a kind of *fiction*, a byproduct of the presented singular interests and opinions. Participants in Better Reykjavik have a sense that they want a high-quality community but do not wish to invest too much time in it without considering their benefits. They are independent individuals, each pursuing their own goals and satisfying their own needs, with the collective working for the common good emerging somewhat incidentally. Identity, subjection to group interests, and shared collective goals are of secondary importance here and do not directly influence the participants. However, they may indirectly impact collective memory, which we will revisit later.

In this model, despite the goal of the group being to amass a large collection of facts, it turns out that building only one set of arguments can be insufficient, and it would be better to create alternative sets. In the Deliberatorium, the emphasis is placed on capturing the multiple threads of the debate. Different positions, if arguments fuel them, are not excluded from the discussion under the pretext of seeking consensus. This leads to *incremental creativity*, where new ideas are largely inspired by familiarizing oneself with the available map of arguments and responding to them. The focus is on gathering a variety of arguments that surface throughout the debate, aiming to capture not just the prevailing viewpoints but also the different strands of reasoning. In this way, we move beyond the deliberative theory that seeks consensus and a uniform set of universally acceptable arguments. We also see the positive effects that can arise from confronting opposing positions. Conflicts, permissible as a useful tool, should concern not people but positions in the debate, hence the proposed anonymization of authorship of arguments on this platform. Constructive conflict can play a beneficial role by integrating criticism into idea generation.

Thick cognitive processes come into play when shared values and interests facilitate close collaboration within a group, a dynamic that cannot be achieved when there is no mutual interaction among individuals as the individuals communicate with the platform rather than with each other. In earlier public policy studies, examining such processes was considered a mistake, but in the current post-positivist approach, the consideration of ethical values, personal sympathies, and emotions has been accepted. As we recall, the positivist separation of facts from values, with

an exclusive focus on facts, was put on the shelf in contemporary policymaking. However, an open question remains: what influences online collectives to establish a common identity and shared values, and how does this affect collective thinking?

In this context, Decide Madrid presents an interesting case. On the one hand, we observe mass participation, and on the other hand, we can see identifiable minorities of highly active individuals who engage extensively in collaboration. Each proposal has its dedicated community space where individuals are invited to collaborate, and it is within these spaces that some of the best solutions are often formulated. The collaborating groups are often formed outside the platform through natural human interactions and join as pre-existing collectives. However, online activities enable the expansion of the engaged group and the creation of new social bonds focused on addressing societal issues. The latest implementation of natural language processing technologies in Decide Madrid allows contributions to be grouped based on their content, forming *communities of patterns*, and groups of people focused on specific topics. This approach facilitates the organic emergence of interest groups among unorganized citizens, fostering the creation of bottom-up proposals. Representatives of the local authorities also participate in the proposal creation process to help estimate costs, further expanding the scope of collaboration. The opportunity for such close collaboration also translates into grounding norms and values that unify the community, which become especially apparent when examining the winning ideas.

A similar situation is observed in the BOMK project, where previously existing internet communities of residents enjoyed greater popularity than the official BOMK discussion groups or the interactive *Idea Bank*. It appears that when people choose to engage collaboratively in developing ideas, particularly within grassroots organizations like *Nowohucianie*, it's often due to a deep emotional attachment to their district. This fosters commitment and a sense of shared purpose.

Loomio is a model example of *thick* cognitive processes. The community using this platform typically consists of small to medium-sized high-trust groups characterized by strong interpersonal relationships and a shared consensus on fundamental values. The dynamic decision-making process in Loomio intertwines group discussions closely with the proposed motions, rendering them a vital aspect of the platform's collaborative ethos. By ensuring that each member's input can shape the final outcome, Loomio elevates the quality and acceptance of the decisions, nurturing a robust sense of engagement and solidarity within the group.

An important aspect of projects characterized by *thick* cognitive processes is the cultivation of collective identities. When individuals align with the community's values, it fosters a collective identity that contributes to the overall well-being of the group, promoting a sense of belonging and a shared sense of purpose. In Loomio, thanks to a sense of shared identity, community members have an increased ability to unite for collective action, advocating for common interests or addressing shared challenges. In practice, groups that fail to develop a collective identity often struggle to reach a consensus and eventually dissolve. Analyzed cases demonstrate that when individuals feel a sense of connection, they are more likely to contribute positively to the community, reinforcing a cycle of mutual support. Communities most often

coalesce around shared values and beliefs, as seen in Loomio, Decide Madrid, or BOMK.

Another intriguing trait shared by nearly all the discussed projects is their genesis within social movements that challenge the traditional political establishment. In the case of Decide Madrid, it is the M15 Indignados movement; the beginnings of Loomio are linked to the Occupy Movement; Better Reykjavik emerged in response to the economic crisis and the victory of the satirical *Best Party* in the local election; and even Deliberatorium has a history of implementation for the Italian Democratic Party, intended as an alternative to old, corruption-ridden political parties. Contestation was a powerful motivator for the initial engagement of collectives in these projects, and the expected *independence* from authorities consequently added vigor to the projects and built their positive image. Undoubtedly, the tradition of contestation and the sense of independence impact the formation of group identity in these projects. For instance, one of the compelling motivations in Decide Madrid was to establish a community that self-initiates from the grassroots level and independently addresses issues, aiming to transform society instead of relying on ineffective government institutions. Furthermore, in Decide Madrid and other projects, it has been observed that strong opinions and biases are not viewed as obstacles to participation but as elements of active involvement. Coherent groups typically unite through strong emotions stemming from contestation and pride in their shared achievements, further bolstering their identity.

Collective identity is essential for a group's success but can also lead to conflict. Such conflicts may threaten the group's cohesion and even its existence. Addressing this challenge is possible with an appropriate group and member maturity level. For a group to effectively exhibit collective intelligence, its members must, at least to a certain extent, trust each other and collectively value the discussed topics. Collectives bound by such ties are aware of their identity and structure; however, this does not mean that there is no room for self-organization. This includes the natural emergence of leaders. The structure in groups with a common identity is usually flexible and does not impose a strictly defined leader but is open to providing space for those who engage most actively to demonstrate their leadership.

Our evaluation shows that strong emotions, conflicts, and contestation can benefit smaller groups, as existing interpersonal bonds within them help maintain the group as a whole. There is typically less emotional involvement in larger groups and projects where data collection is the primary focus and where the emphasis lies on facts and arguments. Care for the rationality of arguments and the pursuit of impartiality work particularly well in sufficiently large communities, where the number of participants allows for a more precise mapping of issues, enabling the construction of various problem-solving paths. On the other hand, the development of collective decision-making heuristics functions better in smaller communities, where identity, shared values, and the associated narrative are potent forces that deeply engage participants in the group's work.

In the field of public policy, due to the limited ability to assess the quality of a selected solution (remember the discussions about the War on Poverty?), retaining the entire argumentative structure of discussion in a sort of common memory is

highly valuable. This allows individuals to have alternative solutions readily available, trace connections between arguments, and model causal relationships (which may advance technologically in the near future with the development of AI based on large language models). However, Mercier and Landemore (2012) argue that even in the case of moral and complex policy decisions, some criteria for better outcomes are possible. Although the actual success of a selected policy may have to await future outcomes, Mercier and Landemore claim that the “epistemic bases” of decisions are themselves indirect measures of the possible success of the policy selected. According to Capella, “deliberating groups can be effective in advancing the epistemic bases for good decisions and for enhancing the quality of decisions (...) concerning public issues” (2017). This measure of opinion quality can be called evaluating *argument repertoire* (Cappella et al., 2002). The number of arguments connected with a given problem can be treated as a measure of the quality of the debate and, therefore, an indicator of increased intelligence in group deliberations. Generating arguments and counterarguments requires people to envision conditions that would prove or falsify their explanations, prompting them to consider different perspectives and solutions to the problem.

The Deliberatorium and Better Reykjavik explored technical solutions to enable the creation of extensive argumentation structures. These projects aimed to transform the pool of ideas generated during the initial stages of the debate into a well-structured and well-supported framework for decision-making. Especially in the Deliberatorium, the refined ideas and arguments are organized into topically structured tree-like frameworks, consisting of a hierarchy of questions to be answered, possible answers for these questions, and the supporting arguments for each answer (Klein, 2011a, b). This collective work can address complex problems by breaking them down into manageable components. Using such structured frameworks aligns with the principles of transparency and accountability, as it allows for the traceability of decisions back to the underlying arguments and evidence. Clarity helps participants systematically explore and evaluate various facets of the problem or decision at hand. On the other hand, traceability is essential for maintaining trust and ensuring the credibility of online platforms in the context of public debate.

However, for the *epistemic base* of the debate to be preserved, collective intelligence must grapple with a significant issue, namely the process of *collective memory* (CM). Utilizing collective memory offers an opportunity to avoid systematic errors in collective reasoning. By referencing memory, we can verify what worked in the past and how it worked, comparing these observations with our current situation. Accumulated knowledge can enhance the functioning of the entire cognitive system by facilitating continuous learning, receiving feedback, and making forecasts. This forms the basis for progressively efficient intelligence throughout the entire community.

As we recall, the concept of collective memory has been acknowledged as crucial for CI since as early as 1999, when Heylighen introduced it in his theoretical model. In this model, Heylighen envisioned the entire cognitive operations as the construction of collective mental maps. He meant preserving a structural model for the debated problem in memory: “Various discoveries by members of the collective are being registered and stored in the [collective mental map] so that the information will

remain available for as long as necessary.” This map could take the form of “a registry of events or an edited collection of notes; it is a highly selective representation of features relevant to problem-solving” (Heylighen, 1999).

As revealed by the evaluations described in this chapter, incorporating CM as part of the collective thinking process is still not obvious. In fact, it is an oft-neglected area whose potential remains untapped. Insights from specialists involved in projects like Decide Madrid and Loomio indicate that the participants frequently repeat similar mistakes in successive iterations, and new participants seldom leverage past experiences. Furthermore, after engaging in the processes once, it is common for them to leave the platform. For instance, as observed in the BOMK project, the outcomes of negotiations conducted by municipal officials—where opinions are aggregated and duplicate projects are merged—could serve as an extraordinary source of accumulated practical wisdom based on experience. However, the resource is underutilized, and new users do not benefit from the experience of previous years. Even in projects such as Deliberatorium and Better Reykjavik, which place a significant emphasis on memorializing the results achieved, there is a prevailing belief that observations and conclusions arising from debates are often irretrievably lost.

The identified general problem with CM is particularly significant, as many researchers believe that memory-related cognitive processes are crucial for the entire intelligence system for both individual and collective intelligence. For, as J. Hawkins (2021) notices, memory allows every thinking being to effectively manage the energy of the entire system. Thanks to memory, we do not waste energy analyzing the same activities over and over again. Attention and focus are *rare* valuable resources, so thanks to the use of behavioral patterns stored in memory gained through feedback and learning, we do not distract attention from unnecessary details. The same, as Mulgan notes, applies to collective intelligence: memory and learning are the most important cognitive factors, thanks to which we can accumulate the knowledge verified by practice, and, consequently, we are “rethinking how we think” (Mulgan, 2018, p. 73). Landemore highlights that when the intelligence of a collective is extending not just through space (including many people) but also through time (including the knowledge of more than one generation), room for memory and experience is made. A crucial element of intelligence is the ability to learn from one’s mistakes, what “requires the ability to store and process knowledge’s of one’s past” (2012b, p. 240–241). Thinking is, therefore, largely associated with recalling patterns stored in memory and comparing them to current challenges. Nevertheless, placing this process at the center of collective thinking has not been successful in any of the online projects utilizing CI that I am aware of.

Collective memory is still an untapped field within CI. Nevertheless, due to the ongoing proliferation of artificial intelligence employing large language models, significant new opportunities for its advancement emerge. Certainly, many questions persist: What should be preserved in this memory—everything or selected threads? What criteria should guide the formation of collective memory, and how should it be practically utilized in collective thinking? Finally, how will current issues related to the digital public sphere, such as the erosion of traditional media, the privatization of debate by social media giants, behavioral targeting, opinion polarization, and the

flood of fake news, affect the shaping of collective memory and online debate in general? Answers to these questions, likely to emerge in the near future, will have a significant impact on enhancing the quality of intelligent collective thinking. We will discuss these topics in the next chapter.

References

Adenskog, M. (2018). After the equilibrium: Democratic innovations and long-term institutional development in the city of Reykjavik. *Analyse & Kritik*, 40(1), 31–54. <https://doi.org/10.1515/auk-2018-0002>

Aggarwal, I., Woolley, A. W., Chabris, C. F., & Malone, T. W. (2019). The impact of cognitive style diversity on implicit learning in teams. *Frontiers in Psychology*, 10, 112. <https://doi.org/10.3389/fpsyg.2019.00112>

Aitamurto, T. (2014). *Crowdsourced off-road traffic law experiment in Finland*. Parliament of Finland.

Aitamurto, T., & Chen, K. (2017). The value of crowdsourcing in public policymaking: Epistemic, democratic and economic value. *The Theory and Practice of Legislation*, 5(1), 55–72. <https://doi.org/10.1080/20508840.2017.1282665>

Aitamurto, T., Landemore, H., & Galli, J. S. (2017). Unmasking the crowd: Participants' motivation factors, expectations, and profile in a crowdsourced law reform. *Information, Communication & Society*, 20(8), 1239–1260. <https://doi.org/10.1080/1369118X.2016.1228993>

Almirall, E., Lee, M., & Majchrzak, A. (2014). Open innovation requires integrated competition-community ecosystems: Lessons learned from civic open innovation. *Business Horizons*, 57(3), 391–400. <https://doi.org/10.1016/j.bushor.2013.12.009>

Ambrosino, A., Birk, S. C., Mingardi, C., Nikolou, G., Spielberger, L., Thevenin, E., & Trinkle, A. (2023). Youth and democracy: Digital opportunities for the future of participation. In G. Rouet, & T. Côme (Eds.), *Participatory and digital democracy at the local level. Contributions to Political Science*. Springer, Cham. https://doi.org/10.1007/978-3-031-20943-7_5.

Anastasio, T. J. (2022). Deriving testable hypotheses through an analogy between individual and collective memory. In S. M. O'Mara (Ed.), *Progress in brain research* (Vol. 274, pp. 31–70). Elsevier.

Andler, D. (2012). What has collective wisdom to do with wisdom? In H. Landemore, & J. Elster (Eds.), *Collective wisdom*. Cambridge University Press.

Arana-Catania, M., van Lier, F. A., Procter, R., Tkachenko, N., He, Y., Zubiaga, A., & Liakata, M. M. (2021). Citizen participation and machine learning for a better democracy. *Digital Government: Research and Practice*, 2(3), Article 27. <https://doi.org/10.1145/3452118>.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence (Ed.) *The psychology of learning and motivation* (Vol. 2). Academic Press.

Bachleitner, K. (2022). Collective memory and the social creation of identities: Linking the past with the present and future. In S. M. O'Mara (Ed.), *Progress in brain research* (Vol. 274). Elsevier.

Baregheh, A., Rowley, J., & Sambrook, S. (2009). "Towards a multidisciplinary definition of innovation", *Management Decision*, 47(8), 1323–1339. <https://doi.org/10.1108/00251740910984578>

Barometr Krakowski. (2018). Raport badawczy: Badanie socjologiczne przeprowadzone w Krakowie maj-lipiec 2018. <http://barometrkrakowski.pl/wp-content/uploads/2019/02/Barometr-Krakowski-2018.-Raport-badawczy.pdf>.

Bartlett, R., & Deserii, M. (2016). Loomio and the problem of deliberation. *Open Democracy*. <https://www.opendemocracy.net/en/digitaliberties/loomio-and-problem-of-deliberation/>

Becker, J., Porter, E., & Centola, D. (2019). The wisdom of partisan crowds. *Proceedings of the National Academy of Sciences*, 116(22), 201817195. <https://doi.org/10.1073/pnas.1817195116>

Bigham, J., Bernstein, M., & Adar, E. (2015). Human-computer interaction and collective intelligence. In T. Malone & M. Bernstein (Eds.), *Handbook of collective intelligence*. MIT University Press.

Bojic, I., Marra, G., & Naydenova, V. (2016). Online tools for public engagement: Case studies from Reykjavik. [arXiv:1611.08981](https://arxiv.org/abs/1611.08981)

Bonabeau, E. (2009). Decisions 2.0: The power of collective intelligence. *MIT Sloan Management Review*, 50(2), 45–52.

Bose, T., Reina, A., & Marshall, J. A. R. (2017). Collective decision-making. *Current Opinion in Behavioural Sciences*, 16, 30–34. <https://doi.org/10.1016/j.cobeha.2017.03.004>

Carey, D. (2009). *Iceland: The financial and economic crisis*. OECD Economics Department Working Papers, No. 725, OECD Publishing, Paris, <https://doi.org/10.1787/221071065826>

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, 54, 1–22. <https://doi.org/10.1037/h0046743>

Cappella, J. N., Price, V., & Nir, L. (2002). Argument repertoire as a reliable and valid measure of opinion quality: Electronic dialogue in campaign 2000. *Political Communication*, 19(1), 73–93.

Capella, J. N., Zweng, J., & Price, V. (2017). Collective intelligence: The wisdom and foolishness of deliberating groups. In K. Kenski & K. H. Jamieson (Eds.), *The Oxford handbook of political communication* (1st edn.). Oxford University Press.

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Sage.

Chen, K., & Aitamurto, T. (2019). Barriers for crowd's impact in crowdsourced policymaking: Civic data overload and filter hierarchy. *International Public Management Journal*, 22(1), 99–126. <https://doi.org/10.1080/10967494.2018.1488780>

Chikarsal, P., Tomprou, M., Kim, Y. J., Woolley, A., & Dabbish, L. (2017). Deep structures of collaboration: physiological correlates of collective intelligence and group satisfaction. In *Proceedings of the 20th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW 2017)*.

Citizens.is. Better Reykjavik. (2023). https://citizens.is/portfolio_page/better_reykjavik/

Clark, B. Y., Zingale, N., & Logan, J. (2017). Intelligence and information gathering through deliberative crowdsourcing. *Journal of Public and Nonprofit Affairs*, 3(1), 55–78. <https://doi.org/10.20899/jpna.3.1.55-78>

Collins, B., Marichal, J., & Neve, R. (2020). The social media commons: Public sphere, agonism, and algorithmic obligation. *Journal of Information Technology & Politics*, 17(4), 409–425. <https://doi.org/10.1080/19331681.2020.1742266>

Corbin, J., & Strauss, A. (2008). *Basics of qualitative research* (3rd ed.). Sage.

Crowdsourcing Better Education Policy in Reykjavik. (2019). Centre for Public Impact. <https://www.centreforpublicimpact.org/case-study/crowdsourcing-better-education-policy-reykjavik>.

Darwin, C. (1998). *The descent of man and selection in relation to sex* (Original work published 1871). Prometheus Books.

Dawkins, R. (1976). *The selfish gene*. Oxford University Press.

Durkheim, E. (1971). *Elementary forms of the religious life* (J.W. Swain, Trans.). Free Press. Original work published 1912.

Džinić, J., Svidroňová, M. M., & Markowska-Bzducha, E. (2016). Participatory Budgeting: A Comparative Study of Croatia, Poland and Slovakia. *NISPAcee Journal of Public Administration and Policy*, 9(1), 31–56. <https://doi.org/10.2478/nispa-2016-0009>

Eisenhardt, K. M. (1989). Building theories from case study research. *The Academy of Management Review*, 14(4), 532–550.

Engel, D., & Malone, T. W. (2018). Integrated information as a metric for group interaction. *PLoS ONE*, 13(10), e0205335. <https://doi.org/10.1371/journal.pone.0205335>

Epp, D. A. (2017). Public policy and the wisdom of crowds. *Cognitive Systems Research*, 43, 53–61. <https://doi.org/10.1016/j.cogsys.2017.01.002>

Fernández-Martínez, J., López-Sánchez, M., Aguilar, J. A. R., Rubio, D. S., & Nemegyei, B. Z. (2018). Co-designing participatory tools for a new age: A proposal for combining collective and artificial intelligences. *International Journal of Public Administration in the Digital Age*, 5, 17.

Finley, K. (2014). Out in the open: Occupy wall street reincarnated as open source software. *Wired*. <https://www.wired.com/2014/04/loomio/>

Flavián, C., Guinalíu, M., & Gurrea, R. (2006). The role played by perceived usability, satisfaction and consumer trust on website loyalty. *Information Management*, 43, 1–14.

Fleming, S. M., & Frith, C. D. (2014). *The cognitive neuroscience of metacognition*. Springer.

Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience*, 8, 443. <https://doi.org/10.3389/fnhum.2014.00443>

Gardner, H. (1993). *Multiple intelligences: The theory in practice*. Basic Books.

Goldstone, R. L., & Theiner, G. (2017). The multiple, interacting levels of cognitive systems (MILCS) perspective on group cognition. *Philosophical Psychology*, 30(3), 334–368. <https://doi.org/10.1080/09515089.2017.1295635>

Guth, K. L., & Brabham, D. C. (2017). Finding the diamond in the rough: Exploring communication and platform in crowdsourcing performance. *Communication Monographs*, 84(4), 510–533. <https://doi.org/10.1080/03637751.2017.1359748>

Habermas, J. (1984). *The theory of communicative action*. Beacon Press.

Haidt, J. (2013). *The righteous mind: Why good people are divided by politics and religion*. Penguin Books.

Halbwachs, M. (1992). The reconstruction of the past (L. A. Coser, Trans.). In: Coser, L.A. (Ed.), *On collective memory*. University of Chicago Press, pp. 46–52 (Original work published 1950). <https://doi.org/10.7208/chicago/9780226774497.001.0001>

Hardas, M. S., & Purvis, L. (2012). Bayesian vote weighting in crowdsourcing systems. In N.T. Nguyen, K. Hoang, & P. Jędrzejowicz (Eds.), *Computational collective intelligence: Technologies and applications* (ICCCI 2012). Lecture Notes in Computer Science (Vol. 7653). Springer. https://doi.org/10.1007/978-3-642-34630-9_20

Hawkins, J. (2021). *A thousand brains: A new theory of intelligence*. Basic Books.

Hayek, F. (1945). The use of knowledge in society. *The American Economic Review*, 35(4), 519–530.

Hebb, D. O. (1949). *The organization of behavior: A neuropsychological theory*. Wiley & Sons.

Heylighen, F. (1999). Collective intelligence and its implementation on the web: Algorithms to develop a collective mental map. *Computational & Mathematical Organization Theory*, 5(3), 253–280.

Heylighen, F., Heath, M., & Van Overwalle, F. (2004). The emergence of distributed cognition: A conceptual framework. *Proceedings of collective intentionality IV*, Siena, Italy.

Hogan, M., Ojo, A., Harney, O., Ruijer, E., Meijer, A., Andriessen, J., Pardijs, M., Boscolo, P., Boscolo, E., Satta, M., et al. (2017). Governance, transparency and the collaborative design of open data collaboration platforms: Understanding barriers, options, and needs. In A. Ojo & J. Millard (Eds.), *Government 3.0—next generation government technology infrastructure and services*. Springer.

Hong, L., & Page, S. (2004). Groups of diverse problem-solvers can outperform groups of high-ability problem-solvers. *PNAS*, 101, 16385–16389.

Horn, J. L. (1969). Intelligence: Why it grows, why it declines. *Transaction*, 4, 23–31.

Huebner, B. (2013). *Macrocognition*. Oxford University Press.

Iacuzzil, S., Massaro, M., & Garlatti, A. (2020). Value creation through collective intelligence: Managing intellectual capital. *The Electronic Journal of Knowledge Management*, 18(1), 68–79.

Iandoli, L., Quinto, I., Spada, P., Klein, M., & Calabretta, R. (2018). Supporting argumentation in online political debate: Evidence from an experiment of collective deliberation. *New Media & Society*, 20(4), 1320–1341. <https://doi.org/10.1177/1461444817691509>

Kahneman, D., Krakauer, D. C., Sibony, O., Sunstein, C., & Wolpert, D. (2022). An exchange of letters on the role of noise in collective intelligence. *Collective Intelligence*, 1(1). <https://doi.org/10.1177/26339137221078593>

Kerzner, H. (2017). Project management organisational structures. In H. Kerzner (Ed.), *Project management case studies*. Wiley. <https://doi.org/10.1002/9781119389040.ch4>

Kirschner, P. A., Buckingham Shum, S., & Carr, C. S. (2003). *Visualizing argumentation: Software tools for collaborative and educational sense-making*. Springer-Verlag.

Klein, M. (2007). *How to harvest collective wisdom for complex problems: An introduction to the MIT Deliberatorium*. <https://doi.org/10.13140/RG.2.2.32743.24489>.

Klein, M. (2011a). Enabling large-scale deliberation using attention-mediation metrics. *Computer Supported Cooperative Work*. <https://doi.org/10.2139/ssrn.1837707>

Klein, M. (2011b). The MIT deliberatorium—enabling large-scale deliberation about complex systemic problems. In *Proceedings of the 3rd international conference on agents and artificial intelligence (ICAART 2011)* Vol 1—Artificial Intelligence (pp. 15–24). Rome, Italy.

Klein, M. (2017). Towards crowd-scale deliberation. <https://ssrn.com/abstract=2987624>

Klein, M. (2021). Crowd-Scale Deliberation for Group Decision-Making. In: Kilgour, D. M., Eden, C. (eds) *Handbook of Group Decision and Negotiation*. Springer, Cham. https://doi.org/10.1007/978-3-030-49629-6_40

Klein, M., Spada, P., & Paulson, L. (2023). From shouting matches to argument maps. In S. Boucher, C. A. Hallin, & L. Paulson (Eds.), *The Routledge handbook of collective intelligence for democracy and governance* (pp. 190–201). Routledge. <https://doi.org/10.4324/9781003215929-14>

Kuhn, D. (1991). *The skills of argument*. Cambridge University Press.

Landemore, H. (2012). Collective wisdom: Old and new. In H. Landemore & J. Elster (Eds.), *Collective wisdom*. Cambridge University Press.

Landemore, H. (2012b). *Democratic reason: Politics, collective intelligence, and the rule of the many*. Princeton University Press: Princeton, NJ, United States.

Landemore, H. (2020). *Open democracy: Reinventing popular rule for the twenty-first century*. Princeton University Press.

Leitner, K. H., Warnke, P., & Rhomberg, W. (2016). New forms of innovation: Critical issues for future pathways. *Foresight*, 18(3), 224–237. <https://doi.org/10.1108/FS-07-2014-0050>

Lenart-Ganssiniec, R., & Sułkowski, Ł. (2018). Crowdsourcing—a new paradigm of organisational learning of public organisations. *Sustainability*, 10, 3359.

Levy, P. (1997). *Collective intelligence: Mankind's emerging world in cyberspace*. Plenum Press.

Levy, P. (1998). *Becoming virtual: Reality in the digital age*. Plenum Press.

Linders, D. (2012). From e-government to we-government: defining a typology for citizen coproduction in the age of social media. *Government Information Quarterly*, 29, 446–454.

Loomio Blog: Political Parties (2023). <https://www.loomio.com/blog/tags/political-parties>

Lund, A. E., & Russell, C. (2022). What is the relationship between collective memory and metacognition? In S. M. O'Mara (Ed.), *Progress in brain research* (Vol. 274, pp. 31–70). Elsevier.

Malone, T. (2015). Conclusion. In T. Malone & M.S. Bernstein (Eds.), *Handbook of Collective Intelligence*. The MIT Press: Cambridge, MA, United States.

Malone, T. (2018). *Superminds: The surprising power of people and computers thinking together*. Little, Brown and Co.

Malone, T. W., Laubacher, R., & Dellarocas, C. (2010). The collective intelligence genome. *MIT Sloan Management Review*, 51(3), 21–31.

Medialab Prado (2019). *Future democracies*. Laboratory of Collective Intelligence for Participatory Democracy. Medialab Prado. <https://archive.org/details/FutureDemocraciesLCPD>

Mercier, H., & Landemore, H. (2012). Reasoning is for arguing: Understanding the successes and failures of deliberation. *Political Psychology*, 33(2), 243–258.

Mergel, I. (2015). Opening government: Designing open innovation processes to collaborate with external problem solvers. *Social Science Computer Review*, 33(5), 599–612. <https://doi.org/10.1177/0894439314560851>

Miller J. (2011). Will Extremists Hijack occupy wall street? *New York Times*. http://www.nytimes.com/2011/10/26/opinion/will-extremists-hijack-occupy-wall-street.html?_r=0

Moher, D., Liberati, A., Tetzla, J., Altman, D.G., & Group, P. (2009). Preferred Reporting Items for Systematic Reviews and Meta Analyses: The PRISMA statement. *Ann. Intern. Med.*, 151(4).

Mountcastle, V. B. (1978). An organizing principle for cerebral function: the unit model and the distributed system. In G. M. Edelman, V. B. Mountcastle (Eds.), *The mindful brain*. MIT Press.

Mulgan, G. (2018). *Big mind: How collective intelligence can change our world*. Princeton University Press.

