INSIDE GENAI

n° 05

Responsible GenAI Use: Addressing Legal and Ethical Challenges

CRIF

Together to the next level

Table of Contents

1 Introduction

This document explores the legal and ethical challenges of generative AI, focusing on global regulatory approaches, intellectual property, data privacy, and accountability. It highlights the importance of responsible AI use and outlines CRIF's commitment to human-centric, transparent, and fair AI practices.

When Alan Turing posed the question "Can machines think?" in 1950, he could not have envisioned that seventy-five years later machines would not only attempt to answer his question, but also generate complex responses with far-reaching social, legal, and ethical implications.

The release of ChatGPT in November 2022 marked a watershed moment in artificial intelligence, introducing the public to the power of generative AI (GenAI). Unlike traditional AI systems, GenAI models, such as OpenAI's GPT, Google's Gemini (formerly PaLM), and Stability AI's Stable Diffusion, are capable of generating human-like content, including text, images, and codes. Despite its vast potential, GenAI's rapid development and deployment have outpaced the evolution of adequate legal and ethical frameworks.

Within this context, the commitment to the responsible use of GenAI is imperative, given its integration into our professional and personal lives. By prioritizing human values, safeguarding privacy, and fostering accountability, we can harness the power of AI to build a better, fairer future for all. At CRIF, we are developing a **framework for the responsible use of AI** that addresses both legal and ethical challenges, and is rooted in integrity, transparency, fairness, and a human-centric innovation ecosystem.

2 International Legal Frameworks

Countries around the world are adopting different strategies to regulate artificial intelligence: The EU has taken a comprehensive legal approach, while others like the US, China, and the UK follow more decentralized, ideological, or flexible models. These divergent paths highlight the complexity of achieving global alignment in AI governance.

The European Union has emerged as a global leader in the regulation of artificial intelligence, particularly through the adoption of the **AI Act** in 2024. This comprehensive legal framework introduces a risk-based approach that classifies AI systems based on their potential impacts on safety, fundamental rights, and public trust. Under this framework, higher-risk systems are subject to stricter obligations, including requirements for transparency, high-quality datasets, risk management prior to AI deployment, and post-market monitoring procedures.

