

Decoding Global AI Governance: A Computational Linguistic Analysis of National Regulations

Eryclis Silva

University of Illinois at Urbana-Champaign
ersilva2@illinois.edu

Abstract

Artificial Intelligence (AI) regulations addressing rights, usage limits, bias, and accountability mechanisms have been adopted by governments worldwide. My research examines global AI governance through a comparative analysis of regulatory documents from over 60 countries. Using mixed methods, including advanced computational linguistics techniques to support qualitative analysis, this work evaluates changes in AI regulation over time and similarities in policy priorities. Our findings reveal significant alignment with EU regulations, highlighting its influence on global governance, and indicate regulatory capture by industry in the Americas and decolonial dissent in Africa. This research contributes to the multi-modal understanding of AI governance, offers practical policy recommendations, and outlines future research directions, including ongoing data collection and exploration of public dialogue's impact on regulatory frameworks.

Introduction

Artificial Intelligence (AI) is experiencing exponential growth, significantly outpacing regulatory efforts to govern its development and application. This research explores the global governance of AI through a comparative analysis of regulatory documents from 64 countries, as well as the European Union (EU) and African Union (AU). The gap between governance and technology is not unique to AI, but it presents substantial challenges. Historical research indicates that technological change often exceeds the pace of policy interventions (Edgerton 2010; Jamison 1989), reflecting a persistent struggle to align emerging technologies with social norms and human rights (Wernick and Artyushina 2023; Bench-Capon and Modgil 2017). This disconnect is exacerbated by lawmakers' lack of technical expertise and digital literacy (Khan et al. 2023; Crowley et al. 2021), and the pervasive influence of industry regulatory capture. In this context, my research aims to address the complexities of AI governance, which can be understood as an assemblage of four interrelated modes: law and policy, technical design, social norms and practices, and market dynamics. Each of these modes evolves semi-independently (Bygstad and Hanseth 2010; Tasca and Piselli 2019), leading to unintended outcomes and interactions that complicate effective

governance (Tasca and Piselli 2019; Avila-Robinson and Miyazaki 2011). Building on past literature, I recognize several key factors influencing technology governance:

1. A historical bias towards U.S. interests in technology regulation (Caselli and Coleman 2006; Cavoukian 2014).
2. The collaborative power of frameworks like the EU's data protection efforts on global governance.
3. The undemocratic impact of industry self-regulation.

Research Objectives

The central question of my research is: How are different countries and regions approaching AI regulation, and what factors are influencing these approaches? To address this, I have established the following specific research objectives:

1. To examine the similarities and differences in global AI regulations.
2. To identify the key influencers and factors shaping emerging AI regulations.
3. To uncover potential biases in the perspectives represented in these regulations.

Through these objectives, I aim to contribute to a more comprehensive understanding of the current state of global AI governance. My research employs a mixed-methods approach, leveraging both qualitative and quantitative analyses to evaluate:

- Changes in proposed AI regulation over time.
- Topical and document similarities in policy priorities.
- The use of hedging and uncertainty language that may impact enforceability (Grattet and Jenness 2005; Pollach 2007).
- The interests reflected in policy parameters.

By addressing these objectives, my research will not only provide practical policy recommendations but also expand the multi-modal conceptualization of AI governance and present a research agenda for future studies in this critical field.

Methodology

We compiled a comprehensive dataset of 1065 documents, including proposed or passed legislation and strategy documents, covering 64 countries and multinational governance

bodies such as the EU and AU. This collection was sourced from the OECD database on AI policy, national policy databases, and news alerts on emerging AI regulations. We extracted full text from these documents using simple text extraction and optical character recognition (OCR) when necessary. Metadata related to the introduction, modification, or passage of policies, as well as their authors, was also collected. Where possible, we gathered full versioning history.

Our focus has been on formal documentation that reflects national AI policies or strategies, dedicated regulatory frameworks, or any AI-related regulations. We ensured comprehensive coverage by considering official documents in English, Spanish, and Portuguese. The data collection methodology varied by region, encompassing South America, North America, Australia, Asia, Europe, and Africa. In terms of data analysis, we employed a mixed-methods approach that combined qualitative and quantitative techniques to gain a nuanced understanding of AI governance.

