International Workshop
Shaping the Future of Artificial Intelligence and Regulation

18-19 September 2025
Montréal, Québec

Summary Report

With the support of

International Workshop
Shaping the Future of AI and Regulation

Second edition of the Quebec-Oxford-France workshop series “Shaping the Future of AI”.

Date: 18-19 September 2025
Venue: Court of Appeal of Quebec (100 rue Notre Dame Est, Montréal) - Room RC 22
Partners: IVADO, Université de Montréal, Maison française d’Oxford, University of Oxford, British Consulate-General Montréal, Délégation générale du Québec à Londres, CIFAR
Number of invitees: 15-20 in total (from Quebec, UK and France), with invited PhD students and post-docs
Convenors:
Prof. Angeliki Kerasidou (University of Oxford, Ethox Centre)
Prof. Catherine Régis (Université de Montréal, IVADO, Mila)
Prof. Célia Zolynski (Université Paris 1 Panthéon-Sorbonne, Observatoire de l'IA de Paris 1)
Coordinators and editorial team:
Gëlle Foucault (Université de Montréal)
Antoine Congost (IVADO)

Context

This event is the second edition of the Quebec-Oxford-France workshop in the series “Shaping the Future of AI”, which builds on comparative and interdisciplinary perspectives to explore key societal opportunities and challenges related to the development and deployment of AI.

Over the past few years, we have witnessed the development of various AI regulations worldwide, alongside global governance initiatives aimed at ensuring that individuals, businesses, and governments can harness the benefits of this transformative technology while mitigating the risks it poses, particularly to human rights, the environment, and democracies. The European Union’s AI Act stands out as a prominent example of such regulatory efforts, with the potential to shape businesses globally that operate within the EU. In contrast, the UK has opted for a “pro-innovation”, sector-specific regulatory approach, while Canada has focused on an agile, high-impact AI systems specific model (though these efforts have stalled with the prorogation of Parliament in 2025).

Comparing these three approaches provides valuable insights into the legal, social, political, and economic factors that shape them, offering guidance for defining future steps in the regulatory landscape, including at the implementation level. Furthermore, these national initiatives unfold within the broader context of global efforts to establish common redlines, bridge the digital divide, and enhance regulatory interoperability between countries. The coordination of national and international normative efforts is a complex, ongoing challenge that requires thoughtful academic and multistakeholder perspectives to guide governments, international organizations, and global alliances like the G7 and G20.

It is therefore timely to organize this invitation-only workshop bringing together leading AI researchers from Quebec, Oxford and France to share the latest developments and challenges on AI regulation at a national and international levels.

The workshop will cover three main topics:

- **Comparative AI regulatory approaches**
- **Challenges in implementing AI regulations**
- **Current initiatives and geopolitical considerations for global AI governance**

Programme

Thursday, 18th September

8:45 - 9:10	Breakfast
9:10 - 9:30	**Opening remarks**: Prof. Catherine Régis; Frédéric Tremblay, Director General of the Deputy Ministry for American Relations, Economic Affairs and Strategic Intelligence (Québec Ministry of International Relations and La Francophonie); Mario Riverot-Huguet, Head of Science and Technology (British Consulate-General Montréal)
9:30 - 11:30	***First session:*** **Comparative AI regulatory approaches in each country** Session chaired by Prof. Catherine Régis (Université de Montréal, IVADO, Mila)

Prof. Célia Zolynski (Université Paris 1 Panthéon-Sorbonne, Observatoire de l'IA de Paris 1)

***Title*:** Deepfakes captured by AI regulation: more than a legal issue, an interdisciplinary challenge

***Abstract*:** Deepfakes illustrate both the need to adopt appropriate AI regulations and the difficulty of designing it in such a way as to achieve the desired objective, take into account the various issues raised and reconcile legal standards with technical solutions. The study of deepfakes through the framework established by the AI Act also calls for an interdisciplinary analysis (incorporating Philosophy, Arts, and History, among other fields) to ensure that these various requirements are properly met.

Prof. Rebecca Williams (University of Oxford) - Online

Title: Comparative AI regulatory approaches in each country

Abstract: The UK's present approach to regulating AI might be characterised as watching and waiting. Outside the EU and thus not bound by the EU AI Act the UK is not rushing to legislate, and the Data Use and Access Act 2025 if anything reduces the protections of the (UK)GDPR. One potential result of this position is that it may be left to existing legal rules to adapt to cover the challenges raised by AI. But this is not necessarily problematic. Focusing on the example of algorithmic decision-making, the law already has a set of rules tailor-made to control abuse of disparate power and to ensure transparent and fair decision-making in the form of the rules of judicial review. If we can adapt our existing rules, the need for new ones may become less pressing.

Prof. Pierre Larouche (Université de Montréal)

Title: Competition, Regulation and Innovation in AI geopolitics - Canadian perspective

Abstract: I begin with a theoretical framework to make sense of current debates in AI governance. In a nutshell, this involves, in substance, a set of bilateral tradeoffs between risk regulation, competition (including industrial policy) and innovation, together with some fundamental institutional design decisions. The current regulatory approaches of leading jurisdictions are mapped onto this framework, in order to produce a structured comparative account and to give substance to the current geopolitical challenges. This reveals, among others, that the abrupt shift in US policy with the new administration also brought the EU in a pivot position to set the future evolution of global AI governance with its next moves.This provides a backdrop for the Canadian perspective on AI governance, as it has evolved so far and as it will be reframed by the new government.

11:30 - 11:45	Break
11:45 - 12:30	***Lightning talks:*** Latin American and South Korean regulatory approaches

Prof. Melissa Hyeshun Yoon (Hanyang University)

Title: South Korea's AI Governance Framework: Balancing Innovation and Regulation through the AI Basic Law

Abstract: This presentation examines South Korea's comprehensive approach to AI governance through the recently enacted "Framework Act on the Promotion of Artificial Intelligence Development and the Establishment of a Trusted Foundation" (commonly referred to as the "AI Basic Law") and its ongoing implementation decree preparations. I will analyze how South Korea is attempting to balance technological innovation with ethical considerations and regulatory oversight, drawing comparisons with other Asian approaches, particularly Japan's AI Promotion Act where relevant. The presentation will highlight key provisions of the Korean framework, including risk-based regulatory approaches, AI ethics guidelines, and mechanisms for public-private collaboration. I will also discuss the challenges and opportunities in implementing this framework within South Korea's unique technological and social context, offering insights for international regulatory harmonization efforts.

Prof. Juan David Gutiérrez Rodriguez (Universidad de los Andes) - Online

Title: Emerging AI Regulatory Approaches in the Global South

Abstract: We are witnessing a global trend of growing interest in introducing regulations that address artificial intelligence (AI). For example, after creating a novel database that maps AI bills and regulations, we documented over 600 regulatory instruments submitted, discussed, and/or approved in twenty-five Latin American and Caribbean countries and territories. This paper examines the rules and regulatory projects that directly and indirectly address AI development, acquisition, adoption, deployment, and use in the Global South. The text characterizes diverse regulatory tools (e.g., audits, transparency instruments, etc.) and nine AI regulatory approaches: principles-based, standards-based, agile approaches, facilitator approaches, adaptive approaches, mandatory disclosure approaches, rights-based, risks-based, and liability approaches. Finally, the paper discusses the policy and political challenges associated with implementing AI regulation in the Global South.

12:30 - 13:45 Lunch

14:00 - 16:00 ***Second Session:*** **Exploring the implementation challenges in AI regulation**

Session chaired by Prof. Angeliki Kerasidou (University of Oxford, Ethox Centre)

Prof. Benjamin Guedj (INRIA, University College London) - Online

Title: When Law Meets Code: Technical Hurdles in Implementing AI Regulation

Abstract: Efforts to regulate AI often run into a fundamental difficulty: the gap between high-level legal principles and the technical realities of AI systems. As a machine learning (ML) researcher, I will highlight why core implementation challenges — such as defining transparency, auditing complex models, ensuring robustness under distributional shifts, and certifying compliance at scale — resist simple solutions. These challenges are not only technical but also shape what kinds of regulation are feasible in practice. My aim is to shed light on where regulation collides with current ML capabilities, and to outline opportunities for collaboration between regulators, technologists, and researchers to make regulation both effective and realistic.

Prof. Christian Gagné (Université Laval, IVADO, Mila)

Title: The Case for National AIs

Abstract: The considerable advances of artificial intelligence in the last few years, in particular with Large Language Models (LLMs) and other Foundational Models (FMs), have announced a period of important technological advances that are already impacting significantly the economy and society. However, these technological advances were controlled mostly, until recently, by Big Tech American companies. Given the significant turmoil we have seen since the recent US presidential election, there is a significant erosion of thrust that has led to question our current dependencies from US technological companies regarding artificial intelligence. The capacity to develop a stronger digital sovereignty leads to the idea of having national AIs, with LLMs and FMs that are built by and for citizens of a given nation, better reflecting their culture, values, and languages while being developed and deployed on local technological infrastructures. In this presentation, I will develop the case for such national AIs, the surrounding technological and societal context, and the conditions required for achieving them.

Alexei Grinbaum (CEA-Saclay)

***Title*:** From AI ethics to AI regulation and back: operationalizing the AI Act recital 27 (abstract coming soon)

Abstract: I will describe the context of AIOLIA project training in AI ethics, starting from the sources of ethical tension in AI system design and all the way down to the tensions concerning the research exception in the EU AI Act. I will then briefly introduce the AIOLIA training module.

16:00 - 16:20 Break

16:20 - 17:35 ***Lightning talks:*** Insights from a judge and two lawyers

Session chaired by Honorable Judge Benoît Moore (Quebec Court of Appeal)

Honorable Judge Simon Ruel (Quebec Court of Appeal)

Title: From Promise to Peril: The Uses and Regulation of AI by the Judiciary

Abstract: The judiciary faces a dual challenge with respect to the use of AI. On the one hand, AI systems can strengthen justice by making it faster, more accessible, and more consistent. However, it can also threaten justice by introducing bias, eroding confidentiality, or undermining judicial independence and impartiality. The central question is not whether AI will enter courtrooms. It already has, at least to some extent, in Quebec and Canada. The key issue is how AI will be integrated, regulated, and controlled so that it enhances rather than compromises the legitimacy of judicial decision-making.

Paul Gagnon and Misha Benjamin (BCF)

Title: News from the front – Navigating AI regulation in practice

Abstract: This session aims to highlight key learnings and emerging trends from two leading attorneys in the field of AI. With an international practice representing both AI providers and adopters, Misha and Paul will discuss how regulation is shaping contract negotiations and AI product design. The session also aims to explore the goals and impacts of emerging AI regulation such as: (i) regulation as a competitive moat for Big Tech; (ii) regulation as a driver of innovation; and (iii) the impact of local regulation on companies with global reach and ambitions. Bringing practical and hands-on experience, the two speakers aim to highlight limits and opportunities found in emerging AI regulation.

