Cambridge
Forum

AN EDITORIAL

Comparative perspectives on the regulation of large language models

Cristina Poncibò [iD] and Martin Ebers

Law, University of Turin and Collegio Carlo Alberto, Torino, Italy
Corresponding author: Cristina Poncibò; Email: cristina.poncibo@unito.it

1. Introduction

Large language models (LLMs) represent one of the most significant technological advancements in recent decades, offering transformative capabilities in natural language processing and content generation.[1] Their development has far-reaching implications across technological, economic and societal domains, simultaneously creating opportunities for innovation and posing profound challenges for governance and regulation. As LLMs become integral to various sectors, from education to healthcare[2] to entertainment, regulators are scrambling to establish frameworks that ensure their safe and ethical use.

Our issue primarily examines the private ordering, regulatory responses and normative frameworks for LLMs from a comparative law perspective, with a particular focus on the European Union (EU), the United States (US) and China.[3] An introductory part preliminarily explores the technical principles that underpin LLMs[4] as well as their epistemological foundations. It also addresses key sector-specific legal challenges posed by LLMs, including their implications for criminal law, data protection and copyright law.

2. Understanding LLMs as new epistemic agents

Paseri and *Durante*'s paper argues that understanding LLMs as new epistemic agents-entities capable of generating knowledge and representations distinct from human reasoning-is crucial for identifying and addressing their risks and opportunities.

One of the key aspects explored is the way LLMs categorise and interpret data, which diverges fundamentally from human legal reasoning. LLMs process vast amounts of information through patterns and statistical correlations, often yielding outputs that lack the intentionality and contextual awareness inherent to human cognition. This difference raises pressing questions about the trustworthiness

---

[1]Wachter, S., Mittelstadt, B., & Russell, C. (2024). Do large language models have a legal duty to tell the truth? *Royal Society Open Science*. Retrieved from http://dx.doi.org/10.2139/ssrn.4771884.

[2]Wang, D., & Zhang, S. (2024). Large language models in medical and healthcare fields: Applications, advances, and challenges. *Artificial Intelligence Review*, 57, 299. doi:10.1007/s10462-024-10921-0.

[3]Some of the papers referred to in this editorial will be published at a later date.

[4]The introductory paper to this themed issue sets the stage for an exploration of LLMs through a technical lens, emphasising the interconnectedness of their technological capacities with the societal contexts in which they operate.

and reliability of artificial intelligence (AI)-generated outputs, particularly in contexts where nuanced interpretation is essential, such as legal decision-making.

Another dimension examined by the authors is the standard of trustworthiness applied towards LLMs compared to human legal judgements. Legal systems traditionally rely on processes like evidence evaluation, argumentation and adherence to procedural safeguards to establish trust in legal systems. In contrast, the opacity of LLMs' decision-making processes complicates the application of similar benchmarks. The authors emphasise the need for context-sensitive assessments that account for these differences, ensuring that the deployment of LLMs in legal contexts does not erode foundational principles of fairness and accountability.

Lastly, *Paseri* and *Durante* consider the impact of LLMs on the diversity of thought within legal and ethical systems. While these models have the potential to foster innovation by processing and synthesising information from diverse sources, over-reliance on their outputs may inadvertently suppress interpretative diversity. This risk stems from the tendency of algorithmic systems to optimise for patterns and consensus, which can marginalise unconventional or minority perspectives. The authors conclude that safeguarding the pluralism of legal and ethical discourse requires careful scrutiny of how LLMs are integrated into decision-making processes, with attention to their potential to both enhance and constrain the diversity of interpretations and approaches.

3. Regulating LLMs: Comparative perspectives

The regulatory responses to LLMs in the EU, the U.S., Singapore and China reflect the differing legal, cultural and political environments of each region.

