

AI in Regulation



This issue covers:

- the role of AI in regulatory work
- the importance of transparency
- the 2 types of AI most relevant to regulators

Read this alongside other issues in the bite-sized AI guidance series

Regulators exercise the state's power to compel and make decisions that have real consequences for people and organisations. They make decisions that affect people's rights, safety, and livelihoods. That creates distinct considerations for how AI is used.

AI can help regulatory work, not take it over

Regulatory practice depends on something AI cannot provide: careful judgement, legal interpretation, discretion, and decisions that affect people's rights and public trust. AI can handle the fast work of scanning, sorting, and summarising, but the slow work of weighing evidence, applying the law, and being accountable to the public remains with people. That is why human oversight is not optional ([see the **Organisational integration is key issue**](#)).

AI also introduces new categories of risk ([see the **Opportunities and risks issue**](#)). Bias can be hidden inside a model's training data and invisible in its outputs. Decisions can drift toward automation without anyone making a conscious choice to allow it. AI outputs can be difficult to explain to the people they affect, creating real challenges for transparency and appeal rights.

None of this transfers responsibility away from your organisation. Accountability remains with the legally authorised decision-makers, not with the tool and not with the vendor. Getting the governance arrangements right from the start makes this easier to uphold in practice.

Regulatory leaders will want to be confident they understand what AI is doing in their systems, and why, before placing significant weight on its outputs ([see the **Steps to success issue**](#)).

Be transparent in how AI is being used in decision-making

The bar for ethical AI use in a regulatory context is higher than in many other settings. Getting the technical and governance foundations right, as earlier sections describe, is necessary but not sufficient. Regulators also need to be clear about the principles that will guide how AI is used, and where it should not be used at all.

When using AI, regulators must set clear ethical boundaries that protect public trust and ensure decisions remain lawful, fair, explainable, and accountable.

In practice this means paying particular attention to transparency and explainability, and to fairness, bias, and privacy. Regulators have a duty to explain how and why decisions are made. That duty does not change when AI supports or informs those decisions. Where AI is used:

- regulators should be open about when AI is involved, what role it plays, and how particular models work
- decision-makers must understand how AI outputs were generated and their limitations
- affected parties should be able to understand how AI-informed decisions were reached.
- transparency supports fairness, protects public trust, and reduces legal and reputational risk.

Where AI has influenced a decision affecting a member of the public, that person has a right to seek information about how that decision was reached. This is an existing obligation under the Official Information Act 1982, and it has direct implications for how AI-supported processes are designed and documented.

Not all AI is the same, and the differences matter for regulators

The two types most relevant to regulatory work are predictive AI and generative AI.

Predictive AI learns patterns from existing data to classify, score, or forecast outcomes. A system that flags high-risk licence applications based on historical compliance data is an example. These systems can be powerful, but they inherit the biases of the data they were trained on. If past decisions were inconsistent or unfair, a predictive AI system will tend to repeat and amplify those patterns.

Generative AI produces new content, including text, summaries, draft documents, and responses to questions. Large Language Models (LLMs) like those behind tools such as Copilot and ChatGPT are generative AI. They are well suited to tasks like summarising information, drafting correspondence, or identifying themes across large volumes of text. However, they can also produce outputs that sound confident but are factually wrong. This is known as hallucination, where the system generates plausible-sounding content that has no basis in fact. Without human review, this risk can go undetected.

Both types of AI share an important characteristic: they learn from data, and their outputs reflect the quality, completeness, and representativeness of that data. Incomplete or biased data produces unreliable output.

To get the full set of practical steps for how regulatory leaders can lead AI innovation with confidence, check out the full guidance: [**Responsible AI in Action**](#).