Qualitative analysis involved structured coding of policy parameters using the open-source tool Taguette. This content analysis was oriented in theories of policy development (Heikkila and Cairney 2018), institutional analysis (Siddiki et al. 2022), and information governance (Sanfilippo and Liu 2023). We applied the established and validated integrated Governing Knowledge Commons and Contextual Integrity (GKC-CI) framework (Shvartzshnaider, Sanfilippo, and Apthorpe 2022) to examine the structure and strength of institutions as a hierarchy of strategies, norms, and rules related to data and information processes within the policies. This framework allowed us to identify specific data flows and transformations as articulated in the policy documents, focusing on elements such as data subjects, data senders, data recipients, data types, aims, conditions, consequences, and the language used. We also examined the governance of individual components and entire data flows.

Our quantitative analysis utilized advanced natural language processing (NLP) techniques, taking advantage from vectorization methods such as Term Frequency-Inverse Document Frequency (TF-IDF) and word embeddings from models like Word2Vec and BERT, capturing both the frequency and contextual meanings of terms within the documents. For topic modeling, we applied dimensionality reduction using Uniform Manifold Approximation and Projection (UMAP) and clustering through Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) to identify distinct topics within our dataset. After vectorizing the documents, we calculated similarity scores for each document pair using cosine similarity, resulting in a similarity matrix that quantifies the relationships between documents. We visualized these findings through heatmaps to illustrate document similarities and created 2D scatter plots to demonstrate how documents cluster by topic.

By integrating these qualitative and quantitative approaches, our analysis provided a detailed and comprehensive overview of the current landscape of global AI governance, revealing both shared characteristics and notable differences in policy strategies across various regions and governance frameworks.

Current Progress

Our research has identified two main approaches to AI regulation: “soft laws” and “hard laws”. National AI strategies, often aligned with the OECD AI Principles, serve as “soft laws”. These value-driven documents emphasize development, innovation, and global political interests, with provisions for research funding. In contrast, “hard laws” provide comprehensive frameworks for AI regulation, addressing themes such as development, ethical standards, and fundamental rights. Some legislative initiatives propose oversight bodies for AI advancements. While national strategies constitute the majority of collected policy documents, with 71% of institutional statements coded as strategies, 26% as norms, and 3% as rules, hard laws exhibit a different distribution (39% strategies, 47% norms, 14% rules). Tentativeness scores, measuring (un)certainly in enforcement, are higher for hard laws (0.56) than soft laws (0.31), indicating that hard laws have less clearly defined enforcement mechanisms than national strategies.

Moving forward, our analysis reveals significant alignment between AI regulations in various regions and the European Union (EU), reflecting shared priorities and influences. Many countries have adopted provisions from the EU’s AI Act even before its formal passage, demonstrating the EU’s substantial impact on global AI governance. In Africa, the alignment is particularly strong, with an 87% similarity score between the AI regulations of the EU and the African Union (AU). Nigeria and Kenya also show high alignment with the EU, with similarity scores of 85%. This indicates a significant influence of EU policies on African AI regulations. In North America, the AI regulations of the U.S. and Canada show an 83% similarity score, while the U.S. and EU align at 81%. The highest alignment is observed between Canada and the EU, with an 87% similarity score, highlighting the EU’s influence on North American AI policies. Comparative patterns in other regions show varied degrees of alignment. Australia’s AI regulations align with the EU at 74%, indicating a moderate influence. In Asia, Singapore exhibits an 88% similarity score with the EU, reflecting a high level of influence. East Asian countries also show significant alignment: China and Japan (85%), China and the EU (85%), Japan and the EU (82%), and Korea and the EU (83%). The moderate alignment between China and Korea (71%) suggests a more balanced, context-specific influence. These findings underscore the EU’s pivotal role in shaping AI governance frameworks worldwide, with varying degrees of influence across different regions.

Future Work

While our research provides valuable insights into global AI governance, it also opens up several avenues for further exploration. The comprehensive dataset we have compiled serves as a foundation for ongoing analysis and future studies. One important direction is to continuously update this dataset with new policy documents to track the evolving regulatory landscape. Additionally, future research should investigate the influence of public dialogue, including traditional and social media, on shaping regulatory aims and ap-

proaches. Understanding how public discourse impacts policy can provide deeper insights into the motivations behind regulatory decisions.

