Chapter 1: AI Policy Diffusion: Mapping Text Similarity and Regulation Networks Worldwide

Abstract

AI presents enormous opportunities but also poses significant risks. In response, national governments and multilateral organizations have adopted a range of laws, regulations, and initiatives to address concerns related to AI. What explains the explosive growth of AI policies regarding ethics and human rights? Why do some countries align while others diverge in their policy objectives? This chapter examines both the timing of AI policy adoption and the similarity of language used in national initiatives. I argue that shared IO memberships and AI adoption rates are associated with greater convergence in policy language. To analyze these patterns, I leverage AI policy objectives as indicators of regulatory priorities and apply a keyword-based approach using the OECD dataset of nearly 900 national policies. Through topic modeling and network analysis, I capture descriptive trends distinguishing social protection from economic objectives. I also conduct dyadic regression analyses on the frequency of specific keywords within the ethics and human rights category — namely, ethics, privacy, non-discrimination, transparency, accountability, and safety. The findings suggest that AI ethical and human rights norms have diffused globally within a short period. Regarding the correlates of specific keywords, shared IO membership and the use of AI in law enforcement consistently correlate with greater textual similarity across all categories. This chapter maps the evolving landscape of AI governance through policy language, offering a detailed account of the mechanisms behind policy diffusion and policy alignment across countries.

1.1 Introduction

The OECD defines AI as a machine-based system. With some objectives, it receives inputs and generates outputs such as predictions, content, recommendations, and decisions, with varying levels of autonomy and adaptiveness.[1] Because AI systems can impact physical and virtual environments at scale, the potential benefits and risks they pose to social and economic life are quantitatively different from those of other emerging technologies. While AI has existed for some time, only recently have governments moved to regulate the technology. Since the release of ChatGPT in November 2022, the discourse around AI regulation has intensified further. What explains the rapid increase in AI policies regarding privacy, non-discrimination, and transparency within a few years? How do countries converge or diverge in their AI policy objectives?

National regulators around the world face the delicate balance between human rights and the development and use of AI. As countries have taken steps to mitigate the risks, various regulatory models have emerged in AI governance with different intentions and priorities. One area of tension is between economic and social objectives. AI is expected to transform the economy while causing significant disruptions. For example, the EU's emphasis on fundamental rights contrasts with the US's innovation-driven approach.[2] U.S. regulators have focused on leveraging existing regulatory frameworks and developing voluntary standards. However, the EU prefers an ex-ante regulatory approach and has proposed a comprehensive framework to regulate AI across sectors. As AI superpowers compete for global influence, examining how the rest of the countries approach AI regulation offers valuable insights. Another area of competing demands is social protection and security. Authoritarian countries have adopted a coercion-based approach to enhance state capacity and public safety. In democracies, divergences have also arisen regarding the form and frequency of rights protection.

Even within AI ethics and human rights policies, there are a range of principles and priorities. AI poses risks to personal data protection and privacy. When algorithms are used to profile, it raises concerns about discrimination. Transparency and responsible disclosure are also commonly discussed in AI ethical principles, so stakeholders are aware of their interactions with AI systems

[1] See https://oecd.ai/en/ai-principles

[2] See https://www.csis.org/analysis/whats-ahead-cooperative-regulatory-agenda-artificial-intelligence

and provided with relevant information to understand the outputs. As AI systems may cause harm, what actors should be accountable for the proper functioning is a highly contested question. Additionally, as incidents related to cyberattacks and critical infrastructure have increased, regulators have moved to prevent misuse and adverse outcomes and to ensure the robustness and security of AI systems.

To examine the variation in AI policy objectives, I focus on the roles of IO networks and AI adoption rates. Regulation in new policy domains often spreads through IO networks, as several IOs have taken the initiative to set agendas and facilitate information exchange on ethical and human rights norms. Meanwhile, the widespread adoption of AI creates regulatory demands to address the harms and risks associated with the technology. In democracies especially, higher adoption rates of AI systems in government agencies and law enforcement tend to generate greater public pressure for accountability.

To investigate the recent explosive growth in AI regulation, I examine the extent to which these regulatory priorities are addressed in policy objectives. The analysis leverages the OECD dataset of national AI policies, which covers nearly 900 AI policies in 70 countries and the EU. Specifically, I conduct keyword-based text analysis of AI policy objectives to identify temporal patterns. While AI policies on ethics and human rights are often perceived as vague and aspirational, little is known about the specific content discussed in these documents. Policy objectives, in particular, present noteworthy information about the motivations behind regulation, the values they aim to uphold, and the harms they seek to prevent.

In this chapter, I examine the timing of adoption, the similarity of language, and the structure of AI regulation networks. The analysis is divided into two sections. First, I employ a keyword-based approach to reveal descriptive patterns in AI regulation. Using topic modeling, I identified temporal trends across three themes: ethics and human rights, market and innovation, and policy and regulation. To assess the text similarity between countries, I present network models based on the overlap of ethics and human rights keywords over time. Second, I conduct dyad-level regression analyses to explore the correlates of specific keyword usage — focusing on ethics, privacy, fairness, transparency, accountability, and safety. Finally, I present qualitative evidence of concrete policies that reflect these principles.

The descriptive patterns indicate that discussions around ethics and human rights increased between 2017 and 2021, while the emphasis on market and economic concerns gradually declined. The EU and several regional powers have played a crucial role in setting the agenda within this AI policy domain. Among EU countries, France and Germany have acted early and consistently, alongside others such as the Netherlands, Denmark, and Sweden. A number of EU countries act as followers, shaped by the EU's unified regulatory framework. At the dyad level, shared IO membership and the presence of AI applications in law enforcement are consistently associated with higher similarity across all six categories. In particular, greater similarity in AI capacity between the two countries is linked to more frequent discussions of privacy, transparency, and security. Similar profiles in ethnic fractionalization correspond with greater attention to fairness, bias, and discrimination, while similarity in corruption levels is associated with more frequent discussions of transparency and accountability.

In summary, this paper represents a pioneering effort and rich empirical patterns to systematically examine both aspirational initiatives and binding legislation through analysis of AI policy objectives. It contributes to the broader literature on policy diffusion by offering a keyword-based text analysis of regulatory priorities across countries. This study lays important groundwork for future research in a rapidly evolving and promising area of inquiry.

1.2 Background: AI governance and human rights protection

Given the broad scope of AI applications, AI regulation impacts a variety of actors, including private companies, research institutes, government entities, and civil society groups. AI policies also span different sectoral orientations, ranging from technology and education to infrastructure and social welfare. Additionally, AI policies take various forms: some adopt ethical frameworks to guide technological developments, while others aim to pass binding rules and establish new regulatory agencies.

While automated systems have significantly improved service efficiency and access to opportunities, concerns about AI's negative impacts persist. Excessive data collection and behavioral targeting by large tech firms infringe on data privacy. Both biased data and flawed algorithms can generate discriminatory outputs. With their increased use in judicial systems, a lack of transparency

and intrusive surveillance raise concerns about safeguarding civil rights and liberties.[3]

As AI systems become more prevalent, numerous government agencies, private firms, research institutes, and NGOs have published AI ethics guidelines. A few international principles were proposed and adopted as benchmarks. Notable examples include the EU's ethics guidelines for trustworthy AI (2018)[4], OECD AI principles (2019)[5], and UNESCO's recommendation on the ethics of artificial intelligence (2021)[6]. In the analysis of 84 AI ethical guidelines since 2016, the study reveals emerging convergence around five ethical principles, including transparency, justice and fairness, non-maleficence, responsibility, and privacy (Jobin, Ienca and Vayena 2019). Human rights framework is a common theme in many AI ethical guidelines (Fukuda-Parr and Gibbons 2021). The UN High Commissioner on Human Rights has warned that using AI systems for forecasting and profiling could negatively impact the right to privacy, a fair trial, freedom from arbitrary arrest and detention, and the right to life.[7]

In addition to AI ethics, I focus on the rights to privacy and non-discrimination. Protecting digital privacy is essential for effective AI governance. AI systems are characterized by their autonomy and reliance on personal data (European Commission 2021). The processing of extensive personal data in developing AI models has raised concerns about its impact on privacy rights and safety.[8] The EU has been a global leader in data protection and digital markets, advocating for a regulatory framework that respects human rights values. The EU's approach to consumer privacy, culminating in the 2018 General Data Protection Regulation (GDPR), has influenced how firms worldwide seek user consent to share, access, and delete data. Many countries are adopting similar regulations at the national level (Engler 2022), and, according to the UNCTAD (2021), 137 of 194 countries have enacted legislation to protect data privacy. Large tech firms also incorporated privacy in voluntary AI guidelines. Google advocated for privacy design principles [9], and IBM listed

[3]UNESCO. "Artificial Intelligence: examples of ethical dilemma". https://www.unesco.org/en/artificial-intelligence/recommendation-ethics/cases

[4]See https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

[5]See https://oecd.ai/en/ai-principles

[6]See https://www.unesco.org/en/artificial-intelligence/recommendation-ethics

[7]Scott Neuman. 'The U.N. warns that AI can pose a threat to human rights.' *NPR*. September 16, 2021.

[8]Cristina Criddle, "Europe's Privacy Watchdog Probes Google's Use of Data for AI Model," *Financial Times*, September 11, 2024, https://www.ft.com/content/9397423a-1737-4ae2-8d8a-3e4301f2c0a5

[9]See https://ai.google/static/documents/EN-AI-Principles.pdf

privacy as one of the five foundational properties for AI ethics.[10]

More recent discussions have focused on AI's potential impacts on people's safety and welfare, as algorithms can perpetuate existing biases and discrimination.[11] The UN Special Rapporteur on extreme poverty has warned that discriminatory AI design in digital welfare programs could affect access to critical resources and opportunities among the most vulnerable populations. [12] Amnesty International research shows that invasive surveillance systems exacerbate racist and discriminatory law enforcement against minorities, as well as abuse against migrants, refugees, and asylum seekers at borders in EU member states.[13] An analysis by U.S. NIST also reveals that facial recognition systems exhibited higher false positives for African and Asian individuals and lower rates for Eastern Europeans (National Institute of Standards and Technology 2019). To protect equal rights, some countries rely on existing laws. For example, the UK's Centre for Data Ethics and Innovation reviewed the application of the Equality Act to automated decision-making and released publications on algorithmic bias. Countries also published numerous reports to assess AI's social impacts on equality and adopt voluntary guidelines for risk mitigation. For instance, New Zealand released an algorithm assessment report in 2018 and made recommendations to improve transparency and accountability in the government's use of AI. [14] The Dutch government issued a 'non-discrimination by design' guideline for government organizations and companies in 2021 and established an algorithmic oversight body in 2023.

While some policies focus specifically on privacy or non-discrimination, the EU's Artificial Intelligence Act has taken a comprehensive approach, becoming the first to propose a legal framework to regulate AI systems. The EU Parliament adopted the AI Act in March 2024, and the Council followed with its approval in May 2024. Depending on the types of AI systems, the obligations will be fully applicable from six months to 24 months after entry into force.[15] A risk-based

[10]See https://www.ibm.com/impact/ai-ethics

[11]Jacob Poushter, Moira Fagan, and Sneha Gubbala. 'Climate Change Remains Top Global Threat Across 19-Country Survey.' *Pew Research Center*, August 31, 2022.

