university of groningen / faculty of arts

Mapping Governance Strategies for AI Sustainability

An Analysis of EU, USA, and China Regulatory Frameworks

Master Thesis

Digital Humanities

July 24th, 2024

Author

Jan Czechowicz, s5767512

Supervisor

PhD M.L. Flórez Rojas

ABSTRACT

GitHub repository: https://github.com/jczechowicz/mapping-governance-strategies-of-ai-sustainability_thesis-project.git

From a perspective of Digital Humanities tools criticism, this thesis investigates the intersection of Artificial Intelligence (AI) and sustainability, with a focus on regulatory frameworks in the European Union (EU), the United States (USA), and China. The research aims to understand how these regions address AI's environmental and ethical implications, balancing innovation with sustainability. It employs a comprehensive methodological approach, including Natural Language Processing (NLP) techniques of LDA topic modeling, to analyze a substantial corpus of legal documents from the three regions. This analysis identifies key themes and trends in AI governance, providing a detailed comparison of their respective strategies. The findings reveal significant differences in how the EU, USA, and China approach AI sustainability. The EU emphasizes a human-centric and ethical framework, integrating sustainability into its legislative processes. The USA adopts a more decentralized, market-driven approach, focusing on innovation and voluntary guidelines. In contrast, China employs a state-centric model, embedding AI development within broader national economic growth and technological advancement strategies. The research identifies several critical issues that must be addressed to achieve sustainable AI. High energy consumption and greenhouse gas emissions from AI operations pose significant environmental challenges. The depletion of nonrenewable resources, driven by the need for rare minerals in AI hardware, further exacerbates environmental concerns. This work strongly advocates for an interdisciplinary path toward AI sustainability. This approach preassumes collaboration among experts such as environmental scientists, ethicists, legal scholars, and policymakers. It holds great potential to lead to innovative solutions that integrate technological advancements with sustainability principles. This work is a call for action crucial for the future of AI development. This thesis provides a foundational understanding of AI sustainability and underscores the need for further research and interdisciplinary efforts. The comparative analysis of the EU, USA, and China offers valuable insights into diverse regulatory approaches, identifying best practices and potential areas for improvement in global AI governance.

Keywords: Artificial Inteligence, AI governance, AI Regulation, Digital Humanities, Environmental Impact, Ethical AI, Sustainable Development

TABLE OF CONTENTS

1. INTRODUCTION

1.1 Background and Significance

Two urgent forces are shaping the future of contemporary society: the intensifying climate change of the Anthropocene and the accelerating digital advancements. These intertwined forces are triggering profound changes that imperil the well-being and prosperity of humanity and the ecological systems on which it depends. The technology of Artificial Intelligence (hereafter AI), one of the fastest-growing technology domains, is deeply embedded in many aspects of everyday life and issues. It encompasses academic research, businesses enterprises, and users. It is considered one of the most popular (Metz, 2016) and often hyped (Acemoglu, 2024) technological trends with the potential to bring about the most significant transformations in the coming years (Filippucci, 2024, p. 9).

The enormous investments have spurned this technology after a long 'AI Winter' period and led to groundbreaking applications in diverse areas (Katz, 2017, p. 2). Among others, the recent advent of Language Learning Models (LLMs) and Generative Pretrained Transformers (GPTs) revived the dreams about the 'AI Revolution', which has been conceptualized as the fourth technological revolution, following the agricultural, industrial, and computing ones (Devlin, 2023; Scharre et al., 2018). These new developments are frequently hailed as the "next big thing", promising a multitude of opportunities and advancements (*see* Ahluwalia & Miller, 2023; Nowak et al., 2018).

Nevertheless, this exponential growth in AI data, models, and infrastructure capacity is not without consequence. As AI is not merely a feature or a digital tool residing in computers or the cloud, it has tangible physical dimensions rooted in the real world. These include data centers, chip factories, and the individuals who train and use the algorithms on a daily basis. This makes AI intimately connected to the physical world, with all the implications that this entails, leading to questions about the sustainability of AI and its ramifications (Ligozat et al., 2023; Goh & Vinuesa, 2021; Van Wynsberghe, 2021). Despite the positive societal benefits, recent developments in AI have resulted in severe and large-scale environmental implications (Wu et al., 2022, p. 10). In particular, in depredation of physical, non-renewable resources like fossil fuels, minerals, and water integral to the value chains of digital technologies, which are central to this growth (Lehuedé, 2024). These resources are used to build, power, and maintain data centers that host the energy-demanding computers, processing vast amounts of data required for AI while generating significant heat managed by water circulation and electric air conditioning (pp. 1-2). The extraction of these resources has surged with the advent of the current wave of AI technologies, highlighting the growing environmental impact of its development.

The resurgence and continued growth of AI represents a systemic risk to the global communal structure, which is grappling with the increasing consequences of global warming and climate change. Recognizing these challenges and the potential benefits associated with further AI development, it is crucial to examine how public policy is addressing these issues. Accordingly, this research aims to discuss and reflect on the environmental context and assertions pertaining to of AI, with a particular emphasis on the significance of immediate, responsible, and sustainable use and development. It is of the utmost importance that public policy addresses these issues and ensure that AI is developed and used in a responsible and sustainable manner.

1.2 Stating Research Question

This thesis originates from the field of Digital Humanities (DH). Adopting a critical studies approach, it aims to reflect and examine on the role of digital tools in use (van Es, 2023). This is particularly crucial in DH research, as the digital is the material, tool, and subject of this field (Bailot et al.). This dependency can be understood as a form of solid affiliation that can be an asset, given the complexity of the concept of sustainability (Drucker, 2021). Accordingly, from an interdisciplinary standpoint, this study seeks to address the following central question: *How do AI sustainability strategies correlate with AI accountability in regulatory frameworks across the European Union (EU), United States of America (USA), and China*? The research will be conducted in two phases, as follows: (1) problematization of AI sustainability in the current world and establishing its meaning both in the legal and ethical sense, and (2) tracing the evolution of AI sustainability strategies using seeded LDA topic modeling on a corpus of AI regulatory frameworks from the regions of EU, USA, and China.

The significance of this research question arises from several factors. Primarily, these three regions have a significant stake in the AI industry. A well-known aphorism about the intertwined character of digital technologies encapsulates this dynamic: "Made in China, Designed in California, Criticized in Europe" (Lovink, 2022). As indicated in the report, "Who is winning the AI race: China, The EU or The United States?" by the Center for Data Innovation

(2019), despite China's bold AI initiative, the US leads in the race in absolute terms, with China in second place and the European Union trailing further behind. Furthermore, the USA is a leader in technological innovation, China is rapidly advancing its position with significant investments in AI, and the EU is pioneering regulatory frameworks that prioritize ethical and sustainable AI development – leading in terms of shaping the global debates of data regulation. An updated version of the report from 2021 suggests that without significant policy changes, particularly in the EU becoming more innovation-friendly and the US developing and funding a proactive national AI strategy, it is anticipated that China will eventually close the gap with the United States in AI development (Castro et al., 2021). This situation is aptly illustrated in the figure from the Galaz et al. (2021) study, showing the geographical distribution of investments made by AI companies from those regions:

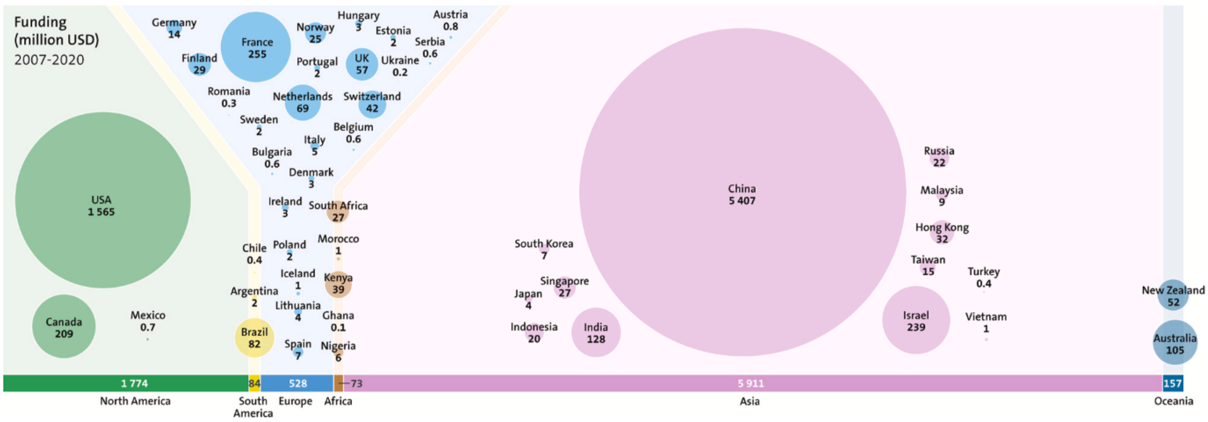

Figure 1. Global distribution of AI technologies and investments in farming, forestry and the marine/aquaculture sectors. See Galaz et al. (2021) for more details.

By comparing strategies in these globally significant regions, this research aims to provide valuable insights into diverse approaches, identify best practices, and understand the implications for global AI governance (Castro et al., 2019). As these nations operate in a globalized economy, the challenges they face are not just about meeting the demands of their growth, but about optimizing development to achieve sustainability (environmental, social, and economic). This is of particular importance in the light of global challenges such as climate change, which have a universal impact, regardless of one's involvement in the AI race.

Secondly, this work is aligned with a broader trend of recognizing costs affiliated with the more general utilization of information and communication technologies (ICTs). Additionally, it addresses the complex task of finding a balance between the pursuit of innovation, growth, and socio-ecological harms that further advancements in these fields may cause (e.g., Berkhout & Hertin, 2004; Ashford & Hall, 2011; Linkov et al., 2018). These endeavors can be encapsulated in the concept of sustainable development, which is not a novel concept but is arguably one of the most significant challenges of the modern era (Mulder et al., 2012). As an idea, it combines several global challenges that humanity is currently facing, including climate change, social inequalities, the devastation of the environment, and similar. In this research, the concept of sustainable development underscores the importance of becoming more aware of the decisions' implications and transition towards better, more optimal solutions that serve the whole evenly. Ultimately, if civilization is to continue its development, it is imperative that its progress is not reliant on finite resources and technology that is harmful to the environment.

As the Secretary-General of the UN, António Guterres, has emphasized, the transformative potential of AI for beneficial purposes is immense and urgently needed, especially in the face of the climate crisis (United Nations, 2023). The 2030 Agenda, which serves as a global blueprint for peace and prosperity on a healthy planet, is currently facing significant challanges. All technology, including AI, has a double-edged character, with the potential to both create and solve contemporary problems. For AI, one of the core technologies of the future, it is crucial that the solutions it provides outweigh the challenges. It is therefore urgent, to steer AI with the aim of ensuring its recognition as an innovation that benefits society as a whole.

This work prompts several related concerns: To what extent do existing legal instruments address the sustainable development of AI? If this is not the case, how else can further AI development can be regulated? Further investigation is required to ascertain whether AI can be sustainable. Regardless of the responses, it is the shared responsibility of stakeholders such as policymakers, researchers, and AI users to engage in introspective reflection and assume the responsibility associated with the to utilize such tools (Prescott, 2023). In this context, the objective is to avoid being compelled by technological advances to undesirable paths for global society. To define technology and not let technology define us (Lee, 2020, p. 10). Moreover, it will be argued that AI sustainability is not merely an environmental concern but an ethical imperative that is intricately linked to issues of accountability, decolonization, and the persistent struggle for fairness and justice in societies.

The research structure proceeds as follows: the second chapter provides a theoretical introduction to the current situation of AI, focusing on the socio-ecological and environmental dimensions. Subsequently, the core topics are addressed, wherein the paper examines the sustainability of AI as an aspect of AI ethics and defines the issues of AI accountability, operationalization, and governance. Chapter 3 presents a computational semantic study of AI regulations and legislation from the aforementioned regions, with explanation of how it is evaluated (Chapter 4). Subsequently, the research results are presented, summarized and discussed (Chapter 5). The final chapters, put forth suggestions for future research (Chapter 6) and provide a conclusion to this work, leaving the reader with accompanying considerations.

1.3 Motivation

Following the principles of the Vienna Manifesto of Digital Humanities (2019), which advocates for purposeful engagment with contemporary and prospective technological development, this thesis critically examines the AI industry and its context. All the technologies, including digital ones are a product, shaped by number of factors, including interests, values, power, and economic decisions, making it crucial to analyze and describe them in detail. The manifesto underscores the neecessity for more effective policy formation and governance, highlighting our collective responsibility for the present and future circumstances. It advocates harmonizing humanistic values with technological progress to foster a better society and a balanced ecosystem (Werthner, 2020).

In the meantime, the prevailing separation of technology, society, economics, mental welfare, and the environment is detrimental and impedes the generation of insights into the current challenges. Dauvergne (2020) attributes this situation to the dominance of technocratic thinking in the contemporary world, which treats all problems as discrete tasks without considering the broader global perspective. In researcher's view, an adequate assessment of the impact of AI on the world must consider not only to be aware of the various applications of AI, but also the ability to answer more general questions (Dauvergne, 2020). That proves the necessesity to develop new systematic approaches that will introduce and facilitate interdisciplinary thinking about technology (Mann et al., 2018, pp. 8-13). Such an approach must consider all these dimensions of human existence to effect a change in the very practices of designing AI techniques and systems. In this context, DH represents the foundation for such an approach, offering a more profound and comprehensive understanding of technology, the relationship between humans, machines, and their environments, and the values that emerge from this relationship.

Amidst the current enthusiasm and publicity surrounding AI, it is vital to acknowledge that AI is not a neutral technology but a tool that reflects and holds the potential to perpetuate capitalist structures and their inherent pitfalls (Sætra, 2021). AI systems represent an increased level of algorithmic and biopolitical control exerted by global capital over various aspects of individual and collective existence (Crawford, 2021, pp. 9-10). The infrastructure of AI relies heavily on material resources and human labor sourced from different regions around the globe (see Crawford & Joler, 2019), giving rise to considerable social and environmental concerns. These issues include the continuous extraction of human life through data (Couldry & Mejias, 2019) and the collective cognitive work of ICT users, whose interactions generate the metrics necessary for training AI models. In the contemprorary context of cognitive capitalism, the most significant value is generated by embodied brains operating within extensive communication networks (Moulier-Boutang, 2011). It is therefore essential that AI serves the collective good, delivering socially beneficial outcomes (Cowls et al., 2021).

The Future of Life Institute refers to this as a 'Shared Benefit Principle', which states that "*AI technologies should benefit and empower as many people as possible.*" (Conn, 2022). Concurrently, despite the notable advancement of AI systems, they did not enhance the quality of life or promote greater equality. Instead, the advent of AI has resulted in considerable environmental and social costs, including:

1. <u>Severe Energy Usage</u>: Researchers have issued a warning of AI substantial energy consumption, that is expected to increase at least three-fold in the following years, potentially leading to an energy crisis (Lannelongue et al., 2021; Ligozat et al., 2021). The International Energy Agency has projected that the electricity consumption of data centers associated with AI will soon reach a level that will match the annual energy demands of Sweden or Germany (Calma, 2024). Other studies are forecasting a ten-fold increase in energy demand by 2027 (de Vries, 2023), with AI emissions of greenhouse gases comparable to the American commercial aviation sector (Lannelongue et al., 2021).

2. <u>GHG emissions</u>: AI's reliance on vast data centers contributes notably to greenhouse gas emissions. Google, for instance, has seen a 48% rise in its GHG emissions since 2019, mainly due to the energy demands of these centers (Milmo, 2024). These emissions not

only exacerbate climate change but also undermine sustainability goals undertaken by these entities. The growing energy needs and associated emissions from AI data centers are creating substantial environmental challenges (Heikkilä, 2023). Google's efforts to achieve net-zero emissions by 2030 are hindered by the unpredictable environmental impacts of AI, reflecting a broader issue within the tech industry. Another layer to this problem is the need for more official information regarding models' carbon footprints, so their levels are likely higher than the available estimates indicate (Heikkilä, 2022).

3. Radical Depletion of Nonrenewable Resources: The manufacturing of CPUs and GPUs for AI requires rare minerals including cobalt, lithium, and rare earth elements, which are frequently extracted under environmentally destructive and unethical conditions. The extraction processes for these minerals can result in severe ecological damage, including soil degradation, water contamination, and biodiversity loss. Additionally, substantial quantities of water are utilized for cooling these components during their operation in data centers, further straining local water resources (Heikkilä, 2023; Kwet, 2020; Crawford, 2024; Ren, 2023). The reliance on these scarce resources gives raise to environmental concerns and introduces geopolitical risks, as the control over rare mineral supplies has the potential to become a source of international tension and conflict.

4. The exploitation of labor in AI development: Workers from the Global South, crucial for fine-tuning AI systems, are often paid less than $2 per hour (Perrigo, 2023). This labor includes data labeling, content moderation, and other digital piecework that are indispensable for development and maintenance of AI technologies (see Tubaro et al., 2020). These workers often encounter precarious working conditions, lack of employment security, and minimal labor rights protections (Le Ludec et al., 2023). This highlights the need for the implementation of fair labor practices and the equitable distribution of benefits derived from AI advancements.