Mimi Zou and *Lu Zhang* provide a detailed analysis of China's regulatory response to generative AI (GenAI), with a focus on the Interim Measures for the Management of Generative AI Services introduced in July 2023 ("Interim Measures"). These Measures mark a significant step in China's efforts to regulate the rapid development and deployment of GenAI technologies, including LLMs.[5] The authors emphasise the dual objectives underpinning these Measures: promoting technological innovation to maintain China's competitive edge in the global AI race, while simultaneously addressing the risks posed by these technologies – particularly, in terms of content safety, national security and social stability. Indeed, the Interim Measures reflect a cautious but ambitious approach, establishing rules that require GenAI providers to ensure their services adhere to China's strict content regulation standards. This includes mechanisms to prevent the generation of harmful or illegal content and mandates that service providers conduct security assessments before deploying their technologies. At the same time, the Measures seek to foster innovation by providing regulatory clarity and encouraging enterprises to contribute to the development of a domestic AI ecosystem. This balancing act, however, reveals inherent tensions between technological growth and the need to mitigate societal and political risks, a challenge not unique to China, but emblematic of the global struggle to regulate LLMs. The authors argue that China's regulatory framework, despite its focus on content control and security, could influence global AI governance in significant ways.

From a comparative perspective, *Jane Loo* and *Jason Allen Grant* also analyse Singapore's evolving AI governance framework, addressing the country's innovative approach to regulating AI technologies. They examine Singapore's AI strategies in terms of ethics, governance and regulatory frameworks, proposing a dynamic and adaptive model that could serve as a global reference. Their work exemplifies the need for flexible governance in the fast-paced field of AI.

The analysis of these authors immediately invites a comparative reflection on the regulatory approaches of other major players in AI governance, notably the U.S. and the EU.

---

[5] The *Interim Measures for the Regulation of Generative AI Services* (Chinese: 生成式人工智能服务管理暂行办法) represent a set of guidelines introduced by China to oversee the use of generative artificial intelligence services accessible to the public. They took effect on 15 August 2023.

The U.S. has so far taken a *laissez-faire* approach to AI regulation, with an emphasis on fostering innovation and allowing the private sector significant latitude in the development and deployment of AI technologies. While federal agencies, such as the Federal Trade Commission, have issued guidance on AI-related issues like bias and transparency,[6] the U.S. lacks a comprehensive federal framework akin to China's Measures or the EU's AI Act. This fragmented regulatory environment reflects the U.S.'s preference for innovation-driven policies, but leaves gaps in addressing systemic risks and ensuring accountability.

In contrast, the EU has adopted a more prescriptive and precautionary approach, exemplified by the EU's AI Act.[7] In this respect, both *Ugo Pagallo* and *Sebastian Hallensleben* criticise the EU's AI Act, albeit from different perspectives. According to *Pagallo*, the normative attempt of EU legislators to govern misuses and overuses of LLMs with the AI Act is ill suited, leading to issues of competitiveness, legal uncertainty, bureaucratic burdens and the threat of over-frequent revisions of the law to tackle advancements of technology. *Hallensleben*, on the other hand, focuses on the EU's AI Act's reliance on harmonised standards which provide a presumption of conformity with regulation – highlighting that standardisation of GenAI is still in its infancy due to the lack of technical maturity – both at European and international levels. Additionally, *Teresa Rodríguez de las Heras Ballell* specifically explores the AI Act and the proposed EU liability rules for AI. She provides an in-depth analysis of the evolving rules on AI governance, with a particular focus on liability keeping pace with technological advances.

The EU's framework seeks to establish a risk-based classification system for AI applications, imposing stringent requirements on high-risk systems, including those used in sensitive areas like healthcare, law enforcement and content moderation. Unlike China's approach, which is heavily influenced by state interests in content control and security, the EU emphasises fundamental rights, transparency and accountability as central pillars of its regulatory strategy.[8] The EU's AI Act, however, is not without its critics – who argue that its detailed requirements coupled with its limited reliance on a truly risk-based approach, could stifle innovation and impose significant compliance costs on businesses.[9]

Apart from these individual jurisdictions, there is also the question of how to coordinate the activities of different enforcement bodies – both at national and international levels. As AI impacts the core mandate of market authorities dealing with financial stability, data protection, competition and telecommunications, *Oscar Borgogno* calls in his paper for the creation of a coordinated monitoring scheme to unify supervisors and market authorities in addressing the cross-industry challenges posed by LLMs deployment. Using examples such as the UK's Digital Regulation Cooperation Forum[10] and the EU's AI Act implementation process, the paper illustrates how national and international coordination can foster best practices and regulatory coherence.

