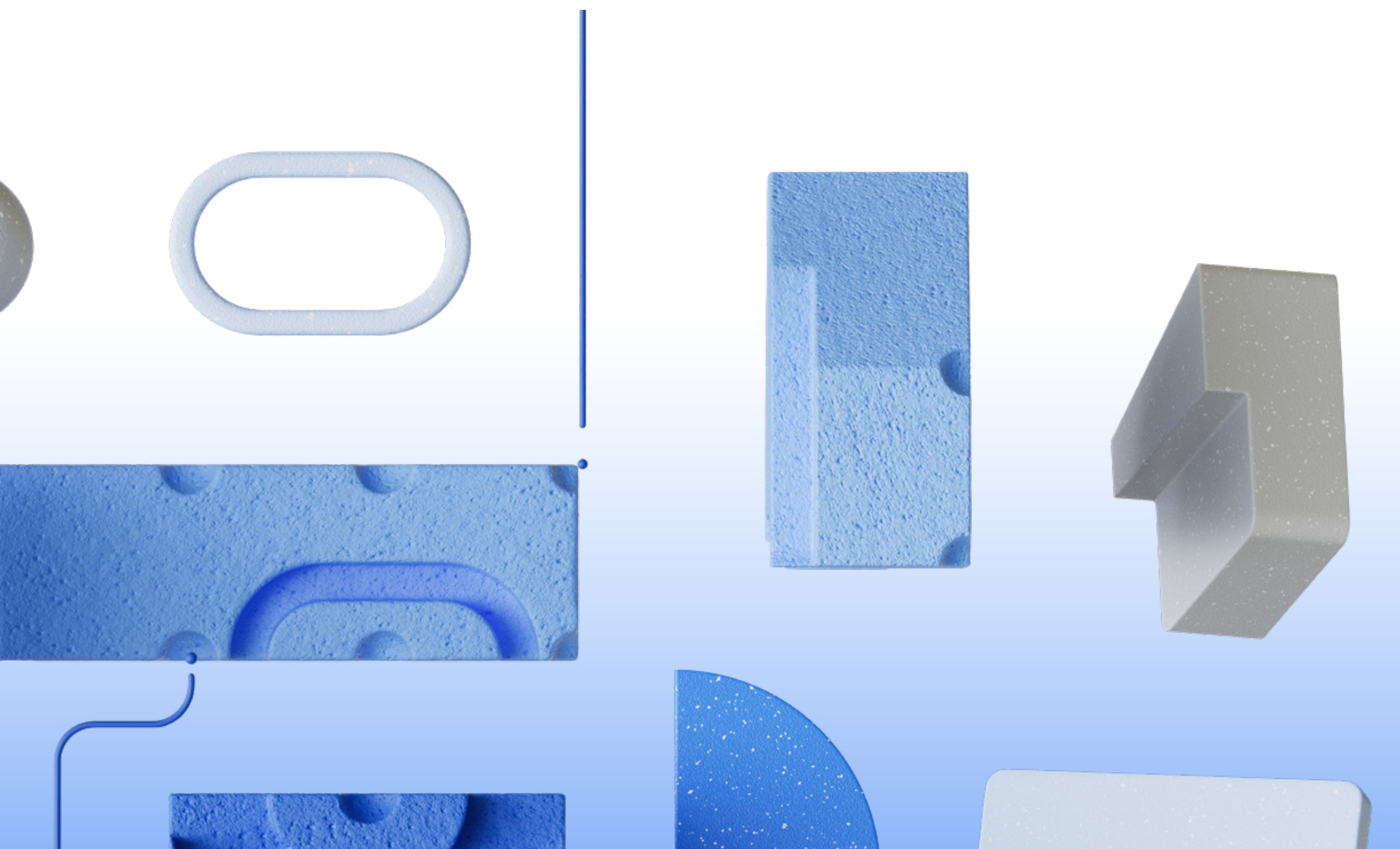
Google DeepMind

Organizing Intelligence

16 big ideas from frontier AI research that will redefine how we build, lead, and scale



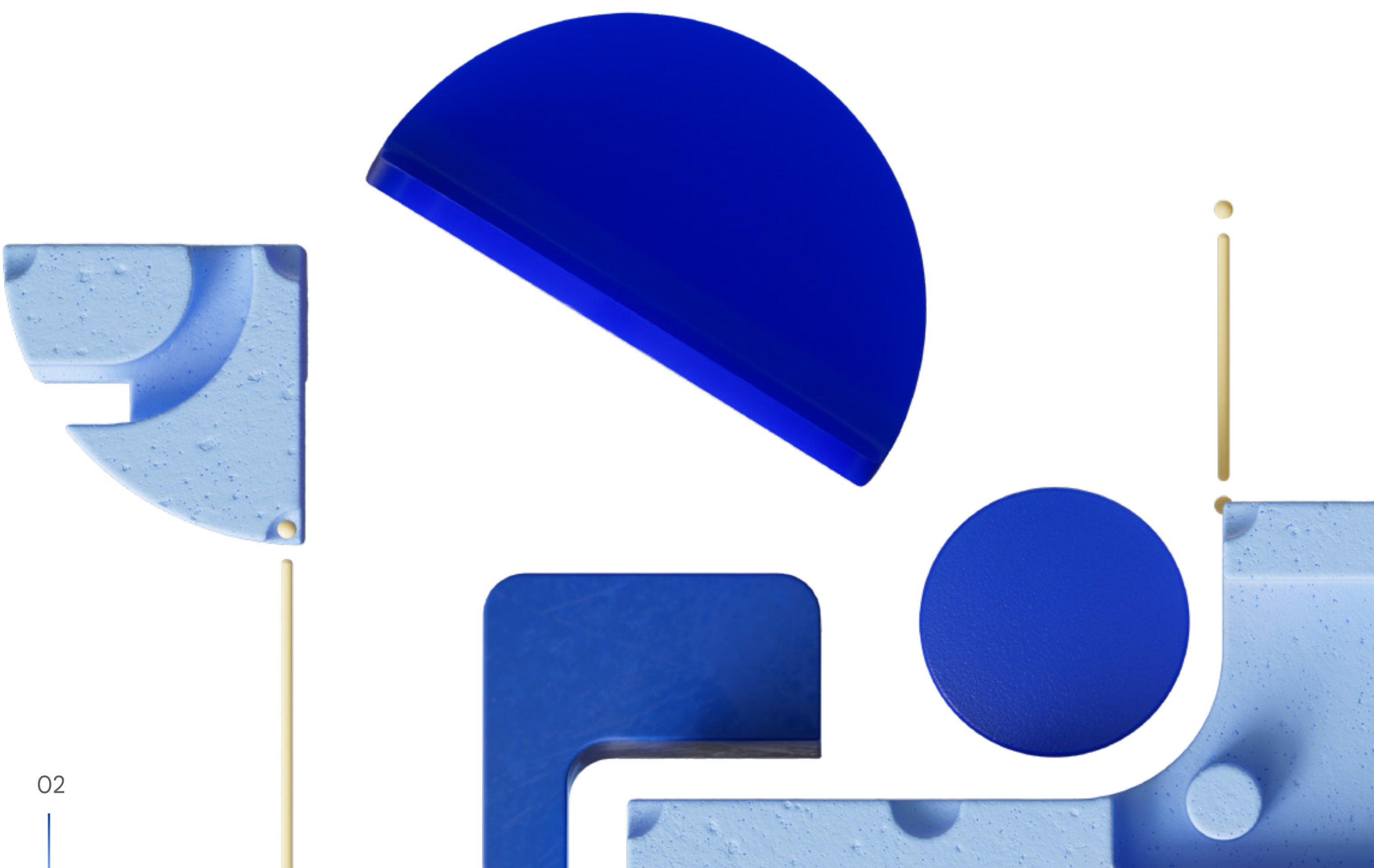
By Martin Gonzalez

Head of Organizational AI Research,
Google DeepMind

Foreword by Simon Bouton

Chief Experience Officer,
Google DeepMind

A special thanks to Muhua Huang, Stanford Graduate School of Business doctoral researcher, whose partnership was valuable in the conceptualization and research of this research primer.



Executive Summary: 16 big ideas that will redefine how we build, lead, and scale

Organizing Intelligence offers a glimpse into the frontier of AI and organizations, where some ideas are deliberately provocative and others are still taking shape.

This snapshot provides a panoramic view of the rich spectrum of research happening at the intersection of computer and management science, organized into three parts:

Part 1
AI for Organizations

Explores research pushing the boundaries of AI as a transformative tool

Page 07

1: The Generality Frontier: Moving from narrow to general intelligence

The transition from narrow AI to artificial general intelligence (AGI) marks a shift from using task-specific tools to working with systems that are at least as capable as humans at most cognitive tasks, including the ability to independently learn entirely new skills.

2: Scaling AI Autonomy: The promise of AI agents and how to govern them

The world is about to experience a transition from low-functioning assistants to highly autonomous agents. What are the trade-offs for top-down governance and the likely impact of human-AI collaboration models?

3: AI Interpretability: Why we need a shared machine-human vocabulary

As AI begins to exceed human intelligence, organizations must move beyond trusting AI output to actively learning from the insights these models generate.

4: Embodied Intelligence: Advances in robotics for the workplace

Take a deep dive into the computational complexity of human movement, the shift toward unified reasoning models, and the future of hybrid workforces with intelligent robots.

5: Invisible Infrastructure with Ambient AI: Computing that recedes into the background of daily life

The original promise of ubiquitous computing is advancing as hardware and AI become seamless extensions of our environment. Here, we examine how new breakthroughs are making the vision of invisible infrastructure possible.

Part 2
AI in Organizations

Summarizes sociotechnical research on the psychological, social, and structural impact of AI

Page 21

6: Deep AI Adoption: Advanced strategies for adopting AI and the risk of work intensification

Successful AI adoption requires shifting from simple task substitution to a product management mindset focused on redesigning entire workflows. Read about the learning tax of new tools and the phenomenon of work intensification.

7: The Expansion and Erosion of Human Skills: How the demand for human skills is shifting

New research reveals how AI can drive skill expansion rather than erosion. Learn how AI is altering talent-market skill demands, the dangers of cognitive surrender, and how to protect against skill erosion.

8: The Future of Jobs: Despecialization and how new jobs are born

Rather than simply automating existing roles, AI will likely redraw specialist boundaries and create entirely new types of jobs. We explore evidence on despecialization and the pathways through which new jobs are created.

9: AI-driven Coordination: AI's impact on organizational life

Research into AI collaboration is coalescing around three areas: social cognitive capabilities of AI, the impact on human coordination, and the efficacy of human-AI hybrid teams.

10: AI for Strategic Decisions: Challenges in implementing AI-powered decision tools

Human bias and rigid structures may stifle AI's usefulness in supporting decisions. We explore three key challenges for leaders: algorithmic aversion, decision tools clashing with the human org chart, and the risk of a monoculture.

11: The Psychology of Work: How professional identity and purpose are shifting

Recent studies are shedding light on how professionals are navigating the psychological shifts brought about by AI, and how they are finding economic viability while preserving pride in their craft.

Part 3
AI as Organizations

Examines the use of AI simulacra as a powerful analytical instrument to decode complex patterns in our organizations

Page 34

12: Generative Simulations: How organizations can rehearse alternative futures

In a niche but growing field, AI simulations deliver benefits over simple modeling. Recent studies have created high-definition mirrors of our organizations.

13: The Simulation of Individuals: How to capture the richness of a human life

Researchers are exploring how to create digital doppelgängers and high-fidelity simulations of individuals. But there's an ongoing debate about the reliability of LLMs for high-stakes human decision-making.

14: Simulating Group Dynamics: Organizations can replicate team dynamics to anticipate real-world risks

Structured AI simulations can now predict team friction points and create virtual rehearsals that bypass the costs of broken chemistry in scenarios like corporate mergers or R&D sprints.

15: Simulating Organizational Behavior: Emergent collective actions at scale

Organizational success often hinges on invisible, emergent interactions that leaders cannot plan. Large-scale simulations uncover how thousands of interactions coalesce into collective behavior.

16: Ethical Frontiers: Simulations present unique questions of misuse

Using AI simulacra in organizations brings about ethical concerns, including consent, overreliance, ownership and manipulation.

Foreword

Organizing Intelligence is a research primer inspired by the AI for Organizations Grand Challenge, an initiative launched in late 2025 by the Stanford Institute for Human-Centered AI and Google DeepMind, two institutions at the forefront of AI development.

We asked a foundational question to the computer and management science communities: What bold research must we undertake to ensure AI acts as a transformative, positive force within our organizations? The challenge inspired scholars from more than 150 universities worldwide to submit proposals, a strong signal of a community ready to shape the future of work.

Our AI researchers are excited to help some selected teams bring their ideas to life, using our organization as their field site and data source. Alongside the judging panel—which includes some of my colleagues from Google DeepMind and professors from Harvard University, Stanford University, UC Berkeley, the University of Chicago, the University of Toronto, and Yale University —we are proud to play a part in driving valuable management research that will benefit the world in the coming years.

At Google DeepMind, our research touches everything from scientific discovery through work on groundbreaking initiatives like AlphaFold and WeatherNext, to industry-leading AI models, such as Gemini and Veo.

Yet, building powerful technology is only half the mandate. Our mission is to build AI responsibly to benefit humanity, including organizations across every sector. We are committed to a future where AI reshapes the workplace to reflect our highest ideals as a society.

This primer captures the breadth of research already in motion, providing a rare overview for business leaders. We look forward to the continued research emerging from this intersection to guide the evolution of our organizations.

Our deepest gratitude goes to Stanford University for a fruitful and enduring collaboration.

Simon Bouton

Chief Experience Officer,
Google DeepMind

Introduction: The Frontier of AI and Organizations

I spend a lot of my time filtering through AI noise, parsing between what is actually changing how we work and what is just clever marketing.

While generally skeptical of hucksters and hype artists, I'm increasingly convinced that AI will rewrite the core assumptions for how we build, lead, and scale our organizations. My guess is that a few decades from now, we'll look back at this period as a watershed moment for organizational research.