Moreover, exploring potential networked impacts among countries could reveal shared governance parameters that transcend geographical boundaries, highlighting the relationships between global AI regulations. Finally, examining the influence of industry on AI governance is crucial; future studies should compare linguistic similarities between global policies and texts from industry sources to uncover patterns of influence. These directions will contribute to a more comprehensive understanding of the complex ecosystem surrounding AI governance and the various factors that shape policy development worldwide.

References

- Avila-Robinson, A.; and Miyazaki, K. 2011. Conceptualization and operationalization of emerging technologies: A complementing approach. In *2011 Proceedings of PICMET'11: Technology Management in the Energy Smart World (PICMET)*, 1–12. IEEE.
- Bench-Capon, T.; and Modgil, S. 2017. Norms and value based reasoning: justifying compliance and violation. *Artificial Intelligence and Law*, 25: 29–64.
- Bygstad, B.; and Hanseth, O. 2010. IT Governance through Regulatory Modalities. Health Care Information Infrastructure and the “Blue Fox” Project. In *Scandinavian Information Systems Research: First Scandinavian Conference on Information Systems, SCIS 2010, Rebild, Denmark, August 20-22, 2010. Proceedings 1*, 50–64. Springer.
- Caselli, F.; and Coleman, W. J. 2006. The world technology frontier. *American Economic Review*, 96(3): 499–522.
- Cavoukian, A. 2014. Evolving FIPPs: proactive approaches to privacy, not privacy paternalism. In *Reforming European Data Protection Law*, 293–309. Springer.
- Crowley, D. M.; Scott, J. T.; Long, E. C.; Green, L.; Israel, A.; Supplee, L.; Jordan, E.; Oliver, K.; Guillot-Wright, S.; Gay, B.; et al. 2021. Lawmakers’ use of scientific evidence can be improved. *Proceedings of the National Academy of Sciences*, 118(9): e2012955118.
- Edgerton, D. 2010. Innovation, technology, or history: What is the historiography of technology about? *Technology and Culture*, 51(3): 680–697.
- Grattet, R.; and Jenness, V. 2005. The reconstitution of law in local settings: Agency discretion, ambiguity, and a surplus of law in the policing of hate crime. *Law & Society Review*, 39(4): 893–942.
- Heikkila, T.; and Cairney, P. 2018. Comparison of theories of the policy process. In *Theories of the policy process*, 301–327. Routledge.
- Jamison, A. 1989. Technology’s theorists: conceptions of innovation in relation to science and technology policy. *Technology and Culture*, 30(3): 505–533.
- Khan, A. A.; Akbar, M. A.; Fahmideh, M.; Liang, P.; Waseem, M.; Ahmad, A.; Niazi, M.; and Abrahamsson, P. 2023. AI ethics: an empirical study on the views of practitioners and lawmakers. *IEEE Transactions on Computational Social Systems*, 10(6): 2971–2984.
- Pollach, I. 2007. What’s wrong with online privacy policies? *Communications of the ACM*, 50(9): 103–108.
- Sanfilippo, M. R.; and Liu, C. 2023. Governing Privacy as Contexts Overlap during Crisis. In *The Usage and Impact of ICTs during the Covid-19 Pandemic*, 46–64. Routledge.
- Shvartzshnaider, Y.; Sanfilippo, M. R.; and Apthorpe, N. 2022. GKC-CI: A unifying framework for contextual norms and information governance. *Journal of the Association for Information Science and Technology*, 73(9): 1297–1313.
- Siddiki, S.; Heikkila, T.; Weible, C. M.; Pacheco-Vega, R.; Carter, D.; Curley, C.; Deslatte, A.; and Bennett, A. 2022. Institutional analysis with the institutional grammar. *Policy Studies Journal*, 50(2): 315–339.
- Tasca, P.; and Piselli, R. 2019. The blockchain paradox. In *Regulating Blockchain: Techno-Social and Legal Challenges*, 27–42. Oxford University Press.
- Wernick, A.; and Artyushina, A. 2023. Future-proofing the city: A human rights-based approach to governing algorithmic, biometric and smart city technologies. *Internet policy review*, 12(1): 1–26.