17:45 - 19:00 Cocktail sponsored by BCF (Room: Salon des avocats)

19:00 Dinner at restaurant Maggie Oakes (426 Place Jacques-Cartier)

Friday, 19th September

9:00 - 9:30 Breakfast

9:30 - 11:30 ***Third session: Global AI governance and geopolitics***

Session chaired by Prof. Célia Zolynski (Université Paris 1 Panthéon-Sorbonne, Observatoire de l'IA de Paris 1)

Benjamin Prud'Homme (Mila)

Title: Global AI Safety and international alliances in a new geopolitical context

Abstract: In this presentation, I will start by reviewing the mandate, structure and content of the International Scientific Report on the Safety of Advanced AI, chaired by Yoshua Bengio. I will then reflect on the politicization of the term "AI Safety", and what this means in the current context of AI development. Finally, I will broaden the conversation to discuss the current shifts in the geopolitics of AI, with an emphasis on the role middle-powers and multilateral organizations could play as we face profound changes in the world order.

Isabella Wilkinson (Chatham House)

Title: Transparency and credible, coherent AI governance

Abstract: As countries, companies and other stakeholders seek to govern AI, transparency has emerged as a central principle and practice. Meaningful transparency is certainly a prerequisite for effective governance. There is growing consensus about its meaning: for example, on aspects of model ('technical') transparency and what constitutes 'public' transparency. However, understandings vary across supranational, multilateral and national governance initiatives. This talk uses AI transparency as a lens for exploring how to overcome emerging issues – fragmentation and incoherence – in global AI governance. It considers the architectures, mechanisms and partnerships required to work towards credibility and coherence, and their durability, both as models advance and amid geopolitical rivalries.

Prof. Catherine Régis (Université de Montréal, IVADO, Mila)

Title: The Creation of the UN Scientific Panel on AI: What does it mean for the Future of AI Governance?

Abstract: In September 2024, the United Nations General Assembly, through its Global Digital Compact, committed to establishing an independent International Scientific Panel on AI within the UN. In the interest of facilitating the United Nations' (UN) formulation of this panel, various actors and organizations have submitted proposals*. Following a period of deliberation, the General Assembly adopted a resolution in August 2025, formally initiating the establishment of the panel. While the precise structure, functioning, financing, and composition of the panel are yet to be delineated, the Resolution specifies that it will be a multidisciplinary, independent, and geographically diverse panel comprising 40 members. It is also understood that this initiative will result in the production of scientific synthesis and analysis of existing research on opportunities, risks, and impacts related to AI. This will be achieved, in part, through the dissemination of one annual "policy-relevant" yet "non-prescriptive" summary report. In this presentation, an exploration will be conducted of the milestones of the Panel, the key normative tensions at stake in achieving the intended results, and the lessons that can be learned from previous experience in global governance.

*See for example: Mila, The Development of the UN Scientific Panel on AI, Policy Paper, Mars 2025.

11:45 - 12:30	Visit of the Court of Appeal of Quebec
12:45- 13:45	Lunch
14:00 - 15:00	**Final words by the convenors & Quebec delicacies**

Prof. Angeliki Kerasidou (University of Oxford, Ethox Centre)

Prof. Catherine Régis (Université de Montréal, IVADO, Mila)

Prof. Célia Zolynski (Université Paris 1 Panthéon-Sorbonne, Observatoire de l'IA de Paris 1)

Workshop Overview – *Shaping the Future of AI and Regulation*

The second edition of the Quebec–Oxford–France workshop series *Shaping the Future of AI* was held on 18–19 September 2025 at the Court of Appeal of Quebec in Montréal. Organized in partnership with IVADO, Université de Montréal, the Maison française d'Oxford, the University of Oxford, the British Consulate-General in Montréal, the Délégation générale du Québec à Londres, and CIFAR, it brought together 18 invited participants from Quebec, the United Kingdom, and France, including PhD students and postdoctoral researchers.

This edition built on comparative and interdisciplinary perspectives to examine the key societal opportunities and challenges raised by the development and deployment of artificial intelligence (AI). Over the past few years, multiple regulatory efforts have emerged worldwide alongside global governance initiatives seeking to ensure that individuals, businesses, and governments can benefit from this transformative technology while limiting the risks it poses to human rights, the environment, and democratic systems. The European Union's AI Act stood out as a landmark initiative, with the potential to shape practices well beyond the EU. By contrast, the United Kingdom pursued a sector-specific "pro-innovation" strategy, while Canada is revisiting an earlier proposal advancing an agile, high-impact model focused on specific AI systems, to give more emphasis to the adoption of trustworthy AI across Canada and the fostering of the Canadian AI ecosystem.

Comparing these approaches provided valuable insights into the legal, social, political, and economic dynamics that underpin regulatory choices and offered guidance for identifying the next steps in the global regulatory landscape, including practical implementation. At the same time, national efforts unfolded within a broader context of international initiatives aimed at defining common red lines, bridging the digital divide, and fostering regulatory interoperability. Coordinating these national and international normative frameworks remains a complex challenge that requires sustained academic and multistakeholder input to guide governments, international organizations, and global alliances such as the G7 and G20.

It was in this context that the workshop convened leading AI scholars and practitioners from Quebec, Oxford, and France to discuss the latest developments and challenges in AI regulation at both national and international levels. Through these exchanges, the event reinforced the importance of comparative dialogue and collaborative reflection to shape the future of AI governance.

Prof. Célia Zolynski — Université Paris 1 Panthéon-Sorbonne, Observatoire de l'IA de Paris 1

Title: *Deepfakes captured by AI regulation: More than a legal issue, an interdisciplinary challenge*

Abstract: Deepfakes illustrate both the need to adopt appropriate AI regulations and the difficulty of designing them in a way that achieves the desired objectives, addresses the various issues raised, and reconciles legal standards with technical solutions. Examining deepfakes through the framework established by the AI Act also calls for an interdisciplinary analysis to ensure that these requirements are properly met.

Summary: Professor Célia Zolynski examined how deepfakes can be addressed through the European Union's regulatory framework, with a particular focus on the AI Act. She emphasized that deepfakes represent a systemic risk capable of undermining elections, eroding public trust, and infringing on individual rights. Their growing realism and accessibility make it crucial to distinguish AI-generated content from authentic human communication. While the AI Act provides legal tools to regulate these practices, it also raises complex questions about effectiveness and enforcement. Drawing on the Act's definition of deepfakes in Article 3, its list of prohibited practices in Article 5, and the transparency obligations outlined in Article 50, she explored the legal architecture designed to confront these risks. She also referred to European Parliament studies, including the 2020 report on deepfakes and the "Children and Deepfakes" study highlighting that 98 percent of non-consensual content targets women and girls, as well as to ongoing European Commission consultations on transparency in AI systems.

Zolynski illustrated that not all deepfakes are prohibited or deemed high-risk under the Act. Instead, the regulation imposes labelling and watermarking obligations, requiring providers and deployers to disclose manipulated content, with exceptions for artistic, satirical, or editorial contexts. She noted the risks of manipulation in democratic processes, such as disinformation and foreign interference, and underscored the alarming rise of sexualized deepfakes, often produced through applications like "Nudify," which target women, minors, and public figures, thereby reinforcing issues of cyber harassment, sextortion, and child sexual abuse material (CSAM). She also stressed the role of complementary frameworks, notably the Digital Services Act, which obliges large platforms to assess systemic risks and implement proportionate mitigation measures, particularly in relation to elections and the protection of minors.

She acknowledged that significant challenges remain. Transparency measures raise technical concerns with watermarking, cognitive limitations in labelling, and persistent risks of false narratives. The applicability of transversal provisions such as Article 5 to harms like CSAM and non-consensual intimate imagery remains debated. National responses in France and the United States offer additional legal avenues, but their scope and consistency are still uncertain. These difficulties illustrate the ongoing tension between protecting fundamental rights and safeguarding artistic and scientific freedoms.

In conclusion, Zolynski argued that addressing deepfakes requires more than legal texts: it calls for an interdisciplinary approach combining law, technology, ethics, and digital literacy. The AI Act, reinforced by the Digital Services Act, constitutes an important step forward, yet its real-world impact will depend on enforcement, international coordination, and the adaptability of regulatory tools to new threats. Protecting democratic processes, shielding vulnerable populations, and sustaining public trust demand continuous vigilance and innovation in governance.

Key Takeaways

- Deepfakes pose systemic risks, from electoral interference to non-consensual sexual content.
- The AI Act defines deepfakes and imposes transparency obligations but does not ban them outright.
- Most deepfake pornography targets women and minors, amplifying gendered harms.
- The Digital Services Act complements the AI Act by requiring platforms to assess and mitigate systemic risks.
- Implementation challenges persist, especially in balancing regulation, freedom of expression, and technical feasibility.

References

European Commission (2025). *Consultation on guidelines and code of practice for transparent AI systems.* https://digital-strategy.ec.europa.eu/en/news/commission-launches-consultation-develop-guidelines-and-code-practice-transparent-ai-systems

European Parliament (2021). *Tackling deepfakes in European policy.* https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU(2021)690039

European Parliament (2023). *Children and Deepfakes.* https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2025)775855

European Union. *Digital Services Act (DSA), Articles 34–35.* https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act_en

Prof. Rebecca Williams — University of Oxford; Pembroke College

Title: *AI regulatory approaches: do we need new rules, or to adapt existing ones?*

Abstract: The UK's present approach to regulating AI might be characterized as watching and waiting. Outside the EU and thus not bound by the EU AI Act the UK is not rushing to legislate, and the Data Use and Access Act 2025 if anything reduces the protections of the (UK)GDPR. One potential result of this position is that it may be left to existing legal rules to adapt to cover the challenges raised by AI. But this is not necessarily problematic. Focusing on the example of algorithmic decision-making, the law already has a set of rules tailor-made to control abuse of disparate power and to ensure transparent and fair decision-making in the form of the rules of judicial review. If we can adapt our existing rules, the need for new ones may become less pressing.

Summary: Professor Rebecca Williams began by setting out the existing regulation on AI globally - the EU AI Act at the supranational level, and its 'Brussels effect' on South Korea, Colorado, and Texas. However, some countries, such as the USA and Brazil, have instead chosen an innovation-first, anti-regulation approach. In comparison, the UK has oscillated between conservatism (such as at the Bletchley summit), and a pro-innovation approach. By way of example, she contrasted the narrowing of Article 22 of the UKGDPR through the Data (Use and Access) Act 2025 and its simultaneous provision to protect subjects of automated decision-making.

Prof. Williams then moved to criticize the simplistic binaries of 'pro-innovation' or 'pro-regulation' approaches. She gave four reasons: firstly, from a UK-specific perspective, key positioning in innovation is regarded as inherently linked to leading in regulation. Secondly, even

in the US, trust is seen as crucial to innovation and fostered by regulation. Third, regulation is not purely hard law - policy guidance may have an impact in practice. Finally, the price of comprehensive legislation may be vagueness, which in turn leads to less guidance. Focusing on the final point, Prof. Williams gave the example of Article 5(1)(c) of the AI Act Act, prohibiting social scoring. The provision contains uncertain terms like 'unrelated' and 'unjustified'.

She then explained that no one piece of top-down regulation can address every need: in particular, regulation tends to be *ex ante,* but in certain use cases, we may wish to respond to *ex post* harms. Regulation requires prioritization and policy decisions, meaning that some issues may fall to the wayside. Further, regardless, it will be necessary to supplement any comprehensive regulation with existing rules. As a result, if we can adapt our existing rules, the need for new rules is less pressing.

