INSTITUTE FOR GLOBAL PROSPERITY

Dissertation Title

A Comparative Analysis of AI/AGI Regulation in the EU and China and the Implications for Global Governance

Student Name: **Priyasha Sharma**

Module Code: **BGLP0014**

Word Count: **15,000**

IGP MSc COURSEWORK

DECLARATION OF OWNERSHIP AND COPYRIGHT FORM

1. DECLARATION OF OWNERSHIP

I confirm that I have read and understood the guidelines on plagiarism produced by IGP and UCL, that I understand the meaning of plagiarism as defined in those guidelines, and that I may be penalised for submitting work that has been plagiarised.

This piece of coursework must be submitted electronically through Turnitin on Moodle by the stipulated deadline. I understand that the coursework cannot be assessed unless it is submitted online and that penalties will be applied for late submissions as per UCL and IGP guidelines unless there is an approved case for Extenuating Circumstances or Reasonable Adjustments.

I declare that all material is entirely my own work except where explicitly, clearly and individually indicated and that all sources used in its preparation and all quotations are clearly cited using a recognised system for referencing and citation. Should this statement prove to be untrue, I recognise the right of the Board of Examiners to recommend disciplinary action in line with UCL regulations.

2. COPYRIGHT

The copyright of the coursework remains with me as its author. However, I understand that anonymised copies may be made available to future students for reference. Please, tick the box if you DO NOT want this report to be made available for teaching purposes.

☐

Abstract

The proliferation of artificial intelligence (AI) systems offers significant benefits but also presents considerable global risks. Despite the international nature of these risks, an internationally accepted governance framework for AI remains absent. This is largely due to geopolitical tensions and the lack of consensus among major global powers on the appropriate scope and methods for regulating AI development. The European Union (EU), China, and the United States are key players in AI regulation, each with divergent approaches. However, common themes can be identified to form the foundation of a global regulatory framework aimed at both managing risks and fostering innovation.

This dissertation conducts a comparative analysis of AI regulations in the EU and China and examines the implications for the development of a global AI governance framework. Using deductive thematic analyses, the study explores the convergences and divergences in their regulatory approaches. While differences arise from China's emphasis on socialist values, national security, minimal focus on individual rights, and geopolitical rivalry with the United States, there are areas of alignment. These include the categorisation of AI systems, safety measures, human-centric design, transparency, professional responsibility, privacy, and accountability.

The study concludes that these shared regulatory principles could serve as a foundation for the development of a multilevel global AI framework. Such a framework should incorporate international, national, and industry-level governance with clearly defined outcomes and responsibilities. The United Nations, as a politically legitimate global forum with broad influence, including among China, is well-positioned to lead this effort.

Acknowledgements

I would like to extend my deepest gratitude to those who have supported and guided me throughout the process of completing this dissertation.

First and foremost, I owe a great deal of thanks to my supervisor, Dr. Christopher Harker, whose invaluable expertise, thoughtful feedback, and consistent encouragement have been central to shaping this work. His insights and advice were indispensable, and I am deeply appreciative of his support.

I would also like to express my sincere appreciation to the module leaders, Dr. Mara Torres and Dr. Yuan He, whose guidance and knowledge have greatly contributed to my academic journey. Their input has been crucial to the development of this dissertation.

A special acknowledgement is reserved for Mr. Robert Whitfield, Chairperson of World One Trust, for generously sharing his expertise in the AI/AGI regulatory space. His collaboration and support under the Community Research Initiative (CRIS) at UCL were instrumental in refining my research focus.

I would also like to recognise the valuable discussions and encouragement from my peers, whose input provided fresh perspectives and motivated me to stay on course.

Finally, my deepest gratitude goes to my family and friends. Their unwavering support, encouragement, and understanding have been a vital source of strength throughout this journey.

Thank you to all who played a part in the successful completion of this dissertation.

Table of Contents

List of Figures and Tables

List of Abbreviations

AI - Artificial Intelligence

AGI - Artificial General Intelligence

CAC - Cyberspace Administration of China

CCP - Chinese Communist Party

CoE - Council of Europe

EU - European Union

G20 - Group of Twenty

G7 - Group of Seven

GDPR - General Data Protection Regulation

GPAI - Global Partnership on AI

HK - Refers to the person providing feedback (assumed "HK" in context refers to feedback from a supervisor)

IEC - International Electrotechnical Commission

ISO - International Organisation for Standardisation

LAWS - Lethal Autonomous Weapons Systems

MAIL - Model Artificial Intelligence Law

MIIT - Ministry of Industry and Information Technology

MPS - Ministry of Public Security

NSCAI - National Security Commission on Artificial Intelligence

OECD - Organisation for Economic Co-operation and Development

OSTP - Office of Science and Technology Policy

PAI - Partnership on AI

PIPL - Personal Information Protection Law

SAMR - State Administration for Market Regulation

SDGs - Sustainable Development Goals

UN - United Nations

UNICRI - United Nations Interregional Crime and Justice Research Institute

UNIDIR - United Nations Institute for Disarmament Research

WEF - World Economic Forum

NSF - National Science Foundation

AI Act - Artificial Intelligence Act (European Union)

GDPR - General Data Protection Regulation (European Union)

AI Bill of Rights - Artificial Intelligence Bill of Rights (United States)

NSCAI - National Security Commission on Artificial Intelligence (United States)

STI - Science, Technology, and Innovation

ELVIS Act - State legislation targeting audio deepfakes in Tennessee (United States)

BRI - Belt and Road Initiative (China)

GPT - Generative Pre-trained Transformer

ISO/IEC - International Organisation for Standardisation / International Electrotechnical Commission

AI Index - Annual report tracking AI developments by Stanford University

NIST - National Institute of Standards and Technology (United States)

OECD AI Principles - Organisation for Economic Co-operation and Development AI Principles

UNICRI - United Nations Interregional Crime and Justice Research Institute

Deepfake - Synthetically generated media created using artificial intelligence

NSF - National Science Foundation (United States)

AI HLEG - High-Level Expert Group on Artificial Intelligence (European Commission)

Chapter 1: Introduction

1. Aim and Objectives

The primary aim of this dissertation is to conduct a rigorous comparative analysis of the regulatory frameworks for Artificial Intelligence (AI) and Artificial General Intelligence (AGI) within the European Union (EU) and China.

This analysis seeks to uncover and articulate the distinct approaches each region employs to manage the dual challenges of fostering innovation and mitigating risks associated with AI/AGI technologies. Through this comparative lens, the study aims to derive actionable insights and lessons that could inform the development of a globally applicable governance framework for AI/AGI.

1.1 Aim of the study

To derive actionable insights and lessons that could inform the development of a globally applicable governance framework for AI/AGI.

1.2 Objectives of the Study

1.2.1 **To analyse the regulatory frameworks in the EU and China by identifying and comparing key themes.**

Using deductive thematic coding, we will systematically identify and compare key thematic areas within these regulatory frameworks. Focus areas will include ethical considerations, data privacy, and security measures, highlighting how each region balances these against the imperatives of technological advancement and innovation.
.

1.2.2 **To develop a comprehensive, cohesive global governance framework for AI/AGI.**

By synthesising the findings from the comparative analysis, we will propose a comprehensive, cohesive global governance framework for AI/AGI. This framework will integrate the strengths and mitigate the weaknesses of the examined regional approaches, aiming to facilitate international cooperation and effective risk management on a global scale.

2. Primary Research Questions

The dissertation is guided by the following primary research questions, which are designed to dissect and understand the complexities of AI/AGI regulation in two of the world's leading regions—the EU and China:

1) **How do the EU and China regulate AI/AGI?**
    a. **How do the EU and China AI/AGI regulations promote innovation, and mitigate risk?**
    b. **How do their regulations compare and contrast with each other?**

2) **What lessons can be learned from the EU and China's approaches to AI/AGI governance that could inform the development of a global governance framework for AGI?**

By addressing these questions, the study aims to contribute to the formulation of a regulatory strategy that not only addresses the specific needs and contexts of different regions but also encompasses a broader global perspective essential for managing the advancing technology of AI/AGI.

3. Importance and Contribution of the Study

This dissertation contributes to academic knowledge by offering a comparative analysis of how the European Union (EU) and China address the dual challenges of fostering innovation and mitigating the risks associated with AI/AGI technologies. Building on the scholarship of Bostrom (2014), who discusses the existential risks posed by AGI, and Müller & Bostrom (2016), who highlight the regulatory challenges surrounding AGI's autonomy, this study critically examines the regulatory frameworks governing AI/AGI in different geopolitical contexts. The study's interdisciplinary approach bridges technology studies, ethics, and regulatory governance, contributing to the development of a comprehensive global governance model for AI/AGI. This follows the principles outlined by Floridi et al. (2018), who emphasise the need for transparency, accountability, and fairness in AI systems while extending these ideas through a comparative evaluation of EU and Chinese strategies. Furthermore, the research contributes to existing discussions in the literature by comparing different cultural, economic, and political approaches to AI governance, as noted by Tallberg et al. (2023) and Goertzel & Pennachin (2007). The study also extends theoretical models of global governance in the context of rapidly evolving AI/AGI technologies, drawing

on global frameworks like those explored by Brundage et al. (2018), Jobin, Ienca, & Vayena (2019), and Hagendorff (2020). By synthesising these insights, the research aims to propose a governance model that integrates the strengths of EU and Chinese approaches, while addressing key challenges such as privacy, security, and ethical concerns. This fills a critical gap in the literature where comparative studies of AI/AGI governance between major global powers remain limited.

This dissertation serves as both a theoretical and practical contribution, advancing academic discussions on AI governance while offering actionable recommendations for policymakers. The dual focus is essential for ensuring that AI/AGI technologies are aligned with ethical standards, societal values, and international cooperation efforts, as emphasised by Fjeld et al. (2020) and Floridi (2021). Ultimately, the study aims to promote a balanced approach that fosters innovation while managing risks on a global scale.

Chapter 2: Literature Review

4. Background and Literature Review

The rapid advancement of Artificial Intelligence (AI) and its evolution towards Artificial General Intelligence (AGI) poses unique challenges that necessitate robust and adaptable governance frameworks (Tallberg, et. al, 2023). This research is driven by the critical need to examine and understand how such technologies, which hold the potential to surpass human capabilities and control, are regulated across different global contexts.

AI and AGI represent transformative technologies that promise substantial benefits but also pose significant risks, including ethical dilemmas, data privacy issues, and potential economic disruptions. The European Union, known for its comprehensive privacy laws and ethical standards, offers a contrast to China's more pragmatic, development-focused AI governance model. This contrast provides fertile ground for analysis to understand how different cultural, economic, and political contexts influence AI regulation.

4.1 The Evolution from AI to AGI: Capabilities and Challenges

AI encompasses technologies that perform tasks requiring human intelligence, such as visual perception, decision-making, and language translation. Narrow AI excels in specific tasks but cannot generalise across different domains (Russell & Norvig, 2016). In contrast, AGI, also referred to as broad AI, aims to understand, learn, and apply knowledge across diverse tasks, mirroring human cognitive abilities (Goertzel & Pennachin, 2007). This distinction is crucial, as AGI's potential for autonomy and decision-making presents unique regulatory challenges (Bostrom, 2014).

