AI, Power and Privacy: building trust for the future

Ivana Bartoletti

Keynote for the ICO Practitioner Conference, 14 October 2025

Introduction

Thank you to the Information Commissioner's Office for the invitation and for the work you do every day to protect privacy and enable responsible innovation in the UK.

Artificial intelligence offers extraordinary possibilities when trust is built in from the start, and your efforts are vital to making that trust real.

I came to privacy through human rights law, where I saw early on that **privacy and equity are inseparable**.
The people who feel the weight of data collection most acutely are almost always those already disadvantaged:

- families relying on loyalty schemes to save a few pounds and surrendering detailed shopping histories,
- women crossing borders for reproductive care and leaving digital traces,
- young activists exercising their right to protest and finding their movements tracked.

Privacy is not an abstract technical issue: it is a question of fairness and power.

Today I want to connect the sweeping headlines about artificial intelligence with the day to day reality you all know well: data governance, risk assessments, and the practical business of keeping people safe.
We hear that AI is racing toward artificial general intelligence and that privacy rules must relax because the tide cannot be held back.
Yet we also face constant breaches and very real human harms: facial recognition surveillance, intrusive workplace monitoring, threats to democratic debate.

<u>How do we reconcile these stories?</u>

1. **Cutting through the AI hype**

Speculation about AI's economic promise is intense.
Goldman Sachs asked a pointed question last year: *What trillion-dollar problem will AI solve?*

MIT researchers tracked 300 public AI projects and found **95 percent failed to improve profits**.
The return on investment is far from certain.

That does not mean there are no achievements.

UK researchers recently trained software on 800 stroke scans and tested it on 2,000 more.
The AI proved twice as accurate as specialists in reading scans and could identify when a stroke occurred, lifesaving knowledge.
In automotive manufacturing, AI now predicts mechanical failures and enables advanced driver-assistance systems.
And of course generative AI dominates the headlines, creating both excitement and anxiety.
Sequoia Capital has called these moments *speculative frenzies*, a normal part of technology cycles.

The lesson is to stay level-headed and build what matters after the hype fades.
That is where privacy professionals come in.
We stand between inflated expectations and the messy reality of deployment, often expected to enable rapid change while guarding against the very risks hype tends to ignore.

For example - Anthropic's recent research confirms what many in AI security feared: the **attack surface of large models is effectively infinite** and **shockingly easy to exploit**. Their experiments show that large-scale poisoning, the deliberate manipulation of training data, may require control of as few as **250 documents** to cause measurable corruption. This quantification exposes the fragility of current AI pipelines: the same mathematical dynamics that make hallucinations inevitable also make poisoning, prompt injection, and evasion attacks inescapable. In short, the cost and effort to compromise a model are dramatically lower than expected.

What does it mean for us and our community? It means realising that AI models are vulnerable in ways that traditional systems never were. Unlike human users, AI lacks stateful memory or contextual learning that can defend against manipulation.
Protecting privacy and integrity in this environment requires architectural redesign: provenance control over training data, authenticated data sources, continuous monitoring for anomalies, and strict evaluation of AI vendors' red-teaming and poisoning-resilience capabilities. AI governance must now include explicit **security and trust by design frameworks**, or the privacy foundation of digital systems will not hold.

2. How people really use AI

Consider how people actually engage with AI today.

OpenAI's large-scale study of ChatGPT shows **about 30 percent of use is for work and 70 percent for personal life**.
The greatest transformation is happening at the individual level: employees using chatbots to draft emails, students summarising lectures, families asking for recipes or medical information.

For privacy teams this means:

- **AI literacy is critical.**
  We must help colleagues understand that chats may not be private, that outputs can be wrong, and that personal data shared with a bot may be stored or reused.
- **Enterprise deployment must be secure.**
  Generative AI needs privacy-by-design, but also **dynamic, context-aware access controls**, moving beyond static permission models.
  Intelligent gateways should evaluate each AI interaction against risk factors: data sensitivity, user role, and potential aggregation.
- **Granular oversight is non negotiable.**
  Every AI interaction touching regulated information, health records, financial data, personal identifiers must be automatically checked against applicable law.
  Comprehensive audit logs should document compliance so organisations can respond quickly to regulatory inquiries.