[12]United Nations Human Rights Office of the High Commissioner, "World Stumbling Into Zombie Digital Welfare Dystopia, Warns UN Human Rights Expert," October 2019

[13]https://www.amnesty.org/en/latest/news/2023/06/eu-ai-act-at-risk-as-european-parliament-may-legitimize-abusive-technologies/

[14]See https://data.govt.nz/docs/algorithm-assessment-report/

[15]https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-

approach mandates that AI systems classified to have an 'unacceptable risk,' such as government social scoring, should be prohibited. High-risk sectors will face more stringent obligations and oversight for AI systems that pose significant risks to people's safety or fundamental rights.

1.3 AI regulatory challenges and the diffusion of norms and policies

AI regulation is a timely and important issue area in examining mechanisms of norm and policy diffusion. The policy-making of a particular country is no longer limited to the domestic sphere but is increasingly influenced by transnational factors. With higher adoption rates of AI in government agencies and law enforcement, countries may face a similar set of issues, such as privacy violations or risks from synthetic content. These common challenges create incentives for cooperation and facilitate mutual policy adjustment (Keohane 2005). Shared problems may also encourage countries to learn from one another and generate demand for policy solutions. Moreover, understanding the complexity of AI and evaluating its economic and social impacts presents significant challenges, especially with new AI models being released every few months. Drafting technical standards or designing toolkits requires strong public-private partnerships and inputs from leading AI experts. Countries often lack information on how to regulate AI and may turn to others for knowledge and expertise. This challenge is particularly acute in countries with resource constraints, where bureaucracies have limited technical expertise and lack access to information held by private AI firms.

The diffusion literature identifies various mechanisms of policy interdependence, one of which is competition (Dobbin, Simmons and Garrett 2007). In the context of AI governance, aligning with international norms can foster public trust and attract investment and talent. A well-designed regulatory framework not only enhances trust but can also boost economic performance and improve national competitiveness. As a result, regional powers and economically advanced nations often take the lead in setting international rules and pursuing global leadership. For instance, the UK's AI white paper asserts that a pro-innovation regulatory framework can drive economic

artificial-intelligence

growth and solidify the UK's position as a global leader in AI.[16] Public trust is a common objective in AI regulation frameworks, as seen in Canada's Artificial Intelligence and Data Act (2022),[17] and Singapore's principles for the use of AI in the financial sector.[18]

Political learning occurs as countries observe others, especially early adopters, and evaluate the success or failure of their policy outcomes (Shipan and Volden 2008). High levels of bilateral interactions such as trade and immigration promote the diffusion of human rights standards (Greenhill, Mosley and Prakash 2009). In addition, states are not solely motivated by rationality and material interests but also constrained by the logic of appropriateness and norms (March and Olsen 1998). Countries are more likely to emulate or learn from their sociocultural peers (Simmons and Elkins 2004). Norm entrepreneurs such as IOs and transnational networks also facilitate the spread of norms as countries conform to legitimate AI ethical guidelines (Finnemore and Sikkink 1998). Aspirational principles in ethics and human rights help establish regulatory legitimacy in emerging issue areas. Otherwise, unregulated risks and harm would undermine trust and hinder further development. The transnational influence can facilitate the political mobilization of civil society at the domestic level (Finnemore and Sikkink 1998). Therefore, countries with broader IO membership tend to adopt similar policies over time.

In summary, diffusion mechanisms in AI governance emerge from a range of factors. From the demand side, countries with similar rates of AI adoption may encounter comparable problems and actively seek policy guidance from others. On the supply side, IO networks help disseminate AI ethical guidelines and policy frameworks, often invoking shared values when promoting new rules on AI. In particular, regional powers tend to leverage their alliances and international institutions to reinforce their global leadership and influence. For more reactive countries with limited technical expertise, IOs can help address informational gaps in AI regulation.

H1: Higher AI adoption rates are associated with convergence in AI policy discourse related to ethics and human rights.

[16]See https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach

[17]See https://ised-isde.canada.ca/site/innovation-better-canada/en/canadas-digital-charter-trust-digital-world

[18]See https://www.mas.gov.sg/publications/monographs-or-information-paper/2018/feat

H2: Shared IO membership is associated with greater similarity in AI policy discourse related to ethics and human rights.

1.4 Data and methods

1.4.1 Introducing the dataset

This project leverages the OECD dataset of national AI policies, compiled by the OECD and self-reported by 70 countries and the EU. The dataset covers 882 AI policies, including variables such as country, year, policy name, description, objectives, budget range, and policy areas. Presented in official English translation, this dataset enables cross-national comparison and in-depth analysis of AI policy development. A table of countries and their policy counts is provided in Appendix A.1.

In this chapter, I primarily rely on the **objective** variable from the OECD dataset. The objective for each policy can range from one sentence to multiple paragraphs and serves as a valuable proxy for understanding a country's AI regulatory priorities. In a new and emerging policy domain, there are limited available sources for analyzing the intent of government officials and systematically comparing different types of standards and regulations across countries. The objective is a good indicator of interests and may provide useful information as countries only recently moved to this policy area. For example, the objective for a 2023 executive order in the U.S. notes that the goal is to prevent and remedy discrimination, especially "by protecting the public from algorithmic discrimination." Similarly, one of the objectives of Canada's directive on automated decision-making is that the directive "provides a risk-based approach to ensuring the transparency, accountability, legality, and fairness of automated decisions that affect Canadians." The OECD dataset also includes a similar variable named 'description.' During data cleaning, I transferred relevant text from "description" into "objective" when it aligned with policy objectives.

To explore descriptive trends in AI policy objectives, I use keyword-based topic modeling to calculate the proportion of economic, social, and regulatory keywords over time. I also employ temporal network graphs to measure text similarity related to ethics and human rights across countries. In addition, I conduct dyadic regression analysis on the frequency of specific keywords, including ethics, privacy, non-discrimination, transparency, accountability, and security. Lastly, I compile a table of AI policies that address these specific principles to provide supporting qualitative

evidence.

I include EU policies in the analysis for several reasons. First, EU laws are legally binding for its member states. Its informal rules, such as white papers and guidelines, exert significant influence both within Europe and worldwide. Second, this study primarily employs text similarity measures and network analysis, which utilize a matrix structure to capture relationships between pairs of units. Given the data structure, including the EU in the analysis is both methodologically and substantively appropriate compared to a regression-based approach. However, the dataset's self-reported nature may lead to inconsistencies across countries. Depending on the capacities of the national bureaucracies responsible for reporting, the dataset may not provide a complete picture of the policies. Nevertheless, as a preliminary text analysis, the inclusion of detailed policy objectives offers a useful snapshot into cross-national priorities in AI governance.

1.4.2 Research design: text analysis of policy objectives

AI policies cover a variety of subjects, including ethics, privacy, bias, transparency, innovation, development, etc. AI technologies can significantly increase productivity and yield economic benefits yet undermine privacy rights and equal protection. Therefore, I use a keyword-based approach to capture descriptive trends in subjects concerning AI policies. By generating keyword usage, I identify patterns in the timing of adoption and the similarity of language in AI governance.

To capture variation in regulatory goals, I created three comprehensive keyword topics: ethics & human rights, market & innovation, and policy & regulation. Economic and social objectives are common themes in AI governance. I also include a category for policy and regulation, which can reflect temporal trends in whether a policy is aspirational or incorporates a concrete regulatory framework. The topics of ethics and human rights include trust, security, risk, transparency, human rights, democracy, and privacy. The group on market and innovation emphasizes innovation, economy, market, technology, businesses, and workforce. Policy and regulation topics include terms relevant to the legal and regulatory framework, such as policy and planning, legislative, standardization, compliance, investigation, and prevention (See Appendix A.1 for a complete list of keywords).

First, I use keyword-assisted topic modeling, keyATM, which uses keyword lists as priors

to inform the topic models. The model can generate estimates of keyword proportions by topic. Proportion (%) is defined as the frequency of the keyword divided by the total length of documents (Eshima, Imai and Sasaki 2023). Table 1.1 presents the top 15 words by proportion for each topic, highlighting key themes in discussions related to AI. In the 'Ethics & Human Rights' category, the most frequent word is 'ethical,' often used in contexts such as ethical guidelines, norms, governance, and standards. Additionally, terms related to society appear frequently. Regulators tend to emphasize the social impact of AI, stating in policy objectives that AI systems should benefit people and society. Other notable keywords in this category include 'security,' 'responsible,' and 'principles,' which underscore normative and ethical considerations, while 'protection' highlights concerns related to social protection.

Table 1.1: Top keywords and their proportions (%) by topic

Ethics & Human Rights	**Market & Innovation**	**Policy & Regulation**
ethical (0.51)	development (1.04)	framework (0.29)
society (0.28)	research (0.89)	policy (0.28)
human (0.27)	innovation (0.65)	governance (0.22)
principles (0.27)	technologies (0.62)	legal (0.22)
rights (0.24)	promote (0.48)	regulatory (0.18)
protection (0.23)	sector (0.45)	standards (0.17)
challenges (0.22)	services (0.44)	law (0.13)
security (0.22)	new (0.39)	stakeholders (0.13)
privacy (0.19)	industry (0.32)	regulation (0.11)
responsible (0.18)	economy (0.29)	guidance (0.09)

The dominant keywords in the 'Market & Innovation' category are "development," "research," and "innovation." These terms suggest that economic objectives are primarily focused on advancing AI technologies and fostering innovation within various industries. The emphasis is on promoting economic growth and AI adoption across sectors and services. For the 'Policy & Regulation' category, discussions frequently center around 'regulatory frameworks,' 'governance' structures, 'legal' obligations, and adherence to the rule of 'law.' Safety and standardization are

recurring themes as countries actively work to issue guidance on addressing risks and harm associated with AI. These keywords highlight the growing need for regulatory action to ensure AI technologies are used responsibly.

Second, I employ network analysis to examine text similarities in ethics and human rights policies from 2017 to 2021. The node size represents the count of unique keywords for each country, while node colors differentiate high and low levels of IO membership. The distances between countries are measured based on the number of overlapping related to ethics and human rights.

Third, I conduct dyad-level regression to examine the country's convergence or divergence in AI regulatory priorities. Topic modeling and network analysis examine the general use of keywords within the ethics and human rights category, and dyad regression seeks to understand the variation associated with each specific AI principle. Keywords tied to individual principles, such as privacy, fairness, and transparency, offer a more revealing and precise estimate of policy goals. The dependent variables are the frequency of overlapping keywords in policy objectives, measured at the dyadic level. For instance, if one country mentioned privacy two times, and another country mentioned privacy three times, then the overlapping frequency for the word 'privacy' would be two. The total frequency would be the sum of the frequency of individual keywords within that category. I include six keyword variables:

- Ethics: ethics, ethical
- Privacy: privacy, data protection
- Fairness: non-discrimination, discriminatory, fairness, fair, equal, unequal, unbiased, bias, biased, marginalized
- Transparency: transparency, transparent, explainability
- Security: secure, security, safe, safety
- Accountability: accountability, liability

To explain the variation in keyword similarity and frequency, I include several dyad-level independent variables. **Shared IO membership** is the number of overlapping memberships in

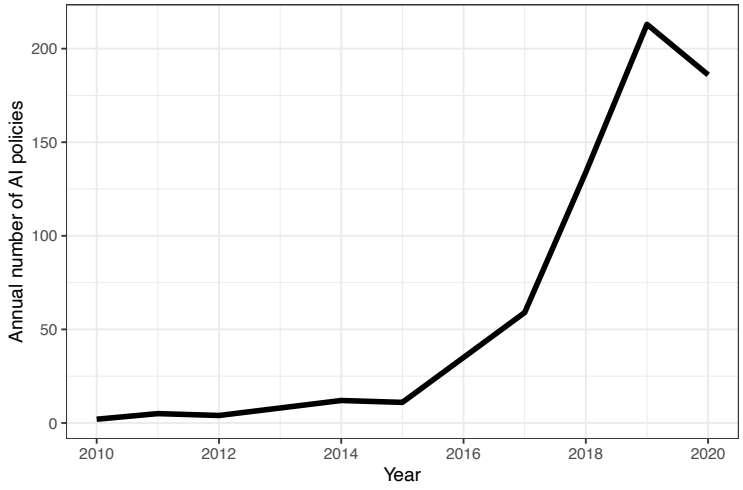
key IOs, including the EU, OECD, G7, G20, and GPAI. This variable captures socialization and institutional embeddedness, which may facilitate convergence in AI policy framing. **Smart policing** is a binary indicator denoting whether both countries have deployed AI in law enforcement (Carnegie Endowment for International Peace 2019).