I think we are entering a period of what philosopher Thomas Kuhn called *extraordinary science*. In his work *The Structure of Scientific Revolutions*, Kuhn noted that most fields move in small, safe increments until they hit a wall where the old rules stop working. This creates a paradigm shift that forces a fundamental rethink of the entire foundation.

Demis Hassabis has described this transition as having "probably ten times the impact the industrial revolution had, ten times faster." Since the entire field of modern organizational research was essentially a response to the industrial revolution, it seems inevitable that we will see a rebirth of the field to meet this new reality.

While research in this field is burgeoning, it is also highly fragmented. Busy executives managing transformation initiatives in their companies rarely have an easy way to see across this fascinating landscape. Fragmentation is a hallmark of extraordinary science; when the old rules break down, new ideas scatter in a thousand directions.

My team and I created this primer to give a panoramic view of that scatter. We intentionally distill multidisciplinary insights into an accessible format. Some of these findings are so new, they're still sitting in preprint repositories like arXiv.org or the Social Science Research Network (SSRN), waiting for peer review. Any flattening of nuance in these summaries is intended to make them readable for a broad, busy audience and is not a disregard for the depth of the original work.

It's easy to get caught up in AI grandiosity. So to avoid a prophetic tone, we don't treat any one study like it's the absolute final word on these big AI questions. Instead, we pull from several sources, at times allowing them to offer counterpoints to one another. While this collection isn't exhaustive—we have surely missed some high-quality research—it represents an educated look at the emerging frontier.

My hope is that these ideas spark a conversation, so we can responsibly shepherd our organizations through this period of extraordinary science.

Martin Gonzalez

Head of Organizational AI Research,
Google DeepMind

Part 01

AI for Organizations

This section explores research pushing the boundaries of AI as a transformative tool and collaborator at work. It offers a view into the evolution of AI from a specialized system into a generalist force, exploring the frontier where artificial general intelligence, high-autonomy agents, embodied AI, and ambient intelligence begin to weave into the very fabric of our physical and social workplaces.

Idea 1

The Generality Frontier: Moving from narrow to general intelligence

TL;DR: The transition from narrow AI to artificial general intelligence (AGI) marks a shift from using task-specific tools to working with systems that are at least as capable as humans at most cognitive tasks, including the ability to independently learn entirely new skills.

Some of the most forward-looking research in AI centers on artificial general intelligence (AGI). But what is it, and how might organizations benefit from it?

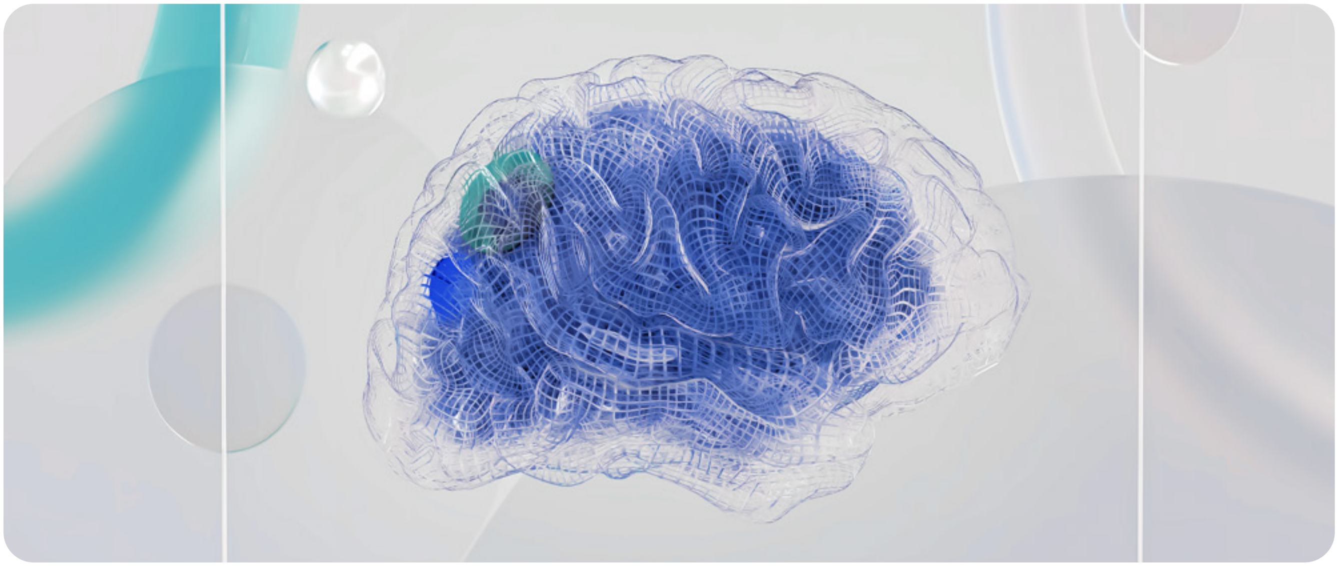
To understand where AI development is headed, let's first consider *narrow intelligence*, which describes the AI deployments we see in the world today. Narrow intelligence systems derive capabilities from specific training: Chatbots that help with language-based tasks are trained on a corpus of natural language, while image generators learn from massive visual datasets. Narrow intelligence is incredibly powerful already, and its progress has been breathtaking. AlphaFold helps researchers find cures for rare diseases by predicting protein structures for drug discovery, Waymo is pioneering autonomous driving with safety rates significantly higher than human drivers, and WeatherNext boosts our ability to predict typhoon pathways and flood risk to help save lives.

But there is more on the horizon. *General intelligence* systems are distinguished by their breadth and ability to cover a wide range of tasks, including metacognitive abilities like learning new skills. Such an AI, at least as capable as humans at most cognitive tasks, could be here within the coming years. Integrated with agentic capabilities, AGI could supercharge systems to understand, reason, plan, and execute actions autonomously.

Google DeepMind researchers have proposed a performance gradient for these systems that's benchmarked against human capabilities, ranging from Level 1 *emerging* performance to Level 5, or *superhuman* performance. While we already see Level 5 superhuman narrow intelligence in systems like AlphaFold, we don't have consensus on whether the recently launched Gemini 3 Deep Think model, or its peers from other frontier labs, has achieved Level 2, or *competent* AGI.

	Narrow AI Clearly scoped task or set of tasks	General AI Wide range of non-physical tasks, including metacognitive tasks like learning new skills
Level 1: Emerging Equal to or somewhat better than an unskilled human	Simple rule-based systems	ChatGPT, Bard, Llama 2, Gemini
Level 2: Competent At least 50th percentile of skilled adults	Toxicity detectors, smart speakers such as Siri (Apple), Alexa (Amazon), Google Assistant, Watson (IBM)	**No consensus of achieving competent AGI as of early 2026**
Level 3: Expert At least 90th percentile of skilled adults	Spelling & grammar checkers such as Grammarly, generative image models	**Not yet achieved as of early 2026**
Level 4: Exceptional At least 99th percentile of skilled adults	Deep Blue (chess), AlphaGo (go)	**Not yet achieved as of early 2026**
Level 5: Superhuman Outperforms 100% of humans	AlphaFold (protein structure prediction), AlphaZero (chess, shogi, go), StockFish (chess)	**Not yet achieved as of early 2026**

From Morris, M. R., et al. (2024). Levels of AGI for operationalizing progress on the path to AGI. Proceedings of ICML 2024. https://doi.org/10.48550/arXiv.2311.02462



The definition of AGI and criteria for identifying it remain subjects of ongoing debate. Key leaders across frontier labs say the path to AGI is not predetermined by computational resources alone; fundamental algorithmic, data, and compute breakthroughs are still required. Several critical factors will shape this transition:

1. The jagged frontier

Current models often exhibit uneven performance, appearing as experts in coding while remaining only *emerging* in factual reasoning. A system only reaches a higher AGI level once it meets the performance threshold across most tasks.

2. Innovation vs. diffusion

Achieving AGI in a lab is only the first step; significant barriers to adoption mean it will take time for this technology to benefit users and organizations, as we're seeing with narrow intelligence systems today.

3. Promises and risks

While the potential is vast, AGI also introduces serious new risks. These could include misuse by bad actors and misalignment, when an AI system takes actions it knows the developer didn't intend. For an extended discussion on these risks, see the referenced articles below.

In contrast to narrow AI, which functions as a digital tool, AGI will be powerful enough to act as a digital collaborator. AGI could help lower barriers to innovation and creativity by enhancing information processing within organizations. It could enable startups to tackle complex challenges previously only addressable by large, well-funded institutions as AGI democratizes access to advanced tools and knowledge.

Learn more

Burnell, R., et al. (2026). *Measuring progress toward AGI: A cognitive framework*. Google DeepMind. https://blog.google/innovation-and-ai/models-and-research/google-deepmind/measuring-agi-cognitive-framework/

Hassabis, D. (2025, July 23). *Demis Hassabis: Future of AI, simulating reality, physics and Games | Lex Fridman Podcast #475* [Video]. YouTube. https://www.youtube.com/watch?v=-HzgcbRXUK8

Legg, S. (2025, December 11). *The arrival of AGI | Shane Legg (co-founder of DeepMind)* [Video]. YouTube. https://www.youtube.com/watch?v=l3u_FAv33G0

Morris, M. R., et al. (2024). Levels of AGI for operationalizing progress on the path to AGI. *Proceedings of ICML 2024*. https://doi.org/10.48550/arXiv.2311.02462

Shah, R., et al. (2025). *An approach to technical AGI safety and security*. arXiv. https://doi.org/10.48550/arXiv.2504.01849

Idea 2

Scaling AI Autonomy: The promise of AI agents and how to govern them

TL;DR: The world is about to experience a transition from low-functioning assistants to highly autonomous agents. What are the trade-offs for top-down governance and the likely impact of human-AI collaboration models?

> Diversify our raw material sourcing to reduce our carbon footprint by
> **25%**
> without increasing costs.

No AI agent available in early 2026 can fully execute a prompt like this. A group of agents would need to research the global supplier market, solicit bids via APIs (a secure digital handshake that allows systems to communicate without exposing their internal architecture), and negotiate contract terms, all while balancing sustainability scores against procurement costs.

But soon highly capable agents will take a user's goals, independently plan and execute multistep strategies, and engage in social and economic actions before reporting back to a human operator after days or weeks. What separates current systems from truly agentic systems is their autonomy and capacity for complex, long-horizon tasks.

One way of understanding the magnitude of the transition is to think about the differences between low- and high-functioning AI agents. While low-functioning systems operate as specialized tools requiring human oversight for every input and output, high-functioning agents navigate complex, multistep projects independently. Tomorrow's systems will move beyond automating simple workflows to engage in sophisticated coordination.

	Low-functioning AI agents	High-functioning AI agents
Autonomy	**High within specific tasks** User in the loop for input, throughput, and output	**High across multi-step, complex tasks** User in the loop for goal setting, with the option to manage specific steps like economic transactions
Task Complexity	**Low** Handles singular task flows	**High** Tasks often involve coordination across multiple systems
Cross-agent Collaboration	**Independent**	**Multi-agent information sharing, collaboration, and negotiation**
Anthropomorphism	Human voice and low-definition features like animated visuals or assigned human roles like customer service or travel agent	Option for high-definition human features, including personality, belief systems, and morals Can include user clones
Example Applications	Customer service chatbots, virtual assistants, automated workflows	Research and development, project management, negotiations, virtual chiefs of staff, individual simulacra

Now is the time for organizational leaders to consider the strategic implications of agentic systems. A foundational paper, "Virtual Agent Economies," suggests asking two questions:

Origins

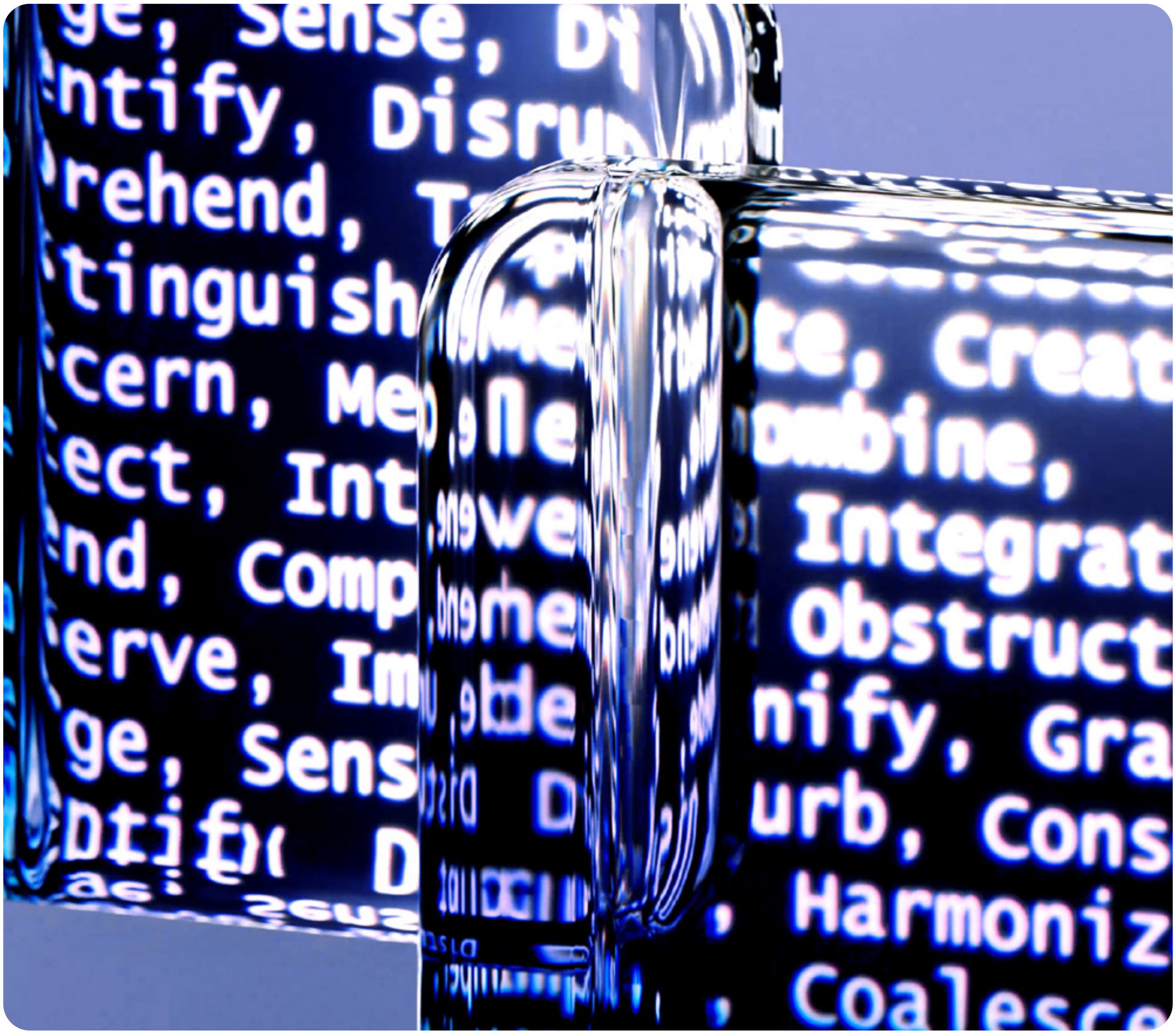
Will AI agents emerge through decentralized, scattered efforts by employees, or will firms design them at an enterprise level?

