TILBURG LAW REVIEW
Journal of International and European Law

Learning from the Ethics of AI – A Research Proposal on Soft Law and Ethics of AI

RESEARCH ARTICLE

JOHAN ROCHEL

ubiquity press

ABSTRACT

This contribution outlines a research proposal combining ethical guidelines on AI and a law-as-data approach. Building upon the definitions of soft law discussed in legal scholarship, it proposes a way of structuring the regulatory landscape on AI and of addressing the question of what is included in the "soft law of AI" today. By adopting a building-blocks approach (combining distinct definitional components of soft law), the paper shows that the state of current soft law on AI depends on which position on international law one defends. Concretely, the paper firstly offers a complete codebook for identifying the different types of soft law. Secondly, it applies this codebook as a proof-of-concept for the research proposal by analyzing 40+ ethical guidelines and by clustering preliminary results according to the actor enacting the guidelines and the legally relevant effects they could deploy. Four paradigmatic types of soft law emerge: statist and international organization soft law, process-oriented soft law, expertise-oriented soft law, and de facto relevant standards soft law. These results illustrate the contributions which are to be expected from a law-as-data research proposal.

CORRESPONDING AUTHOR:

Johan Rochel

Postdoc researcher, EPFL, Lausanne, Switzerland

johan.rochel@epfl.ch

KEYWORDS:

International law; Soft Law; Legal Theory; Technology; AI; Law as Data

TO CITE THIS ARTICLE:

Rochel, Johan, 'Learning from the Ethics of AI – A Research Proposal on Soft Law and Ethics of AI' (2022) 27(1) *Tilburg Law Review* pp. 37–59. DOI: https://doi.org/10.5334/tilr.297

DOI: 10.5334/tilr.297

1 INTRODUCTION

The international regulation of AI is a complex ensemble of actors and norms. Different types of norm produced by different types of actor attempt to regulate AI-powered changes. Especially interesting in the context of AI is that the international regulation landscape entails a number of "ethical" guidelines produced by a variety of actors. These guidelines are produced by companies, NGOs, scientific institutions, states, or groups of states and all claim to influence (to varying degrees) the way AI, as a set of technologies, is defined and regulated.[1] These guidelines are considered a crucial part of the regulatory framework on AI.[2] They have been the topic of social sciences investigations trying to account for the content and elaboration processes of these numerous documents.[3] An important part of this scholarship looks for common themes and substantive values and principles across these documents.

These guidelines are very often qualified as "ethical guidelines" by the actors themselves which try to distinguish them from binding and potentially enforceable commitments. In the meantime, many such actors describe these ethical guidelines as "soft law". On the one side, this might be a rhetorical attempt to give a legal veneer to one's positioning. Ethical content described as being "soft law" seems to be more compelling. On the other hand, the claim might cover some common ground between the efforts by these different actors to provide normative guidance on AI and the evolution of an (international) law on AI.

This paper takes this potential common ground as a starting point for a research proposal. The leading theme of the paper is to ask whether these ethical guidelines can be considered to be "soft law". As it turns out, the question is not quite "whether it is soft law or not", but rather which understanding of soft law outlines, i.e. which mapping of ethical guidelines. In other words, the primary question is how "legal" soft law is; the answer to this question tells us the story of the current state of soft law on AI.

This paper attempts to create the conditions of a dialogue between ethicists dealing with the ethics of AI, social scientists interested in the content and modes of creation of the ethical guidelines, and legal scholars looking (a bit) skeptically at this strange form of normativity claiming to be legal, but not fully so. Readers interested in the ethics or regulation of AI will learn how legal scholars think about soft law and what it takes for ethical guidelines to get legal traction. It will help them structure their arguments in a way which fits legal categories used to address normativity. Legal scholars will look at long-standing controversies in international legal theory from a different law-as-data perspective. The paper claims to be a modest "research proposal" in combining legal theory and a data-based approach in order to provide legal scholarship with greater exposure to the digital humanities, and to strengthen cooperation within communities of researchers in law, philosophy and digital humanities.[4]

---

1 See the overview of these guidelines prepared by OECD AI Policy Observatory https://oecd.ai/dashboards?selectedTab=countries and by Algorithm Watch https://inventory.algorithmwatch.org/.

2 For instance, see C. D. Raab, 'Information Privacy: Ethics and Accountability' in C. Brand (ed), *Ethik in den Kulturen – Kulturen in der Ethik: Eine Festschrift für Regina Ammicht Quinn* (Tübingen, Narr Francke Attempto, 2017); Meg Leta Jones, 'The Right to a Human in the Loop: Political Constructions of Computer Automation and Personhood' (2017) 47 Social Studies of Science 216; Alan F. T. Winfield and Marina Jirotka, 'Ethical Governance is Essential to Building Trust in Robotics and Artificial Intelligence Systems' (2018) 376 Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 1; Miriam C. Buiten, 'Towards Intelligent Regulation of Artificial Intelligence' (2019) 10 European Journal of Risk Regulation 41; Pekka Ala-Pietilä and Nathalie A. Smuha, 'A Framework for Global Cooperation on Artificial Intelligence and its Governance' in B. Braunschweig and M. Ghallab (eds), *Reflections of AI for Humanity* (Springer, 2021).

3 For a meta-analysis of these values and principles, J. Whittlestone and others, 'The Role and Limits of Principles in AI Ethics: Towards a Focus on Tensions' (2019) Proceedings of the 2nd AAAI/ACM Conference on AI, Ethics, and Society 195; Yi Zeng, Lu Enmeng and Huangfu Cunqing, 'Linking Artificial Intelligence Principles' (2019) Proceedings of the AAAI Workshop on Artificial Intelligence Safety 1; L. Floridi and J. Cowls, 'A Unified Framework of Five Principles for AI in Society' (2019) 1 Harvard Data Science Review 1; Anna Jobin, Marcello Ienca and Effy Vayena, 'The Global Landscape of AI Ethics Guidelines' (2019) 1 Nature Machine Intelligence 389; Jessica Fjeld, Nele Achten, Hannah Hilligoss, Adam Nagy, and Madhulika Srikumar, *Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI* (Berkman Klein Center for Internet & Society 2020); Marcello Ienca and Effy Vayena, 'AI Ethics Guidelines: European and Global Perspectives' (2020) Report prepared for the Adhoc Committee on Artificial Intelligence Council of Europe 1.

4 For an overview, Stephen Robertson, 'Digital Humanities' in Simon Stern, Maksymilian Del Mar and Bernadette Meyler (eds), *The Oxford Handbook of Law and Humanities* (Oxford, Oxford University Press, 2020).

DOI: 10.5334/tilr.297

This first section sets the scene of the research proposal by providing some key definitions and by outlining the dataset of 41 ethical guidelines. The second section presents the general strategy of the approach defended in this paper to address the hard law/soft law distinction. A complete codebook for the evaluation of the ethical guidelines composing the dataset is presented. The third section presents preliminary results of the tagging of the dataset. These documents and a table summarizing the tagging results are to be found in annex. Four paradigmatic types of soft law emerge: statist and international organization soft law, process-oriented soft law, expertise-oriented soft law, and de facto relevant standards soft law. The last section discusses methodological challenges which have arisen during the evaluation of the dataset and formulates paths for further research.

2 SETTING THE SCENE OF THE INVESTIGATION

This paper aims at clarifying which part of the ethical guidelines AI can be considered to be "soft law". Central to this objective is the distinction between hard law and other types of norms which (arguably) have legal relevance, namely the broad family of "soft law". This first section aims to clarify some important assumptions of the approach proposed here. It also provides important definitions needed for the interdisciplinary dialogue which this contribution calls for. Finally, it presents the main features of the dataset used in this paper.

2.1 IMPORTANT DEFINITIONS

Firstly, international law on AI is addressed as part of the "regulation" of AI.[5] In that context, regulation is often defined along Black's lines as "the organized attempt to manage risk or behaviour in order to achieve a publicly stated objective or set of objectives".[6] International law is broadly assumed to be one of the types of regulation that operationalizes governance framework.[7] As we will see, contrasting the "law" with non-legal norms (e.g. ethics) and non-normative regulations (e.g. design measures) requires engaging in legal-philosophical debates about the nature of (international) law.[8] This paper can indeed be read as an attempt, against the background of long-standing theoretical debates, to map several understandings of the "legality" of a norm.

Secondly, for this paper, "AI" is broadly defined as any computational system producing intelligent behavior, i.e., complex behavior conducive to reaching goals.[9] In this sense, an 'AI system' is used as an umbrella term for a set of diverse technologies, focused mainly on data-based processing. Clearly enough, the definition of "AI system" must be evolutionary and be

5 For instance, see Raab, 'Information Privacy: Ethics and Accountability' in Brand (ed), *Ethik in den Kulturen – Kulturen in der Ethik: Eine Festschrift für Regina Ammicht Quinn* (Narr Francke Attempto 2017); Jones, 'The Right to a Human in the Loop: Political Constructions of Computer Automation and Personhood' (2017) 47 Social Studies of Science 216; Winfield and Jirotka, 'Ethical Governance is Essential to Building Trust in Robotics and Artificial Intelligence Systems' (2018) 376 Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 1; Buiten, 'Towards Intelligent Regulation of Artificial Intelligence' (2019) 10 European Journal of Risk Regulation 41; Ala-Pietilä and Smuha, 'A Framework for Global Cooperation on Artificial Intelligence and its Governance' in Braunschweig and Ghallab (eds), *Reflections of AI for Humanity* (Springer 2021). See also the "Feasibility Study", Ad hoc committee on AI Council of Europe, December 2020, 18 ff.

6 J. Black, 'Learning from Regulatory Disasters' (2014) 24 LSE Legal Studies Working Paper, 2.

7 For examples of this approach, Margarita Robles Carrillo, 'Artificial Intelligence: From Ethics to Law' (2020) 44 Telecommunications Policy 1; Themistoklis Tzimas, *Legal and Ethical Challenges of Artificial Intelligence from an International Law Perspective* (Springer 2021), 97 ff.

8 Lessig's influential account lists four modalities of regulation: law, social norms, market incentives and architectural measures. Lawrence Lessig, 'The New Chicago School' (1998) 27 The Journal of Legal Studies 661; Lawrence Lessig, 'The Law of the Horse: What Cyberlaw Might Teach' (1999) 113 Harvard Law Review 501. On what non-normative design measures imply for the law, see the work by Roger Brownsword, *Law 3.0 : Rules, Regulation and Technology* (Routledge 2020).

9 For this definition, Vincent C. Müller, 'Ethics of Artificial Intelligence and Robotics' (2020) Stanford Encyclopedia of Philosophy, 1.2. AI-based systems can be purely software-based or can be embedded in physical devices. On the challenges of a legal definition of these socio-technical systems, see Matthew U. Scherer, 'Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies ' (2016) 29 Harvard Journal of Law & Technology 353, 359; Thomas Burri, 'International Law and Artificial Intelligence' (2017) 60 German Yearbook of International Law 91; M. Maas, 'International Law Does Not Compute: Artificial Intelligence and The Development, Displacement or Destruction of the Global Legal Order' (2019) 20 Melbourne Journal of International Law 29, 22 ff.

able to include further technological evolutions.[10] Nevertheless, this does not mean that the concept should be amorphous (encompassing for instance any digital technology).