Regarding control variables, I include **AI capacity**, measuring the absolute difference in log AI capacity scores between country pairs. The **UN voting** is the absolute difference in UN General Assembly ideal point estimates in 2024 from (Bailey, Strezhnev and Voeten 2017). This control accounts for broader foreign policy alignment. **Regulatory quality** measures the absolute difference in perceived regulatory quality derived from the World Bank's Worldwide Governance Indicators (Kaufmann and Kraay 2024). Additionally, I use V-dem liberal democracy (Coppedge et al. 2023) and log GDP per capita (The World Bank 2024) to capture regime type and economic strength. In particular, I include dyadic differences in the Historical Index of Ethnic Fractionalization (Dražanová 2020) as a predictor of fairness-related language based on the expectation that higher levels of ethnic diversity may lead to greater public awareness and demand for non-discrimination rights. For transparency and accountability, I use dyadic differences in corruption scores from V-Dem (Coppedge et al. 2023), as variation in perceived corruption may influence how governments frame issues of oversight and responsibility.

1.5 Results and discussion

Figure 1.1 illustrates the trends in AI policy based on the OECD dataset. AI policies have been steadily increasing since 2010, accompanied by major breakthroughs in AI technologies. The annual number of AI policies has grown exponentially since 2015, with many economically advanced countries increasing R&D investments in AI ecosystems and proposing national AI strategies to realize their benefits. Common AI policy initiatives include research centers and innovation hubs, as well as fostering AI-related capabilities such as genomics, robotics, computing infrastructure, and big data.

Figure 1.1: Annual number of AI policies (2010-2020)



1.5.1 Topic modeling and text similarity

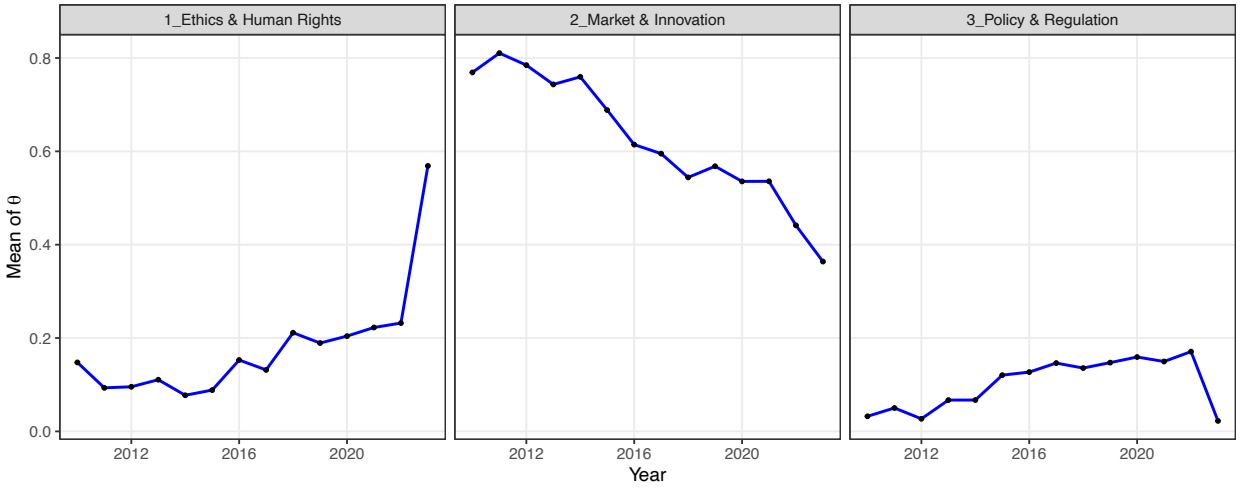
The text analysis of policy objectives highlights a shift in regulatory priorities since 2010. Figure 1.2 presents the estimated distribution of keyword proportions across three topics: ethics and human rights, market and innovation, and policy and regulation. This high proportion reflects the emphasis on enhancing AI capabilities as key components in national AI strategies and policies. During this period, many AI initiatives encouraged investment in AI R&D and data access, allocated funds for computing resources and infrastructure, and facilitated the commercialization of AI, particularly for SMEs and start-ups (Galindo, Perset and Sheeka 2021).

Notably, the innovation and market category exhibits temporal trends opposite to those of human rights and social protection. Since 2010, keywords associated with the market and innovation have declined, while keywords related to ethics and human rights have increased. As the use of AI systems becomes more prevalent, their adverse impacts are gaining greater attention from regulators. Studies have indicated an emerging consensus on AI ethical standards across countries. However, they also critique that countries often use human rights as a rhetorical device rather than establishing enforceable standards and accountability (Fukuda-Parr and Gibbons 2021).

Similarly, keywords related to policy and regulation have shown a steady increase. In the early 2010s, policy objectives were likely more vague and aspirational, but over time, they evolved

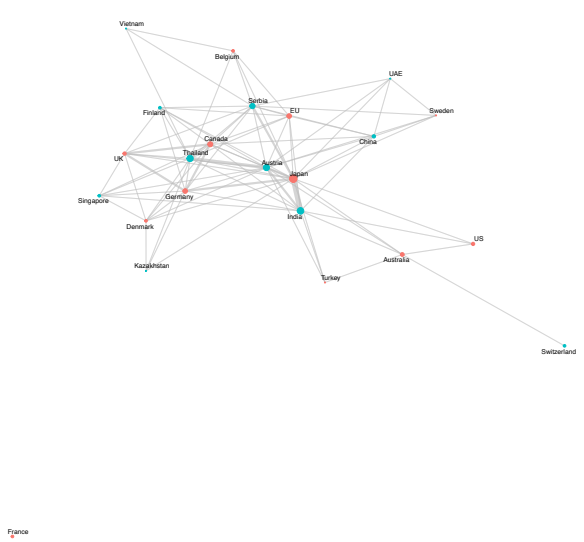
to become more concrete, incorporating implementation procedures and compliance mechanisms. The sharp rise and subsequent dip toward the end of the time series may be attributed to the limited availability of policies in 2023, which could skew the results.

Figure 1.2: Keyword proportion by topic: all AI policies (2010-2023)

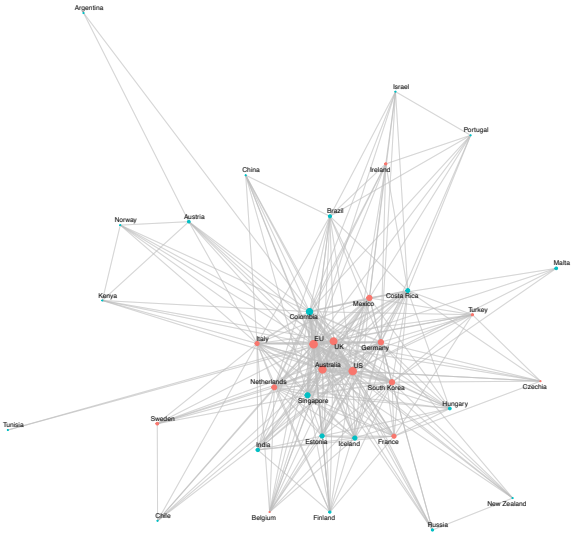


What is driving the rise in ethics and human rights keywords over time? To analyze the convergence of AI ethics and human rights policies, I construct a network based on shared keywords between countries over time. Figure 1.3 displays the text similarity on ethics and human rights from 2017 to 2020. To better illustrate the explosive growth in AI policy adoption, four figures are combined on the same page, and the full-sized versions can refer to Figures A.1, A.2, A.3, and A.4 in the Appendix. As an alternative measure to topic modeling, I manually calculated the unique keyword count related to ethics and human rights. Compared to keyword proportions, this measure highlights the comprehensiveness and technical depth of the regulatory framework and is less skewed by the length of the policy text. For example, many countries include only vague and aspirational principles in their policy objectives, using terms such as ethical, trustworthy, and responsible. In contrast, policies that provide clear guidance for businesses and organizations often incorporate more technical terms such as explainability, controllability, or verifiability.

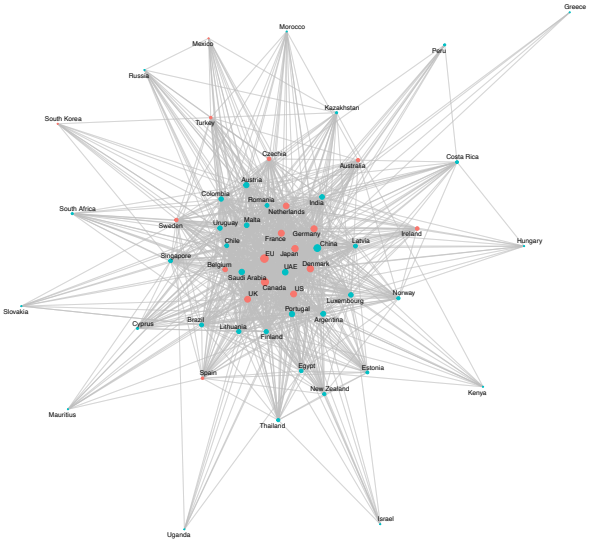
Figure 1.3: Text similarity on ethics and human rights



