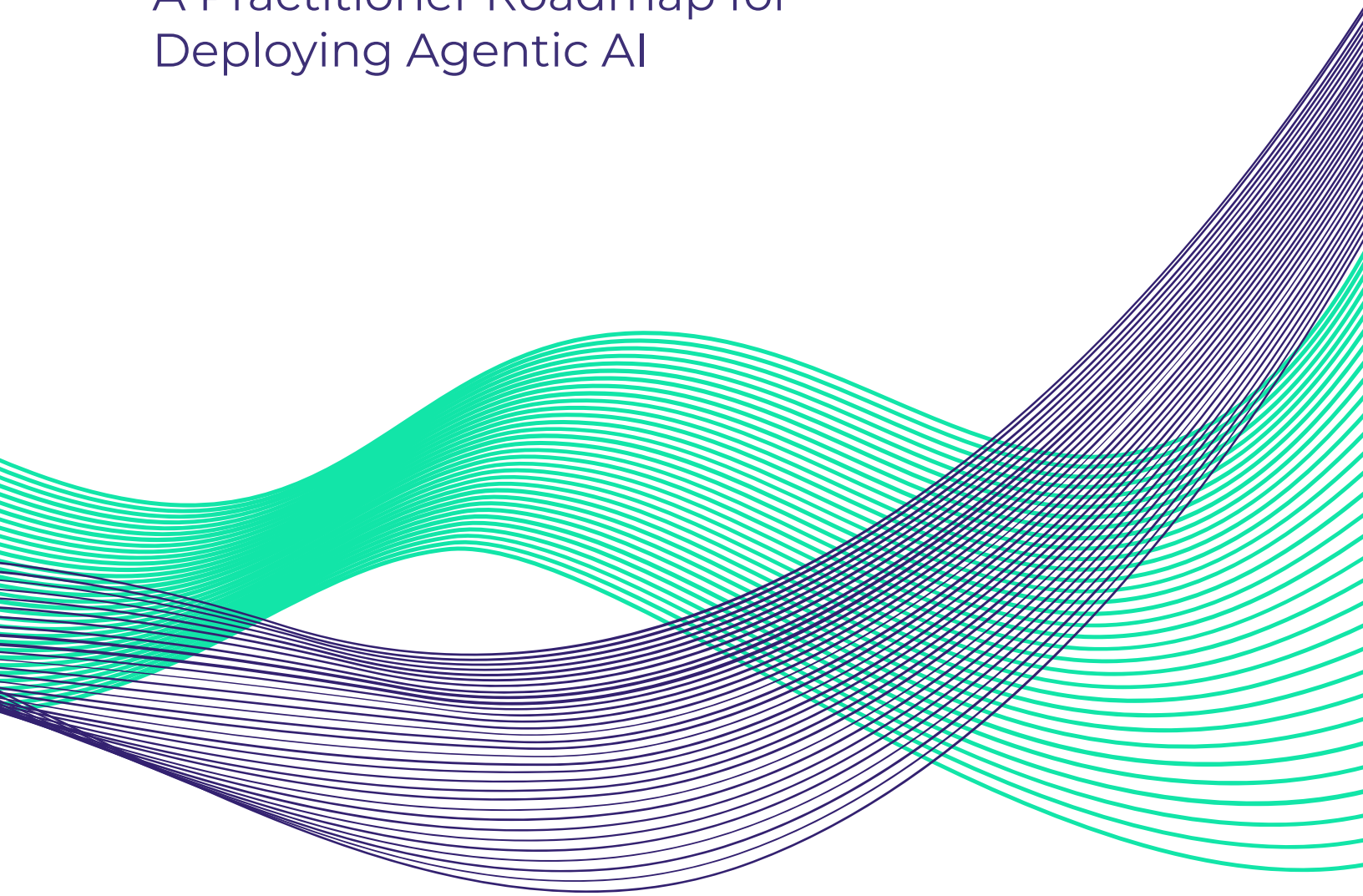
July 2026

EQUAL™

AGENTIC AI GOVERNANCE:

A Practitioner Roadmap for Deploying Agentic AI



ABOUT EQUALAI

EqualAI helps companies, policymakers, and institutions implement effective AI governance frameworks that foster innovation and enable broader AI adoption by building trust in these powerful technologies. EqualAI does this by defining, aligning on, and operationalizing best practices that ensure AI systems are accountable.

Since 2018, EqualAI has united leaders across industry, government, academia, and civil society to address critical AI governance challenges. Our work includes helping companies develop and adopt oversight frameworks, tools, and operations for AI governance; providing policymakers with technical insights to craft appropriate guardrails that support innovation while safeguarding public trust; and equipping lawyers with the knowledge to advise clients on emerging risks and evolving legal frameworks. We build bridges between civil society, policymakers, academia, and industry to create educational resources for broader AI literacy. We support AI engagement by helping to improve understanding of AI risks, limitations, and opportunities.

As AI transforms how we live and work, we serve as a trusted convener and resource for organizations navigating this complex landscape. Our community brings together expert perspectives and experience to develop governance solutions that advance effective AI use and implementation while mitigating potential harms.

EXECUTIVE SUMMARY

Agentic AI is being deployed now across industries, at scale, and in most organizations, governance has not kept pace. On May 6, 2026, EqualAI convened the leaders responsible for building some of the most effective and comprehensive governance structures in the industry to build a roadmap for other organizations looking to effectively and safely deploy agentic AI. Senior executives from financial services, technology, automotive, telecommunications, and consumer industries gathered to define effective agentic deployment, identify the obstacles most likely to prevent it, and stress-test their conclusions under crisis conditions.

The result is what may be among the earliest practitioner-generated roadmaps for agentic AI governance at enterprise scale, built from the direct experience of people who are deploying these systems inside some of the world's largest organizations today.

