Picture credit: Crown Copyright

ENSURING INTEGRITY

Rachel Roumeliotis *reveals how a lack of AI governance threatens to undermine all we're setting out to achieve*

At the height of the pandemic, the development and roll-out of the COVID-19 vaccine put the UK at the forefront of the fight against Coronavirus. Perhaps unbeknownst to much of the public was the fundamental role AI played in this success story. This transformative technology helped researchers confirm the efficacy of the vaccine, but also enabled swift manufacturing while delivering unparalleled insight into the virus' spread. AI's deployment in the life-saving vaccine programme provides a shining example of its role in protecting our population.

However, the use of AI still poses serious ethical and practical questions that governments, academics, citizens and organisations alike are still grappling with. Among the issues at play are how trust, transparency and fairness can be 'built' into AI systems. For example, recent global research found that the UK population is among the most sceptical of AI; just 35 percent said they trust a company using AI as much as they trust a company which does not. Concerns over opaque black-box algorithms to questions regarding the ethical use of personal data and responsibilities related to security and privacy have made AI a hotbed of modern ethical dilemmas.

Regardless of whether AI is developed for military use or to aid a commercial organisation, the issue of governance is crucial

The UK government is well aware of the transformative potential of AI, having recently published its ambitious National AI strategy, but is also keenly aware of the ethical issues raised by its proliferation. The document clearly stated the need for: "strong national and international governance of AI technologies to encourage innovation, investment and protect the public and our fundamental values". Hot on the national strategy's release, in June this year, the UK's Ministry of Defence (MoD) published its own Defence Artificial Intelligence Strategy, firmly positioning its ambitions to pursue: "strategic and operational advantage through AI, while upholding the standards, values and norms of the society we serve, and demonstrating trustworthiness".

MIRRORING HUMANS

When it comes to questions of privacy, security and transparency, nowhere is this more relevant than when it comes to life and death national security scenarios. As the Director of GCHQ recently stated: "Philosophers and data scientists have been grappling with the implications of AI for ethics: how do you ensure that fairness and accountability is embedded in these new systems? How do you prevent AI systems from replicating existing power imbalances and societal discrimination?"

For organisations keen to exploit AI benefits, these may feel like questions that fall outside of their remit. However, whether AI is developed for use on the military's front line or to aid a commercial organisation in its drive to enhance customer experience, the issue of governance is crucial.

According to government analysis, around 15 percent of all businesses have adopted at least one AI technology, with the IT and telecommunications (29.5 percent) and legal (29.2 percent) sectors currently reporting the highest rate of adoption. Spending on AI technologies is expecting to grow 16 percent by 2025. However, despite the burgeoning proliferation of AI in the enterprise, many organisations still lack strong AI governance crucial to ensuring the integrity and security of data-led systems.

In fact, the latest O'Reilly research shows that over half of AI products in production at global organisations still do not have a governance plan overseeing how projects are created, measured and observed.

Deeply concerning is that privacy and security are rated low in the risks organisations evaluate when considering AI applications. These are issues that may directly impact individuals and destroy public trust in these new technologies. AI-empowered organisations report that 'unexpected outcomes' are the most significant risk facing AI projects, followed closely by model interpretability and model degradation, representing business issues. Interpretability, privacy, fairness and safety are all ranked below business risks.

Of course, there may be AI applications where privacy and fairness are not issues (for example, an embedded system to decide whether dishes in your dishwasher are clean). However, companies with AI practices must prioritise the human impact as both an ethical imperative and a core business priority.

As UKRI (UK Research and Innovation) highlights: "responsible use of AI is proving to be a competitive differentiator and key success factor for the adoption of AI technologies. However, cultural challenges, and particularly the lack of trust, are still deemed to be the main obstacles preventing broader and faster adoption of AI."

Lack of governance is not just an ethical concern or privacy issue. Security is also a massive issue, with AI subject to unique risks: data poisoning, malicious inputs that generate false predictions and reverse engineering models to expose private information.