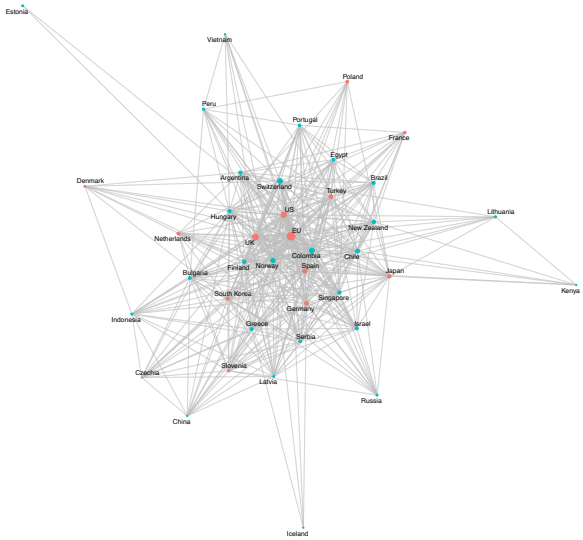
(a) 2017



(b) 2018

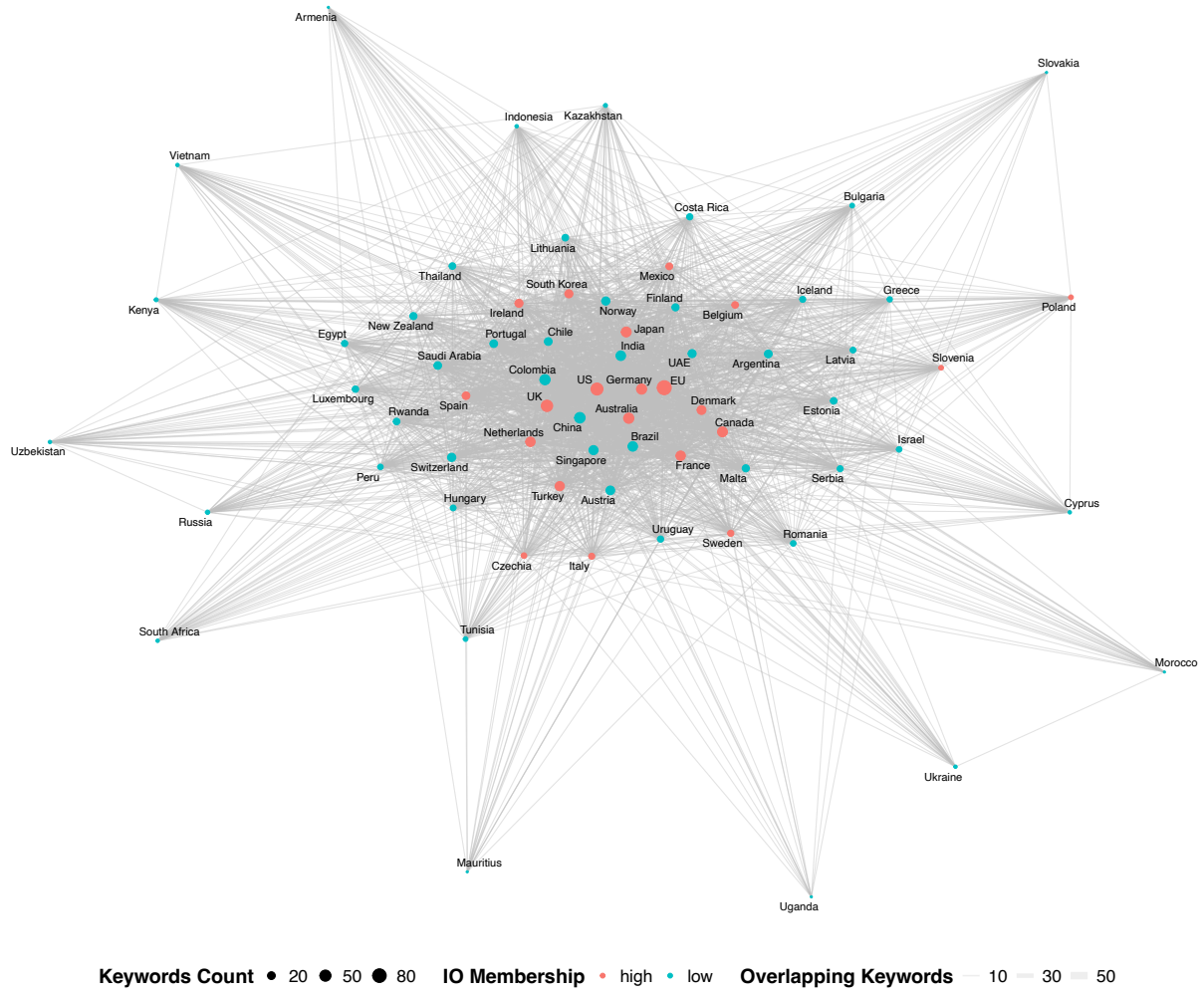


(c) 2019



(d) 2020

Figure 1.4: Text similarity on ethics and human rights (2017-2021)



Each node represents a country. The node size is scaled by the number of unique keywords related to ethics and human rights. Node colors indicate the level of membership in international organizations, including the EU, OECD, G7, G20, and GPAI. Countries participating in more than three IOs are categorized as having high membership, while those involved in fewer than two are classified as having low membership. The edges, or connections between nodes, represent the number of shared keywords between countries or levels of similarity in language. Edge widths

correspond to the level of similarity, with wider lines indicating a greater number of shared keywords. The visualization highlights clusters of countries with similar characteristics.

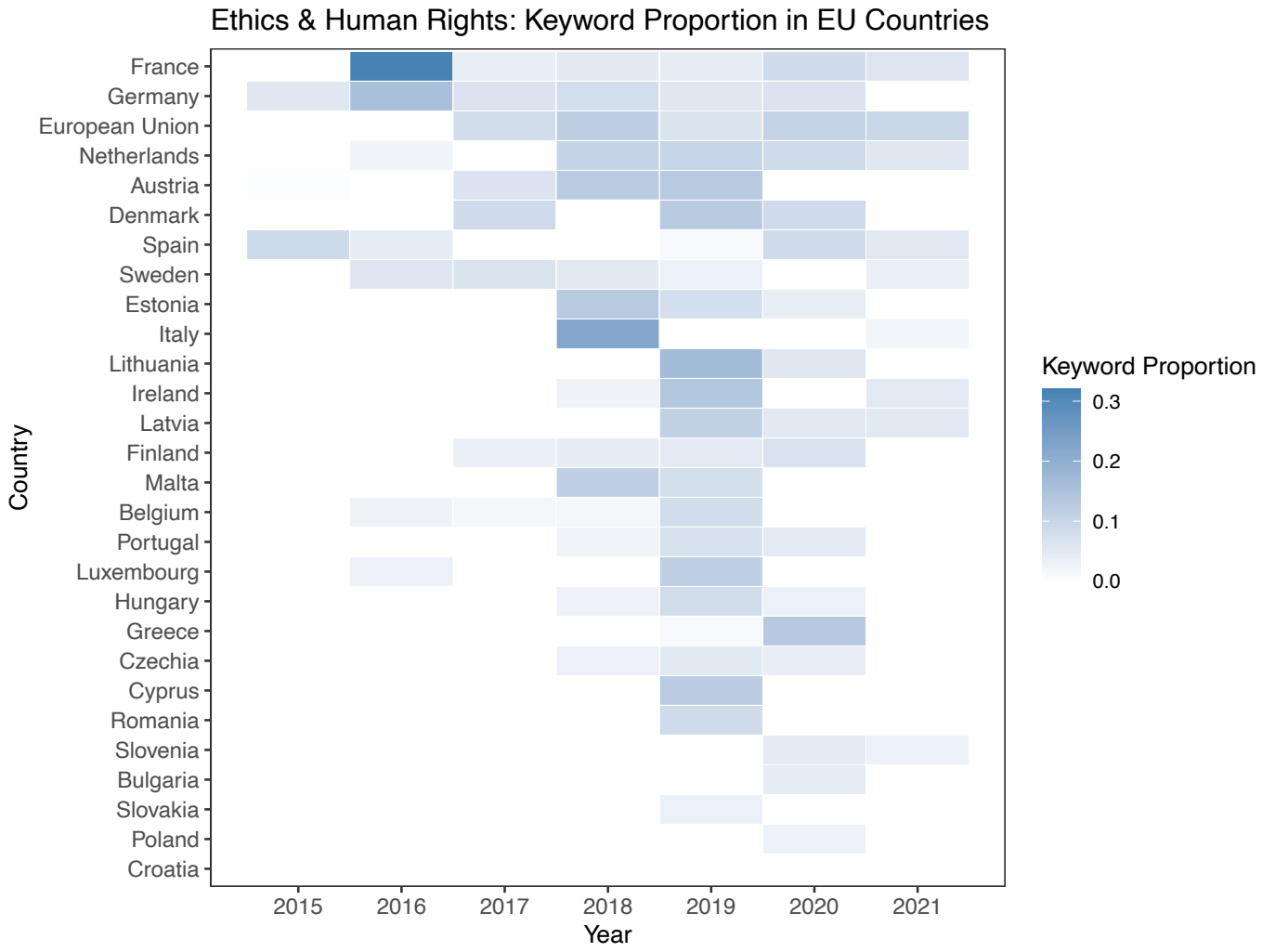
AI norms related to ethics and human rights have diffused to an increasing number of countries within a relatively short period. Between 2017 and 2019, the number of countries discussing ethics and rights protection more than doubled. Over time, countries have incorporated more keywords into their policies and are exhibiting denser interactions within the network. However, the divergence in unique keyword count indicates that many countries address ethics and rights protection in only generic terms, lacking the depth and expertise necessary for substantive impact.

Figure 1.4 presents the aggregated trends across the years 2017 to 2021, highlighting which actors hold central positions in the AI regulation network. Regional powers, particularly G7 countries, have played a significant role in cultivating AI norms. Canada, France, Germany, Japan, the UK, and the US advocated for ethics and human rights associated with AI early and consistently and maintained central positions in the regulation networks. Additionally, the EU has emerged as a global leader in advancing AI ethics and human rights. Countries with broader IO memberships tend to address ethics and human rights earlier and more frequently. Participation in multilateral governance has also encouraged developing countries to engage in discussions on AI ethics. For example, both Egypt and Argentina incorporated UNESCO's ethics recommendations into their national AI policies.

The EU has positioned itself as a proactive first mover, publishing influential AI policy initiatives worldwide. In 2018, the EU adopted the General Data Protection Regulation (GDPR), requiring member states to adhere to its guidelines and implement corresponding policies at the national level. Within the EU, a distinct pattern of first-movers and followers has emerged. Figure 1.5 presents the manually calculated keyword proportion for ethics and human rights in EU countries, ranked by their total proportion of keywords. France and Germany, the de facto leaders of the EU, have exhibited the highest proportions of keywords, adopting policies that mention ethics and rights nearly every year since 2015. They have driven the regional agenda and exerted influence over other member states. In 2015, Germany established a funded expert network called Platform Industry 4.0, and one of its objectives is to identify required actions on standards and norms and provide recommendations for national and international committees. France passed the Digital Republic

Bill in 2016, emphasizing equal rights for internet users. The Netherlands, Austria, Belgium, Sweden, Finland, and Spain are active contributors to AI ethics and rights protection. These countries tend to respond swiftly to international initiatives on AI norms and standardization. They have consistently made reference to fundamental rights in their AI policies and developed institutional mechanisms to ensure compliance.

Figure 1.5: Keyword usage within EU countries (2015-2021)



The EU's regulatory approach has been described as a "super-regulator," spanning from

data protection and antitrust laws to content moderation and AI risks.[19] Nonetheless, various actors have challenged the EU's regulatory power. While civil society groups argue that data rules do not go far enough in protecting personal data, the GDPR has been criticized for placing an excessive burden on SMEs and start-ups. Regarding implementation and compliance, there has also been a lack of clarity over how the rules apply, particularly to emerging technologies.[20] On the other hand, big tech firms argue that regulation stifles innovation and have pressured governments to weaken their proposals. The chief executive of SAP, Europe's largest software firm, has warned the EU against over-regulating AI, stating that it would harm Europe's competitiveness and widen the already large gap with the U.S.[21] During the final negotiation phase of the EU's AI Act, the French, German, and Italian governments advocated for limited regulation of AI foundation models, such as OpenAI's GPT, citing concerns that regulation could undermine their domestic AI industries.[22]

Several countries within the EU bloc act as followers, taking action only when required by law. The years 2018 and 2019 were pivotal for changes in AI governance due to GDPR. For approximately one-third of the countries, keywords related to ethics or human rights first appeared in their policies after 2018. Countries such as Greece, Slovakia, and Croatia have discussed ethics and human rights in their policies in only one year between 2015 and 2021. In summary, the analysis of keyword usage in EU countries reveals how first movers and followers adopt ethics and rights protection. While AI norms and policies are diffusing, the timing and frequency of policy adoption greatly vary across countries.