An emergent approach risks fragmented data governance but allows for high levels of job *crafting*, where employees offload repetitive tasks to retain work they find meaningful. Conversely, top-down systems offer better interoperability and data protection. However, because these systems often tether the firm to a single base model, they may trigger a collapse of diverse thinking—a phenomenon already surfacing in current research.

Degree of separateness

Should agents operate within the human organization, where collaboration is built into the daily flow of work, or should they remain separate, akin to an autonomous exchange governed by strict protocols?

Embedding agents in the human flow creates interface challenges like adoption challenges and *sub-goal optimization*, where employees deep in the organization may tend to prioritize local goals over the firm's global goals. Separating AI agents can mitigate these frictions; however, keeping humans entirely out of the loop risks a deskilling effect, which could have enterprise-level consequences if systems fail and human operators lack the necessary knowledge to troubleshoot them.



Learn more

Tomašev, N., et al. (2025). *Virtual agent economies*. arXiv. https://doi.org/10.48550/arXiv.2509.10147

Hadfield, G. K., & Koh, A. (2025). *An economy of AI agents*. arXiv. https://doi.org/10.48550/arXiv.2509.01063

Sapkota, R., Roumeliotis, K. I., & Karkee, M. (2025). AI Agents vs. Agentic AI: A Conceptual taxonomy, applications and challenges. *Information Fusion*, 126, 103599. https://doi.org/10.1016/j.inffus.2025.103599

Tomašev, N., Franklin, M., & Osindero, S. (2026). Intelligent AI delegation. arXiv. https://doi.org/10.48550/arXiv.2602.11865

Wang, Z. Z., et al. (2025). *How do AI agents do human work? Comparing AI and human workflows across diverse occupations*. arXiv. https://doi.org/10.48550/arXiv.2510.22780

Idea 3

AI Interpretability: Why we need a shared machine-human vocabulary

TL;DR: As AI begins to exceed human intelligence, organizations must move beyond trusting AI output to actively learning from the insights these models generate.

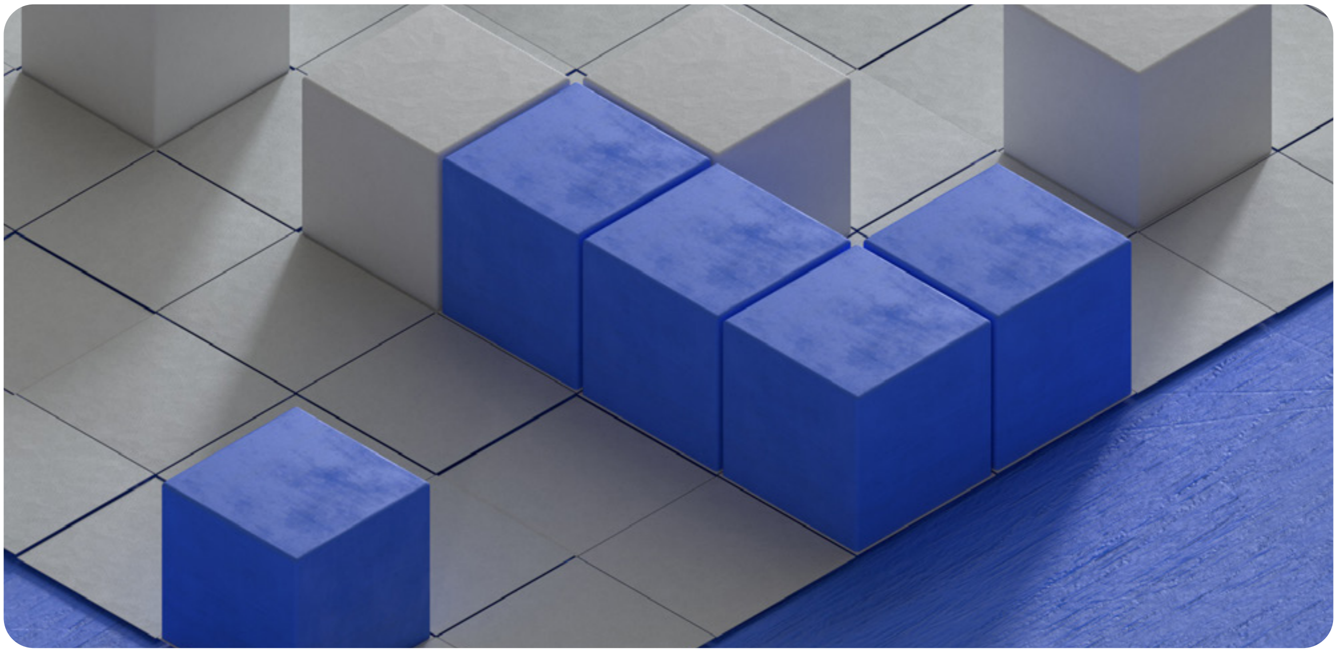
In 2024, four of the world's best chess players—all former or current world champions—sat down not to compete, but to learn. Their teacher was AlphaZero, an AI that mastered chess entirely through self-play, without studying a single human game. The researchers behind the experiment extracted new strategic concepts from AlphaZero's internal representations and presented them to the grandmasters as puzzles.

The results were striking: All four players improved. The concepts they encountered defied conventional chess wisdom. For example, sacrificing the most powerful piece on the board for a subtle long-term advantage, or abandoning a promising attack to quietly reposition across the board. One grandmaster called the ideas "clever but not natural."

If the top experts in a well-understood domain can learn something new from a machine, it raises an important question: As we begin to rely on AI for hiring, strategic, or investment decisions—messier domains where data is noisy and outcomes aren't clear—will the machine be able to tell us what patterns it sees?

AI's ability to recommend is outpacing our ability to understand why. WeatherNext 2, for example, delivers 15-day forecasts with the probabilistic detail that conventional weather models sustain for only 9 to 10 days. Traditionally, weather science typically gains one day of accuracy per decade; this single release represents 50 years of progress. Yet, explaining *how* it reaches a prediction is nearly impossible. Is it wind velocity, the tide, or just random noise?

As scholars James Evans and Eamon Duede recently argued in *Science*, we are witnessing a curious inversion: Our capacity to control nature through AI is outpacing our ability to understand the technology. Organizations are adopting AI tools that deliver strong outcomes, but the reasoning behind those outcomes remains opaque. It's a classic machine learning tradeoff: As predictive power increases, explainability of the results decreases. Modern AI amplifies the complexity and consequences of this known tradeoff. "Trust the output" works until model errors carry real consequences, or someone asks you to explain the AI's decision to a board, a regulator, or a patient.



For these reasons, interpretability is becoming not only a technical concern, but also an organizational one. Future models will need to achieve three levels of interpretability:

1. Humans must be able to understand what the AI is doing.
2. Users must be able to evaluate whether a recommendation fits their context. These are judgment calls that humans must make.
3. At the most ambitious level, people will learn from AI, in the way the grandmasters did.

Reaching the highest level demands new approaches. Research from Google DeepMind suggests that AI concepts don't map neatly onto human words and we need a new vocabulary for humans and machines to communicate. The same paper proposed that the future of interpretability won't be static dashboards, but interactive dialogue, through which AI adapts its explanations to reflect what you already understand, much like a skilled colleague walking you through their reasoning.

Organizations that thrive in the coming decade won't simply deploy AI aggressively; they will also learn from it effectively. Leaders need to treat interpretability as a bridge between machine capability and human judgment, not a checkbox. The chess grandmasters didn't just check AlphaZero's answers. They expanded what they knew. That's the opportunity, and the imperative, for every organization.

Learn more

Evans, J., & Duede, E. (2025). After science. *Science*, 390(6774), eaec7650. https://doi.org/10.1126/science.aec7650

Hewitt, J., Geirhos, R., & Kim, B. (2025). *We can't understand AI using our existing vocabulary*. arXiv. https://doi.org/10.48550/arXiv.2502.07586

Kim, B., et al. (2025). *Because we have LLMs, we can and should pursue agentic interpretability.* arXiv. https://doi.org/10.48550/arXiv.2506.12152

Schut, L., et al. (2025). Bridging the human–AI knowledge gap through concept discovery and transfer in AlphaZero. *Proceedings of the National Academy of Sciences, USA, 122*(13), e2406675122. https://doi.org/10.1073/pnas.2406675122

Idea 4

Embodied Intelligence: Advances in robotics for the workplace

TL;DR: Take a deep dive into the computational complexity of human movement, the shift toward unified reasoning models, and the future of hybrid workforces with intelligent robots.

It's hard to talk about robotics without mentioning Hans Moravec, a robotics pioneer from Carnegie Mellon University. He is best known for an astute observation, now famously called Moravec's paradox, which explains why AI progress appears remarkably advanced in software yet remains comparatively slow in robotics.

Moravec says what seems easy for computers is hard for humans, and what is easy for humans is hard for computers. High-level functions like calculus, chess, and financial modeling require significant cognitive load and training for humans, yet are relatively simple for machines. Conversely, basic actions like walking on uneven surfaces, tying a shoelace, or picking up an M&M are computationally demanding for robots. Moravec attributes this to evolutionary biology. Sensorimotor skills are the result of billions of years of "R&D" through natural selection, whereas abstract logic is a recent, rule-based development that is far easier to translate into computer code.

Recently, however, the robotics industry has experienced a major resurgence as we move beyond rigid, pre-programmed machines toward generalist embodied AI systems that can think and act simultaneously. This shift is driven by Vision-Language-Action (VLA) models, which bridge the gap between high-level reasoning and low-level motor execution by unifying vision, language, and action data into a single framework. VLA models allow robots to understand complex human instructions in context. Think about switching from a train on a fixed track to a driver who can see the road, listen to directions, and navigate a city dynamically. Robots finally have a brain that connects perception directly to movement, rather than having to follow a rigid script.

What lies behind our fascination with humanoid robots? From the moment the word was coined in the 1920s by Czech writer Karel Čapek, robots were envisioned as human-like creatures designed for labor. Little did he know that the field would persist with the humanoid incarnation—as opposed to just robotic hands or vehicular form factors.

Humanoid robots serve a dual purpose: They provide a familiar interface for human coworkers and they act as advanced research platforms. These platforms allow scientists to integrate AI tools and further develop the robot's thinking capabilities. True human learning does not precisely mirror the way large language models function. For instance, a child learns that a pot of boiling water is hot through the direct sensory response of temperature and pain. Many prominent thinkers believe that providing AI with the ability to learn physics via a humanoid body brings us significantly closer to AGI.

What does this mean for organizations? Embodied AI won't be restricted to the manufacturing floor or the warehouse for long. As it integrates into our workplaces, it's important to prepare for the inevitable collaboration challenges that AI-human hybrid teams will bring. (See idea 9: *AI-driven Coordination*.)

Learn more

Gemini Robotics Team. (2025). Gemini Robotics 1.5: *Pushing the Frontier of Generalist Robots with Advanced Embodied Reasoning, Thinking, and Motion Transfer*. arXiv. https://doi.org/10.48550/arXiv.2510.03342

Sermanet, P., et al. (2025). *Generating robot constitutions & benchmarks for semantic safety*. arXiv. https://doi.org/10.48550/arXiv.2503.08663

Moravec, H. (1988). *Mind children: The future of robot and human intelligence*. Harvard University Press. https://www.hup.harvard.edu/books/9780674576186

Kawaharazuka, K., et al. (2025). *Vision-Language-Action Models for Robotics: A Review towards Real-World Applications*. IEEE Access, 13, 162467–162504. https://doi.org/10.1109/access.2025.3609980

Fukuda, T., Dario, P., & Yang, G. (2017). Humanoid robotics—History, current state of the art, and challenges. *Science Robotics*, 2(13), eaar4043. https://doi.org/10.1126/scirobotics.aar4043

Idea 5

Invisible Infrastructure with Ambient AI: Computing that recedes into the background of daily life for the workplace

TL;DR: The original promise of ubiquitous computing is advancing as hardware and AI become seamless extensions of our environment. Here, we examine how new breakthroughs are making the vision of invisible infrastructure possible.

As soon as the personal computer was invented in the 1970s, the creators of the first desktop computer at Xerox's Palo Alto Research Center (PARC) envisioned a new paradigm of ubiquitous computing, also known as pervasive or ambient computing. The Xerox research pioneers imagined a time when computers became invisible, receding into the background of daily life through interactions completely natural to humans.