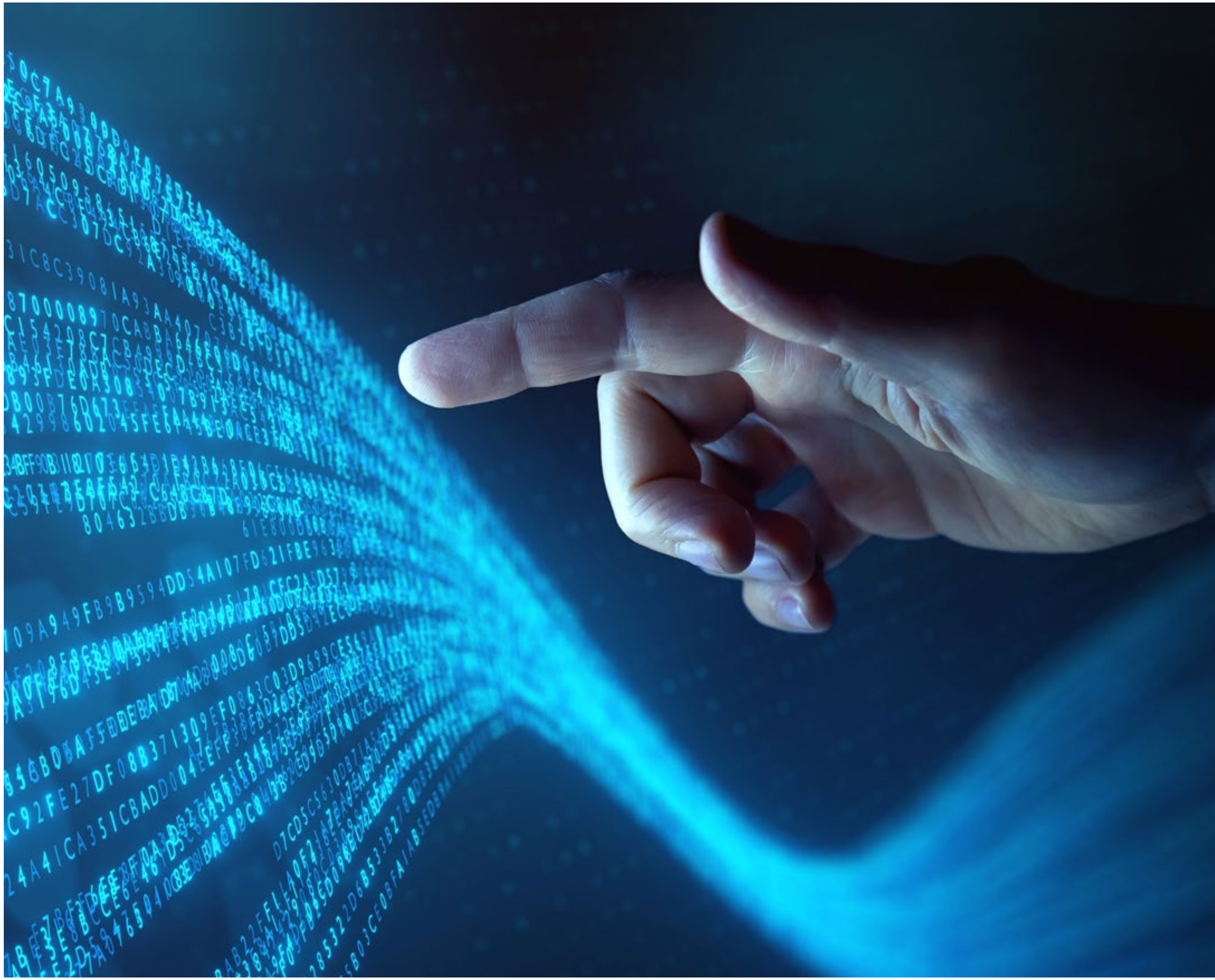
Notably, the AI Act introduces a dedicated regime for General-Purpose AI (GPAI) models. These models are defined in the AI Act by their versatility and the ability to perform a variety of unrelated tasks, rather than being limited to a single function such as facial recognition or language translation. Examples include models such as GPT-5, which can generate text, produce summaries, write code, reason, and process and generate images. Due to their computational power, some GPAI systems pose systemic risks, particularly when embedded into critical infrastructure or decision-making systems. In general, the AI Act requires providers of GPAI models with systemic risk to ensure documentation and traceability of training data. For GPAI models deemed to have systemic impact, there are even stronger obligations such as risk management plans, adversarial cybersecurity testing, and incident notification duties.

By contrast, the United States has adopted a more fragmented regulatory approach. Rather than enacting a unified federal AI law, it relies on the **regulatory mandates of specialized agencies** such as the Federal Trade Commission (FTC), the Food and Drug Administration (FDA), and the National Institute of Standards and Technology (NIST). Each agency provides guidance tailored to a specific domain, such as consumer protection, health technology, or technical standards, but this has led to inconsistencies and regulatory gaps. Several states have proposed their own AI laws, contributing to a patchwork of rules that can create compliance difficulties for developers operating nationwide.

China, another major player in the AI development race, enforces a centralized regulatory model. Its 2023 **Interim Measures for Generative AI Services** require content labeling, adherence to socialist values, and provider responsibility for any harm caused. This framework reflects China's broader governance approach, which emphasizes state control over innovation.

Finally, the United Kingdom has adopted a principles-based, innovation-oriented approach. Through its 2023 **White Paper on AI**, the UK promotes flexibility and collaboration with industry stakeholders while resisting calls for a centralized legal framework. This strategy aims to attract AI investment while encouraging responsible use through soft-law instruments such as voluntary codes of conduct and ethical guidelines.

These **divergent strategies** highlight the geopolitical complexities of AI regulation. Global convergence remains elusive, but international cooperation through the OECD, G7, or bilateral agreements may offer a path toward shared standards without undermining innovation.



3 Copyright and Intellectual Property

Generative AI is reshaping the intellectual property landscape by challenging long-standing legal frameworks. As AI systems increasingly produce creative outputs, questions around authorship, copyright, and data usage have become central to regulatory and ethical debates. Governments and institutions are now working to define new standards and responsibilities in response to these shifts.

Authorship and intellectual property rights are central issues in the GenAI debate. Traditionally, copyright law has been built on the concept of human creativity. GenAI disrupts this foundation by autonomously generating content—ranging from images and music to code and text—that can be indistinguishable from works created by humans.