Analogizing with Robinson's work on criminal law, Prof. Williams explained that a potential solution lies in analogizing to the core: harnessing existing instincts and understanding, and applying them to new scenarios. Her primary example of an existing toolkit which could be adapted was of public law principles, which (amongst other principles) require that decision-makers must have the *vires* to make a decision, follow fair procedure, not fetter their discretion, and only take the right considerations into account. These principles can be analogized with tech regulation. The *vires* requirement is analogous to the requirement of a valid basis for processing data; the requirement to follow a fair procedure is akin to the right to meaningful information, or 'gisting'; and the requirement to only take the right considerations into account can be compared with Article 5(1)(c) on social scoring.

Prof. Williams concluded by considering the ways in which this toolkit could be used. She explained that it could help to interpret existing regulations, review public authorities' actions, and potentially even be used against private parties. In particular, she argued that it may be relevant when addressing the issue of social scoring.

Key Takeaways

- The division between innovation and regulation may be less sharp than we assume.
- Adaptation of existing legal frameworks and rules can help us to supplement new regulation, and reduce the need to regulate extensively to begin with.

- Principles from administrative law are particularly strong candidates for adaptation in the AI context, as they can be analogized with existing requirements in data protection and AI law.

References

European Union. (2024). *Artificial Intelligence Act, Article 5(1)(c)*. https://artificialintelligenceact.eu/article/5/

UK Government. (2025). *Data (Use and Access) Act 2025, c. 18*. https://www.legislation.gov.uk/ukpga/2025/18/enacted

Prof. Pierre Larouche — Université de Montréal

Title: *Competition, Regulation and Innovation in AI Geopolitics – Canadian Perspective*

Abstract: A theoretical framework is used to make sense of current debates in AI governance. At its core, the framework highlights bilateral trade-offs between risk regulation, competition, and innovation, combined with fundamental institutional design choices. The regulatory approaches of leading jurisdictions are then mapped onto this framework to provide a structured comparative perspective and to illuminate the geopolitical challenges that emerge. This analysis shows that the abrupt shift in U.S. policy under the new administration has positioned the European Union as a pivotal actor, with its next regulatory moves likely to influence the trajectory of global AI governance. This international backdrop serves to contextualize the Canadian perspective, both in terms of its evolution to date and the ways it may be reframed by the incoming government.

Summary: Professor Pierre Larouche structured his presentation around three parts: a theoretical framework for comparative AI governance, the shifting dynamics in U.S. and European policy, and the implications for Canada. He introduced a conceptual framework that situates regulation, competition, and innovation as interdependent forces. Protective regulation channels innovation toward safer outcomes, while permissive regulation enables more disruptive advances. Competition fosters diversity and ambition in innovation, yet excessive

concentration risks oligopolistic dominance where states lose traction over firms. The balance between these elements, he argued, defines the space in which effective governance must operate.

Turning to the international stage, Larouche mapped current regulatory approaches onto this framework. The European Union initially positioned the AI Act as a global benchmark, hoping to replicate the “Brussels effect” of the GDPR. However, critiques and the Draghi Report of September 2024 questioned its effectiveness, noting a fear of missing out as other jurisdictions advanced. Meanwhile, the launch of DeepSeek in January 2025 underscored China’s progress, while a new U.S. administration marked a decisive policy reversal. Washington shifted toward industrial policy, open-source ecosystems, and support for startups and academia, while antitrust cases against major tech firms continued. This evolution placed the U.S. at the center of global momentum, while the EU now faces a strategic choice between aspiring to become a “third digital empire” or leading a coalition of non-aligned jurisdictions.

In this comparative landscape, Larouche emphasized that many jurisdictions such as the UK, Japan and Singapore have opted for more cautious and dialogic approaches rather than broad legislative frameworks. These rely on co-regulation and industry engagement, contrasting with the EU’s horizontal model. China, by contrast, has pursued targeted interventions backed by expansive industrial policies. This patchwork of strategies illustrates the geopolitical stakes of AI governance, where competition, innovation and regulation intersect differently across contexts.

Finally, he turned to the Canadian perspective. Canada once sought leadership through initiatives like federal procurement guidelines and active participation in GPAI, but the fate of Bill C-27 remained uncertain during the 2025 election period, leaving the country trailing. The new government has signalled a possible change in direction, with discussions around appointing an AI minister and setting priorities such as scale, adoption, trust, and sovereignty, often summed up in the phrase "light, tight, and right." Canada, however, lacks the market size to position itself as a digital empire and must instead prioritize compatibility with other governance models. Without a clear strategy, it risks becoming a mere satellite of the U.S. digital empire. The best path forward may lie in building coalitions with jurisdictions that seek alternatives to U.S. and EU models, while fostering domestic adoption and maintaining trust. Larouche concluded that for Canada to succeed, policymakers must step beyond traditional comfort zones, balancing regulation with industrial strategy and developing deeper dialogue with industry actors.

Key Takeaways

- Regulation, competition, and innovation are interdependent and must be balanced.
- The U.S. has taken a pivotal role in AI governance, emphasizing industrial policy and open-source ecosystems.
- The EU's AI Act faces skepticism, with doubts about its global influence despite initial ambitions.
- Canada has fallen behind but could align with non-aligned jurisdictions to regain relevance.
- Effective governance requires stepping beyond subsidies and adopting co-regulatory strategies with firms.

References

European Commission (2024). *Draghi Report on EU Competitiveness.* https://commission.europa.eu/topics/eu-competitiveness/draghi-report_en

OECD (2025). *Global Partnership on AI (GPAI) – Integration into OECD Framework.* https://www.oecd.org/en/about/programmes/global-partnership-on-artificial-intelligence.html

U.S. Department of Justice (2025). *Google Search Antitrust Case – Remedies.* https://www.justice.gov/opa/pr/department-justice-wins-significant-remedies-against-google

Prof. Melissa Hyesun Yoon — Hanyang University

Title: *South Korea's AI Governance Framework: Balancing Innovation and Regulation through the AI Basic Law*

Abstract: This presentation examines South Korea's comprehensive approach to AI governance through the recently enacted "Framework Act on the Development of Artificial Intelligence and the Establishment of a Trust-Based Foundation" (commonly referred to as the "AI Basic Law") and its ongoing implementation decree preparations. I will analyze how South Korea is attempting to balance technological innovation with ethical considerations and regulatory oversight, drawing comparisons with other Asian approaches, particularly Japan's AI Promotion Act where relevant. The presentation will highlight key provisions of the Korean framework, including risk-based regulatory approaches, AI ethics guidelines, and mechanisms for public-private collaboration. I will also discuss the challenges and opportunities in implementing this framework within South Korea's unique technological and social context, offering insights for international regulatory harmonization efforts.

Summary: In this talk, Professor Melissa Hyesun Yoon aimed to give a concrete understanding of the South Korean approach to AI regulation and its positive and negative aspects. Whereas other jurisdictions have focused on comprehensive regulation (EU), sectoral self-regulation (the UK, Japan) and technology-specific regulation (China), Korea has adopted an approach of 'balance-seeking selective regulation'.

She began by highlighting Korea's unique position on the global stage: it is the only country outside of the US and China with near-complete sovereign AI capabilities, as it has its own foundation models, training infrastructure, and safety technology. Its sole critical gap is chip technology, as it remains dependent upon imports. With this in mind, Korea has approached regulation with the aim of becoming a Top 3 country in AI, and supported its efforts with significant financial investment and a presidential steering committee.

Prof. Yoon then set out the framework and features of Korea's flagship law: the AI Basic Law (Framework Act on the Development of Artificial Intelligence and the Establishment of a Trust-Based Foundation). The Act aims to balance innovation with safety through risk-based regulation, with an emphasis upon ex-post regulation rather than ex-ante regulation, and

responding to actual, rather than hypothetical, risk. In turn, it hopes to reach a 'Goldilocks' position in comparison with other jurisdictions. She set out three key features of the Act: an equal weight for innovation and trust, the use of selective targeting, and the utilization of mild penalties. She then focused on the tiered regulatory approach within the Act, which distinguishes between high-impact AI, generative AI, high-performance AI, and multilayered systems.

Having established the framework of the Act, Prof. Yoon then moved to address its implementation challenges. Firstly, she argued that it struggles with definition clarity: the definitions of AI systems and high-impact determination processes are circular, and the boundaries are ambiguous. Secondly, the responsibility distribution under the Act fails to recognize the complexity within AI value chains, and the different responsibilities of deployers and developers. Consequently, issues of liability remain uncertain. Thirdly, the Act struggles with global alignment, particularly in light of compatibility with international standards and cross-border enforcement. Finally, the infrastructure for transparency under the AI Basic Law provides for limited public disclosure and stakeholder engagement, meaning that there are gaps in information accessibility.

Prof. Yoon then focused further on the AI Basic Law's 'critical gap': foundation model regulation. Whereas the Act adopts a system-focused approach, the reality of AI value chains is model centric. As a result, although most AI services use foundation models, Korean law excludes these models from regulation. In contrast, providers of cloud systems and API services are subject to full compliance requirements - and Korean companies using these APIs must comply with additional obligations. This leads to asymmetry, wherein companies subject to full liability are likely to struggle to obtain the necessary technical information from foundation model providers due to a lack of transparent information flow.

Finally, Prof. Yoon concluded her presentation by focusing on the key takeaways from the AI Basic Law's implementation. She reiterated the Act's issues, but in turn adopted a broader perspective: although Korea's experience demonstrates the ongoing challenges, other major jurisdictions have fared no better. The attempt to do so simply reveals that every country is struggling with the same impossible tradeoff between innovation and safety.

Key Takeaways

- South Korea aimed to find a 'Goldilocks' position in AI governance, emphasizing negative regulation and responding to actual, rather than hypothetical, risk.
- However, the AI Basic Law struggles with definition clarity, global alignment, transparency and responsibility distribution.
- In particular, it adopts a system-centric, rather than model centric, approach, which fails to acknowledge the practical reality of AI value chains.
- The flaws in the Act reflect a deeper issue facing all jurisdictions: every country is struggling with the same impossible tradeoff between innovation and safety.

Reference

South Korean Ministry of Government Legislation. (2025). Framework *Act on the Development of Artificial Intelligence and the Establishment of a Trust-Based Foundation (*인공지능 발전과 신뢰 기반 조성 등에 관한 기본법*)*. https://www.law.go.kr/lsSc.do?menuId=1&subMenuId=15&tabMenuId=81&query=ai#undefined

Prof. Juan David Gutiérrez — Universidad de los Andes

Title: *Emerging AI Regulatory Approaches in the Global South*

Abstract: We are witnessing a global trend of growing interest in introducing regulations that address AI. For example, after creating a novel database that maps AI bills and regulations, we documented over 600 regulatory instruments submitted, discussed, and/or approved in twenty-five Latin American and Caribbean countries and territories. This paper examines the rules and regulatory projects that directly and indirectly address AI development, acquisition, adoption, deployment, and use in the Global South. The text characterizes diverse regulatory tools (e.g., audits, transparency instruments, etc.) and nine AI regulatory approaches: principles-based, standards-based, agile approaches, facilitator approaches, adaptive

approaches, mandatory disclosure approaches, rights-based, risks-based, and liability approaches. Finally, the paper discusses the policy and political challenges associated with implementing AI regulation in the Global South.