Today, we hold mobile phones below our line of sight. Our heads tilt down and fingers contort, just to keep from dropping the device. In contrast, the Xerox PARC researchers imagined any solid structure—a table, a wall, or a window pane—could be a computer. Interconnecting everyday objects would turn our entire environment into an interactive interface. Rather than carrying a laptop to meeting rooms or coffee shops, you could arrive anywhere and access your data and apps on any available surface. That vision spurred the wave of technologies broadly clustered under the Internet of Things (IoT), which surged in the early 2010s. The IoT era brought smart home devices, Wi-Fi, Bluetooth, cloud computing, and wearables like fitness trackers and smartwatches, with the smartphone used as a universal remote control. But while these connected devices were nothing short of magical, they still did not fully capture the vision of ubiquitous computing. Most lacked intelligence, interactions remained awkward with idiosyncratic voice commands, and plastic boxes sat on tabletops, rarely unplugged from a power source. Generative AI, high-density batteries, and nanosized components like cameras or microdisplays were not yet commercially viable. Neither were neural interfaces or unconventional computing surfaces like smart wood, clothing, or liquid metal. Today, all of this is available and advancing at a blistering pace.

Three ubiquitous computing technologies are pushing toward maturity and commercial availability: wearables, silent speech interfaces, and neural interfaces.

A new era of wearables

Beyond today's smartwatches and rings, a constellation of smart clothing, footwear, and jewelry is emerging. Breakthroughs like Large Sensor Models (LSMs) allow these devices to bridge data gaps when off-body and intuitively understand your activity without manual input. This living wardrobe can detect a developing illness days before a sniffle, recalibrate your shoes mid-stride to save your joints from an injury you haven't even caused yet, or manipulate the digital world without touching a screen—like tapping your jacket to check the weather forecast.

Silent speech interfaces

These interfaces push wearables further by letting you converse with computers without making a sound. Pioneered by MIT Media Lab's AlterEgo, the device wraps around the ear and jaw to intercept faint electrical impulses sent to your vocal muscles as you silently articulate words. By decoding these vibrations, you can message friends or search the web in a crowded room unnoticed. The system whispers back via bone conduction, sending crystal-clear audio directly to your inner ear. This creates a private loop between your internal monologue and the internet, allowing you to navigate the digital world at the speed of thought. While transformative for everyday use, these devices also serve critical medical roles, such as restoring speech for people with neuromuscular disorders.

Neural interfaces

These devices translate the brain's electrical activity—which leaks faintly through the bone of the skull—into digital commands via advanced sensors on the head. For a designer or engineer, for instance, neural interface technology offers the ability to manipulate 3D models or navigate complex datasets by visualizing the movement, effectively turning thoughts into input as today's keyboard and mouse do. Use cases already exist where these interfaces manipulate real-world objects designed to interact with them; in a future with ambient computers, many more things could be controlled through thought.

While many of these profound developments are built for medical use cases, it isn't difficult to see what future workplaces might look like. Ambient AI will cause a shift from fixed workstations to high-bandwidth environments, in which the friction between intent and execution disappears. By replacing physical peripherals with direct neural or muscular inputs, professionals will manage data and collaborate in real time without being tethered to a desk or screen. Workers will no longer be encumbered by human speech, which is linear and low-bandwidth. An invisible infrastructure will allow us to build and solve problems at the speed of thought, fulfilling the promise of ubiquitous computing by letting the hardware vanish into the background.

Learn more

Narayanswamy, G., et al. (2024). *Scaling wearable foundation models*. arXiv. https://doi.org/10.48550/arXiv.2410.13638

Lee, J. Y., et al. (2025). Brain–computer interface control with artificial intelligence copilots. *Nature Machine Intelligence, 7,* 1510–1523. https://doi.org/10.1038/s42256-025-01090-y

Weiser, M. (1999). The computer for the 21 st century. *ACM SIGMOBILE Mobile Computing and Communications Review*, 3(3), 3–11. https://doi.org/10.1145/329124.329126

Olwal, A., & Dementyev, A. (2022). Hidden Interfaces for Ambient Computing: Enabling Interaction in Everyday Materials through High-brightness Visuals on Low-cost Matrix Displays. *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, 1–20*. https://doi.org/10.1145/3491102.3517674

Pearl, D., Intriligator, J., & Liu, X. (2025, February 9). *Seamless Integration: The evolution, design, and future impact of wearable technology*. arXiv. https://doi.org/10.48550/arXiv.2502.05797

Kapur, A., Kapur, S., & Maes, P. (2018). AlterEgo: A personalized wearable silent speech interface. *Proceedings of the 23rd International Conference on Intelligent User Interfaces (IUI '18), USA*, 43–53. https://doi.org/10.1145/3172944.3172977

Jude, J. J., et al. (2026). Restoring rapid natural bimanual typing with a neuroprosthesis after paralysis. *Nature Neuroscience*. https://doi.org/10.1038/s41593-026-02218-y

Part 02

AI in Organizations

This section summarizes sociotechnical research on the psychological, social, and structural evolution of AI at work. It explores the complex effects of the de-routinization of tasks, the crafting of new professional identities, and AI-augmented human decision making and coordination. From the mechanisms of new job creation to the rising primacy of meta-skills, these studies provide a roadmap for navigating the transition from simple tool adoption toward a transformative and human-centric shift in our organizations.

Idea 6

Deep AI Adoption: Advanced strategies for adopting AI and the risk of work intensification

TL;DR: Successful AI adoption requires shifting from simple task substitution to a product management mindset focused on redesigning entire workflows. Read about the learning tax of new tools and the phenomenon of work intensification.

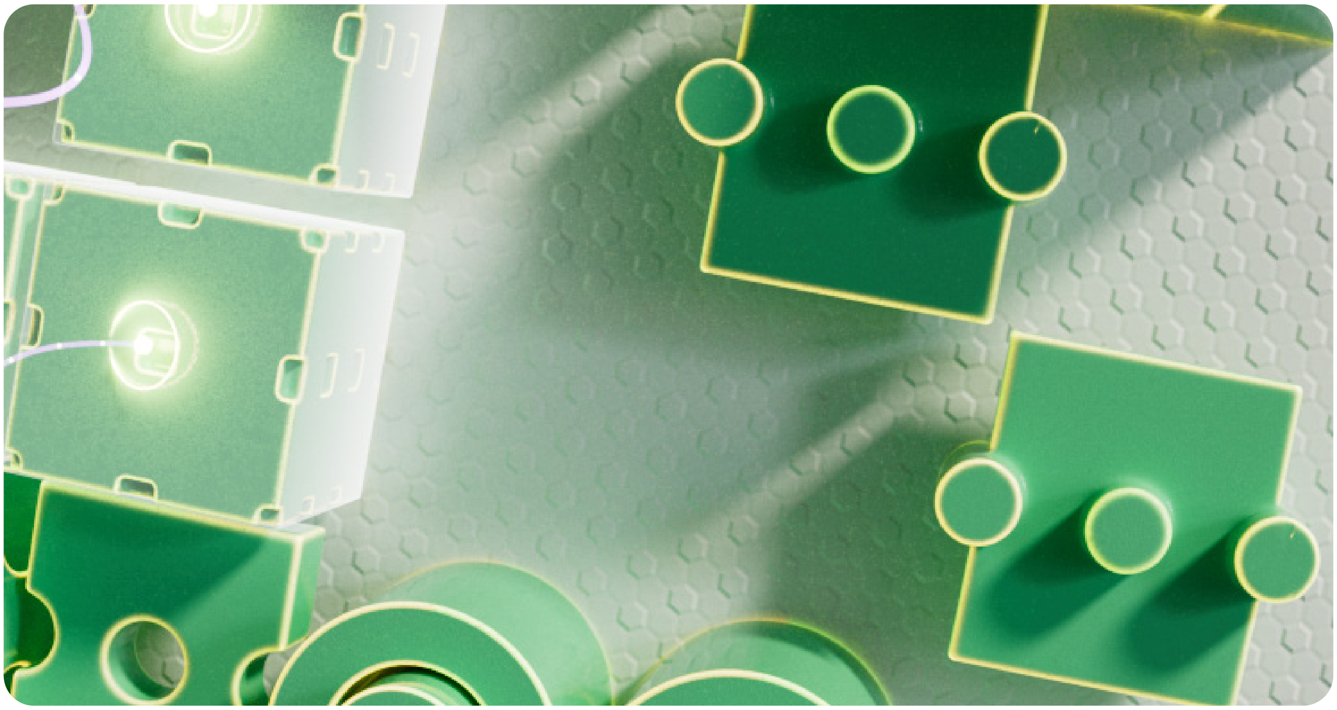
Why do some people become enthusiastic, sustained adopters of AI, while others walk away? An 18-month Stanford study at Google revealed that although most people wanted to find value in AI, many remained stuck in simple substitution. They swapped existing tasks for AI outputs, only to find the learning tax was often greater than the payoff. Crucially, the researchers found that successful adopters didn't just focus on prompt engineering or its more recent sibling, context engineering. Instead, deep AI adopters completely changed how they approached AI—taking inspiration from product management.

As an all-purpose technology, generative AI is like a Swiss Army knife packed with non-obvious functions. A product management mindset helps the user decide which tool is best for the job. Successful AI users identify blockers, understand the technology, and find a strong fit between the two. They focus on redesigning entire workflows rather than seeking quick point solutions.

You might think AI makes work easier and frees up the capacity of the team, but a longitudinal UC Berkeley study found employees who adopted AI saw their work intensify. The intensification grew from the employees' initiative—the study noted that AI use was not mandated—because "AI made 'doing more' feel possible, accessible, and in many cases intrinsically rewarding." The study identified three forms of intensification:

1. **Task expansion** happened when employees stepped into responsibilities previously handled by others because AI made complex tasks feel accessible. This capability reduced dependence on colleagues and allowed individuals to absorb work that might have required additional headcount before. (See idea 7: *The Expansion and Erosion of Human Skills*)

2. **Blurred work/nonwork boundaries** occurred when AI reduced the friction of facing a blank page or an unknown starting point. Employees often slipped the quick AI prompt into moments typically used for breaks, in part because the conversational interface felt more like chatting with a colleague and less like work. This new habit created a workday with fewer pauses and led to more continuous involvement.



3. **Increased multitasking** developed as workers managed multiple active threads or ran parallel agents while waiting for AI outputs. While this rhythm provided a sense of momentum, it demanded frequent attention switching. Constant juggling increased cognitive load, even as AI provided partial offloading within individual tasks.

What it means for organizations:

Start with the work, not AI: The most transformative path to deep adoption begins with a discovery phase to identify which tasks AI can improve. After identifying these high-value problems, teams can start tinkering with solutions.

Build a tinkering culture: Leaders who focus solely on corporate subscriptions that protect data privacy and IT approvals for new tools will miss the mark. They need to demonstrate a credible commitment to the technology; mandates will fail if leaders don't experiment with the new systems themselves.

Watch out for the self-sufficiency spiral: As AI enables more solo and self-contained work, it reduces the need for human collaboration. Organizations must avoid an "alone together" dynamic, where employees no longer engage with their colleagues. A lack of genuine interaction can lead to psychological and cultural damage for both individual workers and organizations.

Learn more

Pratt, A., & Valentine, M. (2026, February 3). To drive AI adoption, build your team's product management skills. *Harvard Business Review*. https://hbr.org/2026/02/to-drive-ai-adoption-build-your-teams-product-management-skills

Ranganathan, A., & Ye, X. M. (2026, February 9). AI doesn't reduce work—It intensifies it. *Harvard Business Review*. https://hbr.org/2026/02/ai-doesnt-reduce-work-it-intensifies-it

Gonzalez, M. (2026, March 19). Five strategies for deeper AI adoption at work. *The Keyword*. https://blog.google/company-news/inside-google/life-at-google/strategies-to-adopt-ai-at-work/

Reif, J. A., Larrick, R. P., & Soll, J. B. (2025). Evidence of a social evaluation penalty for using AI. *Proceedings of the National Academy of Sciences, USA, 122*(19), e2426766122. https://doi.org/10.1073/pnas.2426766122

Idea 7

The Expansion and Erosion of Human Skills: How the demand for human skills is shifting

TL;DR: New research reveals how AI can drive skill expansion rather than erosion. Learn how AI is altering talent market skill demands, the dangers of cognitive surrender, and how to protect against skill erosion.

Empathy, creativity, judgment? Which skills will be valued in the future? Many leaders are worried about the erosion of skills in the workplace; however, new research suggests a path toward skill expansion instead of decline.

A recent study from INSEAD and the London Business School found AI is shifting labor market demand toward higher-order cognitive abilities. An analysis of 8.4 million job postings showed that roles adopting AI saw a 44% increase in demand for cognitive skills like critical thinking, alongside a 4.5% decrease in social skills.

This divergence varies depending on whether AI automates or augments a job. According to another study from Harvard University and Hong Kong University of Science and Technology, automation-prone roles saw a 17% decline in hiring, while augmentation-prone roles experienced a 22% surge. These thriving roles require greater AI literacy, the ability to validate algorithmic outputs, and experience applying contextual judgment. While these two studies offer a snapshot of how employers are grappling with AI today, they are not the final word.

As the landscape continues to shift, this evolution raises a bigger question: What happens to the individuals in these roles, and will their skills inevitably erode? Researchers from The Wharton School have identified a major risk in this transition, called cognitive surrender, which occurs when a person adopts AI output with minimal scrutiny, overriding their own intuition and deliberative reasoning in a near-total transfer of agency. The study suggests that using AI does not always cause skills to erode.

The technology can at times drive skill expansion, depending on three factors:

1. **Method of use:** When AI functions as a tutor rather than a mere tool, it enhances independent human capabilities by building better mental models for complex tasks. For example, medical novices who trained with an AI system providing feedback and clarifying concepts showed lasting performance gains even after the tool was removed. Interrogating AI outputs to understand its reasoning allows users to expand their contextual expertise and professional judgment.

2. **Individual psychological traits:** People with a high need for cognition, defined as a tendency to engage in and enjoy effortful thinking, are more likely to resist uncritical surrender to AI. Similarly, those with higher fluid intelligence, or the capacity to solve novel problems and reason in new situations, are better able to offload tasks strategically while maintaining cognitive control.

3. **Environmental factors:** Receiving detailed feedback and operating under performance-based incentives can reactivate human deliberation. Conversely, extreme time pressure can suppress internal oversight and make cognitive surrender more likely.