1.5.2 Dyad regression on specific keywords

Table 1.2 presents dyadic regression results examining similarity across specific keywords within the ethics and human rights category. The dependent variable is the frequency of overlapping

[19] Marietje Schaake, "Europe's 'Super-Regulator' Role Is Under Threat," Financial Times, January 15, 2025, https://www.ft.com/content/ce0d64b5-192e-48ae-a219-041874b580f2.

[20] Javier Espinoza, "EU Admits It Has Been Hard to Implement GDPR," Financial Times, June 23, 2020, https://www.ft.com/content/66668ba9-706a-483d-b24a-18cfbca142bf.

[21] Joe Miller, "SAP Chief Warns EU Against Over-Regulating Artificial Intelligence," Financial Times, September 20, 2024, https://www.ft.com/content/9db8fe6d-3f8a-4886-a439-c23faf459c23.

[22] Billy Perrigo, "E.U.'s AI Regulation Could Be Softened After Pushback," Time, August 1, 2023, https://time.com/6338602/eu-ai-regulation-foundation-models/.

Table 1.2: Dyad regression results by specific keywords

	Dependent variable:					
	Ethics	Privacy	Fairness	Transparency	Accountability	Security
	(1)	(2)	(3)	(4)	(5)	(6)
Shared IO	0.590***	0.168***	0.070***	0.171***	0.046***	0.310***
	(0.047)	(0.019)	(0.014)	(0.018)	(0.007)	(0.024)
SmartPol	0.364***	0.121***	0.172***	0.339***	0.096***	0.330***
	(0.090)	(0.036)	(0.026)	(0.035)	(0.014)	(0.045)
UN vote diff	0.193***	0.022	0.038**	0.064***	0.039***	0.136***
	(0.059)	(0.024)	(0.016)	(0.022)	(0.009)	(0.030)
GDP diff	−0.270***	0.134***	0.004	0.051**	−0.010	0.014
	(0.056)	(0.023)	(0.017)	(0.022)	(0.009)	(0.028)
AI capacity diff	0.055	−0.072**	0.014	−0.148***	−0.022	−0.091**
	(0.088)	(0.035)	(0.026)	(0.034)	(0.014)	(0.044)
Libdem diff	0.521**	0.330***	0.151***			0.527***
	(0.202)	(0.081)	(0.055)			(0.102)
Ethnicity diff			−0.169***			
			(0.056)			
Corruption diff				−0.616***	−0.08 ***	
				(0.093)	(0.040)	
Reg.Quality diff	0.190**	−0.107***	−0.026	0.162***	0.035*	−0.038
	(0.084)	(0.034)	(0.023)	(0.036)	(0.015)	(0.042)
Constant	0.362***	0.030	0.017	0.177***	−0.038**	−0.102*
	(0.109)	(0.044)	(0.034)	(0.041)	(0.017)	(0.055)
Observations	1,830	1,830	1,431	1,830	1,830	1,830
R^2	0.129	0.064	0.063	0.150	0.088	0.139
Adjusted R^2	0.126	0.061	0.058	0.147	0.084	0.136
Note:						*p<0.1; **p<0.05; ***p<0.01

keywords between country pairs. Shared IO Membership is consistently and significantly associated with higher similarity across all keyword categories, with strong effects on ethics and security. This suggests that IO networks play an important role in norm convergence, and countries are more likely to align in ethics and security principles. Smart policing presence is also positively correlated with similarity across all six categories, with stronger coefficient estimates for ethics, transparency, and security. Countries implementing AI in law enforcement tend to align their language as they may face similar concerns or regulatory challenges.

The absolute difference between the two countries' ideal point estimates in the UN General Assembly voting shows positive effects on ethics, fairness, transparency, and accountability. This indicates that a larger difference in foreign policy preferences is actually associated with a wider divergence in AI norms. AI capacity difference is negatively associated with privacy, transparency, and security. Similar country profiles in AI research, development, and commercialization may encourage more convergence in these domains. Divergence in GDP per capita is negatively associated with discussions on ethics yet positively associated with privacy and transparency. This indicates that while GDP is an indicator of generic terms such as ethics, it is not a similarity predictor for specific norms. In addition, regime type is not associated with greater convergence in normative discourse. Instead, similarities in the corruption index lead to greater convergence in transparency and accountability, emphasizing the effect of institutional trust and the rule of law. In particular, the smaller gap in the ethnic fractionalization index significantly increases the similarity in fairness language. Countries with similar characteristics in ethnic diversity are more likely to discuss fairness or non-discrimination principles.

Table 1.3: AI policy and objectives

Country	Year	Policy Name	Privacy	Fairness	Transparency	Accountability
Germany	2016	Ethical Guidelines For Self-Driving Cars				
EU	2018	Ethics Guidelines on AI	✓	✓	✓	✓
	2018	Declaration on AI in the Nordic-Baltic Region (Estonia, Iceland, Denmark, Latvia, and others)	✓			
India	2018	AI Standardisation Committee	✓			
Mexico	2018	Principles and Impact Analysis Guide for AI Use	✓	✓	✓	
S. Korea	2018	Ethics Guidelines for Intelligent Society			✓	✓
Singapore	2018	Principles for Fairness and Transparency in AI		✓	✓	✓
Australia	2019	Australia's AI Ethics Framework	✓	✓	✓	✓
Canada	2019	Directive on Automated Decision-Making		✓	✓	✓
Canada	2019	Digital Charter	✓		✓	✓
China	2019	Beijing Consensus on AI and Education		✓		
China	2019	Governance Principles for Responsible AI	✓	✓		
EU	2019	Comprehensive Policy on AI and Robotics			✓	
Germany	2019	FAIR Forward – AI for All	✓			
Hungary	2019	AI Ethical Guidelines				
Ireland	2019	Data Sharing And Governance Act	✓			
Japan	2019	AI Utilisation Guidelines	✓	✓	✓	✓
Malta	2019	Malta's Ethical AI Framework				
New Zealand	2019	Algorithm Charter for Aotearoa			✓	✓
Norway	2019	AI and Privacy Report	✓		✓	
Saudi Arabia	2019	Saudi Data and AI Authority		✓	✓	✓
UAE	2019	AI Principles and Ethics for the Emirate Of Dubai		✓	✓	✓
UK	2019	Online Harms White Paper				✓
US	2019	Memorandum to Heads of Agencies on Regulatory and Non-Regulatory Approaches to AI	✓			

To examine the substance of ethics and human rights policies with greater precision, I compiled a table of AI policies referencing four specific keywords—privacy, fairness/non-discrimination, transparency, and accountability—or include ethics in their title. Although the policies vary in their use of hard and soft law, the comparison focuses on regulatory intent and priorities. Privacy and non-discrimination are two fundamental rights that can be negatively impacted by AI systems. Transparency and accountability are exemplary AI ethical principles and are frequently included in ethics guidelines. Table 1.3 and 1.4 list the policy names and corresponding countries from 2018 to 2020. The analysis is limited to AI-specific policies with a regulatory focus, excluding policies primarily aimed at promoting AI, such as national AI strategies. This list may not be exhaustive, as it is based on policy objectives rather than original documents. However, it aims to provide qualitative evidence of what these policies actually entail, complementing the earlier keyword usage analysis.

The table suggests that in 2018 and 2019, democratic and economically advanced countries were at the forefront of AI regulation, adopting policies that addressed multiple principles, such as privacy, fairness, transparency, and accountability. The EU, Canada, France, Germany, and the UK have positioned themselves as global agenda-setters, establishing comprehensive governance frameworks. Similarly, Asian countries such as Singapore, Japan, and South Korea developed detailed and technical guidelines for businesses. According to the Global AI Index, these countries rank among the top ten in terms of AI capacity. [23] Meanwhile, the AI superpowers — the US and China — are engaged in an AI race and may refrain from stringent regulations to maintain competitiveness and leadership. Nonetheless, the EU and other wealthy democratic countries have both the capacity and motivation to safeguard fundamental rights and develop multifaceted AI policies early on.

In the Middle East, Saudi Arabia established the Saudi Data and AI Authority in 2019, while the UAE released AI principles and ethics for the Emirate of Dubai. Among developing countries, Brazil and Egypt adopted personal data protection laws in 2020, followed by Rwanda in 2021, signaling the expanding influence of the EU's GDPR model. This suggests the diffusion

[23] See https://www.tortoisemedia.com/intelligence/global-ai

of AI-related norms and regulations, where developing countries align with established global rules to facilitate trade and attract investment. This may speak to the de facto 'Brussels Effect', the phenomenon in which large companies must comply with GDPR rules within the EU, leading firms to adopt these regulations internationally. This, in turn, influences firms and NGOs lobbying local governments to pass similar regulations (Tamim 2024).

Table 1.4: AI policy and objectives

Country	Year	Policy Name	Privacy	Fairness	Transparency	Accountability
Brazil	2020	General Law on Protection of Personal Data	✓			
Colombia	2020	Ethical Framework for Artificial Intelligence in Colombia				
Denmark	2020	Law on the Disclosure of Data Ethics Policy			✓	
Egypt	2020	Personal Data Protection Law	✓			
EU	2020	Juri Reports on "Making AI European"	✓	✓	✓	✓
EU	2020	The Robustness and Explainability of Artificial Intelligence			✓	
EU	2020	White Paper on Artificial Intelligence	✓	✓	✓	✓
France	2020	AI Sandbox Program of the National Data Protection Authority	✓			
Japan	2020	Machine Learning Quality Management Guideline				✓
S. Korea	2020	Human-Centered National Guidelines for AI Ethics				
Singapore	2020	AI Ethics and Governance Body of Knowledge	✓	✓		
Switzerland	2020	Guidelines on AI			✓	
UK	2020	Data Ethics and AI Guidance Landscape				
UK	2020	Guidance on AI and Data Protection	✓			
UK	2020	Review Into Bias in Algorithmic Decision-Making		✓		✓

1.6 Conclusion

The literature has established that liberal democracies are more likely to protect human rights. However, the binary classification of regime types does not fully capture AI's threats to societies nor the variation in human rights protection within democratic countries. The scope of human rights violations related to AI systems extends beyond mass surveillance and online censorship. The lack of privacy protection and algorithmic bias are becoming increasingly prevalent without adequate regulatory oversight. AI governance involves complex and interconnected issues, making it challenging to accurately assess the varying levels of rights protection in AI policies worldwide.

This descriptive analysis examines the dynamics of AI governance over time, exploring when AI regulation first emerged and how it has diffused. This project aims to make conceptual and measurement contributions to this emerging field and engage in meaningful debates about the political consequences of increasing automation and its impact on globalization. Furthermore, this paper explores how governments balance social protection with innovation in the age of automated systems and the role of international actors in bringing countries together to initiate discussions and exchange ideas. Presenting the empirical research of AI through text analysis provides insights into the government's regulatory preferences and sheds light on the timeline of policy adoption.