Summary: In this presentation, Professor Juan David Gutiérrez aimed to examine how to regulate for the emergence of AI: focusing on the differing regulatory approaches across countries in the Global South. In doing so, he built upon his work with both UNESCO and the Universidad de los Andes.

Prof. Gutiérrez began by establishing his definition of regulation: binding rules issued by public bodies. Regulation sits alongside a variety of other AI Governance Instruments used by States, including case law policies, guidelines and other 'soft law', and informal rules. He then moved on to address the State's multifaceted relationship with technology. At once, it is a regulator and supervisor, facilitator and enabler, developer and buyer, and deployer and end user. It must thus attempt to use regulation to tackle these respective roles.

Prof. Gutiérrez then moved on to an overview of the global conversation on regulation, and established that the number of AI-related bills passed into law globally has increased in recent years (Stanford Institute for Human-Centered Artificial Intelligence (HAI), 2025). Focusing on Latin America and the Caribbean, both regions have seen an explosion of regulation, with over 600 AI-related regulatory instruments across the two (Gutiérrez and Hurtado, 2025). This was predominantly focused in five countries: Brazil, Mexico, Argentina, Colombia, and Peru.

He then established that AI-related regulatory instruments are not the monopoly of legislative bodies. Whilst they may be legislative, they are also potentially executive or judicial. For example, Peru's government issued a decree developing its congress-issued national AI law, and Brazil's electoral body (which is judicial in nature) issued a general regulation on the use of generative AI in the context of electoral processes.

Prof. Gutiérrez then examined the years in which different regulatory instruments in Latin America and the Caribbean started their regulatory process, aiming to capture the level of conversion of AI. He concluded that the explosion of regulation began in 2023, and was fully realized in 2024, with 2025 likely to match or outperform 2024.

Finally, Prof. Gutiérrez set out nine different emerging AI regulatory approaches, based on his work with UNESCO. These were: (i) principles-based, (ii) standards-based, (iii) agile and experimentalist, (iv) facilitating and enabling, (v) adapting existing laws, (vi) access to information and transparency mandates, (vii) risk-based, (viii) rights-based, and (ix) liability.

He concluded that these differing approaches indicate that there is no one-size-fits-all answer to AI regulation. Rather than focusing solely on the EU, Chinese, or US models, it is important to recognize the alternative paths which are being explored by different jurisdictions. In turn, the debate on AI regulation reflects deeper questions: those of the type of State we want to have, the type of citizen-state relationship we aspire towards, and the society we ultimately want to live in.

Key Takeaways

- The State has a complicated relationship with AI: it is at once a buyer, facilitator, deployer and supervisor.
- AI regulation is not necessarily the monopoly of legislative bodies - rather, it may be judicial or executive.
- Across Latin America and the Caribbean, the numbers of AI-related regulatory instruments have spiked since 2023.
- It is important to acknowledge the different regulatory approaches outside of the US, China, and the EU - and the way in which the debate around AI regulation reflects deeper issues of society and democracy.

References

Gutiérrez, J., & Hurtado, M. (2025). *Gestión del conocimiento en la era digital: Tendencias, retos y oportunidades en el desarrollo empresarial*. https://dialnet.unirioja.es/servlet/articulo?codigo=10020901

Stanford University, Human-Centered Artificial Intelligence. (2025). *The 2025 AI Index Report*. https://hai.stanford.edu/ai-index/2025-ai-index-report

UNESCO. (2024). *Consultation paper on AI regulation: Emerging approaches across the world*. https://unesdoc.unesco.org/ark:/48223/pf0000390979

Prof. Benjamin Guedj — University College London; Inria

Title: *When Law Meets Code: Technical Hurdles in Implementing AI Regulation*

Abstract: Efforts to regulate AI often run into a fundamental difficulty: the gap between high-level legal principles and the technical realities of AI systems. As a machine learning (ML) researcher, I will highlight why core implementation challenges — such as defining transparency, auditing complex models, ensuring robustness under distributional shifts, and certifying compliance at scale — resist simple solutions. These challenges are not only technical but also shape what kinds of regulation are feasible in practice. My aim is to shed light on where regulation collides with current ML capabilities, and to outline opportunities for collaboration between regulators, technologists, and researchers to make regulation both effective and realistic.

Summary : Professor Benjamin Guedj explored the complex "Regulation-Reality Gap" that arises when attempting to tackle the complex task of translating regulatory principles into technical reality. He introduces the issue, which stems from how principles endorsed in regulatory frameworks or guidelines, are usually hard to implement in code. Concepts that bear societal importance such as fairness, transparency, safety, or accountability, do not have a unified and clear mathematical definition. This creates a first dimension of complexity for the question of pragmatically enforcing governance principles. The implementation of these principles in machine learning tools is made harder by the complexity and changing nature of systems. The concept of explainability exemplifies some of these issues. Explainability, as Professor Guedj points out, can map to various technical practices: from saliency maps, to counterfactuals and feature attributions. It can also be quantified by an array of available metrics.

In other words, we have observed time and time again that technology evolves much faster than the law. Thus, the two are not consistent; and laws can become unworkable, due to their vagueness. This calls for governance to lessen their detachment from technical reality, and perhaps offer higher flexibility.

Professor Guedj goes on to lay out the core challenges faced during implementation of regulatory principles. Among them, the issues of transparency and explainability: as the speaker points out, deep models make explainability harder because their complex probabilistic

predictions are uninterpretable for humans, something known as the black-box problem. Robustness to distributional shift is another important challenge in implementation, for which the speaker recommends regulation to mandate stress testing across realistic distributional shifts and post-deployment monitoring: crucial practices that vague terms might not require. The speaker also highlights the lack of a well-established audit framework for AI systems, making auditing and verification harder to implement without, for example, compromising data privacy. Concerns around bias and fairness are also very relevant, considering the multitude of competing definitions that fall under these umbrella terms. Finally, another important question is that of the scalability of compliance: as model or dataset sizes grow, so do compliance costs, making it harder especially for lower-scale organizations or companies to absorb the cost of compliance.

From these technical considerations, Professor Guedj highlights some implications they hold for regulation. He points out the importance of balancing ambition with feasibility, which can be made easier in a couple of ways. The importance of focusing on outcomes and properties (say, robustness) over brittle checklists, and of prioritizing transparency of processes and evidence, rather than a single explanatory method, both play crucial roles in this. As he also points out, iteration can play a key role: phased obligations, sandboxes and post-deployment monitoring can all indeed be highly beneficial.

Highlighting the intrinsic link between society, technology and law, Professor Guedj concludes by putting forth implementation as the step at which regulation either fails or succeeds. Thus, technical reality should shape what is enforceable and useful. For this to happen successfully, collaboration across disciplines is essential. The law needs to meet code to enable responsible deployment of AI Systems.

Key Takeaways

- There is a persistent *regulation–reality gap*: legal principles like fairness or transparency lack precise technical definitions, making enforcement difficult.
- Transparency and explainability remain unresolved due to the black-box nature of deep models and competing technical methods.

- Robustness to distributional shifts and the absence of standardized audit frameworks are major hurdles for safe and accountable AI.
- Compliance costs scale with model size, creating disproportionate burdens on smaller organizations.
- Effective AI regulation requires balancing ambition with feasibility through outcome-based rules, iterative approaches (e.g., sandboxes), and strong collaboration between regulators and technologists.

Prof. Christian Gagné — Université Laval; Institut intelligence et données

Title: *The Case for National AIs*

Abstract: The considerable advances of AI in the last few years, in particular with Large Language Models (LLMs) and other Foundational Models (FMs), have announced a period of important technological advances that are already significantly impacting the economy and society. However, these technological advances were controlled mostly, until recently, by Big Tech American companies. Given the significant turmoil we have seen since the recent US presidential election, there is a significant erosion of thrust that has led to question our current dependencies from US technological companies regarding AI. The capacity to develop a stronger digital sovereignty leads to the idea of having national AIs, with LLMs and FMs that are built by and for citizens of a given nation, better reflecting their culture, values, and languages while being developed and deployed on local technological infrastructures. In this presentation, I will develop the case for such national AIs, the surrounding technological and societal context, and the conditions required for achieving them.

Summary: Professor Christian Gagné explores some of the recent milestones in AI development, and how their unfolding can motivate new avenues, such as the development of national AI systems. Indeed, major breakthroughs such as Large Language Models (LLMs) and Foundation Models (FMs) have had a transformative impact, possibly the biggest one since the World Wide Web's appearance in the mid 1990s. However, resulting advanced systems are mostly controlled by a few Big Tech companies based in the United States. The models of these extremely wealthy and powerful entities present a strong bias towards Anglo-American culture, as they develop models trained on the web's content. One reason for this concentration of

power is the scarcity and expensive nature of computation resources and machine learning expertise. Or so it mainly was, until the Chinese company DeepSeek presented their highly performant models, despite having access to less advanced hardware than US-based companies such as OpenAI or Google. This shows it would be possible for national AIs to emerge; systems that could reflect local culture and values, while supporting digital sovereignty and allowing the development of local expertise.

As Professor Gagné points out, data played a fundamental role in the revolutionary advances that have brought LLMs and FMs. This data is often based on web-scraped content (and, potentially, additional sources), which, at scale, is quite unresolved legally. Numerous cases have now pointed out how web scraping is often disrespectful of the law. While the scaling law for LLMs states that more data requires more computing power, the belief that only Big Tech can develop competitive models is erroneous. Expertise can be equally important, especially for certain topics organized in tight research circles; and computational capacities are also key.

The speaker goes on to discuss questions of confidentiality and sovereignty. As he points out, the current international state increasingly seems to near the end of *pax Americana,* the period of relative peace that followed World War II, promoting liberal democracy in a movement led by the US. In these circumstances, there is a need to reduce reliance on the US, especially in AI. Since LLMs, which are now mainly developed in the US, collect data from its users, confidentiality also becomes a concern. Under the Patriot Act and the Cloud Act, the US government can even access cloud-stored information on anyone – even non-US users. As such, national AI initiatives could substantially reduce such concerns by reducing reliance on the US tech sector. As Professor Gagné points out, LLMs and FMs are still in such an early stage that it would be possible to catch up and develop local technology for a global impact: an opportunity to promote digital sovereignty.

Diving into how such local development could take place, the speaker starts by highlighting the strong open science culture in machine learning research: from open-sourcing code, papers and models, to providing information for transparency and reproducibility. Another key question is the representation of national cultures and languages: current LLMs, with their Anglo-American bias, understand French less than English: let alone lower-resource dialects and other regional aspects. Future LLMs could be adapted to reflect and support these local cultures, and serve as building blocks to build a variety of adapted tools. Finally, regarding access to large-scale data, Professor Gagné points to some open sources, such as Common Crawl web graphs, while reminding the importance of using good scraping and collection practices, such as traceability, right of removal, and intellectual property. To address these, some initiatives such as Quebec's National Archives (BanQ)'s development of a 'local' dataset can serve as alternatives.