Learn more

Gulati, P., Marchetti, A., Puranam, P., & Sevcenko, V. (2025). *Generative AI adoption and higher order skills*. arXiv. https://doi.org/10.48550/arXiv.2503.09212

Chen, W. X., Srinivasan, S., & Zakerinia, S. (2024). *Displacement or complementarity? The labor market impact of Generative AI (Working Paper No. 25039).* Harvard Business School. https://www.hbs.edu/faculty/Pages/item.aspx?num=67045

3: Shaw, S. D., & Nave, G. (2026, January 12). *Thinking—Fast, Slow, and Artificial: How AI is Reshaping Human Reasoning and the Rise of Cognitive Surrender* (SSRN Working Paper No. 6097646). https://doi.org/10.2139/ssrn.6097646

He, V. F., Li, S., Puranam, P., & Lin, F. (2025). *Predictive AI can support human learning while preserving error diversity*. arXiv. https://doi.org/10.48550/arXiv.2502.16411

Uzzi, B. (2026). AI can spark creativity—if we ask it how, not what, to think. Nature, 649(8097), 529-529. https://doi.org/10.1038/d41586-026-00049-2

Idea 8

The Future of Jobs: Despecialization and how new jobs are born

TL;DR: Rather than simply automating existing roles, AI will likely redraw specialist boundaries and create entirely new types of jobs. We explore evidence on despecialization and the pathways through which new

Jobs in a startup evolve as it becomes a large company. In the beginning, it's "all hands on deck." Everyone wears 10 hats to keep the lights on. As the company scales, those hats become heavy. The company hires specialists in finance, engineering, and sales. Each professional wears one hat. Eventually, the organization becomes a giant machine of deep experts who coordinate across silos.

Are we about to see a massive U-turn? Large organizations could see a collapse of specialization because AI endows users with multifunctional capabilities. But the current evidence is mixed:

- **Despecialization is likely:** Harvard University experiments suggest that AI can democratize expertise and allow people to work across functional boundaries. By swapping roles between R&D and commercial professionals at Procter & Gamble, the researchers demonstrated that traditional boundaries can dissolve.
- **Despecialization has limits:** Stanford and Harvard researchers argue that expertise distance matters. Those in adjacent roles, such as frontend and backend engineers, can use AI to bridge into new areas. However, occupational outsiders (e.g., moving from legal to finance) hit a *GenAI wall:* the point at which AI no longer narrows the expertise gap. Researchers also distinguish between ideation and implementation. Bridging expertise is harder with implementation tasks, which require tacit knowledge necessary to ensure quality.

These two contrasting effects may be a feature of the technology's jagged frontier. Also, they only consider the effect on jobs as we currently define them.

What new jobs will emerge as a result of AI? Although no structured study has yet answered this question, we can look at new job creation in the past to see if the same patterns will repeat.

Four mechanisms could be relevant to AI's impact on jobs:

1. **Technology-created bottlenecks:** New augmenting technology often requires human expertise to manage the new output. For example, naval officers emerged from the necessity of operating large ships with complex navigation, and circuit layout designers followed the invention of the circuit board. The role of prompt engineer has already appeared, and an agentic workflow manager will likely follow. Film production houses using generative media have reinvented the computer graphics team to be involved from the earliest phases of ideation and world-building, not just in the final stages of post-production.

2. **New problems from new technologies:** New occupations can emerge from *task vacancies*, where different groups compete for ownership of a new problem. Cybersecurity, for instance, became a field where police, IT experts, and lawyers all competed for ownership until a new type of role was created.

3. **Job crafting:** Organizations sometimes see grassroots efforts of individuals engaging in *task crafting* (performing duties outside their job description) and *relational crafting* (building new networks). If enough people shape their roles in the same way, a new job title eventually emerges. Customer success and community manager roles both began this way, as sales and marketing professionals reshaped their daily work.

4. **Legitimizing nonwork:** Collectives may mobilize to transform activities previously considered nonwork into paid occupations. The needs for childcare and elderly care created the roles of nannies and caregivers, while household chores gave way to personal concierges. Perhaps AI will allow casual hobbies or volunteer activities to become legitimate new jobs.

Learn more

Acemoglu, D., & Restrepo, P. (2019). Automation and New Tasks: How technology displaces and reinstates labor. *The Journal of Economic Perspectives, 33(2)*, 3–30. https://doi.org/10.1257/jep.33.2.3

Anteby, M., Chan, C. K., & DiBenigno, J. (2015). Three lenses on occupations and professions in organizations: becoming, doing, and relating. *Academy of Management Annals, 10*(1), 183–244. https://doi.org/10.1080/19416520.2016.1120962

Autor, D., Chin, C., Salomons, A., & Seegmiller, B. (2024). New Frontiers: The Origins *and Content of New Work, 1940–2018. The Quarterly Journal of Economics, 139*(3), 1399–1465. https://doi.org/10.1093/qje/qjae008

Dell'Acqua, F., et al. (2025). *The cybernetic teammate: A field experiment on generative AI reshaping teamwork* (SSRN Working Paper No. 5188231). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5188231

Vendraminelli, L., et al. (2025). *The GenAI wall effect: Examining the limits to horizontal expertise* (SSRN Working Paper No. 5462694). http://dx.doi.org/10.2139/ssrn.5462694

Wrzesniewski, A., & Dutton, J. E. (2001). Crafting a job: Revisioning employees as active crafters of their work. *The Academy of Management Review, 26*(2), 179–201. https://doi.org/10.2307/259118

Idea 9

AI-driven Coordination: AI's impact on organizational life

TL;DR: Research into AI collaboration is coalescing around three areas: social cognitive capabilities of AI, the impact on human coordination, and the efficacy of human-AI

The vast landscape of AI and collaboration is organizing around three research questions:

Does AI have the social capabilities to be a good collaborator?

Many researchers are exploring whether LLMs can develop what psychologists call *theory of mind*. This is a fundamental collaboration skill through which people attribute mental states like beliefs, intents, and knowledge to themselves and others. Theory of mind is the social engine that allows us to coordinate actions effectively. It is often measured by its complexity, known as levels of recursive reasoning:

- The first order of complexity is self-reflection: "I think that X."
- The second order is predicting another's belief: "I think that you think X."
- The third order is social triangulation: "I think that you believe she knows X."

Current research is finding that AI can develop these social capabilities. By accurately imputing both the cognitive and emotional states of users, LLMs are evolving into empathetic partners capable of true collaboration.

Does using AI help or hurt human collaboration?

While there is little empirical research within companies on this question, we have evidence from collective deliberations in a political setting. In one study from Google DeepMind, an AI designed to find common ground on splintered political opinions outperformed human mediators in generating statements that participants collectively endorsed. Political polarization is notoriously intractable because confirmation bias leads people to dig in when challenged.

It is easy to see how this system could resolve coordination challenges in organizations, particularly when progress is stalled by knowledge silos, hidden incentives, or personality quirks.

This finding echoes another argument that AI's primary value for organizations lies in its ability to slash translation costs or the manual effort required to turn one team's outputs into another's inputs. By transferring knowledge between departments, companies can achieve coordination without the need for a common vocabulary.

But there is a compelling counterargument. Researchers from Harvard and Carnegie Mellon argue that by drastically lowering the cost of thinking, AI inadvertently undermines mental proof—the observable effort we use to verify unobservable traits like skill, intent, or integrity. In low-trust environments, potential collaborators look for these proofs to decide how much to cooperate and when none exist, they instinctively retreat. The authors conclude that by making thinking easier, AI can make cooperation harder.

Are hybrid human-AI teams superior to human-only teams?

A groundbreaking meta-analysis of human-AI teams offers a nuanced conclusion. By synthesizing data from 106 experimental studies, the researchers evaluated how machine learning and generative AI actually impact team performance. Overall, hybrid teams performed better than human-only teams. However, meta-analysis reveals that human-AI collaboration is not always additive. In cases where the AI outperformed the human, the hybrid team's results were paradoxically inferior to those of the AI working alone. This result suggests that when AI capabilities are significantly higher than a human's capabilities at a specific task, forcing a human-in-the-loop approach decreases overall performance. While technological advances may shift these benchmarks since the study's 2024 publication, entrenched human biases could sustain this performance gap even as AI capabilities accelerate. (See idea 10: *AI for Strategic Decisions*.)

Ultimately, leaders must decide where to optimize for raw performance and where other critical factors—like human oversight and long-term skill retention—justify the tax of a human-in-the-loop design.

Learn more

Tessler, M. H., et al. (2024). AI can help humans find common ground in democratic deliberation. *Science, 386*(6719), eadq2852. https://doi.org/10.1126/science.adq2852

Wojtowicz, Z., & DeDeo, S. (2024). *Undermining mental proof: How AI can make cooperation harder by making thinking easier.* arXiv. https://doi.org/10.48550/arXiv.2407.14452

Vaccaro, M., Almaatouq, A., & Malone, T. W. (2024). When combinations of humans and AI are useful: A systematic review and meta-analysis. *Nature Human Behaviour, 8*(12), 2293–2303. https://doi.org/10.1038/s41562-024-02024-1

Brooks, J. A., et al. (2024). Deep learning reveals what facial expressions mean to people in different cultures. *iScience, 27*(3), 109175. https://doi.org/10.1016/j.isci.2024.109175

Choudary, S. P. (2026, February 25). AI's big payoff is coordination, not automation. *Harvard Business Review*. https://hbr.org/2026/02/ais-big-payoff-is-coordination-not-automation

Goldberg, A., & Srivastava, S. B. (2024). How Artificial Intelligence Can Enrich Our Understanding of Organizational Culture. *Management and Business Review*, 4(2), 32-37. https://doi.org/10.1177/26941058241301622

Kumar, A., et al. (2026). When large language models are reliable for judging empathic communication. *Nature Machine Intelligence, 8*(2), 173–185. https://doi.org/10.1038/s42256-025-01169-6

Lu, L., Tormala, Z. L., & Duhachek, A. (2025). How AI sources can increase openness to opposing views. *Scientific Reports, 15*, 17170. https://doi.org/10.1038/s41598-025-00791-z

Schmutz, J. B., et al. (2024). AI-teaming: Redefining collaboration in the digital era. *Current Opinion in Psychology, 58*, 101837. https://doi.org/10.1016/j.copsyc.2024.101837

Street, W., et al. (2024). *LLMs achieve adult human performance on higher-order theory of mind tasks*. arXiv. https://doi.org/10.48550/arXiv.2405.18870

Idea 10

AI for Strategic Decisions: Challenges in implementing AI-powered decision tools

TL;DR: Human bias and rigid structures may stifle AI's usefulness in supporting decisions. We explore three key challenges for leaders: algorithmic aversion, decision tools clashing with the human org chart, and the risk of a monoculture.

Decision making is a prediction problem that AI promises to solve with surgical precision. Yet, emerging research shows human bias and rigid organizational structures can stifle AI's superpower. Three research streams offer actionable implications for leaders:

Algorithm aversion bias

Algorithm aversion is a behavioral anomaly —studied in traditional machine learning and more recently in generative systems—where people prefer human forecasters or decision-makers to superior, yet imperfect, algorithms. This bias exists for several reasons, including people's desire for agency, negative moral or emotional reactions (especially when decisions carry major consequences for other people), and a belief that human experts possess unique knowledge uncaptured by AI systems. Most interestingly, researchers note an *asymmetrical forgiveness*, or a willingness to believe human error is perfectable over time, whereas machine error is fundamentally fixed. To address this bias, studies suggest building systems with modifiability (even slight user adjustments significantly boost adoption), improving explainability to clarify black-box analysis (see idea 3: AI Interpretability), and anthropomorphizing systems to foster human-like trust.

Decision tools that conflict with the human org chart

Choosing which AI-powered decision support system to deploy is just as important as *where* it is deployed. One study by Stanford researchers found that aligning these tools with the human org chart creates problems. Human silos, organized around legacy structures or managerial convenience, rarely optimize the data structure for AI to be effective. Moreover, decision-makers deep in the organization tend to optimize for role- or team-specific goals and feel little agency to go beyond silos. Researchers call this the *leaf node* problem. Decisions become trapped at the bottom of the hierarchy, isolated from the broader organizational context that's needed for optimal decisions.

Since AI does not always consider human silos when offering strategic guidance, it's important to deploy the systems at a level of the organization to maximize company-level goals.

AI-assisted organizational learning that levies a tax on collective imagination

In management science, decision making, strategy, and organizational learning are tightly intertwined, forming the backbone of how a company adapts. New research from New York University and MIT illuminates how generative AI supports organizational learning, including knowledge search, creation, transfer, and forgetting.

Without AI, these processes are constrained by knowledge scarcity, when expertise is difficult to obtain or synthesize. Through *generative organizational learning*, AI creates knowledge abundance through its democratizing and conversational capabilities, allowing decision makers to co-create solutions in areas that were previously the domain of specialized experts.

This trend may come with a hidden tax on collective diversity, however. While AI makes individual output more creative, another study[3] from the University College London and University of Exeter found it causes collective output to become significantly more similar, reducing total novelty across the organization.

To avoid a monoculture of ideas, research suggests implementing a sequential decision paradigm—where decision makers record independent thoughts before consulting AI.

Learn more

Valentine, M. A., et al. (2024). *The Algorithm and the Org Chart: How Algorithms Can Conflict with Organizational Structures. Proceedings of the ACM on Human-Computer Interaction, 8(CSCW2), 1–31.* https://doi.org/10.1145/3686903

Wiesenfeld, B. M., & Kellogg, K. (2025). Generative organizational learning: Affordances for new modes of knowledge search, creation, transfer, and forgetting with LLMs *(NYU Stern School of Business Research Paper No. 6221878)*. https://doi.org/10.2139/ssrn.6221878

Doshi, A. R., & Hauser, O. P. (2024). Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances, 10*(28), eadn5290. https://doi.org/10.1126/sciadv.adn5290

Sunstein, C., & Gaffe, J. (2025). An anatomy of algorithm aversion. *Science and Technology Law Review, 26*(1). https://doi.org/10.52214/stlr.v26i1.13339

Mahmud, H., et al. (2021). *What influences algorithmic decision-making? A systematic literature review on algorithm aversion. Technological Forecasting and Social Change, 175*, 121390. https://doi.org/10.1016/j.techfore.2021.121390

Idea 11

The Psychology of Work: How professional identity and purpose are shifting

TL;DR: Recent studies are shedding light on how professionals are navigating the psychological shifts brought about by AI, and how they are finding economic viability while preserving pride in their craft.