The preliminary comparative analysis of these regulatory regimes highlights the importance of fostering an international dialogue on LLMs regulation, as the differing approaches underline the need for a more harmonised global framework that can address cross-border challenges posed by GenAI technologies. Against this background, *Hiroki Habuka* and *David U. Socol de la Osa* discuss the enhancements and next steps of the G7 Hiroshima AI process with a focus on advancing human

---

[6]Federal Trade Commission. (2020, April 8). Using artificial intelligence and algorithms. Retrieved from https://www.ftc.gov/business-guidance/blog/2020/04/using-artificial-intelligence-algorithms.

[7]Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act).

[8]The EU has established two other regulatory frameworks that may also be relevant to LLMs, such as the General Data Protection Regulation (GDPR) and, more recently, the Digital Services Act (DSA).

[9]Cf. Ebers, M. (2024). Truly risk-based regulation of artificial intelligence: How to implement the EU's AI Act. *European Journal of Risk Regulation*. https://doi.org/10.1017/err.2024.78.

[10]The Digital Regulation Cooperation Forum (DRCF). https://www.gov.uk/government/collections/the-digital-regulation-cooperation-forum.

rights, democracy and the rule of law. In particular, their article examines how the G7's initiatives on AI can contribute to a common framework that promotes global standards and values in the regulation of GenAI technologies, emphasising the importance of international cooperation in shaping the future of AI governance.

4. Private ordering and public governance of LLMs

The regulation of LLMs also involves navigating a complex divide between private ordering – where platforms and industry actors set their own rules – and public governance – where state authorities intervene to protect the public interest. From the role of platform terms and conditions in shaping the use of GenAI to the regulatory frameworks emerging in public administration and criminal law, three authors shed light on how both private and public entities are grappling with the challenges of governing LLMs in an increasingly AI-driven world.

From this perspective, *Lilian Edwards, Igor Szpotakowski, Gabriele Cifrodelli, Joséphine Sangaré* and *James Stewart* examine how platforms are shaping the use of GenAI through self-regulation, critically assessing the balance between private control and the need for state intervention. *Sophie Weerts* explores the evolving regulatory landscapes in the U.S. and EU, discussing how governments are adapting to the integration of AI in public administration while ensuring transparency and accountability. *Beatrice Panattoni* brings a legal perspective, investigating the intersection of GenAI and criminal law, highlighting the challenges of responsibility, liability and the role of AI in criminal justice. Together, these authors provide valuable insights into the ongoing struggle to balance private and public governance in the regulation of LLMs.

5. GenAI and data protection

Both *Elana Zeide* and *Hannah Ruschemeier* offer insights into the regulation of LLMs with a focus on privacy, each from a distinct regulatory perspective, thereby highlighting the comparative challenges of aligning AI technologies with privacy frameworks. While *Zeide* examines the evolving privacy norms in the U.S. and advocates for a revised framework to address the specific risks posed by GenAI, *Ruschemeier* explores the tensions between LLMs and the EU General Data Protection Regulation (GDPR),[11] emphasising the challenges of reconciling these emerging technologies with European privacy laws. Particularly, *Ruschemeier*'s article points out a series of conflicts that challenge the coexistence of these two frameworks.