In a similar wave, Switzerland recently came out with a national AI initiative, where two universities collaborated on building a public and fully open infrastructure. Such initiatives may stand as examples as we try to move forward through the current state of geopolitical chaos; one in which the capacity to develop national AIs, can become a matter of economic security. The speaker finally calls on Canada, France and the UK to step up as leaders in this initiative, and figure out its feasibility, especially while aligning with environmental regulations.

Key Takeaways

- Recent breakthroughs in LLMs and FMs, though dominated by U.S. Big Tech, open opportunities for national AI systems that reflect local cultures, values, and languages.
- Digital sovereignty is a central motivation: reliance on U.S. platforms raises concerns over cultural bias, confidentiality, and exposure to laws such as the Patriot Act and Cloud Act.
- Data remains a cornerstone for AI, but issues of legality, scraping practices, and intellectual property demand stronger governance and responsible collection.
- Open science, national datasets (e.g., BanQ), and initiatives like Switzerland's public AI infrastructure show viable paths for building local capacity.
- Developing national AIs is both a strategic and geopolitical issue, requiring leadership from countries like Canada, France, and the UK, while balancing innovation with environmental sustainability.

Alexei Grinbaum — Research Director, CEA-Saclay

Title: *From AI ethics to AI regulation and back: operationalizing the AI Act recital 27*

Abstract*:* I will describe the context of AIOLIA project training in AI ethics, starting from the sources of ethical tension in AI system design and all the way down to the tensions concerning the research exception in the EU AI Act. I will then briefly introduce the AIOLIA training module.

Summary: The speaker began by giving an overview of the EU AI Act, and its timeline for implementation. In particular, he highlighted the combination of AI literacy rules, codes of practice, and high-risk rules across Annex III and Annex I categories, and their staggered introduction until August 2027.

He then moved to discuss Article 2 of the AI Act, and its exclusion of AI models which are 'specifically developed and put into service for the sole purpose of scientific research and development' (Art 2.6). He explained that it is difficult to conceptualize an AI system which would fall under this category, at least if we define it strictly. Any originally scientific or open-access model has the potential for later commercialization. Referencing his previous work, the speaker analogized with the dual-use military/civil concern approach taken to other regulatory frameworks, such as those for biotechnology (Grinbaum and Adomaiyis, 2024).

The speaker then addressed the issues with the definition of an 'AI system' under the February 2025 Commission Guidelines for the AI Act. In particular, he highlighted the dissonance between what many researchers would have considered to constitute an 'AI system', and the finalized definition in the Commission's guidance, focusing on the position of Bayesian learning, knowledge representation and reasoning, and time series analysis and forecasting.

He then discussed the role of ethical principles in the EU's original 2019-2020 guidelines, and the seven core principles which were ultimately included in Recital 27 of the AI Act: human agency and oversight; technical robustness and safety; privacy and data governance; transparency; diversity; nondiscrimination and fairness; societal and environmental wellbeing; and accountability. He discussed the influence of the Independent High-Level Expert Group on Artificial Intelligence, and their Assessment List for Trustworthy Artificial Intelligence (ALTAI). Horizon Europe's approach to AI ethics integrates these ethical considerations into research projects, applying the same ethics by design and ethics of use approaches.

Further focusing on these seven principles, the speaker discussed the tension inherent in applying them 'by design', and the way in which prioritizing one may require sacrificing another.

Taking the example of face recognition technology, he explained that the requirements of security and privacy are inherently at odds in this context. Whereas prioritizing security would entail recording as many parameters as possible, and potentially applying them in contexts such as neighbourhood control, this is contrary to a privacy-centric approach. This is particularly true in light of the fact that the meaning of many parameters formulated by face recognition neural networks is unknown.

He then moved to a second example of disease recognition, focusing on the way in which the definition of these ethical principles may shift depending on the context in which they are applied. In this particular context, the meaning of 'explainability' could differ greatly depending on the reason *why* explainability is necessary - for instance, explainability from a patient's perspective is different to explainability from a debugging perspective.

He finally discussed his previous work addressing the link between bioethics and AI ethics (Aucouturier and Grinbaum, 2025), and the eight-part checklist used to identify and select serious and complex issues. He highlighted the AIOLIA framework, and the way in which it could be used to classify the risk for each requirement in differing scenarios. He then discussed the balance between ensuring compliance with principles of ethics by design, and ensuring the efficacy of the AI system itself, giving the example of virtual friends and the tension in determining their place on the spectrum of 'tool' and 'friend'.

Key Takeaways

- Both the AI Act and the Commission's guidance contains a lack of clarity and uncertain definitions.
- Ethical principles and an emphasis on ethics by design have been incorporated throughout AI regulation.
- The specific definition of each principle will vary depending on the use case and individual context.
- Frameworks such as AIOLIA can be used to classify scenarios in depth.

References

AIOLIA Project. (2025). *AIOLIA project*. https://aiolia.eu/

Aucouturier, J.-J., & Grinbaum, A. (2025). Training bioethics professionals in AI ethics: A framework. *Journal of Law, Medicine & Ethics*. https://www.cambridge.org/core/journals/journal-of-law-medicine-and-ethics/article/training-bioethics-professionals-in-ai-ethics-a-framework/B3066FEED17A41A2D962ABC239455B1F

European Commission. (2025). *Guidelines on the definition of an artificial intelligence system*. https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-ai-system-definition-facilitate-first-ai-acts-rules-application

Grinbaum, A., & Adoimatis, A. (2024). Dual use concerns of generative AI and large language models. *Journal of Cyber Policy*. https://www.tandfonline.com/doi/full/10.1080/23299460.2024.2304381#abstract

Szegedy, C., Zaremba, W., Sutskever, I., Bruna, J., Erhan, D., Goodfellow, I., & Fergus, R. (2013). Intriguing properties of neural networks. *arXiv*. https://arxiv.org/abs/1312.6199

The Honourable Judge Simon Ruel — Québec Court of Appeal

Title: *From Promise to Peril: The Uses and Regulation of AI by the Judiciary*

Abstract: The judiciary faces a dual challenge with respect to the use of AI. On the one hand, AI systems can strengthen justice by making it faster, more accessible, and more consistent. However, it can also threaten justice by introducing bias, eroding confidentiality, or undermining judicial independence and impartiality. The central question is not whether AI will enter courtrooms. It already has, at least to some extent, in Quebec and Canada. The key issue is how AI will be integrated, regulated, and controlled so that it enhances rather than compromises the legitimacy of judicial decision-making.

The full text of this presentation is available **at this link**.

Summary: Judge Simon Ruel offered a comprehensive reflection on the opportunities and challenges that AI presents for the judiciary. His presentation examined how AI could both enhance and endanger the administration of justice, emphasizing the urgent need for deliberate and ethically grounded integration.

He began by outlining the potential benefits of AI in judicial decision-making. When properly designed and deployed, AI systems could assist judges by automating repetitive or technical tasks, thereby allowing them to concentrate on their essential role of weighing arguments, exercising judgment, and articulating the reasoning behind their decisions. One of the most promising applications lies in legal research and analysis. AI has the capacity to process and synthesize vast bodies of case law, statutes, and doctrine, which could, in principle, increase the consistency and predictability of judgments. This is particularly relevant in common law systems, where judges must ensure that similar cases are treated consistently in accordance with the principle of parity. Such research is time-consuming, and therefore an ideal candidate for AI support.

However, current tools remain inadequate for the Canadian legal context. Publicly accessible systems such as ChatGPT or Copilot are not trained on Canadian legal databases, including the jurisprudence of the Supreme Court of Canada, nor on Quebec's civil law corpus. Even commercial systems face significant limitations, including a lack of bilingual and cross-jurisdictional capabilities. These gaps create blind spots that undermine reliability, particularly in a bijural and bilingual jurisdiction such as Quebec.

Beyond research, AI could play an important role in administrative and evidentiary support. Judges often face immense volumes of documents, ranging from written submissions and expert reports to satellite imagery and social media content. AI can assist in organizing and summarizing such material, making complex or high-volume cases more manageable and reducing the chronic backlogs that threaten access to timely justice. In Quebec, hearings are still not automatically transcribed, which creates costs and delays for appeals. AI-based transcription and anonymization could improve accessibility, speed, and clarity in both official languages, without replacing human review. In specialized tribunals dealing with standardized cases such as tenancy, small claims, or social security, AI might also assist in generating draft decisions, provided that judges retain full control over reasoning and outcomes.

Judge Ruel emphasized that these potential advantages cannot be separated from profound ethical and governance challenges. The first concern is the preservation of the fundamental values of justice: fairness, independence, impartiality, equality, and respect for human dignity. Judicial reasoning is inherently human, rooted in empathy, moral discernment, and contextual understanding. No algorithm can replicate these qualities, and delegating judgment to machines would erode the human dimension of justice, reducing decisions to mechanical outputs detached from compassion and nuance. For that reason, human oversight must remain constant, especially in any moderate or high-risk use of AI.

He also highlighted the importance of confidentiality, security, and sovereignty. Judicial data often include sealed records, confidential evidence, and sensitive testimonies that must remain strictly protected. To prevent AI models from inadvertently learning from such material, secure environments will be required to ensure that sensitive data do not enrich algorithmic systems. Equally, questions of data ownership and storage are crucial. If AI infrastructures are controlled by foreign providers, the independence of Canadian courts could be compromised. The Canadian Judicial Council has made it clear that all classified judicial data must remain within Canadian jurisdiction, a principle that must extend to AI systems to safeguard judicial independence.

Transparency and accountability represent another central concern. Judicial reasoning depends on traceability and justification, which means that opaque systems are fundamentally incompatible with judicial standards. Judges must be able to verify, explain, and, if necessary, challenge the reasoning behind AI-generated outputs. This is particularly important given the growing number of AI hallucinations, where systems fabricate citations or misrepresent legal precedent, undermining credibility and public confidence.

For these reasons, continuous education is essential. Technological literacy is now part of judicial ethics. Judges must understand the limits and biases of AI systems as well as the

implications of data provenance and prompt design. Training and awareness should therefore become integral to judicial education to ensure responsible and informed use.

Turning to Canada and Quebec, Judge Ruel observed that the integration of AI could significantly improve access to justice, particularly for self-represented litigants who might use AI to conduct legal research or draft documents. Yet he also noted that institutional inertia, incomplete digitization of court records, and fragmented technological infrastructures hinder progress. In Quebec, the judiciary does not have full control over its technological environment, which remains under provincial administration. This dependence on shared digital systems limits the autonomy of courts and delays innovation. By contrast, other provinces that have achieved greater administrative independence over technology have been able to advance more rapidly in modernizing their judicial operations.

Judge Ruel situated these national challenges within a broader international context. Several jurisdictions have already begun experimenting with AI in judicial systems. China has implemented Intelligent Trial 1.0, a system that automates case classification and document management. Singapore uses the Intelligent Court Transcription System to transcribe hearings in real time. India's Supreme Court employs SUPACE, a platform that assists in cataloguing precedents and processing case materials. Brazil's Supreme Federal Court uses the VICTOR system to organize appeals efficiently, while in the United States, the National Center for State Courts has created an AI Sandbox to allow judges to explore these technologies in a secure environment. These examples illustrate that AI can be responsibly integrated into judicial systems when supported by robust ethical, institutional, and legal frameworks.