AGI's transformative potential spans various sectors not limited to but including healthcare, finance, manufacturing, and logistics, promising significant advancements in healthcare and finance. In healthcare, AGI could enhance diagnostic accuracy and personalised medicine (Esteva et al., 2019). In finance, AGI's capabilities could improve risk assessment and fraud detection, leading to more efficient operations (Davenport & Ronanki, 2018). However, the autonomous nature of AGI also raises ethical concerns, including accountability, transparency, and bias (Müller & Bostrom, 2016). Scholars emphasise the need for regulatory frameworks to balance these impacts with ethical standards (Floridi et al., 2018).

The emergence of AGI necessitates robust regulatory measures to mitigate associated risks. Concerns include privacy breaches, where sensitive data handled by AGI could be compromised, and security vulnerabilities that could be exploited to manipulate or disrupt AGI operations (Brundage et al., 2018). Regulatory frameworks must establish standards for transparency, accountability, and fairness in AI development and deployment (Jobin, Ienca, & Vayena, 2019). Accountability in AGI systems are crucial for ensuring that actions taken by these systems are traceable and that operators can be held responsible for the outcomes (Müller & Bostrom, 2016). Transparency is needed to ensure that stakeholders understand how decisions are made by AGI systems, which is essential for public trust and governance (Jobin, Ienca, & Vayena, 2019). Additionally, mitigating bias is imperative to prevent discriminatory outcomes and ensure fairness in automated decisions, which is especially challenging as biases can be embedded in the data used by AGI systems (Floridi et al., 2018).

Furthermore, ethical dilemmas such as the potential for AGI to exacerbate job displacement due to automation, the exacerbation of existing biases, and the erosion of public trust highlight the need for comprehensive regulatory frameworks. These frameworks must establish clear standards for transparency, accountability, and fairness in AI development and deployment to ensure that AGI aligns with societal values and ethical principles (Brynjolfsson & McAfee, 2014).

4.2 Tracing the Path: Historical Perspectives on AI and AGI Regulations

The evolution of AI regulations has been marked by incremental advancements, beginning with basic guidelines and eventually evolving into more structured regulatory frameworks. In the early stages, AI development was primarily guided by ethical codes and informal standards established by research institutions and industry bodies (Calo, 2017). However, as AI technologies advanced and their societal impacts became more apparent, governments increasingly recognised the need for formal legislative measures to oversee AI development and deployment (Cath, 2018). As reported by Stanford University's 2023 AI Index, the number of legislative bills referencing "artificial intelligence" across 127 surveyed nations saw a sharp rise, growing from just one in 2016 to 37 by 2022 (Artificial Intelligence Index Report 2023).

While national and regional efforts have made substantial progress, there have also been attempts to create global frameworks for AI governance. The idea of a global governance body to regulate AI development was proposed as early as 2017. Notable global initiatives include the OECD AI Principles (adopted in 2019), the G20 AI Principles (adopted in 2019), and the World Economic Forum’s ten 'AI Government Procurement Guidelines' (issued in 2019) (Campbell, 2019). Other significant initiatives include the Global Partnership on AI (GPAI), launched in 2020, and UNESCO’s international instrument on the ethics of AI, introduced in 2021. These global efforts

signal the growing recognition of AI as a technology that transcends national boundaries and requires coordinated international governance.

Over the last decade, many national and regional authorities have developed comprehensive strategies, action plans, and policy papers on AI and AGI. These documents often cover a broad range of topics, including regulation and governance, industrial strategy, research, talent development, and infrastructure (Bradford, 2023). The three largest economies—China, the USA, and the EU—have adopted distinct approaches to AI regulation. The USA follows a market-driven approach, China opts for a state-driven model, and the EU pursues a rights-driven approach (Campbell, 2019).

In Europe, individual countries within the EU have developed their own national AI strategies, complementing the broader European strategy. The General Data Protection Regulation (GDPR), introduced in 2016, stands as a significant milestone in AI regulation, setting comprehensive standards for data privacy and protection, which directly affect AI systems handling personal data (Voigt & Von dem Bussche, 2017). Additionally, the European Commission's publications, such as the "Ethics Guidelines for Trustworthy Artificial Intelligence" and "Policy and Investment Recommendations for Trustworthy AI," published in 2019, have set ethical standards for AI governance (European Commission, 2020). The AI Act, first proposed in 2021 and adopted in May 2024, marked a pivotal moment in Europe’s regulatory approach by introducing a risk-based framework for AI regulation, categorising AI into minimal, limited, high, and unacceptable risk categories, while specifically addressing general-purpose AI systems like ChatGPT.

The United States has also taken an active approach to AI regulation since 2016. The National Science and Technology Council's report, *Preparing for the Future of Artificial Intelligence* (2016), underscored the need for continued AI development with few regulatory restrictions. In 2019, the White House Office of Science and Technology Policy published a draft document titled "Guidance for Regulation of Artificial Intelligence Applications," which proposed ten key principles for federal agencies to consider when regulating AI. The National Security Commission on Artificial Intelligence’s 2021 report further solidified the USA’s stance on AI by recommending substantial investments in AI technologies to align AI uses with national values (NSCAI, 2021). President Biden’s *AI Bill of Rights*, released in 2022, and his Executive Order on *Safe, Secure, and Trustworthy Artificial Intelligence* in 2023, addressed core issues like algorithmic bias, data privacy, and the safety of AI systems (White House, 2022). Additionally, individual states, such as California and Utah, have passed AI-specific bills aimed at addressing emerging AI challenges, including deepfakes and voice cloning technologies.

In contrast, China’s regulatory landscape was relatively lax before 2020, as the government prioritised AI development over regulation. The *New Generation AI Development Plan*, released by the State Council in 2017, was designed to promote rapid AI growth (CCP, 2020). Early regulatory frameworks, such as the *Governance Principles for New Generation AI* (2019), provided general guidelines for responsible AI governance, emphasising privacy, security, and agile governance. However, as AI technologies advanced, China introduced more targeted regulatory frameworks. The 2021 *Provisions on the Management of Algorithmic Recommendations* marked China’s first binding regulation on algorithms, motivated by concerns over online content dissemination and the ethical implications of AI-driven recommendations (CAC, 2022). The draft *Measures for the Management of Generative AI Services*, released in 2023, focused on text generation and training data, underscoring China’s commitment to managing both the technical and ethical dimensions of AI (Natlareview, 2023).

China's *Model Artificial Intelligence Law (MAIL)*, released for comments in April 2024, continues the trend of regulating AI with a focus on ensuring accuracy and control over training data and content generation. Additionally, the *Algorithmic Accountability Act* introduced measures to promote transparency and accountability in automated decision-making systems, emphasising the need for formalised AI oversight (Diakopoulos, 2016). Together, these legislative actions underscore China's transition from a largely laissez-faire approach to a more controlled regulatory framework that balances innovation with governance.

These legislative developments across different jurisdictions reflect a growing consensus on the importance of formal AI oversight, though the approaches differ significantly due to varying political, economic, and cultural factors.

4.3 Broader Studies of Policy Making and Innovation in EU and China

The contrasting approaches to AI governance between the European Union (EU) and China epitomise the impact of distinct socio-political environments on regulatory frameworks. The EU’s regulatory architecture, particularly exemplified by the General Data Protection Regulation (GDPR), places a significant emphasis on individual rights, advocating for stringent measures regarding transparency, accountability, and the safeguarding of personal data (Voigt & Von dem Bussche, 2017). The forthcoming AI Act continues this trajectory towards comprehensive regulation, aiming to standardise AI practices across member states by classifying AI systems based on their risk level and instituting stringent controls on high-risk applications (European Commission, 2021). This serves as a protective mechanism for citizens and sets a normative standard that could influence global AI governance.

However, this may appear as pre-emptive of the results in some aspects of AI governance, particularly when considering the EU's focus on fostering trust and accountability in AI. To avoid overgeneralisation before empirical analysis, the role of the EU AI Act should be contextualised as a regulatory model that aims to reconcile innovation with strict governance, without assuming its outcomes on innovation are inherently restrictive, as the outcomes may vary depending on the sectors and risk levels involved.

In stark contrast, China's approach to AI governance, particularly encapsulated in the New Generation Artificial Intelligence Development Plan, focuses on rapid technological deployment and economic gains. While this plan acknowledges the ethical use of AI, its emphasis on pragmatic, state-led innovation frequently prioritises economic development over individual privacy concerns (Liang et al., 2018). This model of governance reflects a strategic imperative for China to leverage AI as a tool for achieving global leadership, with less stringent regulatory measures compared to those seen in the EU (Lee, 2018).

The divergence between the EU and Chinese governance models reflects deeper philosophical, political, and cultural orientations towards technology and state control. The EU's approach, deeply rooted in democratic ideals, focuses on the protection of individual rights, ethical considerations, and the development of a trustworthy AI ecosystem. This is evidenced by the detailed procedural standards within the AI Act and GDPR, which both aim to hold AI systems and their developers accountable, safeguarding citizens from potential risks (Floridi et al., 2018). In contrast, China's model, characterised by a top-down, state-controlled strategy, aligns with its broader utilitarian approach, where innovation is guided by state interests to promote economic growth and maintain social stability (Liang et al., 2018; Lee, 2018). This approach prioritises rapid deployment and economic competitiveness but raises questions regarding privacy, human rights, and ethical considerations (Bayamlioglu et al., 2018).

Scholars highlight that both models present their own sets of advantages and limitations. The EU's regulatory framework, while ensuring the protection of individual rights and ethics, could potentially slow down the pace of AI innovation due to bureaucratic hurdles and the associated costs of compliance (Floridi et al., 2018). On the other hand, China's flexible, innovation-driven strategy fosters rapid technological growth but raises significant concerns regarding the infringement of privacy and the potential for state surveillance (Ding, 2018). This disparity is not merely a reflection of technological priorities but also demonstrates how governance structures can shape the trajectory of technological development. Scholars argue that while the EU's regulatory approach may slow down innovation in some domains, it sets a global

benchmark for ethical AI deployment, fostering long-term sustainability through public trust and broad acceptance (Floridi et al., 2018; Voigt & Von dem Bussche, 2017).

Conversely, China's approach demonstrates how less restrictive regulatory environments, combined with aggressive state support, can drive technological advancements and ensure competitiveness in the global AI race (Ding, 2018; Lee, 2018). However, this model's emphasis on economic utility over ethical considerations has been criticised for its potential long-term societal risks, such as erosion of privacy and individual freedoms (Bayamlioglu et al., 2018).

The juxtaposition of these two models suggests that a balanced approach could offer a more comprehensive framework for global AI governance. Such a framework might integrate the EU's ethical rigour and strong regulatory oversight with China's innovation-centric policies to create a regulatory structure that supports both responsible innovation and global competitiveness (Allen, 2019). This balance would be critical for addressing the global challenges posed by AI technologies while also fostering sustainable development.

4.4 Academic Debates on Responsible Innovation and Public Deliberation in AI

The concept of responsible innovation has become increasingly crucial in ensuring that AI technologies develop in alignment with societal values and ethical standards. Stilgoe, Owen, and Macnaghten (2013) propose a framework for responsible innovation based on four key principles: anticipation, reflexivity, inclusivity, and responsiveness. These principles serve as a guide for embedding ethical considerations into technological innovation processes.