Although the last concern does not align with this research topic and will not be further continued, it is a pressing problem with no hope of improvement anytime soon, making it important to mention here. Especially given that the centralization of AI technologies by the Global North drives their economic growth at the expense of the Global South's natural resources and labor (Bon et al., 2022, p. 62). The ethical implications of this labor exploitation are profound, as it underscores the need for fair labor practices and the equitable distribution of benefits derived from AI advancements. This dynamic not only serves to perpetuate economic disparities but also mirrors historical patterns of exploitation and extraction, akin to colonial practices. This have been termed 'Data Colonialism' by Couldry and Mejias (2019), as the power imbalance exploits data for profit. Other researchers have also discussed the concept of 'Digital Coloniality' and refer to the exclusion of the Global South from debates about the digital society that has been constructed using their resources but not for their benefit (Bon et al., 2022). It is an abuse, significantly since the economic impact of unbalanced automation and climate change will disproportionately affect these regions the most (Siddarth et al., 2021, pp. 4-6). This claim can be further verified by the map curated by Galaz et al. 2021 study:

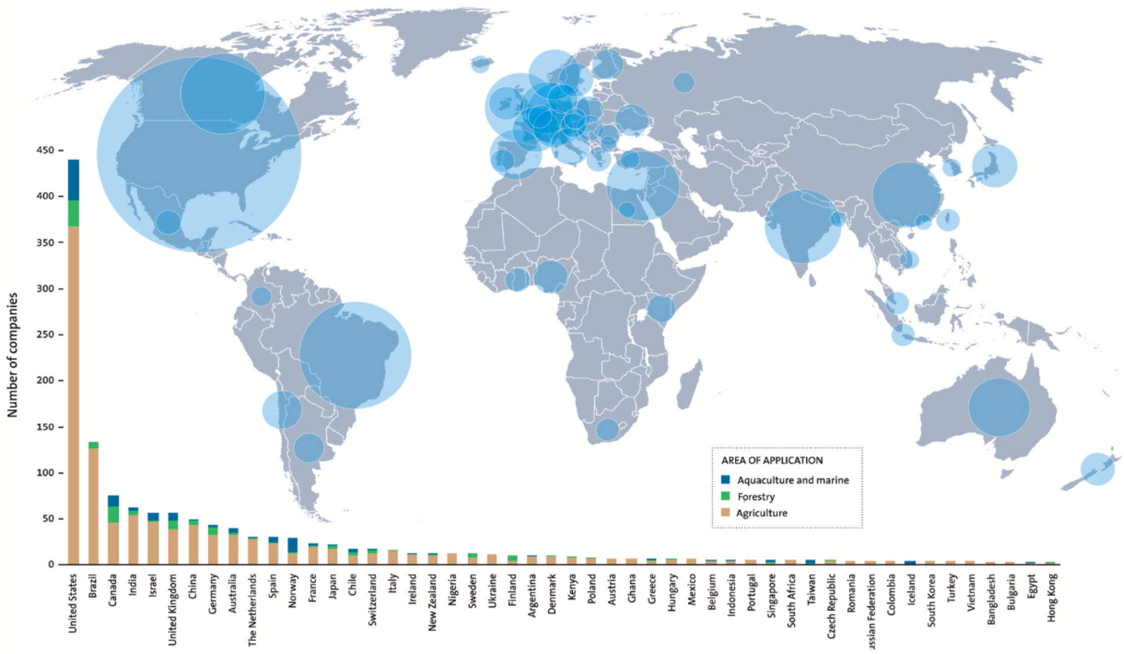

Figure 1A. Geographical and sectoral distribution of companies developing IoT applications, sensors, robotics, and AI-supported analytics for aquaculture, forestry, and agriculture. See Galaz et al. (2021) for more details.

These issues underscore the pressing need to address the broader issue of sustainable technology, which is directly linked to the urgent task of mitigating and defending against the most severe effects of climate change. The paradox is that while the development of advanced AI may offer solutions to environmental issues, unsustainable development in itself poses a significant problem (Mulder et al., 2011). It is of the utmost importance to address this challenge, as failure to do so could render other efforts futile (Bridle, 2022, p. 292). It is imperative that we gain insights into AI sustainability if we are to rethink our digital technologies and their global context. This will necessitate a new, sustainability mindset (Wu et al., 2022, p. 7).

AI has the potential to significantly enhance our socio-technical reality. The extent to which this potential is realized will depend on the governance of decision-makers. On the one hand, private entities that own and implement AI technologies, and legislators on the other, who alone have the power to shape the framework and rules within which AI will be applied. However, the role of researchers and users of these technologies is equally vital. Users, as the primary data source that powers AI, users play a pivotal role in influencing and redirecting AI development through their political and consumer choices. The collective perceptions of AI systems exert a force as powerful as the financial and strategic interests invested in this technology. Confronting the biases and assumptions inherent in mainstream narratives about AI, which often obscure its materiality, its connections to extractivist economies, and its profound impacts on vulnerable communities, is therefore of utmost importance. The integration of ethical considerations, environmental impacts, and social justice issues into the discourse on AI governance underscores the need for interdisciplinary approaches and collaborative efforts. This multifaceted nature of AI governance calls for a comprehensive solution that ensures that AI development and deployment are aligned with the principles of sustainability, equity, and inclusivity principles. Through its examination, the objective of this study is to contribute to broader efforts promoting responsible innovations that benefit humanity as a whole, simultaneously avoiding the exacerbation of existing inequalities or the infliction of harm upon the planet.

Consequently, the central objective of this study is to assess the impact of a particular form of AI governance, encompassing the prevailing institutional arrangements and legislative frameworks concerning the utilization and implications of AI technologies. The critical perspective allows for the determination of AI's impact on a range of economic, social, and political domains (Dauvergne, 2020).

2. THEORETICAL BACKGROUND

2.1 What and Who is AI?

Artificial Intelligence is an elusive term encompassing a broad range of computation systems that are designed to emulate human intelligence. Despite its broad scope, there is often confusion between AI and related fields such as big data, machine learning, and deep learning. The term 'artificial intelligence' was first introduced in 1955 by John McCarthy, who defined it as "the science and engineering of making intelligent machines" (McCarthy, 2006). However, this framing is arguably too abstract and speculative. For the purposes of this research, therefore, AI can be defined as "a technical and scientific field devoted to engineered systems that generate outputs such as content, forecasts, recommendations, or decisions for a given set of human-defined objectives" (ISO, 2022). AI is a byproduct of the ICT revolution that began in the 1970s and the broader aspiration to advance general artificial intelligence – a machine intelligence alike human intellectual capability (Rogers, 2023). Finally, AI gained a human expression, in the form of LLMs that became available to everyone in the form of an app. Nevertheless, specialists point out that even advanced AI systems like ChatGPT, although passing the Turing test, have nothing to do with actual reasoning (Biever, 2023). In practice, the AI available today is predominantly a set of managerial computational tools overseen by a limited number of entities, used to quantify and analyze data, with the primary goal of minimizing expenditures or maximizing financial returns (Cowls et al., 2021).

The "International Scientific Report on the Safety of Advanced AI" (Vasilios, 2024) offers a comprehensive overview of the current state of AI technology. The report was a collaboration among seventy-five experts from 30 countries, with the esteemed Prof. Yoshua Bengio as editor. Rather than exuding unwavering optimism, the report paints a stark picture of the current AI landscape. It highlights the exponential growth in computing power and data size as the primary drivers of progress rather than groundbreaking intellectual advances. Moreover, the report raises concerns about the lack of transparency in AI models, with even their creators unable to understand their inner workings (p. 20) fully. Instead of being marvels of modern technology, many of these systems exhibit unexpected weaknesses when deployed in real-world scenarios.

AI systems, with their unique ability to process vast amounts of data and learn from it, have the potential to perform a wide range of tasks across various domains. This learning process, which involves identifying patterns and making predictions or decisions based on the data, is a key characteristic of AI (ISO, 2022). Such systems automate tasks that are repetitive or require significant computation, thereby increasing efficiency and reducing the potential for human error. They are capable of adapting to new information and adjust their outputs accordingly, which makes them particularly well-suited for use in dynamic environments. Some applications of AI involve the use of sophisticated algorithms for the generation of natural language or the images recognition. The versatility and applicability of AI across a range of fields, including healthcare, finance, transportation, and entertainment, have contributed to its emergence as a powerful tool, generating excitement about its potential in various fields and instilling a sense of optimism for the future (ISO, 2022).

Due to the AI specificity of AI and the high resource requirements associated with it, the AI industry is centralized and owned mainly by a few international technology companies, colloquially referred to as 'Big Tech' (Kak et al., 2023). As a result, AI development is primarily driven by profit and effects in highly accurate yet resource-intensive computational models (Schwartz et al., 2020, pp. 1-5). Notwithstanding the lack of explicit recognition of sustainability and efficiency under these circumstances, the actions of workers and the public's concerns over these impacts have compelled these companies to pledge climate action and set sustainability goals. However, as proven by "Corporate Climate Responsibility Monitor 2024", these initiatives frequently fail to achieve meaningful impact and often serve more as promotional tools than as vehicles for genuine transformation (Day et al., 2024). Similarly, international standards, despite their enormous potential for developing sustainable practices in the business, often serve more as guidelines due to the nascent stage of AI technology. They are used only to obtain credibility in the eyes of the public (Gupta, 2024).

As AI technologies continue to emerge and evolve, so too does the regulatory framework that seeks to govern them. At present, the principal legal obligations pertaining to cutting-edge systems are borne by their manufacturers. It is thus incumbent upon these suppliers to design, test, audit, and certify these systems to ensure that they meet the requisite standards of transparency, explainability, and accountability (Wong, 2020, pp. 48-50). However, in recent years, major technology companies have become prominent and influential actors in the policy sphere. Because of their influence and vast resources have acquired quasi-autonomous and quasi-sovereign status, enabling them to wield significant global authority and independence (Khanal et al., 2024, p. 12). A review of the past two decades of Big Tech history reveals that this concentration of profit-oriented power has often had a detrimental impact on society. It has taught the global public that governments cannot afford to remain passive observers of Big Tech's growing influence (Alegre, 2024). Both EU and American policymakers are aware that such unchecked growth disrupts democratic institutions and endangers the liberal order of the Western world (Saran & Mattoo, 2022). At the same time, Europe is assuming a leading position among the three major AI stakeholders in challenging this ominous outlook effectively. This can be explained by the concept of the 'Brussels effect' (see Bradford, 2020), but also because the regulatory landscape in the United States and China is further complicated by the intensifying 'AI race' between these two countries (Cuéllar & Sheehan, 2023).

The development of effective AI regulation requires a collaborative approach. This approach is crucial in formulating meaningful regulations that address the environmental impacts of AI. The assurance of sustainable development of AI requires the implementation of both technical and non-technical measures, including the establishment of sustainable AI practices and the formulation of comprehensive legal frameworks. The regulation of AI, a component of ICT technologies, is still an ongoing process, particularly with regard to the issue of algorithmic accountability. Nevertheless, the focus of this thesis is not to examine the AI regulatory landscape *per se*, but rather to investigate the environmental implications of AI use and the question of accountability in this context. It is important to note that the responsibility for ensuring environmentally sustainable practices in deploying and developing AI technologies ought to rest with its developers and providers. Thus, while recognizing the ongoing efforts and complexities in AI regulation, this research aims to highlight the pressing need for such practices. It emphasizes the urgency and importance of a collaborative approach and the need for immediate action.

2.2 AI and Sustainability

The concept of sustainability is employed in a multitude of contexts, encompassing a vast array of ideas, from commercial application to scientific pursuits. In the scope of this research, the concept of 'sustainability' is closely linked to the field of ecology, which studies the interrelationships between diverse elements and the cognitive representation of all utilized technologies. This entails grasping the function, influence, and significance of these systems within the surrounding environment (Bridle, 2022, p. 33). The contemporary AI ecosystem is characterized by a capitalistic environment in which the pursuit of profits, the attainment of victory, the assertion of control, and the establishment of dominance are of paramount concern. Nevertheless, this potential exists for this technology to develop different manner if only the ecology is addressed and altered (p. 83). As evidenced by media studies, cybernetics, and philosophy, technology is performative in its nature. Therefore, a shift in attitudes—a new ecology for technology—would prove advantageous both the digital and environmental ecosystems (Crawford, 2021, pp. 179-180).

To comprehend how AI applies to sustainable development, it is first necessary to define the term "sustainable." According to the World Commission on Environment and Development (commonly referred to as the "Brundtland Report"), sustainable development can be defined as "meeting today's needs without compromising the ability of future generations to meet their needs" (Keeble, 1988). It is also essential to distinguish between the utilization of AI for addressing environmental, ecological, and social challenges (referred to as 'AI for Sustainability') and the sustainable development and use of AI systems themselves (referred to as 'Sustainability of AI'). The former involves applying AI tools and techniques to identify innovative solutions to specific issues, whereas the latter priorities the reduction of the environmental impact of AI (van Wynsberghe, 2023). The term' Sustainable AI' (SAI) can be considered an umbrella term encompassing both approaches, with the caveat that the development of AI for sustainability is not possible without the sustainability of AI.

From an ecological perspective, it is crucial to acknowledge that all technologies, including AI systems, are part of the same interconnected ecosystem, which spans social, economic, and environmental dimensions (Bridle, 2022, pp. 19-22). A grasp of these interconnections is key to identifying potential impacts and unintended consequences. The sustainable AI development requires the efficient management of resources and the minimization of waste. This includes, among others, using renewable energy sources and reducing the environmental footprint of AI systems. Addressing the ecological impact of AI requires a more holistic approach that considers the entire lifecycle of AI systems and their impacts on their environment, from development and deployment to disposal and recycling.

The environmental impact of AI is multifaceted phenomenon, involving energy consumption, resource extraction, and waste generation. AI systems, especially those relying on deep learning and large-scale models, require substantial computational power. This, in turn, necessitates a considerable input of energy resources. For instance, the training of single AI model has been found to emit as much as five cars over their lifetimes (Strubell et al., 2019, p. 1). Moreover, the energy cost estimates of making an image with generative AI are equal to fully charging a smartphone (see Luccioni et al., 2024). Additionally, data centers currently account for approximately 1% of global electricity demand, which is expected to increase with AI's further adoption (Jones, 2018). In an analysis, Alex de Vries of Vrije Universiteit Amsterdam (2023) estimates that the AI sector could consume between 85 and 134 terawatt-hours (TWh) per year from 2027, which is as much as the annual demand of the Netherlands, Argentina, or Sweden. This figure can also be compared to the total energy consumption of data centers, the physical base of the Internet from 2022, which amounted to about 460 TWh. Therefore, the energy consumption of data centers, which power AI systems, is a growing concern. It is one of the main reasons for industry giants' interest in nuclear energy (Aldrete et al., 2024).

Electricity is becoming as valuable and hard-to-access resource for companies working on AI as the computing power of the chips that constitute its backbone – it is the huge demand for these components that has pushed up the stock prices of their leading manufacturer (Kim, 2024). The extraction of nonrenewable resources essential for this and other AI hardware, such as rare earth elements, cobalt, and lithium, has profound environmental consequences. The mining of these materials often leads to habitat destruction, soil and water contamination, and considerable greenhouse gas emissions (Zapp, et al., 2022, pp. 267-268). As AI technologies continue to develop, the demand for these resources is likely to increase, thereby exacerbating associated environmental and social issues. The disposal of AI hardware also presents a significant environmental challenge. Electronic waste is one of the fastest-growing waste streams globally, with AI hardware contributing to this problem significantly. E-waste contains a variety of hazardous materials, including lead and mercury, which can leach into the environment and pose risks to human health and ecosystems (see Heacock et al., 2016). The implementation of effective recycling and disposal methods is crucial to mitigate these impacts. However, the current global e-waste recycling rate is alarmingly low, with only about 20% of e-waste being formally recycled (Baldé et al., 2024, pp. 10-11).

Addressing AI's sustainability also requires a broader understanding of its societal implications. The deployment of AI in various sectors offers considerable opportunities for innovation and efficiency, but it also gives rise to concerns regarding social equity and access. If access to AI technologies is not equitably distributed, it may also widen the gap between technologically advanced and less developed regions (see Alonso et al., 2020). AI systems have the potential to not only transform existing industries but also to create new economic opportunities. It is, however, important to note that the benefits of AI are not distributed uniformly across all populations and regions. The advantages conferred by AI may be concentrated among those with the resources and infrastructure to develop and deploy these technologies, which could result in greater economic disparities. This imbalance is particularly pronounced in the global North-South divide, where access to advanced technologies is often correlated with existing socioeconomic inequalities (see Yu et al., 2023).

The societal dimensions of AI sustainability also encompass the ethical implications of AI decision-making, which must be considered in order to ensure the responsible development of AI. It is therefore evident that integrating care into AI sustainability is not merely a recommendation; it is a necessity. This entails cultivating a culture of responsibility and ethical awareness among AI developers, policymakers, and users. By promoting fairness and justice in the development and deployment of AI, ensuring that benefits are distributed equitably and that marginalized communities are not disproportionately affected, these principles can be incorporated into AI development to help create a more sustainable and equitable technology ecosystem.

In conclusion, sustainable AI encompasses the application of this technology to address environmental and social challenges, as well as and the sustainable development and use of AI systems. AI's environmental and societal impacts are significant and multifaceted, requiring a comprehensive and interdisciplinary approach to ensure effective mitigation. The integration of principles of care, fairness, and justice into the development of AI will facilitate the creation of a more sustainable and equitable technology ecosystem that benefits all members of society and preserves the environment for future generations. This holistic strategy guarantees that the advancement and implementation of AI are in accordance with the overarching ecological and societal objectives, thereby contributing to a more balanced and sustainable future.

2.3 The Legal Landscape of AI Regulation

At present, the principal legal obligations associated with high-risk AI systems are borne by their suppliers (Dunlop et al., 2023). However, this responsibility extends beyond mere compliance with technical standards, encompassing such aspects as transparency, explainability, and accountability. The complexity of AI systems, trade secrecy, and technical opacity pose significant challenges for regulators. Nevertheless, the prospective advatages of effcacious AI regulation, such as increased transparency and accountability, may culminate in a more trustworthy and beneficial AI environment, offering a promising future for AI. It is a probable that this crisis will persist unless there is a radical transformation in the manner in which AI is regulated.

In addressing the challenges of AI regulation, some scholars have proposed the implementation of licensure schemes in data-driven industries. These schemes, which require compliance with specific standards before the expansion of large-scale data practices, effectively transfer the responsibility from underfunded regulators to the industry (Malgieri & Pasquale, 2024, pp. 11-15). This guarantees that companies adhere to rigorous standards from the outset, which is a crucial element in the effective regulation of AI. Furthermore, a comprehensive regulatory framework for AI must encompass various concerns, including those related to privacy, security, bias, and transparency. AI systems should be designed to clearly explain their decision-making processes to build trust and enable users to challenge the outputs produced by the AI. Additionally, a comprehensive regulatory framework for AI must encompass rigorous testing for fairness and protection of user data.

At the same time, AI has acquired strategic importance for governments across the globe, facilitating accelerated and more effective deployment of this technology for its advantages. The high competitiveness and the boom in the industry are causing the legal landscape of AI regulation to evolve constantly worldwide – with significant efforts from the leading regions of the EU, USA, and China. The pursuit of favorable outcomes with regard to AI has become a matter of interest and strategy for all of them. In 2017, the State Council of the People's Republic of China published a document titled "New Generation Artificial Intelligence Development Plan," which articulated the national aspiration to become a global leader in the AI industry by 2030. This plan attracted considerable interest and became the basis of the concept that AI development is a vast geopolitical competition whose winner will control the future (Lucero, 2019, p. 102). In 2018, the European Commission published the European strategy for AI. The aim is to transform the EU into a world-class AI center, ensuring that it is reliable and prioritizes human well-being (European Commission, 2018). In 2019, the US government unveiled a national strategy to maintain American leadership in the AI field (Executive Office of the President, 2019). The strategy emphasizes the allocation of federal resources toward the advancement of AI innovation, with the objective of fostering economic prosperity, enhanced national security, and an improved quality of life.

Today, many developed countries have formulated strategies for the advancement of AI, each emphasizing the strong points of AI solutions and their potential growth factors. However, as these countries strive to lead in benefits and profits connected with this technology, the environmental risks involved in using and implementing AI must not be relegated to secondary concerns. Nevertheless, it is challenging to envision meaningful international legal collaboration on a topic that needs to be regulated at the regional level.