Appendix A: Appendix for Chapter 1

A.1 Keyword lists

- **Ethics and HumanRights**: accessible, accountability, accurate, bias, balanced, challenge, challenges, centric, centred, citizens, civil, comprehensive, confidentiality, consumer, consumers, controllability, contestability, credible, democracy, democratic, democratisation, discrimination, discriminatory, diverse, diversity, dignity, equal, equally, equitable, exclusionary, ethics, ethical, explainability, explainable, fair, fairness, flexibility, freedom, freedoms, fundamental, guidelines, harm, holistic, human, humans, impact, inclusive, inclusion, inclusiveness, integrity, interoperability, justice, liable, liability, liberty, liberties, marginalized, norm, norms, nondiscrimination, personal, principles, privacy, protection, protect, protecting, protected, proportionate, reliable, reliability, responsible, responsibility, responsibilities, responsibly, resilience, right, rights, risk, risks, robust, robustness, safe, safeguard, safeguards, safeguarding, safety, secure, security, societal, social, society, stereotypes, sustainable, traceability, transparent, transparency, trust, trusts, trusted, trustworthiness, trustworthy, uncertainty, unequal, unbiased, usable, user-centered, users, value, values, verifiable, vulnerabilities, wellbeing.

- **Market and Innovation**: academia, access, accelerate, advance, advanced, benefit, breakthroughs, business, businesses, capability, capabilities, collaboration, commercial, competition, competitive, competitiveness, computational, computing, companies, competence, defence, defense, deployment, deploying, development, discovery, digitization, economic, economy, ecosystem, ecosystems, education, educational, employment, entrepreneurship, entrepreneurs, excellence, expert, experts, expertise, funding, grant, growing, growth, industry, industrial, infrastructure, innovation, innovative, intellectual, investment, investments, investors, job, jobs, knowledge, labour, leadership, manufacturing, machine, market, markets, network, new, patent, production, productivity, procurement, products, projects, publications, research, researchers, resources, revolution, science, scientific, sciences, sector, sectoral, sectors, service, services, skills, skilled, startups, talent, technical, technological, technologies,

technology, trade, training, transformed, transformation, vision, workforce.

- **Policy and Regulation**: assessments, commission, compliance, consultations, coordinate, coordinated, coordination, council, directive, disciplines, enforcement, evaluate, framework, governance, guidance, investigate, investigation, law, laws, lawful, legal, legality, legislation, mitigate, mitigating, mitigation, monitor, oversight, plan, platform, policy, prevent, procedures, proposal, redress, regulate, regulation, regulatory, recommendations, report, resolution, respect, review, sandbox, sandboxes, standards, standardisation, standardization, stakeholders, verification.

Figure A.1: Text similarity on ethics and human rights: unique keyword count (2017)

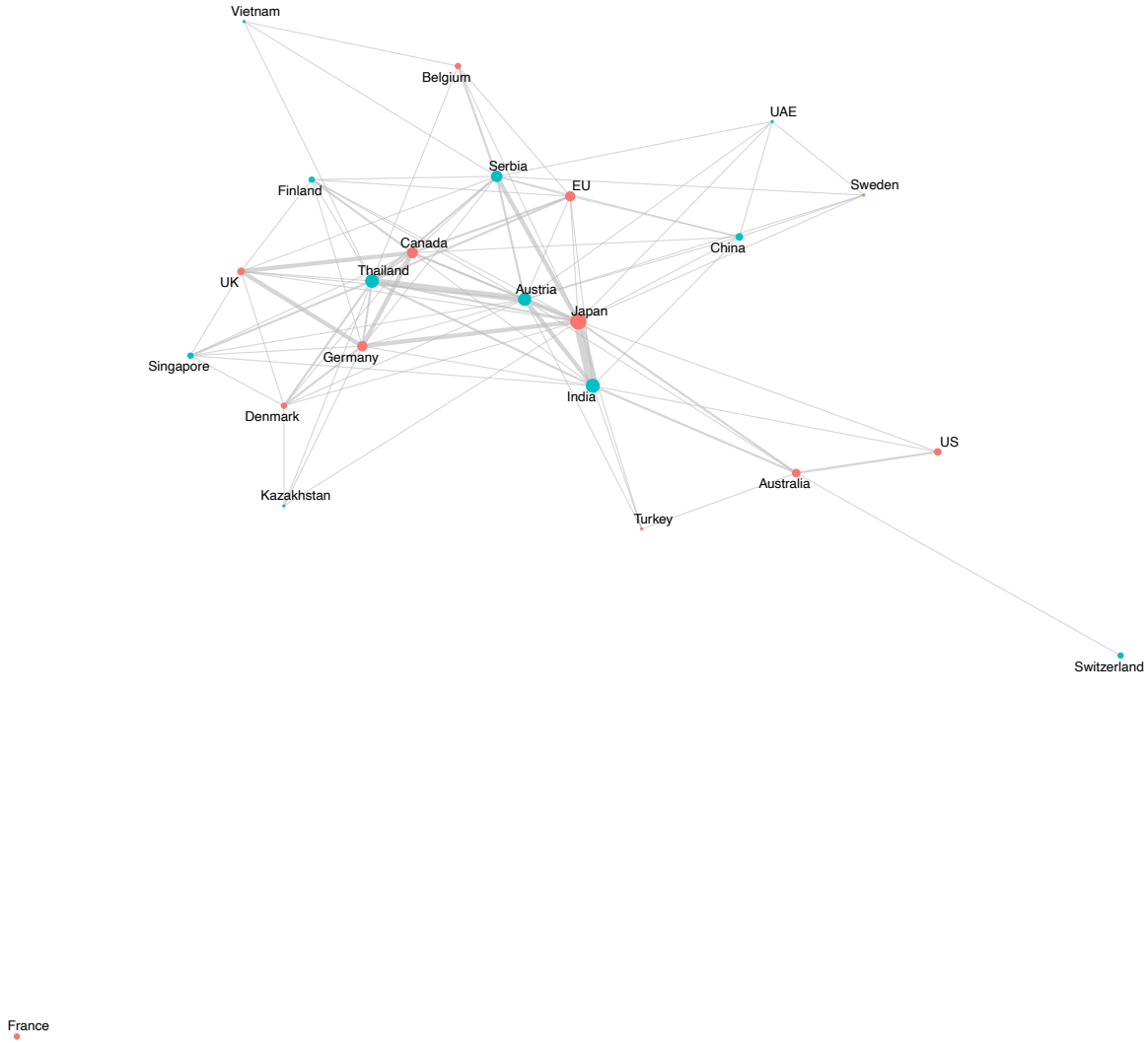


Figure A.2: Text similarity on ethics and human rights: unique keyword count (2018)

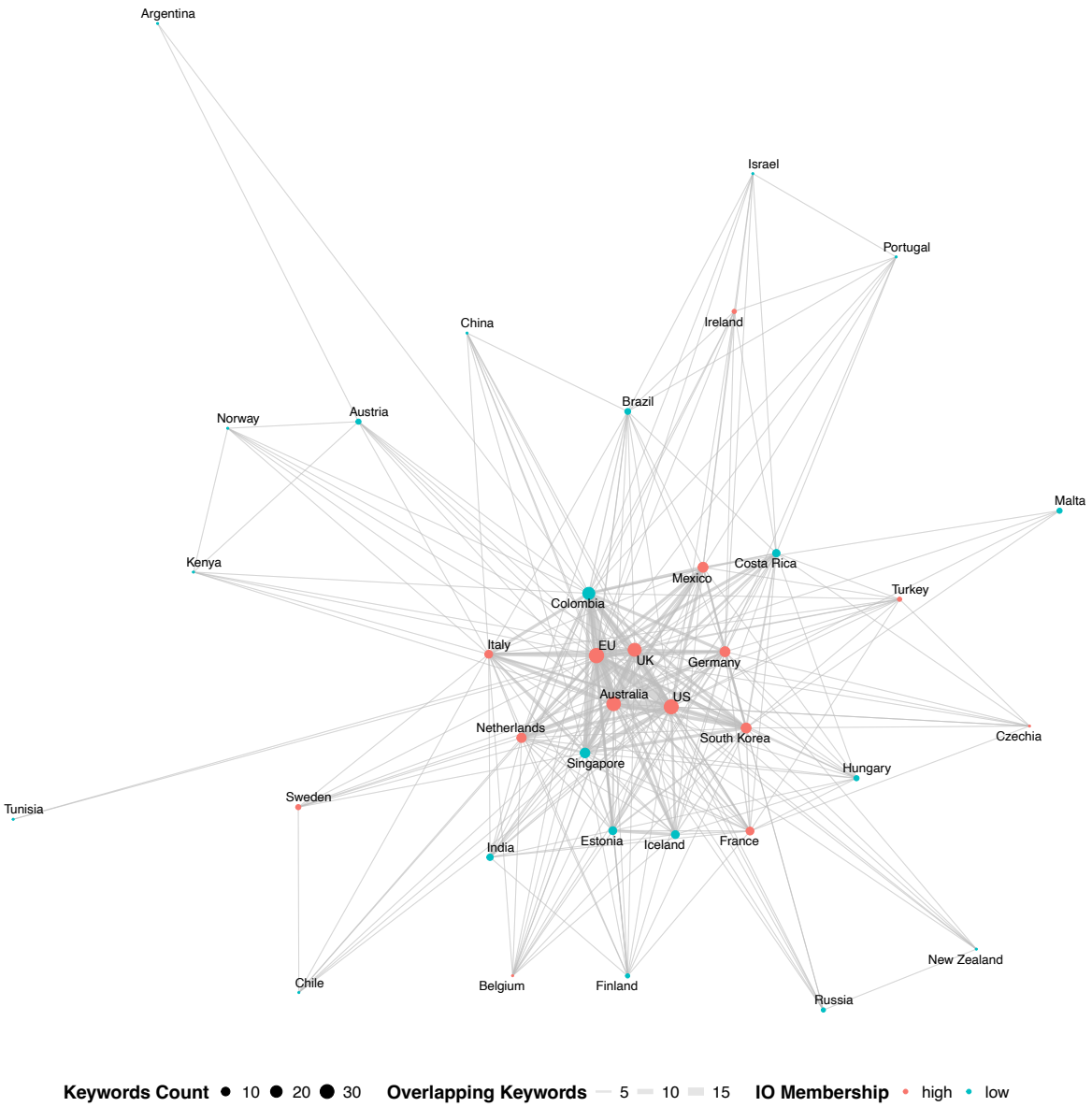


Figure A.3: Text similarity on ethics and human rights: unique keyword count (2019)