In conclusion, Judge Ruel described AI as both a promise and a peril for the judiciary. Used wisely, it can strengthen access to justice, reduce delays, and allow judges to focus on their essential human function: judging. Used carelessly, it risks undermining the very foundations of fairness, independence, and public trust. The judiciary must therefore approach AI with caution and deliberation, modernizing to meet public expectations without compromising the human essence of justice. Properly designed and governed, AI can become a valuable ally in making justice more accessible, transparent, and resilient.

Key Takeaways

- The question is no longer whether AI will enter the courtroom, but how it will be governed to ensure it strengthens rather than undermines justice.
- AI can assist judges with legal research, case law analysis, and document management, improving efficiency and access to justice while helping to reduce court backlogs.
- Its deployment raises critical issues of data sovereignty, confidentiality, and the preservation of the human and ethical foundations of judicial reasoning.
- Judicial independence, fairness, and accountability must remain non-negotiable. AI should support, not replace, human judgment and empathy.
- The judiciary should embrace innovation deliberately and cautiously, ensuring that transparency and the rule of law remain at the heart of AI integration.

Paul Gagnon & Misha Benjamin — Partners, Technology and Artificial Intelligence Group, BCF

Title: *News from the front – Navigating AI regulation in practice*

Abstract: This session aims to highlight key learnings and emerging trends from two leading attorneys in the field of AI. With an international practice representing both AI providers and adopters, Misha and Paul will discuss how regulation is shaping contract negotiations and AI product design. The session also aims to explore the goals and impacts of emerging AI regulation such as: (i) regulation as a competitive moat for Big Tech; (ii) regulation as a driver of innovation; and (iii) the impact of local regulation on companies with global reach and ambitions. Bringing practical and hands-on experience, the two speakers aim to highlight limits and opportunities found in emerging AI regulation.

Summary: The speakers began by detailing their original experience working with AI-based startups, almost nine years ago. At the time, there was no established playbook for the issues, but the same topics were relevant then as today: not only the flagship Acts, but also issues of responsible deployment. They explained how regulatory requirements such as privacy by design

and guardrails are not opposed to commercial success: rather, the best companies aim to incorporate them for the outset.

They then explained the lack of clarity regarding regulations which apply to the deployment and operation of technology, particularly in data-rich environments. More clarity is needed from existing regulators: taking automated banking as an example, they explained that the comprehensive guidelines on decision-making should be taken as good practice. This in turn linked into the question of whether AI should be regulated as an object at all.

On the flip side, however, the speakers explained that innovators don't build new products with regulation in mind - it is difficult to strike the balance between ensuring compliance down the line and stifling innovation. Smart regulation may tie into innovation, but it is still imposed.

The speakers then detailed a shift in tone that they had observed in practice, from arguments that AI could *not* be regulated to a belief on the part of major players that regulation was necessary. Guardrails in privacy were used as an incentive to capture audiences from rival products: the first companies to make commitments to produce consumer data were rewarded with growth in their consumer bases. Other companies would then respond, leading to dialogue. However, now that many consumers have been locked into their product of choice, companies have shifted back. A similar pattern has been seen in respect of copyright infringement.

The speakers then explained that the discourse around regulation fails to reflect the fact that the debate is about both art and science, and a question of what actual human oversight looks like. Even if an optimal oversight mechanism can be established, human questions remain: questions of staffing, allocation of resources, and organizational oversight. Further questions arise in relation to liability. As a result, AI cannot be viewed solely as an object of technology - but also one of human resources and litigation.

They then moved on to discuss the use of existing vehicles to legislate and regulate in response to AI. In order to regulate efficiently, rules have been introduced into ill-fitting bodies of law: for instance, consumer protection has been tied to privacy under the GDPR, and competition law has attempted to regulate tech-specific realms. The practical considerations of which existing vessels AI regulation can be attached to is accompanied by a lack of distributed expertise across different regulators, despite the fact that they should all have a uniform, coordinated approach. This has a knock-on effect in respect of remedies: if AI regulation is shoehorned into an existing field of law, that field must already have established suitable remedies for this specific use case. In turn, issues arise in enforcement. A lack of manpower, expertise and funding leads to a lack

of meaningful enforcement. This risks regulation only existing for those who are already legally literate and careful - rather than everyone.

The speakers concluded by assessing the current state of AI regulation as a whole, arguing that it should not solely be viewed as a risk. Rather, it is important to assess the risks at hand, but adjust based on each individual customer's personal tolerance.

Key Takeaways

- Regulation may incentivize and support innovation, but is ultimately imposed - it is difficult to strike the balance between supporting new ideas and ensuring compliance once a product is scaled up.
- Major players in AI have engaged in dialogue, adjusting the levels of consumer protection they provide in order to incentivize users to shift from rival companies. However, once they acquire a captive user base, they shift back.
- Existing bodies of law have been used to shoehorn in regulation on AI, but this has knock-on effects further down the line, particularly in respect of remedies and effective enforcement.
- From an advisory perspective, the approach taken to AI regulation should be adjusted based on each individual customer's risk tolerance.

Benjamin Prud'homme — Vice President, Public Policy, Safety and Global Affairs, Mila

Title: *Global AI safety and international alliances in a new geopolitical context*

Abstract: This presentation reviews the mandate, structure, and content of the *International Scientific Report on the Safety of Advanced AI*, chaired by Yoshua Bengio. It highlights the rapid yet uncertain trajectory of general-purpose AI, as well as the politicization of the term "AI safety" and its implications in the current context of AI development. The discussion also considers the evolving geopolitics of AI, with a focus on the role of middle powers and multilateral organizations in shaping global governance as the world order undergoes profound changes.

Summary: Benjamin Prud'homme presented the *International Scientific Report on the Safety of Advanced AI*, chaired by Yoshua Bengio. He stressed that the report is not a strategy but a scientific assessment aimed at informing policymakers. Commissioned after the 2023 AI Safety Summit at Bletchley Park, its mandate comes from around 30 countries along with the EU and the UN. Over 70 experts contributed, focusing on three guiding questions: what general-purpose AI systems can currently do, what risks they pose, and what mitigation techniques are available.

He outlined the report's findings on capabilities. General-purpose AI is advancing rapidly, with notable improvements in reasoning and programming. Recent trends include "inference scaling" and the development of AI agents capable of browsing, coding, and research tasks, though they still struggle with complex multi-step reasoning. The trajectory of progress remains highly uncertain. Some experts see a slow evolution, while others warn of breakthroughs that could accelerate development dramatically, including advances that might themselves increase the speed of future progress.

The report identified a wide spectrum of risks. Well-established harms include scams, biased outputs, and the creation of child sexual abuse material (CSAM). Emerging risks include biological misuse, cyberattacks, persuasion and strategic behaviours, and the possible erosion of human control. Experts disagree on timelines: some consider such threats decades away, while others warn of societal-scale harms within years. Prud'homme emphasized the dilemma facing policymakers: act preemptively on limited evidence or risk being unprepared for rapid and

disruptive developments. The debate around open-weight models illustrates this challenge, as they foster transparency and research but also facilitate malicious use.

In conclusion, he underlined the uncertainty of AI trajectories and the dependence of outcomes on societal and governmental choices. Both highly positive and highly negative futures remain possible. The report calls for stronger international collaboration, with the UK continuing to host the secretariat and Bengio remaining as chair through 2025. Prud'homme also reflected on the politicization of the term "AI safety" and its implications for global debate. He highlighted the need to involve middle powers and multilateral organizations in governance discussions, while recognizing that not all regions are equally affected. For many in the Global South, issues around large language models are less relevant than immediate concerns such as access, inequality, or different sectoral priorities, underscoring the need for inclusive approaches.

Key Takeaways

- The *International Scientific Report* provides scientific evidence, not strategy or prescriptions.
- AI capabilities are advancing rapidly, with trends such as inference scaling and AI agents.
- Risks include scams, CSAM, bias, biological misuse, cyberattacks, and potential loss of control.
- Policymakers face an "evidence dilemma": act on limited proof or risk being unprepared.
- Inclusive global cooperation is essential, as risks and priorities vary across regions.

References

International Scientific Panel on AI Safety. (2024). *The International Scientific Report on the Safety of Advanced AI: Interim report* (Y. Bengio, Chair). https://yoshuabengio.org/2024/06/19/the-international-scientific-report-on-the-safety-of-advanced-ai/

UK Government. (2023). *AI Safety Summit 2023, Bletchley Park*. GOV.UK. https://www.gov.uk/government/topical-events/ai-safety-summit-2023

Isabella Wilkinson — Research Fellow, Digital Society Programme, Chatham House

Title: *Transparency and Credible, Coherent AI Governance*

Abstract: As countries, companies and other stakeholders seek to govern AI, transparency has emerged as a principle and practice, and as a prerequisite for effective governance. There is growing consensus about its meaning: for example, on aspects of model ('technical') transparency and what constitutes 'public' transparency. However, understandings vary across supranational, multilateral and national governance initiatives. This talk uses AI transparency as a lens for exploring how to overcome emerging issues – fragmentation and incoherence – in global AI governance. It considers the architectures, mechanisms and partnerships required to work towards credibility and coherence, and their durability, both as models advance and amid geopolitical rivalries.

Summary: In this talk, Isabella Wilkinson offered, from a think tank perspective, a detailed exploration of the concept of transparency, regarded as a prerequisite for global AI governance. A cornerstone of international approaches - whether within the OECD, AI summits, the European Union, or the United Nations - transparency has become an indispensable dimension of global AI governance. Yet, as she underlined, two major challenges remain: the lack of coherence and the lack of clarity surrounding this concept, both of which create challenges in terms of interoperability (between governance approaches) and enforceability. Her presentation was thus structured around a central question: What are the steps needed to promote coherence regarding transparency requirements, and why do they matter for effective AI governance?

Seeking to provide avenues for reflection on this question, Isabella Wilkinson oriented her presentation along two lines: first, by exploring the theoretical and practical contours of the concept of transparency and second, by presenting research findings on transparency and coherence.

She began by highlighting the complexity of transparency. Drawing on the literature (e.g. on information asymmetries and approaches to transparency as a virtue, relation and system) and practitioner approaches, she argued that in the AI context, transparency is understood on two levels: technical and public. The latter is aimed at democratizing the development and deployment of technology. It calls for a dynamic form of contextualization that reflects the needs and resources of non-expert actors while keeping pace with technological advances. On

this point, she left the audience with an open question: should the definition of transparency itself be continuously updated?

To illustrate her argument, Isabella Wilkinson moved from the general to the specific by focusing on how transparency has been defined in the European Union's approach to AI governance. She emphasized that transparency is not only referenced but also explicitly defined in the *EU AI Act* (Preamble, 27). The Code of Practice on General-Purpose AI elaborated on obligations for model providers on technical and public transparency. Isabella walked through the strength of this approach (e.g. by adopting a 'lifecycle' approach to transparency requirements (branching from upstream to model to downstream) and its shortcomings (e.g. a watered down definition of public transparency which would benefit from further clarity).