Mulgan, G. (2019). *Social innovation: How societies find the power to change*. Policy Press.

Neisser, U. (1967). *Cognitive psychology*. Prentice-Hall Inc.

Noveck, B. S. (2015). *Smart citizens, smarter state: The technologies of expertise and the future of governing*. Harvard University Press.

Noveck, B. S., Harvey, R., & Dinesh, A. (2019). *The open policymaking playbook*. The Governance Lab, New York University. <https://thegovlab.org/static/files/publications/openpolicymaking-april29.pdf>

Noveck, B. S., Dinesh, A., Muñozcano, B. R., Gambrell, D., Joerger, G., Gimeno, E., Konopacki, M., Ryan, M., Kornberg, M., Harvey, R., DeJohn, S., & Alsina, V. (2020). *Crowdlaw for congress: Strategies for 21st century lawmaking*. The Governance Lab, New York University. https://congress.crowd.law/files/crowdlaw_playbook_Oct2020.pdf

Observatory of Public Sector Innovation. Case Study Library—Better Reykjavik (2023). <https://oecd-opsi.org/innovations/better-reykjavik/>

Olszowski, R. (2015). Loomio. Collective Decision-Making. In A. Turczyn (Ed.), *E-government 2.0 in practice* (pp. 65–90). Wyższa Szkoła Europejska im ks. J. Tischnera. <https://doi.org/10.6084/m9.figshare.13649972.v1>

Olszowski, R., Pięta, P., Baran, S., & Chmielowski, M. (2021). Organisational Structure and Created Values. Review of Methods of Studying Collective Intelligence in Policymaking. *Entropy*, 23(11), 1391. <https://doi.org/10.3390/e23111391>

Origgi, G. (2012). Designing Wisdom through the Web. In H. Landemore, & J. Elster (Eds.), *Collective wisdom*. Cambridge University Press.

Ostrow, R. (2017). Featured commitment—Madrid. *Open Government Partnership*. <https://www.opengovpartnership.org/stories/featured-commitment-madrid/>.

Paulus, P. B., & Nijstad, B. A. (2003). Group creativity: Common themes and future directions. In P. B. Paulus & B. A. Nijstad (Eds.), *Group creativity: Innovation through collaboration* (pp. 327–348). Oxford University Press.

Presupuestos Participativos. (2023). La Edición 2021–2022 de los Presupuestos Participativos de Madrid ya Tiene Proyectos Ganadores. Decide Madrid. https://decide.madrid.es/presupuestos#budget_investments_list

Prpić, J., Taeihagh, A., & Melton, J. (2015). The fundamentals of policy crowdsourcing. *Policy & Internet*, 7(3), 340–361. <https://ssrn.com/abstract=2648778>

Raport z ewaluacji procesu wdrażania VII edycji budżetu obywatelskiego miasta Krakowa (2020). <https://budzet.krakow.pl/zalacznik/375080>

Raport Ewaluacyjny. VI edycja Budżetu Obywatelskiego Miasta Krakowa (2019). <https://www.bip.krakow.pl/plik.php?zid=259425&wer=0&new=t&mode=shw>

Raport Ewaluacyjny. Vedycja Budżetu Obywatelskiego Miasta Krakowa (2018). <https://budzet.krakow.pl/zalacznik/331234>

Raymond, E. S. (1999). *Surprised by Wealth, Linux Weekly News (1999-12-10)*, Retrieved from <https://lwn.net/1999/1216/a/esr-rich.html>

Raymond, E. S. (2001). *The Cathedral and the Bazaar: Musings on Linux and Open Source by an accidental revolutionary*. O'Reilly Media.

Royo, S., Pina, V., & Garcia-Rayado, J. (2020). Decide Madrid: A critical analysis of an award-winning e-participation initiative. *Sustainability*, 12(4), 1674. MDPI AG. <https://doi.org/10.3390/su12041674>

Russell, N. (2006). Collective memory before and after Halbwachs. *The French Review*, 79(4), 792–804.

Ryan, M., Gambrell, D., & Noveck, B. S. (2020). *Using collective intelligence to solve public problems*. Nesta.

Saebø, Ø., Rose, J., & Flak, L. S. (2008). The shape of eparticipation: Characterizing an emerging research area. *Government Information Quarterly*, 25, 400–428.

Sandel, M. J. (2012). *What money can't buy: The moral limits of markets*. Farrar.

Skaržauskienė, A. (2015). *Social technologies and collective intelligence*. Mykolas Romeris University.

Sifry, M. L. (2013). Can social software change the world? *Loomio Just Might*. forums.e-democracy.org/groups/oddi/messages/topic/6PsAkxxvxwpJylz7LOJqV

Silvertown, J. (2009). A new dawn for citizen science. *Trends in Ecology & Evolution*, 24(9), 467–471.

Stahl, G. (2006). *Group cognition: Computer support for building collaborative knowledge*. MIT Press.

Stahl, G. (2009). *Studying virtual math teams*. Springer Verlag.

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of intelligence*. Cambridge University Press.

Sternberg, R. J., & Sternberg, K. (2012). *Cognitive psychology* (6th ed.). Wadsworth, Cengage Learning: Belmont CA, United States.

Steyvers, M., & Miller, B. (2015). Cognition and collective intelligence. In T. Malone & M.S. Bernstein (Eds.), *Handbook of collective intelligence*. The MIT Press.

Surowiecki, J. (2005). *The wisdom of crowds*. Anchor Books.

Taeihagh, A. (2017). Crowdsourcing: A new tool for policy-making? *Policy Sciences*, 50(4), 629–647. <https://doi.org/10.1007/s11077-017-9303-3>

Theiner, G., & Goldstone, R. L. (2010). Recognizing group cognition. *Cognitive Systems Research*, 11, 378–395.

Tomasello, M. (2014). *A natural history of human thinking*. Harvard University Press.

Torraco, R. J. (2005). Writing integrative literature reviews: Guidelines and examples. *Human Resource Development Review*, 4, 356–367. <https://doi.org/10.1177/1534484305278283>

Tulving, E. (1983). *Elements of episodic memory*. Oxford University Press.

Verhulst, S. G., Zahuranc, A. J., & Young, A. (2019). Identifying citizens' needs by combining AI and CI. *New York University*. https://thegovlab.org/static/files/publications/CI-AI_oct2019.pdf

Vincenzo, I., Giannoccaro, I., Carbone, G., & Grigolini, P. (2017). Criticality triggers the emergence of collective intelligence in groups. *Physical Review E*, 96, 022309.

Wampler, B. (2007). A guide to participatory budgeting. In A. Shah (Ed.), *Participatory budgeting* (pp. 21–54). The World Bank. <https://doi.org/10.1596/978-0-8213-6923-4>

White, J. (2007). *Managing information in the public sector*. M.E. Sharpe.

Wijnhoven, F., Ehrenhard, M., & Kuhn, J. (2015). Open government objectives and participation motivations. *Government Information Quarterly*, 32(1), 30–42. <https://doi.org/10.1016/j.giq.2014.10.002>

Williams, G. C. (1966). *Adaptation and natural selection: A critique of some current evolutionary thought*. Princeton University Press.

Wspólne Projekty BO (2020). Podgórze Duchackie 2020. <https://www.facebook.com/groups/BO2020.XI.Podgorze.Duchackie>

Wydział Polityki Społecznej i Zdrowia Miasta Krakowa. (2020). Budżet Obywatelski Miasta Krakowa 2020. <https://budzet.krakow.pl/zalacznik/372206>

Yin, R. K. (2009). *Case study research: Design and methods* (4th ed.). Sage.

Yu, L., Nickerson, J. V., & Sakamoto, Y. (2012). *Collective creativity: Where we are and where we might go*. Preprint at [arXiv:1204.3890](https://arxiv.org/abs/1204.3890)

Chapter 4

Online Public Debate. How Can We Make It More Intelligent?



A crucial element of intelligence is the ability to learn from one's mistakes.

H. Landemore (2012, p. 240)

4.1 The Online Public Sphere and its Problems

4.1.1 *Erosion of the Traditional Public Sphere*

In the preceding chapter, we examined various online projects dedicated to specific aspects of public policy that were conducted with the involvement of open communities. The online debate on public issues has a much broader dimension than specialized projects, however. The Internet has democratized the dissemination of information. With platforms like YouTube, blogs, and social media sites, more people than ever have the ability to share their views and contribute to public discourse. In the twenty-first century, modern Internet media, particularly social networks, have become a substitute for the traditional public sphere. This space serves as the arena for civic debates, which have multifaceted effects on policymaking.

The dominant views within social sciences regarding the role of civic debates in policymaking were shaped by the opposition to the purely aggregative model of Joseph Schumpeter and Robert Dahl, who gave a public opinion in the role of aggregation of preferences and interests existing in society rather than creating any new values in the debate (Paxton, 2020). In contrast to this approach, a debate-oriented model was proposed by Jürgen Habermas and John Rawls, emphasizing the importance of debate in shaping civic attitudes, involvement in community life, and creating social bonds. As conceptualized by Habermas, the public sphere refers to the realm within social life where public opinion can be formed and accessible to all. It is a space where individuals unite to have dialogues about societal problems and eventually influence political action. We can call it a realm of public political discourse that acts as an intermediary among other societal subsystems, such as

the economy, politics, culture, and personal life. In the archetypal representation of the public sphere, it facilitates “critical public debate” (Habermas, 1991, p. 52). This sphere can be defined as the domain of discourse where individuals exchange opinions and engage as legal subjects, citizens, economic actors, and/or members of families and communities. Within this contested participatory site, these actors form a political body, negotiate, and contest various aspects of political and social life (Somers, 1993). This dynamic discourse consequently shapes shared norms of consensus and agreement, forming a collective public opinion that can influence institutional systems (Pietrzyk-Reeves, 2012).

The relationship between the public sphere and public policies is deep and interwoven, with one often influencing the other in various capacities. Public policies do not exist in an ideological vacuum. The public sphere serves as the arena for their formation, a place for debate, where opinions are shaped, and various stakeholders influence policy decisions. From the nineteenth-century coffee house debates to mass media, community forums, town halls, and NGOs, individuals discuss their views, get exposed to diverse perspectives, and gradually form opinions. Public policies are often swayed or shaped by this collective public sentiment, indicating the importance of an active and engaged citizenry. On the other hand, once a policy is implemented, the public sphere acts as a feedback mechanism. People’s experiences, critiques, and suggestions related to the policy are aired out in the public domain. Governments can then take this feedback and refine the policies or institute new ones, thus ensuring that public policies remain responsive and dynamic. A vibrant public sphere can also serve as a check on the misuse of power. When policies are perceived as unjust, discriminatory, or favoring a particular group, citizens can mobilize, discuss, and protest. For instance, the 2018 Yellow Vests Movement in France, which began in response to rising fuel prices and living costs, expanded into a broader movement against economic inequality and President Macron’s policies. Similarly, the 2019–2020 demonstrations in Hong Kong, aimed at preserving the city’s autonomy and civil liberties, stand as another testament to how the public sphere pushes for policy changes in the face of perceived encroachments by Mainland China.

In recent years, it has become clear that the traditionally understood *public sphere*, comprised of media such as newspapers, radio, and television, is eroding. Old broadcasters are becoming a thing of the past, giving way to online media, and old forms of civic activity in public spaces are also fundamentally changing. At the same time, due to the increasing number of social areas that the state deals with, ties between the administration and citizens are being severed. Areas of social life that were once less interesting to the government and more closely connected to the experiences of individual citizens are now becoming more distant from them. The shaping of public policies, which, according to the original ideas of Habermas and Rawls, was supposed to be based on a well-developed public sphere, has now, in practice, become reliant on arbitrary decisions made by state administrations, resembling a bureaucratic machine rather than an intelligent mechanism. Elected politicians also frequently fail to exert genuine control over the bureaucratic apparatus, becoming dependent on the operational modes of the institutions.

The erosion of the traditional public sphere applies to most countries to a greater or lesser extent. Traditional media outlets, which once played a crucial role in the public sphere, are increasingly being purchased by large corporations or business tycoons. This consolidation of media power poses a significant threat. Firstly, these monopolistic tendencies restrict the diversity of content available to the public, fostering a homogenized perspective rather than a rich tapestry of varied views. Secondly, such a concentration jeopardizes journalistic autonomy, with editorial decisions possibly bending to cater to corporate interests rather than unbiased reporting. In Australia, for instance, News Corp, spearheaded by Rupert Murdoch, exerts considerable dominance over the media landscape, shaping much of the nation's discourse and steering political sentiment. The company's influence has sparked debates on media pluralism and democratic health. In the United States, on the other hand, entities like Sinclair Broadcast Group own numerous local television stations and occasionally disseminate centralized messaging, which can limit local autonomy and raise concerns about nationally coordinated narratives on local platforms. Most importantly, local news outlets, which traditionally covered community-level issues, are declining or shutting down. This reduces the public's engagement with local affairs and erodes the local dimension of the public sphere; these local sources, once the heartbeat of community-level discussions and critical checks on local governance, are fading. This decline results in communities being less informed about local happenings, weakening civic participation and lessening local accountability. In the UK, the closing down or reduction in publication frequency of several local newspapers has resulted in *news deserts*, where communities lack any local news coverage. This lack of local reporting makes them more susceptible to misinformation and reduces their engagement in local governance. Similarly, in the US, over a quarter of the country's newspapers have shuttered in the past 15 years. This trend has left many communities without any local news source, impairing the public's knowledge about key local issues and undermining participation at the community level.

The problem with a low level of social capital and low involvement in community affairs is particularly visible in the post-communist countries of Central and Eastern Europe. There, we can observe a low level of trust in political institutions, the loosening of social bonds, weakened norms of consensus and agreement, low quality of public debate (Howard, 2003), and, as many authors believe, the problematic quality of social capital (Chloupkova et al., 2003). For example, in Poland, low levels of civic participation and the poor quality of social capital have become an acute social problem. Poles' social activity (participation in non-governmental organizations) is at 14.2%, making this country the second last among all the European Union Member States. The percentage of people who trust others is at the bottom among the European states and is lower than 20%. This deficit in the civil attitude is particularly acute among young people, whose sense of social community weakens while interest in only their own well-being grows (Czerwiński, 2014). It is essential to note that historically, restrained social activity was partly a result of the mandatory nature of membership in formal associations in communist times. This was further compounded by widespread disappointment stemming from the social, political, and economic transformation outcomes following the fall of communism. Nonetheless,

the evolving socio-political landscape has ushered in new dynamics in recent years. Neighboring Ukraine has witnessed anti-government movements like *Euromaidan* and social unrest due to Russian aggression in Crimea and the Donbas. Similarly, Poland's social fabric has been affected by growing political polarization and, at the same time, reshaped by substantial migrations. These migrations include immigration, primarily from Ukraine, and emigration to Western European nations. These rapid demographic changes have transformed Poland into a multi-ethnic society in a short time, significantly altering the nature of public discourse.

4.1.2 *The Online Public Sphere as an Alternative to the Traditional Public Sphere*

4.1.2.1 The Internet as a Multilevel Platform of Policy Discourse

The online public sphere has been treated as a potential alternative to the traditional one for several years. Nearly simultaneously, discussions regarding the impact of the Internet on the public sphere began with the rise of widespread Internet use in the mid-1990s. There is no doubt that in the digital age, where most of our interactions occur behind screens and within networks, the concept of public discourse has dramatically transformed. Several scholars have debated whether and to what extent digital media constitutes an alternative arena for political discussion. The dominant perspective suggests that online platforms can catalyze social mobilizations beyond mere digital interactions, enabling citizens to influence decision-making and public policy (Bennett, 2012; Dahlberg, 2007a, b). The technological change has a strong impact on the social fabric and, in many cases, modifies it irreversibly. Researchers persist in investigating and deciphering the evolving shifts in society and politics aligned with the ongoing metamorphoses of the media environment.

The online public sphere offers unprecedented opportunities for public participation and outreach. Information shared by online friends carries more weight for many people than a formal announcement by an official on mainstream media. Embracing the digital realm brings forth challenges such as maintaining content integrity, protecting against misuse, and addressing the potential spread of harmful content, whether it's libel, pornographic material, or extremist propaganda. The policy of data openness and transparency can also be a double-edged sword, promoting accountability on the one hand while, on the other hand, highlighting areas of concern that could be used as ammunition by political rivals.

But let's look on the bright side: thanks to the online public sphere, the citizens are actively participating in sharing real-time updates on important events, from natural calamities to corruption, thereby acting as on-the-ground reporters. This interactive dynamic boosts public participation and augments the government's ability to react promptly and effectively. Professionals in public administration are keen to tap into the enhanced civic engagement that these platforms offer. However, they

are wary of the implications for information governance and the necessity to strike a balance between openness and safeguarding privacy. For instance, platforms like Facebook, Twitter (now rebranded to X), and Instagram have become essential tools for national endeavors such as disaster response and public awareness campaigns. During emergencies like the 2005 Hurricane Katrina, the 2011 Tōhoku earthquake in Japan, or the 2018 California wildfires, and from February 2022, the reporting of war crimes committed by Russians in Ukraine, social media platforms facilitated the rapid spread of crucial information. They provided platforms where volunteers could organize aid, arrange shelter, share imagery, chart out affected areas, and even plan logistics for affected individuals (Hansen et al., 2020).

We can view the process of transitioning public debate to the Internet as an opportunity to elevate it to a higher level (Castells, 2012) or, conversely, as a threat to democracy and freedom of speech (Cohen & Fung 2021, p. 23). No matter our view, it would be a mistake to underestimate the depth of the Internet media's contribution to societal change. One of the leading researchers in digital era governance and politics, Helen Margetts, is aware of the threats but maintains a modestly positive view of the condition of the online public sphere. She contends that while social media aids collective action through numerous "tiny acts of participation," it also reshapes the dynamics of political discourse, introducing a new "chaotic pluralism," although the ramifications of this shift remain somewhat elusive (Margetts et al., 2016). It has become clear that there is no single, definitive *effect*: the utilization and outcomes of the internet for political, and indeed all purposes, are always dependent on various factors (Dahlgren, 2018). Digital tools enable unprecedented levels of communication and data exchange among people globally. These technologies allow individuals to create, distribute, and collaborate on content while tapping into a vast reservoir of information. Platforms like Google and Facebook seem to liberate information and people from the constraints of old institutions, allowing us unrestricted access to and processing of information whenever and however we choose. They overcome traditional barriers like geography and expense. Social media platforms, in particular, have been instrumental in facilitating social movements such as the Arab Spring, France's Yellow Vests, Hong Kong's Umbrella Revolution, and various protests against sanitary restrictions in the U.S. and Europe (Bernholz et al., 2021 p. 1).

Due to the societal changes stemming from the proliferation of technology, the Internet has become a multilevel platform of policy discourse, eliminating the previous unilateralism and asymmetry in the relation between a sender (political actor) and a recipient (citizens), which was reinforced by the traditional media and the political practice to date. Thanks to the impact of the Internet, the idea of communalism and participative action is undergoing a redefinition, and the Internet has become a meta-medium, which—more than any previous mass media—changes its nature (Peisert & Stachura, 2011). At its core, the primary technological shift from the mass media era to the digital age is the transition from a broadcast (one-to-many) communication style to a networked (many-to-many) approach. This change comes with negligible incremental costs for information dissemination and communication. The digital framework of the public domain is characterized by this unique

information flow, with a significantly larger number of content creators and distributors. As a result, consumers have a broader range of information sources and types to choose from. Additionally, specific content can be tailored and directed toward individual users or user groups by content providers, advertisers, social media platforms, and other entities (Cohen & Fung, 2021). The blurring of the traditional border between senders and receivers and the spread of new forms of deliberation accompany this change. Digital technologies facilitate the creation of virtual communities that go beyond conventional community borders, allowing individuals worldwide to join based on shared interests. The online, participative model of social life shows a relatively low entry barrier for civic involvement. Technology contributes to the generation of both new areas of the public sphere and new ways of deliberation. Online, the lines blur between various social behaviors, roles, systems, and audiences, allowing individuals to engage in multiple roles across diverse practices and public spheres, all through a single online profile.

Internet media, especially those identified with the umbrella term *social media*, are altering the conventional dynamics between media creation and consumption. Social media refers to a broad category that encompasses social networking sites like Facebook, X/Twitter, and LinkedIn; video hosting platforms like YouTube; image-sharing services like Flickr and Instagram; blogging sites; messaging applications like Telegram; and platforms for live streaming and collaborative wikis including Wikipedia. Unlike traditional broadcasting, where production and consumption are distinct roles, the Internet also enables consumers to act as information producers, commonly known as *prosumers*. Platforms like YouTube, a leader in video content publishing, exemplify this shift by turning audiences into creators. This transformation not only broadens the educational and informational role of public service broadcasting but also evolves it into a participatory mandate. In this setting, participation signifies providing an online platform that allows citizens to publicly share their own audio-visual content (Fuchs, 2023, p. 250). In today's world, it is essential for politicians, political parties, NGOs, and social movements to maintain a presence on social media platforms.

However, it should be noted that while digital communication technologies generally increase the potential for participating in political debate, most citizens consume rather than produce information, and only a minority become *prosumers* (Sampedro & Martinez Avidad, 2018). The hopes expressed by Yochai Benkler in the early stages of social media development, suggesting they would allow a mass reorientation of citizens from passive listeners to potential speakers (Benkler, 2006, p. 213), have been realized for only a small percentage of internet users. The majority remain receivers, not broadcasters, though they do exhibit signs of activity (e.g., liking and retweeting).

Another problem is the public sphere fragmentation. Dahlgren (2005) asserts that in today's digital era, the concept of a *singular public sphere* has been replaced by multiple, fragmented public spheres. While the rise of the Internet and platforms like Twitter or Facebook have intensified this fragmentation, they are not its sole cause. Even though these varied spheres signify a weakened unified public sphere, they do not negate its existence entirely. Discussions on social media mold our general

political perspectives. Moreover, the beliefs and views formed online, especially regarding opponents, extend beyond the digital realm (Collins et al., 2020).

The online public sphere's most evident strength lies in its accessibility and reach. In contrast to traditional media that were bound by geography, the digital realm offers instant global access. Such interactions foster a more globalized form of public discourse, introducing varied cultural and socio-political perspectives into the mix (Cohen & Fung, 2021). Additionally, the online environment provides a platform for voices that may have been marginalized or suppressed in twentieth-century mass media settings. Empirical findings of Gentzkow and Shapiro proved that online media does better than offline media in exposing people to varied news sources, allowing them to see views that are distant from their own (Farrel & Schwartzenberg, 2021). The anonymity that the internet offers can embolden individuals to share personal experiences, perspectives, and critiques without the fear of immediate backlash or societal judgment. This dispersion of voice ensures a more diverse and inclusive discourse, pivotal for a holistic public sphere. At first glance, the digital public sphere seems to offer distinct benefits in terms of freedom of expression and the diversity of viewpoints presented. Traditional mass media is often limited to information which is produced by journalists and editors working within certain ideological and social frameworks. The technical capabilities of the digital sphere allow for a much wider array of voices to be heard by large audiences. Consequently, this increased level of expression results in a broader range of information sources and viewpoints that people can easily access, either intentionally or incidentally, within the public domain (Cohen & Fung, 2021).

While in the era before the Internet, the mass media-driven public sphere communicated with a distinct editorial tone and perspective, platforms like Facebook, YouTube, and X/Twitter allow users to distribute content, consume it, and connect with each other in a completely free manner, at least in theory. However, social media platforms are not just simple conduits for information. They engage in activities like curating, moderating, amplifying, or even manipulating user behavior (which we will discuss later), but this role is different from the editorial role of traditional mass media (Cohen & Fung, 2021). Certainly, the digitally mediated public sphere provides access to a wider spectrum of perspectives, though it also raises concerns about a novel form of *covert* manipulation, where it is unclear who ultimately takes responsibility for the information and the resulting behaviors. In the old-style media, it was clear that the editorial staff was responsible for the quality of information, although they supplemented it with their own point of view. Now, it is not clear at all.

The process of public sphere transformation is often perceived as a transition from the strong-gatekeeper model of the large mass media semi-oligopolistic organizations, especially TV broadcast networks and large-circulation newspapers, in which a relatively small number of organizations dominated the production of news, to a model of decentralized broadcasters under algorithmic curations (Farrell & Schwartzenberg, 2021). Strong gatekeeper systems inherently limit the inclusion of diverse voices in political discussions. Those with unconventional political views often find it challenging to gain media visibility in such systems. On the other hand,

weak gatekeeper systems do not impose a rigid societal consensus. They accommodate a wide array of individuals with varying beliefs about truth, allowing them to actively participate in public discourse and even establish their own media outlets or publishing platforms. In an environment dominated by weak gatekeepers, where practically everyone can be a content creator or distributor, a broader spectrum of individuals can attain a position of knowledge authority. As traditional indicators of expertise become less influential in determining who can enter the media landscape, there is potential for the rise of new, alternative sources.

4.1.2.2 Social Network Analysis as a Method of Calculating the Features of an Online Debate

Another distinguishing feature of the online public sphere as a space for debate, interaction, and social relations is its significantly advanced ability to track *hard* data concerning collectives far beyond what was possible in the traditional public sphere. In the pre-internet era, social interaction studies had to rely on selective surveys or were conducted within chosen *representative* communities. In contrast, phenomena occurring in this new kind of public sphere are highly measurable, and the results can pertain to the entire population. The substantial societal changes resulting from online interactions are parallel to the progress in research tools and methods. These advancements are geared toward comprehending and documenting various social media structures and the critical events that can influence these collective frameworks.

An advantage of using social media data in public sphere research is the possibility of obtaining data more quickly compared to conducting surveys or interviews. This allows for the analysis of emerging events, crises, or political challenges almost immediately after their occurrence. Researchers of online social networks have the ability to investigate interactions, responses, and attitudes of large groups of people without worrying whether the respondents' memory is failing them or whether the study itself distorts the responses, as was the case in past survey research.

Social network analysis is a fairly new scientific approach for examining and interpreting the connections between various online entities. This method offers effective techniques for encapsulating networks and pinpointing key figures or elements that hold influential roles within a grid of connections: "Social network analysis offers a systematic method to evaluate social media efforts, replacing anecdotes with scientifically based evidence" (Hansen et al., 2020, p. 9). Social scientists have dedicated many years to refining techniques for social network analysis and visualization. The foundational concepts and methods of this field trace their roots back to the 1930s when Jacob Moreno and his collaborators at New York University, Columbia, and Harvard embarked on pioneering work. Their study of relationships among members of a football team laid the groundwork for the core components of modern social network theory, including measures, maps, and models. In 1934, Moreno introduced a classification technique that involved calculating interpersonal relationships using

metrics and algorithms. This technique not only provided valuable insights into individual connections but also offered a means to analyze important network properties of the entire community represented on a graph (Moreno, 1934).

Social network analysis, previously limited in scope and reliant on manually gathered and processed data regarding social connections, saw a resurgence with the rise of computer networks. The increasing tracking and recording of our interactions and affiliations by mobile devices and social media platforms have made network analysis progressively more valuable. This analytical approach, employing mathematical techniques to systematically understand networks, introduces the concept of vertices (representing individuals) linked by edges (symbolizing social connections like friendships). Network metrics play a crucial role in pinpointing individuals occupying central or influential positions within a network, identifying closely interconnected subgroups (known as network clusters), and discerning key characteristics of the network's overall structure, such as its density.

Network data is typically represented in one of two formats: an edge list or a matrix. In directed networks, edges have distinct points of origin and destination, while undirected networks lack such directionality. Edges, also referred to as links, ties, connections, or relationships, serve as the connective threads of networks, linking two vertices together. These edges can represent various types of relationships, such as proximity, collaboration, kinship, friendship, partnerships, citations, transactions, or shared attributes. Weighted networks feature edges with associated values, providing additional nuance to relationships. Often, edges correspond to the fundamental elements of social media platforms, like friends on social networking sites or actions like forwarding posts, replies, and quotes. These edges establish connections between vertices, which are also known as nodes or agents. Typically, these vertices represent individuals or social structures, such as workgroups or organizations. Associated with vertices is attribute data, which can describe demographic characteristics of individuals, such as age, gender, or race, as well as data related to their interaction with a system, such as the number of logins, messages posted, or edits made. Additionally, attributes can include metrics like the number of followers each user has, providing insights into popularity within the network, and the number of people each user follows (Hansen et al., 2020).

The strongest advantage of this method in the context of online public discourse analysis lies in its use of universal, quantifiable metrics. For example, one such metric applicable to the entire network is *density*, which measures the level of connectivity among a group of vertices by calculating the percentage of observed connections compared to the maximum possible connections if everyone were connected to everyone else. Other metrics that apply to the entire network include modularity, geodesic distance, and reciprocated edge ratio. Many metrics are also computed for each individual vertex in the network. For instance, various *centrality* measures, each based on objective criteria, gauge the significance or centrality of a vertex within the network. One such measure is *betweenness centrality*, which focuses on connection paths—measuring the shortest number of neighbor-to-neighbor hops required to connect two individuals who are not direct neighbors (White & Borgatti, 1994). The shortest distance between two people is known as *geodesic distance*.