The US Copyright Office has taken a firm stance on this issue: **Works lacking human authorship are not eligible for copyright protection**. This principle was reaffirmed in *Thaler v. Perlmutter*, where Dr. Thaler, a computer scientist, attributed authorship of an artwork to the operation of software. The court held that a work generated by an AI system without human input cannot be copyrighted. Consequently, companies using GenAI must carefully document the extent of human involvement if they intend to claim copyright over its outputs.

Moreover, the training phase of GenAI systems raises significant legal complexities. These models are typically trained on massive datasets, including copyrighted materials scraped from websites, repositories, and platforms, often without the creators' explicit consent. This has triggered several lawsuits, most notably *Getty Images v. Stability AI*, where the image licensing company alleges unauthorized use of millions of proprietary images to train diffusion models. The outcomes of these legal battles will likely shape the future of intellectual property rights in the GenAI era.

To address these tensions, the EU has introduced the **General-Purpose AI Code of Practice**. This voluntary code, developed through a multi-stakeholder process, offers practical guidance to help providers comply with the AI Act, particularly in the areas of transparency, copyright, and systemic risk management.

The copyright chapter outlines concrete measures for aligning with EU copyright law. It encourages providers to obtain licenses for copyrighted training materials where feasible, prioritize the use of openly licensed or public domain data, use synthetic datasets when appropriate, and maintain clear documentation of data sources and usage rights. These practices aim to reduce the risk of intellectual property violations and promote ethical data use.

Within the European Parliament, rapporteur Axel Voss has proposed legislative changes to clarify rights and responsibilities under EU copyright law. His draft report advocates creating a new legal status for AI-generated content and imposing duties on model providers to disclose training data sources and obtain appropriate usage rights. However, EU institutions have not yet decided if and how a dedicated legal framework will be adopted.

4 AI and Data Privacy

GenAI models may rely on vast datasets that include personal information collected without consent. This raises serious legal and ethical challenges. To mitigate these risks, experts emphasize anonymization, transparency, and privacy-by-design in AI development.

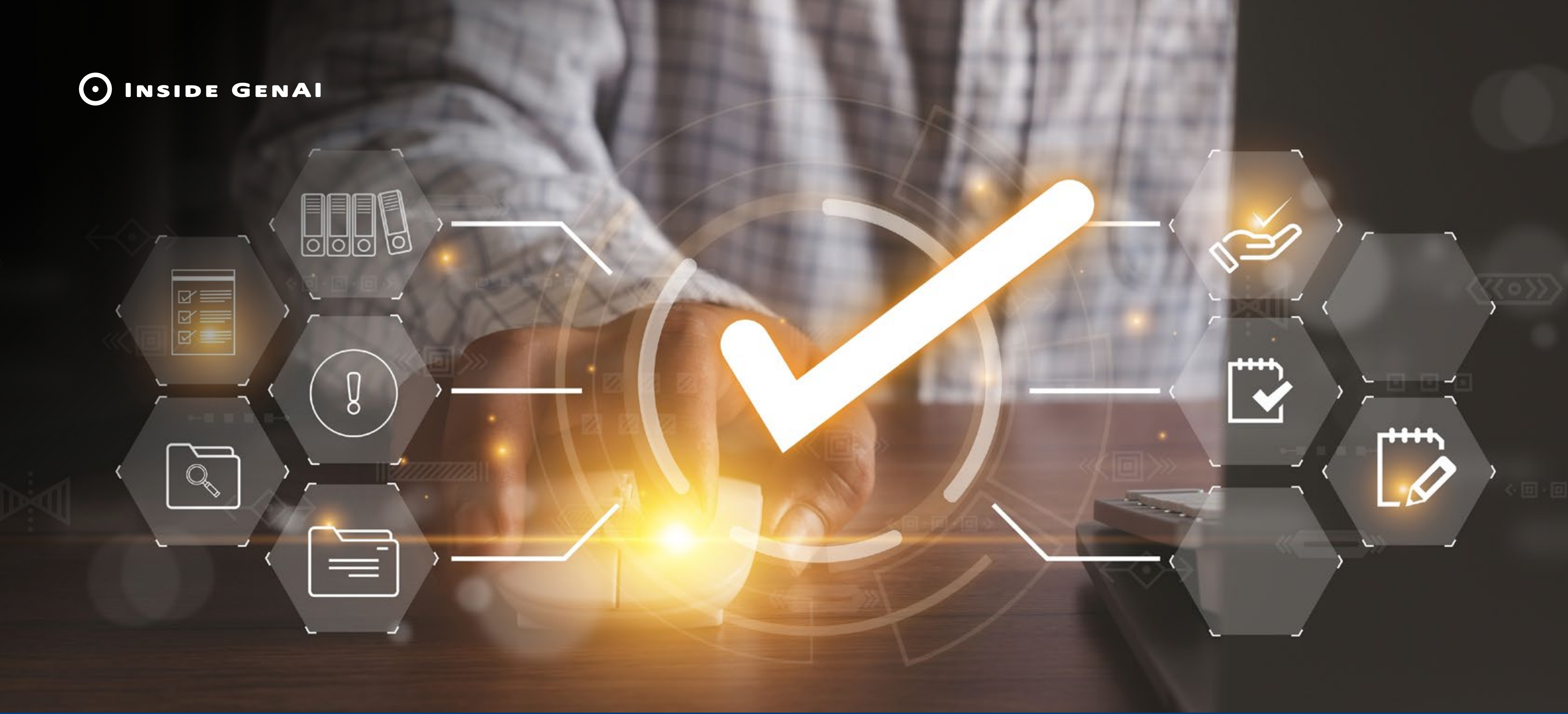
Data privacy remains one of the most sensitive and contested issues in GenAI regulation. Large-scale AI models are trained on datasets that may contain personal, sensitive, or confidential information scraped from blogs, social media, forums, news sites, and databases. This often occurs without the knowledge or consent of the individuals whose data is used.