Anticipation involves identifying and managing potential impacts, risks, and ethical issues proactively before they manifest. In the context of AI, anticipation requires stakeholders to consider long-term consequences, both intended and unintended, such as the potential for bias in algorithmic decision-making or the displacement of human labor. This is especially pertinent in the development of AGI, where the broader societal and existential risks may only become evident after deployment. Anticipation ensures that innovation does not progress at the expense of societal well-being (Stilgoe et al., 2013; Bostrom, 2014).

Reflexivity calls for innovators to critically evaluate their own assumptions, values, and practices, fostering a culture of ethical self-awareness within development teams. Reflexivity in AI governance highlights the necessity for continual introspection among AI developers, policymakers, and stakeholders to question not only what is being developed but also how and why it is being developed (Owen et al., 2013). Reflexive governance can help address the inherent biases that arise from the data used to train AI systems or from the institutional norms that shape innovation trajectories. Scholars

such as Fisher and Rip (2013) emphasise the challenge of embedding reflexivity in environments dominated by commercial and competitive pressures, which often limit the capacity for ethical reflection.

Responsiveness requires adaptability and the willingness to modify innovation practices in response to new information or societal concerns. For AI governance, responsiveness involves creating regulatory and ethical frameworks that can evolve as AI technologies and their societal impacts become clearer. The dynamic nature of AI demands that governance frameworks be flexible enough to respond to emerging risks while still fostering innovation. This is particularly challenging in the fast-paced AI industry, where regulation often lags behind technological developments (Cath, 2018).

While these principles provide a theoretically robust framework, their practical implementation in the AI industry faces significant challenges. The rapid pace of AI development often prioritises speed and market competitiveness over comprehensive ethical deliberation (Fisher & Rip, 2013). This tension between innovation and ethical responsibility underscores the need for governance frameworks that balance the demands of technological advancement with ethical rigor.

Inclusivity, the most commonly discussed principle, emphasises the engagement of a broad range of stakeholders, including marginalised groups that may be disproportionately affected by AI technologies. Effective inclusivity requires integrating diverse perspectives, particularly from those who are not traditionally represented in the technology development process, such as minority communities, civil society organisations, and laypersons (Stilgoe et al., 2013). Inclusivity in AI governance enhances legitimacy and promotes fairness, ensuring that AI systems serve the interests of all societal groups. However, Fisher and Rip (2013) argue that inclusivity remains underdeveloped in many AI innovation processes, as public participation is often superficial and lacks substantive influence over decision-making.

Public deliberation is another critical element of responsible AI governance, closely tied to inclusivity. Macnaghten, Kearnes, and Wynne (2005) contend that engaging the public in discussions about emerging technologies ensures that developments reflect societal values. Public deliberation fosters transparency, legitimacy, and trust in AI governance processes. However, as Jasanoff (2016) notes, many public engagement efforts have been criticised for being perfunctory, serving as procedural formalities rather than genuine efforts to incorporate public input. This critique calls for more meaningful public engagement mechanisms that empower stakeholders to shape the course of AI development.

Recent scholarship has highlighted the broader complexities and challenges surrounding responsible innovation and public deliberation in AI governance. Rather than focusing solely on individual studies, the central concerns across this body of literature revolve around accountability, transparency, fairness, and the long-term

ethical risks posed by AI. Cath et al. (2018) and Binns (2021) emphasise the need for governance models that ensure transparency and accountability, particularly in the use of algorithmic decision-making systems. These scholars argue that opaque AI systems can perpetuate biases and inequalities, thereby undermining the ethical use of AI. Furthermore, existential risks associated with superintelligent AI, as discussed by Bostrom (2014) and Yudkowsky (2006), underscore the necessity for robust, forward-looking governance frameworks capable of mitigating catastrophic outcomes. Together, these debates illustrate the urgent need for interdisciplinary approaches to AI governance that integrate ethical, social, and technical considerations.

Integrating diverse disciplinary perspectives is essential for developing adaptive and comprehensive AI governance frameworks. Haenlein et al. (2019) advocate for a multi-disciplinary approach that involves technologists, ethicists, sociologists, and policymakers, recognising the multifaceted nature of AI's impact on society. By drawing on insights from various fields, a multi-disciplinary governance model can anticipate and address complex ethical, social, and technical challenges, ensuring that AI technologies evolve in a manner that is both innovative and socially responsible.

The concept of **responsible innovation**, when combined with meaningful **public deliberation**, forms the backbone of a robust AI governance framework. Together, these principles ensure that AI development aligns with societal values and ethical standards, fostering trust and social acceptance. To navigate the challenges posed by AI development, particularly with regard to AGI, governance frameworks must balance the demands of rapid technological innovation with the imperatives of ethical responsibility and public accountability.

4.5 Navigating Ethical Quandaries in AI Regulation

As AI systems become increasingly embedded in critical decision-making processes, the ethical principles of transparency, accountability, and fairness are central to the development and implementation of regulatory frameworks. These principles are also deeply intertwined with the framework of responsible innovation presented earlier, where **anticipation**, **reflexivity**, **inclusivity**, and **responsiveness** form the foundation for navigating the ethical quandaries posed by AI. The principles of transparency, accountability, and fairness provide a practical manifestation of how responsible innovation is operationalised within AI regulations (Stilgoe, Owen, & Macnaghten, 2013).

Transparency in AI regulation involves making AI systems explainable and understandable to both users and regulators. This is critical in fostering public trust and ensuring that the systems operate in line with societal values. Floridi et al. (2018) argue that transparency allows stakeholders to scrutinize AI systems, identifying

potential biases and ensuring that decisions can be understood and questioned. However, transparency in AI is a complex challenge because many AI systems, particularly those involving deep learning, operate as "black boxes," where even the developers cannot fully explain how decisions are made (Burrell, 2016). The absence of interpretability generates concerns regarding the possibility of attaining complete transparency in AI systems, particularly within critical sectors such as healthcare and law enforcement.

Accountability is equally vital in AI governance. The principle dictates that developers and operators of AI systems must be responsible for the actions and decisions generated by these systems. According to Jobin, Ienca, and Vayena (2019), accountability ensures that there are mechanisms in place to identify who is responsible when AI systems cause harm or produce undesirable outcomes. The **reflexivity** aspect of responsible innovation connects to this, requiring AI developers and operators to remain critically aware of their ethical responsibilities and the broader societal impacts of their technologies (Owen et al., 2013). However, as AI systems become more autonomous, questions arise about how to assign responsibility when decisions are made without direct human input. Scholars like Danks and London (2017) highlight this dilemma, emphasising the need for clearer regulatory frameworks that address the distribution of responsibility between human operators and AI systems.

Fairness is perhaps the most pressing ethical concern in AI regulation. AI systems, if left unchecked, can perpetuate and even exacerbate societal biases, leading to unfair treatment of certain groups. Cases of algorithmic bias have surfaced in areas such as hiring, criminal justice, and access to financial services (O'Neil, 2016). For example, facial recognition systems have been found to have higher error rates when identifying individuals from certain ethnic backgrounds (Buolamwini & Gebru, 2018). Fairness in AI, therefore, extends beyond simply ensuring equal treatment; it requires ongoing scrutiny of the data, algorithms, and outcomes to prevent discriminatory practices. This resonates with the **anticipation** aspect of responsible innovation, where potential biases and harms are proactively identified and mitigated (Stilgoe et al., 2013). Danks and London (2017) argue that fairness also involves creating mechanisms for those adversely affected by AI decisions to seek redress, thus linking fairness with accountability.

The **EU AI Act** represents one of the most comprehensive efforts to embed these ethical principles into law. The Act emphasises the importance of transparency, accountability, and human oversight, particularly for high-risk AI systems (European Commission, 2021). It establishes requirements for documenting the data used in AI training, mandates clear explanations of how AI systems make decisions, and places

strict requirements on systems that significantly affect individuals' rights. This is reflective of the EU's broader approach to AI governance, which prioritises human rights and ethical considerations over purely economic or technological concerns (Floridi et al., 2018). Yet, while the EU AI Act sets ambitious standards, its implementation poses challenges, especially in ensuring compliance across diverse industries and technologies.

Despite the clear ethical frameworks that exist, significant challenges remain in addressing **algorithmic bias**, **privacy breaches**, and **maintaining human oversight**. Algorithmic bias continues to produce discriminatory outcomes, as demonstrated by hiring algorithms that favor certain demographic groups over others, and predictive policing systems that disproportionately target minority communities (O'Neil, 2016). The challenge of eliminating bias is compounded by the fact that AI systems often rely on historical data, which reflects existing societal biases. While bias detection and mitigation techniques are being developed, Veale and Binns (2017) argue that these techniques are not foolproof and require continuous monitoring and adjustment.

Privacy breaches are another significant concern in AI regulation. AI systems often process large amounts of personal data, raising the risk of privacy violations if adequate safeguards are not in place. Privacy-preserving technologies such as differential privacy and federated learning offer promising solutions to protect individuals' data while still allowing AI systems to learn from it (Crawford & Calo, 2016). However, these technologies are still in the early stages of adoption and may not be sufficient to address all privacy concerns, particularly as AI systems become more pervasive and integrated into everyday life.

To address these challenges, ongoing developments in **bias detection**, **privacy-preserving technologies**, and **frameworks for human oversight** are critical. Whittaker et al. (2018) emphasise the importance of case studies and empirical research in illustrating the effectiveness of various regulatory responses. For example, bias audits in AI systems have proven effective in identifying and correcting discriminatory outcomes, while human-in-the-loop systems have enhanced accountability by ensuring that critical decisions made by AI systems are overseen by human operators (Whittaker et al., 2018).

In conclusion, navigating the ethical challenges in AI regulation requires a multifaceted approach that integrates transparency, accountability, and fairness with the broader principles of responsible innovation. While there is growing recognition of the importance of these ethical principles, significant work remains to be done to ensure that AI systems are both effective and aligned with societal values. As AI continues to evolve, regulatory frameworks must be agile and adaptive, capable of addressing emerging ethical dilemmas while promoting innovation.

Chapter 3: Methodology

5. Methodology

5.1 Research Approach and Design

This study adopts a qualitative, comparative analysis approach using deductive thematic analysis to examine the AI/AGI regulatory frameworks of the European Union (EU) and China. This approach is chosen for its strength in highlighting the similarities and differences between the two governance models and how these models address the dual challenges of fostering innovation while managing risks. Comparative analysis is particularly effective when examining complex policy environments across different cultural, political, and legal contexts, allowing for a rich understanding of the respective regulatory practices (Miles, Huberman, & Saldana, 2014). This research design adheres to the principles of "multiplism" in policy analysis (Cook, 1985). “Multiplism” emphasises incorporating diverse perspectives and acknowledges potential biases within analytical techniques.

The decision to use deductive thematic analysis is grounded in the work of Braun and Clarke (2006), who argue that thematic analysis is a flexible method suited to identifying patterns within data. In this study, a predefined set of codes and themes, adapted from the 2020 Berkman Klein Center meta-review, which have been built upon and augmented through our analysis, provides a structured framework to compare the regulatory approaches. These themes include privacy, accountability, transparency, fairness, human-centric principles, and state support for innovation, which are essential for understanding how each region navigates the complexities of AI/AGI governance. By applying these predefined themes, this study systematically identifies and compares how both regions address key governance challenges related to AI/AGI, particularly concerning risk management and fostering innovation. The deductive approach allows for a targeted exploration of these themes based on existing literature, ensuring that the analysis is aligned with the study’s objectives (Braun & Clarke, 2006).