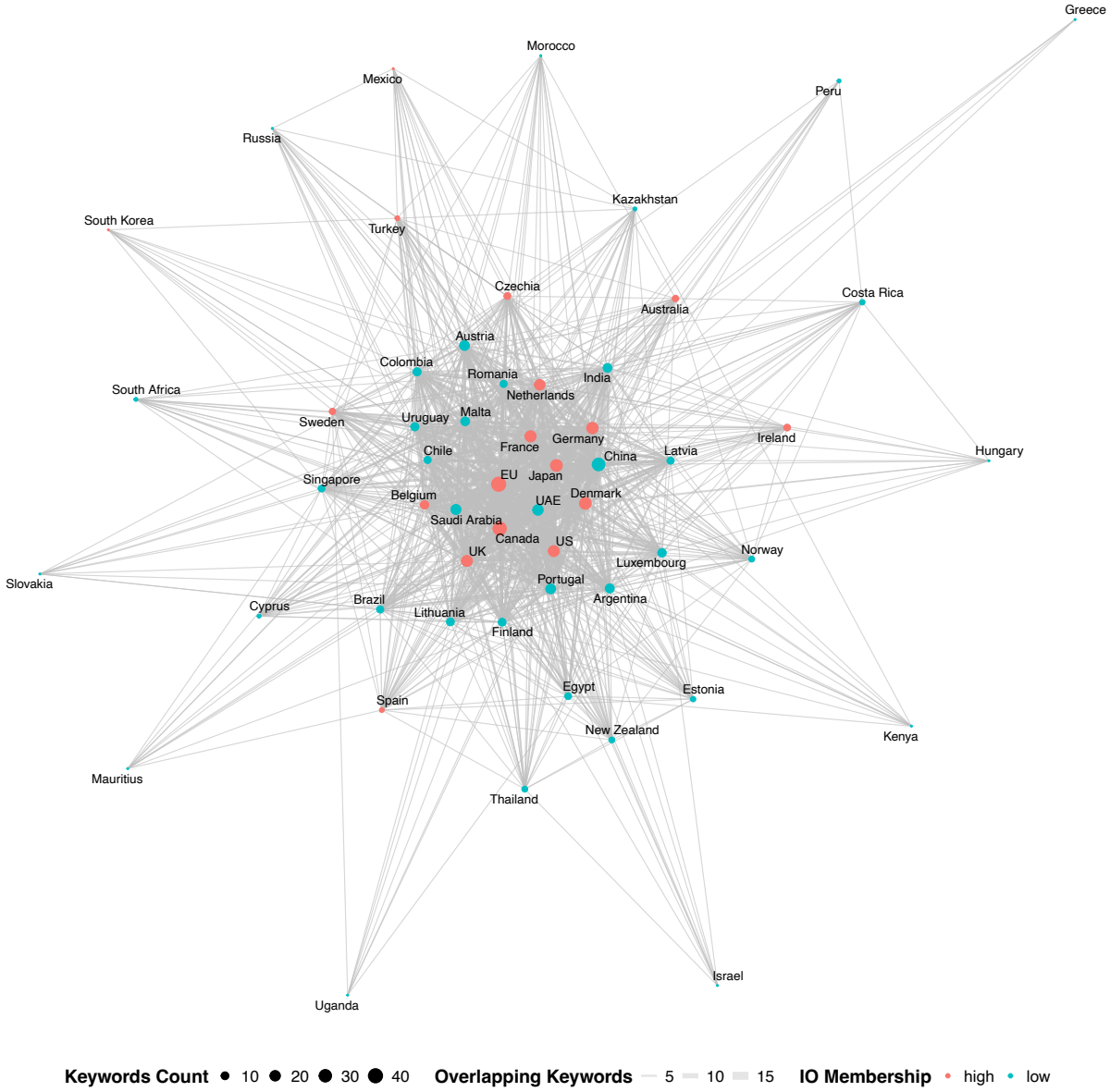


Figure A.4: Text similarity on ethics and human rights: unique keyword count (2020)

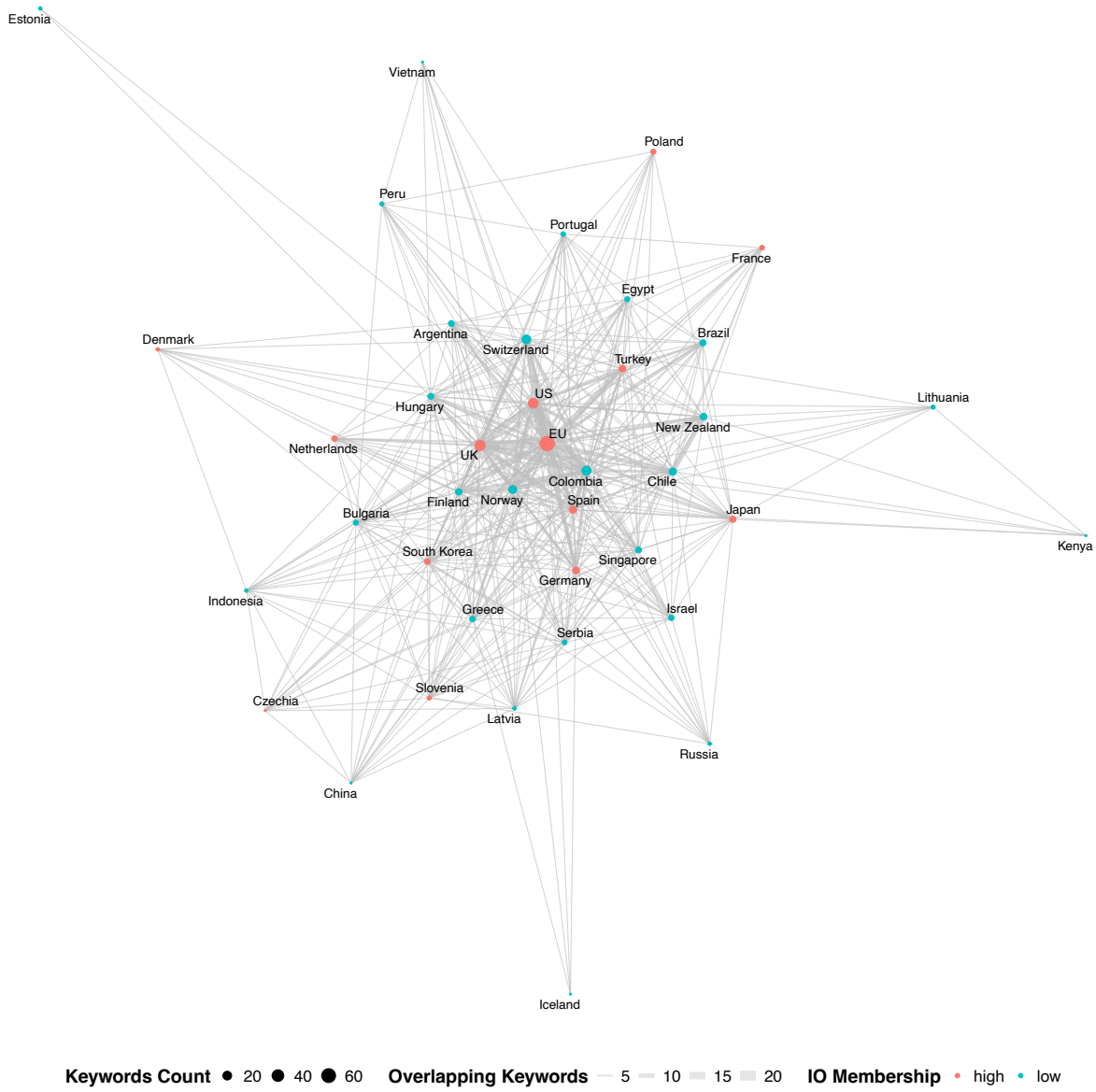


Table A.1: AI policy count by country in the OECD dataset

Country	**Count**	**Country**	**Count**
Argentina	10	Malta	5
Armenia	2	Mauritius	2
Australia	32	Mexico	6
Austria	9	Morocco	4
Belgium	22	Netherlands	12
Brazil	12	New Zealand	8
Bulgaria	2	Nigeria	2
Canada	14	Norway	19
Chile	11	Peru	9
China	21	Poland	4
Colombia	30	Portugal	11
Costa Rica	7	Romania	3
Croatia	1	Russia	11
Cyprus	2	Rwanda	6
Czech Republic	8	Saudi Arabia	5
Denmark	12	Serbia	18
Egypt	7	Singapore	25
Estonia	9	Slovakia	2
European Union	60	Slovenia	6
Finland	12	South Africa	3
France	34	South Korea	14
Germany	33	Spain	17
Greece	3	Sweden	13
Hungary	14	Switzerland	6
Iceland	4	Thailand	5
India	23	Tunisia	7
Indonesia	1	Turkey	32
Ireland	8	Uganda	3
Israel	8	Ukraine	1
Italy	10	United Arab Emirates	8
Japan	23	United Kingdom	55
Kazakhstan	7	United States	75
Kenya	6	Uruguay	4
Latvia	5	Uzbekistan	3
Lithuania	4	Vietnam	6
Luxembourg	6		

References

Abbott, Kenneth W and Duncan Snidal. 2000. "Hard and Soft Law in International Governance." *International Organization* 54(3):421–456.

Abbott, Kenneth W, Robert O Keohane, Andrew Moravcsik, Anne-Marie Slaughter and Duncan Snidal. 2000. "The Concept of Legalization." *International Organization* 54(3):401–419.

Anderljung, Markus, Joslyn Barnhart, Anton Korinek, Jade Leung, Cullen O'Keefe, Jess Whittlestone, Shahar Avin, Miles Brundage, Justin Bullock, Duncan Cass-Beggs et al. 2023. "Frontier AI Regulation: Managing Emerging Risks to Public Safety." *arXiv preprint arXiv:2307.03718* .

Arel-Bundock, Vincent. 2017. "The Unintended Consequences of Bilateralism: Treaty Shopping and International Tax Policy." *International Organization* 71(2):349–371.

Baccini, Leonardo, Pablo M Pinto and Stephen Weymouth. 2017. "The Distributional Consequences of Preferential Trade Liberalization: Firm-Level Evidence." *International Organization* .

Bailey, Michael A, Anton Strezhnev and Erik Voeten. 2017. "Estimating Dynamic State Preferences from United Nations Voting Data." *Journal of Conflict Resolution* 61(2):430–456.

Baker, Andrew. 2010. "Restraining Regulatory Capture? Anglo-America, Crisis Politics and Trajectories of Change in Global Financial Governance." *International Affairs* 86(3):647–663.

Barnett, Michael N. and Martha Finnemore. 1999. "The Politics, Power, and Pathologies of International Organizations." *International Organization* 53(4):699–732.

Barnett, Michael and Raymond Duvall. 2005. "Power in International Politics." *International Organization* 59(1).

Bush, Sarah Sunn and Pär Zetterberg. 2021. "Gender Quotas and International Reputation." *American Journal of Political Science* 65(2):326–341.

Cao, Xun. 2010. "Networks as Channels of Policy Diffusion: Explaining Worldwide Changes in Capital Taxation, 1998–2006." *International Studies Quarterly* 54(3):823–854.

Carnegie Endowment for International Peace. 2019. "AI Global Surveillance (AIGS) Index." https://carnegieendowment.org/files/AI_Global_Surveillance_Index1.pdf.

Checkel, Jeffrey T. 2001. "Why Comply? Social Learning and European Identity Change." *International Organization* 55:553–588.

Checkel, Jeffrey T. 2005. "International institutions and socialization in Europe: Introduction and framework." *Internatioinal Organization* 59(04).

Chen, Zhisheng. 2023. "Ethics and Discrimination in Artificial Intelligence-Enabled Recruitment Practices." *Humanities and Social Sciences Communications* 10(1):1–12.

Chiba, Daina, Jesse C Johnson and Brett Ashley Leeds. 2015. "Careful Commitments: Democratic States and Alliance Design." *The Journal of Politics* 77(4):968–982.

Cole, W. 2015. "Mind the gap: State capacity and the implementation of human rights treaties." *Internatioinal Organization* 69:405–441.

Coppedge, Michael, John Gerring, Carl Henrik Knutsen et al. 2023. "V-Dem Country-Year Dataset v13.".

Dafoe, Allan. 2018. "AI Governance: A Research Agenda.".

Davenport, Christian. 2007. "State Repression and Political Order." *Annual Review of Political Science* 10(1):1–23.

De Vries, Catherine E., Sara B. Hobolt and Stefanie Walter. 2021. "Politicizing International Cooperation: The Mass Public, Political Entrepreneurs, and Political Opportunity Structures." *International Organization* 75(2):306–322.