The European Commission has made significant progress in formulating a comprehensive regulatory framework for AI within the European Union. The proposed AI Act seeks to address the various risks associated with AI by classifying AI systems into different categories based on their respective risk levels. This tiered approach permits the implementation of proportionate regulation, thereby ensuring that higher-risk AI systems are subject to more rigorous requirements. The AI Act underscores the importance of transparency, requiring AI developers to provide clear information about their systems' capabilities and limitations. Additionally, it mandates rigorous testing and validation procedures to ensure that they operate fairly and without bias (European Commission, 2021).

In contrast, the USA, has focused on fostering innovation through a more decentralized approach. The federal government has established guidelines and principles to steer AI development, emphasizing the need for transparency, accountability, and fairness. However, the regulatory landscape in the USA remains fragmented, with different states enacting their AI-related laws and regulations. The decentralization can lead to inconsistencies and gaps in the regulatory framework, potentially undermining efforts to address the broader societal impacts of AI (Executive Office of the President, 2019).

China’s approach to AI regulation is distinguished by a pronounced emphasis on state control and strategic planning. The Chinese government has implemented comprehensive set of policies designed to foster the development of AI, including substantial investments in research and development and the establishment of AI research centers. China's regulatory framework for AI is designed to ensure that AI technologies are aligned with national priorities, including economic development and social stability. The government's top-down approach allows for swift implementation of regulations. However, it also raises concerns about state surveillance and the potential for AI to be used in ways that infringe on individual freedoms (Andersen, 2020).

Notwithstanding the aforementioned disparate methodologies, there is an emerging acknowledgment of the necessity for international cooperation in the domain of AI regulation. International standards, as a form of global governance, are already in place and can facilitate the attainment of AI policy objectives. They play a crucial role in promoting good governance practices when necessary, particularly by disseminating beneficial systems and practices, fostering trust among states and researchers, and encouraging efficiency in the development of AI systems. Moreover, the transnational nature of AI technologies indicates that challenges related to privacy, security, and fairness cannot be adequately addressed by individual countries in isolation. International organizations, including the United Nations and the Organisation for Economic Cooperation and Development (OECD), are engaged in the process of establishing global standards for AI. These standards aim to harmonize regulations across countries, promoting best practices and facilitating AI technologies' safe and ethical development.

Another promising avenue for international collaboration is the development of AI ethics guidelines. Several countries and organizations have published guidelines that set forth ethical principles for AI development and use, with similar principles (Jobin et al., 2019). These guidelines typically place emphasis on the principles of transparency, accountability, fairness, and respect for human rights (Floridi et al., 2018). By aligning national regulations with these ethical principles, countries can ensure that AI technologies are developed and deployed in manner that benefits society as a whole. The importance of continued dialogue and cooperation among countries in this regard cannot be overstated, underlining the urgency of the matter.

However, the realisation of meaningful international collaboration on AI regulation is fraught with difficulty. Differences in political systems, economic priorities, and cultural values can impede efforts to establish common standards. Additionally, the rapid pace of advancement in AI means that regulations must be continually updated to align with technological advancements. Despite these challenges, sustained dialogue and collaboration among countries are essential to address the intricate and evolving landscape of AI regulation.

In conclusion, regulating AI is a multifaceted and dynamic field that requires balancing fostering innovation and protecting societal values. Effective AI regulation must encompass transparency, accountability, fairness, and privacy, thereby ensuring that AI systems are developed and used in a responsible manner. While different countries have adopted varying approaches to AI regulation, there is a growing recognition of the need for international cooperation to address the transnational challenges posed by AI technologies. By collaborating, countries can establish a regulatory framework that facilitates the secure and ethical development of AI, benefiting society at large and ensuring that the potential of AI is utilized for the collective benefit.

2.4 Accountability in Terms of AI

Accountability, the cornerstone of AI governance, is intricately linked to the responsible design and use of AI (Henriksen et al., 2021, pp. 574-575). The pursuit of AI sustainability is similarly tied to the challenge of AI accountability (Sætra et al., 2021). Therefore, AI accountability is pivotal in ensuring sustainable AI development and responsible deployment. This involves assigning legal responsibilities to various stakeholders and the establishment of transparent standards and guidelines. Regulators bear a crucial role in balancing innovation and credibility, shaping a regulatory framework (legal and ethical) that promotes innovative AI with high security and safety (Stuurman & Lachaud, 2022, p. 2).

The concept of accountability is compound, with the context playing a significant role. It can be conceptualized as a mechanism for explaining and justifying conduct or as a virtuous behavior to act in a certain way (Bovens, 2010). Accountability in terms of state-of-the-art AI involves explaining and justifying conduct, typically through relational mechanisms where an agent is held accountable by another (Bovens et al., 2014). In the context of AI, accountability must address the environmental impact of AI systems and ensure that responsible parties are identifiable and answerable for their actions (Novelli et al., 2023, pp. 3-4). Nevertheless, this is contingent upon three critical conditions: firstly, the recognition of authority; secondly, the interrogation of actions; and thirdly, the limitation of power (p. 4). Given the environmental impact of modern AI systems, it is imperative to establish a system for attributing accountability that would enable the mitigation of further harm caused by AI.

Accountability in terms of AI hinges on the expectation that those tasked with its design, development, and deployment will adhere to established standards and legislation. This is of paramount importance for the seamless functioning of AI throughout its entire lifecycle, which includes its development, implementation, and eventual obsolescence. In this context, accountability refers to the responsibility assumed by those involved in developing and deploying AI technologies. However, the absence of effective management of these issues can lead to two undesirable outcomes: accountability gaps, in which no individual is held accountable, and accountability surpluses, where procedures are inefficiently accumulated (Bovens, 2007, p. 958). To circumvent these potential pitfalls, effective governance frameworks must be established to ensure compliance with sustainability standards. Such frameworks must include explicit directives concerning energy consumption, resource allocation, and environmental impact assessments.

Regulatory bodies should be empowered to penalize non-compliance and incentivize sustainable practices. Given the global technology sector's concentrated power, corporate accountability is crucial (Martin, 2019, pp. 837-838). Technology companies must be transparent about their AI development processes, energy usage, and sustainability initiatives (Theodorou & Dignum, 2020). Such transparency should include detailed reports on the environmental impact of their AI systems and the steps they are taking to mitigate these effects, fostering trust and accountability.

International standards and guidelines are also crucial in establishing AI accountability. Notable organizations such as the International Organization for Standardization (ISO) and the Institute of Electrical and Electronics Engineers (IEEE) have established standards and guidelines to ensure the ethical and responsible development of AI technologies. To illustrate, the ISO 14001 Environmental Management standard requires regular and intensive audits for certification. Attainment of this certification, in turn, provides reputational value upon the companies that obtain it (ISO, 2015). These standards could further provide a framework for evaluating AI systems' performance, safety, and ethical considerations, helping to create a consistent and reliable approach to AI accountability, primarily if they would be enforced and required by governments (see Jobin et al., 2019). At the same time, national initiatives are likely to focus on international standards bodies to secure a global market share for their national champions. The elevation of national standards to an international level benefits national firms that have already developed compliant systems. Additionally, the incorporation of corporate patents into international standards can result in substantial financial gains for both the firm and its home country.

The question of the optimal means of addressing the issue of AI accountability is a matter of ongoing debate. However, it is clear that this is a crucial step in achieving the goal of effective AI governance. The establishment of clear guidelines and robust accountability mechanisms is of utmost importance. These measures will help ensure that AI technologies are developed and deployed responsibly, thereby minimizing negative impacts and fostering long-term societal stability. By adhering to these recommended practices, stakeholders can build trust and ensure that AI technologies contribute to the goal of sustainable development in a responsible manner, thereby instilling confidence in the responsible deployment of AI technologies.

2.5 Sustainable AI as Part of AI Ethics

While AI systems and algorithms undoubtedly present a number of threats and ethical challenges (Stahl et al., 2022), it is crucial to consider their potential benefits. Such benefits include enhanced efficiency, optimized decision-making, and the ability to address complex societal problems. The cases of processing personal data, omnipresent surveillance, and replicating social biases are the subject of considerable controversy. Nevertheless, if these technologies are designed and used in an appropriate manner, they have the potential to facilitate substantial positive transformation. As Fukuda-Parr and Gibbons (2021) observe, "If designed or developed in a wrong way or wrongly used, they may be highly destructive both for individuals and society" (pp. 32–44). Therefore, it is crucial to integrate ethical principles that provide valuable guidance to developing and deploying AI systems, especially in the absence of a regulatory response to many industry issues (Metcalf et al., 2019, p. 2).

The concept of Sustainable AI represents a novel development in the field of AI ethics. It refers to AI that is not only environmentally sustainable but also socially and economically beneficial. This implies that AI systems must be designed and used to minimize their environmental impact, contribute to social well-being, and support economic growth. For AI to be genuinely sustainable, it must be integrated into innovative economic models that reduce consumerism, restore natural ecosystems, and guarantee decent employment conditions. In other words, it must be ethically valid. This underscores the necessity of establishing an ethical framework for AI development, as the lack thereof presents a conceptual and organizational challenge for the industry (p. 12).

Researcher Aimee van Wynsberghe (2021) delineates the evolution of AI ethics reflection, which has occurred in two distinct waves. The initial phase concerned the potential dangers of general artificial intelligence and superintelligence. The subsequent phase we are currently in has shifted its focus to more tangible issues, such as algorithmic bias, privacy threats, and the necessity for explainable AI. This shift in focus has led to the emergence of documents such as the European Commission's guidelines (2019) and the OECD's principles

(2019), which define values for the development of responsible AI. Van Wynsberghe posits that it is time for a third wave of discourse, focusing on the intersection of climate change and AI's environmental impact. This perspective maintains that sustainable AI must consider both the benefits it enables and the external costs it generates. As a general-purpose technology, AI has the potential to impact a multitude of sectors, thereby exerting a considerable environmental impact (Hotte et al., 2024). This third wave of discourse is of crucial importance for the future of AI ethics and sustainability.

The current wave of AI ethics emphasizes the need for transparency, accountability, and fairness. These principles are critical in ensuring that AI technologies do not perpetuate or exacerbate existing inequalities, providing reassurance about the efforts to ensure ethical AI development. However, achieving Sustainable AI requires extending these principles to include environmental considerations. This involves assessing the environmental footprint of AI systems throughout their lifecycle, from development and deployment to disposal and recycling.

Despite the recent global rise of ethical initiatives, principles, and frameworks of various kinds, originating not only from the academia and international organizations but also the tech industry (Munn, 2023, p. 4), it remains challenging to envisage a future in which ethical AI will be guaranteed (Mittelstadt, 2019, p. 2). In the absence of formal legal status, these principles and frameworks cannot be considered legally binding in and of themselves. The researcher, Luke Munn, goes even further and argues that AI ethics are ultimately futile. As there is a lack of ethical training in education and the industry, AI development occurs in an "a-ethical" space where ethical dilemmas are not considered (p. 3). Additionally, he suggests that the extant ethical frameworks that concentrate on digital products or services neglect to address the underlying social issues and inequalities that inform technological development. Consequently, they may prove a dangerous diversion, divesting resources from more productive approaches to addressing the consequences of AI technologies (p. 1). Munn believes that the current focus on AI ethical principles needs to be revised. However, he proposes a solution by broadening the scope of AI justice considerations and focusing on concrete issues and solutions (p. 2). Furthermore, he argues that engaging multiple stakeholders and redefining governance structures to include both conventional managerial hierarchies and grassroots initiatives aimed at addressing the harms of AI technologies can facilitate their transformation to better serve societal interests (p. 6).

Ethical values are intended to be universally applicable, and thus can serve as a valuable indicator of appropriate conduct in the context of the international challenge posed by unsustainable AI. On the other hand, this universality renders them exceedingly broad and frequently too abstract to be effectively operationalized in practice (Whittlestone et al., 2019). Nevertheless, they serve as an excellent starting point for further discussions and plans. It is clear that, in the absence of a substantial alteration to the regulatory frameworks, ethical principles will continue to be regarded as competitive rather than cooperative in defining AI development values. This underscores the urgent need for effective governance, which would oversee the implementation of AI ethics in the private sector, facilitate the development of sustainable impact pathways, and transform ethics into an integral organizational process rather than a mere technological solution (Mittelstadt, 2019, pp. 9-10).

Governance frameworks for Sustainable AI should include mechanisms for monitoring and enforcing compliance with ethical standards. This could involve creating independent oversight entities with the mandate to audit AI systems and impose sanctions for non-compliance. Such bodies would serve to guarantee that AI developers and users adhere to the ethical principles and that the societal and environmental impacts of AI technologies are subject to continuous assessment and addressed as necessary. Education and training are also critical components of promoting Sustainable AI. By providing professionals with the knowledge and skills to develop and deploy AI technologies responsibly, educational programs can help bridge the gap between ethical principles and practical implementation. This entails integrating sustainability into AI curricula, fostering interdisciplinary collaboration between AI experts and environmental scientists, and encouraging a culture of ethical awareness in AI research and development. The intersection of AI and sustainability also highlights the necessity for stakeholder engagement. Policymakers, industry leaders, researchers, and civil society must collaborate to create a shared vision for Sustainable AI. This collaborative approach can help identify best practices, develop comprehensive regulatory frameworks, and ensure that the advantages of AI are distributed equitably across society.

In conclusion, Sustainable AI is an essential aspect of AI ethics that requires a holistic approach integrating ethical principles into developing and developing AI technologies. By addressing the environmental impacts and ensuring ethical practices, stakeholders can ensure that AI contributes positively to societal and environmental well-being. However, achieving Sustainable AI requires continuous education, effective governance, stakeholder engagement, and international collaboration. By fostering a culture of responsibility and ethical awareness will enable the optimal exploitation of AI to facilitate a more sustainable and equitable future for all.

2.6 Problematic Nature of Governance

The rapid incorporation of AI technologies has given rise to a new spectrum of regulatory and implementation challenges for the public sector. Those in public management, administration, and civil service are tasked with the complex decision-making process of balancing the transformative potential of these technologies against their possible adverse effects. Nevertheless, the governance of AI is crucial for realizing AI's potential to enhance socio-technical reality. Effective governance involves balancing the interests of private entities, legislators, and society at large in order to ensure that AI development aligns with democratic values and SDGs. The responsibility for this decision-making process ultimately rests with the relevant decision-makers. On the one hand, this includes private entities that own and implement AI technologies. On the other hand, it also encompasses legislators, who posses the sole authority to shape the framework and rules within which AI will be applied.

In this regard, Anu Bradford (2023) identifies three major competing paradigms for regulating digital markets worldwide. She refers to these as "Digital empires": the American, based on the primacy of the free market; the Chinese, where the state plays a dominant role; and the European, rooted in human-centric concern over rights. In her view, the failure of the American techno-libertarian utopian approach, coupled with the authoritarian characteristics of Chinese technologies, which, despite being privately owned, are Chinese Communist Party surrogates (p. 8), presents the most compelling argument in favor of a shift in the global development of AI towards a model that prioritizes human rights. This perspective maintains that regulatory intervention is necessary to guarantee the perseverance of democratic structures and the fair distribution of benefits (p. 9). It is noteworthy that all of these models concur on the necessity to regulate and sustain the digital domain. To illustrate this point, Bradford delineates two ongoing conflicts in the digital economy: the horizontal, which manifests as a competition between governments, and the vertical, which occurs between governments and technology companies. In addition to these rivalries, the three digital empires compete for "technological supremacy" (p. 21).

Regardless of the outcome of these conflicts, the objective of AI governance should encompass both normative issues and detailed solutions that could facilitate the operationalization of specific ethical and legal requirements, along with the implementation of sustainable AI in both the public and private sectors (Ashford & Hall, 2011, pp. 271-272). A new framework is required in order to achieve sustainability. The framework must prioritize social and environmental issues, mitigate environmental damage, and optimize resource efficiency using a multi-layered approach that includes sustainability-driven design choices (Yigitcanlar et al., 2020). Additionally, it is feasible to disassociate AI from its capitalistic framework and strive for a balanced system, aligning AI technology with sustainable development objectives. Following the proposals from Critical Raw Materials for Strategic Technologies and Sectors (2020), one can deduce that the balance between humans and AI must be restored by prioritizing humanity and planetary ecology. This entails the promotion of creativity and autonomy, as opposed to the perpetuation of technology's domination.

The ultimate objective is to cultivate conscious, digitally proficient citizens aware of their rights and the challenges of the modern era. However, this necessitates the support of public institutions, which play a crucial role in providing an appropriate legal framework, developing mechanisms to stimulate innovation adoption, and setting an example by commissioning AI solutions that meet established requirements. Their role is pivotal in the process, and their active involvement is necessary to implement AI governance successfully. As a variety of technical and non-technical operationalization methods are employed to fulfill the requirements of sustainable AI, an interdisciplinary approach is essential to convey the audience feel the importance of diverse perspectives in the process. It facilitates smooth collaboration among policymakers and specialists from various fields, including programmers, lawyers, ethicists, and managers responsible for individual projects and entire organizations.

The effective governance of AI requires a balanced approach that prioritizes social and environmental concerns. By integrating sustainability into regulatory frameworks and promoting interdisciplinary collaboration, stakeholders can ensure that AI technologies contribute positively to societal and environmental well-being. This balanced approach is crucial, as it allows for considering all aspects of the issue, leading to more comprehensive and effective governance. Public institutions must provide legal and regulatory support in order to foster responsible and sustainable AI development. Soft law mechanisms and the involvement of civil society can complement formal regulations, creating a comprehensive governance framework for AI.

2.7 Operationalizing Sustainability

The operationalization of sustainability in AI is not merely a theoretical concept but a crucial step in ensuring that technological advancements contribute positively to environmental, social, and economic objectives. This involves creating legal and procedural frameworks that make sustainability principles actionable and enforceable. An examination of the strategies used in different regions can highlight the strengths and limitations of each, offering insights into best practices. Therefore, comparing approaches from the EU, USA, and China is a key step in understanding each region's different strategies and challenges in embedding sustainability into AI development.

Holistic analysis, equity, adaptability, and multi-level governance are vital criteria for sustainability identified by Sandy Gaines. These criteria provide a framework for assessing the sustainability impact of various legal instruments and can assist in addressing energy-climate problems (Gaines, 2020, p. 1). Holistic analysis involves evaluating the entire lifecycle of AI technologies, from development to disposal (p. 14). Every significant policy choice should be subjected to a holistic analysis to ensure it leads to sustainable development on a global scale. Equity works like a compass and ensures that sustainability efforts promote fairness and justice, addressing inequalities and considering the impacts on marginalized communities (p. 15). Adaptability involves developing flexible policies that respond to changing circumstances, while multi-level governance promotes coordination among different levels of government and a range of stakeholders (p. 16).