In 2016, Go world champion Lee Sedol faced a personal crisis after a defeat by AlphaGo, an AI system trained on a dataset of 30 million moves from 160,000 games by advanced non-professional players. He later remarked that "losing to AI, in a sense, meant my entire world was collapsing. I could no longer enjoy the game, so I retired." In retirement, Sedol redirected his expertise by writing books, founding a children's academy, and teaching board game design at a national university. Meanwhile, as inspired by AI play, the number of Go players has increased and player skills have improved. Sedol's journey of meaning displacement is one many professionals are beginning to navigate today.

Sedol's experience foreshadowed career reinvention strategies seen in recent studies of writers and dancers adapting to generative AI. The subjects of these studies often engage in an inner battle to preserve their professional identity, using behaviors that help them assert their human craft. These include making their human labor visible by, for instance, using personal anecdotes or misplaced humor to differentiate from AI output, or make it clear no AI was used to appeal to an audience who values 100% human work. The trend is part of a broader concept of *identity* work, a psychological process through which individuals protect or restructure their professional image in response to AI.

While many creative professionals resist AI, others have begun to embrace the technology. Several insights emerge from this transition:

- **Technical vanguardism:** Some professionals position themselves at the forefront of innovation by participating in projects to evaluate and refine generative AI models, defining themselves as collaborators and co-creators of the technology.

- **Redefining success:** To protect their professional self-worth, AI embracers resist vanity metrics (like views or likes on social media) and instead establish internal, self-defined standards for evaluating the quality and value of their work.

- **The AI ghostwriting effect:** Those who use AI often downplay its role in their work. Positioning AI as a ghostwriter allows individuals to leverage the efficiency of the tool while publicly preserving their identity as the sole creative engine.

- **Cultivating spaces of transcendence:** Perhaps most significantly, many creators purposefully cultivate sacred spaces—such as a private dance studio or a disconnected writing room—to feel connected to the purity of their craft. These spaces allow for the emergence of insight into their work that AI cannot replicate.

While these studies focus on creative fields, they show us what is likely happening in many other industries. AI is changing how we see ourselves and the role work plays in our lives. Leaders who only focus on the tech and ignore the psychological side of this shift will struggle. To make this transition successful, we need a human-centered approach that supports people as they find their new place in an AI world.

Learn more

Jiang, W. & Varma, P. (n.d.). *Dancing with AI: The role of meta-agency in the reconstruction of viable and meaningful careers amid technological change* [Unpublished manuscript]. INSEAD. https://sites.google.com/view/winnie-jiang/working-projects

Varanasi, R. A., Wiesenfeld, B. M., & Nov, O. (2025). AI rivalry embracing generative AI are reshaping the writing profession. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems, Japan, 1198*, 1–19. https://doi.org/10.1145/3706598.3714035

Part 03

AI as Organizations

This section examines the use of AI simulacra as a powerful analytical instrument to decode complex patterns in our organizations. It explores the frontier of generative agent simulations, when thousands of AI agents interact within high-fidelity digital environments. By modeling the ripple effects of interaction dynamics and world-building, this research stream offers a transformative way to test organizational ideas and predict systemic outcomes before they manifest in the physical world.

Idea 12

Generative Simulations: How organizations can rehearse alternative futures

TL;DR: In a niche but growing field, AI simulations deliver benefits over simple modeling. Recent studies have created high-definition mirrors of our organizations.

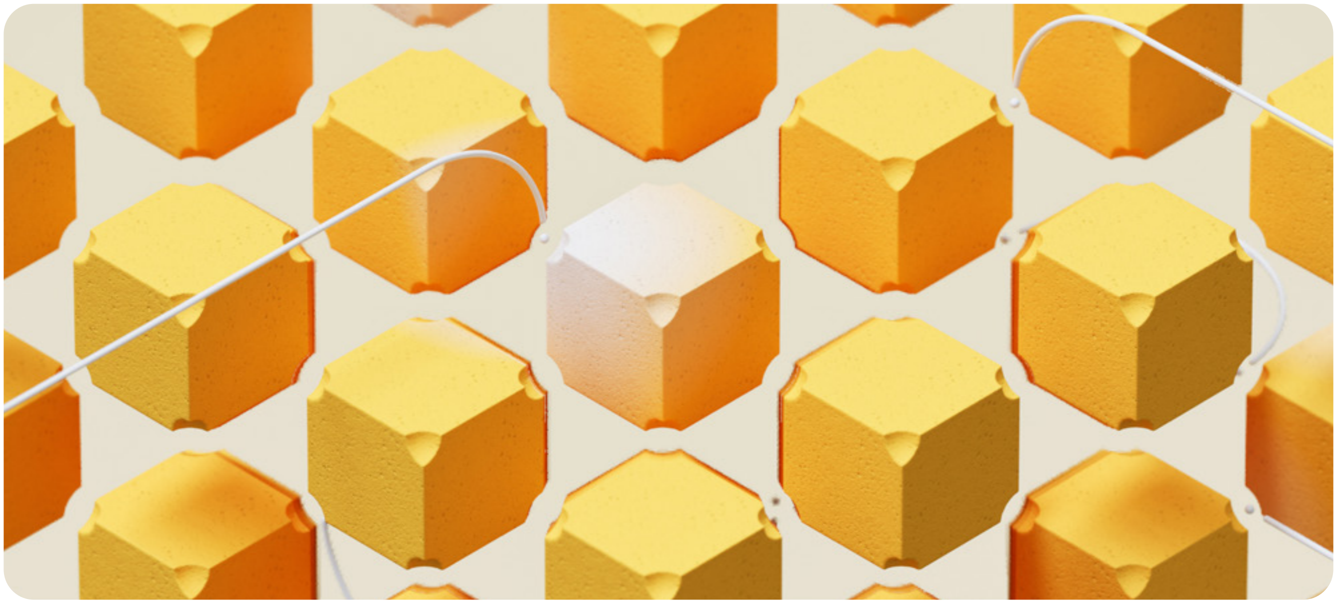
There are two ways of thinking about the future. A prediction tells you what will probably happen based on patterns from the past. A simulation asks something harder: Why does it happen, and what would change if the rules were different? Agent-based modeling offers a real test of understanding patterns by building a world from scratch, populating it with actors, giving them behaviors and beliefs, and watching what emerges. When your outcomes match reality, you aren't just describing the world—you're explaining it.

Until now, simulations relied on oversimplifications so extreme that pioneers described them as "doing violence to reality." Early models gave each agent a handful of if-then rules (e.g., move toward food, avoid predators, copy your neighbor). While useful, early simulations were far from realistic.

The arrival of large language models changed everything. Recent research shows what this looks like in practice: 25 AI agents, each with a name, job, and memory, placed inside a simulated town and left to their own devices. No one scripted what would happen. And yet, recognizable human dynamics emerged: Gossip spread, friendships formed, a party was organized, rumors took on lives of their own. The agents reflected, made plans, and reacted to each other in ways that felt less like a computer program and more like a society. It's a significant shift for organizations because for the first time, the "people" inside a simulation can behave like real humans.

Still, a simulation is only as trustworthy as its assumptions. AI agents are not universal stand-ins for every individual in an organization; they reflect the perspectives dominant in their training data while often glossing over under-represented experiences.

Leaders should remember, a simulation of the workforce is not an exact replica. It is a detailed, structured sketch. And this is what makes simulation so valuable: It forces you to make assumptions explicit, examine them, and consider when they might be incorrect. Unlike a gut feeling or a strategy deck, a simulation is an argument you can interrogate. You can change an assumption and immediately see what breaks. You can probe the edges of your own thinking in ways that are rarely possible in a boardroom. A generative



simulation is a new kind of organizational mirror, one that reflects not only what you expect to happen, but also the hidden logic underneath those expectations.

Imagine being able to stress-test high-stakes decisions through pre-mortems at scale, surfacing critical failure modes and challenging assumptions before a single resource is committed. By modeling how new policies ripple through the informal networks of an organization, you can anticipate second-order effects like cultural resistance or clever workarounds, essentially running a dress rehearsal for internal change before it hits the real world.

This predictive power extends to the external market as well. What if you were to create AI simulations of your competitor's organizations, replacing static strategy snapshots with living maps that simulate the dynamic, accumulating interactions? Ultimately, AI simulations will allow you to navigate the complexities of both human behavior and competitive shifts with a degree of foresight that intuition alone could never provide.

Learn more

Park, J. S., O'Brien, J., Cai, C. J., Morris, M. R., Liang, P., & Bernstein, M. S. (2023). Generative agents: Interactive simulacra of human behavior. *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology, USA, 2*, 1–22. https://doi.org/10.1145/3586183.3606763

Epstein, J. M. (2007). *Generative social science: Studies in agent-based computational modeling*. Princeton University Press. https://press.princeton.edu/books/ebook/9781400842872/generative-social-science-pdf

Kozlowski, A. C., & Evans, J. (2025). Simulating subjects: The promise and peril of Artificial Intelligence stand-ins for social agents and interactions. *Sociological Methods & Research, 54*(3), 1017-1073. https://doi.org/10.1177/00491241251337316

Idea 13

The Simulation of Individuals: How to capture the richness of a human life

TL;DR: Researchers are exploring how to create digital doppelgängers and high-fidelity simulations of individuals. But there's an ongoing debate about the reliability of LLMs for high-stakes human decision-making.

Pop culture has long been fascinated by the doppelgänger—a person's mirror image that shares a striking, near-identical resemblance. Now, a growing community of computer science and psychology scholars is turning this folklore into an engineering challenge. The first riddle they must solve is what inputs are required to capture the richness of an individual.

Are humans the sum of their personality tests or their curated anecdotes? Most people assume we are the ultimate authorities on our own lives; yet, we are frequently the least reliable narrators of our own stories. What if our true identities could be found by triangulating external data points, such as the patterns of our actions and observations from our closest collaborators, along with a self-report of the private beliefs and values that are difficult to discern?

One Stanford study, "Generative Agent Simulations of 1,000 People," took a pragmatic first step. Rather than attempting a complete mapping of all the data points for one's life, researchers focused on the raw material of an interview transcript. By running two-hour interviews with more than a thousand real people, the team built AI agents that replicated participants with 85 percent accuracy—matching the consistency with which study participants answered a survey two weeks later. While that implies a notable error rate, this simplified study indicates the research space is fertile ground for future work.

Other studies have taken an even more pragmatic approach. In the real world, a transcribed two-hour interview is rare, and researchers would need to work from self-disclosed personality types, artifacts of written documents or presentations produced by the participants, or even general impressions developed from meetings. Some researchers believe sparse input could be enough to build agents from structured psychometric profiles rather than rich life narratives.

Another study, "Designing AI-Agents With Personalities," found that when agents are built from formal personality inventories instead of loose adjectives, they make different decisions on risk-taking and moral dilemmas, in patterns that match real human data.

Both startups and larger tech companies are starting to take this challenge seriously. They are finding ways to enrich datasets on individuals in an effort to understand what individuals will buy or how they'll vote. For internal operations, companies are creating AI versions of key decision-makers to simulate and prepare for high-stakes meetings. Other organizations are looking to use these models to simulate employee reactions to a new change, program, or policy.

Some AI researchers are unconvinced that the LLMs that exist in April 2026 are capable of building out simulations that organizations can use reliably to make real-world, high-stakes decisions. However, these emerging practices raise big questions around data privacy and potential misuse. (See idea 16: *Ethical Frontiers*.)

Learn more

Park, J. S., Zou, C. Q., Shaw, A., Hill, B. M., Cai, C., Morris, M. R., Willer, R., Liang, P., & Bernstein, M. S. (2024). *Generative agent simulations of 1,000 people*. arXiv. https://doi.org/10.48550/arXiv.2411.10109

Huang, M., Zhang, X., Soto, C., & Evans, J. (2026). Designing AI-Agents with personalities: A psychometric approach. *Personality Science, 7*. https://doi.org/10.1177/27000710251406471

Hu, Q., Xiao, Q., Cao, H., & Shen, H. (2025). *When Your Boss Is an AI Bot: Exploring Opportunities and Risks of Manager Clone Agents in the Future Workplace*. arXiv. https://doi.org/10.48550/arXiv.2509.10993

Kim, J., Evans, J., & Schein, A. (2025, April 24-28). *Linear representations of political perspective emerge in large language models* [Conference paper]. ICLR 2025, Singapore. https://doi.org/10.48550/arXiv.2503.02080

Idea 14

Simulating Group Dynamics: Organizations can replicate team dynamics to anticipate real-world risks

TL;DR: Structured AI simulations can now predict team friction points and create virtual rehearsals that bypass the costs of broken chemistry in scenarios like corporate mergers or R&D sprints.

You can assemble a team of brilliant individuals and still watch them fail. A landmark study in *Science* found that a team's collective intelligence was not correlated with the average or maximum IQ of its members. Social sensitivity of group members and equal conversational turn-taking predicted performance more accurately.

When social interaction patterns break down, the consequences are difficult to ignore. A Harvard study showed that when teams lack psychological safety, they stop learning: Members withhold ideas, hide errors, and avoid the interpersonal risks that productive collaboration requires. Eventually, projects stall, people leave, and mistakes compound in silence. The organization then experiences delayed launches, failed integrations, and turnover that leaders often misattribute to individual underperformance rather than to the team itself. The problem is rarely talent; it's interaction. But what if you could detect these failure modes before a team ever meets by using AI to set up a team simulation? A Stanford study recently showed this is possible, with stunning results. The researchers built a Virtual Lab—a group of AI agents playing an immunologist, a computational biologist, a machine learning specialist, and a scientific critic. The agents held structured meetings, debated strategy, and collectively designed a pipeline for creating nanobody therapeutics against SARS-CoV-2. From 92 in silico designs, real-world laboratory testing identified two candidates that proved promising. According to the study, team design emerged as a key success factor: Having distinct roles, structured deliberation, and a dedicated critic who forced the group to challenge weak ideas early, rather than converge on a comfortable consensus, led to the new discoveries.