These tasks may sound technical, but they have deep human implications.

If the biggest revolution is personal, then every misstep is personal too.

3. The limits of existing law

The law must now face a new reality.
The Internet of Things once worried us with devices that *transmitted* data.
Today's AI systems *initiate* interaction, simulate emotion, and build personalised engagement loops.

Children, and how they interact with this technology, are a clear example of this.
A 2024 study showed that **children aged three to six were more likely to trust a friendly robot than a human being**.
These systems are not just products: they are companions, teachers, even friends. The popularity of AI companions has increased dramatically in recent years: of the 337 active and revenue generating companion apps worldwide, 128 were released in 2025 alone. There is no doubt that emotional intimacy is the next lucrative frontier.

Two concerns follow:

1. **Legislation lags design.**
   Current data-protection laws regulate how information is handled but rarely address the **psychological impact** of tools designed to mimic empathy.
   The EU AI Act and Online Safety Act make a start by recognising manipulative design as a risk, but gaps remain.

2. **Techno-legal solutionism has limits.**
   Privacy-by-design is essential but cannot answer every question.
   Should some data be collected at all?
   If we outsource summarising and organising information to machines, what happens to the cognitive skills those tasks build?
   Mandating privacy enhancing technologies may reduce harm, but it cannot decide which technologies society should embrace in the first place.

Privacy teams often bear the brunt of these dilemmas, expected to engineer social outcomes through technical controls when what is really needed is public debate and political choice.

4. AI Agents: the next frontier

Chatbots may be experiments – **AI agents** are something else.
They can act autonomously across systems and databases.
Integrated into operating systems, they raise serious questions:

- Is privacy possible at the application layer when an agent can roam across multiple services?
- How do we prevent silent aggregation of data as agents chain tasks together?

Before adoption races ahead, we need firm controls:

- Developers must be able to designate certain applications as off-limits to agents.
- Transparency must be mandatory, so users know when an agent is acting on their behalf.
- Strong security must wrap every interaction so we can anticipate, not merely react to, vulnerabilities.

Booking a flight or a haircut can never take precedence over the right to privacy.

5. Privacy professionals as enablers

In many organisations privacy teams work alongside HR, security, and governance groups to shape responsible AI adoption.
It can feel as though not everything is about privacy, yet almost every serious AI harm is a privacy harm:

- discrimination driven by biased data,
- exposure of sensitive health or financial information,
- surveillance that chills free expression.

Our job is not to block progress.
It is to **enable sustainable innovation** by insisting on, and designing, guardrails that protect people and keep trust intact.

That role requires courage.
It means asking whether we should collect a dataset, not only how we protect it.
It means reminding boards and executives that privacy is a foundation of competitive advantage, not a drag on productivity.

6. Building a global framework

We must also lift our gaze beyond national boundaries.

AI has become a geopolitical instrument. At its core, it is **compute, models, data, and cloud**, and control over these elements now translates into political power.

Infrastructure sovereignty has turned into a form of geopolitical agency, signalling a decisive shift towards the politicisation of transnational data governance.

Although the United States and the European Union differ in their methods, both increasingly frame their digital strategies in the language of sovereignty, moving away from earlier ideals of regulatory convergence and multistakeholder collaboration.

AI is, therefore, far more than a technological question. Its ability to reshape economies, influence democratic systems, and redefine labour markets makes coherent global governance indispensable.

Institutions such as the **United Nations**, the **Council of Europe** (with which the ICO cooperates), national regulators, and private actors all share responsibility for setting clear boundaries

- banning manipulative AI applications,
- enforcing transparency for high-risk systems, and
- building mechanisms for cross-border accountability.

Privacy professionals must help shape that global framework.

Closing

Let me leave you with three thoughts:

- **Stay level-headed.** The AI boom will evolve, but the need for privacy and equity will only grow.
- **Lead with literacy.** Equip your organisations and communities to use AI wisely and critically.
- **Shape the debate.** Privacy professionals are not spoilers. We are the enablers of innovation that lasts.

Artificial intelligence will not define our future by itself.

The choices we make about privacy will.

Thank you.