Thirdly, given this definition, it is important to make clear that the relevant framing does not focus on regulating AI per se, but rather the socio-technical changes which AI makes possible. As proposed by Crootof and Aar, the challenge is to identify and to address "tech-created legal uncertainties".[11] In that sense, this paper is not to be understood as an emanation of an approach which considers AI to be a kind of exceptional technology which requires specific legal norms. This trap of "exceptionalism" is well-known in the context of technology law and ethics, i.e. the tendency to over-react to novel technologies, remain domain-specific, and reinvent the wheel at each innovation.[12] In general, the approach outlined *infra* does not deliver a normative argument for the creation of AI-specific international legal norms. It aims at taking a rather descriptive approach based on the current reality of the international regulation of AI. Nevertheless, in attempting to describe the current state of international law on AI – paying attention to making its own methodological presuppositions as explicit as possible – this broadly descriptive approach contributes to its further structuration and emergence and must, as such, be understood as part of a normative endeavor.

2.2 DESCRIPTION OF THE DATASET

As mentioned in the introduction, this paper focuses on the so-called "ethical guidelines" formulated on AI by a variety of different actors. The main objective of this paper is to apply a law-as-data approach to the guidelines, thereby trying to provide a proof-of-concept for this methodological approach.

The list of 41 documents considered in the dataset is to be found in the annex. To achieve the objective stated in the introduction, namely bringing structure into the ethical/soft law claims raised by many AI guidelines, the starting point of the analysis was to look for an existing dataset claiming to gather ethical/soft law norms. For this proof-of-concept, I draw upon the methodology defended by Fjeld et al. in their overview of the AI ethics landscape.[13] To briefly summarize their methodology, they include general documents in their dataset based on three main criteria. Firstly, they exclusively include documents that are normative. They exclude all documents that simply describe how an AI system works or list advantages and disadvantages of an AI system. Secondly, they exclude documents which call for a specific action without specifying normative elements at stake or documents which call for the elaboration of principles without proposing any. Thirdly, they exclude certain early instances of legislation and regulation, mainly because it would make side-by-side comparison with the other documents difficult. This point needs to be considered differently in the perspective of the objective of this research proposal. If the objective is to identify interesting patterns in soft law norms on AI, early (in the sense of incremental) instances of legislation should be taken into account. Indeed, as explained below, it is one of the possible understandings of soft law's legal relevance (incremental legal norm). To address this challenge, normative documents produced by the UN since 2020 have been added to the dataset. Furthermore, newer documents have been added

10 For general overviews on AI and the law, see Grégoire Loiseau and Alexandra Bensamoun (eds), *Droit de l'intelligence artificielle* (LGDJ 2019); Alberto De Franceschi and Reiner Schulze (eds), *Digital Revolution - New Challenges for Law* (C.H.Beck 2019); Harry Surden, 'Artificial Intelligence and Law: An Overview' (2019) 35 Georgia State University Law Review 1305; T. Wischmeyer and T. Rademacher (eds), *Regulating Artificial Intelligence* (Springer 2020); Harry Surden, 'Ethics of AI in Law: Basic Questions ' in Markus Dirk Dubber, Frank Pasquale and Sunit Das (eds), *The Oxford Handbook of Ethics of AI* (Oxford, Oxford University Press, 2020). AI can be seen as part of a broader informational challenge. See Julie E. Cohen, *Between Truth and Power : The Legal Constructions of Informational Capitalism* (Oxford University Press 2019).

11 Rebecca Crootof and BJ Ard, 'Structuring TechLaw' (2021) Harvard Journal of Law & Technology 1, 3.

12 Easterbrook famously referred to the danger of drafting law for the "horse", Frank H. Easterbrook, 'Cyberspace and the Law of the Horse' (1996) University of Chicago Legal Forum 207. For alternative approaches, see Meg Leta Jones, 'Does Technology Drive Law? The Dilemma of Technological Exceptionalism in Cyberlaw' (2018) 101 Journal of Law, Technology and Policy 249; Hin-Yan Liu and others, 'Artificial Intelligence and Legal Disruption: A New Model for Analysis' (2020) 12 Law, Innovation and Technology 205. For references and discussion of examples, see Crootof and Ard, 'Structuring TechLaw' (2021) Harvard Journal of Law & Technology 1, 6 ff.

13 Fjeld, *Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI* (Berkman Klein Center for Internet & Society 2020), 11 ff.

DOI: 10.5334/tilr.297

to the original dataset (composed of documents made public before 2020). Wherever possible, a new version of the documents has been used.

It is important to stress that the general objective of this contribution does have an impact on the creation standards for the dataset. By contrast to Fjeld et al., who wanted to map an emerging consensus on values and principles, here there is no strong requirement to claim exhaustivity in the list of documents identified. The analysis presented below could be expanded by integrating additional documents. The development of the dataset can be done in an aggregative logic. The mapping proposed and the preliminary conclusions presented do not focus primarily on the content of the norms, but rather on their conceptualization with respect to the soft/hard law distinction. This general objective makes the approach less dependent upon exhaustivity.

One last caveat on the definition of the dataset is important. Each guideline identified can be tagged as a whole, but also, depending on its richness and complexity, can be considered as a set of data points. In this case, the norms entailed by a guideline (i.e. a document) should each be considered a different data point to be tagged. To take a concrete example, the EU "Ethics Guidelines for Trustworthy AI" is a 41-page document. It includes an important number of norms proposed as part of a body of norms on AI. Many of the data points included in this document may have the same characteristics, though this need not always be the case. For instance, the information about the enactor of the norm will remain the same (the special experts' group mandated by the EU Commission), but the feature of a specific norm (e.g. the way it is promoted, presented or justified) could change for each data point. The methodology outlined below could work with a minimal definition of norm as being equated with a prescription on the conception/development/use of an AI system. For a detailed analysis, every single prescription of a large document could be considered a norm (i.e. a data-point to be accounted for). For instance, each distinct component of trustworthiness mentioned in the EU Guidelines could, in a prescriptive sense, be interpreted as a single norm prescribing a specific conception/development/use of an AI system. This caveat is important for further research which would go into the details of each document of the dataset and which could cooperate extensively with social sciences. For this proof-of-concept contribution, the document level shall prove sufficient.

3 WHERE DOES SOFT LAW BEGIN?

Based on these definitions and clarifications, this section outlines a methodological approach for structuring the soft law norms identified. This approach should be understood as a proposal for further research to be conducted in this field and as a contribution to methodological discussions in legal and ethical scholarship. This section 1) locates the law-as-data approach chosen in the legal literature, 2) presents the parameters used to evaluate the dataset (in form of a codebook) and 3) discusses methods to apply these parameters to the 41 documents of the dataset.

3.1 A LAW-AS-DATA APPROACH

A law-as-data approach can be considered a sub-part of empirically minded digital humanities applied to the law. In that sense, it refers to any approach that treats law as data, structures legal material as data, and then analyses these data to retrieve information (about e.g. content of the norms, law-making processes, interpretation, adjudication and enforcement).[14] In the context of public international law, it is important to acknowledge that these approaches are challenging to use. In contrast to tax law for instance, public international law does not offer large and structured datasets. This explains why current international law projects tend to

14 UrsKa Šadl and Henrik Palmer Olsen, 'Can Quantitative Methods Complement Doctrinal Legal Studies? Using Citation Network and Corpus Linguistic Analysis to Understand International Courts' (2017) 30 Leiden Journal of International Law 327; Michael A. Livermore and Daniel N. Rockmore (eds), *Law as Data: Computation, Text, and the Future of Legal Analysis* (Santa Fe Institute Press 2019); Wolfgang Alschner, 'The Computational Analysis of International Law' in Rossana Deplano and Nicholas Tsagourias (eds), *Research Methods in International Law: A Handbook* (Elgar, 2020); Jens Frankenreiter and Michael A. Livermore, 'Computational Methods in Legal Analysis' (2020) 16 Annual Review of Law and Social Science 39.

DOI: 10.5334/tilr.297

focus on questions with existing structured datasets, such as case-law analysis[15] or free-trade agreements' analysis.[16]

Law-as-data approaches start with data-based methods for structuring the body of legal norms according to pre-defined categories. The objective here is a structured dataset whose creation forces scholars to explicitly present and justify the categorizations and parameters they use to assess a body of norms. This is why law-as-data approaches require legal-philosophical scholarship. They need well-constructed conceptual delimitations which in turn need to be embedded into a conceptual and normative framework.

Based on a structured dataset, it is possible to try out diverse analytical tools. Short of large structured datasets, it is hardly interesting with respect to the ratio investment/return to use Natural language processing (NLP)[17] and machine learning[18] tools. The claim is not that this is impossible, but rather that the promised return is not worthwhile. This might be different with network analysis tools which can be used even on a small dataset.[19] These network tools are used to account for putative normative relations within the dataset. Tools such as node measures (identifying important points of convergence such as sources or norms), and network measures (quantification of network properties, e.g., as structuring important parts of the dataset around specific norms/sources) are interesting in looking for patterns within the dataset.[20]

The use of these approaches in (international) law has given rise to a major debate on the methodological challenges faced by legal scholarship.[21] Firstly, these tools are seen as being in tension with theoretical and doctrinal competences.[22] The approach defended here clearly needs to be firmly combined with legal theory. Secondly, broader criticisms bear upon the desire to model and predict human interactions using computational tools and, more fundamentally, use mathematical-statistical approaches to provide a unified grammar for complex realities.[23] This criticism echoes the broader debate on empirical approaches to the law.[24] The approach

15 Niccolò Ridi, 'The Shape and Structure of the 'Usable Past': An Empirical Analysis of the Use of Precedent in International Adjudication' (2019) 10 Journal of International Dispute Settlement 200; Mattias Derlén and Johan Lindholm, 'Is it Good Law? Network Analysis and the CJEU's Internal Market Jurisprudence' (2017) 20 Journal of International Economic Law 257.

16 Todd Allee, Manfred Elsig and Andrew Lugg, 'The Ties between the World Trade Organization and Preferential Trade Agreements: A Textual Analysis' (2017) 20 Journal of International Economic Law 333; Joost Pauwelyn and Wolfgang Alschner, 'Forget about the WTO: The Network of Relations between PTAs and double PTAs' in Andreas Dür and Manfred Elsig (eds), *Trade Cooperation: The Purpose, Design and Effects of Preferential Trade Agreements: Volume undefined: World Trade Forum* (Cambridge, Cambridge University Press, 2015).

17 Wolfgang Alschner and Damien Charlotin, 'The Growing Complexity of the International Court of Justice's Self-Citation Network' (2018) 29 European Journal of International Law 83; Marc L Busch and Krzysztof J Pelc, 'Words Matter: How WTO Rulings Handle Controversy' (2019) 63 International Studies Quarterly 464.