For organizations, the Stanford study suggests that team chemistry can be modeled and tested before anyone is hired or merged into a new unit. Platforms like Concordia, developed at Google DeepMind, already make this possible.

By using AI agents to simulate personalities and structural roles—like a devil's advocate or customer proxy—organizations can spot friction points with zero stakes. If integrating a complex corporate merger, modeling these dynamics in advance could save the company from massive, real-world costs of failed integrations, delayed projects, and broken chemistry.

Collective success depends not only on talent, but also on social factors. And for the first time, team dynamics are something you can rehearse.

Learn more

Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., & Malone, T. W. (2010). Evidence for a collective intelligence factor in the performance of human groups. *Science, 330*(6004), 686–688. https://doi.org/10.1126/science.1193147

Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly, 44*(2), 350-383. https://doi.org/10.2307/2666999

Swanson, K., Wu, W., Bulaong, N. L., Pak, J. E., & Zou, J. Y. (2025). The virtual lab of AI agents designs new SARS-CoV-2 nanobodies. *Nature, 646*(8085), 716–723. https://doi.org/10.1038/s41586-025-09442-9

Vezhnevets, A. S., Agapiou, J. P., Aharon, A., Ziv, R., Matyas, J., Duéñez-Guzmán, E. A., Cunningham, W. A., Osindero, S., Karmon, D., & Leibo, J. Z. (2023). *Generative agent-based modeling with actions grounded in physical, social, or digital space using Concordia*. arXiv. https://doi.org/10.48550/arXiv.2312.03664

Kim, J., Lai, S., Scherrer, N., Agüera y Arcas, B., & Evans, J. (2026). *Reasoning models generate societies of thought*. arXiv. https://doi.org/10.48550/arXiv.2601.10825

Idea 15

Simulating Organizational Behavior: Emergent collective actions at scale

TL;DR: Organizational success often hinges on invisible, emergent interactions that leaders cannot plan. Large-scale simulations uncover how thousands of interactions coalesce into collective behavior.

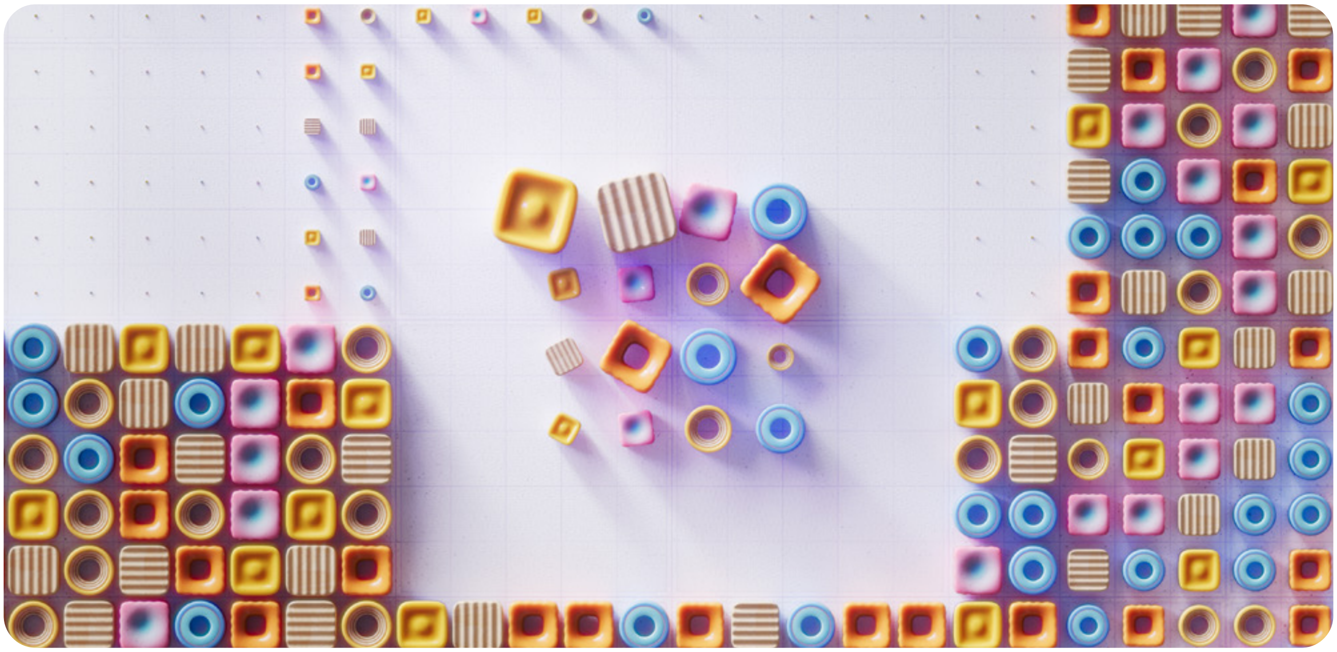
Imagine you are a CEO 72 hours from announcing a return-to-office mandate. Your team has stress-tested the logic, prepared the FAQ, and lined up your champions. What you cannot rehearse, however, is the invisible friction that could hinder success: the Slack threads you will never see, the hallway conversations that rewrite your message, and the resentment that hardens before leadership even senses a problem. No one decides that a backlash will form. It emerges from many people reacting not only to the policy, but also to each other.

While organizations use sophisticated tools to design strategy, they have almost no way to predict how a new strategy will propagate through the human networks that carry it out. We have already seen the power of simulating individuals and teams, but the potential of AI scales dramatically when we model entire organizations to discover what computational social scientists call *emergence*—the way thousands of private, simple interactions unexpectedly coalesce into collective patterns of belief and behavior that can either propel a strategy forward or quietly dismantle it.

A recent study in *Science Advances* showed this kind of emergence can be simulated. The researchers placed populations of LLM agents in repeated interaction games. The agents spontaneously developed shared conventions and collective biases appeared at the population level that no individual agent exhibited. A small committed minority could tip the entire group toward a new norm. The dynamics looked less like a computer program and more like a culture forming in real time.

Moving beyond toy laboratory tasks, recent research tackles complex, societal dynamics. A University of Chicago team used LLM simulations to forecast COVID-19 polarization in the U.S. as it happened. Remarkably, the results tracked belief divergence in real time, a feat traditional sentiment surveys can only capture after the damage is done.

Understanding these cascading effects, though, is only half the prize. How might you help a group reduce polarization with the aid of AI? An MIT study showed that AI-driven dialogue could durably reduce conspiracy beliefs. Using a pipeline of real-time interactions with an AI chatbot, researchers found that bespoke, evidence-



based dialogues reduced participants' belief in conspiracies by an average of 20%. This effect was durable—lasting at least two months—far outlasting most human-led interventions.

Together, these studies offer a glimpse into what could be possible for organizations that lean into AI simulations: Get ready to simulate how changes cascade through your workforce, identify where dynamics go wrong, and design targeted corrections that travel the same channels. For any leader who has watched an internal debate calcify into factions or a change management effort fall flat, these new capabilities replace costly trial-and-error with predictive foresight.

As of early 2026, this idea has not yet been brought to commercial scale, save for a couple of early startups. One of the greatest challenges ahead is validating these simulations as high-fidelity tools that bridge the reality gap between modeled behavior and real-world organizational dynamics. We are excited to see this work evolve and become a staple for organizations everywhere.

Learn more

Ashery, A. F., Aiello, L. M., & Baronchelli, A. (2025). Emergent social conventions and collective bias in LLM populations. *Science Advances, 11*(20), eadu9368. https://doi.org/10.1126/sciadv.adu9368

Kozlowski, A. C., Kwon, H., & Evans, J. A. (2024). *In silico sociology: Forecasting COVID-19 polarization with large language models*. arXiv. https://doi.org/10.48550/arXiv.2407.11190

Hackenburg, K., Tappin, B. M., Hewitt, L., Saunders, E., Black, S., Lin, H., Fist, C., Margetts, H., Rand, D. G., & Summerfield, C. (2025). The levers of political persuasion with conversational artificial intelligence. *Science, 390*(6777), eaea3884. https://doi.org/10.1126/science.aea3884

Costello, T. H., Pennycook, G., & Rand, D. G. (2024). Durably reducing conspiracy beliefs through dialogues with AI. *Science, 385*(6714), eadq1814. https://doi.org/10.1126/science.adq1814

Idea 16

Ethical Frontiers: Simulations present unique questions of misuse

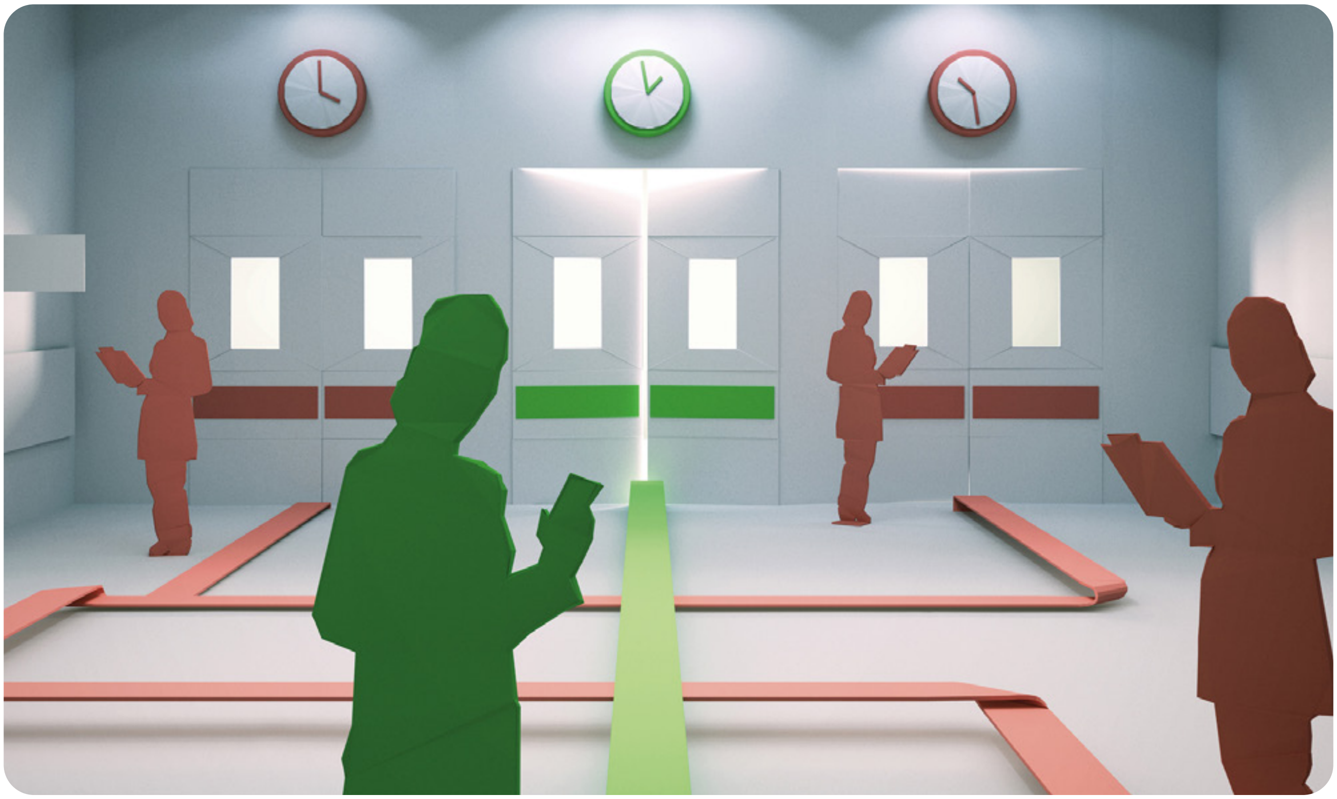
TL;DR: Using AI simulacra in organizations brings about ethical concerns, including consent, overreliance, ownership, and manipulation.

The simulations discussed thus far represent the current frontier of the field. While the AI community continues to debate whether models are yet capable of true organizational replication, we need to consider the risks surfacing as this technology advances. Four primary concerns have emerged from the current studies on the ethics of AI simulations and advanced agents:

1. **Consent:** When a company builds an AI replica of a senior leader to pre-test how she would react to a business review presentation or new org structure, it is predicting reactions that she never made. Researchers flag reputational harm and identity distortion as distinct dangers that emerge whenever AI starts speaking as someone who is real. These risks are compounded by workplace hierarchies, where employees may feel pressured to consent and ultimately lose the power to audit or challenge their digital doubles.

2. **Overreliance:** When probabilistic outputs influence real-world decisions, organizations can get into trouble. If leaders use AI like a crystal ball, they'll make decisions based on events and circumstances that haven't happened. Employees can be judged, passed over, or replaced based on actions they never took and thoughts they never had. Since these models can only draw from cold correlations, they may leave out the human nuance that drives everyday workplace behavior.

3. **Ownership:** If a simulation is built on an individual employee, who owns it—the company or the person? When employees leave the organization, do they have the right to take their digital twins with them, similar to how employees in many countries have a legal right to access and port their human resources data? How the ownership question gets resolved will dictate who controls the model, whether employees must opt in, and if workers have the right to audit or calibrate outputs with which they disagree.



4. **Manipulation:** AI-driven dialogue has a proven capacity to shift human beliefs in ways that last. In the wrong hands, a simulation designed to anticipate resistance can become a tool for psychological manipulation. The same technology that lets you rehearse a more collaborative future also lets you rehearse a more obedient one, blurring the line where creating alignment around shared goals ends and coercive influence begins.

The papers listed in the 'Learn more' box offer a more thorough discussion of the risks of AI simulations. The trajectory of this technology depends entirely on the choices that leaders, developers, and regulators make today. We hope this overview serves as a catalyst for the coordination needed to ensure that the technology reflects the world we want to inhabit.