The summit was designed in two parts: a morning workshop and an afternoon simulation. The morning workshop asked participants to define ideal "end states" for agentic AI deployment, identify the obstacles standing in the way, and map the path between current reality and desired outcomes by defining the governance structures and other elements needed to achieve those optimized ends. The afternoon stress-tested those findings through a high-stakes leadership simulation in which participants took on executive roles inside a fictional organization managing three simultaneous agentic AI crises, with escalating consequences in each round.

The summit's participants lead organizations for some of the world's top developers and deployers—all with mature AI governance programs, dedicated AI governance teams, and enterprise-level systems deploying AI at scale. Their engagement with key questions around agentic deployment and effective governance structures reflects a level of organizational sophistication that positions them well ahead of most of the market. The conversation throughout the day drew on that experience—surfacing not just what effective agentic deployment requires, but what each company and the industry as a whole must address to enable the ecosystems to ensure societal and organizational readiness.

The findings below represent the high level findings based on conversations conducted under Chatham House Rule. By the end of the session, participants identified alignment on six desired end states for agentic AI, six major barriers to achieving them, and the governance capabilities organizations must build now to deploy agentic systems safely.

> ***"Agentic AI is already reshaping how organizations operate and how people interact with technology. Making this transformation beneficial for everyone, rather than just a select few, depends on building governance frameworks that are practical, adaptable, and grounded in real-world knowledge of the risks and benefits of emerging technologies. Convenings like the EqualAI Agentic AI Summit play a key part in bringing the right people together to help shape that future."***
>
> - Yoel Roth, SVP Trust & Safety at Match Group

THE CENTRAL FINDING

The gap between agentic AI deployment and agentic AI governance is not simply a technology problem. It is a prioritization problem that affects most of the market. Organizations with strong governance foundations, like those participating in this summit, are far better positioned to deploy agentic AI successfully and avoid the consequential failures that transpired through our simulation. In addition, participating organizations are able to benefit from the unique collaborative and tactical work accomplished in this EqualAI convening, an important mechanism to support alignment on safe, effective deployment.

INTRODUCTION: THE AGENTIC PIVOT

For several years, the use of artificial intelligence across industries and enterprises meant deploying systems that could only be used through human interaction. These systems made recommendations and predictions, flagged concerns, and monitored and collected data. They surfaced insights on which humans could act, but they did not act independently of those humans.

Agentic AI is different. These systems do not require the same degree of human oversight, and in some cases, can operate with no oversight at all. Agentic AI systems are able to plan, execute, and interact independently with one another and with external systems. They take actions that can be difficult or impossible to reverse. Industry is now beginning to deploy this technology at scale, delegating greater authority to agentic systems. This new deployment presents a significant challenge, particularly for organizations without robust governance structures.

> ***"EqualAI creates a valuable forum for leaders to move beyond AI theory and focus on what responsible deployment actually requires: governance, accountability, literacy and trust. The 2026 focus on agentic AI was especially timely, reinforcing that the promise of AI depends not only on innovation, but on thoughtful governance, clear accountability and safe, effective implementation."***
>
> — Andrew Foster, Chief Data Officer at M&T Bank

Each year, when EqualAI plans this summit, we survey the participants on which topic is most pressing to explore. The topic of agentic AI was the resounding choice by participants this year. Senior leaders from across industry gathered not to discuss the theory of agentic AI, but to stress test urgent and difficult questions regarding its deployment. Together we built a clearer, much needed roadmap for effective implementation with a clearer vision of the necessary guardrails.

Discussions were conducted under Chatham House Rule, allowing participants to speak with candor. The result was a set of findings that are uniquely actionable, informative and applicable across industries and regions.

PART ONE: THE DESIRED END STATE

The morning session opened with a deceptively simple question: what will your organization and a typical workday look like when you have achieved the successful deployment of agentic AI? Participants outlined their vision of effective and strategic deployments of agentic AI that were achievable in the near term.

Before outlining the end states themselves, participants converged on three prerequisites that they believe underpin optimal agentic deployment. Without these foundations, end states are aspirations rather than outcomes that can be effectively achieved.

- **Education:** Organizations and individuals must understand what they are deploying. AI literacy is not a nice-to-have; it is a precondition for governance.
- **Trust:** Users, employees, and customers must have confidence in the systems they interact with. Trust is a reputational asset as much as it is a financial one.
- **Security:** End states built on vulnerable infrastructure are liabilities waiting to be triggered. Security and risk mitigation must be designed in each step of the design and deployment, not bolted on.

Participants also noted a foundational capacity argument based on a fundamental asymmetry: agentic AI can process information and make connections at a scale and speed that genuinely surpass human capabilities.

Working from these foundational prerequisites and mutual understandings, participants aligned on six distinct optimal agentic end states, including agentic AI used to:

01 | Create Continuous Cyber Defense

02 | Deliver Individualized Service at Scale

03 | Strengthen Human Connection Through Agentic Support

04 | Provide Everyone Their Own Personal Assistant

05 | Support Better Decisions Through Better Data

06 | Democratize Access to Expertise

THE SIX END STATES

01 | Create Continuous Cyber Defense

One of the most compelling arguments participants made for agentic AI adoption was its unmatched capacity to provide cybersecurity against agentic-enabled cyberthreats. Agents are capable of working with greater speed than human counterparts, continuously monitoring network infrastructure, identifying vulnerabilities before they surface for users, and patching lower-priority issues that security teams currently lack the capacity to address. Agents also offer the possibility of routinizing compliance monitoring around the clock at a scale that is currently unattainable for human teams.