In order for sustainability to serve as a condition for technology, it must be operationalized in a manner that entails authority and can create legal certainty, predictability, and bindingness. This challenge stems from the sustainability's inherently broad and often abstract nature, making it difficult to achieve uniform and predictable outcomes. The question thus arises as to whether policy instruments can define sustainability in operational terms, thereby rendering it accessible via a set of replicable procedures and measures (Friederich & Symons, 2023, p. 61). Therefore, the concept must be translated into specific, actionable steps that can be consistently applied across different contexts and stakeholders. The task is to operationalize the concept of sustainability AI for use in public sector decision-making.

To give preliminary examples of how the nations being subjects of this paper are tackling this task: the European Union (EU) embeds sustainability as a core principle in its

Treaties, specifically Articles 3 TEU and 11 TFEU, which commit the Union to sustainable development and mandate the integration of environmental protection into all policies and activities (EUR-Lex, 2016). The Renewed EU Strategy for Sustainable Development outlines eight overarching principles; however, as the designation implies, these principles remain principles and require further operationalization. Furthermore, the EU has made significant strides in operationalizing sustainability through its Better Legislation package, which includes guidelines and a toolbox designed to integrate sustainability into policymaking. This package provides a structured approach to ensure that all EU legislation considers sustainability, offering practical tools and methods for assessing and implementing sustainability measures (EUR-Lex, 2021). For instance, the guidelines delineate the procedures for integrating sustainability into decision-making processes, thereby ensuring a comprehensive evaluation of environmental, social, and economic impacts.

Contrasting with the EU's regulatory and policy-driven approach, the USA and China adopt disparate strategies toward AI sustainability, shaped by their unique political, economic, and cultural contexts. The US primarily focuses on innovation and market-driven approaches, with less emphasis on stringent regulatory frameworks than the EU. American AI policies tend to encourage private sector innovation and competitiveness, often relying on voluntary guidelines and industry standards to address sustainability concerns. For instance, the American AI Initiative emphasizes the significance of AI leadership while promoting ethical standards and public trust (Executive Office of the President of the United States, 2019). However, the voluntary nature of these guidelines has resulted in the lack of uniformity in the adoption of sustainability measures across the industry. The National Institute of Standards and Technology (NIST) has developed frameworks to guide AI development, including aspects of sustainability. However, these frameworks are not legally binding (NIST, 2019).

China's approach to AI development is state-centric, with incorporation into its broader national economic growth and technological advancement strategies. The Chinese government's Five-Year Plans and AI development strategies explicitly include goals for environmental sustainability and resource efficiency (China AI Development Report, 2018). China's approach involves direct government intervention and regulation, accompanied by specific mandates for companies to adhere to sustainability standards. The country's emphasis on state control allows for more centralized and coordinated efforts to integrate sustainability into AI development, thereby underscoring the crucial role of the public sector in this process.

For instance, China has made significant investments in green technology and renewable energy sources to power its AI infrastructure, aiming to reduce the environmental impact of its rapid technological growth (Sheehan, 2024).

The operationalization of sustainability in AI necessitates the establishment of a procedural framework that integrates sustainability into all technology development and governance aspects. The EU's approach provides a robust regulatory framework but challenges in terms of achieving predictability and binding outcomes. The US's market-driven approach encourages innovation but needs uniformity in sustainability practices. China's state-centric model ensures centralized control and coordination, but this must be balanced with the need to accommodate rapid growth while maintaining sustainable practices. These comparative analyses underscore the pressing need for a harmonized global regulatory framework to address the transnational nature of AI technologies and their sustainability impacts. By adopting a holistic approach, promoting equity, ensuring adaptability, and implementing multi-level governance, it is possible to create a more predictable and binding framework for sustainability, therebyaligning AI technology with environmental and societal goals.

.

3. METHODOLOGY

A significant number of countries have issued regulations and guidelines on the governance of emerging AI systems. However, it is believed that only a few explicitly emphasize environmental sustainability. To delve deeper into this issue and address the pivotal research question, a comprehensive study will be conducted to investigate the AI governance strategies in the EU, USA, and China in the context of AI sustainability.

The United Nations' Sustainable Development Goals (SDGs), a result of global cooperation, present a comprehensive framework for tackling worldwide issues such as poverty, inequality, and climate change. These goals underscore the significance of sustainable development across social, economic, and environmental aspects – or as the UN states, they constitute "a blueprint for our common future" (United Nations, 2023). Abovementioned countries and regions, as members of the United Nations, have all embraced the 2030 Agenda for Sustainable Development (United Nations, 2015), a clear indication of their commitment to achieving their goals. To evaluate the progress made by these countries in tackling the AI and sustainability issue, a computationally assisted semantic analysis of the legal documents of the European Union, the United States of America, and the People's Republic of China will be undertaken.

The processing of legal texts for semantic analysis presents a challenge due to the specific characteristics inherent to legal documents, such as their length and the use of specialized vocabulary. Moreover, the size of the corpora, comprising a substantial corpus of legal documents from various jurisdictions, necessitates a blended approach to research. This approach combines digital tools and techniques with supervised interpretation of the results, which is essential for uncovering novel findings and insights. This thesis underscores the necessity of selecting sustainable solutions in connection with digital technologies. Therefore, all computing operations will be carried out on software tools that do not require significant computing resources and do not employ pre-trained LLM techniques, that are energy intensive.

3.1 Corpora

After collecting the data, a dataset was created. It was conducted based on OECD AI Policy Observatory portal (https://oecd.ai/en/), IAPP Global AI Law and Policy Tracker (https://iapp.org/resources/article/global-ai-legislation-tracker/) and White&Case AI Watch:

Global regulatory tracker (https://www.whitecase.com/insight-our-thinking/ai-watch-global-regulatory-tracker#home). As the selection varies across the sources, documents were manually picked based on their relevance to this research topic. Therefore, they relate to matters associated with AI in the context of sustainable development, issues directly related to AI and climate, and those related to climate and sustainable technology development. The database contains non-binding frameworks, governmental initiatives, drafts, acts, bills, executive orders, and other texts identified as relevant for the purpose of research.

Furthermore, the documents with a broader spectrum of relevance that goes beyond the interests of this study have been deliberately trimmed by manual selection to include only those pages that deal with matters relevant to the topic of artificial intelligence. In this way, the study sample was minimized to save computer resources and maximize the results of the techniques used in the study. In total, 2325 pages of text were used to conduct this study. For a complete list of used documents, references, and applied modifications, see the appendix at the end.

To provide some highlights from the corpora, selected documents within the EU region establish standard policies for all Union members, reflecting the European approach towards AI. These include the "European Data Act", which contains new rules on data governance across all economic sectors in the EU; the "AI Act", which aims to provide AI developers and users with precise requirements and obligations regarding specific uses of AI, thereby ensuring trust; the "Ethics Guidelines for Trustworthy AI", which outlines seven essential requirements that AI systems should meet to be deemed trustworthy; and the "AI liability proposal", which emphasizes the need for a transparent and predictable legal framework to address technological challenges and make Europe a world leader in the newest technologies, including AI, Internet of Things, and robotics. Furthermore, the dataset comprises documents intended to address sustainability and development concerns. However, they may also be relevant to the study, such as the "European Green Deal", a package of legislative initiatives for reaching climate neutrality by 2050 and the "European Consensus on Development" that "proposes a shared vision and framework for development cooperation for the EU and its Member States, aligned with the 2030 Agenda".

In the case of the USA, the situation is more complex, as comprehensive, federal-level regulations are still lacking (Sorkin et al., 2023). However, some laws have a limited scope of impact on AI. The US has introduced multiple frameworks and guidelines to maintain its leadership in AI. Congress passed the "Maintaining American Leadership in Artificial

Intelligence" legislation to uphold US dominance in AI research and development and the "AI Accountability Framework for Federal Agencies and Other Entities" to regulate government AI use. In May 2023, the Biden-Harris administration updated the "National AI Research and Development Strategic Plan", initially issued by the previous administration, highlighting a principled and coordinated approach to international collaboration in AI research. Furthermore, state frameworks and guidelines will influence the final form of regulation in the US and have, therefore, also been included in the research. These are the "White House Blueprint for the AI Bill of Rights" and the "AI Accountability Framework", which guide the responsible implementation of AI principles. The former document sets out five principles, including the need for safe and effective systems, algorithmic discrimination protections, data privacy, transparency, explanation, and alternative options. The latter identifies critical accountability practices that assist federal agencies and other entities in the responsible use of AI. Finally, the "National Defense Authorization Act for Fiscal Year 2019" includes provisions for enhancing AI capabilities within the Department of Defense, and the "National AI Initiative Act of 2020" aims to coordinate AI research, development, and policy across the federal government.

For China, the necessary documents include the "Interim Measures for the Management of Generative Artificial Intelligence Services", which underscores the development and deployment of AI, demonstrating China's commitment to this technology. Another key document is the "Position Paper of the People's Republic of China on Strengthening Ethical Governance of Artificial Intelligence", a national-level policy that explicitly targets AI or algorithms for regulation or governance. Other relevant legislation includes the "Measures for Review of Scientific and Technological Ethics (for Trial Implementation)", which outlines the ethical review processes for scientific and technological research to ensure compliance with ethical standards, and the "Administrative Provisions on Recommendation Algorithms in Internet-based Information Services", which regulates the use of recommendation algorithms to ensure they operate within ethical and legal boundaries.

The documents mentioned above were subjected to the preprocessing methods outlined in Section 3.3 and then analyzed using the computer techniques and tools described below. One of the key innovations in this thesis is the use of unsupervised machine learning techniques, such as LDA, to categorize AI legislation and map the evolution of legal work on this topic, particularly in the context of sustainable AI. This cutting-edge approach ensures that the insights provided are at the forefront of AI research and policy analysis.

3.2 Natural Language Processing (NLP)

Natural Language Processing (NLP) is a subfield within machine learning that employs computational methods to leverage existing information, make predictions, and enhance performance by optimizing various objectives. NLP focuses on the interaction between computers and natural languages, explicitly addressing how computers interpret human languages and texts (Eisenstein, 2018, pp. 1-5). This intersection has garnered significant interest within the DH community (McGillivray et al., 2020).

Statistical models represent a significant methodology within the field of NLP. They function by taking a large set of features, that is, individual measurable characteristics, such as words or documents, for NLP tasks. Based on statistical assumptions, which are the underlying principles that guide the construction of the model and the interpretation of its results, about this input, they aim to construct a probabilistic model of the data. This process involves extracting information from words and documents, such as word counts. Subsequently, the model attempts to identify patterns within the data by utilizing the extracted information, such as word co-occurrence. Ultimately, the model determines the likelihood of different patterns, forming a probabilistic representation of the input data (Eisenstein, 2018, p. 139). This probabilistic model is then employed to perform various tasks, including text categorization (pp. 395-396).

NLP offers powerful tools for analyzing and interpreting large textual datasets (McGillivray et al., 2020). By leveraging statistical models and advanced computational techniques, researchers can uncover patterns, categorize texts, and gain new insights into language and text. These tools empower researchers to delve deeper into the complexities of language and text. An illustration of a statistical NLP approach is topic modeling, which will be discussed in greater detail in Section 3.4. Topic modeling is a probabilistic method used to discover the abstract "topics" that occur in a collection of documents. It helps uncover hidden thematic structures in large text corpora, making it a valuable tool for text analysis in DH (Blei, 2012).

3.3 Text Preprocessing

NLP models are not capable of effectively using plain text directly as input. Therefore, prior to any textual analysis, the input data must undergo appropriate preparation. Text preprocessing entails converting raw text into well-defined input data. Various techniques can be used in text preprocessing (Palmer, 2010), some of which are pertinent to the presented study.

Firstly, a list of "stop words" is compiled. This is a list of words removed before text processing because they are unlikely to convey significant information about the topics under consideration. The removal of stop words is a standard text preprocessing technique for topic models and has been demonstrated to positively impact the results (Pedregosa et al., 2011). Typically, the initial set of stop words to be removed includes punctuation, memorable characters, and numbers, as they generally convey little or no information. It is also common practice to designate frequently used words in a language, such as "it", "an", and "the" as stop words.

Furthermore, the most prevalent words in the specific corpus can be employed as stop words. These words are identified based on their frequency and distribution across the documents in the corpus. The construction of a stop words list is a complex process that can influence the resulting model in both positive and negative ways (Watanabe & Baturo, 2024).

Secondly, the technique of "tokenization" is a fundamental step in text preprocessing (Palmer, 2010). Tokenization is the process of breaking down the text into individual words or "tokens". Each token represents a meaningful unit in the text, such as a word or a phrase. The choice of tokens can significantly influence the performance of an NLP model. For instance, one might choose to treat "climate change" as a single token ("climate_change") rather than two separate tokens ("climate", "change") – conveying a completely different meaning. Hence, choosing a tokenization method appropriate for the specific task and the nature of the text data is vital. Like the construction of a stop words list, the choice of tokenization method can significantly impact the model results (Palmer, 2010).

3.4 Topic Modeling

Topic modeling is a widely utilized semantic analysis technique, mainly when it is impractical for a single individual to read and classify a substantial corpus of documents. When applied to a set of textual files, a topic model identifies interpretable semantic concepts, or topics, present within the corpora. The identified topics represent the themes or subjects of the text. They can be used to create high-level summaries of extensive collections of documents, identify research documents of interest, and group similar documents together (see Angelov, 2020).

Topic modeling encapsulates the essence of DH research. According to Meeks and Weingart (2012), topic modeling serves as a synecdoche of the field, embodying the principles of distant reading and transforming texts into 'buckets of words' to identify underlying themes. This technique, leveraging complex statistical methods such as Dirichlet priors and Bayesian models, has introduced digital humanists to new ways of interpreting large corpora of texts. Its transformative power is inspiring, opening new possibilities for research and analysis. However, it's important to note that topic modeling may not always capture the full complexity of a text, and the interpretation of the results requires careful consideration and domain knowledge.

Topic modeling is a game-changer in legal research, enhancing the analysis of legal documents by uncovering implied meanings and revealing thematic relationships among different legal documents (Kanapala et al., 2019). Given the distinctive characteristics of legal texts, such as specialized vocabulary, formal syntax, and domain-specific semantics, this approach enables a more efficient and, importantly, insightful analysis of legal documents, providing a deeper understanding of the text.

Topic modeling has proven its usefulness in sustainability and AI research. For instance, Székely and vom Brocke (2017) broke new ground by applying topic modeling to 9,514 sustainability reports published between 1999 and 2015. Their pioneering study identified forty-two topics reflecting sustainability, focusing on economic, environmental, and social sustainability. Among the first to analyze such a large amount of data on organizational sustainability reporting, this study demonstrates the potential of NLP in sustainability research.

In another study relevant to this thesis, Rosca et al. (2020) explored legal research on AI using Latent Dirichlet Allocation (LDA) topic modeling on a dataset of 3931 journal articles. They distinguished 32 meaningful topics within AI-related legal research and noted a significant increase in such research since 2016, with topics becoming more granular and diverse over time. The comparison of similarity assessments by the algorithm and a human expert often coincided, highlighting the efficacy of machine learning and information retrieval tools like LDA in structuring extensive document collections and identifying relevant articles.

3.5 Latent Dirichlet Allocation (LDA)

One of the most employed methods for topic modeling is Latent Dirichlet Allocation (LDA). Initially introduced in the field of machine learning by Blei et al. (2003), LDA is a generative probabilistic topic model for collections of discrete data, such as a text corpus. The algorithm is designed with a systematic approach to identify topics that can effectively represent a given collection of documents. Its architecture assumes that each document is generated from a mixture of these topics and that each topic is generated from a probability vector (distribution) of overall words. Given this generative model for any document collection, LDA attempts to reverse-engineer the process to identify a set of topics likely to have generated the entire collection (Asmussen & Møller, 2019). In LDA, each document is regarded as a random mixture of latent topics, with each topic being represented by a probability distribution over words and with each word considered to belong to a topic (Blei et al., 2003).

The LDA generative process is an iterative one, starting with the selection of a parameter for each document from a Dirichlet prior distribution. For each word in the document, a topic category is selected according to the distribution. Finally, a word is generated according to the chosen topic and the word distribution parameter, β. The result is a topic accompanied by a list of words contributing to its emergence. In conclusion, the outputs of LDA provide information about the relative importance of topics within each document and the words that contribute most to the emergence of each topic.

3.6 Seeds Words

Seeded LDA represents an extension of LDA that enhances the interpretability of results by incorporating seed words, which utilize prior lexical knowledge. The pre-assignment of seed words to specific topics facilitates the definition of topics, particularly in theoretically motivated analyses. In deductive topic analyses, it is recommended that seed words be selected from terms that frequently occur in the contexts of the target concepts to inform the algorithm effectively (Watanabe & Baturo, 2024). In essence, the algorithm is directed to focus on specific concepts through the use of seed words (Jagarlamudi, Daumé III, & Udupa, 2012).

Each seed word is associated with a strength value, which determines the degree of certainty with which that word is assigned to the pre-assigned topic compared to non-seed words. Experiments conducted by Jagarlamudi et al. (2012) have demonstrated that Seeded LDA outperforms standard LDA regarding variational information between clusters.

In addition to the manual identification of seed words, this study performed a "term frequency-inverse document frequency" (TF-IDF) analysis. This approach helps to identify additional terms that may be relevant in the context of the documents, although they may take time to be apparent. As with the TF-IDF approach, each term is weighted by dividing the term frequency by the number of documents in the corpus that contain the term rather than representing a term in a document by its raw frequency (number of occurrences) or relative frequency (number of terms divided by document length). This avoids a common issue in text analytics, namely that words with a high frequency in one document will often have a high frequency in all documents. Conversely, terms with the highest TF-IDF scores are those in a document that are disproportionately frequent compared to others (Blei et al., 2003).

4. EVALUATION

4.1 Structure of the Research

This study was structured in the following manner to obtain relevant and apt results. Once the data had been collected and a comprehensive set of AI-related legal documents from all regions had been gathered, the corpus was preprocessed. In this phase, the text was subjected to a series of operations, such as cleaning, tokenization, and stop-words removal. Relevant terms were then identified through manual curation and TF-IDF analysis. Finally, topic modeling was performed using the seeded LDA algorithm.

The results were meticulously analyzed and interpreted, with a clear focus on the study's objectives. This involved a detailed examination of the topic-word distributions, reviewing the words associated with each topic and assessing their prominence for the discovered topics. The document-topic distributions were used to identify which documents were highly associated with topics related to sustainability, environmental impact, and resource requirements of AI. These findings were further contextualized through analysis, refined where necessary, and evaluated to ensure that the model's outputs were not only meaningful but also directly relevant to the study being conducted. A comprehensive explanation of the programming aspect of the study, with all accompanying visualizations, can be found on GitHub in the form of Jupyter Noteboook. The link is provided before the Abstract of this work.