5.2 Data Collection

The data for this study consists of a comprehensive collection of official regulatory documents, policy briefs, and legislative texts from both the EU and China (Table 1).

These documents are publicly available and were selected from the past decade to ensure relevance. The data focuses on key thematic areas such as ethical considerations, technological advancement, risk management, and enforcement mechanisms.

Table 1: List of Documents Included in Thematic Analysis

Country/Region	**Document Title**	**Year Released**
EU	EU Artificial Intelligence Act	June 2024
	Digital Services Act	Oct 2022
	Digital Markets Act	Sept 2022
	Ethics guidelines for trustworthy AI (European Commission)	April 2019
	General Data Protection Regulation	April 2016
China	Model Artificial Intelligence Law (MAIL) v.2.0	April 2024
	Artificial Intelligence Law of the People's Republic of China (Draft for Suggestions from Scholars)	May 2024
	Basic Safety Requirements for Generative Artificial Intelligence Services	April 2024
	Guidelines for the Construction of a Comprehensive Standardization System for the National Artificial Intelligence Industry (Draft for Feedback)	Jan 2024
	Provisions on the Administration of Deep Synthesis Internet Information Services	Nov 2022
	Opinions on Strengthening the Management of Science and Technology Ethics	March 2022
	Provisions on the Management of Algorithmic Recommendations in Internet Information Services	Dec 2021
	Ethical Norms for New Generation Artificial Intelligence Released.	Sept 2021
	Artificial Intelligence Standardization White Paper (2021 Edition)	July 2021
	The PRC Personal Information Protection Law (PIPL)	Nov 2021
	Governance Principles for a New Generation of Artificial Intelligence: Develop Responsible Artificial Intelligence	June 2019

	Beijing AI Principles	2019

These documents were systematically selected to cover a wide spectrum of regulatory focus areas, ensuring a robust comparison of AI/AGI governance frameworks.

5.3. Data Analysis

This study employs a **deductive thematic analysis** to compare the AI/AGI regulatory frameworks in the EU and China. The thematic analysis approach is guided by pre-established themes derived from existing literature on AI governance, such as transparency, accountability, privacy, security, and innovation (Fjeld et al., 2020; Floridi et al., 2018). The analysis will focus on how these themes are addressed within the regulatory frameworks of both regions.

The key steps in the deductive thematic analysis process are as follows (Figure 1):

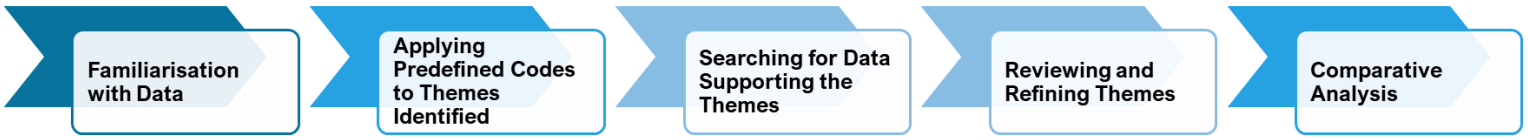

***Figure 12**: Data Analysis Process based on Braun & Clark's (2006) Guidelines. SOURCE: Author*

1) Familiarisation with Data:

This step involves thoroughly reviewing all collected documents and literature on AI/AGI governance in the EU and China. The goal is to gain an understanding of how each document aligns with the predefined themes derived from previous studies (e.g., ethical standards, risk management).

2) Applying Predefined Codes:

In deductive thematic analysis, the next step is to apply a priori codes based (refer to Table 2 for the exhaustive list) on themes identified in the literature (Braun & Clarke, 2006). The codes (e.g., "transparency," "innovation," and "accountability") will be systematically applied to the data. This process ensures that the analysis remains focused on relevant regulatory areas and allows for direct comparison between the EU and China's approaches.

3) Searching for Data Supporting the Themes:

With the predefined codes applied, the data will be reviewed to identify specific examples and discussions that support or contradict each theme. For instance, within the theme of "accountability," relevant regulations and policies from both regions will be compared to evaluate how accountability is enforced in each jurisdiction.

4) Reviewing and Refining Themes:

In this step, the initial coded data will be reviewed to ensure that the themes are well-supported by the data and are relevant to the research questions. Any discrepancies or patterns that do not fit the predefined themes will be noted, but the focus remains on confirming or refining the predefined themes.

5) Comparative Analysis:

The final step involves synthesising the findings and comparing how the EU and China handle each theme. This comparative analysis will reveal both the similarities and differences in how these regions approach the regulation of AI/AGI, particularly regarding ethical governance, privacy protection, and risk management.

5.4 Limitations

Despite its strengths, the deductive thematic analysis approach presents several limitations that must be acknowledged.

First, **reliance on publicly available data introduces a potential limitation**. Regulatory documents, policy briefs, and scholarly literature form the primary data sources for this study, which means that internal, unpublished policy developments—especially in a rapidly evolving field like AI—may not be captured. The study is therefore constrained by the availability of data that reflects the most current state of AI governance. While the chosen documents are comprehensive, they may not fully represent emerging trends or policy shifts that are happening behind closed doors, particularly in fast-moving technological environments.

Second, **researcher bias in the application of predefined codes is an inherent risk** in deductive thematic analysis. Although the use of a priori codes ensures consistency, the manual coding process is subject to the researcher’s interpretation of how the data aligns with these codes. The potential for subjective bias in coding and theme identification can influence the findings. To mitigate this, cross-validation of codes will be employed to ensure that the application of themes remains consistent, and regular peer reviews will help reduce bias in the coding process. Nevertheless, complete elimination of subjectivity in manual coding is unlikely, which remains a limitation of the method.

Third, the **rapid evolution of AI governance** is a significant limitation. The regulatory landscape for AI/AGI is constantly changing, with new policies and frameworks being introduced at both national and international levels. The findings of this study, therefore, represent a snapshot of current governance structures, which may be superseded by future developments. While the research offers valuable insights into the present state of AI governance in the EU and China, it may not fully capture future shifts that could alter the regulatory environment significantly.

Finally, the **cultural and political contexts of the EU and China** pose an additional challenge. The deductive thematic approach, while structured, may not fully account for the nuanced, region-specific political and cultural factors that influence AI/AGI governance. While the predefined themes provide a useful framework for comparison, they may overlook subtle dynamics within each region's political and regulatory environment that affect how policies are shaped and implemented. For instance, China's centralised governance model and focus on economic development may interact with AI regulation in ways that are not fully captured by a focus on themes like privacy or accountability. Similarly, the EU's emphasis on human rights and data protection may involve more complex considerations than what is captured by the predefined themes.

Chapter 4: Findings and Analysis

6. Developed Themes and Codes

This section outlines the key themes and codes developed through the deductive thematic analysis, based on the 2020 Berkman Klein Center for Internet & Society meta-review of existing sets of principles (such as the Asilomar Principles and Beijing Principles). This analysis represents the core aspects of AI governance in both the EU and China. These themes encompass critical regulatory areas, ranging from risk management and safety to fostering innovation and ensuring accountability.

To better understand how these themes interrelate, Figure 2 presents a thematic map, visually depicting the relationships between the identified themes. The figure highlights how certain themes, such as Safety and Security, are closely linked with broader concerns like Accountability and Professional Responsibility, while themes related to Promoting Innovation are supported by Clear Governance Mechanisms.

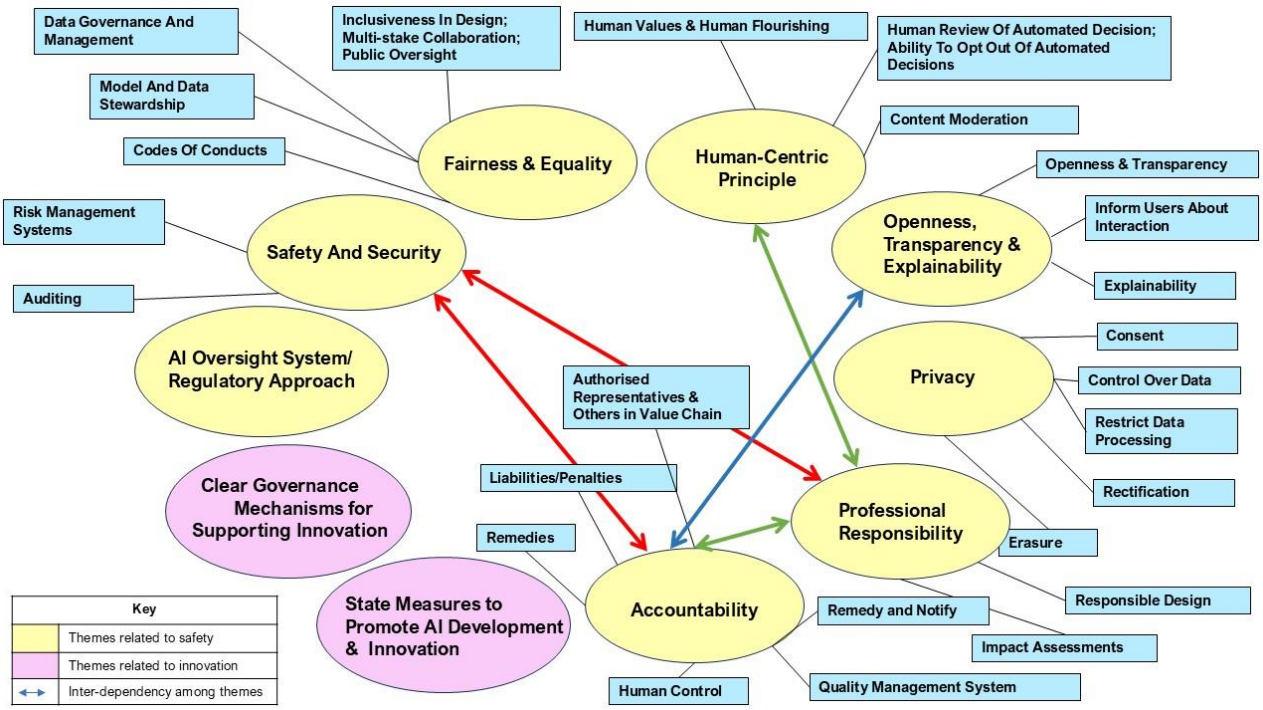

Figure 23: *Map of Thematic Codes Created and their Inter-Relationships. SOURCE: Author*

This mapping helps to clarify the complex web of dependencies between governance principles, demonstrating how themes associated with Risk Management (e.g., Safety,

Fairness, Accountability) overlap with those promoting Innovation (e.g., State Support for AI Development, Governance Mechanisms). The arrows in the figure indicate these interdependencies, illustrating how achieving goals in one area (e.g., Safety) may directly impact others (e.g., Professional Responsibility).

Table 2 below provides a detailed breakdown of the themes and associated codes that were generated through this analysis, with each code representing specific regulatory measures or governance principles related to either risk management or innovation support. The codes have been adjusted to fit the comparative analysis of the EU and Chinese regulatory frameworks for AI governance (Fjeld et al., 2020).