In the final part of her presentation, Isabella Wilkinson reflected on the pathways to greater coherence between diverse approaches to AI transparency. She stressed the importance of building infrastructures capable of bridging binding and non-binding frameworks, fostering exchanges between stakeholders and regulatory bodies, and drawing lessons from non-regulatory platforms that support the dissemination and socialization of norms and best practices. She further underscored the need for more social science research (e.g. surveying the meaning of transparency in different contexts) for scientific research on transparency indicators.

Ultimately, this presentation brought to light one of the major challenges of global AI governance: defining and operationalizing key concepts, such as transparency, in ways that ensure their clarity, interoperability, and enforceability. Advancing toward good AI governance requires addressing these issues head-on, as their implications extend far beyond transparency.

Key Takeaways

- Transparency has become a cornerstone of global AI governance but remains fragmented and inconsistently defined across jurisdictions.
- It operates on two levels: *technical* (algorithmic) transparency and *public* transparency for democratizing AI deployment.
- The EU AI Act provides a definition, but persistent gaps led to the drafting of the 2025 General-Purpose AI Code of Practice, which itself faces weaknesses.

- Achieving coherence requires infrastructures linking binding and non-binding frameworks, stronger stakeholder engagement, and better dissemination of best practices.
- A major challenge is operationalizing transparency so it is clear, interoperable, and enforceable, ensuring durability amid technological and geopolitical shifts.

References

Bommasani, R. et al. (2024). "The Foundation Model Transparency Index v1.1 May 2024", Center for Research on Foundation Model, https://crfm.stanford.edu/fmti/paper.pdf

Felzmann, H. et al. (2020). "Towards Transparency by Design for Artificial Intelligence", Science and Engineering Ethics, vol. 26, https://link.springer.com/article/10.1007/S11948-020-00276-4

European Commission. (2025). *General-Purpose AI Code of Practice*. Retrieved from https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai

European Council and European Parliament. (2024). *Artificial Intelligence Act.* Accessible : https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202401689

Prof. Catherine Régis — Université de Montréal; IVADO; CIFAR

Title: *The Creation of the UN Scientific Panel on AI: Implications for the Future of AI Governance*

Abstract: In September 2024, the United Nations General Assembly, through its Global Digital Compact, committed to establishing an independent International Scientific Panel on AI within the UN. In the interest of facilitating the UN formulation of this panel, various actors and organizations have submitted proposals. Following a period of deliberation, the General Assembly adopted a resolution in August 2025, formally initiating the establishment of the panel. While the precise structure, functioning, financing, and composition of the panel are yet to be delineated, the Resolution specifies that it will be a multidisciplinary, independent, and geographically diverse panel comprising 40 members. It is also understood that this initiative will result in the production of scientific synthesis and analysis of existing research on

opportunities, risks, and impacts related to AI. This will be achieved, in part, through the dissemination of one annual "policy-relevant" yet "non-prescriptive" summary report. In this presentation, an exploration will be conducted of the milestones of the Panel, the key normative tensions at stake in achieving the intended results, and the lessons that can be learned from previous experience in global governance.

Summary: Professor Catherine Régis examined one of the most recent chapters in global AI governance, namely the establishment of the UN Scientific Panel on AI. She places it at the heart of one of her central concerns: how to bridge science and policymaking, particularly in a context where science may serve as a strategic lever (an approach of smart power) to exert influence on policy in the absence of appetite for regulatory action. To explore this new yet still emerging UN initiative, Régis first situated it within the broader landscape of global AI governance, then outlined its currently known contours, and finally integrated it into a more holistic approach inspired by earlier experiences in global governance.

As she emphasized, the creation of the UN Scientific Panel on AI is far from the first international initiative in this domain. Rather, it follows on from several prior efforts aimed at the collective management of AI-related risks and the maximization of its benefits for all. Despite the well-known challenges facing multilateralism in regulating AI—namely the unprecedented pace of technological development and significant geopolitical tensions—the UN arena has given rise to the proposal for an Independent International Scientific Panel on AI within the UN (as part of the *Global Digital Compact*, 2024), culminating in its formal adoption on 26 August 2025 (General Assembly Resolution 79/235).

Still under construction, only two features are currently known regarding the panel: its objectives (to produce evidence-based scientific assessments synthesizing and analyzing existing research on AI's opportunities, risks, and impacts) and its structure (two co-chairs, one of whom must come from a "developing country," and forty members appointed by the General Assembly on the basis of expertise and inclusivity, with full disclosure of all conflicts of interest). Importantly, the panel is not tasked with generating new research but with synthesizing existing knowledge, notably in the form of an annual report that is "policy-relevant but non-prescriptive," an expression she finds particularly interesting. Finally, she noted that the panel's mandate explicitly excludes the military domain - an omission she deplored, given the international community's apparent "powerlessness" in this sphere.

Building on these elements, Catherine Régis observed that while it is still too early to determine what exactly can be expected from such an initiative, lessons may nonetheless be drawn from previous experiences in global governance, both within and beyond the field of AI, as well as

from existing academic research. To this end, she proposed examining the Intergovernmental Panel on Climate Change and the *International AI Safety Report* (2025), both of which demonstrate the capacity of the scientific community to influence global governance despite certain weaknesses. She underlined, for instance, that there may be a trade-off between scientific rigour and policy responsiveness, and likewise between timely assessment and inclusivity. It is then drawing on knowledge from academic research that Régis integrated into her reflections three areas of knowledge: the use of international norms (including scientific content as a vector of influence), the contemporary strategies of autocratic regimes in constructing a form of "hollow multilateralism," and the role of epistemic communities which are particularly relevant in contexts of uncertainty, such as that surrounding AI and its trajectory.

In conclusion, Régis argued that the panel holds considerable promise for global AI governance, but highlighted several remaining grey areas: the criteria of expertise applied in selecting its members, the management of conflicts of interest, the modalities for consulting the private sector and civil society, and the means by which the panel will engage with the political community - particularly with the Global Dialogue created alongside it. A new stone has thus been laid in the edifice of global AI governance, and its impact will only become clear once the panel is operational. In a context where multilateralism and international institutions are under heavy strain, the international community is in need of demonstrable successes; this panel may yet prove to be a success in the difficult endeavour of regulating AI on a global scale.

Key Takeaways

- The UN General Assembly formally adopted the creation of the International Scientific Panel on AI in August 2025.
- The panel will be multidisciplinary, independent, and geographically diverse, with 40 members and two co-chairs (one from a developing country).
- Its mandate is to synthesize existing research, not produce new studies, delivering one annual "policy-relevant but non-prescriptive" report.
- Lessons from past global governance bodies (e.g., IPCC, International AI Safety Report) highlight tensions between rigour, inclusivity, and timely responsiveness.
- Key unresolved challenges include membership criteria, conflict of interest management, engagement with civil society and industry, and the omission of military AI.

References

Intergovernmental Panel on Climate Change. (n.d.). *Assessment reports*. Retrieved from https://www.ipcc.ch/reports/

International Scientific Panel on AI Safety. (2024). *The International Scientific Report on the Safety of Advanced AI: Interim report* (Y. Bengio, Chair). Retrieved from https://yoshuabengio.org/2024/06/19/the-international-scientific-report-on-the-safety-of-advanced-ai/

United Nations. (2024). *Global Digital Compact*. United Nations. Retrieved from https://www.un.org/techenvoy/global-digital-compact

United Nations General Assembly. (2025, August 26). *Resolution 79/235: Establishment of an International Scientific Panel on Artificial Intelligence*. United Nations. Retrieved from https://digitallibrary.un.org/

Conclusion – Next Generation Perspectives

As future professionals, we would like to reflect on the key lessons we took away from the conference and how these insights can inform our role in shaping responsible AI governance. From students in law, political science, and computer science, our group brings together diverse fields that converge on pressing questions of regulation and governance.

The conference showed us how legal, technical, and geopolitical perspectives must come together to address challenges such as systemic risks, manipulated media, and global power imbalances. The clarity of the presentations, as well as the comparisons drawn between approaches in Europe, North America, South America, and Asia, underscored the importance of inclusive and comparative perspectives in tackling shared problems. We saw that regulatory and governance choices made today will directly shape the professional opportunities, civic responsibilities, and ethical frameworks within which our generation will operate.

Another key lesson was that AI governance is not limited to passing binding rules. It also requires sustained dialogue across disciplines and the creation of bridges with other fields that have faced global governance challenges. Despite the magnitude of the obstacles, the symposium stressed the value of keeping the conversation open and ensuring a plurality of voices in shaping solutions.

The event also made clear that every jurisdiction grapples with the same tension: how to balance innovation with regulation. Different perspectives, from ethicists to lawyers, suggested that new forms of governance will be required, combining binding rules with voluntary frameworks and blending concrete provisions with broader ethical principles. Finally, the multidisciplinarity of the discussions gave us confidence that the field is moving beyond vague debates toward a deeper and more precise understanding of AI's risks, from specific harms such as synthetic media to broader questions of international governance and existential threats. It was equally encouraging to see recognition of the geopolitical dimension of AI, particularly the concentration of power in its development, which is central to any risk assessment.

For us, the conclusion is clear: the governance of AI must remain multidimensional, inclusive, and globally informed. The choices made now will define the environment in which this and next generations will live and work, and it is our responsibility to carry forward this dialogue with accountability, equity, and interdisciplinary collaboration at its core.

Biographies of the Speakers

(listed in order of appearance)

Prof. Catherine Régis — Université de Montréal; IVADO; CIFAR

Catherine Régis is a Professor of Law at the University of Montreal (UdeM), an Associate Academic Member at Mila (Quebec AI Institute), the Co-Director of the Canadian Institute for AI Safety research program at CIFAR, and Director of International Policy and Social Innovation at IVADO. She holds both a Canada CIFAR AI Chair and a Chair in Scientific Diplomacy and Global AI Governance (Fonds de recherche du Québec), and is a Senior Research Associate at the University of Cambridge Jesus College's Intellectual Forum. From 2021 to 2023, she served as Associate Vice-President for Strategic Planning and Responsible Digital Innovation at UdeM and from 2021-2023 she was the Co-chair of the Working Group on Responsible AI for the Global Partnership on AI. Her work focuses on responsible AI governance and regulation at national and international levels, with an emphasis on human rights, equitable benefit sharing, and applications in health care and justice.

Prof. Angeliki Kerasidou — University of Oxford, Ethox Centre

Angeliki Kerasidou is a Senior Fellow in the Nuffield Department of Population Health at the Ethox Centre and a Research Fellow at the Wellcome Centre for Ethics and Humanities, University of Oxford. She studied theology and philosophy in Greece, Germany, and the UK, and received her DPhil from Oxford in 2009. Her research examines ethical issues raised by new technologies and socio-economic change in biomedical research and clinical practice. She is currently investigating the ethics of artificial intelligence in population health, focusing on accuracy, efficiency, and

relational and epistemic trust, and leads an international collaboration on empathetic health-care systems.

Prof. Célia Zolynski — Université Paris 1 Panthéon-Sorbonne; Observatoire de l'IA de Paris 1

Célia Zolynski is a Professor of Private Law at the Sorbonne Law School (Université Paris 1 Panthéon-Sorbonne) and Co-Director of the DreDIS Research Department at IRJS. Agrégée in Private Law and Criminal Sciences with a PhD from Université Panthéon-Assas, she directs the Master II in Law of Creation and Digital and co-directs the Master I in Digital Law. She coordinates the Paris 1 Observatory on Artificial Intelligence and serves on several national committees on digital ethics and rights. Her research focuses on digital law, intellectual property, fundamental rights, and AI governance.