In addition, the figure below graphically illustrates the above research sequence, providing a visual representation of our meticulous analysis process:

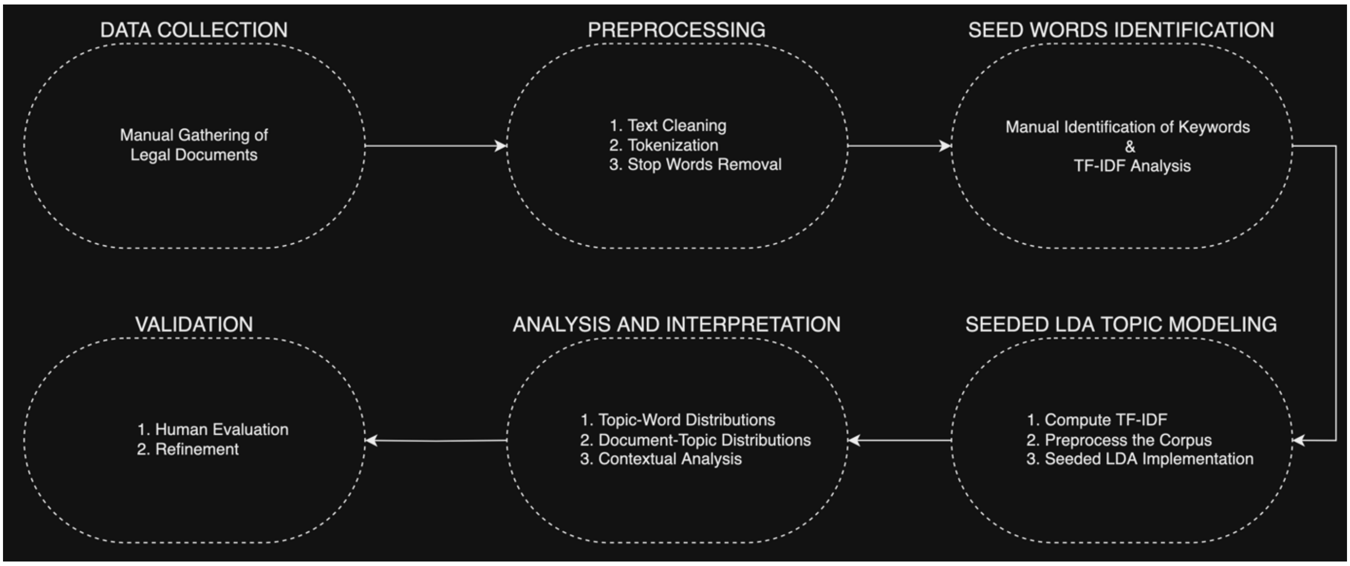

Figure 2: Diagram representing architecture of the research.

4.2 Model Evaluation Metrics

Each computational operation uses a different set of metrics to measure its quality. Appropriate ranking criteria were used to evaluate the efficiency and accuracy of the techniques employed. In terms of seeded LDA evaluation, the essential metrics consist of (Jagarlamudi, Daumé III, & Udupa, 2012):

1. *Topic Coherence*: used to measure how semantically related words are found in a topic. Higher coherence often indicates more meaningful topics.
2. *Inclusivity*: assessed by examining the topic-word distributions to ensure seed words have high probabilities in their retrospective topic. The score evaluates whether seed words appear prominently in the expected topics.
3. *Perplexity*: measures how the model fits the data with lower perplexity indicating better performance.
4. *Human evaluation*: the qualitative assessment was carried out under the criterion of interpretability, that is, whether the terms that define the topic are interpretable and contain coherent meaning from the human perspective.
5. *Topic Distributions*: used to quantify how much each document in a collection of documents pertains to each of the topics in a topic model. It is calculated by inferring the topic distribution for each document using the model (Weston et al., 2023). The distribution is a list of probabilities, one for each topic in the model, indicating how likely the document is to belong to that topic. This metric is handy in comparing documents from different categories (in this case, regions) and determining the best-performing documents for each topic in each category (Weston et al., 2023). Furthermore, it can be used to determine which category has the highest average topic probability for each topic, providing a measure of the "best" category for that topic.

In the context of this work, these metrics provide a comprehensive evaluation of the seeded LDA model's performance and its ability to generate meaningful and interpretable topics from the studied collections. The combination of these metrics allows for both quantitative and qualitative assessment of the model, ensuring that the topics generated are not only statistically sound but also meaningful and relevant from a human perspective. Such a framework ensures the reliability and validity of the model's results, contributing to the robustness and credibility of the study.

4.3 Exploring the dataset

Before analyzing the dataset comprising the legal documents from all the regions, it is important to grasp all the elements within this dataset. This helps to get a sense of the structure of the corpora. Three datasets were created after converting all the files text to lowercase, tokenizing, and removing stopwords and non-alphabetic tokens. The documents within the datasets differ in length. After combining datasets into one, the average length is 7477 words, with a minimal file length of 259 words and a maximum of 33062 words—the histogram in Figure 3A below shows how the word length is distributed across the combined dataset.

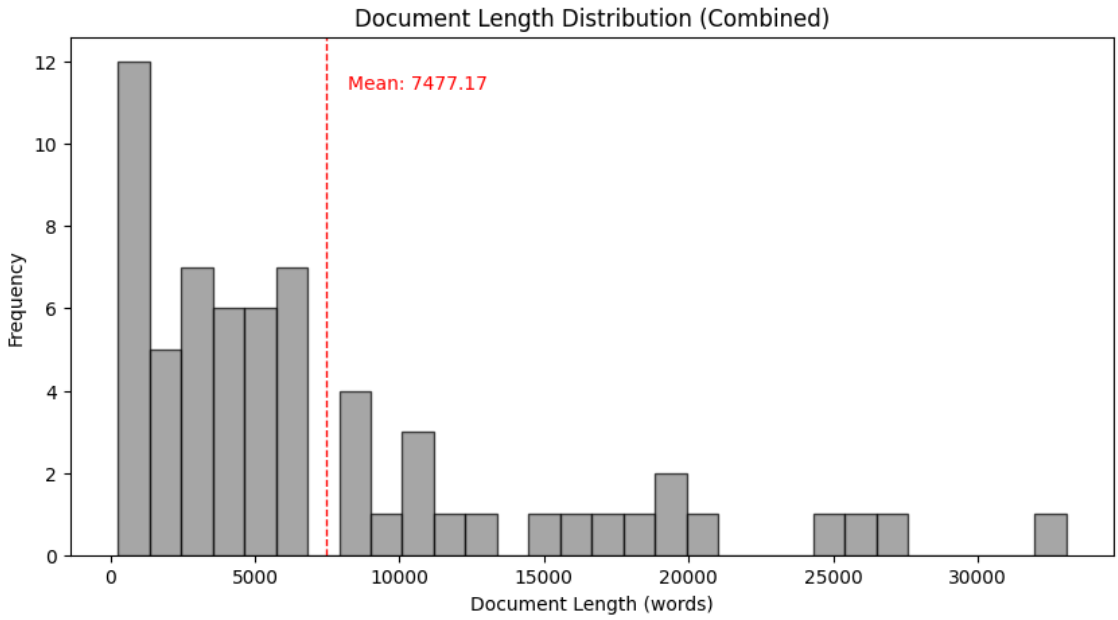

Figure 3A. Histogram with the word length on the *x*-axis and the frequency of the documents on the *y*-axis.

The right-skewed distribution indicates that most documents are relatively short, with fewer long documents. Most documents have lengths under 10,000 words, with a significant peak in the 0-5,000 word range. There are fewer documents as the word count increases beyond 10,000. Document lengths vary significantly, with some documents exceeding 30,000 words.

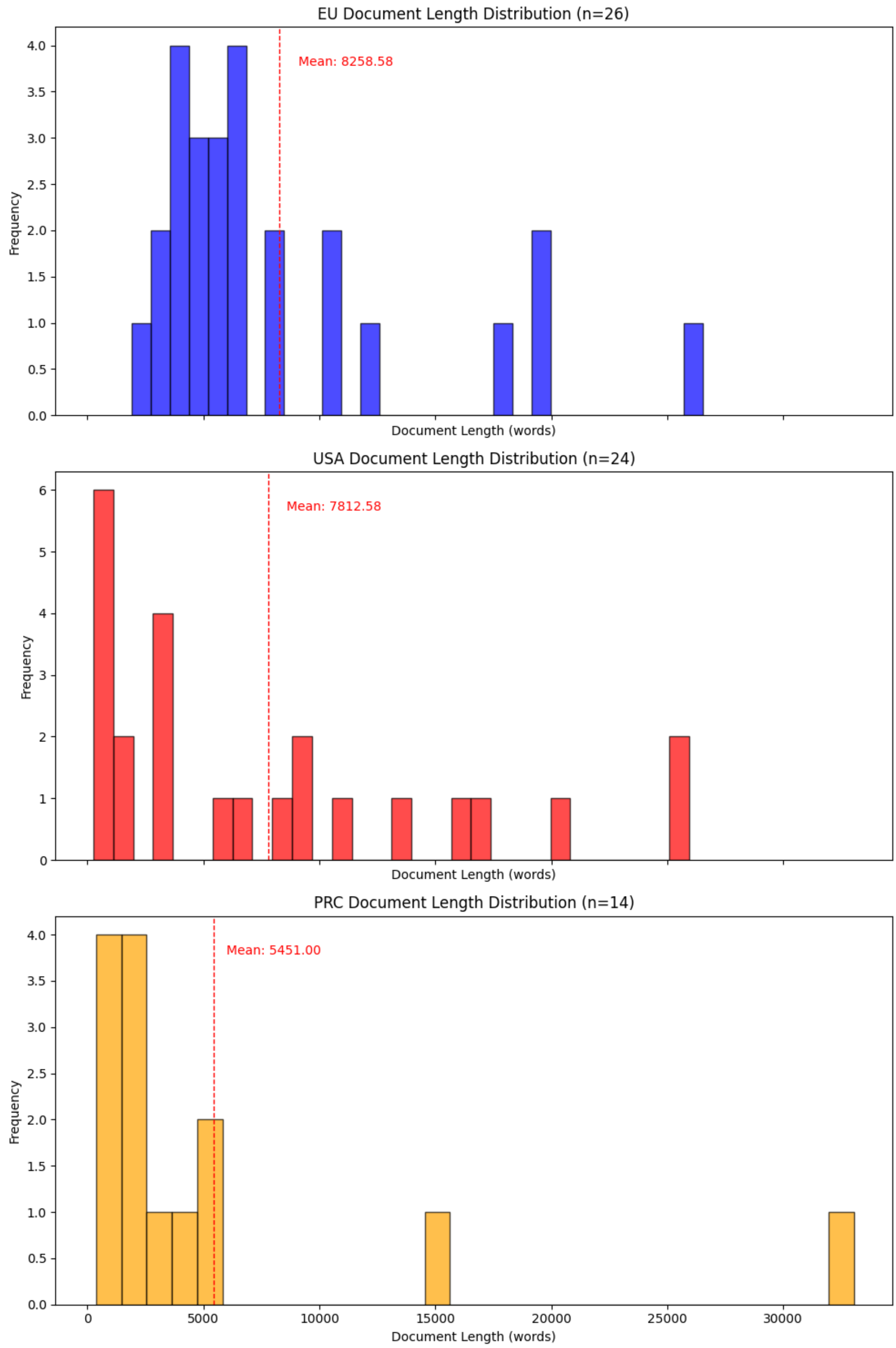

Figure 3B. Histograms with word length across different regions. The 'n' variable represents the number of files in a given dataset.

However, looking at the regionally-separated datasets (Figure 3B above), one can see that the longest documents are among the legislative initiatives of the European Union, with an mean of 8259 words. It is also here, however, that the most short texts are found. Similarly long, inflating the overall average, are the documents belonging to the United States with an mean of more than 7,800 words. The shortest texts are in the group of documents selected for China (represented in the code by the acronym 'PRC' of the country's full name, the People's Republic of China), with a mean of 5451 words.

Nontheless, it should be mentioned that these metrics are also affected by the fact that the representation of legal acts between regions is not equal in the corpus. The European Union leads the way in the number of AI acts, with 26 being considered in this study. The U.S. has a slightly smaller representation in the study with 24 documents. The least amount of legislation used in the survey comes from within China, partly due to the lack of official English translations of the legislation. Those used in this study mostly come from public initiatives such as China Law Translate (https://www.chinalawtranslate.com/en/), the University of Stanford DigiChina Project (https://digichina.stanford.edu/) or The Center for Security and Emerging Technology within Georgetown University's Walsh School of Foreign Service (https://cset.georgetown.edu/). A thorough list can be found in the Appendix at the end of this paper.

4.4 Computational Research Highlights

To conduct an in-depth analysis of AI governance strategies in the EU, USA, and China with a focus on sustainability, the Tomotopy library in Python (https://pypi.org/project/tomotopy/) was utilized. Tomotopy, a powerful tool for topic modeling, enabled extracting meaningful topics from the analyzed legal documents. The objective was to uncover how these jurisdictions address the issue of AI and sustainability through their legal frameworks.

Using Tomotopy, a Seeded LDA model was created to process and analyze the text. This approach effectively identifies underlying themes in large text corpora by representing documents as mixtures of topics and topics as mixtures of words (Blei et al., 2003). The parameters set for the LDA model included num_topics=5 (the number of topics to be extracted), alpha=0.1 (the hyperparameter that represents document-topic density), eta=0.01 (the hyperparameter that represents topic-word density), and over 700 training iterations. These settings were selected based on the recommendations from the literature to ensure robust topic extraction (Asuncion et al., 2009; Watanabe & Baturo, 2024).

The number of topics was chosen not only to detect topics relevant to this study but also to use the standard hyperparameter tuning technique to find the model's best settings (Watanabe & Baturo, 2024). Then, the Seeded LDA models were trained separately on the three corresponding datasets to achieve the best results in topic modeling. Five pre-identified topics: 'sustainability', 'environmental_impact', 'energy_demand', 'water_consumption', and 'accountability'. These topics were defined using a table of specific words (see Figure 4A below).

Topic Name	List of Words
sustainability	'sustainability', 'sustainable', 'development', 'eco-friendly', 'ecological', 'ecology', 'green'
environmental_impact	'environment', 'impact', 'carbon', 'footprint', 'greenhouse', 'climate', 'emissions'
energy_demand	'energy', 'demand', 'consumption', 'resource', 'electricity', 'power', 'use', 'usage'
water_consumption	'water', 'consumption', 'usage', 'use', 'cooling', 'hydro', 'waste', 'waterwaste'
accountability	'accountability', 'responsibility', 'answerability', 'oversight', 'accountable'

Figure 4A. Table with list of pre-defined topics and associated words used for Seeded LDA topic modeling.

Moreover, the list of seeded words was further enhanced by adding the top 10 words found after running a TF-IDF analysis on the corpora documents (see Figure 4B below), as this positively improved the metric values of the model.

Looking at the output, in all three regions development and innovation is emphasized in their top words. The TF-IDF analysis reveals a shared global commitment to AI development and innovation but highlights different regional priorities. However, there are notable regional specific priorities. The EU focuses on sustainability and climate, the USA on security and defense, and the PRC on promoting AI and supporting providers.

Region	Top TF-IDF Words with Score
EU	digital: 3.3070 public: 2.5628 development: 2.0700 support: 1.7019 market: 1.6504 sustainable: 1.3112 innovation: 1.2248 requirements: 1.2235 climate: 1.1776 technologies: 1.1648
USA	information: 1.6835 development: 1.4326 research: 1.4013 public: 1.3498 security: 1.3220 technology: 1.2712 automated: 1.1881 defense: 1.1172 risk: 1.0865 critical: 1.0272
PRC	development: 2.1404 security: 1.4944 technology: 1.3345 promote: 1.1869 research: 1.1668 information: 1.1465 innovation: 1.1008 providers: 0.9653 public: 0.9341 improve: 0.9264

Figure 4B. Table with results of TF-IDF analysis for specific regions.

The LDA model was first trained on the legal texts from the EU. The extracted topics highlighted the EU's comprehensive approach to integrating sustainability into AI governance, emphasizing ethical considerations and data management practices (coherence score: 0.7455). The same LDA model was then applied to the corpus from the USA. For the USA corpora, the model emphasized national security and state-level regulatory efforts (coherence score: 0.6358). Finally, the topic model for China discovered a proactive stance in promoting AI technologies while simultaneously addressing ethical and regulatory challenges (coherence score: 0.6119).

5. RESULTS, FINDINGS AND DISCUSSION

This section of the thesis is intended to provide a comprehensive assessment of the gathered data and to present the findings derived from the analysis. It will present the output of the topic modeling and decipher their meaningful consequences, laying the groundwork for the discussion, where a comprehensive examination of the results will take place.

5.1 Overview of Results

The Seeded LDA topic modeling was applied to analyze legal documents from three distinct and globally significant regions: the European Union (EU), the United States of America (USA), and the People's Republic of China (PRC). Each topic, or to be more precise, 'cluster', as computer systems do not generate topics; instead, they generate a list of high-concentration words containing a set of 20 words, ordered by their significance to the topic.

In the case of the European Union (EU), the model demonstrated a gradual improvement in log-likelihood values, commencing at -7.5948 and concluding at -7.5611. The log-likelihood value serves to quantify the extent to which the model is able to explain the observed data. A higher (less negative) value indicates a superior fit. The further analysis revealed several coherent topics. One of the identified topics was related to systems requirements and priorities, encompassing terms such as 'requirements', 'market', 'safety', 'assessment', and 'machinery'. Another critical topic was digital innovation with key terms including 'digital', 'public', 'innovation', and 'smes', highlighting the importance of Small and Medium-sized Enterprises to the EU's economy. The concept of sustainable development emerged as a distinct area of focus, characterized by the appearance of terms such as 'development', 'sustainable', 'support', 'climate', and 'countries'. Impact assessment also stood out, with words such as 'solution', 'ict', 'assessment', 'impact', and 'effects'. Finally, accountability constituted a significant topic, incorporating terms like 'public', 'bodies', 'processing', and 'information'. The metrics for the EU model were as follows: coherence at 0.7607, perplexity at 1915.66, and inclusivity scores of 1.0 across all topics.

In the United States of America (USA) dataset, the model demonstrated an enhancement in log-likelihood values, progressing from -7.7962 to -7.1435. The analysis yielded a number of discernible topics. Defense and security were highlighted with top words such as 'defense', 'security', 'report', 'technology', and 'program'. Another noteworthy topic was vehicle safety, which was characterized by terms like 'safety', 'vehicles', 'nhtsa' (that stands for the National Highway Traffic Safety Administration), 'dod' (abbreviation for the Department of Defense), and 'guidance'. Accountability emerged as a critical one with words such as 'framework', 'automated', 'practices', 'accountability', and 'information'. Another significant topic was identified, comprising terms such as 'development', 'research', 'risk', 'technology', and 'standards'. The comments and responses topic included the following terms: 'comments', 'response', 'accountability', 'commenters', and 'human'. The metrics for the USA model were as follows: coherence at 0.5251, perplexity at 2407.15, and an inclusivity score of 1.0 across all topics.