Closeness centrality assesses the distance from strategically positioned individuals, while eigenvector metrics evaluate the quality of connections. To analyze user clustering, subgroups are separated, taking into consideration the frequency of mutual communication. This process identifies opinion leaders and the most prominent topics discussed within these subgroups. Clustering analysis can be carried out using the Clauset-Newman-Moore algorithm (Clauset et al., 2004) or, alternatively, the Wakita-Tsurumi algorithm (Wakita & Tsurumi, 2007). An example of a graph used in social network analysis is presented in Fig. 4.1.

Several metrics in social network analysis are valuable for estimating collective intelligence in online debates as they measure various aspects of network structure, influence, and information flow. Metrics like *betweenness centrality* and *closeness centrality* help identify influential individuals in the debate. If the debate is too unbalanced and dominated by only a few individuals, it can negatively affect the intelligence of the group. Users with high *betweenness centrality* act as bridges connecting different parts of the network and can control information flow. Understanding the subgroups revealed by clustering algorithms sheds light on diversity and echo chamber dynamics. For instance, in a controversial topic discussion forum, clustering

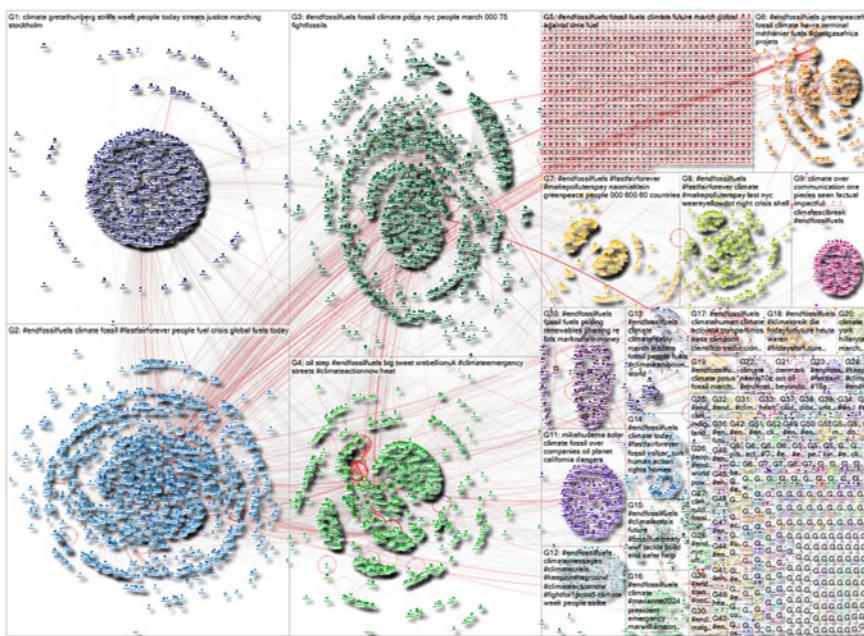


Fig. 4.1 An example of a graph used in social network analysis. This graph, generated using NodeXL software, illustrates a debate regarding fossil fuels that took place on the X/Twitter platform in September 2023. The graph was laid out using the Harel-Koren Fast Multiscale layout algorithm, and the vertices were grouped by cluster using the Clauset-Newman-Moore algorithm. Source <https://nodexlgraphgallery.org/Pages/Graph.aspx?graphID=292661>

analysis may unveil distinct communities of like-minded users. The metric of reciprocated edge ratio gauges the extent to which connections are reciprocated among users. A productive debate often features a higher reciprocated edge ratio, indicating that users engage in reciprocal discussions and build shared understanding. Higher density, on the other hand, implies smoother information exchange, while lower density may indicate polarization. A high-density climate change debate suggests active information sharing, while low-density hints at isolated groups. Generally, a more connected, diverse, and reciprocally engaged debate network can be conducive to collective intelligence. However, interpreting these metrics requires careful consideration of the specific debate context and participants, as their implications may vary based on the nature of the discussion and platform. We will revisit this topic later in the chapter and provide examples of how this method of analysis is employed to evaluate collective intelligence in social media debates.

4.1.3 *The Problems of the Online Public Sphere*

4.1.3.1 **Public or Non-public?**

While many believe that user-generated content provides everyone with a voice on social media, thereby contributing to a more diverse and participatory public discourse, the reality is nuanced. Although the Internet indeed facilitates easy content creation and sharing, and we witness the emergence of *prosumption* (where media consumers also become content producers), disparities exist in the visibility and attention garnered by specific users. Entertainment or propaganda tends to overshadow educational and political discourse, making social media lean more toward digital sensationalism than serious discussion. Numerous instances show that internet platforms have not truly fostered a participatory environment; instead, media giants and celebrities largely dominate online attention and the digital public space. (Fuchs, 2023, p. 287). Habermas, one of the most renowned theorists of the public sphere, expresses skepticism about the Internet's role in its enhancement. He believes it divides the public into a myriad of specific “issue publics” (Habermas, 2006). In a more recent analysis, Habermas views certain studies on the public domain as evidence supporting his belief that the Internet and social media have led to “semi-public, fragmented, and self-circulating discussion,” thus distorting the public sphere (Fuchs, 2023, p. 294).

In the mass-media-based, strong-gatekeeper model of the public sphere, it was clear who held the responsibility for disseminating or withholding specific information. In contrast, the contemporary model is less transparent, with unspoken rules. While it may appear that there is unrestrained freedom to publish online, this freedom is often overshadowed by the temptation of *clickbait* headlines—enticing titles designed to attract readers to superficially satisfying narratives. Furthermore,

the undisclosed affiliations and biases of major internet corporations further complicate matters. A small number of social media providers and digital companies, sometimes in collaboration with governments, exert significant control over the majority of online interactions. Borrowing from Habermas's terminology, these corporate algorithms have effectively *colonized* the digital public space to serve the interests of global corporations and influential political entities. Tech companies employ artificial intelligence and extensive datasets to promote their products, while users of these platforms tend to prioritize their personal interests over engaging in meaningful civic discourse (Sampedro & Martinez Avidad, 2018).

The influence and value of social media platforms will grow as more people use them, thanks to network effects. This phenomenon has led to the dominance of global giants like Meta (the parent company of Facebook, Instagram, and WhatsApp) and Alphabet (the parent company of Google and YouTube). These corporations serve as primary gateways to online information, wielding significant power. From both individual and national perspectives, a handful of these companies play a pivotal role in shaping our digital information environment through their algorithms. These firms' guidelines have a greater influence on freedom of speech than decisions made by national governments. As Timothy Garton Ash (2017) puts it, these major tech corporations are the "private superpowers" of today. However, such immense power also carries risks, including the potential for misuse, manipulation, and widespread surveillance (Bernholz et al., 2021, p.2). Given the rapid pace of digital information consumption, our attention spans are stretched thin by a constant influx of shallow content delivered at high speed. This leaves little time and space for self-reflection to detect potential manipulation or for meaningful conversations on social media.

Unlike the traditional public sphere, where journalists and editors curated stories to highlight what they considered essential for public discourse, the digital public sphere often amplifies distractions and targeted hostility. Furthermore, the "signal-to-noise" ratio in the digital realm may be less favorable than in the era of mass media. Consequently, navigating the digital public sphere requires substantial effort to distinguish reliable information from propaganda, advertisements, or mere junk (Cohen & Fung, 2021). Despite apparent progress in terms of diversity, expression, and accessibility, the digital public sphere is frequently filled with irrelevant chatter, misleading information, and hateful comments. The compartmentalization of differing viewpoints, each reinforcing its own beliefs in opposition to others, can potentially limit the benefits that seem to come with the digital landscape.

Finally, a critical aspect to consider when evaluating whether the digital space for debate truly functions as a public sphere is the ownership of the platforms where most discussions take place. Although, in theory, there is flexibility to switch platforms and choose one's preferred environment for online discussions, in practice, a small number of major players in the social media market hold dominant positions. These platforms are profit-driven entities, with their primary goal being user engagement and ad revenue rather than fostering public discourse. This commercial focus can often clash with the ideal of an impartial public sphere. Furthermore, the data-centric approach of these platforms raises concerns about user privacy, as individuals often

lack full control over their data, potentially dissuading some from engaging openly and honestly.

Interestingly, both Facebook and Twitter's management teams have long aimed to portray their platforms as genuine, free public spheres—virtual agoras extending ancient patterns of debate. This portrayal has persisted despite their content oversight (sometimes extensive), covert practices like shadowbanning, or the manipulation of user behavior through behavioral targeting. Facebook founder Mark Zuckerberg has consistently expressed his desire for the platform to foster an informed and civically engaged public, often asking, "How do we help people build a civically engaged community?" (Collins et al., 2020). Similarly, in his September 2018 testimony before the United States Senate Committee on Intelligence, Jack Dorsey, the long-time CEO of Twitter, described the platform as a means to "serve the public conversation." He referred to Twitter as a "global town square" and a "trusted and healthy place that supports free and open democratic debate" (Collins et al., 2020; Dorsey, 2018).

From a legal perspective, both X/Twitter and Facebook undeniably belong to the private sector, and they own all the content published on their platforms according to their terms of service. Considering this, if we discuss the structural transformation of the public sphere, Frank Pasquale, a professor of law at Brooklyn Law School, in his book *The Black Box Society: The Secret Algorithms That Control Money and Information*, claims that this transformation involves commodifying the platforms of public talk and using automated recommendation systems to prioritize content presentation (2015). Following this line of thought, the phenomenon of moving public debate to these platforms could be termed the privatization of public discourse by private companies. However, there are suggestions that, due to their widespread presence and indispensability, social media platforms like Facebook and X/Twitter have essentially evolved into public utilities. As the use of these platforms becomes less of an option and more of a necessity, the argument goes that they should be subject to public regulations, similar to other essential utilities that citizens rely on (Collins et al., 2020). Barber (2011) contends that as corporations privatize significant portions of the Internet, its potential as a forum for free speech diminishes. He expresses concern that such privatization may pave the way for monopolization, which, in turn, can result in unequal access to crucial civic and cultural resources.

So, how do we resolve this dilemma, public or private? If we treat discourse on social media as a new form of the public sphere without reservations, it should certainly be subject to certain rules and regulations, imposing values such as impartiality, fairness, and freedom of speech. But how can we justify the obligation for private entities to implement such norms when the platform and content belong to them? (Collins et al., 2020). They proposed that drawing on Elinor Ostrom's study (1990), debates on social media platforms should be treated as common-pool resources. This theory originally pertained to the wise use of resources accessible to all members of society, such as pastures or fishing grounds, where one user's consumption diminishes another user's consumption. In such scenarios, it is challenging, albeit not impossible, to exclude potential beneficiaries from utilizing the resource. Resources that are rival (where one user's consumption impacts another's)

and non-excludable (where it is challenging to prevent individuals from accessing them) warrant special attention. As early as 1990, Ostrom suggested the potential of online public discourse to be viewed as a common-pool resource (Ostrom, 1990). The concept of the internet as a *digital commons*, a communal resource where information, tools, and spaces are accessible to everyone but must also be safeguarded as a shared value, has been explored and endorsed by scholars like Lessig (1999), Benkler (2006), and others. Collins and his collaborators further expanded on this idea, taking into account the contemporary influence of social media giants in online discussions.

According to these scholars, Twitter and Facebook's success is partly attributed to their ability to leverage the benefits of a healthy public discourse, a characteristic of developed civil societies. Since both companies are headquartered in the United States, both could *free-ride* on the benefits of a developed public sphere in this country, so they could use it to fill their platforms with content. However, as Collins argues, the public sphere, like other commons, can be over-exploited. Toxic conversations on Twitter and Facebook deplete this valuable resource, making it inaccessible for current and future users and diverting the attention of engaged citizens. In the realm of social media, attention has become a limited and crucial resource. The concept of *attention economics* views attention as a scarce commodity amid information overload, seeking efficient ways to manage and allocate it. The scarcity of attention becomes apparent in the presence of toxic discussions and information noise on social media platforms, leading to cognitive overload and distracting us from constructive debates on important public issues. When faced with a cacophony of voices, opinions, and misinformation, sifting through and identifying what deserves our focus becomes challenging. The inundation of toxic discourse on social media not only disperses our attention but also hinders our ability to engage effectively and participate in meaningful societal conversations. Safeguarding and effectively directing our attention become essential to ensure productive engagement in public debates. Therefore, platforms like Facebook risk over-exploiting the public sphere, akin to depleting a common resource. Toxic discussions limit civic engagement and degrade the quality of public discourse, both online and offline. Consequently, it is crucial to protect and manage our attention to enable productive participation in public debates.

4.1.3.2 Behavioral Targeting, Opinion Polarization, and Conditioning Human Behavior

One of the most significant concerns regarding the role of social media in public debates is their practice of behavioral targeting and the resulting potential for manipulating human behavior. Shoshana Zuboff, a renowned social psychologist, offers a compelling critique of this phenomenon in her book *The Age of Surveillance Capitalism*. Behavioral targeting involves the collection and analysis of user data,

including browsing habits, preferences, and online interactions, to deliver personalized content and advertisements to individuals. On the surface, this may seem harmless and even convenient, as it aims to provide internet users with content tailored to their interests. However, Zuboff argues that beneath this seemingly innocuous exterior lies a more perilous reality. According to Zuboff, companies like Google and Facebook have transformed user data into a valuable asset, commodifying human experiences for profit. They employ sophisticated algorithms and predictive analytics to anticipate users' desires and manipulate their behavior to maximize engagement and, consequently, advertising revenue. With access to behavioral data, Google can discern what a specific individual thinks, feels, and does at a given time and place, enabling them to tailor contextual advertisements accordingly. Users' lives are represented through this behavioral data, and Google's predictive capabilities directly correlate with the volume of data they accumulate.

A particularly alarming consequence of behavioral targeting is the erosion of user privacy. According to Zuboff, individuals are subjected to continuous surveillance, often without their explicit consent or knowledge. Social media platforms track users across the web, collecting vast amounts of data, including personal information, which is then used to create detailed profiles. This data is not only employed for advertising purposes but can also be leveraged for various other objectives, such as political manipulation and social control. Zuboff argues that behavioral targeting creates a *shadow text*, an alternative reality constructed by algorithms that shape the information users encounter. This, she contends, leads to the formation of *filter bubbles* or *echo chambers*, where individuals are primarily exposed to content that reinforces their existing beliefs and opinions, resulting in an echo of similar views. This phenomenon further disconnects political discussions from reality. In these online filter bubbles, users are divided into fragmented groups structured around shared opinions, where any disagreements are either non-existent or consciously evaded. This selective exposure can polarize society and hinder meaningful discourse, as users are less likely to encounter diverse perspectives.

Eli Pariser (2011) primarily associates this phenomenon with the rapid proliferation of machine learning-based business models used by companies like Google and Facebook. It is driven by the interplay between individual preferences and machine learning algorithms, resulting in algorithms consistently removing content that contradicts individuals' existing beliefs. For example, if a liberal user is less likely to click on news articles shared by conservative friends compared to those shared by fellow liberals, the news feed is likely to reduce the presence of conservative-leaning stories gradually. This can create a self-reinforcing pattern, leading to the formation of a bubble where challenging content is screened out. As a result, individuals may only encounter information or viewpoints that align with their pre-existing beliefs. Furthermore, individuals may remain unaware of this bias if they heavily rely on algorithmic sources for their information consumption.

A slightly different perspective on the phenomenon of online polarization is the “bottom-up” approach. According to Cass Sunstein (2017), the nature of online communication inherently fosters group polarization. This is because the online environment streamlines the process for individuals to connect with peers of similar

mindsets while distancing themselves from opposing views, thus intensifying group extremes. Sunstein argued that the rise in homophily, where individuals with similar mindsets congregate and primarily consume information that aligns with their pre-existing biases, promotes heightened extremism. This potentially leads to many people leaning toward online communities and conversations that resonate with their perspectives, thereby accentuating their extremist leanings through continuous engagement with like-minded individuals. Regardless of whether we attribute the trend of opinion polarization in online communication to inherent homophily or to algorithm-induced information bubbles, it's indisputable that major social media platforms benefit from this phenomenon. They reinforce user dependency on their services and inadvertently promote polarization through their operational mechanisms.

Another troubling consequence highlighted by Zuboff is the manipulation of human behavior. She contends that companies like Google and Facebook are becoming increasingly adept at gathering and processing what she terms "behavioral surplus." They leverage this data to make more precise predictions about users' behavior. Corporations are not only focused on extensive data collection but also on its depth and breadth, tapping into users' personal data through the Internet of Things. This behavioral surplus is crucial for creating products that aim to predict behavior with near certainty. Yet, Zuboff suggests that it doesn't end there; true forecasting of behavior means having control over it. Social media platforms utilize psychological insights to exploit human vulnerabilities, often nudging users toward spending more time online and engaging with specific content. By feeding on people's insecurities, fears, and desires, these platforms can exert undue influence over users' decisions, ultimately compromising their autonomy. With access to behavioral data, companies like Google can discern what an individual is thinking, feeling, and doing at a specific time and location, allowing them to tailor contextual advertisements to that person. Users' lives are represented through this behavioral data, and the predictive capabilities of these companies are directly proportional to the amount of data they collect.

The goals and operations of automated behavioral modification were designed and controlled by companies to achieve their own revenue and growth objectives. Zuboff argues that this technology can modify human behavior as easily as adjusting the operation of devices. Context can be designed around specific actions, enforcing change by reinforcing particular behaviors—a concept B. F. Skinner, often regarded as the father of behaviorism, termed *operant conditioning*. Zuboff cites numerous statements from high-ranking Silicon Valley employees. One of them claims, "Conditioning at scale is essential to the new science of massively engineered human behavior. The goal of everything we do is to change people's actual behavior at scale. We want to figure out the construction of changing a person's behavior, and then we want to change how lots of people are making their day-to-day decisions" (Zuboff, 2019). Similarly, Andrew Ledyina, a former Facebook product manager, stated that "the fundamental purpose of most people at Facebook working on data is 'to influence and alter people's mood and behavior'" (Hill, 2014).

In swift succession, it became apparent that such a powerful tool for manipulating people's decisions could also be applied in the political realm. As early as 2012, researchers affiliated with Facebook published an article in the prestigious journal *Nature*, provocatively titled "A 61-million-person experiment in social influence and political mobilization" (Bond et al., 2012). In this controlled, randomized study conducted on the eve of the 2010 US congressional elections, researchers experimentally manipulated the content displayed to nearly 61 million Facebook users: "The 'social message group' ($n = 60,055,176$) was shown a statement [encouraging to participate in elections] at the top of their 'News Feed.' This message encouraged the user to vote, provided a link to find local polling places, showed a clickable button reading 'I Voted,' showed a counter indicating how many other Facebook users had previously reported voting, and displayed up to six small randomly selected 'profile pictures' of the user's Facebook friends who had already clicked the 'I Voted' button (...). The 'informational message group' ($n = 611,044$) was shown the message, poll information, counter, and button, but they were not shown any faces of friends. The control group ($n = 613,096$) did not receive any message at the top of their News Feed" (Bond et al., 2012). The results revealed that users who received a message containing pictures of their friends were about 2% more inclined to click the 'I voted' button compared to those who received just the basic information. According to the researchers, this experiment realistically resulted in approximately 340,000 additional people going to the polls who otherwise wouldn't have voted. The researchers openly admitted that this experiment demonstrated "the importance of social influence for effecting behavior change" (Bond et al., 2012).

The seemingly innocuous and apolitical project sparked an intense debate as experts and the general public began to realize the unprecedented persuasive power of social media and their ability to influence behavior. Jonathan Zittrain, a specialist in internet law at Harvard, posited that one can imagine a situation where Facebook discreetly manipulates public debate and election results using means its users cannot detect or control. He accused Facebook of "digital gerrymandering: the selective presentation of information by an intermediary to meet its agenda rather than to serve its users." Zittrain complained that social media platforms, instead of presenting information impartially and encouraging users to debate and draw their own conclusions, "turned out to be shaping my experience according to its political agenda" (Zittrain, 2014). These allegations were fully confirmed a few years later in an affair involving Facebook users. However, the company collecting the data and influencing the users was, in that case, an external entity. The case became known as the Cambridge Analytica scandal and was described in detail by the company's former employee, Christopher Wylie (2019).

4.1.3.3 The Manipulation and Distortion of Online Debates: Cambridge Analytica and Twitter Files

Cambridge Analytica (CA) was a British political consulting company that specialized in data mining, data brokerage, and data analysis for electoral processes. In

2014, a researcher named Aleksandr Kogan developed an app called “thisisyourdigitallife”. While it was billed as a personality quiz for academic research, the app collected data from Facebook users and also scraped data from the profiles of the quiz-takers friends, leveraging Facebook’s loose data-sharing policies at the time. As a result, data was harvested from approximately 87 million users, the majority of whom had neither consented to nor been aware of such extensive data mining.

Cambridge Analytica acquired this data and used advanced data modeling to create detailed profiles of these users, targeting them with highly personalized political ads. The company claimed their methods could effectively “micro-target” individuals to influence their opinions on political matters. Most importantly for us, CA was hired to oversee major political campaigns, including a campaign related to the Brexit referendum and several election campaigns (e.g., Donald Trump’s presidential campaign in the US). The scandal came to light in 2018 when whistleblowers, including Christopher Wylie, and investigative reports by media outlets like The Guardian and The New York Times exposed the data misuse (Wylie, 2019).

In the wake of the CA scandal, Facebook came under severe scrutiny, experiencing widespread public backlash. Following the revelation of the scandal, the company’s stock price plummeted. Subsequently, they were slapped with multiple fines, notably a \$5 billion penalty from the Federal Trade Commission (FTC) due to privacy infringements. A pivotal moment during this period was Mark Zuckerberg’s testimony before the U.S. Congress in April 2018. In response to the unfolding events, Facebook implemented several modifications to its platform to avert similar future breaches, such as limiting third-party app access to user data and introducing a tool allowing users to view which apps could access their data.

The public’s reaction to Cambridge Analytica’s activities and its connections with Facebook was expressed through both demonstrations and online actions. One notable example is the use of the hashtag #DeleteFacebook on Twitter, which encouraged users to either leave Facebook or cease using it for a significant duration. However, an analysis of the Facebook user data suggests that the campaign was not successful. Despite the #DeleteFacebook campaign in 2018, Facebook did not experience a significant drop in user numbers. There was a slight dip in the user growth rate, potentially attributable to the campaign (Mills, 2020), but this trend quickly reverted to its prior state.

Even though the Cambridge Analytica scandal eroded trust in Facebook and other social media platforms for many users, making them more vigilant about the risks of targeted content manipulation, it seems that even several years after the incident, online public discourse remains susceptible to abuse. However, the nature of these abuses has shifted. In 2018, Facebook clamped down on access to its application programming interface (API), and in 2023, Twitter followed suit. These measures have significantly curtailed the influence of private entities seeking to mold public discourse on social media in accordance with their agendas. Yet, increased regulatory pressures by governmental bodies and platforms’ moderation practices, combined with covert forms of censorship, have become grounds for new abuses and different forms of shaping online discourse in line with political agendas, as exemplified by the recent ‘Twitter Files’ case.



Fig. 4.2 Elon Musk's post on collective mind on the X/Twitter platform. *Source* <https://twitter.com/elonmusk/status/1650703841468219394>

The *Twitter Files* refers to a collection of internal Twitter documents unveiled between December 2022 and March 2023. Comprising mainly of correspondence of prominent figures within the company, these documents surfaced in the aftermath of Elon Musk's famous acquisition of the platform. Specifically, *Twitter Files* shed light on nineteen instances of controversial actions undertaken by Twitter's former leadership. In November 2022, a mere month following Musk's official assumption of control over Twitter, the newly minted platform owner announced plans to unveil a collection of Twitter's confidential documents concerning the stifling of free speech and the distortion of open discourse. Musk articulated his stance by stating, "The collective mind that is this platform requires more signals and less noise." This implies that fostering freer and less distorted expression will enhance the quality of collective thinking on Twitter (see Fig. 4.2). He also commented that the public had a right to understand the events that took place under Twitter's preceding leadership (Propper, 2022).

The publication of these papers was orchestrated in tandem with Musk by the journalists Matt Taibbi and Bari Weiss. Taibbi described the information in the files as a "Frankenstein tale of a human-built mechanism (...) one of the world's largest and most influential social media platforms (...) grown out [of] the control of its designer" (Grynbaum, 2022). According to Taibbi, this collection of papers suggests that although Twitter acknowledged moderation requests from several political factions, the platform's politically biased staff influenced a noticeable selectivity and censorship in presenting information, favoring one side of the dispute. According to Weiss, decisions were made "in secret [to] actively limit the visibility of entire accounts or even trending topics" without notifying users. Weiss perceives this as covert censorship (Picchi, 2022). These practices altered the nature of the public discourse by algorithmically controlling the reach, dissemination, and presentation of people's speech in a manner that is challenging to detect.

The most surprising disclosures in the Twitter Files, primarily reported by Taibbi and Lee Fang, relate to the extent of the FBI's and Pentagon's efforts to influence particular users' visibility on the platform. According to Taibbi, there were over 150 emails exchanged between Twitter official Roth and the FBI from January 2020 to November 2022. While some of these inquiries seemed to be routine investigations,

many were requests for the company to restrict accounts identified by the FBI as spreading misinformation. However, Taibbi claims that these actions were not based on misinformation but rather on a disagreement with a specific political perspective. Taibbi has accused the agency of acting as if it were a *subsidiary* of Twitter, revealing systematic flagging of Twitter users by FBI agents for *potentially violative content* related to the 2020 election. The emails revealed that an FBI special agent was in regular contact with Twitter about tweets the agency wanted removed. The FBI flagged several accounts to Twitter's Trust and Safety Team for allegedly distributing election *misinformation*, which appeared to be more akin to endorsing candidates from a specific political group over another. Furthermore, Twitter was found to be actively collaborating with the Pentagon to boost U.S. government-approved accounts, often in Arabic or Russian, promoting agreeable or fabricated viewpoints. This phenomenon can be described as Orwellian "centralized content moderation" of political information (Wallace-Wells, 2023).

Another controversial topic concerns the practice of *shadowbanning*. For years, Twitter authorities dismissed allegations from various politicians who believed they were being *shadowbanned*, a term typically used for Twitter users unaware that their visibility on the platform has been suppressed. The *Twitter Files* unveiled the actual *blacklisting* of notable individuals whose opinions were not in alignment with the platform's political stance. This list includes conservative politicians and journalists like Dan Bongino and Charlie Kirk, as well as Dr. Jay Bhattacharya from Stanford University, a consistent critic of the collective mentality regarding COVID, who voiced his opposition to lockdowns. A well-known case of content interference was the infamous *Hunter Biden laptop controversy*. The New York Post released a story suggesting that emails on Hunter Biden's laptop revealed that he had connected a Ukrainian businessman with his father. Instantly upon its emergence, Twitter management unilaterally judged the story as untrustworthy and halted its spread. Without considering any possibility of contesting this decision, Twitter not only suspended the newspaper's account but also the accounts of individuals who circulated the story, including the White House press secretary, Kayleigh McEnany. The conclusion is that Twitter has granted itself the right to mute news that did not agree with its political optics, gaining more influence in this respect than prominent government representatives.

The question arises: Has revealing the controversies around debate distortion on social media had a positive influence on freedom of speech and the quality of online discourse? Or has it, on the contrary, further eroded trust in these platforms, consequently limiting debate? The exposure of Cambridge Analytica's practices was undoubtedly shocking to the public. However, whether this has led to changes in the practices of platform operators and users remains unclear. Facebook's pledge to restrict third-party access to behavioral data seems unlikely to have altered its practices of content targeting and user conditioning, although it limited this possibility to external companies. Meanwhile, Elon Musk's revelations about Twitter's censorship and biases prompt questions about whether his own management might continue similar practices, albeit with a different ideological slant. Nonetheless, it appears that X—the rebranded Twitter—now exhibits considerably more freedom of

expression. This is evident in the near-complete dismantling of the platform's content moderation department and the shift to algorithm-based speech control. Musk has consistently presented himself as a proponent of open debate, arguing that it will unleash the collective intelligence of the platform's users (refer to Fig. 4.2). Only time will tell if this approach is successful, as unfettered freedom of speech carries its risks, including high levels of informational noise and the spread of misinformation.