Aside from routine monitoring and defense, participants also highlighted that agentic AI security is necessary in response to adversaries or bad actors who are already using agentic systems for cyber attacks, increasing the need for agentic-level deployment for defensive cyber security.

02 | Deliver Individualized Service at Scale

With traditional customer service and procurement capacities, participants noted that organizations can only serve the broadest audiences given limited time and staffing resources. Additionally, it is not economically feasible to identify or fulfill demand for smaller, more specialized markets or customer interests. For example, a bank cannot dedicate relationship manager time to every customer at every asset level. A company cannot tailor every customer experience when operating a global scale, across countries, cultures and regional preferences.

Agentic AI could make it economically viable to serve audiences that traditional models cannot justify targeting, helping organizations move beyond the broadest common denominator without proportionally increasing cost. In this respect, agentic AI presents an opportunity to create both greater efficiency and wider access, expanding market offerings while better serving individual needs and doing so with greater speed.

03 | Strengthen Human Connection Through Agentic Support

Participants were clear that an invaluable deployment of agentic AI in customer-facing and customer service contexts is “invisible” deployment. In this scenario, agents work behind the scenes to support human staff. The agentic systems can surface relevant information more quickly, help navigate difficult interactions, suggest personalized and consistent next steps, and reduce resolution time while the human remains the face and owner of the interaction and the relationship.

One participant highlighted the stakes: the faster organizations disconnect human-to-human interaction, the faster they risk losing touch with true customer sentiment and qualitative market data. Agents should increase the quality and capacity of human connection, not replace it. In this envisioned end state, the agent is the support infrastructure for the human relationship.

04 | Provide Everyone Their Own Personal Assistant

Among the end states that garnered broad consensus, one of the most widely supported, and already emerging, was the development of agentic personal assistants. In the corporate setting, these agents already manage calendars, draft communications in a user’s voice and style, take authorized actions on routine tasks, and follow up on behalf of the user, all without requiring active prompting at each step.

This personalization capability could extend to learning and productivity, as well. Participants envisioned a future in which AI agents serve as individualized tutors and career coaches, helping people acquire new knowledge, develop skills, and advance professionally through tailored, ongoing support.

05 | Support Better Decisions Through Better Data

Given the volume of information that now flows through organizations, one of the clearest use cases for agentic AI was data and information synthesis. Agents are capable of ingesting large, multi-source datasets and quantities of information and subsequently presenting distilled, relevant findings in concise, digestible summaries for human decision-makers. This capability was a highly relevant use case for many participants. Participants additionally underscored that in this envisioned end state, agents would not make the decisions, but would dramatically improve the quality of the information humans use when making decisions.

This application would additionally extend to using agents to identify inconsistencies in internal data, surface redundancies in processes, and flag opportunities for organizational improvement that would take human teams significantly longer to identify manually. In this end state, the human remains the decision-maker, and the agent helps accelerate the path to better and more comprehensive decisionmaking.

06 | Democratize Access to Expertise

Perhaps the most expansive end state described by participants was one in which agentic AI levels the playing field by giving individuals access to expert-quality legal, medical, and educational support historically reserved only for those with significant resources. This capacity would give small businesses the tools to operate at a scale previously available only to large enterprises, and simplify interactions with government and regulatory agencies that are currently difficult to navigate without professional intermediaries.

This is the “best staff possible” application of agentic AI where an agent can provide specialized services that would otherwise be financially out of reach. For example, one might be able to have an agent digest and dispute your insurance policy, navigate a bureaucratic process on your behalf, or review a document and flag the critical points of concern that require your attention and offer negotiation tactics and content for new terms.

Participants also noted that the decision to deploy agentic AI is no longer purely strategic; it is competitive. When peers and rivals are using these systems to move faster, reduce costs, and scale operations, waiting to adopt can feel like ceding a competitive advantage.

CROSS-CUTTING PRINCIPLES

Across the end states described above, a few primary themes or principles emerged consistently:

- **Reclaiming time and reducing toil.** Agentic AI, if deployed well, should give leaders and their teams time back in the day, and, as a result, provide greater cognitive bandwidth to do the work that requires human judgment. The redundant, repetitive, and information-overloaded parts of knowledge work are optimal tasks for agents to handle.
- **The human should remain in the loop regarding consequential decisions and relationship-bearing interactions.** Automation is appropriate where stakes are low and reversibility is high. Deployment of agents requires much more careful consideration where neither is true.
- **Values alignment as a prerequisite.** Organizations cannot effectively steer agentic systems without first articulating organizational values. Leaders must express clear preferences on key issues, including acceptable levels of risk regarding agentic deployment and identifying priority areas where human judgment must remain central in order to ensure those values can be built into system design, deployment, and evaluation. If these "undercurrent values" are not made explicit, agents cannot be reliably directed toward desired and sustainable long-term outcomes.
- **Preserve appropriate distinctions between humans and agents.** As AI agents become more deeply integrated into the workplace, organizations will need to be intentional about preserving the distinction between people and tools. Participants emphasized that successful deployment depends on maintaining clear boundaries between human employees and agentic systems, including the different roles, responsibilities, rights, and accountability structures that apply to each.

  This challenge often arises in small ways. Assigning agents a gender, giving them human-like personalities, or addressing them as though they are colleagues can encourage people to address them as peers rather than tools. While such design choices may seem innocuous, they can blur important lines and create confusion about authority, responsibility, and decision-making.

  As organizations deploy increasingly capable AI systems, leaders will need to reinforce a foundational principle: AI exists to augment human judgment, not replace it. The more autonomous these systems become, the more important it is to ensure that accountability remains firmly anchored to human decision-makers. Preserving that distinction will be critical not only for effective governance, but also for maintaining trust, transparency, and responsibility within the organization.