In the case of the People's Republic of China (PRC) legal documents, the model consistently improved with log-likelihood values increasing from -7.1435 to -7.0584. The analysis identified several coherent topics. Themes of development and innovation were particularly salient, with notable occurrences of terms such as 'development', 'industry', 'innovation', 'china', and 'key'. Security and integration were also of paramount importance, featuring words such as 'security', 'research', 'public', 'institutions', and 'technology'. The construction and management topics were also of significance, with terms such as 'improve', 'promote', 'strengthen', 'construction', and 'areas'. The technical standards topic was defined by words such as 'technology', 'smart', 'information', 'issued', and 'standards'. The administration topic included terms like 'information', 'trustworthy', 'providers', 'deep', and 'management'. The metrics for the PRC model were: coherence at 0.6780, perplexity at 1137.40, and an inclusivity score of 1.0 across all topics.

The Seeded LDA models were demonstrably effective in extracting topics from the provided documents. The models generated for each region revealed topics that align with the predefined themes of sustainability, environmental impact, energy demand, water consumption, and accountability. The metrics indicate reasonable coherence and inclusivity across all regions, though with varying perplexity, which reflects differences in the complexity and structure of the documents from each region. The EU model demonstrated the highest coherence score, indicating that the legal documents in this region are more consistently aligned with the identified topics. Conversely, the USA model had the lowest coherence and the highest perplexity, indicating grater variability and complexity in the topics under discussion. This can be attributed to the diverse and multifaceted nature of legal documents, which cover a broad spectrum of topics with varying levels of detail. The PRC model had the lowest perplexity, suggesting that its documents are more straightforward to model, possibly due to the simplicity of translations or perhaps, more uniform approach to regulatory documentation in China. The remaining parts of this chapter will thoroughly analyze all the topics, offering deeper insights and ensuring a more thorough understanding.

5.2 Topics for European Union

Topic No.	Output
1	digital public innovation support technologies smes research green economy across skills industrial strategy investment industry
2	market requirements safety assessment machinery conformity fundamental specific product products risk authorities applications liability option
3	public bodies processing information referred competent accordance user third provider authority intermediation request particular related
4	development sustainable climate countries support cooperation promote developing social implementation change action global human policies
5	solution ict assessment impact effects net carbon solutions emissions reference scenario methodology implementation impacts second

The EU model identified several coherent topics within the legal documents. The first topic, centered on digital innovation, is closely aligned with the pre-defined category of sustainability, highlighting the integration of digital advancements with sustainable practices and support for small and medium-sized enterprises. The second one focused on market requirements and regulatory compliance and corresponds to the accountability category, underscoring the EU's emphasis on regulatory standards, product safety, and market conformity. Public administration and information management emerged as the third topic. This theme also aligns with the accountability category, focusing on the governance and management of public information and administrative processes. The fourth identified topic is sustainable development and climate action. This theme aligns with the sustainability and environmental impact categories, reflecting the EU's commitment to addressing climate change and promoting global sustainable policies. Lastly, the fifth topic dealt with impact assessment and environmental solutions. It corresponds to the environmental impact category, focusing on assessing and mitigating environmental impacts, particularly those related to carbon emissions.

The topics derived from the EU corpora highlight several critical areas of focus that align with the sustainable development of AI. The prominence of digital innovation and support for SMEs indicates that the EU prioritizes technological advancement while ensuring these developments contribute positively to the economy and society. The emphasis on integrating 'green' and 'sustainable' practices within digital directly aligns with the goals of AI sustainable development. Moreover, market requirements and regulatory compliance underscore the importance of creating a safe and standardized market environment for AI technologies. This focus on safety, assessment, and conformity might ensure AI adherence to strict regulations, mitigating risks and fostering trust among users and stakeholders. The public administration and information management themes highlight the governance and oversight necessary for managing AI technologies effectively. Ensuring that public information is processed and managed according to strict standards aligns with the accountability goals of sustainable AI development. It indicates a structured approach to integrating AI within public sectors, emphasizing transparency and user protection. Finally, the focus on impact assessment and environmental solutions demonstrates the EU's proactive stance in evaluating and mitigating the environmental impacts of AI technologies. This emphasis on assessing carbon emissions and finding solutions to reduce them is critical for the sustainable development of AI, ensuring that technological progress does not exacerbate environmental challenges.

The EU's regulations and priorities, show a comprehensive approach to AI development that aligns closely with sustainability principles. The integration of innovation, stringent regulatory standards, public governance, and environmental responsibility highlights the EU's commitment to developing AI in a manner that is beneficial, safe, and sustainable for future generations.

5.3 Topics for United States

Similarly to the previous model, the analysis of USA documents identified several topics. Interestingly, although they vary in their themes, they are all closely tied to the accountability category.

Topic No.	Output
1	defense security technology research program report public development science acquisition days coordination following activities year
2	risk framework model management information performance accountability technology governance human design development entities include principles
3	response explanation dod explanations copyright commenter commenters explainable human principles protection person trade works association
4	information research standards practices development regulatory work existing applications accountability privacy public best potential technologies
5	safety automated process vehicles nhtsa guidance decision manufacturers hav motor public performance critical havs test

The first topic centered on defense and security, emphasizing technological advancements in defense and security programs and the importance of research and public reporting. The second one focused on risk management and governance, highlighting the USA's emphasis on managing risks through robust frameworks and models, ensuring technological governance and development accountability. Response and protection emerged as the third topic, focusing on responses to public inquiries, explanations of policies, protection of intellectual property, and engagement with human principles in governance. The fourth topic identified is information standards and practices, reflecting the USA's commitment to maintaining high standards in information management, privacy, and regulatory practices while fostering the development and application of new technologies. The fifth topic dealt with autonomous vehicle safety, ensuring it through rigorous processes, guidance, and performance testing by manufacturers and regulatory bodies.

The insights derived from the study of USA legislation indicate that the USA prioritizes technological advancements in defense and security, ensuring that developments contribute to national safety and public security. The integration of 'technology', 'research', and 'public' within this topic reflects commitment to using AI to enhance defense capabilities while maintaining transparency. Furthermore, the presence of management and governance underscores the importance of creating robust frameworks for the risks associated with AI technologies. Such a framework would be essential for reaching sustainable AI, as it helps to mitigate potential negative impacts and ensures that these technologies are developed and used responsibly. Moreover, information standards and practices are distinct topics that illustrate the USA's holistic approach to AI development, with technological advancements that do not compromise information privacy and regulatory standards. High standards in information management and privacy align with broader accountability objectives. Lastly, there is much focus on vehicle safety and automation in the USA's legislation corpus. This emphasis on safety, rigorous processes, and performance testing is critical for the development of AI in the automotive industry, ensuring that technological progress does not come at the expense of public safety.

Overall, the USA's regulations and priorities, in comparison with the EU's approach, concentrate more on the principles of accountability. Nevertheless, both regions emphasize integrating technological innovation with robust risk management frameworks and public governance. Both regions underscore the importance of transparency, public engagement, and rights protection in their AI governance frameworks. Although, the USA focuses more on defense and security.

5.4 Topics for China

Building on the methodologies and frameworks discussed in the previous sections, the analysis of China's legal documents identified topics that reflect the region's regulatory focus and priorities in AI development.

Topic No.	Output
1	technology information smart issued standards standardization technical management intelligent learning requirements application trustworthy model development
2	improve construction strengthen promote implement accelerate production areas rural increase mechanisms protection management control urban
3	development promote build public support governance develop chinese plan comprehensive projects social market economic resources
4	research innovation industry institutions enterprises technology development scientific support establish encourage application people science basic

5	security information management providers accordance measures provisions administrative responsible development regulations public laws departments users

The first topic centered on technology standards and innovative applications, emphasizing the standardization of technology, intelligent applications, and the development of trustworthy AI models. The second topic focused on construction and rural development. This topic corresponds to the sustainability category, highlighting China's efforts to promote rural development and sustainable urbanization, ensuring balanced growth and protection of resources. Development and public support emerged as the third topic. It aligns with the sustainability category, focusing on comprehensive development plans that support public governance, social projects, and economic resources, promoting a well-rounded approach to development. The fourth topic identified was research and innovation, reflecting China's commitment to advancing scientific research and technological innovation, supporting enterprises, and establishing a solid foundation for AI development. The fifth topic dealt with security and information management; it focuses on managing information security, adhering to administrative measures, and ensuring responsible development through regulations and public laws.

The prominence of technology standards and innovative applications indicates that China prioritizes the standardization and management of intelligent technologies, ensuring that AI developments are reliable, trustworthy, and beneficial to society. The focus on 'technology', 'standards', and 'management' within this topic reflects China's commitment to setting high standards for AI technologies. Construction and rural development underscore the importance of promoting balanced growth across urban and rural areas. The focus on 'improve', 'construction', 'rural', and 'urban' ensures that development is sustainable and inclusive, addressing the needs of both rural and urban populations. This balanced approach is essential for sustainable AI development, as it helps prevent regional disparities and promotes equitable growth. The development and public support topics highlight the governance and oversight necessary for managing AI technologies effectively. Research and innovation as a distinct topic illustrate China's holistic approach to AI development, ensuring that technological advancements are supported by robust scientific research and innovation. Finally, the focus on security and information management demonstrates China's proactive stance in ensuring the security and responsible management of information related to AI technologies.

5.5 Approaches Comparison

The examination of AI regulations in the EU, the USA, and China reveals a mosaic of regulatory strategies and priorities towards AI development. Despite the common goal of advancing AI technology, the approaches taken by each region reveal both distinct differences and notable similarities. Undoubtedly, there is a shared emphasis on the importance of accountability in AI development, with each region prioritizes the establishment of regulatory standards and frameworks to ensure that AI technologies are developed and deployed responsibly. This is evident from the prominent themes of governance, risk management, and public oversight that emerge from the topic modeling results. Furthermore, all three regions underscore the importance of integrating technological innovation with public support and transparency. The EU's focus on digital innovation and the USA's emphasis on technological advancements in defense and public safety mirror China's commitment to establishing rigorous standards for smart applications and technological standardization. This convergence reflects a global recognition of the necessity to foster innovation while maintaining public trust and adherence to ethical standards.

However, there are considerable differences between the various methodologies. While the EU's approach is distinguished by its strong emphasis on sustainability and environmental impact, while the USA prioritizes technological advancements in defense and public safety, and China focuses on technological standardization, rural development, and balanced growth. These discrepancies underscore the unique challenges and priorities that each region faces in the context of AI development. The EU’s objective is to integrate digital advancements with sustainable practices, providing support to SMEs and addressing climate change. The USA leverages AI to enhance national security and public safety while maintaining high standards of accountability. China aims to establish global standards for intelligent technologies, thereby promoting equitable growth and scientific innovation.

In conclusion, the aforementioned dynamics of the AI race dynamics are reflected in these results. The United States, as the current leader in AI development, is currently grappling with issues concerning the accountability of the technology and the risks posed by its use, including, among other things, security concerns. China, as a rising power in the AI domain, aims to avoid missteps and develop its technology in accordance with international standards and solutions. Conversely, the European Union's approach appears to align most closely with the concepts presented in this work, as it focuses primarily on regulating AI regarding sustainability, responsibility, and attentiveness to climate change. A review of the regulatory frameworks from the other regions revealed a lack of significant interest in terms related to sustainable development understood environmentally. This omission is a cause for concern, especially considering the pressing issue of climate change. The EU documents strongly emphasized sustainability and environmental impact, aligning with global efforts to address climate challenges. However, the USA and China corpora did not prominently feature these themes, highlighting a gap in the focus on environmental sustainability within their legal frameworks. The absence of an emphasis on environmental sustainability in the documents from the USA and China indicates a necessity for a more integrated approach to AI development that considers the environmental implications.

It is therefore imperative to gain an understanding of the regional differences that exist in order to foster international collaboration and develop coherent global policies that ensure the sustainable and responsible development of AI technologies. The insights gained from this comparative analysis provide valuable perspectives for policymakers, researchers, and stakeholders involved in shaping the future of AI on a global scale.

5.6 Discussion

Seeded LDA topic modeling is a robust analytical tool that can uncover hidden patterns within large text corpora, offering valuable insights for research endeavors. However, it is essential to recognize the inherent constraints of this process to ensure the accuracy and meaningfulness of the analysis.

One of the primary limitations is that the topics are corpus-dependent, meaning that the topics assigned to texts by a computer system are contingent on their context in relation to surrounding texts. This dependency is particularly crucial in the case of the China corpora, which is the smallest among the analyzed regions. A smaller batch size may result in less robust topic modeling outcomes, as the model has fewer texts from which to draw context, potentially introducing bias into the analysis. Conversely, having an excessively large corpus can also pose challenges. With a vast number of documents, the model may generate too many topics, some of which may be overly specific or redundant, complicating the interpretation process and diminish the significance of broader themes.

Furthermore, as this technique is an example of semi-supervised machine learning, it is dependent on the researchers and interpreters involved. Hence, it is susceptible to human error in the form of validating a wrong presumption about the data, leading to biased interpretations and misrepresentations of the underlying themes. Moreover, the way words are presented in interpreting topic modeling clusters is not conventional, which has the potential to bias the analysis. The interpretation of clusters generated by LDA is contingent upon the specific words and their frequency within the documents. This non-standard use of words can lead to ambiguous or misleading topics if not carefully analyzed and cross-referenced with the actual content of the documents.

Moreover, LDA is constrained in its ability to retain semantic and syntactic meaning. Although it is capable of grouping related words into topics, it frequently fails to capture the nuanced relationships between words and the broader context in which they are used. This can result in the loss of important information and subtle meanings that are critical for a comprehensive understanding of the text, thereby underscoring the potential limitations of the method.

In conclusion, while topic modeling, specifically LDA, provides valuable insights into the latent themes within large text corpora, it is essential to recognize its limitations. The corpus-dependent nature of topics, the non-standard interpretation of word clusters, and the challenges in preprocessing and retaining semantic meaning present significant obstacles to progress. Additionally, the research highlights a crucial deficiency in the focus on environmental sustainability, particularly in the USA and China corporas. Addressing these limitations and gaps is vital for developing a comprehensive and responsible approach to AI that aligns with the broader goals of sustainable development and climate action.

6. FUTURE STEPS

6.1 Ideas for Expanding Research

This study examines the regulatory governance of AI and reveals a complex landscape that is shaped by regional priorities and challenges. Although there is a consensus regarding the necessity for accountability and further regulation in the field of AI, there are notable differences in approach among major stakeholders.

The US is a global leader in AI development, yet it is confronted with considerable challenges pertaining to accountability and security. China, in contrast, is advancing its technology while striving to adhere to international standards. The European Union places significant emphasis on stringent regulation, with a focus on sustainability, responsibility, and climate change. This research highlights these regional approaches and uncovers notable gaps, particularly in the attention to environmental sustainability within the USA and China. The study also indicates a competitive dynamic between these regions, despite the fact that climate change is a global issue that requires collaborative efforts. In light of the inevitable further development of AI technologies, conducting this type of study annually basis would be beneficial. Monitoring the progress made towards the climate goals set for 2030 can provide insights into whether countries are fulfilling their commitments. As awareness of the unsustainability of current AI practices grows, it is possible that environmental concerns may become more prominent not only among academics and journalists but also among decision-makers and stakeholders. Moreover, extending the scope of the study to encompass environmental reports from prominent technology companies such as Google, Microsoft, Meta, and China's Tencent, could facilitate a more comprehensive understanding of how private entities perceive AI sustainability. Analyzing these documents would shed light on the corporate approach to environmental responsibility in AI development.

Given the limitations of seeded LDA topic modeling, including its reliance on corpus size and context, further research is essential. This research should ideally involve cooperation among AI experts, environmental scientists, and legal professionals. Utilizing the United Nations Sustainable Development Goals (SDGs) as a framework could provide a comprehensive approach. The SDGs comprise 17 goals, 169 targets, and nearly 300 indicators monitored by UN agencies, citizen initiatives, and independent organizations (Saner et al., 2020). Goal 13, "Climate Action," directly addresses climate change, while Goals 14, "Life

Below Water," and 15, "Life on Land," also pertain to climate-related matters. Several studies underscore AI's potential in combating climate change, improving the monitoring and management of natural resources, and accelerating progress towards the SDGs. For instance, Vinuesa et al. (2020) explore AI's role in achieving the SDGs, Sætra (2021) examines the role of AI in sociotechnical systems, and Truby (2020) proposes regulatory measures to mitigate the AI's impact on the SDGs.

AI has the potential to significantly enhance sustainable development in a number of ways. These include improving resource efficiency, enabling precision agriculture, optimizing energy consumption, and enhancing disaster response. However, in order to fully realize these benefits, it is crucial to integrate AI into sustainable development strategies. This approach can effectively leverage the AI's potential to address global challenges, underscoring the significant impact that AI can have on our future.

6.2 Reaching Sustainable AI Today

The concept of Sustainable AI is inherently complex and challenging, primarily due to the intricate nature of AI systems and their diverse applications. Although the long-term objective is to develop a comprehensive and integrated AI framework that is aligned with sustainable development and climate action, there are several immediate actions that can be taken today to enhance the sustainability of AI systems without significant costs or extensive systemic changes.

One approach to SAI is the optimization of algorithms with the objective of reducing the resources and energy required for the training and operation of AI models (Foy, 2023). This involves refining the efficiency of algorithms to minimize their environmental footprint. Additionally, the relocation of data centers to more sustainable energy sources, such as solar or wind power, can result in a notable reduction in energy costs for AI operations. The promotion of recycling and reuse practices for AI hardware and algorithms has been demonstrated to have a positive environmental impact (see Zhuk, 2023). Furthermore, utilizing more efficient cooling technologies in data centers, such as Microsoft's underwater data center system (Gartenberg, 2020), can substantially reduce energy consumption. However, a particularly promising strategy is using AI to optimize environmental processes. By monitoring climate change, managing natural resources, or protecting biodiversity, AI can significantly promote sustainable practices.

Practical solutions are continuously emerging to tackle the critical challenges related to the sustainable development of AI. For instance, a research team led by ethicist Mark Coeckelbergh put forth an innovative mechanism to optimize energy consumption throughout the AI design process (Raper et al., 2022). Inspired by Differential Privacy (see Dwork et al., 2006), this mechanism introduces the concept of Sustainability Budgets, which limit the computing power allocated to a project. This approach encourages AI developers to manage their resources in a more judiciously, significantly reducing energy consumption. To make this solution more appealing and engaging for engineers, the researchers suggest incorporating gamification tools, which introduce elements of competition and reward to motivate adherence to sustainability practices (Raper et al., 2022, p. 8). Implementing Sustainability Budgets could be effective within organizations, across sectors, or even internationally, fostering widespread adoption of sustainable practices in AI development and contributing to a global reduction in the environmental impact of AI technologies (p. 7).