In the EU, a central concern lies in the difficulty of ensuring data subject rights within the expansive data ecosystems underpinning LLM development. Fundamental rights such as the right to be forgotten and data access rights present significant hurdles when applied to GenAI. For instance, the extensive use of large-scale data scraping to train LLMs often involves personal data, making it difficult to guarantee the erasure of such data upon request or to provide individuals with transparent access to their personal information within these vast and opaque datasets. She further examines how LLMs challenge the core principles of data protection enshrined in the GDPR. Principles such as data minimisation and purpose limitation – designed to ensure that only necessary data is collected and used for clearly defined objectives – clash with the broad (and often indiscriminate) data requirements of LLMs. These models thrive on the vastness and variety of their training data, often exceeding what would traditionally be considered necessary for their stated purposes. Transparency, another cornerstone of GDPR, is similarly complicated by the "black-box" nature of LLMs, which obscures how data are processed and how outputs are generated. These challenges highlight a misalignment between the fundamental ethos of data protection law and the technical realities of LLM development.

---

[11]Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data. *Official Journal of the European Union, L 119*, 1–88. Retrieved from http://data.europa.eu/eli/reg/2016/679/oj.

Adding to these complexities is the issue of categorising data. Distinguishing between personal and non-personal data is becoming increasingly fraught as LLMs process vast datasets that frequently include a blend of both. The blurred boundaries between these categories' complicate compliance with GDPR, particularly when GenAI unintentionally outputs information that could be linked back to an identifiable individual. Ruschemeier points out that such scenarios expose gaps in the current regulatory framework, as traditional distinctions between data types struggle to accommodate the realities of GenAI processing. Her paper underscores the urgency of addressing these gaps.

6. Copyright on GenAI outputs

Jerome De Cooman's contribution tackles the economic and legal challenges posed by the indistinguishability of GenAI outputs from human-made works. This indistinguishability creates market inefficiencies, particularly an "information asymmetry" that risks devaluing high-quality human creations. His analysis addresses the increasingly blurred lines between human creativity and machine-generated outputs, proposing an innovative framework to navigate these complexities. Drawing inspiration from the EU rules of origin and the "substantial transformation" test, *De Cooman* argues for a systematic approach to delineating human contributions within the creative process. This framework seeks to establish clear criteria for identifying when a human creator has exercised free and creative choices, thereby meeting the originality threshold required for copyright protection. Central to this framework is the emphasis on human contribution and originality. In creative industries, originality serves as a cornerstone of intellectual property rights, safeguarding the unique input of human creators. *De Cooman* highlights the importance of clearly distinguishing human-authored works from machine-generated outputs, not merely to uphold legal standards but also to preserve the intrinsic value associated with human creativity. This differentiation becomes particularly critical as GenAI systems, including LLMs, continue to produce content that is often indistinguishable from human-created works, challenging traditional notions of authorship and originality. The proposed framework also addresses the economic dimensions of this issue by providing for safeguards against market valuation. Research indicates a clear consumer preference for human-authored works, which typically command higher market valuations than their machine-generated counterparts. However, this preference is undermined when consumers cannot readily identify the origin of creative outputs, leading to a "lemons problem" where the perceived quality of all works – human- or machine-made – declines. By introducing mechanisms to reliably attribute human authorship, the framework aims to protect the economic viability of original creations, ensuring that human creativity retains its market distinctiveness in an era increasingly dominated by GenAI. Through this analysis, *De Cooman* underscores the importance of establishing clear standards for differentiating human and machine contributions.

7. Conclusion

This themed issue of the Cambridge Forum on AI: Law and Governance embarks on an interdisciplinary exploration of the profound legal implications of LLMs. From understanding LLMs as unique epistemic agents to addressing specific legal issues, such as data protection and copyright law, the contributions collectively highlight the complexities of integrating LLMs into society while maintaining core principles of fairness, accountability and human creativity. We sincerely hope that this inaugural issue serves as a catalyst for deeper reflection and dialogue on the transformative implications of GenAI within the legal and regulatory landscape. By offering a comparative perspective on its challenges and opportunities, we aim to set the stage for informed and constructive discussions that not only clarify the current state of affairs but also inspire innovative solutions.