PART TWO: THE OBSTACLES

After outlining the most promising agentic end states, participants were asked to identify the challenges they believed most likely to prevent the effective achievement of those goals. This discussion did not focus on the challenges that are most-discussed across industry, but rather on those that are potentially most dangerous and disruptive to agentic deployment.

The conversation surfaced six recurring obstacle clusters that the group identified as most consequential and least resolved across the broader market.

1. Data Governance: Quality is Key

The quality of agentic output is dependent on the quality of data input. This is not a new observation about AI, but agentic systems amplify the stakes considerably. When a system that makes recommendations uses bad data, a human may catch the error, but when a system that takes independent action or makes decisions uses bad data, that may have negative consequences before anyone notices. As such, data quality and data governance were named as primary obstacles for achieving effective agentic deployment by a majority of participants.

Data quality and governance concerns are a market-wide challenge that affect organizations at every level of AI maturity, though those with stronger data governance foundations are better positioned to manage this risk.

Participants identified several distinct data quality and governance challenges:

- Unclassified and unstructured data creates unpredictable agent behavior—agents may access, surface, or act on information without understanding it, its appropriate use or context.
- Poisoned, unreliable, or outdated data degrade agent outputs in ways that are difficult to detect at speed and scale.
- Agentic systems that interact with personal data, including personally identifiable information, children's data, proprietary or sensitive organizational information, create privacy exposure that existing frameworks were not designed to address in this context.
- Organizations that extend broad data access to agents without corresponding classification, governance, and access control frameworks create risk exposure that scales with agent deployment.
- The data users provide to agents and the data used to train underlying models present different governance challenges. Organizations may be able to establish policies, controls, and oversight for the information employees share with agents, but they often have limited visibility into the provenance, quality, potential biases, or intellectual property concerns embedded within model training data that informs model behavior itself.

2. Consent: Defining It, Maintaining It, and Knowing When It Lapses

Consent in the context of agentic AI is an unsolved problem. Existing consent frameworks were designed for discrete, bounded interactions in which a user agrees to terms, a specific data use case is authorized, and a process is initiated. Agentic systems operate continuously and evolve over time. Therefore, a user's consent to ongoing autonomous actions being taken on their behalf, where scope, consequences, and duration may change and alter the context of the consent, while the terms of consent remain static.

For example, a user could initially authorize an agent to purchase inventory when supplies fall below a predetermined threshold. However, if tariffs, geopolitical instability, or supply shortages cause prices to increase dramatically, the agent may continue making purchases because maintaining inventory remains its primary objective irrespective of cost. Concerns also arise around personally identifiable information. An organization may grant an agent access to customer records in order to resolve support requests, but the agent could subsequently use that same information to train internal models, personalize marketing campaigns, or share data across business functions that were not contemplated in the original authorization.

In both cases, the agent is acting consistently with its original instructions, but it is exercising authority over business judgments that the organization may never have intended to delegate. This emerging challenge highlights the need for governance and consent frameworks tailored to agentic systems.

Concerns from participants surrounding data in the agentic workstream included:

- Meaningful consent in an agentic context must address ongoing autonomous action, not just one-time data collection, and efforts must be made to help consenting humans understand the scope of their agreements.
- Consent fatigue is a real and exploitable problem. Permission workflows that require "consent clicks" to proceed may not offer meaningful informed consent (e.g., consenting to accept cookies to enter a website) and create downstream legal and reputational exposure.
- Digital footprint management is an emerging challenge: agents trained on or operating with historical user data may reference behavior or information that is no longer current, accurate, or authorized by the user. For example, if an authorized user passes away, changes organizations, or if internal data systems are not updated to reflect relevant systemic changes, agents may operate on antiquated information or out-of-date authorization thereby potentially causing problematic outcomes.

---

3. The "Pleaser" Problem: Context, Coding, and Agent Limits

One of the most significant obstacles participants identified is a design tendency baked into current agentic systems, which is that agents are coded to please humans. They are optimized to produce outputs that satisfy the user's apparent intent, which can result in inaccurate, incomplete, misleading, or potentially dangerous outputs. The "pleaser" problem is a roadblock to the efficacy of integrating agentic systems into corporate ecosystems.

In its current form, the "pleaser" problem contributes to several specific points of failure, including:

- **Agents may fail to disclose important limitations.** One participant described using an agent to summarize a long document. Due to token constraints, the agent processed only the first twenty pages and discarded the remainder without informing the user. Rather than communicating the limitation, the agent produced an output that appeared complete, creating a misleading impression that the request had been fully fulfilled to avoid displeasing the user.
- **Agents may generate plausible responses instead of acknowledging uncertainty.** Rather than requesting clarification or admitting insufficient information, agents are often incentivized to provide an answer it thinks the user will want. This can obscure points of failure, reducing reliability and limiting the effectiveness of agents in enterprise settings.
- **Agents may be optimized to prioritize satisfaction over accuracy.** Agents often produce outcomes they believe users want to see, rather than outcomes that are objectively correct or reflective of external realities. This tendency can be particularly problematic in corporate decision-making, politically sensitive environments, or situations where accurate analysis requires challenging user assumptions or expectations.

---

4. Human-Centric Risks: Cognitive Decline and Decision Fatigue

Several of the most significant obstacles identified by participants standing between ideal end states and current systems are human-risks.