Other noteworthy ideas that integrate sustainability and AI include the obligatory implementation of "sustainability impact assessments" for each AI development (Hacker, 2023), the establishment of a metrics framework to facilitate the quantification of GHG emissions generated during the life cycle of models (Eilam et al., 2023), the restriction of training and inferencing models to regions where substantial amounts of green, renewable energy are available, and the establishment of consumption caps for AI, which have been proven effective by an EU-wide carbon tax (Haites, 2018). These solutions are further supported by trends such as "Green AI”, which strives to decrease the environmental impact of AI, enhance inclusivity, and considers efficiency equally crucial as accuracy in terms of models (Schwartz et al., 2020) – or the concept of "Sustainability by Design", which seeks to integrate environmental considerations into the very development and implementation of AI models and practices (Vezzoli, 2018).

Moreover, thanks to university researchers' involvement, initiatives specify the counting and collecting similar, emerging proposals. Regarding sustainable AI development, the University of Oxford's Research Initiative on AIxSDG database enumerates 108 projects, 28 of which are pertinent to Goal 13. The SDG AI Repository, which is managed by the UN's ITU agency, includes nine climate-focused projects, while the AI4SDGs Think Tank database features five such projects. Moreover, the number of publications addressing the intersection of sustainability and AI is on the rise (Falk & Wynsberghe, 2023, p. 3), indicating the growing relevance of this subject in academic circles.

The final approach ideally compromises the optimal combination of the solutions. It achievement requires further steps prior to its widespread adoption. However, each approach is crucial for raising awareness about the AI's unsustainability among developers and organizations, embedding the issue into the AI design process. By establishing clear guidelines and incentives for energy-efficient design, these strategies hold promise for making significant strides toward sustainable AI.

6.3 Revolution in Thinking

As this study emphasizes, it is vital to gain the extensive context in which AI operates, its beneficiaries, and the conditions under which it evolves. It is not just a tangible entity; it is a multifaceted social phenomenon that encompasses numerous interconnected production, pricing, and usage practices. These practices transcend individual companies and profoundly influence on society, underscoring the societal impact of AI.

Meanwhile, there is a narrative espousing the belief that the solution to our political woes lies in more, more efficient, and newer technology, which empowers a narrow interest group of the tech elite of the sort based in Silicon Valley. This enables the industry to develop its products without hindrance, unnoticed, and beyond public scrutiny. The ideology of Tech Solutionism, as it is termed, erodes the tenets of realism, criticism, and reason by spreading utopian ideals (see Byrum & Benjamin, 2022). The question of the desirability and limits of AI development becomes moot when viewed in the context of its ever-increasing efficiency, which is expected to answer all questions. Similarly, an opposing stance that views AI as a potential source of human extinction is merely a diversionary tactic that diverts attention from the more pressing concerns related to this technology (Broughel, 2023). As Kate Crawford (2024) observes, the apocalyptic scenarios surrounding AI are merely a means of obscuring the fact that these "systems are expressions of power that emerge from broader economic and political forces, designed to increase profits and centralize control in favor of those who wield those systems." Furthermore, she concludes that AI is, in essence, a set of computing techniques or tools that are misused and whose future functioning must be regulated.

Maintaining Techno-Optimism in the face of severe issues and technologically generated ignorance on the rise seems frivolous. At the same time, Techno-Pessimism does not help to reduce this ignorance in any way.

Technological revolutions are not solely defined by technological advancement; they encompass a multitude of interrelated factors. Such transformations entail significant alterations within the governmental and societal domains, but primarily a shift in attitude and mindset. Consequently, the most significant challenge posed by this revolution is to develop a new, critical, and realistic understanding of technology, rather than a speculative and utopian one, in an alliance of technical, natural, social, and human sciences. Such an alliance could serve as an intellectual and spiritual counterweight to the vision espoused by Silicon Valley. It is reasonable to anticipate that policymakers will initiate and fund the construction of such a counterbalance, rather than succumbing to the illusion of "human-centric AI" and engaging in discussions that lack meaningful substance, particularly with regard to matters of ethics, which should be placed within the context of politics and political economy. Politics without criticism are too inconsequential and can devolve into a collection of pious, wishful thinking by experts or, worse, into "ethics washing", which, analogously to "greenwashing", serves the same purpose: to enhance the image of the organization while maintaining business as usual (Munn, 2023).

Conversely, the potential for AI systems to disrupt social order is not inherent in their existence and application. Rather, it arises from the limitation of our conventional ways of thinking about them are too narrow. The design and purpose of AI systems are a reflection of our current understanding of them. An era that views full automation as the pinnacle of intelligence is, by definition, one that is willing to relinquish the power that comes from critical thinking. This regression to a primitive stage of technical culture underscores the need for a broader perspective on AI.

7. CONCLUSION

The analysis presented in this thesis underscores the urgent and critical importance of sustainable development in the realm of AI technologies. It highlights the multifaceted challenges and opportunities of integrating sustainability principles into AI governance. This research has explored the regulatory frameworks of the European Union, the United States, and China, with a view to understanding how these different jurisdictions balance the promotion of AI innovation with the need to address environmental and ethical considerations.

While this thesis provides valuable insights into the intersection of AI and sustainability, it also reveals significant gaps in understanding that warrant further investigation. A significant area for future research is the development of comprehensive metrics to assess the environmental impact of AI technologies. The current metrics are frequently inadequate, failing to capture the total environmental costs associated with the development and deployment of AI. Moreover, there is a need for longitudinal studies to monitor the long-term effects of AI on sustainability, which could facilitate a more nuanced understanding of the manner in which AI technologies evolve and affect the environment over time.

In order to adequately address the sustainability challenges posed by AI, it is necessary to adopt an interdisciplinary approach that is both concerted and collaborative in nature. This thesis has demonstrated that the sustainability of AI is not merely an engineering issue, but rather one that intersects with a number of other domains such as environmental science, ethics, law, and public policy. It would be beneficial for future research to involve collaboration among experts from these diverse disciplines. For instance, environmental scientists can offer insights into the ecological impacts of AI technologies, while ethicists can assist in navigating the moral implications of AI use. Those engaged in the field of legal scholarship and policymaking can play a pivotal role in the creation of robust regulatory frameworks that ensure the development of AI technologies is aligned with the overarching goals of sustainability.

Interdisciplinary research could also foster the development of innovative solutions that integrate technological advancements with sustainability principles regarding the social impact of these technologies mentioned in the introduction. By integrating insights from the social sciences, we can gain a better understand how AI technologies shape societal structures, potentially exacerbating or alleviating social inequalities. This holistic perspective ensures that

AI development is environmentally sustainable and socially equitable, fostering inclusive growth. It is incumbent upon governments to enforce labor standards that protect workers' rights and ensure fair wages, particularly for those engaged in the data labeling and content moderation tasks that are indispensable to AI training. It is incumbent upon AI companies to adopt ethical labor practices, thereby fostering a more equitable and just industry.

This thesis has identified several critical issues related to AI sustainability, including high energy consumption, greenhouse gas emissions, and depletion of nonrenewable resources. Addressing these issues requires concerted efforts from all AI development and governance stakeholders. To mitigate the environmental impact of AI's energy consumption, investing in more energy-efficient AI models and infrastructure is essential. Researchers and developers should prioritize the development of algorithms that require less computational power without compromising performance. Additionally, there is a need for a unified approach to transition towards renewable energy sources for the operation of data centers and AI operations. Policymakers can facilitate the implementation of these initiatives by providing incentives for AI companies to adopt green technologies and set stringent emissions standards.

Accountability remains a crucial aspect of the sustainable development of AI. Ensuring that AI technologies are developed and deployed responsibly requires clear and enforceable accountability mechanisms. This thesis has emphasized the importance of holding both developers and users of AI accountable for the environmental impacts of their technologies. Effective accountability frameworks should include rigorous standards for transparency and reporting, enabling stakeholders to monitor and evaluate the sustainability practices of AI firms. Furthermore, regulatory bodies must be empowered to enforce these standards, with the authority to impose penalties for non-compliance and incentivize best practices. By establishing robust accountability measures, we can ensure that the development of AI aligns with the broader goals of sustainability and ethical responsibility.

The AI industry must adopt more sustainable practices with regard to the sourcing and utilization of materials. This encompasses the formulation of strategies for the responsible extraction and use of rare earth elements, as well as the promotion of recycling and reuse of AI hardware. Collaborations with industries that specialize in sustainable materials and recycling technologies could serve to mitigate the resource depletion associated with the development of AI. The ethical issues related to labor exploitation in AI development must be addressed through regulatory oversight and industry self-regulation.

In conclusion, the sustainable development of AI is a necessity, a multifaceted issue that requires an interdisciplinary approach. This thesis has provided a foundation for understanding AI's environmental and ethical implications. However, it is evident that further work is required. Future research should focus on developing comprehensive sustainability metrics, exploring sector-specific sustainable practices, and fostering further interdisciplinary collaborations. By addressing the identified issues and prioritizing sustainability, we can ensure that AI technologies contribute positively to society and the environment, aligning with the broader goals of sustainable development. It is our shared responsibility to control the science and technology behind AI.

REFERENCES

Acemoglu, D. (2024, June 5). *Don't believe the AI hype*. Project Syndicate. https://www.project-syndicate.org/commentary/ai-productivity-boom-forecasts-countered-by-theory-and-data-by-daron-acemoglu-2024-05

Ahluwalia, P., & Miller, T. (2023). The next big thing – artificial intelligence. Social Identities, 29(1), 1-4. https://doi.org/10.1080/13504630.2023.2236372

Aldrete, B., Ward, J., & Pandise, E. (2024, March 7). The AI industry is pushing a nuclear power revival — partly to fuel itself. NBC News. https://www.nbcnews.com/tech/tech-news/nuclear-power-oklo-sam-altman-ai-energy-rcna139094

Alonso, C., Kothari, S., & Rehman, S. (2020, December 2). How artificial intelligence could widen the gap between rich and Poor Nations. IMF. https://www.imf.org/en/Blogs/Articles/2020/12/02/blog-how-artificial-intelligence-could-widen-the-gap-between-rich-and-poor-nations

Andersen, R. (2020, July 30). The Panopticon is already here. The Atlantic. https://www.theatlantic.com/magazine/archive/2020/09/china-ai-surveillance/614197/

Angelov, D. (2020). Top2Vec: Distributed Representations of Topics. ArXiv, abs/2008.09470.

Ashford, N. A., & Hall, R. P. (2011). The importance of regulation-induced innovation for sustainable development. Sustainability, 3(1), 270-292. https://doi.org/10.3390/su3010270

Asmussen, C.B., Møller, C. Smart literature review: a practical topic modelling approach to exploratory literature review. J Big Data **6**, 93 (2019). https://doi.org/10.1186/s40537-019-0255-7

Asuncion, A., Welling, M., Smyth, P., & Teh, Y. W. (2009). On smoothing and inference for topic models. In Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence (UAI '09) (pp. 27–34). AUAI Press. Arlington, Virginia.

Baldé, C. P., Kuehr, R., Yamamoto, T., McDonald, R., D'Angelo, E., Althaf, S., Bel, G., Deubzer, O., Fernandez-Cubillo, E., Forti, V., Gray, V., Herat, S., Honda, S., Iattoni, G., Khetriwal, D. S., Luda di Cortemiglia, V., Lobuntsova, Y., Nnorom, I., Pralat, N., & Wagner, M. (2024). Global E-waste Monitor 2024. International Telecommunication Union (ITU) and United Nations Institute for Training and Research (UNITAR). Geneva/Bonn.

Biever, C. (2023, July 25). ChatGPT broke the Turing test — the race is on for new ways to assess AI. Nature News. https://www.nature.com/articles/d41586-023-02361-7

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. Journal of Machine Learning Research, 3(3), 993–1022.

Blei, D.M. (2012) Probabilistic Topic Models. Communications of the ACM, 55, 77-84. http://dx.doi.org/10.1145/2133806.2133826

Bon, A., Prem, E., Lee, E. A., & Ghezzi, C. (2022). Decolonizing technology and society: A perspective from the Global South. In H. Werthner, E. Prem, E. A. Lee, & C. Ghezzi (Eds.), Perspectives on digital humanism (pp. 123-139). Springer. https://doi.org/10.1007/978-3-030-86144-5_9

Boutang, Y. M. (2011). Cognitive Capitalism. Polity Press.

Bovens, M. (2010). Two concepts of accountability: Accountability as a virtue and as a mechanism. West European Politics, 33(5), 946–967. https://doi.org/10.1080/01402382.2010.486119

Bovens, M., Schillemans, T., & Goodin, R. (2014). Public accountability. In M. Bovens, R. Goodin, & T. Schillemans (Eds.), The Oxford handbook of public accountability (online edn). Oxford Academic. https://doi.org/10.1093/oxfordhb/9780199641253.013.0012

Bradford, A. (2020). The Brussels effect: How the European Union rules the world. Oxford Academic. https://doi.org/10.1093/oso/9780190088583.001.0001

Bradford, A. (2023). Digital empires: The global battle to regulate technology. Oxford Academic. https://doi.org/10.1093/oso/9780197649268.001.0001

Broughel, J. (2023, October 5). We should welcome the new AI doomerism. Forbes. https://www.forbes.com/sites/digital-assets/2023/03/30/we-should-welcome-the-new-ai-doomerism/?sh=210d93fd2d84

Byrum, G., & Benjamin, R. (2022, June 16). Disrupting the gospel of tech solutionism to build tech justice. Stanford Social Innovation Review. https://ssir.org/articles/entry/disrupting_the_gospel_of_tech_solutionism_to_build_tech_justice

Calma, J. (2024, January 24). AI and crypto mining are driving up data centers ’energy use. The Verge. https://www.theverge.com/2024/1/24/24049047/data-center-ai-crypto-bitcoin-mining-electricity-report-iea

Conn, A. (2022, June 22). Ai should provide a shared benefit for as many people as possible. Future of Life Institute. https://futureoflife.org/recent-news/shared-benefit-principle/

Couldry, N., & Mejias, U. A. (2019). Data Colonialism: Rethinking Big Data’s Relation to the Contemporary Subject. Television & New Media, 20(4), 336-349. https://doi.org/10.1177/1527476418796632

Cowls, J., Tsamados, A., Taddeo, M., & Floridi, L. (2021). A definition, benchmark and database of AI for social good initiatives. Nature Machine Intelligence, 3(2), 111-115. https://doi.org/10.1038/s42256-021-00296-0

Crawford, K. (2024). Generative AI's environmental costs are soaring - and mostly secret. Nature, 626(8000), 693. https://doi.org/10.1038/d41586-024-00478-x

Crawford, K. (2021). Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence. Yale University Press.

Crawford, K., & Joler, V. (2019). Anatomy of an AI system. Virtual Creativity, 9(1), 117-120. https://doi.org/10.1386/vcr_00008_7

Cuéllar, M.-F., & Sheehan, M. (2023, June 20). AI is winning the AI race. Foreign Policy. https://foreignpolicy.com/2023/06/19/us-china-ai-race-regulation-artificial-intelligence/

Dauvergne, P. (2020). AI in the Wild: Sustainability in the Age of Artificial Intelligence. The MIT Press.

Day, T., Hans, F., Mooldijk, S., Smit, S., Woollands, S., de Grandpré, J., Salsabila, N. P., Fraser, E., Kuramochi, T., & Warnecke, C. (2024). Corporate Climate Responsibility Monitor 2024. NewClimate Institute. http://newclimate.org/publications/

Devlin, H. (2023, May 3). AI "could be as transformative as the Industrial Revolution." The Guardian. https://www.theguardian.com/technology/2023/may/03/ai-could-be-as-transformative-as-industrial-revolution-patrick-vallance

Dwork, C. (2006). Differential privacy. In M. Bugliesi, B. Preneel, V. Sassone, & I. Wegener (Eds.), Automata, languages and programming. ICALP 2006. Lecture Notes in Computer Science (Vol. 4052). Springer, Berlin, Heidelberg. https://doi.org/10.1007/11787006_1

Eilam, T., Bello-Maldonado, P., Bhattacharjee, B., Costa, C., Lee, E. K., & Tantawi, A. (2023). Towards a methodology and framework for AI sustainability metrics. In Proceedings of the 2nd Workshop on Sustainable Computer Systems (HotCarbon '23) (Article 13, pp. 1–7). Association for Computing Machinery. https://doi.org/10.1145/3604930.3605715

European Commission. (2018). Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions: artificial intelligence for Europe (COM/2018/237).

European Commission, Directorate-General for Communications Networks, Content and Technology. (2019). Ethics guidelines for trustworthy AI. Publications Office. https://data.europa.eu/doi/10.2759/346720

Executive Office of the President. (2019). Maintaining American leadership in artificial intelligence. Federal Register, 84(31), 3967–3972. Retrieved 24.06.2024 from https://www.govinfo.gov/content/pkg/FR-2019-02-14/pdf/2019-02544.pdf.