The **General Data Protection Regulation (GDPR)** sets a high standard for data protection in the EU. It requires personal data to be processed lawfully, fairly, and transparently, in line with principles such as data minimization and purpose limitation, and it grants user rights such as access, rectification, erasure, and objection. However, most GenAI providers cannot determine the origin or content of all training data, making compliance technically and legally challenging.

A major concern is **model memorization**: GenAI models can inadvertently retain training data, meaning that personal information (such as names, phone numbers, or health records) may appear in the output when prompted. This has been observed in several studies and reported by users during interactions with large language models.

In the US, the California Consumer Privacy Act and other state laws offer similar rights but suffer from inconsistent enforcement and limited scope. The **absence of a federal AI bill or privacy law** exacerbates regulatory uncertainty.

Experts advocate for privacy-preserving technologies, such as data anonymization techniques, to mitigate risks. Internal governance mechanisms such as audit trails, dataset documentation, and human review are also essential to uphold privacy-by-design principles in GenAI systems.



5 Liability and Accountability

Determining liability for harm caused by GenAI systems is legally complex, especially when outputs are generated autonomously. Both the EU and US have proposed frameworks to address accountability, but gaps remain in assigning clear responsibility.

Determining **legal responsibility** for harm caused by GenAI systems is one of the most complex legal questions facing policymakers today. Traditional liability models that assign fault to a human actor do not readily apply to outputs generated autonomously by machines.

If a chatbot provides harmful medical advice, or if an AI system generates defamatory or discriminatory content, it is unclear whether responsibility lies with the developer, the deployer, or the user.

The EU **AI Act** attempts to address this with obligations for risk assessment, human oversight, incident logging, and post-market monitoring. While these measures increase accountability, they fall short of defining a clear liability regime, especially for general-purpose models used in unpredictable ways.

Alongside the AI Act, the European Commission proposed an **AI Liability Directive** aimed at harmonizing national rules on non-contractual civil liability for harm caused by AI systems. The Directive sought to modernize the EU's liability framework by introducing mechanisms that presume a link between harm and the use of high-risk AI systems when providers fail to meet their obligations, thereby easing the burden of proof for victims. Despite its relevance, the proposal faced political and industry resistance and was ultimately withdrawn from legislative consideration in early 2025. The absence of binding EU-wide rules on AI liability leaves a critical gap in the legal framework, particularly in cases involving autonomous and opaque systems, where traditional fault-based liability models struggle to assign responsibility. As a result, individuals harmed by AI-generated outputs may continue to face significant procedural and evidentiary hurdles when seeking redress.