Table 2: List of Codes and Themes Created

Themes	**Codes**
Themes Related to Risk Management	
1. AI Oversight System/ Regulatory approach	Categorised AI oversight system
2. Safety and Security	Risk management systems / Internal management systems (record keeping of automatically generated logs; technical documentation; education and training of employees; checking accuracy, robustness, and cybersecurity)
	Auditing (post-market surveillance, monitoring, and information sharing)
3. Fairness and Equality	Codes of conduct for fairness, equality, and non-discrimination
	Model and data stewardship
	Appropriate data governance and management practices
	Inclusiveness in design; multi-stake collaboration; Public oversight

4. Human-Centric Principle	Human values and human flourishing
	Human supervision of automated decisions; Ability to opt out of them
	Content moderation
5. Openness, Transparency and Explainability	Openness and transparency
	Information that users are interacting with AI/synthesised content and notification that the content is AI-generated
	Explainability
6. Privacy	Consent
	Control over the use of data
	Ability to restrict data processing
	Right to rectification
	Right to erasure/revocation
7. Professional Responsibility	Responsible design
	Impact assessments
8. Accountability	Quality management system
	Human control/oversight of AI
	Remedy and notify; Corrective actions and duty of information

	Designate authorised representatives for AI developers and providers located outside the state
	Other entities in the AI value chain (besides developers, providers, and authorised representatives)
	Liabilities/Penalties
	Remedies
Themes Related to Supporting Innovation	
9. State Measures to Promote AI Development and Innovations	Government plans for AI development
	Construction of computing infrastructure
	Support for innovation in algorithms, open-source AI, and foundation models
	Support for the construction of foundational and specialised databases
	Develop national integrated big data centre systems
	Promote AI industrial development and the integration/application of AI in various industries
	Support professional talent cultivation institutions and mechanisms
	Provide fiscal and procurement support; allocate special budgets for AI support and development
	Tax credit incentives

	Pilot projects for AI applications in government and public management
	Establish regulatory sandboxes
10. Clear Governance Mechanisms for Supporting Innovation	Organograms
	Advisory forums
	Panel of experts
	Creation of monitoring body, notifying authorities, and conformity assessment bodies
	Database
	Detailed procedures/guidelines for implementation of regulations
	Liability fixation for government bodies
	International cooperation

7. Examination of Themes

Building on the themes outlined in Table 2, this section explores the major areas of AI governance across the EU and China. Key themes like AI oversight, safety and security, fairness, human-centric principles, accountability, and support for innovation will be examined. By comparing these approaches, the analysis will reveal how each region addresses the challenges and opportunities presented by AI, both ethically and practically. A detailed analysis table can be found in Appendix 1, which focuses on the core regulatory documents for both regions.

7.1 Themes Related to Ethics and Risk Management

7.1.1 AI Oversight System/Regulatory Approach

A categorised AI oversight system is necessary to prohibit manipulative, exploitative, or unfair AI practices, while also providing variable levels of regulation based on the perceived risks of non-prohibited AI systems. The risk categories can evolve as AI systems develop, making it crucial to adapt oversight mechanisms over time.

The **EU AI Act (2024)** classifies AI systems based on their potential risk:

- **Prohibited AI Practices**: Systems that exploit vulnerabilities, manipulate behaviour, or cause physical/psychological harm are explicitly banned. This includes biometric categorisation, subliminal techniques, and emotion recognition in sensitive areas like law enforcement, education, and border management.
- **High-Risk AI Systems**: These systems are subject to comprehensive obligations such as registration and post-market monitoring, especially for AI providers. The act mandates “human oversight” and “detailed risk management systems” for these systems.
- **General-Purpose AI Models with Systemic Risk**: These models, even when not tied to specific applications, are regulated due to their broad influence and potential harm.
- **Low-Risk AI Systems**: These systems only have minimal transparency requirements, such as informing users when they interact with AI.

The EU Act strictly prohibits government-run social scoring, which is notably used in China, aiming to protect fundamental rights. This risk-based regulatory framework provides mandatory requirements that adjust according to the category of AI.

In contrast, **China’s Model AI Law (MAIL 2024)** employs a **Licensing Oversight System** for AI systems on the Negative List—which includes systems deemed to pose risks to national security, public interest, or social stability—and a Registry Oversight System for systems outside this list.

- **Article 25** of China's MAIL establishes a **Categorised Oversight System**, where AI products and services on the Negative List require licenses, while those not on the list must register. The **National AI Administrative Authority** updates the **Negative List**, which considers risks to national security, social stability, environmental protection, and other factors.
- **Article 26** prohibits any unlicensed activities related to AI systems on the Negative List, ensuring strict government control over high-risk AI applications.

7.1.2 Safety and Security

Safety refers to ensuring the internal functionality of AI systems without causing unintended harm, while security addresses external threats like cyberattacks or unauthorised access. These two concepts are intertwined, forming the foundation of reliable AI systems.

The theme of **Safety and Security** connects with other key governance themes, such as **Accountability, Professional Responsibility, and Human Control of Technology**. These related themes provide implementation mechanisms for safety and security goals, ensuring AI systems remain trustworthy and resilient throughout their lifecycle.

AI developers and operators are expected to integrate safety and security measures throughout the lifecycle of AI systems. This includes:

- **Record keeping** of automatically generated logs,
- **Technical documentation,**
- **Training and education of employees,**
- **Ensuring accuracy, robustness, and cybersecurity**.
    - **Accuracy**: AI's confidence in correctly classifying data, making predictions, or delivering decisions.
    - **Robustness**: Redundancy measures like backups or fail-safes that ensure continuous operation.
    - **Cybersecurity**: The AI system's resilience against external threats, ensuring data privacy and system integrity.

Developers must also build systems capable of being audited, ensuring continual improvement based on evaluations and feedback. This includes **post-market surveillance, monitoring, and information sharing** on vulnerabilities or attacks, which can be handled by third-party human auditors or even other AI systems.

Both EU and Chinese regulations prioritise safety and security:

- **EU AI Act (2024)** mandates the implementation of a "**Risk Management System"** (Chapter III, Article 9) for high-risk systems. This system must be continuously updated throughout the AI system's lifecycle.
- **Articles 11, 12, 16, and 19** cover **technical documentation**, record keeping, and automatic log generation, all of which are crucial for auditing and oversight.
- **Post-market monitoring** is a requirement in **Chapter IX**, where providers must continue to monitor high-risk AI systems even after deployment. This includes reporting serious incidents and sharing information on vulnerabilities.

- **Article 4** further highlights the importance of **AI literacy**, mandating providers and deployers to ensure their staff are educated on the operations and risks associated with AI.

Similarly, **China's Model AI Law (MAIL 2024)** places significant emphasis on safety and security measures:

- **Article 5** establishes the **Safety and Security Principle**, requiring all AI actors to adopt necessary measures to ensure the safety and security of the systems and related network data.
- **Article 34** mandates **safety/security assessments** before deploying AI systems.
- **Article 46** requires the development of **robust security risk** management systems for foundation models.
- **Article 49** imposes obligations on AI providers to implement **internal management systems**, ensuring compliance with data security, risk control, and quality management standards. Regular **audits and employee education** are mandated to maintain system security and resilience.

Moreover, China's "**Basic Safety Requirements for Generative AI Services" (2024)** lists more than 30 specific safety risks, including algorithmic bias, privacy breaches, and copyright infringement, with additional guidelines specific to the Chinese political context.

7.1.3 Human-Centric Principle/Promotion of Human Values

The **Human-Centric Principle** revolves around the idea that AI systems should be designed and used in ways that align with societal norms, prioritising human values. As the power and prevalence of AI increase, particularly with the emergence of AGI, embedding human priorities, ethical judgment, and decision-making capabilities within AI becomes essential. This ensures that AI systems contribute positively to human welfare, respecting individual rights, cultural values, and societal norms.

This theme includes considerations such as the **review of automated decisions by humans** and the **ability for individuals to opt out of automated decision-making processes**, promoting human oversight in critical decisions that affect people's lives. AI development should focus on **human flourishing** and be conducted with a deep respect for **fundamental rights**, **labour rights**, **privacy**, and the **best interests of humanity**.

The **EU AI Act (2024)** places significant emphasis on the protection of **fundamental rights**.

- **Chapter III (High-Risk AI Systems), Article 27**, mandates a "**fundamental rights impact assessment** for high-risk AI systems". This requires public bodies and private entities providing public services to assess the impact on fundamental rights before deploying AI systems.
- **Chapter X (Codes of Conduct and Guidelines), Article 95-2 (e)**, also promotes the protection of vulnerable groups, ensuring that AI systems do not negatively impact persons with disabilities, or discriminate based on gender.

The **EU framework** prioritises personal and **fundamental rights**, placing a strong emphasis on protecting individuals from the potentially harmful impacts of AI systems. In contrast, the **Chinese regulatory framework** focuses on national security and social stability, reflecting distinct differences in how human-centric principles are interpreted and enforced.

The **Model Artificial Intelligence Law (MAIL 2024)** in China takes a different approach to the human-centric principle, emphasising **socialist values** and national priorities.

- **Chapter 1 (General Provisions), Article 4**, enforces the principle that AI development must always be oriented towards **human benevolence**, ensuring that humanity can supervise and control AI for the promotion of **human welfare**.
- **Article 14** outlines restrictions on AI systems that undermine **national security**, **national unity**, or **public morality**. It prohibits the generation of content that subverts the state, promotes terrorism, or harms the **national image**. This reflects China's focus on maintaining **social stability** and **national interests** through stringent content control.

Additionally, Chinese law incorporates specific mandates related to **intellectual property rights (IPR)**, **business ethics**, and the **protection of consumers and workers**. These provisions are designed to ensure that AI systems respect national values and protect individuals' physical and mental well-being. Similar provisions are outlined in the **Deep Synthesis Services Law (2022)** and **Algorithmic Recommendations Law (2021)**, which impose restrictions on the dissemination of false or harmful information, enforce content moderation, and ensure **public oversight** of AI.

Content moderation is a crucial aspect of China's human-centric approach to AI governance. The **Deep Synthesis Services Law (2022), Chapter II**, mandates that providers of deep synthesis services review user inputs and outputs to prevent the dissemination of harmful or illegal content. In cases where illegal content is detected,

providers are required to report it to the relevant authorities and take actions against the user, including account suspension or closure.

Similarly, the **Algorithmic Recommendations Law (2021), Article 13**, outlines stringent rules for internet news services that utilise AI-driven recommendations. Providers are prohibited from generating or transmitting fake news and are required to ensure that their recommendations align with state-approved sources of information. This regulatory framework reflects China's commitment to **content control** as a means of safeguarding **social order** and **national interests**.

7.1.4 Openness, Transparency and Explainability

A major governance challenge posed by AI is the complexity and opacity of its technology. It is often difficult to fully understand how AI systems work, the data they process, and the decisions they make. These difficulties emphasize the importance of transparency and explainability in AI systems, ensuring they function under ethical oversight and can be monitored and comprehended by those impacted by their operations.

7.1.4 (a) Transparency and Openness

Transparency in AI systems is crucial for fostering trust and ensuring that systems operate in ways that are accountable and open to scrutiny. **Transparency** refers to providing sufficient information about how AI systems function, including access to **data**, **business models**, **algorithms**, and, when applicable, **source code**. This enables external oversight, ensuring that AI systems are not deployed without the knowledge or consent of those affected by them.

The principle of **openness** also requires that users and employees interacting with AI systems are fully informed when they are engaging with AI-generated content.