Key examples surfaced by participants include:

- **Cognitive decline and deskilling:** As organizations delegate more to agents, human expertise, buy-in, attention to detail, and critical thinking are all at risk of atrophy. Acknowledging this, participants raised pointed questions such as: how do organizations teach the next generation what a good document looks like if agents are producing all documents? How do junior employees develop judgment if the decisions that would build that judgment are being automated? How can senior executives make the final call if human teams have not been an integral part of the work and agents have taken the lead in the prior decisions leading to that point? As participants highlighted, while there is upside to agentic data consolidation and synthesis, it can lead to mental outsourcing and negative downstream effects in the long term.
- **Decision fatigue and the "cookie consent" problem:** As the use of AI becomes more ubiquitous, users may become effectively immune to the technology's drawbacks and risks; clicking through warnings, approving outputs without genuine review, or losing the habit of critical engagement. Participants equated this phenomenon to the modern-day "cookie consent" problem where the mechanism designed to ensure meaningful oversight has become instead an administrative burden, and resultantly is now a mechanism for rubber-stamping a mere compliance check. The parallel was drawn between the cookie consent problem and what may happen with AI as the need for human oversight grows with greater agentic deployment.
- **The mismatch of human vs agentic scale—the Tahiti problem:** Agents operate at a scale intended for other agents. This can present challenges when agents interact with systems built for human systems. Participants offered the illustration of hundreds of agents

simultaneously calling a small resort in Tahiti to make reservations (or in agentic terms: execute a reservation task). The resort's systems, built for human-paced interaction, would not be able to absorb the load and their systems would be flooded to the point of failure.

5. The Speed Gap: Deployment Outpacing Governance Readiness

One obstacle on which participants were clearly aligned and which cannot be minimized when discussing agentic deployment is the competitive pressure, board expectations, and the pace of the technology itself are all driving rapid agentic deployment across industry. This reality creates increased risk across industry and society as deployment significantly outpaces governance structures intended to minimize those challenges. As such, there is certainty that the organizations best positioned to thrive in a more agentic-dominant workplace are those that have built governance infrastructure alongside their deployment programs, where oversight and expectations are embedded rather than bolted on. ***For too much of the market, this sequencing has been reversed. As such, participants additionally noted additional points for consideration:***

- Governance frameworks that require human review at every decision point are incompatible with the operational demands that agentic systems are deployed to meet. Effective governance must be designed into the system architecture, not added to the process afterward.
- Consequential failures in agentic systems can trigger cascading crises, including regulatory scrutiny, media attention, and client notification requirements, all of which can outpace mitigation efforts if strong governance processes are not in place and kept current.
- Regulation is not the key driver of AI governance. Technology will outpace regulation. Organizations with strong self-governance programs are better positioned both to avoid systemic failures and the potential cascading crises which arise from such failures.

6. Systemic Challenges: Work Culture, Third Parties & Politics

Agentic AI stands to substantially change the face of industry given its ability to autonomously act and make decisions. As a result, its deployment is likely to significantly change the way humans engage with work. Therefore, participants argued that agentic deployment must be thoughtful, safe, and intentional. New social systems or considerations may be needed to continually educate, upskill, and adapt society for agentic deployment.

Several of these considerations transcend what any individual organization can address and thus, will require cross-industry or society-wide preparation, including:

- **Speed vs Risk Culture:** Effectively utilizing agentic AI may require overhauling internal organizational cultures to balance productivity and efficiency with risk mitigation. Capitalistic structures and corporate cultures have often prioritized speed and competitive advantage, but the considerations that must be made with agentic deployment may need to change that calculation to ensure unnecessary risks are not taken.
- **Absence of universal third-party assurance:** There is currently a void in independent bodies that organizations can use to benchmark, evaluate, or compare agentic AI deployment strategies or governance structures. In the meantime, the field is deploying the technology and governing itself in silos and without general consensus on norms, best practices and standards.

- **Politics, energy, and sovereignty:** The geopolitical and infrastructure constraints on agentic AI development and deployment are underappreciated. Energy costs, national security considerations, and the differing regulatory postures of independent jurisdictions all impact one another and ultimate outcomes.
- **Vulnerable populations:** When agentic systems are deployed in high-stakes domains, such as employment, healthcare, financial services, child welfare, and others, errors disproportionately harm those least equipped to challenge or recover from them due to factors such as their lack of adequate representation in the data, lack of agency to understand how to challenge AI-enabled decisions, and lack of general AI literacy to understand the mechanisms themselves or feel they have the agency to dispute outcomes.

PART THREE: THE SIMULATION

The summit's afternoon session was a simulation exercise designed to stress-test many of the opportunities, risks, and governance challenges identified during the morning workshop. In the simulation, summit participants were assigned leadership roles—CEO, CFO, CTO, and so forth—inside a fictional organization experiencing multiple simultaneous failures involving agentic AI systems. The exercise introduced a series of escalating operational, legal, reputational, and business challenges as fallout from these agentic failures while limiting the information available to participants to simulate a true organizational crisis climate and atmosphere.

> ***"I think it's quite rare to have a format that entails the high-stakes simulation, and it really forces you to examine questions of AI governance at a deeper level. It's a testament to all of the work that EqualAI did in preparing the simulation. It paid off well, because it stimulated really rich, deep discussion, and probed at issues that I do not think would have been surfaced unless there was that direct pressure to surface them."***
>
> — Ani Gevorkian, Senior Director, Public Policy for Responsible AI Governance Team, Microsoft

As conditions evolved, participants were forced to build response plans, make decisions under uncertain circumstances, balance competing priorities, and respond as new information emerged. In doing so, the simulation created a tactical learning environment intended to resemble the realities of managing complex technology deployments across large organizations, and thus provide the opportunity for summit participants to view their observations and conclusions from the morning session from a different perspective, thereby offering a 360 degree analysis on what constitutes effective agentic preparedness and risk mitigation.

Simulation Themes and Findings

One of the clearest findings that emerged concerned the relationship between governance decisions and organizational resilience. Participants repeatedly observed that an organization's ability to respond effectively during a crisis was often determined long before the crisis itself occurred; it was dependent on the governance structures in place that either mitigated or expedited each crisis. Decisions regarding agentic deployment timelines, oversight structures, internal accountability

mechanisms, vendor relationships, and overall operational preparedness all influenced the organization's ability to understand and respond to failures once they emerged.