Learn more

Gabriel, I., et al. (2024). *The ethics of advanced AI assistants.* arXiv. https://doi.org/10.48550/arXiv.2404.16244

Kozlowski, A. C., & Evans, J. (2025). *Simulating Subjects: The promise and peril of Artificial intelligence Stand-Ins for Social Agents and Interactions. Sociological Methods & Research, 54*(3), 1017–1073. https://doi.org/10.1177/00491241251337316

Morris, M. R., & Brubaker, J. R. (2024). *Generative Ghosts: Anticipating benefits and risks of AI afterlives.* arXiv. https://doi.org/10.48550/arXiv.2402.01662

This research primer offers a synthesis of work by pioneers of the field whose intellectual contributions are mapping the current frontier:*

Google DeepMind

A. Stevie Bergman, Abbas Abdolmaleki, Abhishek Jindal, Acorn Pooley, Adil Dostmohamed, Adrian Collister, Adrien Baranes, Agrim Gupta, Alan S. Cowen, Aleksandra Faust, Alessio Quaglino, Alex Bewley, Alex Hofer, Alex Irpan, Alex X. Lee, Alexander Matt Turner, Alexander Sasha Vezhnevets, Alison M. Snyder, Allan Dafoe, Allan Zhou, Amaris Paryag, Anca Dragan, Andrew Marmon, Andrew Trask, Anirudha Majumdar, Anna Wang, Annie Xie, Anthony Brohan, Antoine Laurens, Arianna Manzini, Arthur Conmy, Arunkumar Byravan, Ashwin Balakrishna, Assaf Hurwitz Michaely, Atil Iscen, Ayzaan Wahid, Been Kim, Borja Balle, Brandon Hernaez, Caden Lu, Canfer Akbulut, Carl Doersch, Carolina Parada, Carrie Cai, Charles Shu, Chengda Wu, Christine Chan, Chuyuan Fu, Claudio Fantacci, Clement Farabet, Cody Fong, Coline Devin, Conor Griffin, Courtney Biles, Daniel Jarrett, David B. D'Ambrosio, David C. Parkes, David Lindner, David Rendleman, Debidatta Dwibedi, Deepali Jain, Demis Hassabis, Dhruv Shah, Dmitry Kalashnikov, Don Wallace, Donnie Kim, Dushyant Rao, Edgar A. Duéñez-Guzmán, Edward Hughes, Emilio Parisotto, Erik Frey, Fangchen Liu, Federico Casarini, Fei Xia, Francesco Nori, Francesco Romano, Frank Kim, Georgina Evans, Giulia Vezzani, Grace Vesom, Guangyao Zhou, Guy Lever, Hannah Sheahan, Hao-Tien Lewis Chiang, Haroon Qureshi, Harry Law, Hasan Iqbal, Helen King, Helen Ran, Hiu Hong Yu, Ian Storz, Iason Gabriel, Ira Ktena, Isaac R. Galatzer-Levy, Isaac Reid, Isabel Leal, J. Chase Kew, Jacky Liang, Jake Varley, Jascha Sohl-Dickstein, Jasmine Hsu, Jayd Matyas, Jean-Baptiste Alayrac, Jeff Bingham, Jie Tan, Jimmy Yan, Jingwei Zhang, Jinyu Xie, Joel Z. Leibo, John Hewitt, John P. Agapiou, Jonathan Tompson, Jose Enrique Chen, Joshua Ainslie, Joss Moore, Joy Ortiz, Juan Mateos-Garcia, Julian Brown-Cohen, Julian Jacobs, Junkyung Kim, Kanishka Rao, Kate Olszewska, Kathryn Shea, Keerthana Gopalakrishnan, Ken Caluwaerts, Kenneth Oslund, Kevis-Kokitsi Maninis, Konstantinos Bousmalis, Krista Reymann, Krzysztof Choromanski, Ksenia Konyushkova, Laura Graesser, Laura Weidinger, Lauren Kim, Leonard Hasenclever, Li Yang Ku, Li-Heng Lin, Lionel Ho, Lisa Anne Hendricks, London Chappellet-Volpini, Lu Huang, Lucy Campbell-Gillingham, M. Emre Karagozler, Maria Bauza Villalonga, Maribeth Rauh, Marissa Giustina, Markus Wulfmeier, Martin J. Chadwick, Matija Franklin, Matija Kecman, Matthew Botvinick, Meet Kirankumar Dave, Meredith Ringel Morris, Michael Bloesch, Michael Elabd, Michael Henry Tessler, Michael Neunert, Michael Opara, Michiel A. Bakker, Miguel Rivas, Mikel Rodriguez, Mithun George Jacob, Mohit Sharma, Mohit Shridhar, Montserrat Gonzalez Arenas, Murray Shanahan, Nahema Marchal, Nathan Batchelor, Neel Nanda, Nenad Tomašev, Nicolas Heess, Niko Milonopoulos, Noah D. Goodman, Noah Fiedel, Norman Di Palo, Oliver Groth, Oran Kelly, Orhan Firat, Oriol Vinyals, Oscar Chang, Oyvind Tafjord, Pannag R Sanketi, Paul Wohlhart, Peng Xu, Peter Pastor Sampedro, Philemon Brakel, Pierre Sermanet, Pushmeet Kohli, Rachel Sterneck, Radu Soricut, Rajkumar Vasudeva Raju, Raluca Ada Popa, Raphael Koster, Razvan Surdulescu, René Wagner, Rishub Jain, Robert Geirhos, Robert Greig, Robert Moreno, Rohin Shah, Roland Hafner, Ruiqi Gao, Ryan Burnell, Ryan Julian, Sally Jesmonth, Sam Haves, Saminda Abeyruwan, Samuel Albanie, Sandy H. Huang, Saran Tunyasuvunakool, Sasha Brown, Sathish Thoppay Egambaram, Sean Kirmani, Sébastien Krier, Seliem El-Sayed, Senthooran Rajamanoharan, Sergio Martinez, Serkan Cabi, Shane Legg, Shangbang Long, Sharath Maddineni, Sherry Yang, Sichun Xu, Simon Osindero, Sims Witherspoon, Skye Yang, Stefan Krier, Stefan Welker, Stefani Karp, Steph Hughes-Fitt, Stephanie Farquhar, Stephen Emmons, Steven Hansen, Stuart Bridgers, Sudeep Dasari, Sumeet Singh, Tantum Collins, Ted Xiao, Thomas Buschmann, Thomas Kipf, Thomas Lampe, Thomas Power, Tianli Ding, Timi Ijitoye, Tingnan Zhang, Todor Davchev, Tom Erez, Tris Warkentin, Tsang-Wei Edward Lee, Tuan Anh Le, Ulrich Paquet, Verena Rieser, Victoria Krakovna, Vijay Bolina, Vikas Sindhwani, Wenhao Yu, Wentao Yuan, Will Hawkins, William A. Cunningham, William Isaac, Xi Chen, Xiaolin Fang, Yilun Du, Ying Xu, Yuan Yuan, Yuheng Kuang, Yulia Rubanova, Yumeya Yamamori, Yuxiang Yang, Yuxiang Zhou, Zachary Kenton, Zeb Kurth-Nelson, Zhiyuan Zhang, Zhuo Xu

Google Research and Paradigms of Intelligence Teams

Alex Ingerman, Alex Olwal, Alexander Reese, Alison Lentz, Amanda McCroskery, Andrew Barakat, Artem Dementyev, Avia Aharon, Benjamin Barnett, Benjamin Lange, Blaise Agüera y Arcas, Daniel McDuff, Danny Karmon, Geoff Keeling, Girish Narayanswamy, Jake Garrison, Jake Sunshine, James Evans, James Manyika, Jiening Zhan, John Oliver Siy, Junsol Kim, Kumar Ayush, Lize Alberts, Mark Malhotra, Nino Scherrer, Pushmeet Kohli, Reed Enger, Renee Shelby, Ron Ziv, Samy Abdel-Ghaffar, Sarah de Haas, Shrikanth Narayanan, Shun Liao, Shwetak Patel, Shyam Tailor, Tatenda Kanyere, Tim Althoff, Winnie Street, Xin Liu, Xuhai Xu, Yetunde Ibitoye, Yun Liu, Yuzhe Yang

Global Academic Community

American University: Thomas H. Costello

Boston University: Michel Anteby, Pascual Restrepo

Brown University: Claire Nicolas, Hadar Levi-Aharoni, John D. Simeral, Justin J. Jude, Leigh R. Hochberg, Shane B. Allcroft

Carnegie Mellon University: Alexandra Chouldechova, Anita Williams Woolley, Daniel Fried, Graham Neubig, Hans Moravec, Hong Shen, Lingwei Cheng, Qing Hu, Qing Xiao, Simon DeDeo, Zora Zhiruo Wang

Colby College: Christopher Soto

Columbia University: John Hewitt

Cornell University: David G. Rand, Gordon Pennycook, Manoj Karkee, Ranjan Sapkota

Dartmouth College: Sangeet Paul Choudary

Delft University of Technology: Michael Klenk

Duke University: Jack B. Soll, Jessica A. Reif, Richard P. Larrick

Eindhoven University of Technology: Anna-Sophie Ulfert

Emory University: Hancheng Cao

ESSEC Business School: Charles Ayoubi

ETH Zürich: Sophie Kerstan

Harvard University: Amy Edmondson, Annika Hildebrandt, Cass R. Sunstein, Curtis K. Chan, Daniel B. Rubin, David C. Parkes, Edward McFowland III, Fabrizio Dell'Acqua, Hadar Levi-Aharoni, Hila Lifshitz-Assaf, Iavor Bojinov, Jared H. Gaffe, Justin J. Jude, Karim R. Lakhani, Leigh R. Hochberg, Matthew DosSantos DiSorbo, Raffaella Sadun, Saleh Zakerinia, Suraj Srinivasan, Zachary Wojtowicz, Ziv M. Williams

Hong Kong University of Science and Technology: Wilbur Xinyuan Chen

Imperial College London: Guang-Zhong Yang

INSEAD: Phanish Puranam, Piyush Gulati, Preeti Varma, Sihan Li, Victoria Sevcenko, Winnie Jiang

IT University of Copenhagen: Luca Maria Aiello

Johns Hopkins University: Gillian K. Hadfield, Michael McKibben, Pedro Aceves

London Business School: Arianna Marchetti

London School of Economics & Political Science: Ben M. Tappin, Raffaella Sadun

Lappeenranta-Lahti University of Technology: A.K.M. Najmul Islam, Hasan Mahmud, Kari Smolander

Massachusetts Institute of Technology: Abdullah Almaatouq, Alex Pentland, Andrew Koh, Arnav Kapur, Caroline Chin, Christopher F. Chabris, Daron Acemoglu, David Autor, Hause Lin, Julia DiBenigno, Katherine Kellogg, Michelle Vaccaro, Nada Hashmi, Pattie Maes, Shreyas Kapur, Thomas H. Costello, Thomas Malone, Ziv M. Williams

Nagoya University: Toshio Fukuda

New York University: Amy Wrzesniewski, Batia Mishan Wiesenfeld, Joshua M. Epstein, Oded Nov, Rama Adithya Varanasi

Northwestern University: Aakriti Kumar, Aaron Shaw, Brian Uzzi, Bruce L. Lambert, Bryan Seegmiller, Carolyn Q. Zou, Matthew Groh, Nalin Poungpeth

Pennsylvania State University: Erina Farrell

Purdue University: Eamon Duede

Scuola Superiore Sant'Anna, Pisa: Paolo Dario

Stanford University: Alisa D. Levin, Amanda L. Pratt, Amir Goldberg, Arvind Karunakaran, Carolyn Q. Zou, Diyi Yang, Francis R. Willett, James Zou, Joon Sung Park, Joseph C. O'Brien, Kyle Swanson, Louise Lu, Luca Vendraminelli, Luke Hewitt, Melissa A. Valentine, Michael S. Bernstein, Muhua Huang, Omar Shaikh, Percy Liang, Rebecca Hinds, Robb Willer, Yijia Shao, Zakary L. Tormala

Stellenbosch University: Lize Alberts

The Wharton School of the University of Pennsylvania: Ethan R. Mollick, Gideon Nave, Lilach Mollick, Steven D. Shaw

Tufts University: David Pearl, James Intriligator, Xuanjiang Liu

University of California, Davis: David M. Brandman, Maitreyee Wairagkar, Nicholas S. Card, Sergey D. Stavisky

University College London: Anil R. Doshi, Oliver P. Hauser, Piyush Gulati, Vivianna Fang He

University of California, Berkeley: Aruna Ranganathan, Dacher Keltner, Jacob Steinhardt, Jiahai Feng, Maria Monroy, Rebecca Corona, Sameer B. Srivastava, Sangeet Paul Choudary, Stuart Russell, Xingqi Maggie Ye

University of California, Los Angeles: Abhishek Mishra, Brandon McMahan, Brent Gaisford, Chang Xie, Charles Kobashigawa, Johannes Y. Lee, Jonathan C. Kao, Mike Qu, Sangjoon Lee, Xu Yan

University of Amsterdam: Eleni Georganta

University of Chicago: Aaron Schein, Austin C. Kozlowski, Hyunku Kwon, James Evans, Junsol Kim, Shiyang Lai

University of Colorado: Jed R. Brubaker

University of Edinburgh: Shannon Vallor

University of Exeter: Oliver P. Hauser

University of Georgia: Neal Outland

University of Illinois Chicago: Adam Duhachek

University of London: Andrea Baronchelli, Ariel Flint Ashery

University of Michigan: Jane E. Dutton

University of Munich: Benjamin Lange

University of Oxford: Christopher Summerfield, Helen Margetts, Ingmar Posner, Jun Yamada, Kobi Hackenburg, Lisa Schut, Robin I. M. Dunbar

University of Texas at Austin: Yuke Zhu

University of the Peloponnese: Konstantinos I. Roumeliotis

University of Tokyo: Jihoon Oh, Kento Kawaharazuka

University of Toronto: Syed Ishtiaque Ahmed, William A. Cunningham

University of Washington: Benjamin Mako Hill

University of Zurich: Jan B. Schmutz

Utrecht University: Anna Salomons

Warwick Business School: Hila Lifshitz-Assaf

Yale University: Matthew Botvinick, Pascual Restrepo

York University: Xijuan Zhang

Zhejiang University: Jeffrey A. Brooks, Xia Fang

*Author affiliations reflect positions held at the time of publication. For individuals with dual appointments, names appear under each respective institution. Coauthors without university affiliations have been excluded from this list.

Google DeepMind