The **EU AI Act 2024** lays out clear obligations to ensure transparency in AI systems.

- **Chapter IV, Article 50(1)** mandates that AI systems designed to interact directly with individuals must inform the user that they are interacting with an AI system unless it is already obvious. This transparency is crucial for maintaining trust and ensuring informed consent in interactions with AI.

- **Chapter III, Article 26 (6) and (11)** further requires deployers of high-risk AI systems to inform employees and individuals affected by these systems of their deployment and use. This ensures that individuals are aware of the involvement of AI in decisions affecting them, providing them with the opportunity to understand and challenge these decisions if necessary.

The **Model Artificial Intelligence Law (MAIL 2024)** in China also emphasises the importance of transparency and openness in AI systems.

- **Chapter 1, Article 6** outlines the **Principle of Openness, Transparency, and Explainability**, requiring that all AI systems clearly disclose their nature, purposes, and effects. This ensures that both users and the general public understand the role AI systems play in decision-making and content creation.

- **Chapter IV, Article 38** requires developers to disclose to users when they are interacting with AI systems or when content is AI-generated, fostering an environment of transparency in the use of AI.

- **Article 46** extends these principles to the developers of foundational AI models, requiring them to follow openness and transparency in their interactions, preventing **monopolistic practices**.

In the **Deep Synthesis Services Law (2022)** and the **Algorithmic Recommendations Act (2021)**, similar obligations for transparency are laid out, particularly in contexts where AI-generated content might mislead or confuse the public. Providers must prominently label such content, ensuring that the public is aware that they are interacting with AI-generated information.

7.1.4 (b) Explainability

Explainability goes a step further by requiring AI systems not only to be transparent but also to be **understandable** to users and regulators. Explainability refers to the ability to translate **technical concepts** and **decision outputs** of AI systems into formats that are comprehensible and accessible. This is crucial for ensuring accountability, as stakeholders must be able to evaluate and challenge AI systems' decisions when necessary.

Annex IV of the **EU AI Act** provides extensive detail on what must be documented and explained for high-risk AI systems, including the **reasoning behind decisions**, the **mechanisms of the AI system**, and how these decisions were reached. This ensures that AI systems operate with **full accountability** and that their decision-making processes are transparent to users, regulators, and external stakeholders.

Chapter IV, Article 39 of the **Chinese MAIL 2024** mandates that users have the right to request **explanations** from AI providers regarding the decision-making processes and methods used by AI systems. Users also have the right to **lodge complaints** if they find these explanations unsatisfactory. This enshrines the right of individuals to

challenge decisions made by AI systems, a crucial element in ensuring that AI systems are fair and transparent in their operations.

Similarly, **Chapter II of the Algorithmic Recommendations Act 2021**, **Article 12**, encourages providers to optimise the **transparency and explainability** of their AI systems, particularly in how they handle **searches**, **sorting**, and **recommendations**. This provision aims to prevent conflicts and disputes by ensuring that AI-driven outcomes are understandable to those affected by them.

7.1.5 Privacy

Privacy is one of the most pressing concerns in the context of AI, where massive data analytics allow for the collection, storage, and processing of personal information at unprecedented scales. AI systems are used in surveillance, healthcare, advertising, and numerous other areas, making privacy a central issue both in the **visible use** of AI and **behind the scenes**, during the training and development of AI models.

According to **Fjeld et al. (2020)**, privacy concerns in AI governance should cover several key areas:

- **Consent**: Users must be informed about how their data will be used and give explicit permission, either through **notice-and-consent** or **informed consent** mechanisms.
- **Control over data**: Individuals should have some control over how their data is used, whether through **personal tools** or through dedicated systems and institutions.
- **Ability to restrict data processing**: Individuals should have the right to prevent their data from being used in AI systems, either through enforceable legal rights or institutional mechanisms.
- **Right to rectification**: Data subjects should be allowed to correct or update inaccurate or incomplete information.
- **Right to erasure/revocation**: Individuals should have the right to have their personal data removed from AI systems entirely.

The **EU AI Act 2024** places a strong emphasis on privacy throughout the entire lifecycle of AI systems. It recognises privacy as a fundamental right that must be safeguarded.

- **Point 69** of the Act emphasises that "the right to privacy and the protection of personal data must be guaranteed throughout the entire lifecycle of the AI system."

- **Article 2(7)** specifies that **Union law on data protection, privacy, and confidentiality** applies to all personal data processed in connection with the AI Act. This guarantees compliance with existing **GDPR** regulations and ensures that AI systems uphold privacy standards.

- **Article 10(5)(b)** outlines specific privacy-preserving measures for **high-risk AI systems**, including the use of **pseudonymisation** and **technical limitations** on the re-use of personal data.

Furthermore, **Directive 2002/58/EC**, known as the **ePrivacy Directive**, provides additional protections for the confidentiality of communications and the private life of individuals, further reinforcing the EU's stance on privacy.

In **China's Model Artificial Intelligence Law (MAIL 2024)**, privacy is equally significant, though the framework is more aligned with the nation's overarching goals of **social stability and state control**.

- **Chapter 1, Article 14(c)** stipulates that developers must legally protect the **rights and interests** of consumers and workers, ensuring their **privacy**, **honour**, and **personal information** are safeguarded.

The **Deep Synthesis Services Law (2022)**, **Article 14**, specifically mandates consent for the use of biometric information, such as faces and voices. Developers must notify individuals whose personal information is being edited and obtain their **independent consent**.

Chapter IV, Article 29 of the **Algorithmic Recommendations Law (2021)** emphasises maintaining **confidentiality of personal and private information**, further underlining the Chinese government's approach to privacy.

However, **Chapter II, Article 9** of the **Deep Synthesis Services Law** requires **real identity verification** for users of deep synthesis services, indicating a greater degree of control and surveillance by the state, compared to the more individual rights-based approach of the EU.

7.1.6 Professional Responsibility

Professional responsibility is an important ethical consideration, focusing on the duties of those involved in the development, deployment, and regulation of AI systems. This principle emphasises that **professionals** must act conscientiously, adhere to ethical standards, and be **accountable** for the broader social and human rights impacts of AI technologies.

Professional responsibility closely aligns with themes of **accountability** and **human-centric design**, ensuring that AI development is **transparent**, **ethical**, and conducted with foresight into potential consequences.

The **EU AI Act 2024** incorporates the concept of **professional responsibility** in several ways:

- **Chapter X, Article 95** outlines the development of **codes of conduct** for the voluntary application of AI systems, emphasising **ethical design** and **sustainability**. The Act encourages developers to minimise the environmental impact of AI systems through **energy-efficient programming** and **sustainable design**.

- **Chapter III, Article 27** mandates **fundamental rights impact assessments** for high-risk AI systems, ensuring that professional conduct is in line with human rights standards and ethical considerations.

The **Chinese Model AI Law (MAIL 2024)** also emphasises **professional responsibility**, but with a particular focus on **sustainability** and **state interests**.

- **Chapter 1, Article 9** promotes the use of **green principles**, encouraging AI developers to adopt **energy-saving** and **emission-reduction technologies**. This aligns with the Chinese government's broader goals of fostering a **digital ecological civilisation**.

The **Deep Synthesis Services Law (2022), Article 5**, emphasises the need for industry organisations to establish **self-discipline** and **management standards**. Developers and service providers are encouraged to **accept societal oversight** and improve operational standards according to national guidelines.

Article 5 of the Algorithmic Recommendations Act (2021) similarly stresses the importance of **self-discipline** and **ethical responsibility** for those involved in AI systems.

Both frameworks emphasise professional responsibility, but the **EU's approach** is more **human rights-focused** and **ethics-driven**, while **China’s regulations** emphasise **alignment with state goals**, including environmental sustainability and **social discipline**.

7.1.7 Accountability

Accountability in AI governance refers to the assignment of responsibility for the development, deployment, and operation of AI systems. The concept emphasises that

if adverse outcomes or malfunctions occur, there should be mechanisms to identify responsible entities and impose corrective actions or penalties. This theme is closely related to other core principles such as **Safety and Security**, **Transparency and Explainability**, and **Professional Responsibility**.

In the **EU AI Act 2024**, **accountability** is deeply embedded within the requirements for **high-risk AI systems**. The Act requires developers and providers to implement a comprehensive **quality management system** that ensures compliance with the regulatory framework.

- **Article 17** of the Act specifies that providers must establish a **quality management system**, which includes documentation of policies, procedures, and instructions to ensure systematic oversight of AI systems. Part of this system must involve an **accountability framework** that clearly outlines the responsibilities of management and staff, ensuring that everyone involved in the AI lifecycle adheres to the required standards.

- **Article 14** deals specifically with **human oversight** for high-risk systems, mandating that deployers of AI assign natural persons with the necessary competence to oversee AI operations. This ensures that AI systems can be **intervened by humans** when necessary, preventing undesirable outcomes during operation.

- **Article 26(2)** outlines the obligations of deployers to assign competent and trained personnel to oversee the operation of **high-risk AI systems**. The system must be auditable, with logs and data readily available to ensure compliance.

- In **Chapter V, Article 56**, **codes of practice** are outlined for **general-purpose AI models**, encouraging voluntary compliance beyond the formal requirements for high-risk systems.

The EU's approach emphasises **documentation, human oversight**, and **structured accountability frameworks** to ensure that both providers and deployers of AI systems maintain responsibility throughout the AI lifecycle.

In the **Chinese AI regulatory framework**, accountability is also a central theme but is framed within a system that strongly prioritises **state interests** and **national security**. The **Model Artificial Intelligence Law (MAIL 2024)** lays out clear expectations for AI developers, providers, and deployers regarding their responsibility to adhere to both **regulatory** and **state-driven priorities**.

- **Article 7** of the **General Provisions** articulates the **principle of accountability**, stating that all entities involved in AI activities (from research and development to deployment) must be responsible for their respective actions. This accountability framework operates within the context of broader societal and state-driven goals.

- **Article 45(a)** mandates that developers in the **negative list** must establish and operate a **quality management system** in accordance with legal requirements, ensuring that AI systems meet predefined safety and accountability standards.

- **Article 52(c)** adds that providers must establish a **full-lifecycle quality management system** that incorporates human control, ensuring that AI systems can be **intervened by humans** during autonomous operations.

7.1.7 (a) Corrective Actions and Liability

Both the EU and China have provisions for **corrective actions** and **remedies** in cases where AI systems malfunction or violate regulations.

- In the **EU AI Act 2024**, **Article 20** requires providers to take necessary **corrective actions** and inform regulatory authorities about issues arising from the deployment of high-risk AI systems. This includes documentation and **remedy mechanisms** to correct failures or breaches in compliance.

- **Article 16(j)** further specifies that providers must be proactive in taking corrective measures and notifying authorities when issues occur.

Similarly, the **Chinese MAIL 2024** emphasises the need for timely **remedial actions**.

- **Article 37** mandates that developers and providers must rectify any **security defects or vulnerabilities** that arise during the operation of AI systems. The law requires prompt reporting and remedy procedures to address any non-compliance or risk to users or the state.

7.1.7 (b) Authorized Representatives

Both frameworks recognise the need for **authorized representatives** for developers or providers located outside their respective jurisdictions.

- **Article 22(1)** of the **EU AI Act** mandates that providers based outside the EU must appoint an **authorized representative** within the Union to manage compliance and oversee AI operations in the EU.