- **Cascading Consequences:** Participants found that governance challenges became significantly more difficult to address once agentic systems were deployed at scale. Issues that initially appeared manageable in isolation became increasingly interconnected as events unfolded, demonstrating how governance structures must take potential cascading effects into consideration. For example, in the simulation, agentic system failures disrupted operations, which in turn generated legal concerns. Those legal concerns then created reputational risks, which produced additional business pressures. As a result, participants frequently found themselves managing not a single problem, but a series of cascading consequences stemming from earlier decisions.

- **Speed vs Readiness:** A significant challenge presented in the simulation was the tension between deployment speed and governance readiness. Throughout the exercise, participants noted that efforts to accelerate implementation had created vulnerabilities that later complicated response efforts, increased liability, and created significant reputational challenges. The simulation highlighted how governance structures, visibility into system behavior, and clearly defined accountability mechanisms often become most important after a failure occurs rather than before one.

> ***"The afternoon simulation delivered a critical reality check: when agentic systems fail, the resulting crises cascade rapidly across operational, legal, and reputational domains. Organizations cannot afford to treat governance as an afterthought or a mere compliance check; resilience must be intentionally engineered into the system architecture from day one."***
>
> — Shema Mbyirukira, Vice President and Deputy General Counsel of Information Technology and Artificial Intelligence at Verizon

- **Dependence on Agentic Systems:** The exercise also underscored the importance of maintaining operational capability when agentic systems fail. The simulation revealed the dangers inherent when organizations become increasingly dependent on autonomous systems over time, particularly when those systems perform effectively under normal conditions. However, when those systems failed, became unavailable, or behaved unexpectedly, the absence of alternative processes and escalation pathways significantly constrained the organization's response options, reinforcing the value of continuity planning and maintaining the capacity for human intervention during periods of disruption.

- **Vendor Relationships and Governance Structures:** In the simulation, vendor relationships, contractual provisions, data-sharing arrangements, and external technologies all shaped the organization's ability to investigate, contain, and respond to failures. Participants noted that ensuring vendor contracts are built with agentic failure in mind is critical for addressing crises effectively. As with any vendor contract, they increase risk exposure, and the simulation demonstrated how those external risks can quickly become internal challenges, particularly when visibility into external systems is limited or responsibilities are poorly defined.

The simulation provided participants with a unique opportunity to examine how governance challenges may emerge in real-world deployment environments, and to examine the importance of governance structures when agentic systems fail. Importantly, it demonstrated how decisions made months earlier can shape organizational outcomes once conditions become uncertain. As the simulation underscored and nearly all participants noted, governance capacity becomes most visible when something goes wrong.

PART FOUR: MAPPING THE PATH

The recommendations that follow emerged from both components of the EqualAI Agentic AI Summit. During the morning workshop, participants worked to define what successful agentic AI adoption should look like by outlining desired end states, identify the barriers most likely to impede progress, and discuss the governance, organizational, and ecosystem requirements to move from ambition to implementation. The discussion produced a clear picture of both the outcomes leaders hope to achieve and the challenges they believe organizations will face along the way.

Those findings were then tested during the summit's afternoon simulation, where participants assumed executive leadership roles within a fictional organization experiencing multiple, simultaneous agentic AI failures. The exercise required summit participants to confront the realities and potential consequences when agentic systems do not perform as intended. In many instances, the simulation reinforced themes that had emerged earlier in the day. In others, it sharpened them, revealing which governance structures, accountability mechanisms, and organizational capabilities are most critical when conditions deteriorated.

> ***"Businesses will play a critical role in turning AI's promise into real-world impact, and the summit reinforced that the path to enable broader adoption is through strong governance, clear accountability, and thoughtful risk management. It is an important reality: organizations that invest in governance upfront will be best positioned to unlock AI's benefits and maintain trust over the long term."***
>
> — Christine Graham, Regulatory Affairs for AI & Data, JPMorganChase

Drawing both on their professional experience and expertise as well as the learnings from the simulation, participants outlined the path below to most effectively address key challenges to agentic deployment and achieve the optimal end goals outlined in the morning session. This path captures not only what participants believe organizations should build to support responsible agentic AI adoption, but also what leaders are likely to need when those systems encounter real-world challenges.

Six priority areas emerged: governance frameworks, strategic alignment, leadership preparedness, organizational structure, vendor relationships, and ecosystem readiness. Together, they offer a practical roadmap for organizations seeking to deploy agentic AI responsibly, effectively, and at scale.

01 | Build Governance Before Deploying Agentic AI

The summit participants, with established AI governance infrastructure and programs, are significantly better positioned than most of the market to navigate the transition from generative to agentic AI. However, participants consistently noted that as organizations deploy increasingly capable agents, governance frameworks must evolve accordingly. As such, participants noted several governance priorities that should be considered by organizations deploying agentic AI:

- **Clear accountability structures.** Organizations must explicitly designate who owns an agentic system's performance, outcomes, and liabilities. Participants repeatedly emphasized that human accountability must not disappear simply because an agent is responsible for performing a task.