18 Rohan Nanda and others, 'Unsupervised and Supervised Text Similarity Systems for Automated Identification of National Implementing Measures of European Directives' (2019) 27 Artificial Intelligence and Law 199; Masha Medvedeva, Michel Vols and Martijn Wieling, 'Using Machine Learning to Predict Decisions of the European Court of Human Rights' (2020) 28 Artificial Intelligence and Law 237. For references, see Frankenreiter and Livermore, 'Computational Methods in Legal Analysis' (2020) 16 Annual Review of Law and Social Science 39, 45.

19 For an example in international environmental law, see the work on network analysis by Rakhyun E. Kim, 'Is Global Governance Fragmented, Polycentric, or Complex? The State of the Art of the Network Approach' (2020) 22 International Studies Review 903.

20 On this network analysis approach, see R. Whalen, 'Legal Networks: The Promises and Challenges of Legal Network Analysis' (2016) 2016 Michigan State Law Review 539.

21 Tilmann Altwicker, 'International Legal Scholarship and the Challenge of Digitalization' (2019) 18 Chinese Journal of International Law 217, 236 ff. For the argument in favor of methodological diversity in European public law research, see Emanuel V. Towfigh, 'Empirical Arguments in Public Law Doctrine: Should Empirical Legal Studies Make a "Doctrinal Turn"?' (2014) 12 International Journal of Constitutional Law 670.

22 For a similar position, see Hanoch Dagan, Roy Kreitner and Tamar Kricheli-Katz, 'Legal Theory for Legal Empiricists' (2018) 43 Law & Social Inquiry 292.

23 See Mireille Hildebrandt, 'Law as Information in the Era of Data-Driven Agency' (2016) 79 The Modern Law Review 1. Johns goes one step further in arguing that data-mining tools are not just a tool for governance; they are a new type of governance. Fleur Johns, 'Data Mining as Global Governance' in Roger Brownsword, Eloise Scotford and Karen Yeung (eds), *The Oxford Handbook of Law, Regulation and Technology* (Oxford, Oxford University Press, 2017), 777.

24 Gregory Shaffer and Tom Ginsburg, 'The Empirical Turn in International Legal Scholarship' (2012) 106 American Journal of International Law 1; Jakob V. H. Holtermann and Mikael Rask Madsen, 'Toleration, Synthesis or Replacement? The 'Empirical Turn' and its Consequences for the Science of International Law' (2016) 29 Leiden Journal of International Law 1001. This "empirical turn" also involves the use of experiments inspired by psychology. See Jeffrey L. Dunoff and Mark A. Pollack, 'Experimenting with International Law' (2017) 28 European Journal of International Law 1317.

outlined here fully acknowledges the normative nature of the law. Any empirical approach needs to be conceptualized within a broader and openly normative argument. For example, addressing the numerous ethical guidelines on AI requires conceptual clarity as to the assumed definition of soft law, and hence a determined position on the broader issue of the sources of international law. This conceptual and normative positioning will itself be relevant when it comes to using the results of the analysis, for instance as part of an argument relating to the emergence of new legal norms on AI.

3.2 THREE PARAMETERS TO ACCOUNT FOR THE DEFINITIONS OF SOFT LAW

When looking at the 41 documents composing the dataset, the main question which practitioners will ask is: which ones are soft law and which are not? The hypothesis defended here is to address this question using a structured set of parameters capturing the nuances of how soft law is defined. At its core, the distinction between soft and hard law is a discussion of when does a norm become an international legal norm.[25] It relies upon a major point of contention regarding the sources of international law. This topic has been the object of intense scholarship in political science[26] and in international law.[27]

The objective is not to argue for a specific way to understand and justify the distinction between soft and hard law ("the best way to define soft law is X"), but rather to propose categories which are able to account for the main ways in which theorists have distinguished between the two. The objective is to identify the main building blocks used to provide a structured way of comparing different definitions of the soft/hard law distinction.[28]

After a thorough study of the relevant literature, I propose three key parameters to account for the different definitions of the soft/hard law distinction: the binding nature of the norms, the actors enacting the norms, and the legally relevant effects of the norms.[29]

The first parameter looks at the distinction between legally binding and non-binding norms. Its main objective is to exclude norms which clearly qualify as hard law because they are legally binding and to focus the work on the broad family of "soft law". The claim is of course not that hard law norms are irrelevant, quite on the contrary. But their identification and their characterization as legal norms distinct from ethical norms are questions which are more easily settled. International hard law norms are binding because they can be traced back to one of the sources of international law. A norm whose legality relies on one of the sources of international law (as entailed by Art. 38(1) ICJ Statute: Treaty, Custom, General principles) is legally binding.

---

25 On this debate, see Jean D'Aspremont and Samantha Besson (eds), *The Oxford Handbook of the Sources of International Law* (Oxford University Press 2017).

26 For references in political sciences, see Kenneth W. Abbott and Duncan Snidal, 'Hard and Soft Law in International Governance' (2000) 54 International Organization 421; Sylvia I. Karlsson-Vinkhuyzen and Antto Vihma, 'Comparing the Legitimacy and Effectiveness of Global Hard and Soft Law: An Analytical Framework' (2009) 3 Regulation & Governance 400.

27 Among a large body of literature, see D. Thürer, 'Soft Law' (2009) Max Planck Encyclopedia of Public International Law; A. T. Guzman and T. L. Meyer, 'International Soft Law' (2010) 2 Journal of Legal Analysis 171; Alan Boyle, 'Soft Law in International Law-Making' in Malcolm N. Shaw (ed), *International Law* (8 edn, Cambridge University Press, 2017). For theoretical contributions, see Jean D'Aspremont and Tanja Aalberts, 'Symposium on Soft Law' (2012) 25 Leiden Journal of International Law 309; Joost Pauwelyn, Ramses A. Wessel and Jan Wouters, *Informal International Lawmaking* (Oxford University Press 2012). With respect to the governance of AI, Ryan Hagemann, Jennifer Huddleston and Adam D. Thierer, 'Soft Law for Hard Problems: The Governance of Emerging Technologies in an Uncertain Future ' (2018) 17 Colorado Technology Law Journal 37, 42; G. Marchant, '"Soft Law" Governance Of Artificial Intelligence' (2019) AI Pulse 1.

28 For a similar methodological choice, see Karlsson-Vinkhuyzen and Vihma, 'Comparing the Legitimacy and Effectiveness of Global Hard and Soft Law: An Analytical Framework' (2009) 3 Regulation & Governance 400, 402 ff.

29 This proposal is especially based on Dinah Shelton, *Commitment and Compliance: The Role of Non-binding Norms in the International Legal System* (Oxford University Press 2000), 292; Thürer, 'Soft Law' (2009) Max Planck Encyclopedia of Public International Law; László Blutman, 'In the Trap of a Legal Metapher: International Soft Law' (2010) 59 The International and Comparative Law Quarterly 605, 607; Dinah Shelton, 'International Law and "Relative Normativity"' in Malcolm Evans (ed), *International Law* (3 edn, Oxford University Press, 2010), 165; D'Aspremont and Aalberts, 'Symposium on Soft Law' (2012) 25 Leiden Journal of International Law 309, 303; Sumudu Atapattu, 'International Environmental Law and Soft Law: A New Direction or a Contradiction?' in Cecilia M. Bailliet (ed), *Non-State Actors, Soft Law and Protective Regimes: From the Margins* (Cambridge, Cambridge University Press, 2012), 207; Boyle, 'Soft Law in International Law-Making' in Shaw (ed), *International Law* (8 edn, Cambridge University Press 2017), 119.

DOI: 10.5334/tilr.297

This parameter can be criticized in two ways. Firstly, this parameter relies upon a binary model of international law as opposed to a continuum model. Binary models tend to see soft law as being conceptually impossible (either soft or law, but not the two simultaneously), while continuum models see the distinction as being gradual (and often use the image of a penumbra).[30] This criticism is indeed valid, but it should be seen in the context of the law-as-data approach. As explained, the first function of this parameter is to provide a way to set apart norms which clearly qualify as hard law. Indeed, for that reason, all the guidelines considered in the dataset are non-binding and, for that reason, qualify in the first place for a discussion on their potential qualification as soft law. Secondly, the apparent simplicity of the parameter binding/non-binding can be claimed to be misleading. The doctrine of sources in international law reveals important theoretical disputes.[31] This is especially the case when identifying international customary law. Again, the reply refers to the function of this first parameter in the context of the law-as-data analysis. When in doubt, it should be assumed that the norm in question is soft law rather than hard law. In other words, there is a default position on soft law.

The second parameter (actors enacting the norm) raises the challenge of how to define which actors can enact soft law. In his entry "soft law" in the Max Planck Encyclopedia of Public International Law, Thürer says the creation of soft law should be limited to States and "other subjects of international law", meaning mainly intergovernmental organizations.[32] For the sake of the present argument, the working hypothesis on the definition of the other subjects of international law should be broad.[33] The objective here is to structure the dataset, not to take a definitive position on which actors have the legitimacy to create which type of soft law. In this sense, it seems consistent to adopt a broad position and to include international organizations and organized groups of individuals such as civil society, business representatives or standardization organizations (e.g. the IEEE).[34]

This parameter can be criticized in two ways. Firstly, it can be argued from a traditional perspective on the subjects of international law that only States and intergovernmental organizations (mainly UN bodies) should have the competence to enact soft law. This second parameter is, in light of that criticism, related to the sources' doctrine assumed for the first parameter. To define the accepted sources has implications on the actors which qualify as hard law-enactors and, in turn, the potential soft law-enactors. However, to recall, this paper takes a traditional conception (incl. interpretation of the sources) as *one* possible way to define soft law. The overall objective is to account for the majority of existing definitions. It is only in a second step, on the basis of a structured dataset, that we could address the normative question of which actors *should* have the competence to enact soft law. Secondly, the definition of the actors which could enact soft law should not be conflated with an account of customary international law. Customs are one source of international law according to Art. 38(1) ICJ Statute. In short, a custom is a general practice accepted as law. It is composed by States' practices (the behaviors of States) and *opinio juris,* i.e. an attitude towards a legal obligation (States act because they consider the norm as law, and not for instance because of strategic or economic reasons). By contrast, soft law is not a source of international law. As we shall see, soft law can be used as indicator of an emerging consensus on how States concretely behave (assuming consistency between the State's commitment and its actual behavior) and on which norms are considered as legal obligations.

30 For a binary view, see Jan Klabbers, 'The Redundancy of Soft Law' (1996) 65 Nordisk Journal of International Law 167. For a continuum view, see Shelton, 'International Law and "Relative Normativity"' in Evans (ed), *International Law* (3 edn, Oxford University Press 2010).

31 Some authors argue that soft law should itself be considered a new source of international law. See F. A. Cárdenas Castañeda, 'A Call for Rethinking the Sources of International Law: Soft Law and the Other Side of the Coin' (2013) 13 Anuario Mexicano de Derecho Internacional 355.

