Deloitte.

Issue Averted

From AI autonomy to accountable action

When AI acts on its own, who is responsible for mistakes?

When an AI system does more than assist, responsibility for its decisions still sits with the business.

At first, most artificial intelligence (AI) systems were used to support tasks directed by people. Now, agentic AI is beginning to take action on the business's behalf. It can move through tasks, respond to events, and carry out structured work with limited intervention. That is what makes it attractive to organizations under pressure to move faster. It is also what makes human oversight more important.

For Chief Legal Officers (CLOs), the issue is not whether AI can act more independently, but how human review and intervention are built into the process. When an AI system acts within business processes, responsibility does not shift to the technology. The system might get it wrong, but it cannot be held accountable. The organization still owns the outcome. That is why human oversight must remain embedded not only in design and deployment, but also in how the system operates and how its actions are reviewed after the fact. With the right governance, businesses can use agentic AI in a way that is accountable, explainable, and aligned with business objectives, with human judgment retained where it matters most. What matters is not just how the system is built, but how its actions are governed, reviewed, and owned in practice.

What you are trying to avoid

When autonomous systems make consequential decisions, uncertainty over who is responsible can quickly become a business problem.

As agentic AI becomes more capable, the risks become more practical. An AI agent can handle customer interactions, support operational workflows, even act within business processes that once required at least one, maybe more, people. That can improve speed and efficiency. It can also create new exposure when the system makes a decision the business cannot fully justify, explain, reverse, or contain.

So, what you are really trying to avoid is:

Unclear accountability

When an agentic AI system makes a poor decision, responsibility can quickly become blurred. Is the issue rooted in design, deployment, oversight, training data, escalation protocols, or business ownership? In practice, those questions may surface only after something has already gone wrong. They become harder to answer when the business has not decided where review is required, who can intervene, and when a person must step in before the system takes action. The organization's legal function (such as the in-house legal team or external legal advisers) can help determine in advance who is accountable for what, where approval thresholds sit, and when human intervention is required.

Regulatory and contractual exposure

When AI systems act within customer, employee, or commercial processes, legal and regulatory consequences can follow. In financial services, for example, an agentic AI tool that negotiates terms or supports contracting decisions could expose the business to unfavorable commitments or compliance issues. In healthcare, an AI system involved in appointment scheduling or patient-facing processes may create more serious consequences if it misses or mishandles critical information. In higher-risk use cases, a human-in-the-loop approach can help ensure that sensitive decisions, exceptions, and escalations receive the right level of review before they create legal or regulatory consequences.

Loss of trust

Trust may be one of the first things at risk when agentic AI is deployed without clear guardrails. Customers, regulators, employees, and business partners may be more willing to accept autonomous systems when the organization can explain how decisions are made, what controls are in place, and who remains responsible. They are also more likely to trust those systems when the business can show that human oversight remains part of the process where outcomes carry significant consequences. If the organization cannot answer those questions, reputational damage can outlast the original mistake.

How to avoid ugly outcomes

Legal can help turn agentic AI governance into an operating model, with defined roles, clearer escalation paths, and ongoing oversight.

Once you recognize what your organization wants to avoid, how do you respond? The legal department can help shape a practical governance model for agentic AI, one that supports innovation while creating clearer accountability and stronger control. Three actions stand out.

Define accountability before deployment. Legal should be involved early enough to help decide how responsibility will work before an agentic AI system is launched. That includes clarifying what the system is permitted to do, where approval by a team member is required, who owns the business process, and when issues must be escalated. If legal comes in only after the technology is built, the business may be left correcting governance gaps after key design choices have already been made.

This is where the oversight model should be defined. Not every AI-enabled workflow needs the same level of review, but the business should decide in advance which actions can proceed automatically, which require approval, and which demand escalation or override rights or even emergency stop systems. Legal can help determine which uses call for closer review because they affect customers, contracts, sensitive information, regulated activities, or other decisions with meaningful consequences.

How to avoid ugly outcomes

Embed legal into governance and design.
Agentic AI governance works better when Legal is part of the project team, not a checkpoint at the end. The financial services example earlier (page 3) makes the point well: When Legal and Technology teams work together from day one, data handling, consent, accountability, and regulatory constraints can be built into the solution earlier, reducing the need for redesign later.

That same principle applies more broadly. Legal can help shape policies for how agents are used, how their actions are recorded, what controls surround customer-facing or contract-related activities, and how the business responds when the system produces an outcome that needs review. This is also where Legal can help create a shared understanding between the CLO's team, technology leaders, and the business owners responsible for deployment. This is also an opportunity for CLOs to reframe Legal's role within the organization.
When Legal is present at the design stage, it is seen as an enabler rather than a gatekeeper, a distinction that matters both for the quality of the outcome and for Legal's standing in future technology decisions.