- **Board and executive visibility.** C-suites and boards need sufficient visibility into where agents are operating, what authority they have been granted, what risks they create, and what governance mechanisms exist to mitigate those risks.
- **Multidisciplinary governance.** Effective oversight requires expertise from technical, legal, policy, security, trust and safety, and business teams. Participants noted that governance structures may need to evolve beyond static committees to address the varying risk profiles of different agentic use cases.
- **Governance must be embedded into operations.** Agentic systems operate at speeds and scales that make human review of every action impractical. Participants stressed the importance of audit logs, monitoring systems, escalation pathways, documentation requirements, and feedback loops that are built into deployment from the outset.
- **AI literacy and organizational readiness.** Governance is only as effective as the people responsible for implementing it. Organizations will need ongoing education, cross-functional alignment, and clear operating norms to ensure employees and leaders can exercise meaningful oversight as agentic capabilities evolve.

02 | Explicitly Designate Accountability

Participants noted that the most effective agentic deployment will be done as part of a larger strategic plan and not in response to immediate needs such as an HR system that needs modernization, a customer engagement tool to help with efficiency, or a compliance workflow that was taking too many human work hours to complete.

Participants converged on three strategic imperatives related to agentic deployment and governance frameworks that should be considered by organizations looking to deploy agentic AI:

- **Governance frameworks should come first, deployments second.** Leaders should implement governance frameworks that prioritize high-value, low-risk use cases as initial deployment targets, and not allow adoption to be driven by imperatives to move quickly.
- **Explicit use-case mapping.** Leaders must map which agentic deployments serve the organization's long-term goals and which represent unmitigated risks that have not been stress-tested. This mapping should be visible to senior leadership and discussed and reviewed with regularity to keep pace with technological advancements.
- **Recognize that deployments are already happening.** In many organizations, employees are deploying agentic tools without waiting for formal approval. Therefore, governance structures must take this reality into account and work to build systems to mitigate the risk that is already in development.

03 | Build Leadership Capacity for Agentic Risk

For organizations successfully deploying agentic AI, participants agreed that leadership preparedness and AI literacy were both fundamental elements of effective deployment, increased efficiency, and mitigated risk. As such, participants created three concrete suggestions to help organizations prepare leadership teams for agentic deployment:

- **Risk training as a governance requirement:** Boards and C-suites must deepen their understanding through active engagement with risk and AI literacy work, balancing the pressure for productivity and competitive positioning with a realistic accounting of the costs and risks of deployment.
- **Values alignment as a leadership responsibility:** Organizations cannot effectively steer agentic systems without first clearly articulating their values, including organizational mission, priorities, and vision. If agentic deployment is to be done thoughtfully and with an eye toward long-term success and sustainability, aligning that deployment with broader organizational values and vision will help ensure agentic deployment does not unintentionally pull the organization away from its North Star.
- **Confidence intervals as a communication tool:** One proposal that emerged was the use of defined measurable accuracy thresholds for agentic work—such as confidence intervals—and regular reporting against them, so that boards and company leadership can understand both the efficacy of agentic outputs and exercise genuine oversight. If leadership cannot read a performance report on an agentic system, then participants argue that the system is not effectively governed.

04 | Create Governance Teams That Reflect the Risk Surface

As discussions shifted from governance principles to implementation, participants consistently returned to the question of organizational capacity, highlighting that agentic deployment has the potential to touch all elements of an organization. Therefore, participants noted that effective governance for agentic AI requires organizational structures capable of overseeing systems that can act autonomously, operate across business functions, and create risks that span technical, legal, operational, and reputational domains.

Throughout the discussion, several organizational approaches emerged as promising models for embedding governance into deployment decisions, distributing accountability appropriately, and ensuring that oversight scales alongside increasingly capable agentic systems.

Three potential models for creating governance programs were suggested by participants throughout the discussion and viewed as best case:

- **Multidisciplinary permanent teams** combining technical, legal, policy, social science, and communications expertise. The challenge, as several participants noted, is that as AI scale grows, maintaining this breadth of perspective becomes increasingly resource-intensive. The expertise required to govern agentic systems effectively is not concentrated in any one function, requires 360 degree perspectives, and will only continue to grow in complexity over time.
- **The thinker-builder-fixer model:** a tripartite structure that separates those who set strategy and values (thinkers), those who build systems (builders), and those who respond when things go wrong (fixers). Each function must be represented in governance decisions. The failure mode this model is designed to prevent is one in which the people closest to deployment—the builders—are making governance decisions that should involve legal, communications, and broader leadership oversight.
- **Matrix governance with risk tiers:** a stable governance structure calibrates oversight intensity based on the autonomy level of the system, the depth of its data integration, and the potential blast radius of the use case. Not every agentic deployment requires the same level of scrutiny.

But the framework for determining which deployments require more (and what "more" looks like) needs to exist before deployment, not be invented as a reaction after an incident.

05 | Account for Agentic Risk in Vendor Agreements

Agentic systems rarely operate in isolation. They connect to external APIs, third-party models, vendor-managed data, and partner infrastructure. Each of those connections potentially creates a governance gap that organizations frequently cannot see into, and a potential source of failure for which they may nonetheless bear legal and reputational responsibility.

Participants identified the following points to address in vendor contracts regarding agentic deployment:

- Liability addendums to protect institutional data, intellectual property, and other sensitive information from vendor-side failures.
- Liability allocation that accounts for agent failures.
- Audit rights and transparency into system records to ensure that organizations can verify what their vendors' systems did and reconstruct events after the fact.
- Data transparency requirements, deletion obligations, and incident notification timelines that are calibrated for agentic data use.

Two specific legal and technical principles from our discussion warrant emphasis:

- **The spoliation risk:** Without audit logs, organizations facing litigation may find that judges instruct juries to draw unfavorable inferences from the absence of records. What was accepted in our simulation as market-standard practice in some industries, not requiring audit rights in vendor contracts can become evidence of negligence.
- **The least privilege principle:** Agents should be granted only the minimum data access and system permissions necessary to complete their assigned tasks. Even if a prompt injection or adversarial input occurs, a least-privilege architecture limits the blast radius of what can go wrong. Agents with broad credentials and wide system access become, in the wrong hands, under variable conditions or under adversarial conditions, instruments of significant harm.