4.1.3.4 Information Noise on Social Media and the Spread of Fake News

Reducing information noise on contemporary social media and extracting valuable content from online debate is a real challenge for the online public sphere. The pervasiveness of information noise poses significant threats to our information ecosystems, subsequently undermining users' ability to discern fact from fiction. *Information noise*, broadly defined as the overwhelming mass of irrelevant or inaccurate data (Johnson, 2021), is an escalating concern due to its potential to exacerbate societal divisions, impede rational decision-making, and amplify misinformation spread (Le et al., 2020; Wang et al., 2021). Empirical evidence suggests that individuals exposed to high levels of information noise tend to demonstrate a diminished capacity for effectively identifying relevant and accurate information (Bessi et al., 2016; Vosoughi et al., 2018). Research in social psychology and communication has shown that when confronted with an overload of misleading information, humans have been shown to be both irrational and likely to have difficulty distinguishing between what is true and false. These studies have revealed that the human capacity for detecting deception when they are exposed to cognitive overload and information noise is only marginally better than random guessing, with typical accuracy rates ranging from 55 to 58% (Zhou & Zafarani, 2020). A deeper understanding of the mechanisms behind information noise creation and propagation, its impact on individuals and societies, and potential mitigation strategies is critical. In the future, it will be vital for policymakers to work in concert with platform designers, integrating scientific findings into practical applications to design a more transparent and less noise-prone information environment (Zhang et al., 2022).

The spread of open and intentional disinformation, or *fake news* (FN), is one of the most dangerous types of information noise. Fake news is usually defined as “a news article that is intentionally and verifiably false” (Allcott & Gentzkow, 2017) or “an article or message published and disseminated in the media containing false information regardless of the means and motives behind it” (Kshetri & Voas, 2017). FN is seen by many authors and specialists as one of the greatest threats to free speech and independent debate, undermining public trust in governments and causing social unrest (Allcott & Gentzkow, 2017, Pennycook & Rand, 2021). When looking for a method to counter the FN flood effectively, specialists most often point to fact-checking, in which dedicated teams check the credibility of information. However, there exist other approaches. Researchers have identified several distinctive features that can help distinguish fake news from genuine news content. These characteristics encompass various aspects, including writing style and quality, as well as quantitative

factors like word counts and the sentiments expressed. Additionally, the profiles and behaviors of users engaging with fake news, such as those who post, share, like, or comment on it, provide valuable insights that aid in the detection of such disinformation in numerous instances (Zhou & Zafarani, 2020).

Fake news often appears on topics that are critical to the public debate. As a result, they influence the formation of social and political opinions and can influence voters, opinion leaders, and even politicians. Using half-truths, edited videos, and selective reporting, the disinformers push their narrative, and even fact-checking does not always help people understand what is going on. Let us quickly review the topics that have been particularly susceptible to fake news in recent years:

- **Health and disease (coronavirus pandemic, vaccines, etc.).** During the early days of the pandemic, there were widespread claims that certain remedies like drinking bleach, inhaling hot air, or consuming specific herbs could cure or prevent COVID-19. Another example may concern vaccine misinformation: there have been hundreds of posts containing false information about COVID-19 vaccines, including claims that they alter DNA or contain microchips for tracking individuals (Carmichael & Goodman, 2020).
- **Street riots and social unrest.** Images and videos from unrelated incidents or past events are sometimes disseminated, pretending that they are related to current protests or riots, amplifying the sense of chaos and violence. For instance, a widely circulated image during the 2023 French protests against pension reform showed a battered and bloody figure of an elderly man surrounded by aggressive policemen. However, it turned out to be an AI-generated image (Bahl, 2023). Another type of disinformation related to street riots is the claim that pallets of bricks or paving stones were strategically placed at protest sites by the government to brutalize the crowd. Numerous examples of these types of false reports emerged during the 2020 Black Lives Matter protests in the United States. Most often, however, it turned out that the photos were old or the bricks came from a construction site (Lee, 2020).
- **Public figures in controversial situations.** Advanced technology is being used to create deepfake pictures or videos that appear to show public figures doing things they never did. This multimedia can be convincing and has the potential to deceive the public. For example, in the spring of 2023, a series of photos showing former US President Donald Trump detained by the police in dramatic circumstances circulated on social media. In one of the photos, Trump is running away from a police manhunt; in another, he is struggling with a group of officers. Even though the former president had dealings with the prosecutor's office (which gave the photos the appearance of credibility), the spectacular photos, despite their apparent reality, were deepfakes (Aleguas, 2023). (Fig. 4.3).
- **Conspiracy theories.** Conspiracy theories are thriving in the era of social media. Old theories accusing various groups of alleged power over the world are joined by new ones. These include accusations against 5G technology, which combine fears about the technological development per se, with beliefs that this technology is intended to control the minds of innocent people (Ahmed et al., 2020).

- **Military conflicts.** One of the most controversial and emotionally charged fields for spreading fake news and disinformation is events related to military conflicts. The war in Ukraine, which began as a result of Russian aggression in 2022, is marked by numerous mutual accusations from both warring parties regarding the dissemination of fake news. One example is the Bucha massacre, i.e., the mass murder of Ukrainian civilians and prisoners of war by the Russian Armed Forces during the occupation of the city of Bucha. Shortly after this war crime was revealed, Russians and their supporters began disseminating information falsely suggesting the massacre was staged. However, there was no confirmation of these claims (Marchant de Abreu, 2022).

Though spreading false rumors or false propaganda has been a long-standing human activity, modern social media platforms have transformed its pace, reach, breadth, and potential impact. Compared to traditional news outlets like newspapers and television, the production and dissemination of fake news online can occur more



Fig. 4.3 A deepfake depicting the police arrest of Donald Trump, which went viral on social media. Recognizing that this is an image generated by AI is made easier when we notice that the former president has three legs. *Source* <https://twitter.com/EliotHiggins/status/1637927681734987777/photo/1>

rapidly and at a lower cost. Politics in the era of so-called *post-truth*, i.e., an increasing disregard for factual evidence in political discourse (Lockie, 2017), makes it challenging for citizens to discern reliable from unreliable information, leaving them vulnerable to misinformation, disinformation, and propaganda. Fake news appears when individuals and organizations deliberately distort the truth without the platform owner's knowledge or where there is a blatant lack of concern regarding the accuracy of distributed information. Also, many fake articles rely on logical fallacies, such as appeals to emotion, ad hominem attacks, and strawman reasoning. These distortions can appear to be minor, manifesting as a mix of facts and truth, accompanied by an illustration or a headline implying a false interpretation and emotionally charged statements. Nevertheless, it is precisely such *minor* distortions, containing most of the true information with an admixture of falsehoods, that are usually the most difficult to detect and, therefore, the most dangerous. This kind of disinformation can indeed undermine the whole debate.

Furthermore, the primary business strategy of leading tech companies, which was described earlier, i.e., seizing and monetizing users' focus, might amplify these problems. The goal of social media platforms like Facebook or YouTube is to engage users and retain their attention. In this model, where individuals are often confined within filter bubbles and echo chambers, truth frequently takes a back seat, while the paramount focus is on feeding the user content that resonates with or even evokes their emotional states. Many content creators lack the resources to fact-check their information thoroughly. Some seek to attract more views by sharing selective or partial truths, while others, indifferent to the truth, disseminate false information to either promote a sense of unity within a particular group or to incite conflict. The ease of producing and disseminating low-quality information, along with the lack of traditional journalistic standards, contributes significantly to this problem (Cohen & Fung, 2021). Furthermore, the decentralized nature of online information sources, combined with powerful commercial and social incentives to attract attention through unverified news, further exacerbates the issue.

It's important to acknowledge that there have been proposals to utilize collective intelligence for detecting disinformation. The research by Shabani & Sokhn (2018) introduced a model that merges machine learning with crowdsourcing for this task. This approach combines human effort and machine learning to improve the decision-making process in determining the veracity of a message. Here, crowdsourcing is employed to categorize messages, differentiating satire from fake news and identifying content that is challenging for machines to detect. Moreover, in a recent brief article, I discussed the concept of a hybrid model that integrates collective intelligence with style detection and recognition of disinformation propagation patterns (Olszowski, 2021). Regrettably, none of these proposals has yet come to fruition.

4.2 Empirical Research on Online Debates

4.2.1 *How Different is a Policy Debate from Other Debates? Results of Our Experiment*

4.2.1.1 Research Methods and Experiment Preparation

Independent of our perspective on the state of the online public sphere, whether we view it predominantly in terms of its challenges or are optimistic about its capacity for unfettered civic debate, it is undeniable that this form of discourse possesses unique characteristics amenable to empirical study. When I began my research into online public debate in 2014, the opportunity to quantify the extent of various participant behaviors, evaluate their engagement levels, analyze their interactions, and measure the intensity of disputes and controversies arising from argument exchanges was an exciting yet relatively uncharted domain in social sciences. Additionally, I was interested in exploring how these behaviors could influence the quality, maturity, and acceptance of debate outcomes.

When I started my research internship at the MIT Center for Collective Intelligence at the end of 2019, my goal was to gain a better understanding of existing methods for studying these issues and to develop my methodology. I was already familiar with the social network analysis techniques used to study discourse on Twitter and acquainted with the outcomes of a highly intriguing yet regrettably incomplete research project called Catalyst (2015) conducted within the 7th EU Framework Program. Of particular interest to me was the notion of a Catalyst deliberation analytics server, with its principal architect being Mark Klein from MIT (Klein, 2015). Catalyst as a project was not finalized, but based on the published results, I was able to develop a preliminary version of my analytical server. During my internship at MIT, Mark assisted me in refining it, presenting his Deliberatorium system, which I wrote about in the previous chapter. The developed analytics tool allowed not only for tracking likes and connections between users, as is the case in social media, but also for organizing the debate through content classification, highlighting ideas and their associated arguments, and utilizing likes and dislikes to create rankings that result in the selection of winning ideas. Additionally, I could observe what topics participants browsed and estimate how much time they spent analyzing them, track how often they revised their publications, check what reputation they gained in the eyes of others, and use many other variables.

With the help of this tool, I conducted various types of online debates with recruited participants over the following months and monitored their metrics in real-time, which proved to be highly enlightening. In the following sections, I will describe the most intriguing results that emerged from an experiment involving a comparative analysis of a policy issues debate and a business debate conducted with a group of 108 recruited participants. Furthermore, some participants were assigned to a control group, which, as per the design, was meant to discuss neutral topics. The aim was to evaluate these debates, allowing us to capture differences in the behavior

of people and the results they produced. Given that this was during the coronavirus pandemic, health policy appeared to be a fitting subject for public debate. The issue that participants had to tackle was the potential strategy the government should employ in combating the pandemic. In contrast, the business debate revolved around an equally engaging topic: ideas for a business strategy for small companies in the era of the pandemic. Finally, the control group was to discuss a somewhat more entertaining topic related to ideas for a new historical or drama series for a popular streaming platform.

4.2.1.2 Debate Topics and the Issue of Polarization

Contrary to what is commonly seen on social media platforms, debates were conducted in moderately sized groups (30–45 participants) without the use of any algorithms that artificially create information bubbles and polarize participants. Differences in opinions emerged organically and were not artificially amplified, as seen on platforms like Facebook or Twitter, where the user’s newsfeed is manipulated to highlight similar views. The topic of the policy debate was carefully chosen so it did not align directly with the agendas of existing political parties or ideological beliefs. This approach enabled participants to express diverse opinions and mitigate polarization, likewise referred to in this kind of debate as “balkanization,” which involves a tribal division into emotionally opposing factions (Klein, 2015). In Table 4.1, presented below, we can see the top five ideas from each participating group, ranked by the *Community Interest* indicator, which takes into account views, votes, and arguments related to the topic, regardless of whether they were positive or negative. Meanwhile, the *Controversy* metric reproduces both the number of negative reactions and the number of positive reactions if they counteract the negative ones. As we can see, in the business debate, the topic that garnered the most interest (in terms of views, votes, and associated arguments) was not particularly controversial. However, in policy debates and, to some extent, TV series debates, the most controversial topics generated the most interest.

Evidence that there was no particular polarization or balkanization of policy debate participants can be found, for instance, in comparing responses to four of the most popular topics that emerged in this debate. As we can see in Table 4.1, these topics (except *Subsidies and health care reform*) also sparked the most controversy among debate participants. Therefore, I decided to investigate what similarities and differences in the participants’ opinions occurred in relation to particular topics. Did each of these popular and controversial topics reveal groups with mutually exclusive opinions, or maybe the debate was not polarized, and users maintained their independent opinions? The *Policy Debate* analysis of this issue is presented below.

When analyzing the responses of participants, I initially employed the Multidimensional Scaling (MDS) algorithm, a statistical method designed to identify and visualize the similarity of all opinions expressed by users on all discussion topics. In this method, the proximity of points in the visualization indicates a higher degree of similarity in opinions (Wickelmaier, 2003). The results can be seen below in

Table 4.1 The top ideas in each of the groups participating in the experiment due to their *Community Interest* indicator

Idea title	Controversy	Community interest	Views count	Views count/user
(1) Business Debate ('A business strategy for SMEs in times of pandemic')				
Obtaining sector and industry certification	0	3.67	115	3.48
Longtermness	2	3.58	108	3.27
Obtaining a grant for international cooperation in the field of vocational education	2	3.55	114	3.45
Building the owner's personal brand	0	2.94	88	2.67
Finding a market niche	3	2.76	81	2.45
(2) Policy Debate ('A strategy that the government should adopt in the fight against the pandemic')				
Government resignation. The question is, what will it change?	24	6.18	248	5.64
Mass vaccinations	12	5.84	244	5.55
Subsidies and healthcare reform	8	5.8	243	5.52
Mass testing of the country's inhabitants	17	5.32	217	4.93
Mandatory vaccination	17	5.3	216	4.91
(3) Control Group Debate ('Ideas for a historical or drama series for a popular streaming platform')				
The beginnings of Christianity in Poland and Slavic culture	4	4.97	137	4.42
The influence of the Hussite wars on the shape of Europe	3	4.48	136	4.39
The impact of following YouTubers on teenagers	12	4.45	126	4.06
Divorce and its further consequences on family life	10	3.9	111	3.58
Industrial revolution	0	3.81	112	3.61

Figs. 4.4, 4.5, 4.6 and 4.7. As we can see, there was no clear division into two opposing factions, and most users were situated in the center. Next, I visualized the opinions expressed on the four most popular topics. Blue represents negative opinions, and yellow represents positive opinions.

As we can see, even in the case of the two most *extreme* opponents (in the presented example, these were Michał and Tadek), there were instances of converging opinions, and within the heart of the debate, each discussed issue generated slightly different

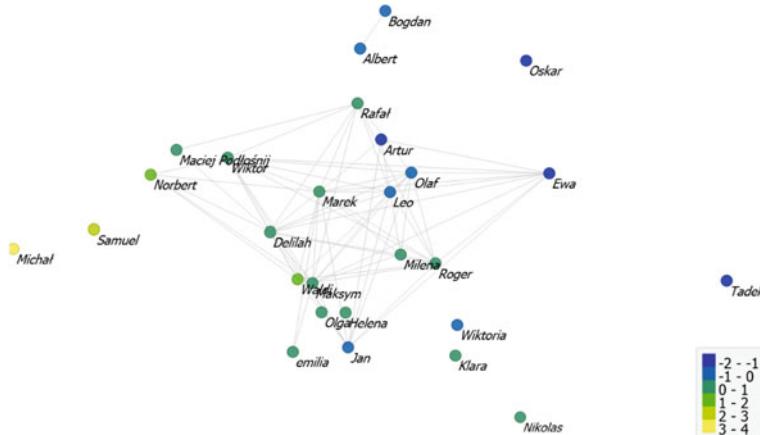


Fig. 4.4 Multidimensional scaling visualizing the convergence of all opinions expressed by Policy Debate participants (the closer the points, the more convergent opinions) and opinions on the idea of *Government Resignation* (blue are negative opinions, yellow—positive)

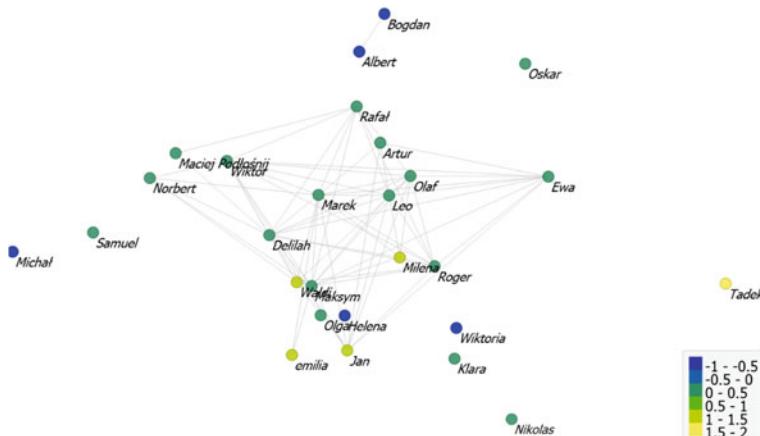


Fig. 4.5 Multidimensional scaling visualizing the convergence of all opinions expressed by Policy Debate participants (the closer the points, the more convergent opinions) and opinions on the idea of *Mass Vaccinations* (blue are negative opinions, yellow—positive)

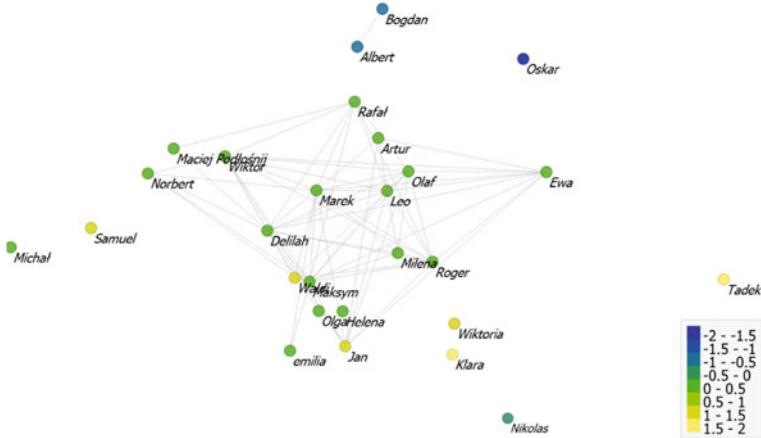


Fig. 4.6 Multidimensional scaling visualizing the convergence of all opinions expressed by Policy Debate participants (the closer the points, the more convergent opinions) and opinions on the idea of *Subsidies and Healthcare Reform* (blue are negative opinions, yellow—positive)

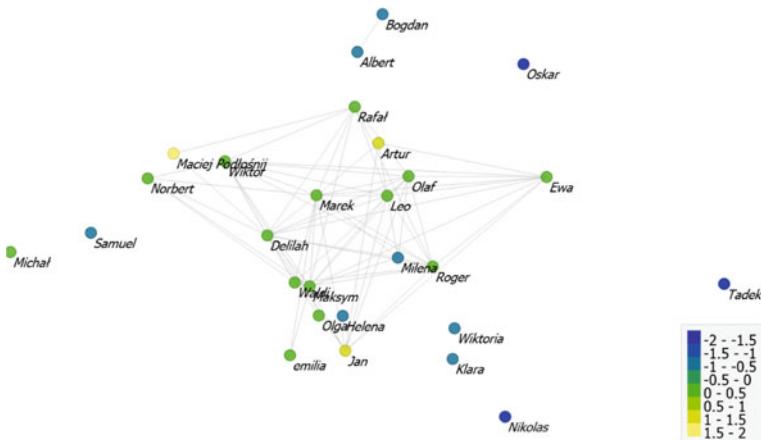


Fig. 4.7 Multidimensional scaling visualizing the convergence of all opinions expressed by Policy Debate participants (the closer the points, the more convergent opinions) and opinions on the idea of *Mandatory Vaccination* (blue are negative opinions, yellow—positive)

divisions. In other debates examined, we saw a similar phenomenon. Based on this, we can deduce that the participants generally retained their individual judgment, and the disagreements primarily pertained to ideas rather than personal issues or groupthink.

4.2.1.3 How Did the Most Engaged Participants Behave?

In analyzing participants' behavior, my initial focus was on measuring their level of involvement in the discussions. My objective was to determine if there were disparities in the dominance of a debate by a few influential participants and, consequently, whether the debates were balanced in terms of the whole community's participation. For this purpose, I devised two metrics. The first, which I named *Engagement*, represents the proportion of posts and arguments a user creates compared to the total in that debate. The second metric, called *Activity*, encompasses a more comprehensive measurement of a user's overall activity on the platform. This metric does not merely include posting but also includes activities such as voting, as well as the time devoted to reviewing and analyzing the contributions of others.

Comparative analysis of these metrics across various debates revealed compelling insights. As we can see in Table 4.2, the debate related to policy issues emerged as the most balanced among all examined discussions, with the top five participants contributing less than half of the total content. This was notably different from the business-related debate and the control group, where the leading five participants accounted for approximately two-thirds of the content. Regarding the *Activity* metric, an even more pronounced disparity was observed among the debates. Participants engaged in policy discussions demonstrated significantly higher levels of interaction with others, showing increased involvement in activities such as reading, analyzing, and voting.

The most significant difference arose regarding conflicts within the debates. The *Conflict* metric addresses the mutual reactions to published content for every possible pair that a particular user has formed with others. Assessing conflict, we consider the difference between positive and negative votes and arguments, both published by a user in relation to her/his counter-debater and vice versa. A *negative* conflict level, indicating a predominance of positive reactions, was a hallmark of the business debate. In contrast, the public affairs debate predominantly featured conflict. In the control group, this was more balanced. Consequently, mutual criticism and stringent evaluations were predominant when policy issues were addressed. This undoubtedly influenced the *Reputation* metric, which displayed the difference between all positive and negative reactions to the content published by a user. A negative level of this metric showed that negative reactions to the content published by a certain user outweighed positive ones. Conversely, *Efficiency*—which I defined as the ratio of a participant's total publications and votes to the time spent assimilating the debate's content—was a strong suit of the business debate. Simply put, users in this debate required less time to articulate their opinions.

Lastly, I delved into the *Density of Social Bonds*, which is the total number of other users with whom a particular user interacted in a debate, divided by the total number of participants in that debate. I also examined the *Intensity of Social Bonds*, which refers to the aggregate total of a user's interactions with others divided by the total number of users in the debate. It turned out that *Density* was fairly consistent across debates, but *Intensity* was typically higher in the policy debate.

Table 4.2 Rankings of the top five users in terms of their engagement in the debate in each of the groups participating in the experiment

Name	Reputation	Engagement [%]	Efficiency	Conflict	Density of social bonds	Intensity of social bonds	Activity
<i>1. Business debate. Top five users in terms of their engagement in the debate (percentage of the debate dominated by a given user)</i>							
Asteria	17	17.24	7.71	-21	0.38	0.84	2.35
Diana	2	13.79	7.5	-12	0.22	0.38	2.81
Ewa	0	13.79	4	0	0.06	0.13	1.04
Wiktoria	0	10.34	4.23	-11	0.28	0.47	3.27
Ala	8	6.9	1.54	-7	0.16	0.28	0.73
Total		62.06					
<i>2. Policy debate. Top five users in terms of their engagement in the debate (percentage of the debate dominated by a given user)</i>							
Tadek	-6	12.75	1.2	10	0.28	0.74	9.08
Klara	15	10.78	1.97	-9	0.4	1.05	7.16
Oskar	-2	9.8	5.38	12	0.23	0.42	4.59
Artur	-14	8.82	5.63	14	0.35	1.16	3.92
Michał	-6	7.84	8.33	11	0.3	0.44	2.6
Total		49.99					
<i>3. Control group debate. Top five users in terms of their engagement in the debate (percentage of the debate dominated by a given user)</i>							
Dorota	5	21.88	6.6	-5	0.37	1.1	3.03
Mariola	-4	15.63	3.27	3	0.3	0.57	3.15
Kamil	0	9.38	4.05	-2	0.1	0.13	1.56
Ewa	12	9.38	5.53	-13	0.37	0.7	1.91
Weronika	0	9.38	0.97	1	0.1	0.17	1.58
Total		65.65					

4.2.1.4 Overall Metrics and Collective Intelligence of the Debates

To compare the characteristics of the analyzed debates, my colleagues and I calculated the average of the most important metrics for each debate, encompassing both user-related and discussion topic-related metrics. Additionally, we adopted several supplementary metrics pertaining to the general properties of the debates. We compiled all this data into a single table, which facilitated the observation of common features and differences (see Table 4.3).

We successfully confirmed that the *Average Topic Controversy* (the level of controversy calculated separately for each topic within a debate, then averaged) and *Average Conflict* were significantly higher in political debates compared to business debates. As shown in Table 4.3, the predominant factors distinguishing policy discussions

Table 4.3 Comparison of the overall metrics of all examined debates

	Debate 1: Business	Debate 2: Public Issues	Debate 3: Control group
Posts	12	19	11
Votes	52↑ 6↓	128↑ 109↓	51↑ 35↓
Arguments	12↑ 5↓	38↑ 45↓	11↑ 10↓
Average efficiency	2.72	1.81	2.01
Average topic controversy	1.17	9.95	4.27
Average conflict	-2.85	2.45	-0.65
Average density of social bonds	0.07	0.1	0.08
Average intensity of social bonds	0.12	0.2	0.15
Average creativity	1.02	2.54	1.21
Average community interest	2.3	4.23	3.45
average activity	0.68	2.45	1.69
Average post views count	25.61	75.43	35.58
Debate maturity	8.08	13.75	11.43
Self-reflection	0.45	0.73	0.58

from business discussions in our experiment were the controversies surrounding the ideas being discussed and the conflicts between participants. This observation aligned with our initial analysis of these metrics within top discussion topics and between top users. The negative conflict level observed in the business debate and control group indicated that agreement among users outweighed disagreement. Notably, the difference in topic controversy between political and business debates was greater than the difference in the conflict that occurred, suggesting that the disputes were not purely personal but rather stemmed from differing opinions on the ideas presented in the debates.

I also observed that the *Group Density of Social Bonds*, i.e., the average number of people each participant interacted with, was similar across all debates. However, the *Group Intensity of Social Bonds* (the average number of interactions per user in a debate) was significantly higher in political discussions compared to the others. This meant that within the existing social network of the debate on public issues, a typical user engaged in more interactions, resulting in a livelier discussion with participants more frequently addressing each other's arguments.

Considering both the theoretical knowledge and the research findings described in Chap. 2, as well as the evaluation of case studies conducted by my research team detailed in Chap. 3, I embarked on identifying phenomena within these debates that could affect the level of collective intelligence revealed in them. Therefore, I was

interested in empirically verifiable behaviors of debate participants that contribute to the development of intelligent outcomes. Regarding the features associated with collective intelligence that we could monitor in our experiment, I focused on *Creativity*, *Activity*, and *Average Community Interest*. Additionally, I was intrigued by the metric of *Efficiency*, although it is not directly related to CI.

The *Creativity* index, calculated individually for each user, primarily depends on the number of ideas published by the user, as well as the quantity of supporting and opposing arguments they post, and partly on the number of edits they make to their own posts. The *Average Creativity* of the entire debate allows us to uncover how proficient the community is at generating novel ideas and solutions and creating innovative output through collaborative labor. As we recall from the analysis of cognitive processes, creativity influences problem-solving abilities. Meanwhile, the *Activity* metric for a user is a combination of published ideas and arguments, participation in voting within the debate, and time spent reading other users' ideas and responding to them. The *Average Activity* in a debate mainly relates to the measure of engaged citizenship, immersed in activity and public deliberation among equals. This underscores the importance of civic activity, as highlighted by researchers studying civil society engaged in public affairs debates. Analyzing the data shown in Table 4.3, we observe that both the *Average Creativity* and *Average Activity* were significantly higher in the debate on public issues than in the other two debates.

Regarding the *Average Community Interest*, this metric reflects the level of interest of participants in the whole debate, as determined by the total number of post views and the sum of arguments and votes associated with the posts. This sum is normalized to account for the number of participants in each debate. This metric can be related to the process of collective sensing, particularly in terms of gathering information and enhancing its quality. Comparing the level of *Average Community Interest* across different debates allows for the verification of whether any debate is dominated by *abandoned* threads that failed to capture anyone's attention or whether all threads are associated with arguments. It is checked whether posts were viewed, rated with up and down votes, and supplemented with arguments, thereby considering as many components of the debate as possible. The comparison of *Average Community Interest* values across debates highlights a significant difference, especially between public issues and business discussions. The debate on public issues garnered considerably more interest, whereas the interest of the community in the business discussion fell below that of the control group.

It is also worth noting the metric of *Efficiency*, which evaluates the ratio of content produced by a user to the time they spent reviewing threads. Here, the advantage of users in the business debate becomes apparent. However, we must remember that efficiency in producing results is not the most important aspect from the perspective of collective intelligence. This efficiency is achieved at the expense of limiting the time devoted to acquiring knowledge through familiarizing oneself with the ideas and arguments of other users. Therefore, *Efficiency* is not our primary concern.

As can be seen, we have managed to highlight many characteristics that defined the conducted debates. However, the most pivotal question I was seeking to address is: How have the observed differences contributed to the level of collective intelligence

achieved, and is it possible to estimate this level using the collected data? As we recall, evaluating the quality of collective work outcomes is challenging. However, I found a useful reference in the concept of *Argument Repertoire* as proposed by Capella et al. This concept centers on the quantity of arguments related to a particular issue, a criterion that these scholars have linked to the quality of reasoning within a debate. In the experiment described here, we applied this concept, associating it with the metric of *Maturity*. *Maturity* allows us to assess the outcomes of the debate, ascertaining whether ideas were adequately *saturated* with arguments or if, on the contrary, ideas devoid of arguments—hence not thoroughly processed by participants—were predominant.