32 Thürer, 'Soft Law' (2009) Max Planck Encyclopedia of Public International Law, § 8.

33 In general, see Malcolm N. Shaw, 'The Subjects of International Law' in Malcolm N. Shaw (ed), *International Law* (8 edn, Cambridge University Press, 2017). On soft law in particular, see Hilary Charlesworth, 'Law-making and Sources' in James Crawford and Martti Koskenniemi (eds), *The Cambridge Companion to International Law* (Cambridge Cambridge University Press, 2012, 198; Thürer, 'Soft Law' (2009) Max Planck Encyclopedia of Public International Law, § 6. The International Law Commission notes that the practices of non-state organizations should only be taken as a subsidiary means of assessment of customs. ILC, "Draft conclusions on the identification of customary international law" (2018).

34 For examples, see Marchant, '"Soft Law" Governance Of Artificial Intelligence' (2019) AI Pulse 1, 5 ff.

DOI: 10.5334/tilr.297

The third parameter is related to the legally relevant effects that soft law norms can deploy. On the one hand, even if it is not legally binding, soft law can create "legal effects". Firstly, this is the case in conjunction with the principle of "good faith" in international law.[35] This principle requires actors not to contradict their own conduct. Soft law chosen by a specific actor is assumed to represent this actor's conduct; it provides evidence of what the actor wants. In that situation, good faith does not transform a soft law norm into a legally binding norm, but it can produce legally relevant expectations for other actors (which assume a general consistency of a State's conduct). A relevant example is the application of good faith to the unilateral declarations made by states.[36] Secondly, as mentioned previously, specific soft law norms can provide evidence of a nascent custom.[37]

On the other hand, soft law norms can produce effects that are not directly legal effects, but which are "legally relevant effects". Firstly, soft law can be an intermediate legal norm found on the development path of an international hard law norm, without having fulfilled all the conditions required for becoming hard law, and perhaps even without having the ambition to do so in the future.[38] Though the norm is not legally binding, actors tend to follow incremental norms as potential forerunners of a hard-law norm.[39] It deploys an aspirational effect through normative pressure on how actors are expected to behave in a specific situation, mainly because this standard of behavior could or has chances to become a binding standard. Secondly, soft law norms can have a legally relevant effect in providing interpretative resources for existing legal norms (both international and domestic law). As Redgwell explains when considering soft law in the context of environmental law, soft law can formulate new general principles which, in turn, can influence the interpretation of existing norms and norms-in-the-making.[40] Thirdly, soft law can prepare the ground for a political discussion on future legal norms. In this sense, the process through which soft law is formulated can be compared to a consultation process that thematizes specific issues, consolidates political opinion, and outlines potential legal paths for addressing these issues.[41]

Taken together for the sake of structuring the dataset, these three parameters are claimed to be able to account for an important number of definitions of the hard/soft law distinction. To recap, these parameters cannot account for the details and richness of each definition. They require a certain form of simplification for the sake of the structuration of the dataset. These general parameters now need to be formulated in form of a codebook used to tag each resource in the dataset.[42]

Codebook: Building Blocks of Soft Law

- Parameter: "Binding status of the norm"
    - Values = ["Legally binding"," Not legally binding"]
        - Corresponding source if binding: Treaty, Custom, General principles

35 Thürer, 'Soft Law' (2009) Max Planck Encyclopedia of Public International Law, § 27.

36 See further Guillaume Futhazar and Anne Peters, 'Good Faith' in Jorge E. Viñuales (ed), *The UN Friendly Relations Declaration at 50: An Assessment of the Fundamental Principles of International Law* (Cambridge, Cambridge University Press, 2020), 198 ff.

37 Charlesworth, 'Law-Making and Sources' in Crawford and Koskenniemi (eds), *The Cambridge Companion to International Law* (Cambridge University Press 2012), 194 ff. Critical, Anthony D'Amato, 'Trashing Customary International Law' (1987) 81 American Journal of International Law 101.

38 For this view, see Catherine Redgwell, 'Sources of International Environmental Law: Formality and Informality in the Dynamic Evolution of International Environmental Law Norms' in Besson Samantha and d'Aspremont Jean (eds), *The Oxford Handbook of the Sources of International Law* (Oxford, Oxford University Press, 2017, 955.

39 Hagemann, Huddleston and Thierer, 'Soft Law for Hard Problems: The Governance of Emerging Technologies in an Uncertain Future ' (2018) 17 Colorado Technology Law Journal 37, 45.

40 Redgwell, 'Sources of International Environmental Law: Formality and Informality in the Dynamic Evolution of International Environmental Law Norms' in Samantha and Jean (eds), *The Oxford Handbook of the Sources of International Law* (Oxford University Press 2017).

41 Atapattu, 'International Environmental Law and Soft Law: A New Direction or a Contradiction?' in Bailliet (ed), *Non-State Actors, Soft Law and Protective Regimes: From the Margins* (Cambridge University Press 2012), 207.

42 For a similar methodological choice, see Tomer Broude, Yoram Z. Haftel and Alexander Thompson, 'The Trans-Pacific Partnership and Regulatory Space: A Comparison of Treaty Texts' (2017) 20 Journal of International Economic Law 391, 395 ff.

DOI: 10.5334/tilr.297

- Parameter: "Creator of the norm"
  - Values = ["States", "International Organizations", "NGO", "Companies", "Research Institutions", "Mix"]
- Parameter: "Special legal relevance of the norm through effects"
  - Values = ["legal effects", "legally relevant effects"]
    - Legal effects: legally relevant commitment taken by a subject of international law (in combination with good faith); evidence of a nascent custom.
    - Legally relevant effects: incremental legal norm, interpretative guidelines, political forum

3.3 ASSESSMENT OF THE THREE PARAMETERS

Assuming that these parameters suitably fulfil the task they are set for, we need to define how to apply these parameters to the 41 guidelines identified. This transition from abstract parameters to measurable elements raises further methodological questions. A perfectly objective assessment is out of reach for any scientific project. What we should aim at is making explicit how the measurement is conceived and how it is done on specific elements.

Firstly, the measurement takes either a yes/no form or a numerical value (between 0 and 1). This transparent and explicit measurement is made to channel disagreements and criticisms on the right level. Three levels of criticisms have to be distinguished: the parameters defined above (the distinctive features of soft law), the measurement elements presented below (is it the best way to measure the parameter?) and the specific instantiation of the measurement (is the document X well evaluated? Why was this numerical value chosen?). It is crucial to keep these three levels of criticism apart. This channeling of disagreements and criticisms is not a weakness of the law-as-data approach, it is rather one of its strengths in contributing to a more specific, empirically informed discussion on the current state of the regulatory landscape on AI.

The two first parameters (binding/non-binding and the creator of the norm) do not raise specific evaluative challenges. The binding/non-binding parameter could indeed lead to disputed issues on the sources of international law. However, as mentioned, the default position if there is doubt on the status of an international norm is set on considering this norm soft law.

To measure the legally relevant effect requires the development of measurement methods. The first category of "legal effects" includes good faith commitments and contributions to the emergence of a custom. The hypothesis is that the two effects should be measured using at least two features:

a) Does the document evaluated represent the position held by the State/organization (or a relevant actor of international law)? This first point is about evaluating the officiality of the document and the capacity of the enactor of the norm to speak in the name of its State/organization. Three values are possible: 1 (strong representation), 0,5 (middle value), 0 (weak representation).
b) Does the language used in the document (incl. reservations) reflect the ambition of the State/organization to make a formal commitment? This second point is about evaluating how affirmative the State/organization is in its position. This complements the information in the first point. A strongly representative position of a State (e.g. by its President) can indeed be formulated in a vague and general way. Three values are possible: 1 (strongly affirmative), 0,5 (middle value), 0 (vague).

In contrast to similar endeavors focusing on the identification of emerging customs, the dimension of consistency of a State's/organization's action across time and distinct thematic commitments is not assessed. This is because this information – though valuable – is not required to structure and map the soft law landscape.

The second category of "legally relevant effects" includes incremental legal norms, interpretative guidelines, and political fora. Firstly, whether something is an incremental legal norm is considered as a yes/no question. It focuses on the theoretical possibility for the document/norm considered to continue its legislative path into hard law. This point should not raise significant controversies as it can be well identified. The political possibility/feasibility is not considered here.

DOI: 10.5334/tilr.297

Secondly, the claim to provide interpretative guidance should be measured using an evaluation of the legitimacy of claimed guidance. Three specifications are proposed: does the actor enacting the norm have a special legitimacy to claim to provide interpretative guidance, namely through expertise, power and/or an inclusive process of norm-creation? The last dimension (norm-creation process) is also used to capture the third element (the legally relevant effect of soft law as a political forum). For each of these three dimensions, three values are possible: 1 (strong expertise/power/inclusive process), 0,5 (middle value), 0 (no expertise/no power/no inclusive process).

Notwithstanding a broader debate on legitimacy in and of international law, the objective is to propose a preliminary way to start distinguishing between different types of norms which claim to provide interpretative guidance. The expertise, the power and the quality of the norm-creating process capture different understandings of the legitimacy which specific actors could use to make their case. Of course, these elements raise important questions about how to define and to measure each specific dimension. Which kind of expertise is relevant, for instance if we have to evaluate ethical guidelines produced by academic research institutions and big tech companies? How to measure the power of an actor, by its population/GDP/revenue for a State or a company? How to define inclusivity in the context of a norm-creation process? These questions cannot be answered in theory here. They have been the object of important scholarship across social sciences. For this proof-of-concept paper, the strategy is to evaluate the dataset and to provide a justification for the evaluation, thereby making explicit which features of the actor/document/process was considered most relevant. Theoretical challenges are not solved by this strategy, but nor are they hidden away.

4 EVALUATION OF THE DATASET

The objective of this section is to exemplify the type of results and further research questions which this research proposal would make possible. This section shows how to apply the parameters presented above to the dataset, and which difficulties need to be addressed in doing so. It 1) presents intermediary results and 2) addresses methodological challenges encountered in the tagging process. This section should be read with the dataset description and the overview table found in the annex.

4.1 INTERMEDIARY RESULTS

This section presents intermediary results generated by the tagging of the dataset. The law-as-data approach brings 4 different paradigmatic types of soft law to light: statist and international organization soft law, process-oriented soft law, expertise-oriented soft law and de facto relevant standards soft law. These results are summarized in the table below.

In order to better understand these 4 types of soft law, we will look at the annotated dataset from two different perspectives. The first perspective focuses on the actors creating soft law, the second on the legally relevant effects of soft law. Eventually, we will discuss the interface between ethics and law as found in the documents of the dataset.

4.1.1 Types of soft law according to the actors

This sub-section presents results according to the perspective of actors creating soft law. The first type of soft law is produced by States. As operationalized through the codebook, States are taken to be relevant when it comes to good-faith-based arguments and to their contribution to the emergence of a custom.