Monitor agentic AI as an ongoing obligation. Governance for agentic AI is not a one-time exercise. Systems evolve, business uses expand, and expectations change. What begins as a narrow use case may take on a broader role over time, or may begin interacting with new categories of data, customers, or transactions. For that reason, legal oversight should be designed as an ongoing process rather than a launch-stage review.

Your dynamic model should include continuous monitoring, explainability, and adaptable policies that evolve as lessons are learned. In practice, this means building in structured review points as the system matures. Monitoring outputs for unexpected patterns, maintaining records that allow the business to explain individual decisions, and revising the governance framework when the systems scope, data sources or user base changes. For CLOs, this means helping the business establish a framework for review, reassessment, and response as agentic AI systems mature in production. Policies that made sense at launch may need revision once the system encounters edge cases in production. The goal is to help the organization scale these tools with stronger control and greater confidence.

What your company gains

Early legal involvement can help organizations adopt agentic AI faster by reducing late-stage redesigns, confusion, and avoidable risk.

The value of stronger governance is not limited to avoiding problems. When legal helps shape the operating model for agentic AI from the start, the business may be able to move faster and with fewer setbacks. Instead of pausing late in the process to address gaps in oversight, accountability, or data handling, teams can build on a more durable foundation from the beginning.

That can produce several benefits. First, it can reduce costly redesign and rework by surfacing risks earlier in the lifecycle. Second, it can improve trust by making autonomous decision-making easier to explain to stakeholders inside and outside the organization. Third, it can create a more sustainable model for scaling AI use cases across functions, because the business is not starting from scratch each time a new system is introduced.

A well-designed human-in-the-loop approach can also help the business combine speed with oversight, allowing teams to move ahead with clearer control over where judgment, review, and intervention remain essential.

For CLOs, there is also a broader opportunity that goes beyond risk management. Agentic AI is changing how business decisions are made, and legal is well placed to shape that change rather than respond to it. Legal can help the organization respond by building governance that is practical, proportionate, and suited to systems that act with increasing independence. That is a different role for legal, and a more valuable one: not as a constraint on innovation but as a function that makes sustainable innovation possible. Done well, that is not simply a governance story. It is a story about how legal helps the business move forward with confidence and about why that contribution matters most before anything goes wrong.

Putting it into practice: Where legal leaders start and who to engage with

Knowing what good governance looks like is one thing. Building it is another. For CLOs and legal leaders who want to move from principle to practice, a few concrete steps can make the difference between governance that exists on paper and governance that actually shapes how agentic AI is used.

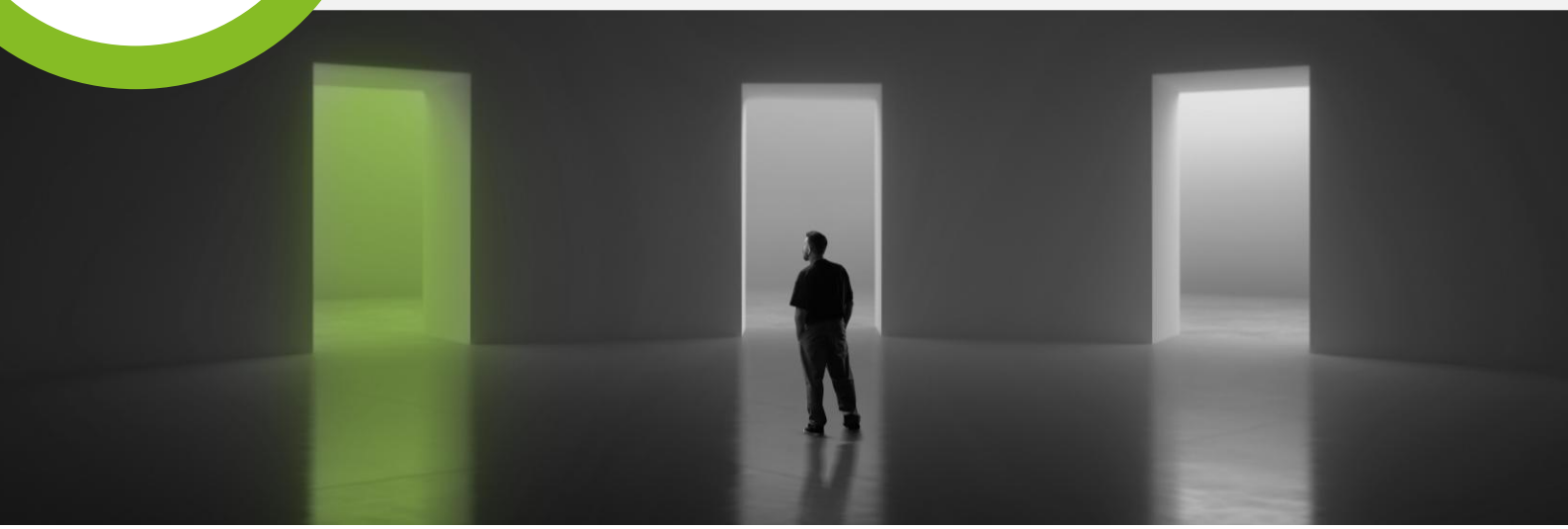
If not already done so, start by mapping the agentic AI systems already in use or in development across the business, many legal teams are surprised by how many exist and how little visibility they or the wider company have over them. From that inventory, identify the highest risk use cases: those touching customers, contracts, regulated activities, or sensitive data. These are the places where accountability frameworks, escalation rights, and human-in-the-loop requirements should be defined first. Next, secure a seat at the table for legal in any AI project team working on agentic systems, not as a final reviewer, but as a contributor from the outset.