Maturity, therefore, relates to the direct outcomes of a debate, with an assessment of these outcomes taking into account both the quantity and the structure of arguments. In my experiment, it was observed that in the policy debate, the *Maturity*, or the level of *saturation* with arguments, was more than twice as high as in the other two debates. An additional measure that can be related to the quality of the developed arguments is *Self-Reflection*. According to our assumptions, a higher frequency of revisions indicates greater care in the content published. This defined reflection and inclination to refine one's own ideas and arguments was also slightly more pronounced in the debate concerning health policy.

What conclusions can we draw from these observations? In the conducted experiment, the debate on policy issues demonstrated an advantage over both the business debate and the control group. However, this does not necessarily mean that the quality of policy debates is always higher in every case, as we will show in the next subsection concerning the debate on Twitter. In the described experiment, specific conditions were established which, though possible to achieve in a laboratory setting, are not always replicable in natural circumstances.

The first thing that becomes apparent is that conflict, disagreement, and mutual criticism are characteristics that often distinguish policy debates from other topics. On the other hand, business debates are more frequently marked by mutual agreement. In the control group, conflict and agreement balance each other out, with a slight predominance of agreement. In the debate concerning policy issues, the controversial nature of the topics discussed is also strongly correlated with *Community Interest*, which is different from what is observed in business debates and in the control group. Controversy attracts attention and arouses interest, and vice versa—popular topics arouse controversy.

A high level of conflict and controversy is not a problem unless it degenerates into *tribal warfare*, where opposing groups completely disagree with each other on all fronts and *exclude* each other from the debate. Conflict can be beneficial for increasing engagement in the debate, stimulating creativity and participant activity, and thereby expanding the *Argument Repertoire* and the *Maturity* of the debate. Therefore, there are situations where conflict leads to an increase in the quality of the debate and a rise in the level of collective intelligence. This is particularly the case when there is no extreme polarization of opinions, transforming debate participants into two separate, non-communicating *tribes*, such as partisan political factions.

Within the described experiment, we have achieved the state where, even in the case of the two most *extreme* opponents, there were instances of converging opinions, and each issue discussed generated slightly different divisions among participants. This allowed us to avoid the trap of extreme polarization and to maintain the positively stimulating nature of the conflict. The participants generally retained their individual judgment, and the disagreements were primarily about ideas rather than personal issues or groupthink. Furthermore, the debate related to policy issues emerged as the most balanced among all examined discussions. In the business and TV series debates, a few individuals clearly dominated, but the policy group was far more balanced. This demonstrates that engagement arising from substantive conflicts about controversial topics can foster greater activity across the entire group, not just among the leaders. In social media, where extreme emotions and mutual exclusion from the debate often dominate, this is much more difficult.

4.2.2 *Online Civic Debate in Polarizing Topics: The Vaccination Mandate*

Since 2018, alongside my studies on collective intelligence in laboratory conditions, I have engaged in empirical research conducted on social media, particularly on Twitter. I have been interested in evaluating discussions about public issues, such as government actions, international policy, and the spread of misinformation. Utilizing research techniques associated with social network analysis, which allows for investigating social structures through networks and graph theory and measuring vertex centrality like betweenness centrality and in-degree, our team conducted several studies. In analyzing the activity of media users in online debates, information flows, and emerging clusters of people with similar activities, I also paid attention to conditions that could foster the emergence of collective intelligence in social media.

One of our most engaging studies involved analyzing Twitter discussions concerning the Polish Ministry of Health's plan to introduce mandatory COVID-19 vaccinations (Olszowski et al., 2022). Since 2019, the ongoing coronavirus pandemic, coupled with the lack of significant progress in its containment, compelled governments of many countries to implement many extraordinary interventions: lockdowns, restrictions on movement and social gatherings, face masking, and physical distancing. One such intervention, introduced in some countries (e.g., Austria) and planned in others, was the obligation to receive vaccinations. This was due to noticeable opposition to vaccinations in certain social groups, stemming from a lack of trust in recently introduced vaccines and doubts about their effectiveness. In Poland, the Ministry of Health announced plans to introduce a vaccination mandate, which triggered a strong public reaction, including on Twitter.

The goal of our research was to examine the network shape that emerged on Polish Twitter around the trials of introducing mandatory vaccinations. We investigated what features characterized the key groups of advocates and opponents of

mandatory vaccinations and how their activity correlated with official epidemiological data concerning the COVID-19 pandemic. Over 70,000 tweets, retweets, and replies published between July and December 2021 were collected and analyzed. In addition, we gathered official statistics from the Polish Ministry of Health on daily vaccination numbers, new cases of infection, deaths caused by COVID-19, number of people in quarantine, number of tests performed, etc. A clustering analysis using the Clauset–Newman–Moore algorithm (Clauset et al., 2004) identified two significant user groups: advocates for and opponents of mandatory vaccination. We also examined temporal trends of tweets, the most commonly used hashtags, the sentiment expressed, and correlations with epidemiological data. The results reveal a substantial degree of polarization, a high intensity of discussion, and a high degree of involvement of Twitter users.

Figure 4.8 presents the examined Twitter users in social network graph clusters. Each node represents a user, and a line between them represents an edge. The size of the nodes is ranked by their *betweenness centrality score* (White & Borgatti, 1994), which measures the influence of a vertex over the flow of information between all other vertices under the assumption that information flows over the shortest paths among them. This graph, in particular, highlights the two most important clusters. The group visualized in dark green (G1) is the largest cluster of the network, consisting of 6520 users, which is 29.94% of the entire sample. The group visualized in dark pink (G2) is the second-largest group, consisting of 5930 users, which is 28.82% of the network. These two most important user groups were of similar size and presented highly polarized opinions. In general, the G1 group mainly consisted of vaccination mandate supporters, while the G2 group mainly consisted of its opponents. The ten most important nodes in terms of *Betweenness Centrality* were numbered, starting with the most central one, which turned out to be the then Minister of Health.

The nature of these groups becomes apparent when we compare their most popular hashtags. Table 4.4 below presents the most frequently used hashtags in group G1. As we can see, the “COVID-19” hashtag (n = 1039), as well as “szczepimysie” (eng. “we vaccinate,” n = 520), were the most popular hashtags. The other popular slogans in this group were general references to the Covid-19 pandemic: “koronawirus,” “corona,” and “covid_19.” The hashtag “dworczyk” relates to Michał Dworczyk, government spokesman, responsible, *inter alia*, for communicating pandemic policies. “Pis” states for the political party *Prawo i Sprawiedliwość* (eng. Law and Justice), which formed the government in Poland at that time and was responsible for pandemic restrictions. The hashtag, translated as “we vaccinate,” is the main slogan for COVID-19 vaccination proponents. An interesting fact is the appearance of the German term *impfpflicht*, which also means mandatory vaccination. The popularity of this hashtag is related to the frequent citation of tweets calling for the introduction of compulsory vaccinations in Germany and Austria. The conducted hashtags review and tweets content review allowed us to define the nature of the G1 group as supporters of the introduction of the vaccination mandate.

As we can see in Table 4.5, in group G2, the most popular hashtag is “stopsegregacjisanitarnej” (eng. “stop sanitary segregation,” n = 2090), which was a slogan used to protest COVID-19 restrictions and to promote the bill of the same

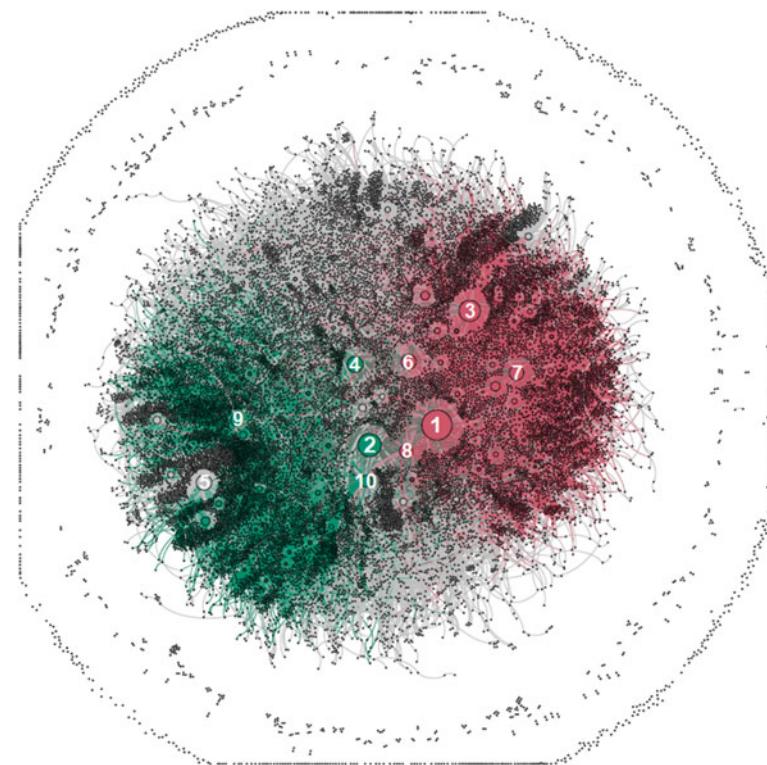


Fig. 4.8 Social network graph of Twitter users debating mandatory COVID-19 vaccination in Poland between 26 July 2021 and 9 December 2021. The two largest groups of users distinguished by the clustering algorithm are G1, marked with dark green (vaccination mandate supporters), and G2, marked with dark pink (vaccination mandate opponents). The ten most influential users, according to the betweenness centrality score, are numbered from 1 to 10 (Olszowski et al., 2022)

Table 4.4 Top 10 hashtags in the group G1 (Olszowski et al., 2022)

Rank	Top hashtags	Number of occurrences
1	covid19	1039
2	szczepimysie (eng. “we vaccinate”)	520
3	Koronawirus	379
4	Dworczyk	283
5	Pis	228
6	szczepimysię (eng. “we vaccinate”)	220
7	covid_19	130
8	Polska (eng. Poland)	118
9	Corona	105
10	impfpflicht (eng. “mandatory vaccination”)	105

Table 4.5 Top 10 hashtags in the group G2 (Olszowski et al., 2022)

Rank	Top hashtags	Number of occurrences
1	stopsegregacjisanitarnej (eng. “stop sanitary segregation”)	2090
2	covid19	534
3	koronawirus (eng “coronavirus”)	340
4	Lextvn	260
5	Konfederacja	220
6	szczepienie (eng. “vaccination”)	160
7	konstytucja (eng. “constitution”)	160
8	USA	129
9	niedzielskidodymisji (eng. “Niedzielski to resign”)	126
10	gotowaniezaby (eng. “boiling frog”)	101

name, submitted for legislation by one of the anti-lockdown political parties. The premise of this bill was “to ban any forms of discrimination against the unvaccinated” (Projekt ustawy “STOP segregacji sanitarnej”, 2021). The general hashtags “covid19,” “koronawirus,” and “szczepienie” (eng. “vaccination”) were also popular, but the true nature of this group is revealed by the names: “konfederacja,” which is the political party strongly opposing vaccination mandate; “konstytucja,” referring to the opinion that the obligation to vaccinate would violate the Polish constitution; “niedzielskidodymisji” calling on the Minister of Health, Adam Niedzielski to resign; and “gotowaniezaby” (eng. “boiling frog”), a well-known aphorism describing a frog being slowly boiled alive, which in this context means the fear of gradual limitation of personal freedom, e.g., by vaccination mandate. To sum up, the analysis of hashtags and tweet contents enabled us to categorize the G2 group as opponents of the vaccination mandate.

Studying the activity of users from the G1 and G2 groups, we primarily analyzed the number of tweets, retweets, and replies published on each day covered in the study by people belonging to each group. The results are shown in Fig. 4.9. The types of user activity that were analyzed can be defined as:

- *Tweets*, i.e., posting new content that includes the personal opinions of the author or links to news articles together with personal comments;
- *Replies*, i.e., direct responses to published tweets, interactions between Twitter users, an activity requiring engagement, exchange of opinions, dialogue, and comments;
- *Retweets*, i.e., forwarding a tweet or reply to reinforce its impact. This is a less engaging form of activity but crucial to understanding the reach of popular tweets.

As can be easily observed, there was a significant increase in activity toward the end of the study period. This correlates with the government’s announcement on December 6, 2021, that the vaccination mandate would be introduced in March 2022 for several professional groups, which led to a great surge in interest in the topic.

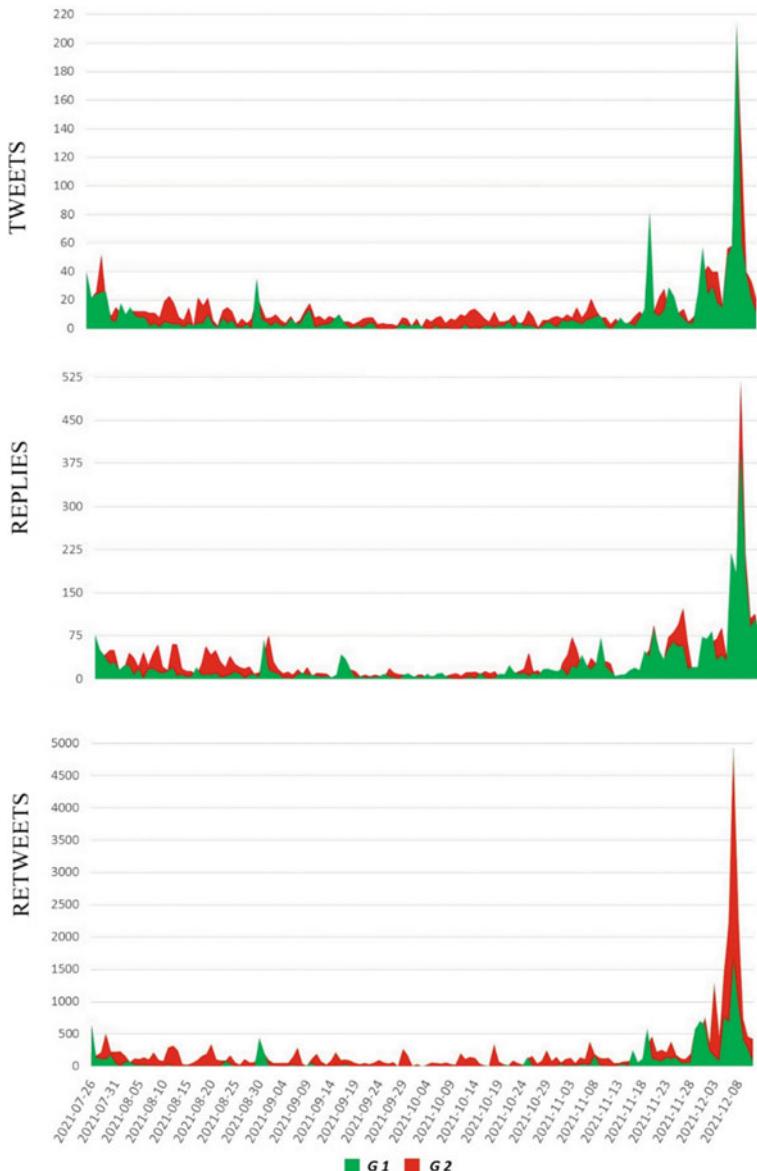


Fig. 4.9 The number of tweets, replies, and retweets published daily by the members of the G1 and G2 groups (Olszowski et al., 2022)

At the same time, we see a very strong polarization. Both supporters and opponents of the vaccination mandate became highly active, though in slightly different ways. Supporters predominated in publishing original tweets, while opponents gained a significant edge in retweeting. Nevertheless, both groups were very active within their own circles. However, how did this increase in activity affect the debate itself?

When conducting research on Twitter, it is challenging to measure the quality of the debate in the same way as we did in laboratory experiments or case studies. We lack information about the quantity and diversity of arguments that appeared in the discussions. However, we can use overall metrics specific to social network studies, such as *Graph Density*, *Modularity*, or *Reciprocity*:

- *Graph Density* is a measure of the number of edges among a group of vertices, compared to the total possible number if every vertex was connected to every other vertex. High graph density means that most individuals are connected to many others, while low graph density suggests that most individuals are not interconnected, implying that the debate occurs within alienated groups.
- *Modularity* is a measure of the fitness of the groups that are created in a clustered network. It quantifies the number of edges that extend from one group to connect with vertices in another group. Low modularity suggests that the clusters or groups formed may be poorly defined or not cohesive. Conversely, high modularity indicates that the groups are well-defined and open to mutual exchange of views.
- *Reciprocity* pertains to the bilateral nature of communication and mutual listening among users. The *Reciprocated Vertex Pair Ratio* is the percentage of vertex pairs that have a reciprocal relationship. This occurs when an edge from Vertex A to Vertex B is complemented by another edge from Vertex B to Vertex A, thus forming a *reciprocated* connection. Additionally, the *Reciprocated Edge Ratio* represents the percentage of edges in the network that have such reciprocal relationships.

How did these overall network metrics change from July to December 2021 in the group we studied? Let's look at the data fluctuation charts shown in Fig. 4.10.

As can be seen, all these values decrease toward the end of the period of interest. This leads to the conclusion that, contrary to the findings in laboratory studies, the intensity of the debate does not contribute to an improvement in its quality. Participants become increasingly distant from each other, more entrenched in their clusters, and less inclined to listen to one another.

Moreover, there is no evidence that the emergence of new medical data about the COVID-19 pandemic and vaccines had any positive impact on the level of debate or mutual listening to arguments. On the contrary, it appears that emerging medical information has no connection with the opinions being presented. How do we know this? One of the final stages of our analysis involved measuring the correlations between epidemiological data and overall network metrics. For this analysis, a set of epidemiological data concerning new COVID-19 cases, casualties, infection tests, etc., was obtained from the Ministry of Health. The results are presented in Fig. 4.11.

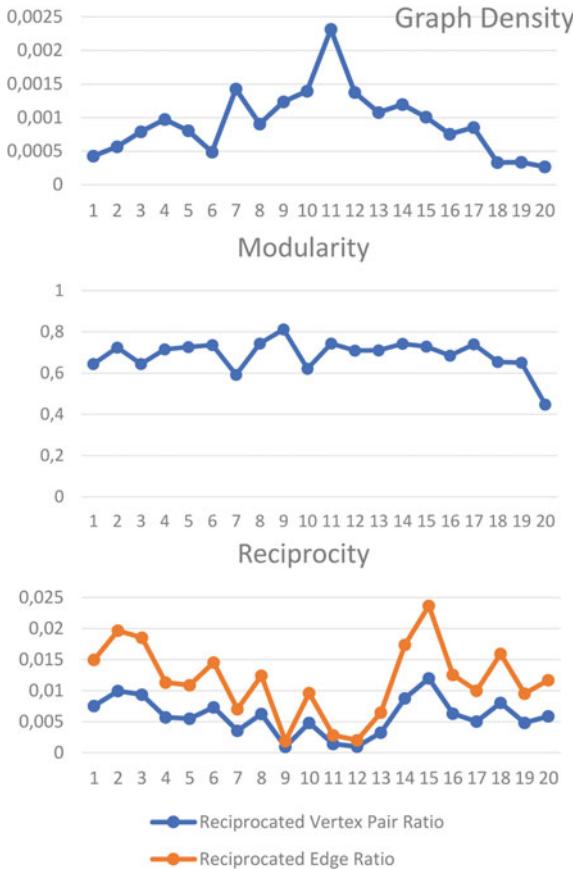


Fig. 4.10 Fluctuation of overall network metrics of the Polish Twitter discussion on vaccination mandate from July to December 2021

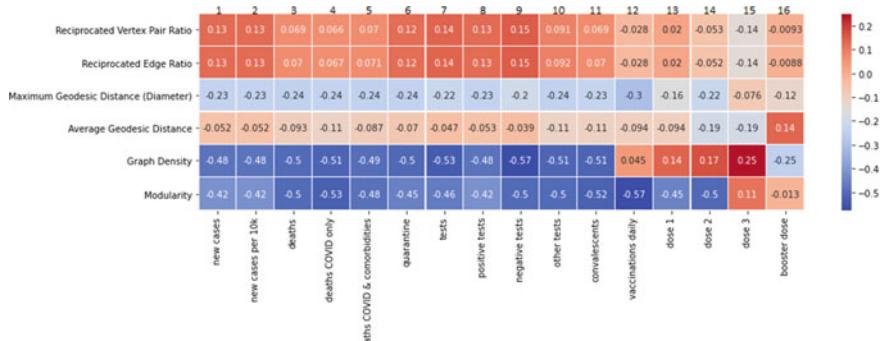


Fig. 4.11 The measures of Pearson's linear correlation between the Covid-19 epidemiological data and the overall network metrics (Olszowski et al., 2022)

The observed negative correlation of *Graph Density* and *Modularity* with most of the epidemiological data indicates that despite an increasing number of facts about the pandemic, there is a fragmentation of the debate and a decline in communication between polarized groups. People are becoming more distant from each other of dissenting views, closing themselves in their own environments. The negligible and essentially insignificant correlation of medical data with *Reciprocity* metrics also does not provide grounds for optimism. As observed in the analyzed debate, the availability of epidemiological data—seemingly non-controversial, non-partisan, and non-emotional facts about the pandemic—does not lead to an increased exchange of opinions and the building of a common knowledge base or reciprocal exchange of arguments. Instead, it deepens differences in interpretation. This observation corroborates earlier studies indicating that in a situation of highly polarized public debate, “science” or “facts” alone do not “resolve” political controversies (nor do they change minds in a linear manner) (Sarewitz, 2000). The lack of correlation with metrics such as *Average Geodesic Distance* or *Reciprocated Vertex Pair Ratio* suggests that the disclosure of epidemiological data did not have any significant impact on the emerging collective intelligence in the debate.

The debate on mandatory vaccinations was characterized by a strong emotional engagement of its participants, which led to increased polarization and the spread of informational noise. The ability to extract and organize arguments from a discussion was limited, making it extremely difficult to raise the level of collective knowledge systematically. There was no opportunity to bring arguments to the forefront, and the most extreme emotional posts gained the largest reach, which we noted in analyzing the most frequently retweeted content. Substantive arguments, when they did appear, were in a clear minority and became invisible, drowned out by extremely polarizing, emotionally charged statements. There was a lack of multifaceted debate, which would have provided an opportunity to refine alternative proposals by considering arguments for and against the discussed issue.

Due to the extreme polarization, a common attitude among participants was antagonism, consisting of discrediting adversaries and denying them the right to express any opinion. Adopting a balanced position, taking into account the reasons of both sides, would probably be equated by the antagonized factions with *betrayal* rather than a serious reflection on public affairs. The conflict turned into a *tribal war*, which pushed collective thinking processes into the background. The lack of a collective platform for cooperation was evident not only in relation to adversaries but also within one’s *own tribe*, as the primary activity was to attack the opposing side and publish *manifestos* full of slogans rather than improving argumentation.

4.2.3 Predicting Stock Prices by Collective Intelligence on Twitter: McDonald's Stock Performance During Russia's Invasion of Ukraine

The debate described in the previous section, conducted on a topic that is emotionally charged and polarizing, is certainly not an example of group interaction where collective intelligence is manifested. In this section, another case of communication on Twitter will be discussed. This case is also emotionally charged but relies less on mutual combat between polarized groups and more on collective mobilization to achieve a common goal. It is an instance of CI emerging from the large numbers of participants: crowd mobilization and processes occurring on a macro scale. Let's delve deeper into how collective intelligence, manifesting in vast, mob-like formations where participants' contributions are aggregated into a common result, can be remarkably powerful under specific conditions. This success is commonly propelled by factors such as significant social mobilization and the absence of severe polarization, which, if present, tends to splinter the collective.

In 2022, the world's public attention shifted toward Eastern Europe. On February 24, Russian troops, amassed along the borders of Ukraine, unexpectedly launched a wide-ranging military offensive. Their aim was to achieve a swift victory over the Ukrainian army and capture the country's capital, Kyiv. Russian President Vladimir Putin announced a *special military operation* under the pretext of supporting the breakaway republics of Donetsk and Luhansk, whose paramilitary forces had been involved in the Donbas conflict with the Ukrainian government since 2014. Putin held irredentist views that challenged Ukraine's right to exist, falsely claiming that Ukraine was governed by neo-Nazis persecuting the ethnic Russian minority.

The initial days of the conflict did not bring spectacular successes for the Russians; instead, they significantly united Ukrainians in resistance against the invaders. Similarly, public opinion in most countries united in protest against the invasion. This strong reaction was reflected in debates on social media. From the beginning, platforms like Twitter, Facebook, and Telegram were flooded with a large volume of comments, frontline reports, and appeals to stop the conflict. In the early days of the war, there were also calls for economic sanctions against Russia aimed at forcing an end to the invasion. These calls were directed at governments, international organizations, and global corporations operating in the Russian market, such as McDonald's and Coca-Cola, urging them to exit and cease doing business with the aggressor quickly.

The significant role that social media played in this war has been the subject of many analyses. The warring parties could instantly publish high-quality reports from the front lines, which were swiftly retweeted or shared. Since anyone with a smartphone could be a *war correspondent*, activities on social media have increased both the possibilities for transmitting information from the battlefield and for spreading false content and manipulating information. It can be said that social media has become another front where the war unfolds.

Our research team, specializing in social network analyses, undertook a project during this period to study the behavior of English-speaking Twitter users. These users had come together in a campaign to condemn the aforementioned corporations, McDonald's and Coca-Cola. They called for these companies to withdraw from Russia and leveraged the threat of consumer boycotts to pressure them (Ahmed et al., 2023). These brands were selected as two of the largest brands trending on Twitter due to users calling for a boycott during the start of the war in Ukraine. These brands were also chosen because both seek to cultivate a global presence on social media to foster consumers' sense of intimacy, trust, and closeness.

We were curious about the proportion of tweets concerning global corporations that contained negative sentiment regarding their presence in Russia. We were also interested in the relationship between these critical tweets and the stock prices of these companies on the New York Stock Exchange, one of the world's largest and most renowned capital markets. We assumed that, although previous research (Bartov et al., 2018; Bollen et al., 2011; Mittal & Goel, 2011; Ranco et al., 2015) has examined Twitter data as a tool for stock market prediction and forecasting in more general terms, there was a lack of research focusing specifically on the relations between citizen-led activism and a company's share price. Our hypothesis was that Twitter users react to global events even more rapidly than the stock indexes and share prices, known for their swift responses. If this could be proven, it would demonstrate that the collective intelligence of Twitter users, operating on a macro scale, can anticipate economic events, a capability that could be highly valuable in predicting global-scale phenomena.

Our research was based on the retrieval and analysis of 725,688 tweets containing keywords linked to McDonald's and Coca-Cola, published between March 1, 2022, and March 16, 2022. Stock market data was collected from Yahoo Finance. In our statistical analysis, we used Twitter-related variables and stock market variables, like open price, close price, and volume. The top 10 most popular tweets on each day (labeled here *topX*) have been analyzed and classified manually by assigning them to either *against*, *in favor*, or *neutral* categories. Tweets categorized as *against* contained negative sentiments about the company or its products, as well as critiques of the company's management decisions or blaming the company for supporting Russian aggression in Ukraine. Tweets that were *in favor* contained a positive sentiment about the company, its products, its sponsoring activities, positive sentiment about working in this company or its products, and tweets supporting companies' management decisions. Tweets categorized as *neutral* did not belong to any of the above categories or contained mixed sentiments.

The first observation was that an increased reaction to political events was primarily evident in tweets concerning McDonald's. As seen in Fig. 4.12, heightened quantities of tweets and their reach, measured by the number of retweets, were noticeable around March 5th and March 8th. The initial peak was caused by escalated criticism, whereas the latter rise was driven by a combination of ongoing critique and numerous positive tweets applauding the companies for their then-recent withdrawal from Russia.

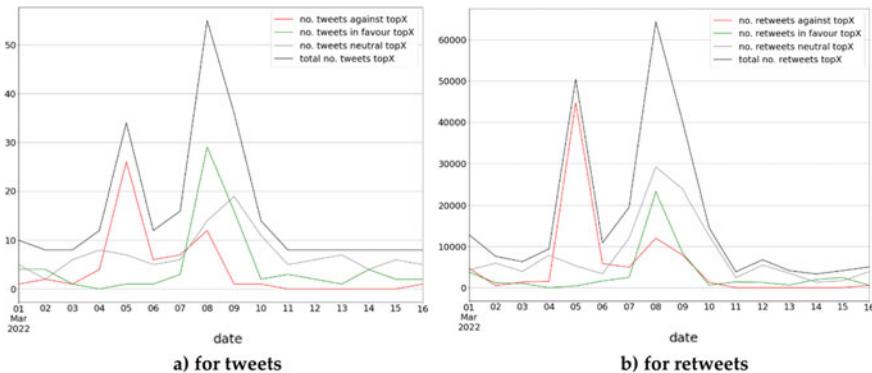


Fig. 4.12 Comparison of the number of tweets and retweets for McDonald's between March 1 and March 16, 2022, distinguishing their sentiments (Ahmed et al., 2023)

However, how did these peaks of interest relate to the stock prices and trading volume? To demonstrate a correlation between them while also accounting for the possibility that events on Twitter may have occurred earlier, we introduced a 1-day lag to the Twitter data for the closing price and index and a 2-day lag for the opening price. The lag length refers to the number of time periods by which we shift one variable backward or forward to measure its relationship with another. This allowed us to show that Twitter activity preceded stock market movements by 1 or 2 days and to determine the degree of correlation between these events. The Spearman method was used to analyze the correlation between Twitter and stock market data.