Beyond the general officiality of a document enacted by a State, the analysis has shown the requirement to assess more precisely the degree of officiality. A good example is the mission conferred to the former French Parliament member and mathematician C. Villani by the French prime minister on behalf of President Macron. This document called “For a Meaningful Artificial Intelligence: Toward a French and European Strategy “(2018) was taken as basis of the French AI strategy “AI for humanity” endorsed by Macron in March 2018. It makes no doubt that the content of the document is the normative positioning closest possible to “the” French position on the international regulation of AI. However, this strategy is not a statute or legally

DOI: 10.5334/tilr.297

binding document in a strong sense, neither domestically nor internationally. Despite this non-bindingness, international partners can expect France to behave in a way which is compatible with the commitments described in this document.[43] The position taken by a State (e.g. the announcement of a national strategy on AI) can have the effect to create expectations for international partners which can be approached using the principle of good faith. In this respect, it is interesting to note, following Deeks, that these international partners will also use AI-powered tools to monitor and assess public positioning by States.[44] Today, as 30 years ago, it is the traditional task of a diplomat to monitor the public positioning of a partner State, but the tools available are completely different (in terms of automatic translation, but also capacity to monitor and compute big amounts of information). Using advanced tools to monitor and evaluate the commitments of future partners is set to become normal diplomatic practice.

The same reflection applies to the legal effect pertaining to customary international law. The position taken by a State can be interpreted as a first evidence of a nascent custom on a specific issue. In that sense, a systematic investigation of the positions taken by States would make the contours of an emerging international consensus appear more clearly (if there is any). Of course, the positioning through strategy document, white paper or declaration is not alone sufficient, it must be shown to correspond to consistent behaviors and *opinio juris*. As we will address below, the object of this emerging consensus must not necessarily be a new norm, it could be an existing norm being interpreted in a new way. Overall, the first type of soft law can be described as "Statist soft law" – it refers mainly to the function of soft law as a form of public commitment by a State which, taken in an aggregate way, represent an important part of the emergence of a customary practice.

The second actors-based type of soft law focuses on international organizations (mainly intergovernmental organizations). It is interesting to start noticing that the two legal effects of the "Statist soft law" (good faith and emergence of a consensus) can be applied by analogy to the international organizations. This reflects the growing role of international organizations as international law subjects. This is well shown by focusing on the UNESCO "Recommendation on the Ethics of Artificial Intelligence".[45] The Recommendation represents the first globally agreed set of norms on the ethics of AI. The Recommendation contributes to the emergence of a consensus. On this dimension, the "International organization soft law" is distinct from the Statist soft law. By its very nature, the position taken by an international organization is the result of a consensus-building mechanism among its members. The position accepted by the UNESCO is not only the position of the international organization, but most importantly of all/the majority of its members. In that sense, the commitments taken by every UNESCO member creates expectations which could be captured by the protection offered to good faith. The UNESCO's Recommendation is an example of a specific soft law which can claim to already represent a global consensus. With respect to our parameters, the soft law produced by international organizations scores highly in terms of inclusive process.

If this example of UNESCO scores high, it means that we could further distinguish weaker versions of it. Soft law guidelines produced by an expert group in the context of the work of an international organization raises a completely different normative claim. This is the difference between the UNESCO's Recommendation and the preparatory work done by the Ad-hoc Experts Group (AHEG) which led to the first document on the ethics of AI in the UNESCO.[46] There is a sense of incrementality within the international organization. This is again different from the example of the UN-Report "The Age of Digital Interdependence" which has been commissioned by the General Secretary from a group of experts and which is not required to be accepted by all Member States. It is rather an attempt by the General Secretary to use his power to steer political attention and goodwill towards asking the "right" questions (agenda-setting). The

43 The other national strategies can be considered in the same way. Their language was graded 0,5 because of the generality and vagueness of the commitments taken.

44 Ashley Deeks, 'High-Tech International Law' (2020) 88 George Washington Law Review 574.

45 This recommendation was adopted by UNESCO's General Conference at its 41st session in November 2021.

46 See the first draft of the Recommendation on the ethics of AI (September 2020, SHS/BIO/AHEG-AI/2020) which was the object of multiple consultation processes among Member States.

objective is not a direct contribution to the emergence of a consensus (like the UNESCO), but rather the political framing of the discussions which, eventually, could lead to a consensus.

The third actors-based type of soft law norm is produced by non-state actors, mainly civil society organizations, research institutions, companies, companies' representatives and professional associations like the IEEE.

Starting with their commonalities, these groups advocate and claim relevance for their guidelines on different grounds. For civil society organizations and research institutions, legitimacy is largely claimed on the basis of expertise and/or inclusive process. As seen for instance in the case of the Montréal Declaration, both claims could converge. Companies and their representative organizations have also tried to base their claims on expertise, sometimes adding the dimension of power. It is not contested that companies bring a certain expertise (especially companies specialized in the development of digital technologies). However, this expertise is found in the context of important business interests. Moreover, certain companies add a dimension of power because of their financial capacity, the number of their customers/users or their capacity to exercise political pressure on specific decision-makers. It is important to note that these two dimensions are difficult to evaluate in the context of authoritarian regimes which exercise a tight political control of civil society, research institutions and companies. The different Chinese elements of the dataset are challenging to assess because of their arguable closeness to the government.

This third type of soft law is the one which moves away from a traditional understanding of soft law produced by States and international organizations. Civil society actors and companies have been given a "0" in the two categories of legal effects (good faith/emergence of custom). While these actors commit themselves publicly to a specific set of values and principles, their commitment has *per se* no direct relevance for international law. Nevertheless, it could have *indirect* relevance in two ways.

On the one hand, it could be relevant through the influence they exercise on States. A powerful company could exercise pressure on their home-base State in order to promote a specific understanding of the legal issues at stake and possible solutions to them. Companies coming together as an industry organization could exercise similar pressure. On the other hand, this commitment could have relevance due to the political forum it is derived from. This is especially true for civil society organizations whose guidelines are the result of inclusive processes. We shall address this specific connection in the next sub-section in showing how guidelines could produce legally relevant effects.

The case of global or regional professional associations like the IEEE has two dimensions. On the one hand, they represent a mix of some of the concerns of companies (mainly business interests mediated by a professional association) and inclusive processes (such as the inclusive consultation mechanisms put in place by the IEEE to elaborate its position). These elements can be accounted for as having indirect relevance (through pressure exercised on States) or as political fora. On the other hand, these professional associations have been promoting a different sort of soft law in the form of technical standards which are defined and accepted as consensus within a specialized community. This latter element is a challenging type of soft law for international law. In contrast to the traditional impact patterns of soft law (through new forms of interpretation and the emergence of new legal norms), the standards defined by the IEEE are *de facto* relevant because a global community of individuals and groups accept them as relevant. This specific type of normativity can be integrated into law when, for instance, a judicial body recognizes these standards as state of the art for a specific setting.[47]

These reflections on this third group of actors illustrate the "building blocks" approach assumed in this work. If one considers soft law to be exclusively about legal effects (which can only be produced by States and intergovernmental organizations), then all documents of the dataset produced by non-States actors do not qualify as soft law. This is *one* possible position. If one considers soft law to be also about legally relevant effects (e.g. political fora and the framing

47 On the different methods to account for technical standards in international law, Andrea Barrios Villarreal, *International Standardization and the Agreement on Technical Barriers to Trade* (Cambridge University Press 2018), 58 ff.

of a specific issue), the guidelines produced by civil society actors and companies could be considered to be soft law.

4.1.2 Types of soft law according to the legally relevant effects

To better grasp why civil society actors and companies could be considered to produce soft law, one needs to focus on the legally relevant effects they could produce. If such a case could be made, it would imply that these actors, while not being a traditional subject of international law, could generate soft law.

Our analysis shows that all documents of the dataset claim to provide interpretative guidance. They want to impact the way AI is regulated. Liu. et al. have usefully conceptualized this impact in three different categories (legal development/replacement/disruption).[48] Legal development mainly bears upon the development of legal norms related to AI-powered sociotechnical changes (through renewed interpretation, revision, or enactment of new legal norms). Legal development is opposed to legal replacement (a situation in which non-legal norms regulate AI-powered changes, such as design/code norms) or to legal disruption (a situation in which the fabric of international law is profoundly challenged by AI). These two latter types of impact are not considered here. In the legal development phase, two main patterns of legally relevant effects are possible for soft law: interpretative guidance for existing norms or enabling the emergence of new norms.

To which extent ethical guidelines can claim to provide interpretative guidance needs to be assessed along three dimensions of legitimacy: expertise, power and inclusive processes. These three dimensions can produce traction to follow one's interpretation. From a normative perspective, the most critical dimension is the one of power. The reference to power (defined as the capacity to impose one's view on others) is difficult to integrate into a conception of international law. In general, international law (like law, generally) is conceived as a mechanism for taming powers.[49] It channels disagreements and disputes in a way that makes sure that grounds other than power are considered relevant. This challenge is even reinforced in the cases in which the power dimension is linked with self-interest in the form of business objectives (e.g. a powerful tech giant). Seen from the perspective of international law, this question leads us to underlying legal-theoretical questions about the legitimacy of international law, its sources and, eventually, soft law. This paper obviously cannot be the place for this debate. However, this paper can be the place to acknowledge that soft law could (also) be about power. As mentioned in the previous sub-section, powerful companies or civil society organizations can influence the way specific legal norms are interpreted. A political sciences approach could try to assess whether, and if yes how, such companies/organizations could use their power to promote specific interpretative patterns, both domestically and internationally.[50] Most importantly, this kind of approach must find out which further actors such as States or international organizations (i.e. traditional norm-setting actors in the sense of the first cluster of results) are the target of these advocacy/lobbying efforts. For instance, the interpretation of existing international humanitarian norms (which are to be applied in the context of armed conflicts) could be the result of pressures exercised from both private companies (such as defense industries) and civil society organizations (such as advocacy campaigns like "Stop killer robots").[51]

The dimensions of expertise and inclusive processes are less problematic to defend as being part of a conception of international law. It does not mean that they are without difficulties, but that, at least conceptually and normatively, it seems plausible to integrate them. For

48 For the use of these concepts "development/replacement/disruption", Liu and others, 'Artificial Intelligence and Legal Disruption: A New Model for Analysis' (2020) 12 Law, Innovation and Technology 205, 15 ff.

49 This is especially true for a broadly constitutionalist approach. See e.g. Nicolas Suzor, 'Digital Constitutionalism: Using the Rule of Law to Evaluate the Legitimacy of Governance by Platforms' (2018) 4 Social Media + Society 1.

50 This connects to investigation on lobbying/advocacy in the context of new AI regulations. See for instance the efforts to influence the EU AI Act. "Big Tech Boost Lobbying Spending in Bruxelles", Politico, March 2022. https://www.politico.eu/article/big-tech-boosts-lobbying-spending-in-brussels/.