Use that position to establish a small number of clear, non-negotiable governance standards, for example, what the system is permitted to do autonomously, what requires human approval, how decisions are logged, and who owns the process when something goes wrong.

Finally, build a rhythm of review into the operating model (e.g., quarterly or at key system milestones) so that governance evolves as the technology does.

None of this requires perfection at launch. It requires a starting point that is deliberate, documented, and owned. Translating these principles into practice is, however, as much a question of capability as it is of process.

Putting it into practice: Where legal leaders start and who to engage with



On the skills side, CLOs and their teams do not need to become technologists, but they do need enough fluency in how agentic AI systems work to ask the right questions: understanding concepts such as autonomous task execution, model behavior, data inputs, and escalation design is increasingly necessary to engage meaningfully with technology and product teams. This might actually mean that the CLO needs to hire those specialists lawyers, but also non-lawyers, to have these conversations and contributions at the right level.

Equally important is the ability to translate legal and regulatory requirements into operational design choices, moving between the language of risk and the language of system architecture, and to lead cross-functionally with data and engineering teams by communicating clearly, building shared understanding, and working through disagreement without defaulting to control.

On the relationships side, the governance model described in this article cannot be built by legal alone: CLOs will need close working relationships with Chief Technology Officers (CTOs) and Chief Information Officers (CIOs), but also those leading engineering and who control how systems are built and deployed.

Collaboration with Chief Risk Officers (CROs), whose frameworks for identifying and managing exposure should sit alongside legal's own accountability structures, and with business unit leaders who are accountable for the commercial, operational, and customer-facing outcomes closest to where agentic AI is deployed is equally important.

CLOs should also invest in their relationship with the board and executive leadership, since agentic AI raises questions about accountability and liability that ultimately sit at the top of the organization, and legal is well placed to help leadership understand what is at stake and what responsible adoption looks like in practice. For CLOs, the opportunity is to lead that effort before the business asks for it, because the organization's that move first on governance will be the ones best placed to move fastest on the technology itself, and the CLOs who build these skills and relationships now will be the ones best positioned to lead when the governance conversation becomes unavoidable.

Authors

Melinda Upton

Partner,
Deloitte Legal
United Kingdom

Richard Punt

Global Legal Leader,
Deloitte Legal
United Kingdom

Sebastiaan ter Wee

Partner,
Deloitte Legal
The Netherlands

Deloitte.

About Deloitte

Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited ("DTTL"), its global network of member firms, and their related entities (collectively, the "Deloitte organization"). DTTL (also referred to as "Deloitte Global") and each of its member firms and related entities are legally separate and independent entities, which cannot obligate or bind each other in respect of third parties. DTTL and each DTTL member firm and related entity is liable only for its own acts and omissions, and not those of each other. DTTL does not provide services to clients. Please see www.deloitte.com/about to learn more.

Deloitte Legal means the legal practices of DTTL member firms, their affiliates or their related entities that provide legal services. The exact nature of these relationships and provision of legal services differs by jurisdiction, to allow compliance with local laws and professional regulations. Each Deloitte Legal practice is legally separate and independent, and cannot obligate any other Deloitte Legal practice. Each Deloitte Legal practice is liable only for its own acts and omissions, and not those of other Deloitte Legal practices. For legal, regulatory and other reasons, not all member firms, their affiliates or their related entities provide legal services or are associated with Deloitte Legal practices.

This communication contains general information only, and none of Deloitte Touche Tohmatsu Limited ("DTTL"), its global network of member firms or their related entities (collectively, the "Deloitte organization") is, by means of this communication, rendering professional advice or services. Before making any decision or taking any action that may affect your finances or your business, you should consult a qualified professional adviser.

No representations, warranties or undertakings (express or implied) are given as to the accuracy or completeness of the information in this communication, and none of DTTL, its member firms, related entities, employees or agents shall be liable or responsible for any loss or damage whatsoever arising directly or indirectly in connection with any person relying on this communication. DTTL and each of its member firms, and their related entities, are legally separate and independent entities.