Figure 4.13 confirms a very strong negative Spearman correlation between the opening price and the total number of tweets and retweets. Furthermore, all methods employed in the analysis for the *topX* posts (i.e., the most popular) confirm a very strong negative correlation between them and the opening price. It's noticeable that the number of retweets criticizing McDonald's and expressing dissatisfaction with its actions in *topX* ("no. retweets against") has a much higher negative correlation with the opening price than the corresponding variables for favorable or neutral tweets ("no. retweets in favor" and "no. retweets neutral"). This underscores a robust relationship between the opening price and Twitter data, indicating that what happened on Twitter in the analyzed period preceded market reactions. For the closing price, we can draw similar conclusions to those for the opening price. In summary, analyzing variables most relevant for assessing the impact of opinions *in favor/against/neutral*, namely the number of retweets categorized by these sentiments, shows that opinions against are negatively correlated with stock prices to a greater extent. This relationship is more pronounced when analyzing opening prices than closing prices. The situation is different in the case of trading volume. Two of the most reliable Twitter variables, namely the total number of tweets and the total number of retweets in *topX*, both confirm a strong positive correlation with volume. This means that increased activity on Twitter preceded a heightened market movement related to the selling of shares.

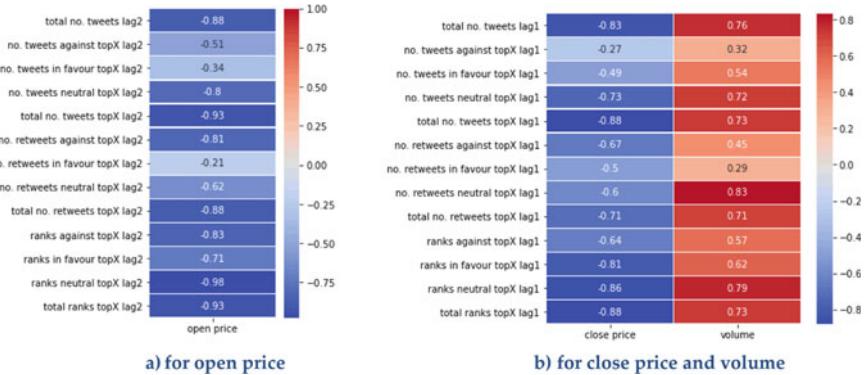


Fig. 4.13 Spearman correlation between Twitter and stock market data for McDonald's (Ahmed et al., 2023)

These phenomena can also be analyzed by tracking time series charts. Figure 4.14 shows plots of McDonald's time-lagged version of the total number of retweets in *topX* in relation to the opening price, closing price, and volume of the company's share prices. Although the closing price reacted much quicker to the unfolding, the relationship between the number of tweets and the open price lasted much longer. The top plot shows that the relationship between the opening price and the total number of retweets was inverse: when one variable decreased, the other increased. This pattern was observed from March 7 to March 15, covering most of the study period. This confirms a very strong link between the volume of tweets for McDonald's and the company's opening stock price. The market is delayed in its response to Twitter activities by approximately two business days. A similar phenomenon occurs in the case of the closing price with a 1-day lag, shown in the middle plot. Regarding the volume variable, presented in the lowest plot, there is a positive correlation with the total number of tweets: when one variable decreases, the other also decreases. This type of relationship was observed from March 4 to March 10. It appears that an increase in the number of tweets led to a rise in volume, but this only persisted until McDonald's suspended its operations in Russia.

Our observations have demonstrated the remarkable ability of large social media crowds to mobilize and exhibit collective intelligence in predicting future phenomena. Intelligent outcomes can emerge from incoherent collectives, confirming Surowiecki's vision of the wisdom of crowds. Of course, it is important to consider the conditions that foster such CI, as it does not manifest in every situation. In the case in question, the emotional load associated with the problem faced by Twitter users was very high, as they were dealing with military aggression against an innocent country, which obviously had a decisive impact on their civic activity and sense of community while defending Ukrainians. On the other hand, the level of complexity of the problem was not high: for most Twitter users, it was clear who was the positive and who was the negative side in this conflict. The controversies that arose were mainly related to the lack of response by the global enterprises to the invasion, so

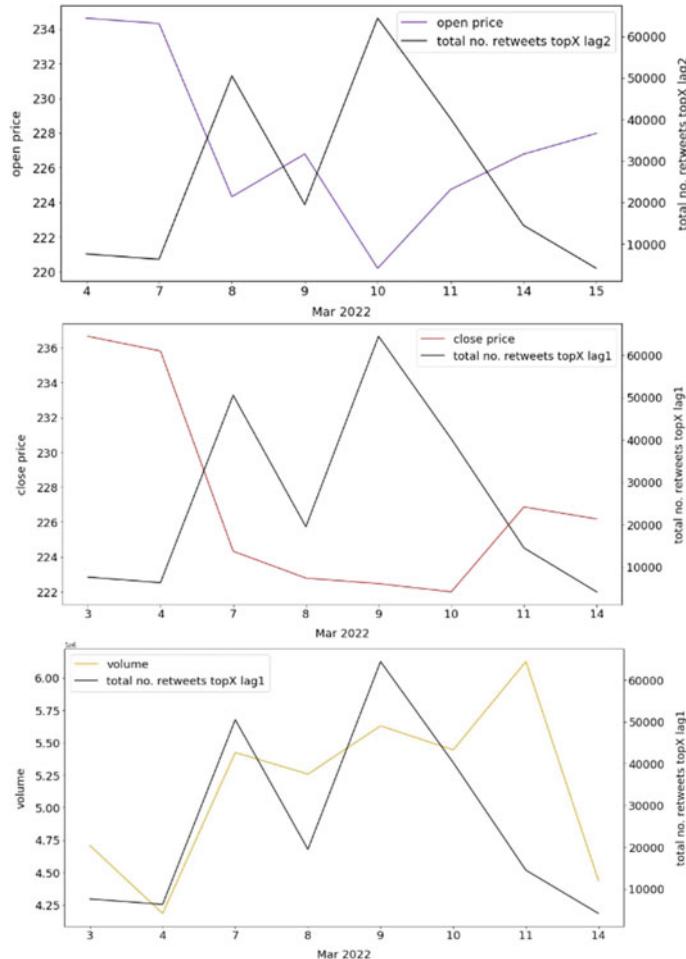


Fig. 4.14 Plots of McDonald's number of retweets with relation to open price, close price, and volume of share prices of this company (Ahmed et al., 2023)

they did not significantly divide the participants in the discussion and did not cause a clear polarization among them. Finally, the timeframe in which these events took place was not overly long, which shielded Twitter users from disinformation and fake news that would undoubtedly have emerged in response to their actions.

4.3 Models of Intelligent Public Debate in an Online Environment

What do the previously presented empirical data and case studies tell us about how to raise the level of intelligence in online debates concerning public issues? We observe that processes grounded in close collaboration, or as Daniel Andler articulates, manifesting *thick* cognitive processes, predominantly take place in small to medium-sized settings characterized by prevailing interpersonal relationships. In such communities, there are the greatest chances for the formation of bonds within the group, mutual inspiration, and a multifaceted debate, allowing for a deeper understanding of the public issue. An obstacle may be the fragmentation and inconsistency of the group, but the previously described experiment proves that even in discussions full of controversies, mutual conflicts, and polemics, the debate does not have to suffer if the polarization is not extreme.

In the case of large platforms, obstacles to debate include extreme emotions, polarization, and information noise, resulting partly from disinformation and behavioral targeting. However, there's also another aspect to consider: CI processes evident in social media platforms like Twitter facilitate the gathering and aggregation of a large volume of opinions. This is particularly useful in situations where extreme polarization is not prevalent and antagonism does not dominate. Certainly, this process is mostly linked to the purely aggregative side of CI. On the other hand, even in the typical social media platform, an actual debate is possible. Subgroups with strong identities and developed interpersonal relationships that are able to manifest *thick* cognitive processes can also emerge.

Enhancing the intelligence of online debates on public issues, though challenging, is certainly achievable. The key factors influencing this enhancement are tied to the degree of participant engagement and independence, the emergence of civic attitudes, the development of a collective identity, and the techniques employed in orchestrating collaborative efforts. Introducing tools to structure the debate effectively warrants serious consideration. These tools could focus on minimizing information noise, extracting substantive arguments, and implementing algorithms that encourage constructive behaviors. Additionally, as emphasized in our case study analysis, harnessing the power of collective memory is of paramount importance. Equally critical is the approach to conflict resolution within these debates, raising the question of whether the deliberative model is the *ideal type* for productive discussions or if viable alternatives exist. These considerations will be explored in the following sections.

4.3.1 Engaged and Independent Online Citizens

The first essential element of an intelligent debate on public issues and government policies is the participants themselves, that is, engaged and responsible citizens.

Active citizenship is closely intertwined with open public discourse, and individuals engaged in the debate feel a much stronger connection to collective decisions than those who remain passive. Civic activity is one of the key components of a post-positivist and open approach to policymaking. Shifting away from a technocratic approach to policymaking, which relies on the analysis of a limited circle of experts, embedding the policy discussion within a wider public context requires the involvement of a substantial number of engaged citizens.

As demonstrated in the case studies analyses, civic engagement is fostered by a sense of citizen duty and a collective pursuit of the common good. This is evident when we examine urban projects discussed in the previous chapter, such as Decide Madrid or Better Reykjavik. The most successful civic initiatives, which gather communities aiming to solve a particular problem, go beyond narrow interests and appeal to values important to the entire community.

However, for individuals within collectives to make valuable contributions to the debate, it is essential that they maintain their independence and resist being dominated by the group. As the experiment described in this chapter showed, independent reasoning allows debate participants to resist groupthink and transcend the *tribal* divisions that extreme polarization in social media often leads to. Members of a community, therefore, should possess the intellectual capabilities, financial means, and time to ensure their independence. The analyzed case studies also demonstrate that participants' civic engagement is contingent upon the degree to which the topic of the debate resonates with their personal experiences and the direct impact it has on their lives. The feeling of commitment to the community is connected to the intertwining of various aspects of one's personal life with that community. The urban layout of my city impacts my health and well-being, equitable laws shape the growth of my business, and a shared foundation of ethical values fosters mutual respect.

Therefore, there is no inherent contradiction between the individualism of citizens and their involvement in the community or between personal interests and the common good. In fact, strong and independent individuals within a group are the cornerstone of an intelligent community. For instance, Friedrich August von Hayek's theory of individualism primarily posits that decision-making fundamentally resides with the individual, not within abstract theoretical constructs. Yet, Hayek also highlighted that individuals are not entirely independent in their decisions, and individualism does not equate to isolation or serve as a justification for selfishness. The Austrian economist was aware of the dangers of societal atomization and knew that individuals detached from strong bonds with communities would actually be less independent. They would be more susceptible to manipulation and unaware instincts. Consequently, the ideal model for collective living hinges on the dynamic interplay between strong, independent individuals and their communities. This necessitates an effort to comprehend the forces that shape human social life. Hayek's call to "attempt to understand the forces that determine the social life of man" while preserving man's independence to "make full use of his knowledge and skill" does not extol isolated or self-contained individuals but rather citizens whose very nature is shaped by their existence in society (Hayek, 1946).

Jaron Lanier, a pioneer in virtual reality and one of the legendary computer scientists from Silicon Valley, believes that finding a proper balance between individualism and social cooperation is the best way to achieve collective intelligence. Lanier asserts:

Every authentic example of collective intelligence that I am aware of also shows how that collective was guided or inspired by well-meaning individuals. These people focused on the collective and, in some cases, also corrected for some of the common hive mind failure modes. The balancing of influence between people and collectives is the heart of the design of democracies, scientific communities, and many other long-standing projects. There's a lot of experience out there to work with. A few of these old ideas provide interesting new ways to approach the question of how to best use the hive mind.

Lanier mentions the free market and scientific communities as examples of cooperation between independent individuals and communities:

What makes a market work, for instance, is the marriage of collective and individual intelligence. A marketplace can't exist only on the basis of having prices determined by competition. It also needs entrepreneurs to come up with the products that are competing in the first place. In other words, clever individuals, the heroes of the marketplace, ask the questions that are answered by collective behavior. (...) Scientific communities likewise achieve quality through a cooperative process that includes checks and balances and ultimately rests on a foundation of goodwill and 'blind' elitism—blind in the sense that ideally anyone can gain entry, but only on the basis of a meritocracy. The tenure system and many other aspects of the academy are designed to support the idea that individual scholars matter, not just the process or the collective (Lanier, 2006)

This is exactly what happens in the Linux development hacker community, where “individual vision and brilliance” can be amplified “through the effective construction of voluntary communities of interest” (Raymond, 2001).

Thus, when applying this model to public life, such an individual is not an entity dominated by the collective nor a recipient of political solutions crafted by the administration. Instead, they are a citizen genuinely engaged in the public sphere, actively contributing to the quality of the entire community. According to John Rawls, preserving individual liberties, as well as the common good, needs active participation: “The safety of democratic liberties, including the liberties of nonpolitical life (...) requires the active participation of citizens who have the political virtues needed to sustain a constitutional regime” (Rawls, 2001). To ensure that citizens’ engagement in public debate is meaningful, it is crucial to perceive them not merely as residents demanding government services but also as co-hosts and stakeholders in the country. Citizens are not supposed to be merely consumers of content but also policy co-creators and amplifiers. With this perspective in mind, classical republicanism gains relevance. Essential elements of republicanism, such as publicity—the condition of being open and public rather than private or personal—and self-government, which implies the ability to co-shape public affairs (Dagger, 2002), appear to be significant indicators for citizens’ participation in online public debates.

Publicity can be understood both as discussing all important issues in public and as “the quality of people who are bound together by a common interest in how they are to live. They will thus have some sense of a (...) public good, and their

business, as a people, will be to govern themselves with an eye to that good" (Dagger, 2006). The republican civic discourse assumes governing in deliberative manner. As Cass Sunstein puts it, "republicans will attempt to design political institutions that promote discussion and debate among the citizenry; they will be hostile to systems that promote lawmaking as 'deals' or bargains among self-interested private groups" (Sunstein, 1988). The civic discourse will only succeed, however, if there is a sufficient supply of civic virtue; otherwise, the debate will be little more than a vain display that merely distracts attention from the *real* politics of bargaining for personal advantage (Dagger, 2006).

The self-government of citizens means protecting the members of society from domination and arbitrary power. Public issues should be discussed openly in order to guard against corruption or manipulation. Transparency and openness in civic discourse are, therefore, measures of protection against situations in which citizens could lose control over public life. Citizen engagement promotes an increase in transparency, which is the crucial value of open policymaking. The other way around, increased transparency also promotes civic engagement because debate participants can see the effects of their actions.

In line with these principles, policymaking, as a matter of public concern, should be openly discussed in public by engaged citizens. Participatory budgets (PB) are an example of this model in an online environment. Transparency in the allocation of public funds encourages active citizen participation. Thanks to openness in debating about the local community's expenses, citizens' sense of self-government increases. One may complain that the scale of this openness is not large if we take into account all public tasks, but it seems adequate in relation to the time possibilities and interests of the average citizen. There are, of course, citizens who can and want to engage to a much greater extent, which becomes possible, for example, in debates on topics related to their specializations (as in the Finnish off-road crowdsourcing described in Chap. 2) or during nationwide crises (as was the case with Iceland in 2010). What is important is that in PB, civic engagement is valued not necessarily because it leads to consensus but for its role in empowering every citizen to actively contribute to the public debate. Participants' engagement in debates seems particularly linked with seeking the most substantive arguments to support their theses and confronting alternative theses.

What other benefits do citizen engagement and independence offer from a policymaking perspective? Participation in the public debate through dedicated online platforms aligns with the current concept of *citizens as coproducers of public services*, as discussed by Helen K. Liu (2021). This approach suggests the potential to involve citizens in improving the quality of public institutions. It requires governmental agencies to issue open calls for online contributions. A crucial role of state administration in influencing such citizen engagement is transparency in accessing public information. According to King's (2007) concept, well-informed citizens can influence government policy when empowered with transparent information and suitable tools. Citizens become insightful when granted access to information and systems, including portals linked to Geographic Information Systems (GIS) and government service rating functions. The more the government perceives citizens as coproducers,

the more crowdsourcing activities can become complementary to government policies. Consequently, citizens can significantly impact policy design and have deeper involvement in policy implementation.

In online debates, citizens primarily invest their own time, which they dedicate to engage in discussions. For this civic engagement to be meaningful, participants must be assured that their involvement is not just an unproductive expression of opinions that get lost in the information noise. A meaningful contribution to public debate necessitates that citizens are able to distinguish authentic information from propaganda. With the vast increase in sources and content, many of which have questionable origins and quality, this task has become both more demanding and more crucial (Cohen & Fung, 2021). Therefore, a key task for the government should be to protect against disinformation and informational manipulation while preserving the debate as open as possible, even to controversial content. Protection against fake news and manipulation seems to be the most important task of public administration to ensure online civic engagement.

4.3.2 Deliberation, Antagonism, or Agonism: In Search of the Model of Online Debate

Now, let us examine the approach to conducting debates that appears most suitable for online discourse on public matters. Primarily, the deliberative approach, frequently referenced in this book, emerges as a foremost consideration. It is based on the works of J. Rawls (1971) and J. Habermas (1989), which some researchers consider one of the “distinct mechanisms responsible for the production of collective intelligence” (Landemore, 2012, p. 89).

Key features of debate in the deliberative model include consensus-oriented communication, reasoned opinion expression, judicious argument, equal participation, independent judgment, critical listening, and sincere decision-making (Stromer-Galley, 2007). A notable aspect of deliberation is its rationalism, where only the force of the better argument prevails. Of particular importance are the substantive value of the arguments (Mansbridge et al., 2010), inclusiveness and diversity (Steenbergen et al., 2003), shared responsibility (Stromer-Galley, 2007), sense of common interest, focus on consensus, and decentralization (Cohen, 2003). In this model, conflict is viewed as a threat arising from mutual non-recognition of validity claims. Group internal stability and shared identification with the result often outweigh the debate’s outcome. The deliberative stance generally assumes the creation of a unified, widely accepted outcome, discounting emotional statements. Advocates of deliberative democracy, influenced by Habermas, promote a form of participation characterized by dialogue in an “ideal situation,” reflecting Rawls’ emphasis on achieving consensus through neutrality and rationality (Paxton, 2020). For example, Jon Elster

(2012, p. 170) supports institutions that “minimize the role of interest” and “minimize the role of passion,” while Bruce Ackerman (1981, p. 21) defines deliberative democracy as engaging in “neutral dialogue.”

According to this model, the legitimacy of decisions stems from the deliberative quality of the processes leading to them. It emphasizes the importance of rational discourse, where participants exchange arguments free from coercion, aiming to reach a consensus that respects the common good. This ideal values the transformative power of dialogue, where through the force of better argument, personal preferences can give way to collective decisions that reflect shared values and mutual understanding. It suggests a shift from decision-making processes dominated by strategic bargaining and the aggregation of preferences to those that prioritize discussion and consensus-building.

One of the key strengths of the deliberative approach is its potential to enhance the legitimacy and acceptance of political decisions. When participants see that their voices are heard and that decisions are made through a fair and reasoned process, their trust in the policymaking system can increase. On the other hand, power imbalances and inequalities can distort deliberative processes. Not all participants have equal opportunities to contribute, and dominant groups may exert undue influence, thereby undermining the inclusivity and fairness of the deliberation. The presence of systemic biases and entrenched interests poses significant obstacles to achieving the ideal of a rational-critical public sphere.

Many researchers consider the deliberative approach to be an obvious normative model for conducting civic debate on the Internet. Antje Gimmmer (2001) claims that we can treat the “internet as a medium of deliberative democracy” that can “augment the public sphere.” Lincoln Dahlberg (2007b) claims that “there is a growing body of Internet-deliberative public sphere research,” referencing numerous works (i.e., Wilhem 2000, Fung & Kedl, 2000; Janssen & Kies, 2005). Stephen Coleman and Jay G. Blumer in *The Internet and Democratic Citizenship* (2009) argue that “it should be at the forefront of research into innovative technologies and cultural forms likely to generate more informed, inclusive, reflective and consequential online deliberation.” Liston et al. (2013) express intent to “bridge the gap between developments in normative deliberative theory and online participation,” and Helene Landemore (2021) argues that digital technologies might “one day allow us to reimagine the possibility of true online deliberation.” Also, among researchers focusing on CI in online public debates, there is a constant interest in the theory of deliberation. For instance, Paolo Spada and Lex Paulson (2023) argue that we can observe a “significant overlap of the categories of deliberation, participation, and collective intelligence;” Helene Landemore (2012, p. 98) argues that “the phenomenon of collective intelligence [is] emerging from inclusive deliberation;” and Joseph Capella contends that deliberation of online collectives can lead to intelligent outcomes (Capella et al., 2017).

However, the debate that actually takes place online is far from the deliberative ideal. It can even be said that there is a predominance of harmful antagonistic talk, especially on social media platforms. The antagonistic nature of most political

discourse on these platforms ultimately harms public conversation. Dahlberg (2007b) quotes the alarming opinion that online debate.

Becomes the arsenal of virtual civil war—civil wars among partisans at all levels. (...) I see conflict. I see an unwillingness to compromise. (...) I fear the extreme erosion of public trust not just in government but also in most things public and political. Instead of encouraging networked citizen participation that improves the public results delivered in our democracies, left to its natural path, the Internet will be used to eliminate forms of constructive civic engagement [of the] 90 percent of citizens.

Furthermore, as Capella et al. (2017) argues, the top-down imposition of norms consistent with a deliberative approach does not necessarily lead to an increase in collective intelligence:

Certain processes common in group deliberation can distort the information available to discussants through suppression of minority opinion, polarization, and the development of risky shifts. In deliberating groups, shared information often dominates or crowds out unshared information, reducing the diversity of information and ensuring that groups do not acquire the full range of information available.

In deliberative debate, information that is commonly known tends to overshadow or suppress information that is not popular, diminishing the diversity of gathered knowledge and preventing groups from accessing the full spectrum of possible solutions.

As Collins et al. (2020) put it, “Talk on social media platforms is so antagonistic that people opt out or disengage from the public conversation. The effect on the common conversation is a reduction in both the number and quality of voices in the discussion.” In the context of social media, where anonymity and the lack of face-to-face interaction can embolden more extreme expressions of disagreement, antagonism can escalate quickly. It can manifest in various forms, such as trolling, harassment, and the spreading of misinformation, all aimed at undermining the validity of the opposing viewpoint rather than engaging in constructive debate.

Antagonism results from the intentional exclusion of the opponent from the debate (Laclau & Mouffe, 2001, p. 129). Participants not only disagree with each other’s views but also engage in actions that discredit, belittle, or attack the opposing side on a personal level. It involves more than just the presence of differing opinions; it is characterized by the intentional exclusion of the opponent from rational discourse, often through emotional insults and attributing malevolent intentions to them. This means discrediting the opposing side’s opinions, considering them entirely unacceptable, and not creating any ground for discussion. The recent phenomenon of *cancel culture*, which entails the ostracism, boycott, or shunning of views that a certain group deems unacceptable, serves here as a striking example. The exclusion engenders conflicts that are antagonistic in nature, signifying the absence of a possibility for a *rational* compromise among participants (Collins et al., 2020). Antagonism undermines the potential for productive political dialogue by creating an environment that is hostile and divisive. Instead of fostering a space for understanding and rational compromise, it polarizes participants, often entrenching them further in their

original positions. This dynamic is particularly problematic when social media algorithms amplify extreme views and create echo chambers that reinforce antagonistic attitudes, further eroding the possibility of meaningful political discourse.

Is there no remedy for the prevailing antagonism on social media? As can be easily observed, deliberative values have a very limited impact on online discussions. However, an alternative model of debate has been proposed in recent years as a different approach to the online public sphere (Collins et al., 2020; Crawford, 2016; Dahlberg, 2007a, 2007b). We can refer to it as the agonistic model, where the key concept is the Greek word *agon*, meaning struggle. However, this struggle adheres to honorable rules inspired by the stance of ancient heroes, the early Olympic games, and the public activity of citizens of ancient Hellas in the agora. This concept, according to Collins et al., “runs contrary to centuries of political theorizing that attempts to reduce conflict in favor of rational consensus and compromise” (Collins et al., 2020). The theoretical foundations of this model are based on the works of Friedrich Nietzsche, Hannah Arendt, Carl Schmitt, and Michel Foucault, with contemporary thinkers referencing this theory, including Chantal Mouffe (2013), William Connolly (2005), David Owen (1995), and Marie Paxton (2020).

Before examining the potential application of agonistic principles to online debates, let us delve briefly into this theory, initially formulated by Friedrich Nietzsche. In much of his writings, Nietzsche emphasizes conflict as a natural element of life, an ineradicable dimension of society. However, *agon* is not a struggle for annihilation; it is rather a contest of heroes who respect and admire each other, even as opponents (Tuncel, 2023). Nietzsche’s agonistic philosophy is based on the concept of *active resistance*, which serves as a strong stimulant for an active life. After all, resistance and conflict “drive even the unskilled man to work” (Acampora, 2018). However, Nietzsche’s exploration of active and reactive forces demonstrates how extreme reactivity can undermine the spirit of agonism (Tuncel, 2023). Although the drive for change is a life-giving force, it should not negate the essence of *life* itself, which includes high ethical values. Instead, through critique, it should continuously renew its true meaning (Siemens, 2021, pp. 35–36). In Nietzsche’s view, contestation is not a total negation but a way of purification and renewal of what exists, and bestowing noble traits upon conflict aims at transforming the destructive aspect of human existence into a constructive one (Siemens, 2021, p. 53). The emotions and ambitions that emerge in contests direct participants toward higher goals “to be like heroes and gods” (Tuncel, 2023). Thus, Nietzsche’s concept of social life is not entirely variable, subjective, and devoid of enduring values, as some authors (e.g., Villa, 1992) believe. Nietzsche argues that “social and cultural goods are produced in agonistic interactions, and that agonistic engagements may be productive means for reproducing values and creating them anew (...). The *agon* (...) is able to produce values, reinvest values, and recreate them,” benefiting not only the involved individuals but also the entire community (Acampora, 2018).

Arendt extends the concept of agonism from outstanding individuals to political communities, referring, like Nietzsche, to ancient Greece. Her concept of performative politics is notably characterized by its agonistic nature, where participants compete for public attention, chances for self-disclosure, and the luminous allure of

glory (Berger, 2011). Arendt perceives the public sphere as the place where citizens vie with each other through public activity: “The public realm itself, the polis, was permeated by a fiercely agonistic spirit, where everybody has constantly to distinguish himself from all others, to show through unique deeds or achievements that he was the best of all others” (1998, p. 41). She also emphasizes the importance of exposure to contrasting viewpoints to foster an “enlarged mentality” (Arendt, 1968, p. 241). This notion refers to the development of a more comprehensive and informed opinion by evaluating a particular issue from alternative perspectives (Capella et al., 2017).

Contemporary authors referring to agonism view debate as a competitive arena, marked by dynamic tension, where participants compete for recognition, precedence, and acclaim (Benhabib, 1992). Contestation is seen as an important incentive and a cure for apathy, engaging people in debate (Paxton, 2020, p. 12). Competition and conflicts are seen not as abuses to eliminate but as core processes (Owen, 1995, p. 161). Both individualistic (Owen, 1995, p. 58) and collective identity (Mouffe, 2012) are valued, and some participants usually gain a dominant position in the group based on prestige (Laclau & Mouffe, 2001, p. 63). Interdependence of different positions is observed, as the alternative solutions are fueled by confrontation and developed in opposition to each other (Connolly, 1991, p. 188). Perseverance in defending an individual position against the majority is valued (Connolly, 2005, p. 123). Emotional statements are acceptable, as completely separating statements from their accompanying emotions is impossible. The opponents should be treated with respect as esteemed adversaries, not enemies (Lowndes & Paxton, 2018). Through collective contestation, participants develop autonomy and group identity relationally. Agonistic relations are practices and specific encounters in which the participants “recognize the legitimacy of their opponents” (Norval 2014, p. 75).