51 On this campaign, consult https://www.stopkillerrobots.org.

instance, the inclusive process dimension can be accounted for from the perspective of a more democratic international law.[52] The expertise dimension could be integrated through interpretation as subsidiary means of interpretation or through the establishment of specialized consulting bodies.[53]

In addition to these three dimensions, the impact of soft law through interpretative guidance raises the question of the target norm (the *interpretandum*). As proposed by d'Aspremont, we need to distinguish between interpretation as content determination (what is the content of a specific norm?) and law ascertainment (whether a given norm can claim to be part of international law).[54] The first dimension is the focus of this cluster and bears upon searching for the meaning of the content of a specific norm. The second dimension was referred to above when dealing with the sources of international law. The analysis of the different documents has shown that the objective is to interpret anew specific types of norms systematically. Numerous guidelines refer to the renewed interpretation of human rights and principles such as liability or accountability.

These privileged types of target norms can be described as foundational norms in their specific legal regimes. As to their function, these general norms grasp and express in legal terms the political and moral values upon which a specific regime is founded.[55] Because of their high level of abstractness, these norms are characterized by the important interpretation and reasoning effort they require.[56] Several international legal norms could appear in this category. This is most importantly the case for general principles relevant to AI. As Redgwell notes in discussing soft law in environmental law, there is some discussion as to which principles exactly should be understood here (general principles of municipal law held by all States, general principles of international law, or general principles of legal systems in general), and about which method of ascertainment is adequate.[57] In addition to these general principles, we might add the category of "values" if they are part of a specific legal regime.[58] Furthermore, human rights should be included in this category of foundational norms.[59] They are formulated as general norms and require a significant interpretative effort to be applied to a specific situation. Furthermore, they have a foundational dimension by concretizing the essential values that a specific legal regime should uphold in a legal norm.

These foundational norms are the usual *interpretandum* for a renewed interpretation. The challenge raised here is how to account for the interpretative guidance in the context of the interpretation methods prescribed by the Vienna Convention on the Interpretation of Treaties. Art. 31(1) VCLT prescribes how to interpret treaty norms by considering them in light of the object and purpose of the treaty, but it does not provide information on how to interpret the norms which themselves represent or are closely linked to this "object and purpose" (such as

52 The idea could be to draw upon demoi-cratic accounts of international law. For this idea in EU law, see Kalypso Nicolaidis, 'The Idea of European Demoicracy' in Julie Dickson and Pavlos Eleftheriadis (eds), *The Philosophical Foundations of European Union* (Oxford University Press, 2012). For international law, S. Besson, 'Institutionalizing Global Demoi-Cracy' in Lukas Meyer (ed), *Justice, Legitimacy and Public International Law* (Cambridge University Press, 2009.

53 Inspiration could be found in the references to expertise in international environmental law, Steinar Andresen, 'The role of scientific expertise in multilateral environmental agreements: influence and effectiveness' in Ellen Hey and others (eds), *The Role of 'Experts' in International and European Decision-Making Processes: Advisors, Decision Makers or Irrelevant Actors?* (Cambridge University Press, 2014).

54 Jean D'Aspremont, 'The Multidimensional Process of Interpretation: Content-Determination and Law-Ascertainment Distinguished' in Andrea Bianchi, Daniel Peat and Matthew Windsor (eds), *Interpretation in International Law* (Oxford University Press, 2015, 114.

55 Samantha Besson, 'General Principles in International Law - Whose Principles?' in Samantha Besson, Pascal Pichonnaz and Marie-Louise Gächter-Alge (eds), *Les principes en droit européen - Principles in European law* (Genève, Schulthess, 2011), 26.

56 Joseph Raz, 'Legal Principles and the Limits of Law' (1972) 81 The Yale Law Journal 823, 838.

57 Catherine Redgwell, 'General Principles of International Law' in Stefan Vogenauer and Stephen Weatherill (eds), *General Principles of Law: European and Comparative Perspectives* (Oxford, Hart, 2017, 10.

58 For instance, in EU law, values refer to Art. 2 TEU and objectives refer to Art. 3(6) TEU. In international law, for example, see Crawford's argument considering "sovereignty" as a legal value. James Crawford, 'Sovereignty as a Legal Value' in James Crawford and Martti Koskenniemi (eds), *The Cambridge Companion to International Law* (Cambridge University Press, 2012).

59 For a similar idea, see Samantha Besson, 'The Law in Human Rights Theory' (2013) 7 Zeitschrift für Menschenrechte – Journal for Human Rights 120, 125.

the objectives, values, and principles of a treaty).[60] For instance, numerous guidelines referring specifically to autonomous weapons systems (AWS) try to provide guidance for the principle of "meaningful human control".[61] This principle could be considered foundational with respect to this specific body of norms. From the perspective of our argument, guidelines claiming to provide interpretative guidance for this concept could be considered soft law because of the legally relevant effect they produce. Indeed, the qualification of the *concrete* effect they produce on the interpretation of what a "meaningful human control" requires could only be assessed in the mid-term, for instance by taking into account the fact that other actors refer to their interpretation (e.g. a State in its diplomatic effort). In the meanwhile – not knowing whether its specific interpretation will "take off" – the guideline (or a specific norm entailed by the guideline) could nevertheless be considered soft law *en puissance*, yet to be realized. To do so allows one to consider the normative richness of the regulatory landscape on AI. The other option would be to count as soft law only interpretative guidance which achieves a certain level of relevance (for instance by being referred to by relevant institutional bodies). This option seems to reflect a traditional understanding of soft law which implies that interpretative guidance needs to be referred to by actors such as States or international organizations in order to be recognized as soft law.

On the emergence of new norms, our analysis has highlighted two different dimensions. The first dimension regards soft law as an incremental norm. The idea is that soft law captures yet-to-be-finalized legislative projects. The analysis has shown that this category was not very useful for structuring the current regulatory landscape on AI. This category was only really useful in the case of the EU-internal process of emergence of a consensus on how to regulate AI. The idea of incrementality within a specific international organization/body can be interesting if one wants to focus on the step-by-step evolution of a specific position. However, this dimension refers more to a political sciences approach capturing the way an international organization/body formulates and matures its position.[62] Of course, depending on the size of the international organization and the range of its members (e.g. virtually universal in the case of the UNESCO), this internal process is akin to the political forum situation we will address below.

The second dimension is about the political forum function, which soft law could fulfil for the emergence of new legal norms. This function is linked to the design of the political process from which legally relevant effects could emerge. Numerous parameters of analysis are thinkable in assessing the normative qualities of this process. This paper focuses on inclusivity as a way of emphasizing that the final content of the document may be less important than the process upon which it relies and the stakeholders it gathers.

It seems clear that the common position of the G20 and the Montréal Declaration could both be said to fulfil a political forum function. But the inclusivity of the processes is very distinct, though both gather diverse actors at a common table. An inclusive process enables the inclusion of more and diverse voices. This bears upon the types of actors participating in the process, but also on their geographical diversity and types of interests for which they advocate. Geographical diversity is especially crucial when it comes to the regulation of digital technologies. In this domain, power imbalances (in terms of technological developments, know-how, capacity to control agenda-setting) are highly relevant.

60 David S. Jonas and Thomas N. Saunders, 'The Object and Purpose of a Treaty: Three Interpretive Methods' (2010) 43 Vanderbilt Journal of Transnational Law 566; Dino Kritsiotis, 'The Object and Purpose of a Treaty's Object and Purpose' in Dino Kritsiotis and Michael J. Bowman (eds), *Conceptual and Contextual Perspectives on the Modern Law of Treaties* (Cambridge, Cambridge University Press, 2018), 277.

61 Important impetus for this principle was given by the NGO "Article 36", specialized in reducing harm from weapons. For the original contribution, Heather M. Roff and Richard Moyes, *Meaningful Human Control, Artificial Intelligence and Autonomous Weapons: Briefing paper prepared for the Informal Meeting of Experts on Lethal Autonomous Weapons Systems, UN Convention on Certain Conventional Weapons* (Article 36 2016). For an overview of the literature and a conceptualization of the term, Filippo Santoni de Sio and Jeroen van den Hoven, 'Meaningful Human Control over Autonomous Systems: A Philosophical Account' (2018) 5 Frontiers in Robotics and AI. For a differentiated approach according to the type of AWS at stake, Daniele Amoroso and Guglielmo Tamburrini, 'Toward a Normative Model of Meaningful Human Control over Weapons Systems' (2021) 35 Ethics & International Affairs 245.

62 In EU law, see the Nihit Goyal, Michael Howlett and Araz Taeihagh, 'Why and how does the regulation of emerging technologies occur? Explaining the adoption of the EU General Data Protection Regulation using the multiple streams framework' (2021) 15 Regulation & Governance 1020.

The process allows stakeholders to frame the relevant questions (in a problem-finding phase), before outlining potential ways to address these questions (in a problem-solving phase). In the first phase, the framing dimension of soft law appears to be crucial.[63] In looking at the documents of the dataset, specific legal issues are especially relevant because they are considered in numerous documents as relevant. Questions such as liability rules and data protection law frame the political discussion by highlighting the "relevant" questions and the "relevant" legal categories for addressing them. The process of defining which are the relevant questions and categories should begin as broadly as possible. This phase is crucial for the capacity of international law to actualize itself in the face of changing circumstances and potential threats to its regulative functions.[64]

Overall, a focus put on the legally relevant effects (mainly through interpretative guidance or as political forum) shows how ethical guidelines could claimed to be soft law from the perspective of the impact they deploy. However, the question remains as to the concrete realization of this impact through traditional actors such as States and international organizations.

Before turning to the final cluster of results, four general types of soft law appear especially relevant for the analysis of the current regulatory landscape on AI. In the continuation of the building blocks approach, these general types should be taken as conceptual lenses useful for looking at the current regulatory landscape.

NAME	STATIST AND INTERNATIONAL ORGANIZATION SOFT LAW	PROCESS-ORIENTED SOFT LAW	EXPERTISE-ORIENTED SOFT LAW	DE FACTO RELEVANT STANDARD SOFT LAW
Main features	- Focus on States/IO - Aggregative dimension (=> emergence of a consensus)	- Focus on legally relevant effect through crystallization of relevant issues/solutions - Claim to identify/structure the relevant issues and/or to provide interpretative guidance	- Focus on legally relevant effect through interpretative guidance	- Technical standards/ rules set by global/ regional organizations
Further research questions	- Empirical investigation on an emerging consensus - Computational tools to further structure and analyse texts (NLP/sentiments analysis, see below)	- Empirical investigation on the impact of specific crystallization processes - Normative work on the quality required for the process	- Definition of expertise (e.g. tensions with economic/political interests) - Normative work on integrating this interpretative guidance into VCLT canons - Empirical investigation on how this expertise enters the law	- Empirical investigation on how these technical standards enter the law - Normative work on the quality required for the standard-setting process

4.1.3 Law and ethics in soft law on AI

A further type of result bears upon the relations between law and ethics as conceptualized in the dataset evaluated. While no document in the dataset explicitly addresses the meta-question of their own positioning,[65] numerous documents of the dataset mention law and/or ethics and their interactions. On the basis of the analysis of these references, two broad conceptions of the relations between ethics and the law seem to emerge. The first conception assumes that ethical and legal requirements (e.g. with respect to the design or use of AI systems) coexist in parallel. These requirements represent and operationalize different types of obligations. Their common ground is the ambition to guide the behavior of the relevant actors. We could speak of two different normativities which coexist.