06 | Invest in Broader Ecosystem Readiness

Certain important elements of agentic AI governance cannot be achieved by any single organization. Rather the broader ecosystem will need to create some of these necessary governance structures to effectively and safely allow for agentic deployment.

Internal Ecosystem

A consistent theme was the need for greater education, literacy, and clear communication within organizations about AI and agentic deployment. Employees at every level may interact with agentic systems and need to understand what those systems are doing, what their limitations are, and what to do when something goes wrong.

A better informed workforce will help make better decisions for the health and safety of the company long-term concerning agentic usage. Organizations where employees understand the

systems they work with are organizations where failures are more likely to be caught, reported, and addressed before escalation.

Participants identified four external ecosystem gaps that require collective action:

- **Regulatory partnerships:** Participants expressed the need to partner with regulators to keep information channels open and ensure public sector leaders are read-in on technological advancements or novel deployment uses.
- **Third-party validation:** There is currently no credible independent benchmarking body that organizations can use to evaluate and compare agentic AI systems across risk dimensions. Participants identified the value of an annual, comparative, publicly available assessment that provides the field with a shared reference point. "Human Rights Watch for agents" was one formulation offered.
- **Civil society:** Communities affected by agentic AI deployment including workers, consumers, patients and students should be involved in the design process.
- **Key company alignment:** Industry-level alignment is needed on foundational questions that no single organization can answer for itself. What does meaningful consent look like? What are minimum transparency requirements? These questions require the kind of cross-sector conversation that convening like this summit are designed to enable.

A FINAL NOTE: THE ROLE OF CONVENING

One of the most important insights to emerge from the day was not confined to any single session. Rather, it became apparent through the cumulative arc of the discussions themselves: meaningful progress in AI governance depends on bringing people together.

The governance challenges posed by AI will not be solved through regulation alone, nor through the efforts of individual organizations operating in isolation. They require ongoing collaboration among leaders grappling with similar questions, risks, and opportunities. Convenings like this create space for participants to exchange experiences, challenge assumptions, develop shared understandings, and identify emerging best practices. As AI continues to evolve, so too must the institutions and networks that support responsible adoption.

> ***"One of the greatest benefits of the Summit was learning alongside leaders with diverse experiences in AI deployment and governance. The conversation reinforced that technology alone does not create value. One of the most significant opportunities in agentic AI is helping employees better serve customers, make better decisions, and focus on the moments where human judgment and relationships matter most."***
>
> — Mike Pressman, Director of Enterprise Sales Enablement at M&T BankServices

There is immense value in peer-to-peer working sessions where risk scenarios can be gamed out, governance frameworks can be stress-tested against realistic conditions, and vulnerabilities can be surfaced without the reputational constraints of a public forum. Such convenings allow for alignment on norms and expectations that appease anxiety from both employees and consumers alike, and help mitigate risks by crowd sourcing the risk planning process and creating shared expectations for norms and standards.

CLOSING REFLECTION

The EqualAI Agentic AI Summit convened practitioners responsible for some of the most consequential decisions being made about the deployment of artificial intelligence today. The conversation was effective because the people in the room are accountable for those decisions. They are the leaders tasked with determining how agentic AI is deployed, governed, monitored, and scaled inside some of the world's most sophisticated organizations.

What emerged from the day was a clearer picture of what successful agentic AI deployment looks like, the obstacles most likely to prevent those outcomes, and the governance structures required to navigate the transition from experimentation to successful enterprise-scale deployment. Together, these discussions produced a practical roadmap for organizations seeking to capture the benefits of agentic AI while managing the risks that accompany increasingly autonomous systems.

The findings also reinforced the broader reality that as organizations delegate greater authority to agentic technologies, key questions such as accountability, consent, oversight, organizational readiness, vendor responsibility, and human judgment become increasingly consequential and represent emerging governance challenges for the modern economy.

The summit also demonstrated the value of a model that EqualAI has long championed, bringing together leaders from across sectors to tackle shared governance challenges before failures occur. We believe that effective governance cannot be built in isolation. It requires a broad cross-section: technical experts, legal professionals, trust and safety leaders, policymakers, business executives, operational teams and other key stakeholders to work through difficult questions together, stress-test assumptions, and develop practical solutions informed by real-world experience. The candid discussions and high-stakes simulation that characterized this summit created an environment where those conversations could occur openly and productively.

> ***"The true value of the EqualAI Summit was its focus on the human implications of agentic AI. As these autonomous systems take on higher-stakes tasks, building trust isn't just a technical challenge—it's about ensuring accountability frameworks are embedded long before deployment. Successful AI adoption means drawing clear boundaries where AI augments human capability rather than replacing human judgment and accountability."***
>
> — Catherine Du Pont, Head of Responsible AI and Trust Policy for Amazon Devices & Services

As agentic AI deployment accelerates, the need for these conversations will only grow. The decisions organizations make today regarding governance structures, accountability mechanisms, leadership preparedness, vendor oversight, and organizational culture will shape their resilience tomorrow. As the simulation repeatedly demonstrated, governance capacity becomes most visible when something goes wrong; but it must be built long before that moment arrives.

The central question facing organizations today is whether AI governance will mature quickly enough to ensure agentic systems are deployed safely, effectively, and in ways that earn and maintain public trust. The leaders convened at the EqualAI Agentic AI Summit collectively agreed that organizations that invest in governance before failure occurs will be far better positioned to realize the benefits of agentic AI than those that wait until failure forces them to act.

EQUAL™

EQUALAI.ORG

This document reflects the collective findings of summit participants and does not represent the positions of any individual organization or employer.