Dobbin, Frank, Beth Simmons and Geoffrey Garrett. 2007. "The global diffusion of public policies: Social construction, coercion, competition, or learning?" *The Annual Review of Sociology* 33(1):449–472.

Donnelly, Jack. 1999. "Human Rights, Democracy, and Development." *Human Rights Quarterly* 21(3):608–632.

Doshi, Rush, Judith G. Kelley and Beth A. Simmons. 2019. "The Power of Ranking: The Ease of Doing Business Indicator and Global Regulatory Behavior." *International Organization* 73(3):611–643.

Dražanová, Lenka. 2020. "Introducing the Historical Index of Ethnic Fractionalization (HIEF) Dataset: Accounting for Longitudinal Changes in Ethnic Diversity." *Journal of Open Humanities Data* 6.

Elkins, Zachary, Andrew T Guzman and Beth A Simmons. 2006. "Competing for Capital: The Diffusion of Bilateral Investment Treaties, 1960–2000." *International Organization* 60(4):811–846.

Engler, Alex. 2022. "The EU AI Act Will Have a Global Impact, but a Limited Brussels Effect." https://www.brookings.edu/research/the-eu-ai-act-will-have-global-impact-but-a-limited-brussels-effect/.

Erman, Eva and Markus Furendal. 2024*a*. "Artificial Intelligence and the Political Legitimacy of Global Governance." *Political Studies* 72(2):421–441.

Erman, Eva and Markus Furendal. 2024*b*. "The Democratization of Global AI Governance and the Role of Tech Companies." *Nature Machine Intelligence* 6(3):246–248.

Eshima, Shusei, Kosuke Imai and Tomoya Sasaki. 2023. "Keyword-assisted topic models." *American Journal of Political Science* .

Espinoza, Javier. 2024. "Europe's Rushed Attempt to Set the Rules for AI." *Financial Times* .

European Commission. 2021. "Proposal for Artificial Intelligence Act." https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32024R1689.

Fearon, James D. 1998. "Bargaining, Enforcement, and International Cooperation." *International Organization* 52(2):269–305.

Finnemore, Martha. 1993. "International Organizations as Teachers of Norms: The United Nations Educational, Scientific, and Cultural Organization and Science Policy." *International Organization* 47(4):565–597.

Finnemore, Martha and Kathryn Sikkink. 1998. "International Norm Dynamics and Political Change." *International Organization* 52(4):887–917.

Fukuda-Parr, Sakiko and Elizabeth Gibbons. 2021. "Emerging Consensus on 'Ethical AI': Human Rights Critique of Stakeholder Guidelines." *Global Policy* 12(S6):32–44.

Galindo, Lourdes, Karen Perset and Fares Sheeka. 2021. "An Overview of National AI Strategies and Policies.".
URL: *https://doi.org/10.1787/c05140d9-en*

Garriga, Ana Carolina. 2016. "Human Rights Regimes, Reputation, and Foreign Direct Investment." *International Studies Quarterly* 60(1):160–172.

Graham, Erin R. and Alexandria Serdaru. 2020. "Power, Control, and the Logic of Substitution in Institutional Design: The Case of International Climate Finance." *International Organization* 74(4):671–706.

Greenhill, Brian, Layna Mosley and Aseem Prakash. 2009. "Trade-based diffusion of labor rights: A panel study, 1986–2002." *American Political Science Review* 103(4):669–690.

Haas, Peter M. 1992. "Epistemic Communities and International Policy Coordination." *International Organization* 46(1):1–35.

Hall, Peter A and David Soskice. 2001. *Varieties of Capitalism: The Institutional Foundations of Comparative Advantage*. London, England: Oxford University Press.

Heinrich, Tobias, Yoshiharu Kobayashi and Leah Long. 2018. "Voters Get What They Want (When They Pay Attention): Human Rights, Policy Benefits, and Foreign Aid." *International Studies Quarterly* 62(1):195–207.

Helmke, Gretchen and Steven Levitsky. 2006. *Informal Institutions and Democracy: Lessons from Latin America*. Baltimore, MD: Johns Hopkins University Press.

Ho, Lewis, Joslyn Barnhart, Robert Trager, Yoshua Bengio, Miles Brundage, Allison Carnegie, Rumman Chowdhury, Allan Dafoe, Gillian Hadfield, Margaret Levi and Duncan Snidal. 2023. "International Institutions for Advanced AI." *arXiv preprint arXiv:2307.03718* .

Hooghe, Liesbet. 2005. "Several Roads Lead to International Norms, but Few via International Socialization: A Case Study of the European Commission." *International Organization* 59(4).

Jobin, Anna, Marcello Ienca and Effy Vayena. 2019. "The Global Landscape of AI Ethics Guidelines." *Nature Machine Intelligence* 1(9):389–399.

Kalyanpur, Nikhil and Abraham L. Newman. 2019. "Mobilizing Market Power: Jurisdictional Expansion as Economic Statecraft." *International Organization* 73(1):1–34.

Kaufmann, Daniel and Aart Kraay. 2024. "Worldwide Governance Indicators, 2024 Update." https://www.worldbank.org/en/publication/worldwide-governance-indicators.

Kelley, Judith. 2008. "Assessing the Complex Evolution of Norms: The Rise of International Election Monitoring." *International Organization* 62(2):221–255.

Keohane, Robert O. 2005. *After Hegemony: Cooperation and Discord in the World Political Economy*. Princeton University Press.

Kim, In Song and Iain Osgood. 2019. "Firms in Trade and Trade Politics." *Annual Review of Political Science* 22(1):399–417.

Konisky, David M and Manuel P Teodoro. 2016. "When governments regulate governments." *American Journal of Political Science* 60(3):559–574.

Koremenos, Barbara. 2005. "Contracting around international uncertainty." *American Political Science Review* 99:549–565.

Koremenos, Barbara, Charles Lipson and Duncan Snidal. 2001. "The Rational Design of International Institutions." *International Organization* 55(4):761–799.

Kreps, Sarah and Doug Kriner. 2023. "How AI Threatens Democracy." *Journal of Democracy* 34(4):122–131.

Mansfield, Edward D. and Nita Rudra. 2021. "Embedded Liberalism in the Digital Era." *International Organization* 75(2):558–585.

Martin, Lisa L. and Beth A. Simmons. 1998. "Theories and Empirical Studies of International Institutions." *International Organization* 52(4):729–757.

National Institute of Standards and Technology. 2019. "Face Recognition Vendor Test." https://www.nist.gov/programs-projects/face-recognition-vendor-test-frvt.

Newman, Abraham and David Bach. 2014. "The European Union as Hardening Agent: Soft Law and the Diffusion of Global Financial Regulation." *Journal of European Public Policy* 21(3):430–452.

Newman, Abraham and Elliot Posner. 2016. "Transnational feedback, soft law, and preferences in global financial regulation." *Review of International Political Economy* 23(1):123–152.

OECD.AI. 2021. "Database of National AI Policies." https://oecd.ai/en/dashboards/overview.

OHCHR. N.d. "International Bill of Human Rights." https://www.ohchr.org/en/what-are-human-rights/international-bill-human-rights.

Rosenbaum, P. and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70:41–55.

Schiff, Daniel, Justin Biddle, Jason Borenstein and Kelly Laas. 2020. What's Next for AI Ethics, Policy, and Governance? A Global Overview. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*. New York, NY: ACM.

Schmitt, Lewin. 2022. "Mapping Global AI Governance: A Nascent Regime in a Fragmented Landscape." *AI and Ethics* 2(2):303–314.

Schopmans, Hendrik R. 2022. From Coded Bias to Existential Threat: Expert Frames and the Epistemic Politics of AI Governance. In *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*. AIES '22 New York, NY: Association for Computing Machinery pp. 627–640.

Shipan, Charles R and Craig Volden. 2008. "The Mechanisms of Policy Diffusion." *American Journal of Political Science* 52(4):840–857.

Sikkink, Kathryn. 1993. "Human Rights, Principled Issue-Networks, and Sovereignty in Latin America." *International Organization* 47(3):411–441.

Simmons, Beth A. 2009. *Mobilizing for Human Rights: International Law in Domestic Politics*. Cambridge University Press.

Simmons, Beth A. and Zachary Elkins. 2004. "The Globalization of Liberalization: Policy Diffusion in the International Political Economy." *American Political Science Review* 98(1):171–189.

Slaughter, Anne-Marie. 1997. "The Real New World Order." https://www.foreignaffairs.com/articles/1997-09-01/real-new-world-order.

Sommerer, Thomas and Jonas Tallberg. 2019. "Diffusion Across International Organizations: Connectivity and Convergence." *International Organization* 73(2):399–433.

Stigler, George J. 2021. The Theory of Economic Regulation. In *The Political Economy*. Routledge pp. 67–81.

Stone, Randall W. 2004. "The Political Economy of IMF Lending in Africa." *American Political Science Review* 98(4):577–591.

Stone, Randall W. 2011. *Controlling Institutions: International Organizations and the Global Economy*. Cambridge University Press.

Tallberg, Jonas, Eva Erman, Markus Furendal, Johannes Geith, Mark Klamberg and Magnus Lundgren. 2023. "The Global Governance of Artificial Intelligence: Next Steps for Empirical and Normative Research." *International Studies Review* 25(3):viad040.

Tallberg, Jonas, Magnus Lundgren, Thomas Sommerer and Theresa Squatrito. 2020. "Why International Organizations Commit to Liberal Norms." *International Studies Quarterly* 64(3):626–640.

Tamim, James. 2024. The Brussels Effect and the GDPR: EU Institutions as Catalysts for Global Data Protection Norms. Technical report.

Tarrow, Sidney. 2005. *The New Transnational Activism*. Cambridge University Press.

The World Bank. 2024. "GDP per Capita (Current US$) — World Bank Data." https://data.worldbank.org/indicator/NY.GDP.PCAP.CD.

Timmermans, Stefan and Steven Epstein. 2010. "A World of Standards but Not a Standard World: Toward a Sociology of Standards and Standardization." *Annual Review of Sociology* 36(1):69–89.

UNCTAD. 2021. "Data Protection and Privacy Legislation Worldwide." https://unctad.org/page/data-protection-and-privacy-legislation-worldwide.

Underhill, Geoffrey R D and Xiaoke Zhang. 2008. "Setting the Rules: Private Power, Political Underpinnings, and Legitimacy in Global Monetary and Financial Governance." *International Affairs* 84(3):535–554.

Vought, Russell T. 2020. "Memorandum for Heads of Executive Departments and Agencies: Guidance for Regulation of Artificial Intelligence Applications." https://www.whitehouse.gov/wp-content/uploads/2020/01/Draft-OMB-Memo-on-Regulation-of-AI-1-7-19.pdf.

Weeks, Ana Catalano. 2018. "Why Are Gender Quota Laws Adopted by Men? The Role of Inter- and Intraparty Competition." *Comparative Political Studies* 51(14):1935–1973.

Westerwinter, Oliver, Kenneth W. Abbott and Thomas Biersteker. 2021. "Informal Governance in World Politics." *Review of International Organizations* 16(1):1–27.

Wilf, Meredith. 2016. "Credibility and Distributional Effects of International Banking Regulations: Evidence from US Bank Stock Returns." *International Organization* 70(4):763–796.

Wilson, James Q. 1980. *The Politics of Regulation*. Basic Books.

Zürn, Michael. 2018. *A Theory of Global Governance: Authority, Legitimacy, and Contestation*. Oxford University Press.