- Similarly, **Article 44** of the **Chinese MAIL 2024** requires that AI developers and providers located outside China must designate **representatives** within the People's Republic of China to handle AI-related affairs and ensure compliance with Chinese regulations.

7.1.7 (c) Liabilities and Penalties

Both the **EU** and **China** have detailed provisions for **penalties** and **legal liabilities** in cases where AI systems breach regulatory requirements or cause harm.

- **Article 99** of the **EU AI Act** outlines penalties that member states can impose on operators, including **fines** and other enforcement measures. The penalties are designed to be **effective, proportionate, and dissuasive** to ensure compliance across the Union.

- **Article 101** provides for fines specifically for **providers of general-purpose AI models**, ensuring that they too are held accountable for violations.

In China, **Chapter VI (Liabilities)** of the **MAIL 2024** similarly provides for a range of penalties, from **warnings** and **fines** to the **revocation of licenses** for entities violating the law.

- **Article 66** outlines **general liabilities**, while **Article 67** focuses on the **revocation of licenses** for AI systems on the negative list.

- **Article 73** introduces the concept of **public interest litigation**, allowing public bodies to file lawsuits on behalf of affected individuals when AI systems infringe upon the rights of multiple people.

7.1.7 (d) Remedies

Both frameworks also provide for **remedies** that individuals can seek if they are harmed by AI systems.

- In the **EU AI Act 2024**, **Article 85** gives individuals the right to lodge complaints with **market surveillance authorities** if they believe an AI system has violated their rights or breached regulations. Additionally, **Article 86** guarantees individuals the **right to explanations** for decisions made by high-risk AI systems that affect them.

- The **Chinese MAIL 2024** provides for **civil tort liability** in **Article 70**, requiring developers and providers to compensate individuals for damages caused by AI systems, except in cases where they can prove they were not at fault.

7.2 Themes Related to Support of Innovation

7.2.1 State Measures to Promote AI Development and Innovations

States play a pivotal role in fostering AI development by devising and implementing strategies that promote innovation and the development of AI technologies. Regulatory sandboxes are one of the key measures to support innovation. These provide controlled environments where AI technologies can be tested and validated before being placed on the market. This helps mitigate risks while encouraging innovation.

- **EU AI Act 2024**: Chapter VI, Article 57, mandates the establishment of AI **regulatory sandboxes** at the national level by Member States. These sandboxes allow AI developers, including startups, to access a controlled environment for innovation, providing priority access to small and medium-sized enterprises (SMEs). Articles 58–63 further elaborate on the operation and governance of these sandboxes.
- **Chinese MAIL 2024**: Chapter V, Article 60, establishes a **regulatory experimental mechanism** for AI, with specific guidelines on participation, risk monitoring, and liability mechanisms. The experimental mechanism allows innovation to flourish in a controlled and monitored setting.

In addition to regulatory sandboxes, both the **EU** and **China** promote a range of state-driven initiatives, such as:

- Developing national **AI development plans**.
- **Constructing computing infrastructure**.
- Supporting **innovation in algorithms, open-source AI, and foundation models**.
- Building **specialised databases** and **national big data centres**.
- Encouraging **AI industrial development** across various sectors.
- **Fiscal and procurement support**, including tax credits and pilot projects for AI applications in public management.
- Establishing **AI special zones** to facilitate focused innovation.
- **Chinese MAIL 2024**: Chapter II outlines various state-led initiatives to support AI innovation. Notably, **Article 18** supports the **supply of data production factors**, a unique feature in Chinese regulations that promotes the aggregation and utilisation of public data for AI applications, expanding the scope of public data supply.

7.2.1 (a) Support for Open-Source AI Models

Both regulatory frameworks recognise the importance of **open-source AI models** for innovation and make provisions to exempt them from certain regulatory obligations, provided they are not classified as high-risk systems.

- **EU AI Act 2024**: Article 2(12) states that the regulation does not apply to AI systems released under **open-source licences**, unless they are placed on the market as **high-risk AI systems**. Specific articles like **Article 25(4)** and **Article 53(2)** waive obligations for open-source AI models.
- **Chinese MAIL 2024**: Chapter V, Article 59, promotes the development of **open-source AI** by formulating compliance guidelines, while **Article 71** provides **liability exemptions** for open-source AI developers.

7.2.2 Clear Governance Mechanisms for Supporting Innovation

Well-structured governance mechanisms are essential to supporting innovation by increasing the ease of doing business. These mechanisms include the creation of advisory forums, expert panels, regulatory authorities, and procedures for conformity assessments.

- **EU AI Act 2024**: Chapter VII provides a comprehensive governance structure, including:
    - **Article 64**: AI Office.
    - **Article 65**: European Artificial Intelligence Board.
    - **Article 66–70**: Advisory forum, scientific panel of experts, and national competent authorities.
    - **Chapter III**: Articles 28–38 outline the process for **notifying authorities**, **conformity assessments**, and **monitoring bodies**.
    - **Chapter VIII**: Article 71 establishes an **EU database** for high-risk AI systems to ensure transparency and traceability.
- **Chinese MAIL 2024**: Chapter I and V elaborate on the state's **governance principle** (Article 3) and outline the responsibilities of the **National AI Administrative Authority** (Article 54). Chinese regulations emphasise the role of the state in balancing **innovation and governance** while ensuring **national security**.

In addition, Chinese regulations provide for a **security review system** (Article 56) for AI technologies that may impact national security, and **Chapter 63** encourages the development of **RegTech** and **ComplianceTech** for AI monitoring and safety.

7.2.2 (a) Procedural Clarity and Conformity

Detailed procedural guidelines and conformity assessment processes are crucial for promoting innovation while maintaining regulatory oversight. This helps eliminate ambiguity and ensures AI systems comply with necessary standards before entering the market.

- **EU AI Act 2024**: Articles 40–49 focus on **harmonised standards**, **conformity assessments**, **certification procedures**, and post-market monitoring. Chapter IX outlines the process for **supervision, enforcement**, and **remedies**.

- **Chinese MAIL 2024**: Articles 47–48 outline the **registry obligations** and procedures for providers. The **Deep Synthesis Services Law 2022** also provides for **formal filings** and **information sharing** for AI service providers (Article 19).

7.2.2 (b) Liability and International Cooperation

Both regulatory frameworks recognise the importance of **liability fixation** for both developers and state bodies, ensuring that the government's actions are also subject to accountability.

1. **EU AI Act 2024**: Article 100 establishes **administrative fines** for Union institutions, bodies, and agencies.

2. **Chinese MAIL 2024**: Article 76 stipulates liability for **state organs** failing to fulfil their obligations, ensuring accountability at every level of AI governance.

Finally, **international cooperation** is critical for AI innovation, especially in a globalized economy where AI developments often transcend national borders.

- **Chinese MAIL 2024**: Article 11 actively encourages **international cooperation**, talent acquisition, and technological collaboration, promoting the formulation of **global AI governance standards**.

- Chapter V, Articles 64–65, provides a framework for **countermeasures and reciprocal actions** against foreign entities that impose discriminatory restrictions on China's AI R&D or trade.

Chapter 6: Discussion and Conclusion

8. Discussion

8.1 Comparative Analysis

The AI regulatory frameworks of the EU and China share similarities but also have fundamental differences in terms of structure, approach, and enforcement. Table 3 provides a comparative analysis of key themes from both regulatory frameworks, showing how each region addresses the governance of AI technologies.

Table 3: Comparison Analysis of EU and China AI Regulations

Theme	EU Regulations	Chinese Regulations
1. General Features		
Approach to AI regulatory framework	General as well as targeting specific distinct AI technologies	General technology neutral
System of AI regulation	Structure- highly centralised and hierarchical; Processes- volatile; Outcome-fragile	Structure- democratic; Processes- stable; Outcome-robust
2. Themes Related to Ethics and Risk Management		
AI oversight system	Categorise AI systems into categories based on perceived risk: Prohibited AI practices; High-Risk AI systems (subject to extensive obligations); General-purpose AI models with systemic risk (moderately regulated); and other AI systems considered low risk (subject to limited regulations) ***Government-run social scoring of the type used in China is banned.***	Licensing oversight" system for "Negative List" systems; and "Registry oversight" system for systems "outside the negative list" Allows all activities except those explicitly prohibited.
2.1 Safety and Security		

Risk Management systems	Yes	Yes
Auditing	Yes (No requirement to align AI system to political ideology (all countries are democratic).	Yes (AI systems need to align with China's political system in view of the tightly censored Chinese internet).
2.2 Fairness and Equality		
Code of conducts for Fairness, equality and Non-discrimination	Yes	Yes
Model and data stewardship	Yes	Yes
Appropriate data governance and management practices	Yes	Yes
Inclusiveness in Design/Multi-stake collaboration/Public oversight	Yes	Yes
2.3 Human-Centric Principle		
Human values and human flourishing	Yes (Lay great emphasis on Fundamental rights, and personal rights of people)	Yes (Lay emphasis on socialistic values, national security, national image etc. Fundamental rights conspicuously missing)
Human review of automated decision; ability to opt out of automated decisions	Yes	Yes
Content moderation	Not prominent	Prominent

2.4 Openness, Transparency and Explainability		
Openness and transparency	Yes	Yes
Inform users about AI interaction	Yes	Yes
Explainability	Yes	Yes
Privacy	Yes	Yes (Except the requirement to identify users of deep synthesis service)
2.5 Professional Responsibility		
Responsible Design	Yes	Yes
Impact Assessments	Yes, including fundamental rights impact assessments.	Yes (No mention of fundamental rights impact assessments)
2.6 Accountability		
Quality management system	Yes	Yes
Human control of AI	Yes	Yes
Remedy and notify	Yes	Yes
Regulate authorised representatives & others in the value chain	Yes	Yes
Liabilities/Penalties	Yes	Yes

Remedies	Yes	Yes
3. Themes Related to Support of Innovation		
State support/measures to promote AI development and innovations	Yes (Regulatory sandboxes; support open-source AI)	Yes (Regulatory sandboxes; support open source AI and foundation models; make and implement plans for development of AI; Construction of computing infrastructure; support innovation in algorithms; support construction of foundational and specialised databases, develop national integrated big data centre system; promote AI industrial development, and integration/application of AI in various industries; support professional talent cultivation institutions and mechanisms; provide fiscal and procurement support, allocate special budgets foray support and development; provide tax credit incentives; agencies to start pilot projects for AI application in govt and public management; establish AI special zones etc.
Clear Governance Mechanisms for Supporting Innovation		
Organograms/advisory forums (panel of experts)/creation of monitoring body/notifying authorities, and conformity assessment bodies, database	Yes	Yes (Government supremacy and finality of decision in such matters is envisaged in Chinese Regulations)
Detailed procedural details	Yes	Yes
Liability fixation for government bodies	Yes	Yes

International cooperation	NA	Yes, including countermeasures against foreign countries/entities in case they impose restrictions against China.

8.1.1 Similarities in EU and Chinese AI regulatory frameworks and policies

The AI regulatory frameworks in both China and the EU converge on promoting responsible AI development and use. They share common principles such as categorised AI systems; safety and security measures; fairness and equality; human centricity; openness and transparency; professional responsibility; privacy; accountability; and specific provisions for promoting AI（Fjeld et al., 2020）.