The first conception is for instance illustrated by the "Beijing AI principles". In this document, the authors explain that "researchers and developers of AI should have sufficient considerations for the potential ethical, legal, and social impacts and risks brought in by their products and

63 The concept of framing is taken from communication studies. Michael J. Carter, 'The Hermeneutics of Frames and Framing: An Examination of the Media's Construction of Reality' (2013) 3 SAGE Open. On framing in the law, Riikka Koulu, 'Human Control over Automation: EU Policy and AI Ethics' (2020) 12 European Journal of Legal Studies 9, 24; Paul James Cardwell and Tamara Hervey, 'Bringing the Technical into the Socio-legal: The Metaphors of Law and Legal Scholarship of a Twenty-First Century European Union' in David Cowan and Daniel Wincott (eds), *Exploring the 'Legal' in Socio-Legal Studies* (London, Palgrave Macmillan UK), 2016.

64 Hin-Yan Liu and M. Maas, ''Solving for X?' Towards a Problem-Finding Framework to Ground Long-Term Governance Strategies for Artificial Intelligence' (2021) 107 Legal Studies Research Paper Series University of Copenhagen Faculty of Law 1.

65 The Unesco ad-hoc Group's work could be considered to come close to this self-reflexivity.

take concrete actions to reduce and avoid them." In a similar way, the G20 declaration on trade and digital economy explains how a digital ecosystem should be organized. It explains in particular that "governments should consider promoting mechanisms, such as data trusts, to support the safe, fair, legal and ethical sharing of data." Interesting in that formulation is the superposition of the different normative demands: ethical and legal demands coexist in a cumulative way.

The second conception takes the legal regime in force as a given contextual element. Ethical norms are used as tools to assess whether the law does what it ought to. In this conception, ethics has to reveal potential failures of the current legal regime. Ethical norms appear as exercising a critical force on existing law, *a minima* in the sense of an assessment of the capacity of the law to guarantee certain principles seen as necessary or desirable. In this conception, ethical norms are often addressed to political and judicial bodies which have the competence to enact new legal norms or revise legal interpretation of existing norms.[66]

The second conception is exemplified by the position expressed by the House of Lords Committee on AI in its report. As the report explains, "there is no consensus regarding the adequacy of existing legislation should AI systems malfunction, underperform or otherwise make erroneous decisions which cause harm. We ask the Law Commission to provide clarity. (...) We recommend that the Law Commission consider the adequacy of existing legislation to address the legal liability issues of AI and, where appropriate, recommend to Government appropriate remedies to ensure that the law is clear in this area."[67] The legal regime is expected to guarantee certain principles, both on a general meta-level (such as clarity) but also with respect to different substantive elements (liability in this example).

With respect to these two categories, soft law as defined along our building blocks approach appears to have two distinct functions. Firstly, as exemplified by the second category, soft law has a critical potential with respect to existing legal norms. Soft law norms are a way of making hard law evolve, by providing renewed interpretative guidance or by representing a political forum in which relevant issues crystallize. The foundational norms described previously can be the entry-door for this critical stance on existing norms. By referring to general principles or human rights, soft law norms function both as a yardstick for the evaluation of existing legal norms and as a tool for developing legal arguments for the further development of the same legal provisions. Following Lenaerts and Gutiérrez-Fons in their argument about the role of foundational norms for the EU, this explains why principles are crucial instruments for a constitutional dialogue which "facilitate the constant renewal of the EU legal order, epitomizing the 'EU's living constitution'".[68] By analogy for the regulation of AI, it can be said that principles have a similar function and, specifically because of this function, these principles are a promising *interpretandum* for soft law norms.

Secondly, soft law norms could be considered as normative vehicles for ethical content. In colloquial terms, it could be argued that ethical content is "upgraded" into the form of legal content through soft law (assuming the law is an upgrade, for instance in terms of traction for the behaviors of relevant actors). Soft law thus plays a specific role in bringing ethical content into the law. As argued above, this happens mainly thanks to the legally relevant effects soft law could deploy (interpretation guidance and political fora). If one defines soft law on the basis of these legally relevant effects, content primarily defined as ethical could find its way into legal content.

4.2 METHODOLOGICAL CHALLENGES AND OUTLOOK

Keeping in mind the objective of outlining a research proposal meant to open original research paths, it is important to pause briefly in order to address the methodological challenges raised by the approach used here. The first challenge bears upon the definition of the three levels required for the tagging: the parameters (the distinctive features of soft law), the measurement

66 For instance, this conception seems to be assumed by the "Montréal Declaration" or the "Universal Guidelines for Artificial Intelligence" of the Public Voice Coalition.

67 UK House of Lords, "AI in the UK: Ready, Willing and Able?", Ex. Summary.

68 Koen Lenaerts and Jose A. Gutiérrez-Fons, 'The Constitutional Allocation of Powers and General Principles of EU Law' (2010) 10 Common Market Law Review 1629, 1169.

DOI: 10.5334/tilr.297

elements (is it the best way to measure the parameter?) and the specific instantiation of the measurement (Is document X well evaluated? Why is this specific numerical value chosen?). One of the challenges associated with these three levels is the difficulty of starting the analysis. These three levels raise deep, normative challenges which often relate to disputes which have characterized legal theory scholarship for decades (the definition of soft law, link between soft law and interpretation, link between soft law and the sources of international law). The ambition of solving these challenges (if this idea of "solving" makes sense in the first place) is impracticable. One has to start the analysis by making one's assumptions as explicit as possible.

Secondly, the actual process of tagging the material raises evaluative challenges. The methodological challenge here is to address ambiguities in the dataset and objectivize evaluative choices in tagging the dataset. Ideal practices would require different researchers to justify disputable annotation decisions, thereby making their decisions understandable and challengeable. However, these ideal practices are often out of reach if you do not have a whole team tasked with this tagging effort.[69] The second-best alternative is (again) transparency. The dataset and the annotation must be accessible and it must be possible to understand why this specific evaluation has been favored. In light of collaborative work opportunities made possible by digital technologies, it is conceivable to consider the evaluation as challengeable work. Like the revision of Wikipedia entries, it would then be possible to challenge a specific evaluation and to provide new insights/justification for the new evaluation, thereby contribution to a community-based effort to improve the understanding of a specific document of the dataset.

Thirdly, it can be argued that the approach chosen is too broad and that, eventually, every publicly expressed normative endeavor (like every ethical guideline considered in the dataset) could be taken as a form of soft law. In the end, this approach seems to conflate ethical and legal arguments or, in other words, the specific legal dimension of soft law seems to disappear. The reply to this line of criticism is two-fold. On the one hand, this impression of everything being soft law comes directly from the building blocks approach chosen. These building blocks reflect definitional parameters which, alone or together, characterize a normative object called soft law. The dispute should therefore bear upon the definition of these parameters. On the other hand, the fact that specific items of the dataset reflect one of these parameters only means that, from the perspective of this parameter used as part of the definition of what is soft law, the item X qualifies as soft law. In the meantime, other dimensions of the item X might be not qualify as soft law.

Overall, this criticism underlines the arguable strength of this approach: soft law should not be considered in all-or-nothing fashion, but rather as a complex ensemble of different elements. Whether a specific guideline qualifies as soft law or not is maybe not the most interesting question, but rather which dimension of the guideline does from the perspective of which definition of soft law.

Furthermore, this "everything-is-soft-law" criticism is right to recall that the results of the tagging efforts are, by definition, descriptive. They cannot and are not meant to solve normative disputes. They can enrich them and offer empirically informed perspectives on the current state of international law on AI. But in order to be fully integrated into a legal argument, further steps are required. It is their interpretation which could make them part of a prescriptive argument.

5 CONCLUSION

This paper has outlined a research proposal to give clearer contours to the international soft law on AI. To do so, it has proposed a law-as-data approach based on the conceptualization of building blocks composing a set of possible soft law definitions. In conclusion, I would like to sketch three interesting paths for future work.

Firstly, both the parameters defined and the dataset considered could be enriched. The parameters could be made more specific with the objective of generating insights relevant to improving one's understanding of the "messy" regulatory landscape on AI. Likewise, the quality of the analysis can be improved by integrating more documents into the dataset.

69 The tagging proposed here was done in three rounds with two research assistants.

DOI: 10.5334/tilr.297

Secondly, with the objective of creating a dataset for international legal scholars, it could be possible to design a usable and community-built database. This hard law/soft law approach could itself be an add-on to existing databases on AI soft law (like the OECD's database). It would be possible to query the database for a specific soft law definition (i.e., a specific combination of the identified parameters, such as "NGO" as actor and "political forum" as legally relevant effect) and to see which elements of the dataset qualify as what type of soft law according to this distinction. This is a first step to providing analytical visualization of this regulatory landscape.[70] These visualizations could contribute to policy-making by making developed insights easily accessible.

Thirdly, as explained above, more exploratory computational tools could be used to look for patterns within the resources of the structured dataset. Network analysis tools could be used to account for putative normative relations within the resources gathered. In addition to the network analysis tools mentioned above, NLP tools such "sentiment analysis" tools could be tried out in a logic of high-risks/high-gains.[71] The tool searches texts and establishes a sentiment score for each resource. This score is taken as an interesting resource to pinpoint specific resources of the dataset, understand their characteristics in the context of a fresh interpretation and try to classify them according to the broad categories of assumptions they refer to when considering AI. This type of approach connects to the ambition to identify and draw upon the way the law is embedded into broader social practices. This work provides a connection to science & technology studies (STS) scholarship.[72] As a prominent example, Jasanoff and Kim have defined socio-technical imaginaries as "collectively imagined forms of social life and social order reflected in the design and fulfilment of nation-specific scientific and/or technological projects".[73] Sentiments analysis tools could help us understand the impact that assumptions on an emerging technology (e.g., seen as a threat or an opportunity) has on the legal interpretation of existing norms.

ANNEX

LIST DOCUMENTS DATASET (ALPHABETICAL ORDER)

- Access Now, 'Human Rights in the Age of Artificial Intelligence' (2018) <https://www.accessnow.org/cms/assets/uploads/2018/11/AI-and-Human-Rights.pdf>
- Amnesty International, Access Now, 'Toronto Declaration: Protecting the Right to Equality and Non-Discrimination in Machine Learning Systems' (2018) <https://www.accessnow.org/cms/assets/uploads/2018/08/The-Toronto-Declaration_ENG_08-2018.pdf>
- Beijing Academy of Artificial Intelligence, 'Beijing AI Principles' (2019) <https://www-pre.baai.ac.cn/news/beijing-ai-principles-en.html>
- British Embassy in Mexico City, 'Artificial Intelligence in Mexico (La Inteligencia Artificial En México)' (2018) <https://docs.wixstatic.com/ugd/7be025_ba24a518a53a4275af4d7ff63b4cf594.pdf>
- Chinese National Governance Committee for the New Generation Artificial Intelligence, led by China's Ministry of Science and Technology, 'Governance Principles for a New Generation of Artificial Intelligence: Develop Responsible Artificial Intelligence' (2019) <http://www.chinadaily.com.cn/a/201906/17/WS5d07486ba3103dbf14328ab7.html>

70 On visual legal analytics tools, see N. Lettieri and D. Malandrino, 'Cartographies of the Legal World. Rise and Challenges of Visual Legal Analytics' (2018) 2018 22nd International Conference Information Visualisation (IV) 241. For examples, see Wolfgang Alschner and Dmitriy Skougarevskiy, 'Mapping the Universe of International Investment Agreements' (2016) 19 Journal of International Economic Law 561, 575 ff. For a seminal article on visualization, see D. Keim and others, 'Visual Analytics: Definition, Process, and Challenges' in A. Kerren and others (eds), *Information Visualization Lecture Notes in Computer Science* (Berlin, Springer, 2008).