AI ADOPTION HAS NOT BEEN MATCHED BY AN INCREASE IN AWARENESS OF ITS RISKS

Research conducted by Forrester Consulting found that 88 percent of security decision-makers believe 'offensive AI' is on the horizon, with as many as two-thirds concerned about AI-led attacks. As always, cybercriminals are just as keen to exploit new technologies as anyone else. However, security remains close to the bottom of the list of perceived AI risks.

With cybercriminals and bad actors surging ahead in their adoption of sophisticated technology, cybersecurity cannot take a back seat in the race to realise AI's promise. It is a vital strand of much-needed AI governance. Governance must rise up the matrix of risk factors for AI projects, becoming a cornerstone of any development and deployment programme.

Governments and international bodies are clear about the need for overarching governance of AI, especially within public services; defence, health and infrastructure. However, individual organisations already benefitting from the technology cannot wait to define their own governance structures. With that in mind, what exactly is AI governance? According to Deloitte, it encompasses a: "wide spectrum of capabilities focused on driving the responsible use of AI. It combines traditional governance constructs (policy, accountability, *etc.*) with differential ones such as ethics review, bias testing and surveillance. The definition comes down to an operational view of AI and has three components: data, technique/algorithm and business context."

In summary: "achieving widespread use of AI requires effective governance of AI through active management of AI risks and implementation of enabling standards and routines".

Without formalising AI governance, organisations are less likely to know when models are becoming stale, results are biased or when data is improperly collected. Companies developing AI systems without stringent governance to tackle these issues are risking their businesses. They leave the way open for AI to effectively take control, with unpredictable results

that could cause irreparable damage to reputation and large legal judgments.

The least of these risks is that legislation will impose governance, and those who have not been practising AI governance of their own will need to catch up. In today's rapidly shifting regulatory landscape, playing catch up is a risk to reputation and business resilience.

THE UK GOVERNMENT IS AWARE OF THE POTENTIAL OF AI, BUT MORE WARY OF THE ETHICAL ISSUES

The reasons for AI governance failure are complex and interconnected. However, it is clear that accelerated AI adoption has not been matched by an increase in education and awareness of its risks. In short; AI is suffering a people problem.

Our research demonstrates significant skills gaps in key technological areas, including ML modelling and data science, data engineering, and the maintenance of business use cases. The AI skills gap is well documented, with much government discussion and policy to drive data skills through focused tertiary education and up/reskilling.

However, technological skills are insufficient to bridge the gap between innovation and governance. It is neither advisable nor fair to leave governance to technical talent alone. Undoubtedly those with the skills to develop AI must also be equipped with the knowledge and values to make decisions and problem solve within the broader context in which they operate. AI governance, after all, is a team effort.

It can also be said that how an organisation governs its use of AI represents its values brought to life. No organisation can claim to respect user privacy or put security first if AI governance is lacking.

As summarised by analysis from PwC: "Data, data use and AI ethics involve more than privacy. Some organisations are adopting principles around explainability, societal benefit and fairness, among other principles. Identify which principles are relevant, and more importantly, what these principles mean to your organisation. Get wide executive agreement on these principles and translate them into concrete standards and procedures for each practice within your organisation to enact trust-driven approaches."

Embedding ethics and security within AI means everyone across the organisation, from CEO to data analyst, CIO to project manager, must engage in AI governance. They must align on why it is that these issues matter and how the organisation's values play out through AI implementations.

Such a strategy starts with empowerment through education, awareness and role-specific training. When it comes to AI, vigilance is a holistic skill that all must master. Frameworks, principles and policies provide the basis for sound innovation, but mean nothing without engaged, educated and empowered humans to bring them to life.

Stepping forward into the age of AI requires focus not only on the potential of the technology itself, but of the development of the people harnessing it. That means education, training, collaboration and asking hard questions now so that robust governance can become a foundation of AI ●

Rachel Roumeliotis is Vice President of Data and AI at O'Reilly.

Two thirds of security decision makers are concerned about AI-led attacks