How can this model of debate be applied to social media, and how can it help us escape the trap of antagonism? Crawford (2016) opened a pathway in this direction by asking, “Can an algorithm be agonistic?” and suggesting that it is indeed possible. Instead of striving for an impossible Rawlsian “overlapping consensus” (Rawls, 1999) standard, we should acknowledge that some conflicts are incommensurable and attempt to transform them into something productive for society. The agonistic model appears to capture essential truths about human nature and the nature of participation in the public sphere that are overlooked by the deliberative approach. Being aware of the dangers posed by unchecked antagonism, agonists propose principles that ensure competition does not lead to chaos but to mutual benefits. Following this lead, Collins et al. (2020) present a very interesting proposal on how, following the guidelines of agonists, to impose an *algorithmic obligation* on major internet platforms. Drawing on Fossen’s (2014) concept of political obligations as practical commitments of entities responsible for the existence of the public sphere, Collins advocates for a “model of agonistic political obligation.” This model obligates users to engage in public discourse, maintaining their views and identity but not depreciating the opponent. It also obligates platforms like Facebook and Twitter to foster highly aspirational, normative public discourse rather than allowing hate speech. Improving the quality of the debate is, at least declaratively, one of the important goals of these platforms.

Twitter, for instance, has committed itself to promoting “healthy debate and critical thinking” (Dorsey, 2018). The primary way social media shape public debate is through their technology, so their task would be to design the agonist algorithms that determine which information is presented to whom and in what manner, thus supporting certain user attitudes. Marichal and Neve (2020) state that “In the asynchronous, relatively anonymous, sentiment-driven discourse of Twitter, aiming for the Habermasian ideal of perfect speech as proposed by normative democratic theories appears highly unrealistic. Agonistic theory, with its recognition of sentiment expression and the creation of networked publics as valid forms of democratic expression, offers a more feasible normative ideal.”

In the proposed agonistic model of debate in social media, Connolly’s advocated “ethos of agonistic respect” comes to the forefront. It is an ethic where “alter-identities foster (...) respect for the differences that constitute them” (Connolly, 1991, p. 166). This approach is grounded in the recognition of the legitimacy of opposing viewpoints and the importance of engaging with them in a constructive and respectful manner. Agonistic respect is an approach to debate that underscores the importance of good-faith arguments and generosity in expressing opposing views. This ethos encourages participants to “test, challenge, and contest pertinent elements in the fundaments of others” (Connolly, 2005, p. 123). It also requires an attempt to understand the opponents’ stance, motivations, and arguments: “You must listen attentively to the one whose faith you contend against” (Connolly, 2004, p. 510), respecting the rules of fair engagement and preserving openness for contestation (Collins et al., 2020). Such an ethos overcomes a primary danger in contests between antagonistic rivals, which Connolly refers to as *existential resentment*—a situation where a group excluded from the discourse leaves the public sphere with a sense of injustice.

Marichal and Neve (2020) believe that it is possible to establish criteria based on the specifics of the ethos of agonistic respect, which can serve as evaluative criteria for a political talk on social media platforms. To achieve this, they conducted an evaluation of the most significant posts from the Twitter gun control debate that took place in February 2018. Through this analysis, they identified seven different modes in which individuals present the value of their opponent’s positions, evaluated the pluralism within the discourse space in the comment threads, and assessed the potential to provoke resentment among adversaries. Subsequently, they ranked the modes of debate in order from most antagonistic/least agonistic to most agonistic/least antagonistic. The seven categories, ranging from least to most agonistic, are: (1) furtive/secretive (accusing opponents of acting in bad faith), (2) cravenly opportunistic (viewing opponents as immoral or indecent), (3) irrational/sentimental (charging opponents with irrationality), (4) willfully ignorant, (5) misunderstanding/misguided, (6) contingently wrong (opponents’ motives might be good but they are impractical), and (7) engaged in the process of reciprocal inquiry (inquiring whether users respect each other’s positions). The results of this work are presented in Table 4.6.

What is striking in this analysis is that Twitter users who demonstrated the highest engagement were much more likely to exhibit antagonistic behavior. However,

Table 4.6 Types of antagonistic and agonistic discourses on Twitter (Marichal and Neve 2020)

Presentation of the opponent's value position	Possibility of ressentiment	Space for value pluralism	Depth of argument	Popularity
Furtive/secretive	High	None	Low	Moderate
Cravenly opportunistic	High	Very limited	Low	High
Irrational/sentimental	Moderate	Somewhat limited	Moderate	High
Willfully ignorant	Moderate	Somewhat limited	Moderate	High
Misunderstanding/misguided	Low	Unlimited	High	Low
Contingently wrong	Low	Unlimited	High	Low
Reciprocal inquiry	Low	Unlimited	High	Low

increasing the level of agonistic respect reduces the possibility of ressentiment, mitigating anxieties and frustrations stemming from the exclusion of certain viewpoints from the debate. As observed in the earlier analyzed Twitter debate on mandatory COVID-19 vaccinations, attempts to exclude opponents of this mandate and question their right to speak increased ressentiment, leading both sides of the dispute to retreat into their bubbles, with fewer connections between them over time. Moreover, what is particularly relevant from the perspective of collective intelligence is that as the level of agonism increases, so does the quality and depth of argumentation, leading to more useful outcomes in online communication.

How can agonistic respect be incorporated into the policy of social media platforms in the form of “algorithmic obligation,” as Collins and his colleagues suggest? Algorithms used by Facebook, X/Twitter, and other platforms, which are responsible for promoting certain posts by increasing their visibility, according to Marichal and Neve (2020), “should emphasize tweets that achieve these [agonistic] objectives.” Making Twitter more agonistic could be achieved by employing natural language processing (NLP) and deep learning technologies to develop algorithms and models capable of introducing counter-perspectives to antagonistic tweets and threads. An NLP model could be tailored to detect keywords associated with discourse and debate, encouraging the opposing side to articulate their arguments more clearly. This model would evaluate each response for its alignment with the original tweet and calculate a cumulative score for the entire thread. Beyond simply scoring responses, the model could also analyze the sentiment of the replies, distinguishing between constructive criticism and mere antagonism. It could further identify patterns of echo chamber formation, where users predominantly interact with like-minded views, and suggest exposure to a broader range of perspectives. By default, Twitter could feature tweets with low bias scores or prompt users to choose whether they wish to *opt-in* to make biased tweets/discussion threads a lower priority in their comment stream (Marichal & Neve, 2020). Moreover, Twitter could introduce features that enable

more nuanced discussions, such as extended rebuttals or dedicated areas for structured debates. These mechanisms would not only improve the quality of discourse but also promote a culture of respect and understanding among users, fostering a more effective and inclusive exchange of ideas.

Agonistic arguments would aim for respectful engagement with adversaries, thereby leading to the development of an intricate structure of arguments based on the views and world perceptions unique to a group's identity. Comparing competing approaches to a problem can broaden our understanding of significant policy issues and thus contribute to a more informed public debate. As Jacquet and Thinyane (2023) assert, online platforms where conflict is agonistic can provide a conducive environment for sparking innovative solutions. Imposed consensus and detachment of presented views from their contextual background can lead to the diminishment of the voice of independent groups and individuals, thereby lowering the level of collective intelligence.

In summary, deliberative theorists argue that discussions should aim for a consensus on political issues, with arguments in the debate being as impartial as possible, detached from the personal emotions of participants, and seeking mutual understanding. On the other hand, agonists have emphasized that emotions and identities are unavoidably intertwined with political discussions and conflicts (Trygvason, 2021). According to the agonistic approach, the claim to impartiality in debate proposed by deliberativists is, in practice, unattainable, and bias or inequality are inherent features of the social realm. Agonism focuses on the opportunities this inequality provides, namely leveraging the effectiveness based on the recognition of one's values that shape the unique identity of the group participating in the debate. Research in online public debates (Capella et al., 2017) has shown that *disagreement* can provide exposure to multiple perspectives and is thus thought to foster the kind of careful reflection needed to arrive at a reasoned opinion. Also, findings from the experiment described in Sect. 4.2.1 confirm that properly managed conflict can lead to an increase in the quality of debate.

It is important to note that while agonistic debate seems like a promising proposal for improving the quality of discussions on public matters on large online platforms, it is not the best model in every case. Conducting debate in a deliberative manner, prioritizing *pure* arguments, and conducting collective analysis aimed at achieving a common outcome can be an effective method of operation in small or medium-sized communities where the bonds cementing the group (e.g., common goals) are present, and a consensus is achievable. Exemplary case studies described in the previous chapter (primarily Deliberatorium and Loomio) show the possibilities of deliberative debate in dedicated platforms. The epistemic values that civic deliberation can bring to policymaking are undeniable. However, we must remember that for organizations, actual deliberation requires specific conditions which are often lacking in free debates conducted on social media. Nonetheless, there are no obstacles to attempting to implement these conditions in dedicated policy projects carried out on specialized platforms.

4.4 Conclusions. Collective Intelligence in Future Policymaking

Time and again, we hear of scientific breakthroughs poised to alter the fate of the world. Such discoveries are most commonly made in fields like medicine, biochemistry, physics, or materials engineering. But could collective intelligence, harnessed through online networks, give us a chance for such a breakthrough in the domain of public policy sciences? Can leveraging the wisdom of crowds in the creation and execution of policy programs lead us to a new model of social life? The answer is both yes and no. Open policymaking, as it stands today, has not revolutionized the public policy. Not yet, at least. However, considering the rapid technological changes that have affected and continue to affect societies globally in recent years, we can assert that the potential of CI remains to be fully unleashed. Large Language Models, operating in the form of services like ChatGPT, exemplify the marriage of AI technology and the collective intelligence of internet users. Yet, for the potential of CI to be truly unlocked, the people interested in public policies, especially political decision-makers, researchers, and civic activists, need to become aware of which social and technical factors significantly influence the level of CI now, and which may do so in the near future. The policymakers also need to understand which collective cognitive processes are effectively realized in the digital space and which still pose challenges.

Adopting an appropriate debate model and maintaining citizen engagement and independence are characteristics of online communication that can increase the level of collective intelligence and, consequently, lead us to more effective and substantively valuable policymaking. What other features of projects utilizing CI, as outlined in this book, can facilitate a more transparent and improved development of public policies in the future? And, going further, which new technologies can aid in bolstering collective thinking processes and be implemented in CI projects? This final section will summarize the key findings in the preceding chapters and delineate the pathways for advancing open policymaking in the coming years.

The first aspect of CI worth special attention, long considered by researchers but still not fully realized in practice, is the mapping of debates and the use of the mental maps created in collective thinking. As Francis Heylighen noted in the 1990s, when the Internet was still a space accessible to a few rather than a common communication channel, the desired outcome of the work of a networked collective is not a single solution but a kind of collective mental map (CMM) (Heylighen, 1999). What exactly is *a mental map*? It is a cognitive or mental representation of an individual's perception of their environment. Each person independently creates a mental map relating to the world around them. The existence of such mental models is grounded in a well-established theory in the social sciences. Psychologists Kahneman and Tversky (1981) introduced the concept of *the framing effect*—a cognitive bias where people make decisions based on the context in which options are presented. The way information is framed significantly influences their decisions and judgments. It is important to remember that “mental maps are not objective reflections of the real

world ‘out there,’ they are individual constructions, based on subjective preferences and experiences” (Heylighen, 1999). According to Heylighen, the quality of the mental maps contributes to a better understanding of the environment, improved capacities for action, and the possibility of achieving our goals.

Kahneman and Tversky’s research demonstrated that human reasoning is not an operation conducted according to the principles of formal logic but rather resembles a simulation of reality: before taking action, we evaluate various available options, imagining what might happen if we choose them. The real art is realizing which thought patterns, or frames, we use in our mental model. Frames direct our mind to what is most important in a given model; they are cognitive shortcuts that shape our thought space, allowing us to think within a specific context and focus on relevant facts while ignoring the irrelevant ones. Frames consist of our values, identity, beliefs, and accumulated experiences. How we act depends on what we can discern from our surroundings with the help of our mental maps. Mental models allow us to recognize patterns and predict outcomes, understand the circumstances we find ourselves in, and comprehend the context particular to a given situation. They also enable us to place facts in new, significant contexts and imbue them with meaning. This type of reasoning, in other words, is the concept of *bounded rationality* that we recall from Chap. 1, fitting into a post-positivist approach to policymaking.

Neuroscience research confirms that we operate with mental maps in our thought processes. Our brains utilize stored sequences of patterns to predict the environment continuously. The brain’s ability to consistently identify patterns, regardless of context, size, or orientation, is critical. However, intelligent thinking is not merely about possessing a large repository of patterns; rather, it is primarily about the capacity to efficiently retrieve and utilize relevant experiences. To find a solution to a problem, neurons *compete* to provide the best-fitting models, conducting a kind of voting to find the best solution (Hawkins, 2021, p. 110). The analogy of the brain as a collective of neurons in relation to collectives composed of many individuals seems apt.

What, then, are collective mental maps? Heylighen defines them as a type of shared memory. Various contributions from members of the collective are recorded and stored in this memory, ensuring the information remains available for as long as necessary. The biological literature provides several examples of swarm intelligence systems that display the characteristics of a collective mental map (CMM): the behavior of social insects (e.g., the foraging of ants, termite nest-site selection and construction) are the most classic examples (Skaržauskienė, 2015). The storage capacity of this memory generally exceeds the capacities of the individual participants’ memories. A CMM is developed by superimposing a number of individual mental maps. Each individual has a slightly different map, and diverse viewpoints can raise the level of group intelligence. This diversity of viewpoints “is healthy since different individuals may complement each other’s weaknesses. Imagine that each individual would have the same mental map. In that case, they would all find the same solutions in the same way” (Heylighen, 1999). Let us recall that the positivist way of thinking about public policy assumed that experts should unite into one team and develop a uniform stance. Relying on precise and specific expert knowledge was

quintessential to the positivist approach to policymaking. In reality, it is different: the strength of collective reasoning lies in the ability to use various frameworks to analyze a phenomenon from different perspectives and, only on this basis, develop a solution (Cukier et al., 2021).

However, a collective mental map is intended to be “not merely a registry of events or an edited collection of notes [but] a highly selective representation of features relevant to problem-solving” (Heylighen, 1999). A mere record of all communications does not constitute a mental map. As time progresses and more people participate in a discussion, the record will expand, making it increasingly challenging to distill any useful guidelines for action. Therefore, the accumulated knowledge must be somehow organized, opinions and arguments grouped, and key thought patterns and frameworks identified. Debate mapping should allow for the organization of discussions and capture of their most important threads, as well as the identification of: key themes of the debate, the most significant flashpoints, the most influential opinions, differences between alternative positions, and key individuals shaping the discussion. Mental maps enable contextual thinking. This means not treating individual facts and statements as isolated data but rather understanding them in the context of the entire set of positions, applied cognitive shortcuts, circumstances referred to, and cultural codes. Creating a CMM involves several steps. It necessitates gathering feedback on the quality of proposed solutions and the adequacy of thought frameworks to situations, as well as optionally averaging opinions and/or selecting the best opinion. This can be compared to the cooperation of ants and the knowledge base they create based on pheromones: “The pheromone network does not record all movements made by all ants; it only registers those collective movements that are likely to help solve the ants’ main problem, finding food” (Heylighen, 1999). Therefore, an integral part of any CMM should be a preference function for choosing the best action or solution.

As we can conclude from analyzing the projects described in this book, the concept of creating collective mental maps has yet to be fully realized. We can still ask, however, how debate mapping has been implemented so far in projects related to public policies. It appears that the projects closest to achieving this concept are those that impose a structured discourse based on the collaborative argumentation of their participants, as seen in Deliberatorium and Better Reykjavik. In these projects, users actively participate in crafting the debate map by selecting a thread to post their comments and choosing a position for or against a given issue. This approach results in the creation of a tree-structured network, where each post embodies a distinct issue (a question awaiting response), an idea (a potential answer to the question), or an argument (supporting or opposing an idea or another argument), thereby documenting the diverse viewpoints represented by the debate participants. The advantage of this solution lies in its ability to visualize the debate as a tree of arguments, enabling an assessment of the quality and volume of arguments linked to specific threads and the level of support expressed by participants for varied positions. However, a challenge arises as users must manually determine which part of the debate map their contributions should belong to. Additionally, this process might necessitate the involvement of debate moderators tasked with aiding participants in deconstructing their contributions into their elemental issues, ideas, and arguments, identifying the

main branches and subsequent sub-branches of the debate, and merging overlapping arguments or threads. Moreover, as some experts in Chap. 3's evaluation have pointed out, this formalized mode of participation can limit user spontaneity and exclude contributions that are not clear-cut arguments for or against, such as comments or unconventional suggestions.

The need to design a technical solution that would automate the construction of deliberation maps has long been a subject of interest for debate organizers. Unfortunately, the technical level of Natural Language Processing (NLP) tools has not been sufficiently high to create maps automatically similar to those used in the Deliberatorium. However, the dynamic development of generative artificial intelligence based on Large Language Models observed since 2022, with ChatGPT being the most notable example, gives hope for significant progress in this area. In the Decide Madrid project, the attempts to implement solutions that search through existing statements and suggest debate threads similar to participants' inputs can be considered a promising step in the right direction.

New technologies leveraging generative AI may soon lead to a significant breakthrough that facilitates better debate mapping and, consequently, the creation of collective mental maps in real-time. An example of such technology is Argument Mining (AM), which involves using computer-associated data science techniques for the study of argumentation patterns and the automatic recognition and extraction of arguments from free text to interconnected argument trees. Recognizing and understanding arguments are central to policymaking and professional activities in all walks of public life. Many experts (e.g., Budzyńska & Villata, 2018; Lippi & Torroni, 2016) agree that the importance of AM is growing due to the explosion of data available online. Analyzing the argumentation in policy debates requires capturing the full complexity of interpersonal connections, social interactions, relations between clusters of arguments, and their interplay. By detecting the argumentative structure of texts and differentiating between claims, premises, and counter-arguments, AM is one of the most promising approaches to combating disinformation and information noise, among the most significant threats to policy debates. Fake news often includes poorly constructed or fallacious arguments. By identifying these anomalies, AM can help assess the credibility of content (Lippi & Torroni, 2016).

From a technical standpoint, Argument Mining (AM) integrates (1) Deep Learning, the training of Large Language Models (primarily based on BERT technology, i.e., Bidirectional Encoder Representations from Transformers) for argument prediction, and (2) the use of sentence semantics to construct a regulatory-ontological system focused on identifying argumentative patterns within the syntactic trees of sentences. These patterns are defined around argument linkers, i.e., words that connect the premises and claims of arguments (see Segura-Tinoco et al., 2022). Only recent advancements in computer hardware and mathematical techniques have enabled researchers to train language models on extensive datasets, tapping into the potential of unsupervised learning and applying it to various NLP tasks. The latest trend in this domain involves pre-training extraordinarily large models using deep neural network representations on an unprecedented scale of data.

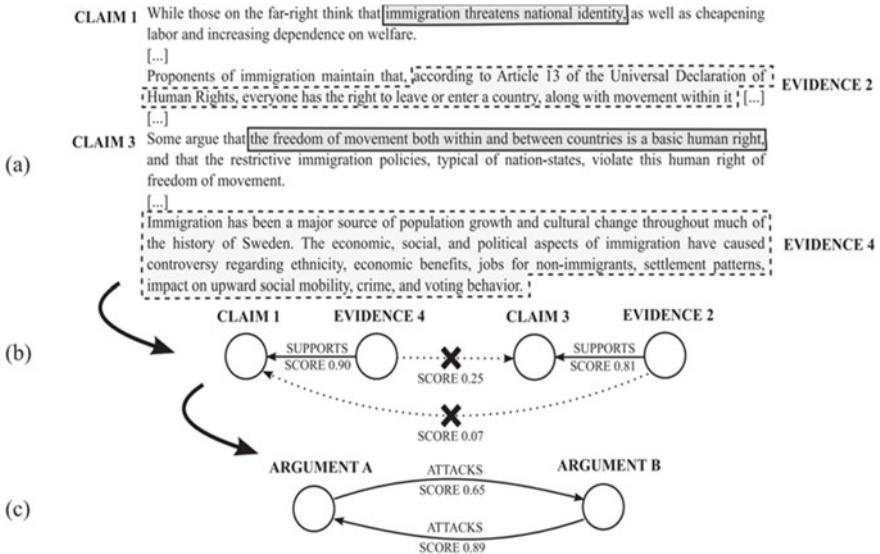


Fig. 4.15 A process of the automatic extraction of arguments from text (Lippi & Torroni, 2016)

Figure 4.15 presents a process of the automatic extraction of arguments from text. Initially, sentences identified as argumentative are pulled out from the input document. Within these sentences, components of arguments, specifically claims and supporting evidence, are pinpointed (as seen in Fig. 4.15a). Following this, relationships between these argument components are predicted (Fig. 4.15b), facilitating the assembly of full-fledged arguments. Ultimately, interconnections between the arguments are deduced to produce a comprehensive argument graph (Fig. 4.15c) (Lippi & Torroni, 2016).

Clearly articulated arguments and well-organized debates form the cornerstone of civilized society and intellectual life. Processes of argumentation run our governments, structure public debates, and frame social beliefs. As demonstrated in this monograph, developing a wide-ranging repertoire of arguments in an open policy-making process presents an opportunity for public policies to harness the potential of CI. Approaches that move beyond the simple collection of opinions and judgments to intricate semantic network representations of complex issues hold real promise to increase the applicability of CI in policymaking (Capella et al., 2017). For this reason, I believe that developing technologies for automatic extraction and organization of arguments could be an important step toward a broader and more effective use of online collective intelligence in shaping public policies. If the concept of open policymaking, understood as a collaborative and transparent approach to public issues, going beyond the engineering model based on experts and presenting diverse viewpoints, is to find broader application in areas beyond urban policy, tools for algorithmically-supported debate analysis would be very useful. Allowing participants full freedom of expression while simultaneously extracting the most important

arguments from their discussions, positioning these arguments in relation to each other, and visualizing them in real time would enhance the transparency of the debate and enable the meaningful involvement of a wide range of stakeholders. It would also facilitate the recording of diverse, distinct argumentation paths in collective mental maps, allowing for the exploration of alternative solutions if the initially selected policy fails.

Another technique that can aid in creating collective mental maps of large crowds and enable the analysis of citizens' spontaneous reactions to public events is social network analysis (SNA), as described in Sect. 4.1.2.2. SNA can be used for creating visualizations of debates through graphs. As we recall, social media users react swiftly to public events, as seen in the case of the consumer boycott associated with the war in Ukraine. Analyzing the boycott organized by users of X/Twitter and its correlation with the stock prices of companies operating in Russia can enable policymakers to discern citizens' opinions more swiftly and respond accordingly. Social media research can be conducted rapidly, with new data potentially gathered daily, and the various states of the debate can be visualized in real-time, allowing for the identification of influential opinions and emerging trends.

The post-positivist approach to policymaking proposed in this book, which goes beyond the positivist, rational-comprehensive model based on experts, gathers a variety of perspectives from citizens. This approach is most effective when adopting an agonistic model of debate, which, in many instances, can produce more insightful outcomes than the deliberative approach. Observing debates conducted agonistically confirms that exposure to contrasting viewpoints fosters an "enlarged mentality" (Arendt, 1968, p. 241). "Disagreement" exposes individuals to multiple perspectives, thereby fostering the careful reflection necessary to form a reasoned opinion. This exposure enables individuals to develop a more comprehensive and informed perspective by examining an issue from alternative viewpoints (Capella et al., 2017). Such a model not only reduces antagonism and diminishes radical polarization but also enhances collective learning, as the greater diversity of connections increases the likelihood of transmitting novel information. As evidence, Capella et al. describe an intriguing study that analyzed data from online group deliberations about the American presidential election in the year 2000. This study reported an interaction effect between the quality of political discussion and the perceived disagreement within discussion networks:

Those with the lowest levels of factual knowledge about political issues are embedded in low-disagreement networks with little political talk. Those in high-disagreement networks or low-disagreement, high-talk networks have elevated issue knowledge. Although these data are about people talking politics with others, disagreement plays a consequential role in advancing issue knowledge measured as accurate responses on political issue questions (Capella et al., 2017)

The agonistic algorithmic obligation for social media corporations, as proposed by Collins et al. (2020), emerges, therefore, as a promising solution.

Many researchers consider the diversity of presented positions and overall high cognitive diversity within a collective as intrinsic characteristics of CI. The plurality of groups and values is seen as legitimate and necessary for one's group identification.

Connolly argues that the process of constituting the identity of any group can only advance by progressively contrasting it, even if only implicitly, with what is different. Identity is ontologically dependent on difference. However, might excessive diversity sometimes be a problem? Let us recall again the study by Aggarwal et al. (2019), which showed that the diversity measure has a curvilinear (inverted U-shaped) relationship with collective intelligence, with high levels of diversity disrupting CI. Excessive diversity also lowers team learning abilities. It was observed that high levels of diversity impede team learning and group coordination, interrupting the flow of CI (Aggarwal et al., 2019). Similar conclusions appeared in the evaluated projects described in Chap. 3: initially, an increase in diversity significantly boosts performance, yet as diversity grows, the incremental benefits of each additional person or group start to wane, and the costs of making the group more diverse increase. Moreover, while diversity is beneficial for collecting a range of arguments, an excessively high level of diversity can lead to difficulties in understanding the structure of argumentation and in making a reasoned evaluation of the debate as a whole.

This implies that to harness the collective intelligence of a community, its members should be able to communicate effectively, share a common set of goals, and understand the same cultural codes. It seems that a certain degree of cohesion within the collective is necessary, and fragmentation into groups with conflicting aims and values can be an obstacle for CI. Consequently, while recognizing the advantages of the agonistic model of debate, it is necessary to express some reservations about those forms of agonism that view the political space as an arena for endless debate without any objective criteria for equity and the common good. Some representatives of contemporary agonism, especially Laclau and Mouffe (2001) and Honig (1993), emphasize the appreciation of marginalized discourses that challenge dominant (*hegemonic*) opinions. In this approach, the promotion of minority, contesting identities is carried out in opposition to an identity grounded in a given group. Hence, the opposition and contestation by minority groups may seem more important than the welfare of the entire community. This represents a grave error, as it risks prioritizing the specific interests of contesting groups over the common good, defined as the welfare of the collective as a whole. The danger lies in a minority identity that exploits the challenge to a major identity, advocating for its particular benefit at the expense of the common good. A delicate balance exists between the pluralism of positions that enriches collective intelligence and the diversity that hampers communication and cooperation, a distinction that policymakers should heed.

In his book *Ultrasociety* (2015), Peter Turchin posits that cooperation is foundational to the success of historically prosperous societies. According to Turchin, the most crucial aspect of cooperation is the cultural transmission of behaviors within societies. This transmission allows people to assimilate certain norms of conduct not only from their parents but also from other members of the community. The development of cooperation among a broad range of individuals, extending beyond kinship ties, was made possible through the intergenerational transfer of certain common moral principles. This echoes the *common cultural patterns* described in Chap. 2, which, according to Michael Tomasello, are transmitted through cultural evolution, where ethics of cooperation were conventionalized, institutionalized, and, having

become normative, led to the creation of collective intentionality (Tomasello, 1999). In discussing the abilities of certain social groups for coordinated, collective action, Turchin refers to the Arabic term *asabiyya*, which signifies a kind of group feeling, a concept of social solidarity with an emphasis on unity, group consciousness, and a sense of shared purpose and social cohesion. Its members can stick together and cooperate, allowing a group to protect itself against enemies. *Asabiyya*, or social cohesion, can be reinforced within a society that exists alongside another competitive society, aligning with the agonistic model of competing communities.

The sense of group solidarity and shared goals, and thus the collective pursuit of the common good, is deeply rooted in communities whose experiences transcend the *here and now*. Landemore highlights that the collective memory of communities, encompassing evolutionarily developed experiences passed down from generation to generation, distinctly influences the level of group intelligence: “When intelligence of a collective is extending not just through space (including many people), but also through time (including the knowledge of more than one generation), the room for both memory and experience is made” (Landemore, 2012, p. 241). The neurobiological findings mentioned earlier in this section confirm that utilizing memory has a decisive influence on creating mental maps and holistic, intelligent thinking. In the case of collectives—similar to individual minds—it can be said that shared memory, which stores the collective’s past experiences, plays a crucial role in collective thinking. Past experiences are preserved as patterns in shared memory, which heavily influence the prediction mechanisms and adjust the system based on new knowledge. The collective mind’s ability to recognize patterns invariantly, meaning recognizing them in various contexts, sizes, or orientations, is fundamental. This leads us to conclude that intelligence of a collective is determined not only by storing many patterns in a shared memory but, more crucially, by efficiently retrieving and referencing experiences that are most pertinent to current challenges.

Maurice Halbwachs’s observations, described in Chap. 3, allow us to understand that collective memory in social and political processes does not refer to a closed past, but rather, it is the single most important perspective (and a strong mental model) through which members of the collective observe present events. Halbwachs situated memory in the present rather than in the past, describing it as a social process that looks forward instead of backward (Bachleitner, 2022). It is the process through which “the past is not preserved but is reconstructed on the basis of the present” (Halbwachs, 1992, p. 40); therefore, both individual and collective memory share the property that they are “not static but dynamic—not fixed but ever-changing. A memory is essentially re-remembered (i.e., reconsolidated) each time it is recalled” (Anastasio, 2022).