Frédéric Tremblay — Director General, Deputy Ministry for American Relations, Economic Affairs and Strategic Intelligence, Québec Ministry of International Relations and La Francophonie

Frédéric Tremblay is the Director General of the Deputy Ministry for American Relations, Economic Affairs and Strategic Intelligence at the Québec Ministry of International Relations and La Francophonie. He previously served as Director of the Québec Office in Washington, Public and Government Affairs Counsellor at the Québec Delegation in Los Angeles, and International Relations Counsellor for North American educational affairs. He holds a Master's in Political Science and a Bachelor's in Communication and Politics from the Université de Montréal. His career spans communications, public affairs, and international relations.

Dr. Mario Rivero-Huguet — Head of Science and Innovation, British Consulate in Montreal

Mario Rivero-Huguet is Head of Science and Innovation at the British Consulate in Montreal, where he leads the UK Science and Innovation Network's activities in Québec and the Atlantic Provinces. He works to establish and strengthen partnerships in fields such as health sciences, aerospace, and clean technology between the UK and eastern Canada. He holds a Master's in Chemistry from the University of Leipzig and a PhD in Environmental Health from McGill University. He has also lectured at McGill University and worked with the Commission for Environmental Cooperation in North America.

Prof. Rebecca Williams — University of Oxford

Rebecca Williams is a Professor of Public and Criminal Law at the University of Oxford and a Fellow of Pembroke College. She studied law at Worcester College, Oxford, before completing a BCL and a PhD at the University of Birmingham. Her teaching focuses on criminal and public law, while her research spans EU and US comparative public law, unjust enrichment, and the intersection of law and computer science. Her current work explores how legal frameworks adapt to technological developments.

Prof. Pierre Larouche — Université de Montréal

Pierre Larouche is a Professor of Law and Innovation at the Université de Montréal's Faculty of Law, specializing in competition law, economic governance, and civil liability in both civil and common law traditions. He holds law degrees from McGill, Bonn, and Maastricht, and previously taught at Tilburg University, where he co-founded the Tilburg Law and Economics Center and launched the Global Law Program. He has also taught at the College of Europe and as a visiting professor at leading universities in North America, Europe, and Asia. His research has influenced European policy in electronic communications and competition law.

Prof. Melissa Hyesun Yoon — Hanyang University

Melissa Hyesun Yoon is a Professor at the Hanyang University School of Law in Seoul, where she has taught administrative law since 2012, and also teaches AI and the law in the School of Artificial Intelligence. Originally trained in biochemistry and physiology, she later pursued law, passed the bar in both the United States and Canada, and practised briefly in Canada and Korea. Her research focuses on administrative law, regulation, and policy in broadcast communications, data, AI, biopharmaceuticals, and nuclear energy. She is the author of several books and articles on AI governance, data justice, and regulatory policy.

Prof. Juan David Gutiérrez — Universidad de los Andes, School of Government

Juan David Gutiérrez is a Professor at the School of Government of Universidad de los Andes in Bogotá. He holds a PhD in Public Policy from the University of Oxford, an LLM in Law and Economics from the Universities of Bologna and Erasmus Rotterdam, and a Master's in Latin American Public Policy from Oxford, as well as a Law degree from Universidad Javeriana. His teaching and research focus on public policy, artificial intelligence, competition and regulation, and natural resource governance.

Prof. Benjamin Guedj — University College London; Inria

Benjamin Guedj is a Professor of Machine Learning and Foundational AI at University College London, Research Director at Inria, and a Turing Fellow at the Alan Turing Institute. He holds a PhD in Mathematics from Sorbonne Université and is Founder and Scientific Director of the Inria London Programme, a joint lab between France and the UK. His research focuses on theoretical machine learning, statistical learning theory, PAC-Bayes, and generalization in deep learning. He is also a Fellow of the ELLIS society and the Royal Statistical Society.

Prof. Christian Gagné — Université Laval; Institut intelligence et données

Christian Gagné is a Professor in the Department of Electrical and Computer Engineering at Université Laval, where he also directs the Institute Intelligence and Data (IID). He holds a Canada CIFAR AI Chair and is an Associate Member of Mila. His research focuses on machine learning and stochastic optimization, including deep neural networks, representation learning, meta-learning, and evolutionary algorithms. He applies these methods to domains such as computer vision, health, energy, and transportation.

Alexei Grinbaum — Research Director, CEA-Saclay

A philosopher and physicist specializing in quantum information theory, he has explored the ethical challenges of emerging technologies since 2003, including nanotechnology, artificial intelligence, and robotics. He is President of the CEA's Operational Committee on Digital Ethics, a member of the French National Digital Ethics Committee (CNPEN), and an expert for the European Commission. His work bridges fundamental science with the ethical implications of technological innovation.

The Honourable Judge Simon Ruel — Québec Court of Appeal

The Honourable Simon Ruel has served as a judge at the Québec Court of Appeal since 2017, after serving at the Superior Court of Québec from 2014 to 2017. He studied law at the Université de Montréal and also holds a Bachelor's in Biochemistry. Before his appointment, he practised mainly in public and administrative law, acted as counsel in several public inquiries, and taught law at the École du Barreau du Québec and Université Laval. He has held leadership roles within the Canadian Judicial Council and has contributed to international initiatives on anti-corruption and international criminal law.

The Honourable Benoît Moore — Québec Court of Appeal

The Honourable Benoît Moore has served as a judge at the Québec Court of Appeal since 2019, after serving at the Québec Superior Court from 2017 to 2019. He earned a Bachelor's and Master's in Law from the Université de Montréal and a D.E.A. in Civil Law from Université Paris 1 Panthéon-Sorbonne. Before his appointment, he was Professor of Law at the Université de Montréal, holding the Jean-Louis Baudouin Chair in Civil Law and serving as Interim Dean and Associate Vice-Rector. An author of leading works on civil liability and obligations, he has lectured internationally and is President of the Québec section of the Association Henri-Capitant and a member of the International Academy of Comparative Law.

Paul Gagnon — Partner, Technology and Artificial Intelligence Group, BCF

Paul Gagnon is a Partner and Co-Leader of the Technology and Artificial Intelligence group at BCF in Montréal. He holds an LL.M. in Intellectual Property and has over ten years of experience in technology, commercial, and intellectual property law. His practice focuses on AI governance, data architecture, and complex commercial negotiations, drawing on previous in-house experience in a leading AI company. He regularly advises clients from start-ups to large enterprises and is co-author of the pioneering Montréal Data License.

Misha Benjamin — Partner, Technology and Artificial Intelligence Group, BCF

Misha Benjamin is a Partner and Co-Leader of the Technology and Artificial Intelligence group at BCF in Montréal. With over ten years of experience, he advises companies of all sizes on technology, data, and AI applications, including transactions and large-scale deployments. He previously held senior roles in leading international companies and start-ups, focusing on software, data use, and regulatory issues. His practice emphasizes practical risk management and strategic support for clients operating in global and highly regulated sectors.

Benjamin Prud'homme — Vice President for Public Policy, Safety and Global Affairs, Mila

Benjamin Prud'homme is the Vice President for Public Policy, Safety, and Global Affairs at Mila. He is a legal expert engaged with the OECD.AI Network of Experts, the UN Advisory Network on AI, and UNESCO's AI Ethics Experts Without Borders. He co-leads international projects on diversity, equality, and advanced AI safety, including contributions to the Global Partnership on AI and the UNESCO–Mila report on AI governance blind spots. A lawyer by training, he serves on the boards of the Canadian Civil Liberties Association, the Observatoire québécois des inégalités, and Legal Aid Montréal.

Isabella Wilkinson — Chatham House, Digital Society Initiative

Isabella Wilkinson is a Research Fellow in the Digital Society Initiative at Chatham House, focusing on security and governance in cyberspace and technology. Her work examines international cyber governance, the online information environment, and advancing responsibility and inclusivity. She previously worked in Chatham House's International Security Programme and contributed to the Journal of Cyber Policy's editorial team. She holds an MA in Democracy and Governance from Georgetown University and a BSc in History and International Relations from the London School of Economics.

Biographies of the Students and IVADO Staff

(listed in alphabetical order)

Halima Bachir is an intern in Knowledge Mobilization at IVADO and a master's student in Public and International Affairs at Université de Montréal. She holds a bachelor's degree in International Relations and International Law from Université du Québec à Montréal (UQAM), where she first oriented her research interests toward digital governance and the regulation of AI.

Antoine Congost is a Knowledge mobilization advisor at IVADO. He holds a bachelor's degree in international studies and a master's degree in political science (Université de Montréal). Previously in charge of AI governance issues at Université de Montréal, he has developed several projects to disseminate and implement the Montreal Declaration on Responsible AI. Passionate about international governance issues, Antoine has also worked at the French Consulate in Montreal, the Secretariat of the Convention on Biological Diversity and the Délégation générale du Québec in Tokyo.

Gaëlle Foucault is a postdoctoral researcher and lecturer in international law at the Faculty of Law, Université de Montréal. Her research, funded by the IVADO Excellence Program, focuses on the global governance of AI. She holds a doctorate in international law from Université de Montréal (Canada), a master's degree in international law from Université Jean-Moulin, Lyon III (France) and a bachelor's degree in law from Université Panthéon-Assas, Paris II (France). She is also coordinator of the H-Pod and is affiliated with Mila–Quebec AI Institute and the Centre de recherche en droit public [Public Law Research Center] of Université de Montréal.

Emma Kondrup is currently a computer science student at McGill University, and an incoming Ph.D. student under Profs. Reihaneh Rabbany and Catherine Régis. Her research interests lie at the intersection of machine learning and socio-legal issues, both exploring AI applications for social good, and areas of tech law, with a focus on global AI governance and bioethics.

Clare Mulrooney is a visiting research intern at Mila–Quebec AI Institute, supervised by Professor Catherine Régis, and an incoming BCL candidate and William Asbrey scholar at St. Edmund Hall, Oxford. She completed her undergraduate degree in Law at Jesus College, Cambridge (2025), and the National University of Singapore. Her research interests lie in the intersection of law and technology, with a current emphasis on the criminalization of artificial image-based sexual abuse and intermediary liability of digital platforms.

Acknowledgments

We thank the convenors, Prof. Catherine Régis, Prof. Angeliki Kerasidou and Prof. Célia Zolynski, for their leadership and vision, as well as the institutional partners IVADO, Université de Montréal, the Maison française d'Oxford, the University of Oxford, the British Consulate-General in Montréal, the Délégation générale du Québec à Londres, and the Canada-CIFAR Chair in AI and Human Rights, for their essential support. We also extend our appreciation to the coordinators and editorial team, Antoine Congost, Gaëlle Foucault, Halima Bachir, Emma Kondrup, and Clare Mulrooney, for their contributions to this booklet, and to the eighteen participants from Quebec, the United Kingdom, and France, whose insights and engagement greatly enriched the workshop.