In the US, the situation is further complicated by Section 230 of the Communications Decency Act, which grants online platforms and service providers immunity from liability for third-party content. Whether this protection applies to AI-generated content remains legally unclear, but it raises concerns that harmful outputs might go unpunished. Some legal experts have proposed a strict liability framework for high-risk applications, meaning providers would be liable regardless of intent or negligence. Others argue for mandatory AI insurance, transparency logs, and public registries for high-impact models.

To help bridge the accountability gap, **companies should adopt internal AI governance systems** including clear usage policies, real-time monitoring of outputs, and mechanisms for redress when harm occurs.

6 Ethical Challenges

The use of GenAI raises significant ethical concerns, including bias reproduction, the creation of false content, and the risk of user-pleasing responses. Responsible adoption requires strong safeguards, along with investment in training and inclusive policies.

GenAI raises **ethical questions** that go beyond legality, challenging core principles of trust, fairness, and human dignity in the digital age.

1 Bias reproduction

GenAI systems learn from large datasets that reflect existing societal inequalities and prejudices. If left uncorrected, these biases can be embedded in outputs, reinforcing stereotypes related to gender, race, class, or disability. This represents one of the most pervasive risks and is particularly concerning in areas such as employment, finance, criminal justice, and education.

2 Hallucinations

The generation of information that is factually incorrect but linguistically convincing is equally problematic. The case of *Steven A. Schwartz*, a New York lawyer who submitted a brief with fictitious citations from ChatGPT, is a textbook example. In Italy, another lawyer relied on ChatGPT to find precedents for a legal case, only to discover that the AI had fabricated them entirely. These “hallucinated” references were grammatically correct and legally plausible, illustrating the danger of misplaced trust.

3 Sycophancy

GenAI models are often trained to optimize for helpfulness, which sometimes means they echo the user’s beliefs or preferences rather than providing objective information. This behavior can mislead users, especially minors or vulnerable users, reinforcing misinformation or emotional bias.

4 Socioeconomic implications

According to the World Economic Forum, up to 40% of global jobs may be impacted by GenAI by 2030. While some roles will be augmented or transformed, others, particularly in the creative industries, customer service, and knowledge work, may disappear entirely.

5 Ethical deployment of GenAI

This requires robust safeguards and significant investment in education and reskilling. Policymakers, developers, and civil society must work together to ensure that GenAI is empowering rather than dehumanizing.

7 Conclusions

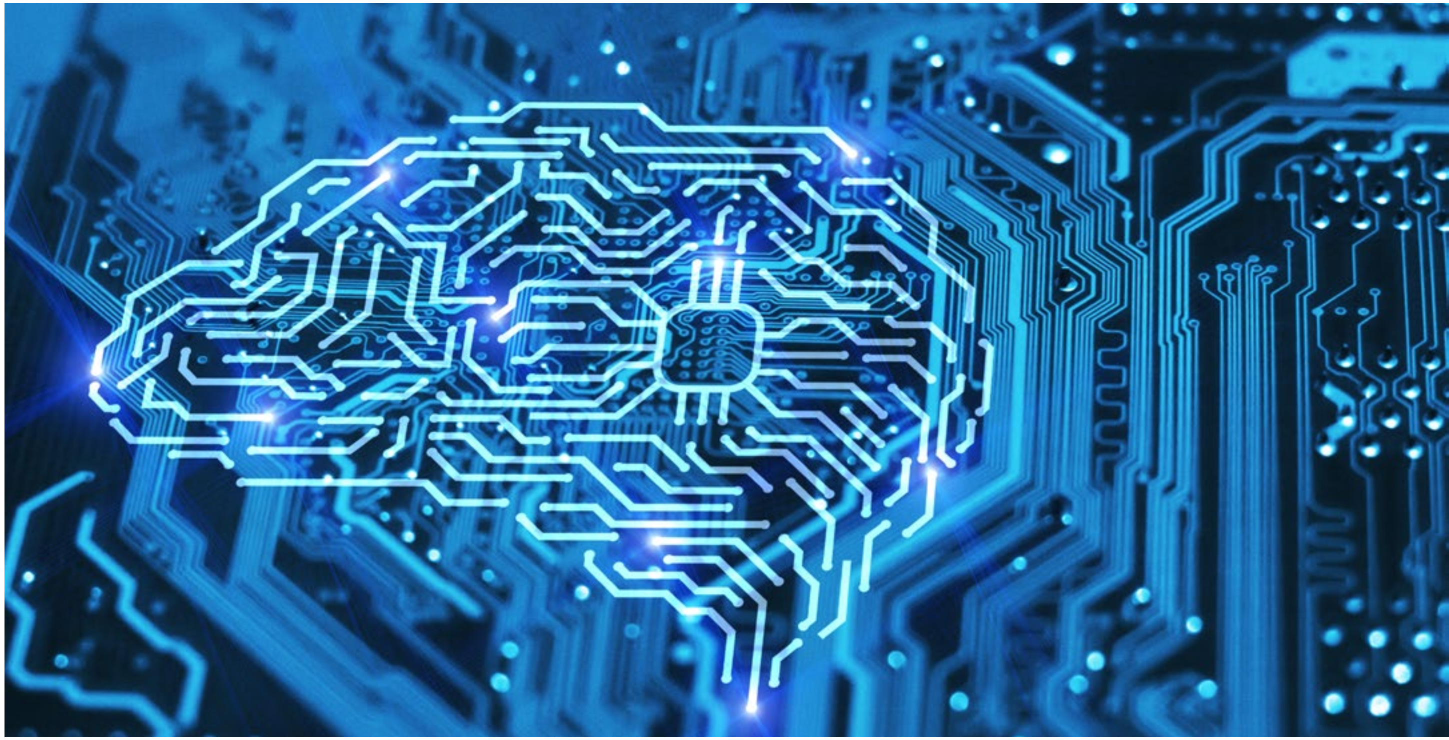
Generative Artificial Intelligence (GenAI) marks a groundbreaking technological shift, but it also raises unresolved legal and ethical challenges. Global regulation remains fragmented, while issues such as fairness, transparency, and rights protection are becoming increasingly critical. At CRIF, we are committed to developing ethical and compliant GenAI systems, with a strong focus on security, privacy, and bias mitigation.