71 Lori Young and Stuart Soroka, 'Affective News: The Automated Coding of Sentiment in Political Texts' (2012) 29 Political Communication 205, 205. For the application to the WTO judgments, see Busch and Pelc, 'Words Matter: How WTO Rulings Handle Controversy' (2019) 63 International Studies Quarterly 464, 473.

72 STS scholarship are often combined with critical legal theory (such as counterfactual approach to promote thinking anew about imaginaries of international law), for instance by Mohsen al Attar, 'The Necessity of Imagination: Using the Counterfactual Method to Overcome International law's Epistemological Limitations' (2021) 33 National Law School of India Review 1. On AI, Christoph Graber, 'How the Law Learns in the Digital Society' (2021) 3 Law, Technology and Humans 12.

73 Sheila Jasanoff and Sang-Hyun Kim, 'Containing the Atom: Sociotechnical Imaginaries and Nuclear Power in the United States and South Korea' (2009) 47 Minerva 119, 120.

- Committee on Standards in Public Life (UK), 'Artificial Intelligence and Public Standards: Report' (2020) <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/868284/Web_Version_AI_and_Public_Standards.PDF>
- Council of Europe, European Commission for the Efficiency of Justice, 'European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment' (2018) <https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c>
- European Commission's High-Level Expert Group on Artificial Intelligence, 'Ethics Guidelines for Trustworthy AI' (2018) <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>
- European Commission, "White Paper On Artificial Intelligence - A European approach to excellence and trust' (2020), <https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf>
- European Group on Ethics in Science and New Technologies, "Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems' (2018) <https://op.europa.eu/en/publication-detail/-/publication/dfebe62e-4ce9-11e8-be1d-01aa75ed71a1/language-en>
- Future of Life Institute, 'Asilomar AI Principles' (2017) <https://futureoflife.org/ai-principles/?cn-reloaded=1>
- G20 Trade Ministers and Digital Economy Ministers, 'G20 Ministerial Statement on Trade and Digital Economy' (2019) <https://www.mofa.go.jp/files/000486596.pdf>
- German Federal Government, 'Update: Artificial Intelligence Strategy' (2020) <https://www.ki-strategie-deutschland.de/files/downloads/Fortschreibung_KI-Strategie_engl.pdf>
- Google, 'AI at Google: Our Principles' (2018) <https://www.blog.google/technology/ai/ai-principles/>
- IDB Inter-American Development Bank, 'fAIr LAC: Adopción ética y responsable de la inteligencia artificial en América Latina y el Caribe' (2020), <https://publications.iadb.org/publications/english/document/fAIr-LAC-Responsible-and-Widespread-Adoption-of-Artificial-Intelligence-in-Latin-America-and-the-Caribbean.pdf>
- IBM, 'IBM Everyday Ethics for AI' (2019) <https://www.ibm.com/watson/assets/duo/pdf/everydayethics.pdf>
- IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, 'Ethically Aligned Design: A Vision for Prioritizing Human Well-Being with Autonomous and Intelligent Systems' (2019) First Edition <https://ethicsinaction.ieee.org/>
- Information Technology Industry Council, 'AI Policy Principles' (2017) <https://www.itic.org/resources/AI-Policy-Principles-FullReport2.pdf>
- International Telecommunication Union, 'AI for Good: Global Summit' (2017) <https://www.itu.int/en/ITU-T/AI/Documents/Report/AI_for_Good_Global_Summit_Report_2017.pdf>
- Japanese Cabinet Office, Council for Science, Technology and Innovation, 'Social Principles of Human-Centric Artificial Intelligence' (2019) <https://www8.cao.go.jp/cstp/english/humancentricai.pdf>
- Microsoft, 'AI Principles' (2018) <https://www.microsoft.com/en-us/ai/our-approach-to-ai>
- Mission assigned by the French Prime Minister, 'For a Meaningful Artificial Intelligence: Toward a French and European Strategy' (2018) <https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf>
- Monetary Authority of Singapore, 'Principles to Promote Fairness, Ethics, Accountability and Transparency (FEAT) in the Use of Artificial Intelligence and Data Analytics in Singapore's Financial Sector' (2019) <http://www.mas.gov.sg/~/media/MAS/News%20and%20Publications/Monographs%20and%20Information%20Papers/FEAT%20Principles%20Final.pdf>
- New York Times' New Work Summit, 'Seeking Ground Rules for AI' (2019) <https://www.nytimes.com/2019/03/01/business/ethical-ai-recommendations.html>
- New-Zealand, 'Algorithm charter for Aotearoa New Zealand' (2020), <https://data.govt.nz/assets/data-ethics/algorithm/Algorithm-Charter-2020_Final-English-1.pdf>
- Niti Aayog, 'National Strategy for Artificial Intelligence: #AI for All (Discussion Paper)' (2018) <https://www.niti.gov.in/writereaddata/files/document_publication/NationalStrategy-for-AI-Discussion-Paper.pdf>
- Organisation for Economic Co-operation and Development, 'Recommendation of the Council on Artificial Intelligence' (2019) <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>

- Partnership on AI, 'Tenets' (2016) <https://www.partnershiponai.org/tenets/>
- Smart Dubai, 'Artificial Intelligence Principles and Ethics' (2019) <https://smartdubai.ae/initiatives/ai-principles-ethics>
- Standard Administration of China, 'White Paper on Artificial Intelligence Standardization' *excerpts in English published by New America* (2018) <https://www.newamerica.org/cybersecurity-initiative/digichina/blog/translation-excerpts-chinas-white-paper-artificial-intelligence-standardization/>
- Telefónica, 'AI Principles of Telefónica' (2018) <https://www.telefonica.com/en/web/responsible-business/our-commitments/ai-principles>
- Telia Company, 'Guiding Principles on Trusted AI Ethics' (2019) <https://www.teliacompany.com/globalassets/telia-company/documents/about-telia-company/public-policy/2018/guiding-principles-on-trusted-ai-ethics.pdf>
- Tencent Institute, 'Six Principles of AI' (2017) <http://www.kejilie.com/iyiou/article/ZRZFn2.html>
- The Public Voice Coalition, 'Universal Guidelines for Artificial Intelligence' (2018) <https://thepublicvoice.org/ai-universal-guidelines/>
- Think 20, 'Future of Work and Education for the Digital Age' (2018) <https://www.g20-insights.org/wp-content/uploads/2018/07/TF1-1-11-Policy-Briefs_T20ARG_Towards-a-G20-Framework-For-Artificial-Intelligence-in-the-Workplace.pdf>
- UN Secretary-General's High-level Panel on Digital Cooperation, 'The Age of Interdependence' (2020) <https://www.un.org/en/pdfs/DigitalCooperation-report-for%20web.pdf>
- UN Secretary-General, 'Roadmap for digital cooperation' (2020) <https://documents-dds-ny.un.org/doc/UNDOC/GEN/N20/102/51/PDF/N2010251.pdf?OpenElement>
- UN Chief Executives Board for Coordination, 'A United Nations system-wide strategic approach and road map for supporting capacity development on artificial intelligence' (2019), <https://unsceb.org/sites/default/files/2020-09/CEB_2019_1_Add-3-EN_0.pdf>
- UK House of Lords, Select Committee on Artificial Intelligence, 'AI in the UK: Ready, Willing and Able?' (2018) Report of Session 2017-19 <https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf>
- UNESCO, 'Recommendation on the Ethics of Artificial Intelligence' (2021) <https://unesdoc.unesco.org/ark:/48223/pf0000379920.page=14>
- UNI Global Union, 'Top 10 Principles for Ethical Artificial Intelligence' (2017) <http://www.thefutureworldofwork.org/media/35420/uni_ethical_ai.pdf>
- United States Executive Office of the President, National Science and Technology Council Committee on Technology, 'Preparing for the Future of Artificial Intelligence' (2016) <https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf>
- University of Montreal, 'Montreal Declaration for a Responsible Development of Artificial Intelligence' (2018) <https://www.montrealdeclaration-responsibleai.com/the-declaration>

With respect to the dataset of Fjeld. et al, the following documents were not considered here:

- Artificial Intelligence Industry Alliance, 'Artificial Intelligence Industry Code of Conduct (Consultation Version)' (2019) https://www.secrss.com/articles/11099 (*no translation available)*
- IA Latam, 'Declaración de Principios Éticos Para La IA de Latinoamérica' (*not found anymore – another Latin-American Declaration was taken)*
- Tencent Institute, 'Six Principles of AI' (2017) http://www.kejilie.com/iyiou/article/ZRZFn2.html (*no translation available)*

The following documents were updated:

- European Commission, 'Artificial Intelligence for Europe: Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee, and the Committee of the Regions' COM (2018) 237 https://ec.europa.eu/digital-single-market/en/news/communication-artificial-intelligence-europe
- German Federal Government, 'Update: Artificial Intelligence Strategy' (2020) <https://www.ki-strategie-deutschland.de/files/downloads/Fortschreibung_KI-Strategie_engl.pdf>

ADDITIONAL FILE

The additional file for this article can be found as follows:

- **Annex table data.** Soft law documents. DOI: https://doi.org/10.5334/tilr.297.s1

ACKNOWLEDGEMENT

I would like to thank the participants of the conference of the working group "Legal Theory" of the European Society for International Law in Stockholm (2021), Céline Braumann, Tilman Altwicker, Julien Briguet, Antoine Dobrynski, Chiara Skirl, and the two anonymous reviewers of this journal for their valuable feedback.

COMPETING INTERESTS

The author has no competing interests to declare.

AUTHOR AFFILIATION

Johan Rochel orcid.org/0000-0002-8123-4057
Postdoc researcher, EPFL, Lausanne, Switzerland

TO CITE THIS ARTICLE:
Rochel, Johan, 'Learning from the Ethics of AI – A Research Proposal on Soft Law and Ethics of AI' (2022) 27(1) *Tilburg Law Review* pp. 37–59. DOI: https://doi.org/10.5334/tilr.297

Published: 17 February 2023

Tilburg Law Review is a peer-reviewed open access journal published by Ubiquity Press.