How is the process of collective memory currently addressed in open policymaking projects? The evaluation of projects described in Chap. 3 reveals that most cognitive processes in online participation function reasonably well. Yet, a discernible flaw in nearly every implemented project is the utilization of collective memory. Project managers frequently recognize the pivotal role of memory in CI, but they generally find themselves without the necessary tools for its effective implementation. Consequently, the experiences collected by project participants are

seldom preserved in a useful form. Although some of these experiences are archived, their accessibility is limited, leading to infrequent use in practice. There exists a short-fall in methodologies for integrating current cognitive processes with the archives of collective memory, such as identifying past issues and solutions that mirror present challenges. High participant turnover further complicates the situation, with many participants often unaware that they are encountering problems that have already been addressed. Furthermore, the effectiveness of previously implemented solutions could be evaluated if the collective memory system was operational.

Hopes for improving this situation have been buoyed by the dynamic development of generative AI using Large Language Models observed since 2022. For instance, the Better Reykjavik project is currently testing an AI Chatbot based on a dedicated LLM, designed to facilitate users' communication in natural language about the accumulated knowledge and past experiences stored in the system. But how does the development of AI technology relate to CI? As Jeff Hawkins (2023) asserts, the recent advances in AI are not based on fundamentally new algorithms. Instead, the progress is largely due to the ability to train AI systems using massive datasets, with the most extensive dataset available being the entirety of text found on the internet. Language AI systems like ChatGPT excel because they are trained with billions of language examples. Going further, Jaron Lanier (2023) believes that the most accurate way to understand technological breakthroughs related to LLMs is to imagine it as an innovative form of social collaboration.

A program like OpenAI's GPT-4, which can write sentences to order, is something like a version of Wikipedia that includes much more data mashed together using statistics. (...) The new programs mashup work done by human minds. What is innovative is that the mashup process has become guided and constrained, so that the results are usable and often striking. This is a significant achievement and worth celebrating—but it can be thought of as illuminating previously hidden concordances between human creations, rather than as the invention of a new mind. After all, what is civilization but social collaboration? Seeing AI as a way of working together, rather than as a technology for creating independent, intelligent beings, may make it less mysterious (Lanier, 2023)

Therefore, ChatGPT and similar models can be seen as a form of social collaboration, the product of the collective intelligence of all internet users whose entire content was used for training. The strength of LLMs lies not in AI algorithms but in the CI that created the opinions, definitions, and innovative ideas, which were then recorded in collective memory. Moreover, as co-creators of this collective intelligence, internet users continuously create new types of data, which are then stored in the collective memory, reflecting the experiences accumulated in the collective thinking process.

In summary, reorganizing the way collective memory is used appears to be one of the most significant challenges currently facing the application of collective intelligence in policymaking. Given that public policies are typically designed for the medium to long term, the ability to utilize accumulated experiences by participants in open policymaking could significantly enhance such projects' quality. However, many questions remain about how this should be done. These include the methods by which opinions and decisions related to public policies should be recorded in

collective memory. The recognition that the durability of a certain norm, custom, or institution serves as a crucial argument for their preservation, and that the past is the most valid means of legitimizing decisions on public matters, has been acknowledged by political thinkers for a long time. The more extended the period over which knowledge is collected, the better it is, as greater accumulated wisdom bolsters the processes of CI. To maintain the continuity of our collective intelligence with the past, it is essential to leverage thoughts, experiences, and knowledge that extend beyond our single generation. Long-term collective memory facilitates continuous learning about the potential effects of policies, which could lead to improved policymaking. This concept is echoed in George Santayana's (1905, p. 284) famous aphorism, "Those who cannot remember the past are condemned to repeat it." The challenge of recording public policy experiences in digitized collective memory, considering the longevity of their application and verifying their societal impact, is both technological and political.

The impact of collective intelligence on policymaking can be significant. CI can provide policymakers with a wider range of perspectives, knowledge, and ideas to inform their decision-making. This can help to ensure that policies are based on a broad range of input and are more likely to reflect the needs and concerns of a diverse range of individuals and communities. Collective intelligence can help to increase accountability in public policies. By providing a platform for open discussion and collaboration, CI can help to ensure that policymaking processes are more inclusive and transparent. This can help to build trust in the political system and increase public confidence in the policies that are being developed and implemented.

The potential for a transformative breakthrough in public policymaking, powered by the synergy of CI and AI solutions, is a tale not yet fully told. It opens the possibility for a future where the collective wisdom of humanity, amplified by the capabilities of AI, guides us toward more effective governance. The journey toward this future is a labyrinth of complexity and potential, where each turn could reveal new pathways to understanding and innovation. In this evolution, the role of technology is not just as a tool but as a catalyst for a deeper, more profound collaboration among the myriad voices that compose the human experience.

References

Acampora, C. (2018). Agonistic communities: Love, war, and spheres of activity. In J. S. Pearson, & H. Siemens (Eds.), *Conflict and contest in Nietzsche's philosophy*. Bloomsbury Academic.

Ackerman, B. (1981). *Social justice in the liberal state*. Yale University Press.

Aggarwal, I., Woolley, A. W., Chabris, C. F., & Malone, T. W. (2019). The impact of cognitive style diversity on implicit learning in teams. *Frontiers in Psychology*, 10, 112. <https://doi.org/10.3389/fpsyg.2019.00112>

Ahmed, W., Hardey, M. (Maz), Olszowski, R., Baran, S., & Fenton, A. (2023). Did the public-driven campaign to Boycott McDonald's and Coca-Cola impact their share price? Manuscript Submitted for Publication.

Ahmed, W., Vidal-Alaball, J., Downing, J., & López Seguí, F. (2020). COVID-19 and the 5G conspiracy theory: Social network analysis of Twitter data. *Journal of Medical Internet Research*, 22(5), e19458. <https://doi.org/10.2196/19458>

Aleguas, S. (2023) The fake arrest of Donald Trump: A deepfake odyssey. <https://levelup.gitconnected.com/the-fake-arrest-of-donald-trump-a-deepfake-odyssey-db3a6c17eba6>

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2), 211–236.

Anastasio, T. J. (2022). Deriving testable hypotheses through an analogy between individual and collective memory. In S. M. O’Mara (Ed.), *Progress in brain research* (Vol. 274, pp. 31–70). Elsevier.

Arendt, H. (1968). Truth and politics. In H. Arendt (Ed.), *Between past and future: Eight exercises in political thought* (pp. 227–264). Viking Press.

Arendt, H. (1998). *The human condition*. The University of Chicago Press.

Ash, T. G. (2017). Pessimism of the intellect, optimism of the will. Timothy Garton Ash on the future of free speech around the world. <https://www.thebritishacademy.ac.uk/blog/pessimism-intellect-optimism-will-timothy-garton-ash-future-free-speech-around-world/>

Bachleitner, K. (2022). Collective memory and the social creation of identities: Linking the past with the present and future. In S. M. O’Mara (Ed.), *Progress in brain research* (Vol. 274). Elsevier.

Bahl, V. (2023). Photo of elderly French protester being ‘beaten’ by police is AI-generated. <https://www.france24.com/en/tv-shows/truth-or-fake/20230330-france-protests-old-man-beaten-by-police-likely-ai-generated>

Barber, B. (2011). Calling all liberals: It’s time to fight. *The Nation*. <https://www.thenation.com/article/archive/calling-all-liberals-its-time-fight/>

Bartov, E., Faurel, L., & Mohanram, P. S. (2018). Can Twitter help predict firm-level earnings and stock returns? *The Accounting Review*, 93(3), 25–57.

Benhabib, S. (1992). Models of public space: Hannah Arendt, the liberal tradition, and Jurgen Habermas. In C. Calhoun (Ed.), *Habermas and the public sphere*. The MIT Press.

Benkler, Y. (2006). *The wealth of networks: How social production transforms markets and freedom*. Yale University Press.

Bennett, W. L. (2012). The personalization of politics: Political identity, social media, and changing patterns of participation. *The Annals of the American Academy of Political and Social Science*, 644, 20–39.

Berger, B. (2011). Political engagement as intrinsic good: Arendt and company. In *Attention deficit democracy: The paradox of civic engagement*. Princeton University Press. <https://doi.org/10.23943/princeton/9780691144689.003.0003>

Bernholz, L., Landemore, H., & Reich, R. (2021). Introduction to: Digital technology and democratic theory. In L. Bernholz, H. Landemore, & R. Reich (Eds.), *Digital technology and democratic theory*. The University of Chicago Press.

Bessi, A., Zollo, F., Vicario, M. D., Puliga, M., Scala, A., Caldarelli, G., Uzzi, B., & Quattrociocchi, W. (2016). Users polarization on Facebook and YouTube. *PLoS ONE*, 11(8), e0159641.

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1), 1–8.

Bond, R., Fariss, C., Jones, J., et al. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature*, 489, 295–298. <https://doi.org/10.1038/nature11421>

Budzynska, K., & Villata, S., et al. (2018). Processing natural language argumentation. In P. Baroni (Ed.), *Handbook of formal argumentation* (pp. 577–627). College Publications.

Capella, J. N., Zweng, J., & Price, V. (2017). Collective intelligence: The wisdom and foolishness of deliberating groups. In K. Kenski, & K. H. Jamieson (Eds.), *The Oxford handbook of political communication* (1st ed.). Oxford University Press.

Carmichael, F., & Goodman, J. (2020). Vaccine rumours debunked: Microchips, ‘altered DNA’ and more. *BBC*. <https://www.bbc.com/news/54893437>

Castells, M. (2012). *Networks of outrage and hope – social movements in the Internet age*. Wiley.

Catalyst Project—FP7. (2015). <https://catalyst-fp7.idea.kmi.open.ac.uk/>

Clauzet, A., Newman, M. E. J., & Moore, C. (2004). Finding community structure in very large networks. *Physical Review E*, 70, 066111. <https://doi.org/10.1103/PhysRevE.70.066111>

Chloupkova, J., Svendsen, G. L. H., & Svendsen, G. T. (2003). Building and destroying social capital: The case of cooperative movements in Denmark and Poland. *Agriculture and Human Values*, 20, 241–252.

Cohen, J. (2003). Deliberation and democratic legitimacy. In D. Matravers, & J. E. Pike (Eds.), *Debates in contemporary political philosophy: An anthology*. Routledge.

Cohen, J., & Fung, A. (2021). Democracy and the digital public sphere. In L. Bernholz, H. Landemore, & R. Reich (Eds.), *Digital technology and democratic theory*. The University of Chicago Press.

Coleman, S., & Blumler, J. G. (2009). *The internet and democratic citizenship: Theory, practice and policy*. Cambridge University Press.

Collins, B., Marichal, J., & Neve, R. (2020). The social media commons: Public sphere, agonism, and algorithmic obligation. *Journal of Information Technology & Politics*, 17(4), 409–425. <https://doi.org/10.1080/19331681.2020.1742266>

Connolly, W. E. (1991). *Identity/Difference: Democratic negotiations of political paradox*. Cornell University Press.

Connolly, W. E. (2004). Response: Realizing agonistic respect. *The Journal of the American Academy of Religion*, 72(2), 507–513.

Connolly, W. E. (2005). Pluralism. *Duke University Press*. <https://doi.org/10.1215/9780822387084>

Crawford, K. (2016). Can an algorithm be agonistic? Ten scenes from life in calculated publics. *Science, Technology, & Human Values*, 41(1), 77–92.

Cukier, K., Mayer-Schönberger, V., & de Véricourt, F. (2021). *Framers: Human advantage in an age of technology and turmoil*. Dutton.

Czerwiński, K. (2014). Przemiany obywatelstwa we współczesnych demokracjach – wyzwania dla edukacji. *Podstawy Edukacji*, 7, 39–59.

Dagger, R. (2002). Republican citizenship. In E. F. Isin & B. S. Turner (Eds.), *Handbook of citizenship studies* (pp. 145–157). Sage Publications.

Dagger, R. (2006). Neo-republicanism and the civic economy. *Politics, Philosophy & Economics*, 5(2), 151–173. <https://doi.org/10.1177/1470594X06064219>

Dahlberg, L. (2007a). The Internet, deliberative democracy, and power: Radicalizing the public sphere. *International Journal of Media & Cultural Politics*, 3(1), 47–64.

Dahlberg, L. (2007b). The internet and discursive exclusion: From Deliberative to agonistic public sphere theory. In L. Dahlberg, E. Siapera (Eds.), *Radical democracy and the internet*. Palgrave Macmillan. https://doi.org/10.1057/9780230592469_8

Dahlgren, P. (2005). The internet, public spheres, and political communication: Dispersion and deliberation. *Political Communication*, 22(2), 147–162. <https://doi.org/10.1080/10584600590933160>

Dahlgren, P. (2018). Public sphere participation online: The ambiguities of affect. *Les Enjeux De L'information Et De La Communication*, 19(1), 5–20. <https://doi.org/10.3917/enic.024.0005>

Dorsey, J. (2018). Testimony to united states house committee on energy and commerce. <https://www.youtube.com/watch?v=41P9cbaWiBc>

Elster, J. (2012). The optimal design of a constituent assembly. In H. Landemore & J. Elster (Eds.), *Collective wisdom: Principles and mechanisms* (pp. 148–172). Cambridge University Press.

Farrell, H., & Schwartzberg, M. (2021). The democratic consequences of the new public sphere. In L. Bernholz, H. Landemore, & R. Reich (Eds.), *Digital technology and democratic theory*. The University of Chicago Press

Fossen, T. (2014). The grammar of political obligation. *Politics, Philosophy & Economics*, 13(3), 215–236. <https://doi.org/10.1177/1470594X13496072>

Fuchs, C. (2023). *Digital democracy and the digital public sphere*. Routledge.

Fung, A. Y. H., & Kedl, K. D. (2000). Representative publics, political discourses and the Internet: A case study of a degenerated public sphere in a Chinese online community. *World Communication*, 29(4), 69–84.

Gimmler, A. (2001). Deliberative democracy, the public sphere and the internet. *Philosophy & Social Criticism*, 27(4), 21–39. <https://doi.org/10.1177/019145370102700402>

Grynbaum, M. M. (2022). Elon Musk, Matt Taibbi, and a very modern media maelstrom. *The New York Times*, December 4, 2022. <https://www.nytimes.com/2022/12/04/business/media/elon-musk-twitter-matt-taibbi.html>

Habermas, J. (1991). *The structural transformation of the public sphere*. MIT Press.

Habermas, J. (2006). Political communication in media society: does democracy still enjoy an epistemic dimension? The impact of normative theory on empirical research. *Communication Theory*, 16(4), 411–426.

Halbwachs, M. (1992). The reconstruction of the past (L.A. Coser, Trans.). In L. A. Coser (Ed.), *On collective memory* (pp. 46–52). University of Chicago Press (Original work published 1950). <https://doi.org/10.7208/chicago/9780226774497.001.0001>

Hansen, D., Shneiderman, B., Smith, M. A., & Himelboim, I. (2020). *Analyzing social media networks with NodeXL: Insights from a connected world*. Morgan Kaufman Publishers.

Hawkins, J. (2021). *A thousand brains: A new theory of intelligence*. Basic Books.

Hawkins, J. (2023). Q & A with Jeff Hawkins on ChatGPT, the Brain, and the Future of AI. <https://medium.com/@Numenta/qa-jeff-hawkins-662107ebaff4>

Hayek, F. A. (1946). *Individualism: True and false* (The twelfth finlay lecture, delivered at University College, Dublin, on December 17, 1945). Hodges, Figgis & Co., Ltd.; B. H. Blackwell Ltd

Heylighen, F. (1999). Collective intelligence and its implementation on the web: algorithms to develop a collective mental map. *Computational and Mathematical Organization Theory*, 5(3), 253–280.

Hill, K. (2014). Ex-Facebook data scientist: Every Facebook user is part of an experiment at some point. *Forbes*. <https://www.forbes.com/sites/kashmirhill/2014/07/07/ex-facebook-data-scientist-every-facebook-user-is-part-of-an-experiment-at-some-point/?sh=6d421118538>

Honig, B. (1993). *Political theory and the displacement of politics*. Cornell University Press.

Howard, M. M. (2003). *The weakness of civil society in post-communist Europe*. Cambridge University Press.

Jacquet, C., & Thinyane, M. (2023). Mobilizing collective intelligence and diversity toward sustainable development goals: From global innovation labs to collective intelligence assemblies for sustainable development. In S. Boucher, C. A. Hallin, & L. Paulson (Eds.), *The Routledge handbook of collective intelligence for democracy and governance*. Taylor & Francis Group.

Janssen, D., & Kies, R. (2005). Online forums and deliberative democracy. *Acta Politica*, 40(3), 384–392.

Johnson, N. F. (2021). The information noise problem in social media. *Journal of Social Media Studies*, 14(3), 45–58.

Klein, M. (2015). *The catalyst deliberation analytics server*. MIT Working Paper. <https://doi.org/10.2139/ssrn.2962524>

King, S. F. (2007). Citizens as customers: Exploring the future of CRM in UK local government. *Government Information Quarterly*, 24(1), 47–63.

Kshetri, N., & Voas, J. (2017). The economics of “Fake News.” *IT Professional*, 6, 8–12.

Landemore, H. (2012). *Democratic reason: Politics, collective intelligence, and the rule of the many*. Oxford University Press.

Landemore, H. (2021). Open democracy and digital technologies. In L. Bernholz, H. Landemore & R. Reich (Eds.), *Digital technology and democratic theory* (pp. 62–89). University of Chicago Press. <https://doi.org/10.7208/9780226748603-003>

Laclau, E., & Mouffe, C. (2001). *Hegemony and socialist strategy: Towards a radical democratic politics*. Verso.

Lanier, J. (2006). Digital Maoism: The hazards of the new online collectivism. *Edge*. https://www.edge.org/conversation/jaron_lanier-digital-maoism-the-hazards-of-the-new-online-collectivism

Lanier, J. (2023). There is no AI. *The New Yorker*. <https://www.newyorker.com/science/annals-of-artificial-intelligence/there-is-no-ai>

Le, Q. V., Mikolov, T., & Sutskever, I. (2020). Information noise in social media: Definitions, measurements, and challenges. *Social Network Analysis and Mining*, 10, 35.

Lee, J. (2020). Were pallets of bricks strategically placed at US Protest Sites? <https://www.snopes.com/fact-check/pallets-of-bricks-protest-sites/>

Lessig, L. (1999). *Code and other laws of cyberspace*. Basic Books.

Liston, V., Harris, C., & O'Toole, M. (2013). Bridging normative democratic theory and internet technologies: A proposal for scaling citizen policy deliberations. *Policy & Internet*, 5, 462–485. <https://doi.org/10.1002/1944-2866.POI349>

Lippi, M., & Torroni, P. (2016). Argumentation mining: State of the art and emerging trends. *ACM Transactions on Internet Technology*, 16(2), Article 10. <https://doi.org/10.1145/2850417>

Liu, H. K. (2021). Crowdsourcing: Citizens as coproducers of public services. *Policy and Internet*, 13, 315–331. <https://doi.org/10.1002/poi3.249>

Lockie, S. (2017). Post-truth politics and the social sciences. *Environmental Sociology*, 3(1), 1–5. <https://doi.org/10.1080/23251042.2016.1273444>

Lowndes, V., & Paxton, M. (2018). Can agonism be institutionalised? Can institutions be agonised? Prospects for democratic design. *British Journal of Politics and International Relations*, 20, 693–710.

Mansbridge, J., Bohman, J., Chambers, S., Estlund, D., Føllesdal, A., Fung, A., Lafont, C., Manin, B., & Martí, J. L. (2010). The place of self-interest and the role of power in deliberative democracy. *The Journal of Political Philosophy*, 18(1), 64–100. <https://doi.org/10.1111/j.1467-9760.2009.00344.x>

Marchant de Abreu, C. (2022). Debunking Russian claims that Bucha killings are staged. *France 24*. <https://www.france24.com/en/tv-shows/truth-or-fake/20220404-debunking-russian-claims-that-bucha-killings-are-staged>

Margetts, H., John, P., Hale, S., & Yasseri, T. (2016). *Political turbulence: How social media shape collective action*. Princeton University Press.

Marichal, J., & Neve, R. (2020). Antagonistic bias: Developing a typology of agonistic talk on Twitter using gun control networks. *Online Information Review*, 44(2), 343–363. <https://doi.org/10.1108/OIR-11-2018-0338>

Mills, S. (2020). #DeleteFacebook: From popular protest to a new model of platform capitalism? *New Political Economy*, 26(5), 851–868. <https://doi.org/10.1080/13563467.2020.1858777>

Mittal, A., & Goel, A. (2011). Stock prediction using Twitter sentiment analysis. <http://cs229.stanford.edu/proj2011/GoelMittal-StockMarketPredictionUsingTwitterSentimentAnalysis.pdf>

Moreno, J. L. (1934). *Who shall survive?: A new approach to the problem of human interrelations*. Nervous and Mental Disease Publishing Co.

Mouffe, C. (2012). An agonistic approach to the future of Europe. *New Literary History*, 43(4), 629–640. <https://doi.org/10.1353/nlh.2012.0038>

Mouffe, C. (2013). *Agonistics: Thinking the world politically*. Verso Books.

Olszowski, R. (2021). Combating fake news with the use of collective intelligence in hybrid system. *Academia Letters*, Article 1984. <https://doi.org/10.20935/AL1984>

Olszowski, R., Zabdyr-Jamróz, M., Baran, S., Pięta, P., & Ahmed, W. (2022). A social network analysis of tweets related to mandatory COVID-19 vaccination in Poland. *Vaccines*, 10(750). <https://doi.org/10.3390/vaccines10050750>

Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action*. Cambridge University Press.

Owen, D. (1995). *Nietzsche, politics and modernity*. Sage Publications.

Pariser, E. (2011). *The filter bubble: What the internet is hiding from you*. Penguin Press.

Pasquale, F. (2015). *The black box society*. Harvard University Press.

Paxton, M. (2020). *Agonistic democracy: Rethinking political institutions in pluralist times*. Routledge.

Peisert A., & Stachura K. (2011). Demokracja deliberacyjna a praktyka społeczna. [in:] red. Anna Olech Partycypacja publiczna. O uczestnictwie obywateli w życiu wspólnoty lokalnej. Fundacja Instytut Spraw Publicznych, Warszawa.

Pennycook, G., & Rand, D. G. (2021). The psychology of fake news. *Trends in Cognitive Sciences*, 25(5), 388–402. <https://doi.org/10.1016/j.tics.2021.02.007>

Picchi, A. (2022). Twitter Files: What they are and why they matter. *CBS News*. <https://www.cbsnews.com/news/twitter-files-matt-taibbi-bari-weiss-michael-shellenberger-elon-musk/>

Pietrzyk-Reeves, D. (2012). Idea społeczeństwa obywatelskiego. Współczesna debata i jej źródła, Wydawnictwo Uniwersytetu Wrocławskiego.

Projekt ustawy “STOP segregacji sanitarnej” (2021). [https://orka.sejm.gov.pl/Druki9ka.nsf/Projekty/9-020-568-2021/\\$file/9-020-568-2021.pdf](https://orka.sejm.gov.pl/Druki9ka.nsf/Projekty/9-020-568-2021/$file/9-020-568-2021.pdf)

Propper, D. (2022). Elon Musk teases details of Twitter suppression of Post’s Hunter Biden scoop ‘soon’. *New York Post*. <https://nypost.com/2022/11/28/musk-says-files-on-twitters-free-speech-suppression-to-be-published/>

Ranco, G., Aleksovski, D., Caldarelli, G., Grčar, M., & Mozetič, I. (2015). The effects of Twitter sentiment on stock price returns. *PLoS ONE*, 10(9), e0138441.

Raymond, E. S. (2001). *The Cathedral and the Bazaar: Musings on Linux and Open Source by an accidental revolutionary*. O'Reilly Media.

Rawls, J. (1999). *A theory of justice*. The Belknap Press of Harvard University Press.

Rawls, J. (2001). *Justice as fairness: A restatement*. Harvard University Press.

Sampedro, V., & Martínez Avidad, M. (2018). The digital public sphere: An alternative and counterhegemonic space? The case of Spain. *International Journal of Communication*, 12, 23–44.

Santayana, G. (1905). *The life of reason*. C. Scribner's Sons.

Sarewitz, D. (2000). Science and environmental policy: An excess of objectivity. In R. Frodeman & V. R. Baker (Eds.), *Earth matters: The earth sciences, philosophy, and the claims of community* (pp. 79–98). Prentice-Hall.

Segura-Tinoco, A., Holgado-Sánchez, A., Cantador, I., Cortés-Cediel, M. E., & Rodríguez Bolívar, M. P. (2022). A conversational agent for argument-driven e-participation. In *The 23rd Annual International Conference on Digital Government Research (dg.o 2022)* (pp. 191–205). Association for Computing Machinery. <https://doi.org/10.1145/3543434.3543447>

Shabani, S., & Sokhn, M. (2018). Hybrid machine-crowd approach for fake news detection. In *2018 IEEE 4th International Conference on Collaboration and Internet Computing (CIC)*. Philadelphia, PA, USA: IEEE. (pp. 299–306) <https://doi.org/10.1109/CIC.2018.00048>

Siemens, H. W. (2021). *Agonal perspectives on Nietzsche's philosophy of critical transvaluation*. De Gruyter.

Skaržauskienė, A. (2015). *Social technologies and collective intelligence*. Mykolas Romeris University.

Steenbergen, M., Bächtiger, A., Spörndli, M., et al. (2003). Measuring political deliberation: A discourse quality index. *Comparative European Politics*, 1, 21–48. <https://doi.org/10.1057/palgrave.cep.6110002>

Somers, M. R. (1993). Citizenship and the place of the public sphere: Law, community, and the political culture in the transition to democracy, *American Sociological Review*, 58(5).

Spada, P., & Paulson, L. (2023). Measuring the effect of collective intelligence processes that leverage participation and deliberation. In S. Boucher, C. A. Hallin, & L. Paulson (Eds.), *The Routledge handbook of collective intelligence for democracy and governance*. Taylor & Francis Group.

Stromer-Galley, J. (2007). Measuring deliberation content: A coding scheme. *Journal of Public Deliberation* 3, 2(1).

Sunstein, C. (1988). Beyond the republican revival. *Yale Law Journal*, 97, 1539–1589.

Sunstein, C. (2017). *#Republic: Divided democracy in the age of social media*. Princeton University Press.

Tomasello, M. (1999). *The cultural origins of human cognition*. Harvard University Press.

Tuncel, Y. (2023). Nietzsche on conflict and Agon. *Nietzsche-Studien*, 52, 381–390. <https://doi.org/10.1515/nietstu-2022-0036>

Turchin, P. (2015). *Ultrasociety: How 10,000 years of war made humans the greatest cooperators on earth*. Beresta Books.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211, 453–458.

Villa, D. R. (1992). Beyond good and evil: Arendt, Nietzsche, and the aestheticization of political action. *Political Theory*, 20, 274–308.

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151.

Wakita, K., & Tsurumi, Y. (2007). Finding community structure in mega-scale social networks. *Proceedings of the 16th International Conference on World Wide Web* (pp. 1275–1276).

Wallace-Wells, B. (2023). What the Twitter files reveal about free speech and social media. *The New Yorker*. <https://www.newyorker.com/news/the-political-scene/what-the-twitter-files-reveal-about-free-speech-and-social-media>

Wang, Y., McKee, M., Torbica, A., & Stuckler, D. (2021). Systematic literature review on the spread of health-related misinformation on social media. *Social Science & Medicine*, 112790.

White, D. R., & Borgatti, S. P. (1994). Betweenness centrality measures for directed graphs. *Social Networks*, 16, 335–346. [https://doi.org/10.1016/0378-8733\(94\)90015-9](https://doi.org/10.1016/0378-8733(94)90015-9)

Wickelmaier, F. (2003). *An introduction to MDS*. Institut for Elektroniske Systemer. Afdeling for Kommunikationsteknologi. Rapport No. R00-6003.

Wilhelm, A. G. (2000). *Democracy in the digital age: Challenges to political life in cyberspace*. Routledge.

Wylie, C. (2019). *Mindf*ck: Cambridge analytica and the plot to break America*. Random House.

Zhang, X., Carpenter, D., & Ko, M. (2022). The role of information noise in social media: perspectives from user behavior and platform design. *Journal of Information Science*, 48(1), 20–35.

Zhou, X., & Zafarani, R. (2020). A survey of fake news: fundamental theories, detection methods, and opportunities. *ACM Computing Surveys*, 53(5), 109.

Zittrain, J. (2014). Engineering an Election. Digital gerrymandering poses a threat to democracy. *Harvard Law Review Forum*, 127(8).

Zuboff, S. (2019). The age of surveillance capitalism. *Public Affairs*.