GenAI is arguably one of the most transformative technological advances of our time. Yet, without robust legal frameworks and ethical safeguards, its rapid growth could deepen inequality, erode public trust, and cause unintentional harm.

The regulation of GenAI remains fragmented across jurisdictions. While the EU is taking the lead with the AI Act, other regions rely on piecemeal or voluntary approaches. Rules on intellectual property, data privacy, and liability are still evolving in response to GenAI's unique challenges.

Equally urgent are the ethical questions: How do we ensure that GenAI reflects fairness, accuracy, and accountability? How do we prevent the erosion of human agency? And how do we guarantee that it is used to serve, rather than replace, human creativity?

At CRIF, we are developing and deploying AI systems that are compliant and **ethical by design**, drawing on interdisciplinary expertise from specialists in law, compliance, and data science. Their work focuses on enhancing bias mitigation and explainability while upholding the highest standards of data privacy and security. We also implement **safeguards** to prevent the generation of harmful, offensive, or misleading content, with moderation systems that are continuously updated to reflect evolving risks. As AI and GenAI advance, so does our commitment to using them responsibly.



References

1. European Commission, 2024. Artificial Intelligence Act (AI Act). Official Journal of the European Union.
2. US Copyright Office, 2023. Policy on Registration of Works Containing AI-Generated Content.
3. Thaler v. Perlmutter, Case No. 1:22-cv-01564 (D.D.C. 2023).
4. Getty Images v. Stability AI, U.K. High Court Filing, 2023.
5. Axel Voss (European Parliament) 2025. Draft Report on Artificial Intelligence and Copyright.
6. China Cyberspace Administration, 2023. Interim Measures for the Management of Generative AI Services.
7. UK Department for Science, Innovation and Technology, 2023. AI Regulation White Paper.
8. Federal Trade Commission (FTC), 2023. Guidance on AI and Algorithmic Fairness.
9. European Data Protection Board (EDPB), 2022. Guidelines on AI and Data Protection.
10. California Consumer Privacy Act (CCPA), 2020. State of California Legislative Information.
11. World Economic Forum, 2023. The Future of Jobs Report.
12. Schwartz, S. A., 2023. Mata v. Avianca Airlines, S.D.N.Y. – legal filing involving fabricated ChatGPT citations.
13. EU General-Purpose AI Code of Practice, 2025.
14. NIST. (2023). AI Risk Management Framework (AI RMF 1.0). US Department of Commerce.

Together to the next level

crif.com