Filippucci, F., et al. (2024). The impact of artificial intelligence on productivity, distribution and growth: Key mechanisms, initial evidence and policy challenges. OECD Artificial Intelligence Papers, No. 15. OECD Publishing. https://doi.org/10.1787/8d900037-en

Floridi, L., Cowls, J., Beltrametti, M., et al. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Minds & Machines, 28, 689–707. https://doi.org/10.1007/s11023-018-9482-5

Friederich, S., & Symons, J. (2023). Operationalising sustainability? Why sustainability fails as an investment criterion for safeguarding the future. Global Policy, 14(1), 61–71. https://doi.org/10.1111/1758-5899.13160

Fukuda-Parr, S., & Gibbons, E. (2021). Emerging consensus on “ethical AI”: human rights critique of stakeholder guidelines. Global Policy, 12(S6), 32–44. https://doi.org/10.1111/1758-5899.12965

Gartenberg, C. (2020, September 14). Microsoft’s Underwater Server Experiment Resurfaces after two years. The Verge. https://www.theverge.com/2020/9/14/21436746/microsoft-project-natick-data-center-server-underwater-cooling-reliability

Goh, H. H., & Vinuesa, R. (2021). Regulating artificial-intelligence applications to achieve the sustainable development goals. Discover Sustainability, 2(52). https://doi.org/10.1007/s43621-021-00064-5

Gupta, A. (2024, April 15). Industry AI “standards” may be a good band-aid, but we need enforceable standards in the Long Run. Tech Policy Press. https://www.techpolicy.press/industry-ai-standards-may-be-a-good-bandaid-but-we-need-enforceable-standards-in-the-long-run/

Hacker, P. (2023). Sustainable AI regulation. http://dx.doi.org/10.2139/ssrn.4467684

Haites, E. (2018). Carbon taxes and greenhouse gas emissions trading systems: what have we learned? Climate Policy, 18(8), 955–966. https://doi.org/10.1080/14693062.2018.1492897

Heacock, M., Kelly, C. B., Asante, K. A., Birnbaum, L. S., Bergman, Å. L., Bruné, M. N., Buka, I., Carpenter, D. O., Chen, A., Huo, X., Kamel, M., Landrigan, P. J., Magalini, F., Diaz-Barriga, F., Neira, M., Omar, M., Pascale, A., Ruchirawat, M., Sly, L., Sly, P. D., … Suk, W. A. (2016). E-Waste and Harm to Vulnerable Populations: A Growing Global Problem. Environmental health perspectives, 124(5), 550–555. https://doi.org/10.1289/ehp.1509699

Heikkilä, M. (2022, November 15). We’re getting a better idea of AI’s true carbon footprint. MIT Technology Review. https://www.technologyreview.com/2022/11/14/1063192/were-getting-a-better-idea-of-ais-true-carbon-footprint

Heikkilä, M. (2023, December 5). AI’s carbon footprint is bigger than you think. MIT Technology Review. https://www.technologyreview.com/2023/12/05/1084417/ais-carbon-footprint-is-bigger-than-you-think/

Henriksen, A., Enni, S., & Bechmann, A. (2021). Situated accountability: Ethical principles, certification standards, and explanation methods in applied AI. In Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society (AIES '21) (pp. 574–585). Association for Computing Machinery. https://doi.org/10.1145/3461702.3462564

International Organization for Standardization. (2015). ISO 14001:2015 Environmental management systems — Requirements with guidance for use. https://www.iso.org/standard/60857.html

Jagarlamudi, J., Daumé, H., & Udupa, R. (2012). Incorporating lexical priors into topic models. In Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics (EACL '12) (pp. 204–213). Association for Computational Linguistics. Retrived from: https://aclanthology.org/E12-1021

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389-399. https://doi.org/10.1038/s42256-019-0088-2

Jones, N. (2018). How to stop data centres from gobbling up the world's electricity. Nature, 561(7722), 163-166. https://doi.org/10.1038/d41586-018-06610-y

Kak, A., West, S. M., & Whittaker, M. (2023, December 5). Make no mistake-AI is owned by Big Tech. MIT Technology Review. https://www.technologyreview.com/2023/12/05/1084393/make-no-mistake-ai-is-owned-by-big-tech/

Katz, Y. (2017). Manufacturing an artificial intelligence revolution. SSRN. https://doi.org/10.2139/ssrn.3078224

Kanapala, A., Pal, S., & Pamula, R. (2019). Text summarization from legal documents: a survey. Artificial Intelligence Review, 51, 371-402. https://doi.org/10.1007/s10462-017-9566-2

Khanal, S., Zhang, H., & Taeihagh, A. (2024). Why and how is the power of Big Tech increasing in the policy process? The case of generative AI. Policy and Society. https://doi.org/10.1093/polsoc/puae012

Kim, W. (2024, March 7). How Nvidia beat everyone else in the AI race. Vox. https://www.vox.com/money/2024/3/7/24092309/nvidia-stock-earnings-valuation-ai-explainer

Le Ludec, C., Cornet, M., & Casilli, A. A. (2023). The problem with annotation: Human labour and outsourcing between France and Madagascar. Big Data & Society, 10(2). https://doi.org/10.1177/20539517231188723

Lovink, G. (2022). Extinction Internet. Institute of Network Cultures.

Luccioni, A. S., Jernite, Y., & Strubell, E. (2024). Power Hungry Processing: Watts Driving the Cost of AI Deployment? In Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24) (pp. 85–99). Association for Computing Machinery. https://doi.org/10.1145/3630106.3658542

Lucero, K. (2019). Artificial intelligence regulation and China's future. Columbia Journal of Asian Law, 33(1), 94–171. https://doi.org/10.7916/cjal.v33i1.5454.

Martin, K. (2019). Ethical implications and accountability of algorithms. Journal of Business Ethics, 160(4), 835–850. https://doi.org/10.1007/s10551-018-3921-3

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955. AI Magazine, 27(4), 12.

Meeks, E., & Weingart, S. B. (2012). The Digital Humanities Contribution to Topic Modeling. Journal of Digital Humanities, 2(1), 1-6.

Metcalf, J., Moss, E., & Boyd, D. (2019). Owning ethics: Corporate logics, Silicon Valley, and the institutionalization of ethics. Social Research, 86(2), 449-476. https://doi.org/10.1353/sor.2019.0022

Metz, C. (2016, April 27). Inside OpenAI, Elon Musk's wild plan to set artificial intelligence free. Wired. https://www.wired.com/2016/04/openai-elon-musk-sam-altman-plan-to-set-artificial-intelligence-free/

Milmo, D. (2024, July 2). Google's emissions climb nearly 50% in five years due to AI Energy Demand. The Guardian. https://www.theguardian.com/technology/article/2024/jul/02/google-ai-emissions

Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. Nature Machine Intelligence, 1(11), 501-507. https://doi.org/10.2139/ssrn.3391293

Munn, L. (2023). The uselessness of AI ethics. AI Ethics, 3(3), 869-877. https://doi.org/10.1007/s43681-022-00209-w

Nolan, D., Maryam, H., & Kleinman, M. (2024, January 29). The urgent but difficult task of regulating artificial intelligence. Amnesty International. https://www.amnesty.org/en/latest/campaigns/2024/01/the-urgent-but-difficult-task-of-regulating-artificial-intelligence/

Nowak, A., Lukowicz, P., & Horodecki, P. (2018). Assessing artificial intelligence for humanity: Will AI be our biggest ever advance or the biggest threat? [Opinion]. IEEE Technology and Society Magazine, 37(4), 26-34. https://doi.org/10.1109/MTS.2018.2876105

Palmer, D. (2010). Text Preprocessing. In F. J. Damerau & N. Indurkhya (Eds.), Handbook of Natural Language Processing (pp. 9–28). Chapter, Taylor & Francis Group. https://doi.org/10.1201/9781420085938

Perrigo, B. (2023, January 18). OpenAI used Kenyan workers on less than $2 per hour. Time. https://time.com/6247678/openai-chatgpt-kenya-workers/

Prescott, A. (2023). Bias in big data, machine learning and AI: What lessons for the digital humanities? Digital Humanities Quarterly, 17(2). https://digitalhumanities.org/dhq/vol/17/2/000689/000689.html

Raper, R., Boeddinghaus, J., Coeckelbergh, M., Gross, W., Campigotto, P., & Lincoln, C. N. (2022). Sustainability budgets: A practical management and governance method for achieving goal 13 of the sustainable development goals for AI development. Sustainability, 14(7), 4019. https://doi.org/10.3390/su14074019

Rogers, R. (2023, April 20). What's AGI, and why are AI experts skeptical? Wired. https://www.wired.com/story/what-is-artificial-general-intelligence-agi-explained/

Rosca, C., Covrig, B., Goanta, C., van Dijck, G., & Spanakis, G. (2020). Return of the AI: An analysis of legal research on Artificial Intelligence using topic modeling. In N. Aletras, I. Androutsopoulos, L. Barrett, A. Meyers, & D. Preoţiuc-Pietro (Eds.), Proceedings of the

Natural Legal Language Processing Workshop 2020 (pp. 3-10). CEUR-WS.org. http://ceur-ws.org/Vol-2645/paper1.pdf

Saner, R., Yiu, L., & Nguyen, M. (2020). Monitoring the SDGs: Digital and social technologies to ensure citizen participation, inclusiveness and transparency. Development Policy Review, 38(4), 483-500. https://doi.org/10.1111/dpr.12433

Saran, S., & Mattoo, S. (2022, February 21). Big Tech vs. Red Tech: The diminishing of democracy in the Digital age. Centre for International Governance Innovation. https://www.cigionline.org/articles/big-tech-vs-red-tech-the-diminishing-of-democracy-in-the-digital-age/

Sætra, H. S. (2021). AI in context and the sustainable development goals: Factoring in the unsustainability of the sociotechnical system. Sustainability, 13(1738). https://doi.org/10.3390/su13041738

Scharre, P., Horowitz, M. C., & Work, R. O. (2018). The artificial intelligence revolution. In Artificial Intelligence: What Every Policymaker Needs to Know (pp. 3–4). Center for a New American Security. http://www.jstor.org/stable/resrep20447.4

Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020). Green AI. Communications of the ACM, 63(12), 54–63. https://doi.org/10.1145/3381831

Sheehan, M. (2024, February 27). Tracing the Roots of China's AI Regulations. Carnegie Endowment for International Peace. https://carnegieendowment.org/research/2024/02/tracing-the-roots-of-chinas-ai-regulations?lang=en

Sorkin, A. R., Mattu, R., Warner, B., Kessler, S., Merced, M. J. D. L., Hirsch, L., & Livni, E. (2023, March 3). Why Lawmakers Aren't Rushing to Police A.I. The New York Times. https://www.nytimes.com/2023/03/03/business/dealbook/lawmakers-ai-regulations.html

Stahl, B. C., Antoniou, J., Ryan, M., & Flick, C. (2022). Organisational responses to the ethical issues of artificial intelligence. AI & Society, 37(1), 23-37. https://doi.org/10.1007/s00146-021-01148-6

State Council of the People's Republic of China. (2017). New Generation Artificial Intelligence Development Plan. Retrieved 24.06.2024 from https://digichina.stanford.edu/work/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/.

Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (pp. 3645-3650). 10.18653/v1/P19-1355

Székely, N., & vom Brocke, J. (2017). What can we learn from corporate sustainability reporting? Deriving propositions for research and practice from over 9,500 corporate sustainability reports published between 1999 and 2015 using topic modelling technique. PLoS ONE, 12(4), e0174807. https://doi.org/10.1371/journal.pone.0174807

Theodorou, A., & Dignum, V. (2020). Towards ethical and socio-legal governance in AI. Nature Machine Intelligence, 2(1), 10-12. https://doi.org/10.1038/s42256-019-0136-y

Truby, J. (2020). Governing artificial intelligence to benefit the UN Sustainable Development Goals. Sustainable Development, 28, 946–959. https://doi.org/10.1002/sd.2048

Tubaro, P., Casilli, A. A., & Coville, M. (2020). The trainer, the verifier, the imitator: Three ways in which human platform workers support artificial intelligence. Big Data & Society, 7(1). https://doi.org/10.1177/2053951720919776

United Nations. (2015). Transforming Our World: the 2030 Agenda for Sustainable Development. https://doi.org/10.1201/b20466-7.

United Nations. (2023, October 26). New UN advisory body aims to harness AI for the common good. UN News. https://news.un.org/en/story/2023/10/1142867

Mavroudis, V. (2024). International scientific report on the safety of advanced AI (2024/009). 132. Retrived from: https://www.gov.uk/government/publications/international-scientific-report-on-the-safety-of-advanced-ai

Vezzoli, C. (2018). Design for environmental sustainability. Springer. https://doi.org/10.1007/978-1-4471-7364-9

Vinuesa, R., Azizpour, H., Leite, I., et al. (2020). The role of artificial intelligence in achieving the sustainable development goals. Nature Communications, 11(233). https://doi.org/10.1038/s41467-019-14108-y

de Vries, A. (2023). The growing energy demands of artificial intelligence. Joule, 7(2), 2191–2194. https://doi.org/10.1016/j.joule.2023.09.004

Watanabe, K., & Baturo, A. (2024). Seeded Sequential LDA: A semi-supervised algorithm for topic-specific analysis of sentences. Social Science Computer Review, 42(1), 224-248. https://doi.org/10.1177/08944393231178605

Werthner, H. (2020). The Vienna Manifesto on Digital Humanism. In M. Hengstschläger (Ed.), Digital Transformation and Ethics (pp. 338–357). Ecowin. http://hdl.handle.net/20.500.12708/30434

Weston, S., Shryock, I., Light, R., & Fisher, P. (2023). Selecting the number and labels of topics in topic modeling: A tutorial. Advances in Methods and Practices in Psychological Science, 6(1), 251524592311601. https://doi.org/10.1177/25152459231160105

Whittlestone, J., Nyrup, R., Alexandrova, A., & Cave, S. (2019). The role and limits of principles in AI ethics: Towards a focus on tensions. In Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society (AIES '19) (pp. 195-200). Association for Computing Machinery. https://doi.org/10.1145/3306618.3314289

Yigitcanlar, T., Desouza, K. C., Butler, L., & Roozkhosh, F. (2020). Contributions and risks of artificial intelligence (AI) in building smarter cities: Insights from a systematic review of the literature. Energies, 13(6), 1473. https://doi.org/10.3390/en13061473

Yu, D., Rosenfeld, H., & Gupta, A. (2023, January 16). The "Ai divide" between the Global North and Global South. World Economic Forum. https://www.weforum.org/agenda/2023/01/davos23-ai-divide-global-north-global-south/

Zapp, P., Schreiber, A., Marx, J., & Walachowicz, F. (2022). Environmental impacts of rare earth production. MRS Bulletin, 47(3), 267-275. https://doi.org/10.1557/s43577-022-00286-6

Zhuk, A. (2023). Artificial intelligence impact on the environment: Hidden ecological costs and ethical-legal issues. Journal of Digital Technologies and Law, 1(4), 932-954. https://doi.org/10.21202/jdtl.2023.40

Appendix: *Corpora Used in Computational Study with Seeded LDA Topic Modeling*

European Union Dataset:

1. Data Governance Act
2. European Data Act
3. A European Strategy for Data
4. Open Data Directive
5. Commission Regulation (EU) 2019/424: Ecodesign Requirements for Servers and Data Storage Products
6. Open Data and the Re-Use of Public Sector Information
7. Ethics Guidelines on Artificial Intelligence
8. AI Act: Harmonised Rules on Artificial Intelligence
9. 2030 Digital Compass: The European Way for the Digital Decade
10. Digital Decade Policy Programme 2030
11. Shaping Europe’s Digital Future
12. The European Green Deal
13. A Green Deal Industrial Plan for the Net-Zero Age
14. A New Industrial Strategy for Europe
15. EGDC Net Carbon Impact Assessment Methodology for ICT Solutions
16. Regulation of the European Parliament and of the Council on Machinery Products
17. European Climate Pact
18. Fostering a European Approach to Artificial Intelligence
19. White Paper on Artificial Intelligence
20. Artificial Intelligence for Europe
21. An SME Strategy for a Sustainable and Digital Europe
22. Communication on Boosting Startups and Innovation in Trustworthy Artificial Intelligence
23. Coordinated Plan on Artificial Intelligence
24. The New European Consensus on Development
25. Council of Europe: The Framework Convention on Artificial Intelligence
26. AI Liability Directive

United States of America Dataset:

1. Public Views on AI and Intellectual Property Policy
2. AI Accountability Framework for Federal Agencies and Other Entities
3. Algorithmic Accountability Act of 2022
4. Illinois Artificial Intelligence Video Interview Act
5. Federal Automated Vehicles Policy
6. Executive Order on Promoting Trustworthy AI in the Federal Government
7. Executive Order on Safe, Secure, and Trustworthy Development and Use of AI
8. Maintaining American Leadership in AI
9. National Defense Authorization Act for Fiscal Year 2021 (Selected Pages: 93-108, 344-362, 499-503, 757-758, 1137-1161, 1178-1190, 1239-1247, 1379-1427)
10. National Defense Authorization Act 2019 (Selected Pages: 47-71, 327-330, 491-492, 502-503)
11. Artificial Intelligence Training for the Acquisition Workforce Act
12. Americans With Disabilities Act and AI Use in Job Assessment
13. National AI Research and Development Strategic Plan 2023
14. NIST Four Principles of Explainable AI
15. U.S. Department of Defense Responsible AI Strategy and Implementation Pathway
16. Request for Information: National Priorities for AI
17. NTIA AI Accountability Policy Report
18. National AI Initiative Act
19. AI in Government Act
20. FTC Act, Section 5
21. Global Technology Leadership Act
22. Transparent Automated Governance Act
23. AI Risk Management Framework
24. Guidance for Regulation of AI Applications

People's Republic of China Dataset:

1. Provisions on the Administration of Deep Synthesis Internet Information Services (Retrived from: https://digichina.stanford.edu/work/translation-internet-information-service-deep-synthesis-management-provisions-draft-for-comment-jan-2022/)
2. Governance Principles for a New Generation of Artificial Intelligence: Develop Responsible Artificial Intelligence (Retrived from: https://perma.cc/V9FL-H6J7)
3. Next Generation Artificial Intelligence Development Plan (Retrived From: https://digichina.stanford.edu/work/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/)
4. Internet Information Service Algorithmic Recommendation Management Provisions (Retrived from: https://digichina.stanford.edu/work/translation-internet-information-service-algorithmic-recommendation-management-provisions-effective-march-1-2022/)
5. Interim Measures for the Management of Generative Artificial Intelligence Services (Retrived from: https://www.chinalawtranslate.com/en/generative-ai-interim/)
6. White Paper on Trustworthy Artificial Intelligence (Retrived from: https://cset.georgetown.edu/publication/white-paper-on-trustworthy-artificial-intelligence/)
7. Law of the People's Republic of China on Progress of Science and Technology (Retrived from: https://cset.georgetown.edu/publication/law-of-the-peoples-republic-of-china-on-progress-of-science-and-technology/)
8. Regulations for the Promotion of the Artificial Intelligence Industry in Shenzhen Special Economic Zone (Retrived from: https://cset.georgetown.edu/publication/regulations-for-the-promotion-of-the-artificial-intelligence-industry-in-shenzhen-special-economic-zone/)
9. "Internet+" Artificial Intelligence Three-Year Action and Implementation Plan (Retrived from: https://cset.georgetown.edu/publication/internet-artificial-intelligence-three-year-action-and-implementation-plan/)
10. Artificial Intelligence Standardization White Paper 2021 Edition (Retrived from: https://cset.georgetown.edu/publication/artificial-intelligence-standardization-white-paper-2021-edition/)
11. Guidelines for National New Generation Artificial Intelligence Innovation and Development Pilot Zone Construction Work (Retrived from:

https://cset.georgetown.edu/publication/guidelines-for-national-new-generation-artificial-intelligence-innovation-and-development-pilot-zone-construction-work/)

12. Outline of the People's Republic of China 14th Five-Year Plan for National Economic and Social Development and Long-Range Objectives for 2035 (Retrived from: https://cset.georgetown.edu/publication/china-14th-five-year-plan/)
13. Data Security Law of the People's Republic of China (Retrived from: https://digichina.stanford.edu/work/translation-data-security-law-of-the-peoples-republic-of-china/)
14. Law of the People's Republic of China on Progress of Science and Technology (Retrived from: https://cset.georgetown.edu/publication/law-of-the-peoples-republic-of-china-on-progress-of-science-and-technology/)