8.1.2 Differences in EU and Chinese AI regulatory frameworks and policies

a) **Differing approaches to AI regulatory frameworks in China and the EU:** China's AI regulatory framework focuses on targeting specific technologies through tailored regulations (such as those governing algorithmic recommendations, deep synthesis technologies including deepfakes, and generative AI services). This approach encourages innovation while maintaining control over AI's development to mitigate associated risks, embedding ethical standards in governance that emphasise national security, public interest, and protection of individual rights（Zhang, 2024）. In contrast, the EU employs a technology-neutral and risk-based approach, systematically categorising AI systems by their associated risks, with stringent regulations for high-risk AI systems to ensure human safety, protection of fundamental rights, and adherence to ethical practices （Allnut & Hardy, 2024）.

b) **Government-run social scoring systems:** The EU Act 2024 categorises AI systems into four categories: Prohibited AI practices, High-Risk AI systems (subject to extensive obligations), General-purpose AI models with systemic risk (moderately regulated), and other low-risk AI systems (subject to limited regulations). Chinese regulations under MAIL 2024 utilise a "Licensing oversight" system for AI technologies on the "Negative List" and a "Registry oversight" for systems outside the Negative List（Zhang, 2024）. Additionally, government-run social scoring systems, such as those used in China, are banned in the EU（Sahin, 2020）. AI tools in China are accused of facilitating digital authoritarianism and accelerating surveillance practices in non-democratic settings, or fragile democracies（Harwell & Duo, 2020）. China has faced accusations of exporting these AI surveillance tools to expand political and economic influence while

opposing democratic governance models (Sahin, 2020).

c) **Government oversight and cooperation:** Both regulatory frameworks emphasise security assessments, audits, and risk evaluations, but Chinese regulations require a more pronounced level of government oversight and collaboration with authorities. While the EU focuses on protecting fundamental rights, China places greater emphasis on socialist values, national security, and maintaining the national image (Fjeld et al., 2020). EU authorities are also more proactive in enforcing regulations, as exemplified by investigations into OpenAI (Allnut & Hardy, 2024). However, China's regulatory regime, though supportive of industries, can result in significant risks due to a lack of strict enforcement (Zhang, 2024).

d) **Monitoring user behaviour and content moderation:** Chinese AI regulations impose direct responsibilities on providers to monitor user behaviour and moderate content, which is absent in EU regulations. For example, China's regulations mandate mechanisms to filter illegal or harmful content, unlike the EU AI Act, which doesn't require policing of user behaviour (National Technical Committee 260, 2024).

e) **Supply of data production factors:** China has a notable advantage in facial recognition technology due to its government-backed partnerships, but lacks strong Chinese language datasets for training AI systems. Companies like Baidu use English-language sources like Reddit and Wikipedia, which are often misaligned with Chinese government censorship requirements (Hale, 2023). To resolve this, the Chinese government coordinates data resource creation for AI training (Zhang, 2024).

f) **Registration for AI applications:** China mandates the registration of all AI systems with authorities, contrasting with the EU's registration of only high-risk AI systems in a public database (Fjeld et al., 2020). This demonstrates China's more authoritative stance on AI oversight, focusing on national security and public order.

g) **Different political systems, different regulatory systems:** The structural difference between Chinese and EU regulations arises from their political systems. China's regulatory system is centralised and hierarchical, characterised by volatility and fragility due to aggressive regulatory action with little resistance from businesses. This leads to cyclical policy changes and uncertainty (Zhang, 2024) (Al Jazeera, 2023). Conversely, the EU's democratic regulatory system provides stability in processes and outcomes, enabling a more predictable regulatory environment (Allnut & Hardy, 2024).

8.2 Proposed Global Governance Framework

While AI offers numerous benefits, it also poses significant risks, such as threatening national security (by democratising capabilities that could be exploited by malicious actors), facilitating unequal economic outcomes (by concentrating market power in the hands of a few companies and countries while displacing jobs in others), and creating undesirable societal conditions (through extractive data practices, reinforcing biased narratives, and environmentally harmful compute requirements) [Roberts et al., 2024]. These risks transcend national borders, emphasising the need for a strong global AI governance framework that accommodates diverse interests without a single sovereign authority. This would allow for cooperative action to maximise AI's benefits while mitigating its risks (Weiss, 2000).

Figure 3 illustrates the **proposed global AI governance framework.** A global AI governance framework must encompass three interrelated levels: international, national, and industry. At the industry level, state-of-the-art practices and codes of conduct for developers and organisations should be established. At the national level, consensus-based standards and specifications should align with global frameworks while addressing local requirements. At the international level, key outcomes should include agreement on globally significant safety and security risks, uniform global standards, interoperability of AI regulatory frameworks, scientific consensus on risk thresholds, and the inclusive development and sharing of AI's benefits（Smith & Crampton, 2024).

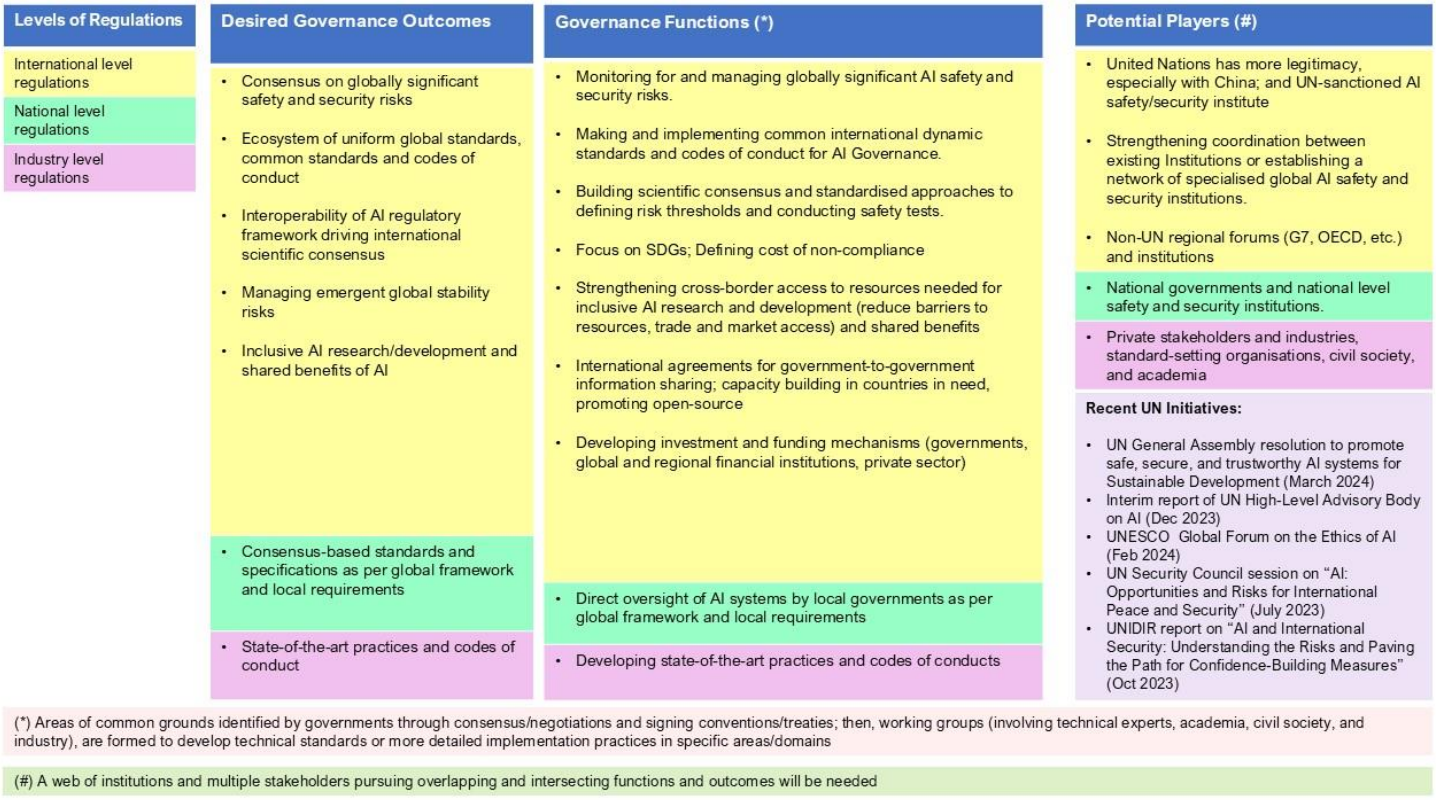

***Figure 3**: Proposed AI/AGI Global Governance Framework*

8.2 (a) Areas of Common Ground and Proposed Governance Functions:

Areas of common ground can be identified by governments through consensus-building, negotiations, and the signing of conventions or treaties. Once consensus is achieved, **working groups** consisting of technical experts, academia, civil society, and industry stakeholders could be established to develop technical standards and more detailed implementation practices in specific areas or domains.

To achieve the desired outcomes, it is essential to focus on specific **governance functions**:

- **Monitoring and managing globally significant AI safety and security risks**
- **Establishing and implementing international dynamic standards and codes of conduct for AI governance**
- **Building scientific consensus and standardised approaches for defining risk thresholds and conducting safety tests**

- **Focusing on the Sustainable Development Goals (SDGs) and defining the costs of non-compliance**
- **Strengthening cross-border access to resources** needed for inclusive AI research and development (by reducing barriers to resources, trade, and market access)
- **Facilitating international agreements for government-to-government information sharing**, capacity building in countries in need, and promoting open-source AI initiatives
- **Developing investment and funding mechanisms** supported by governments, global and regional financial institutions, and the private sector

A multilateral forum, such as the **United Nations (UN)**, seems to be an appropriate institution to facilitate the development, negotiation, and agreement of a **global AI policy framework**. The UN's delegated authority from member states gives it a high degree of procedural legitimacy, though it is still subject to some criticism. In contrast, **regional forums** like the **G7, G20**, and **OECD** may face challenges of legitimacy, given their predominantly Western membership. However, the **OECD** has managed to engage China, which participated in the **UK's AI Safety Summit** (November 2023) and the **World Economic Forum Summit** (January 2024). Nevertheless, China continues to emphasise the importance of leveraging the **UN** for global AI governance (The Cyberspace Administration of China, Global AI Governance Initiative, 2023).

Recent UN initiatives illustrate the organisation's commitment to AI governance, including:

- The **UN General Assembly's resolution** (March 2024) to promote safe, secure, and trustworthy AI systems for sustainable development
- The **Interim Report of the UN High-Level Advisory Body on AI** (December 2023)
- The **UNESCO Global Forum on the Ethics of AI** (February 2024)
- The **UN Security Council session** (July 2023) on "AI: Opportunities and Risks for International Peace and Security"

- The **UNIDIR report on AI and International Security** (October 2023), which explores risks and outlines confidence-building measures for international security

8.2 (b) Strengthening Coordination and the Need for a Decentralised Network

Given the **rapid pace of AI development**, the loosely defined scope of AI (with ongoing disagreements over "field boundaries and what constitutes harm") [Roberts et al, 2024], and its **decentralised nature**, developing a **new centralised AI institution** that covers all aspects of AI could prove problematic. Unlike nuclear material and technology, AI cannot be controlled from a single centralised point.

Therefore, an alternative approach would be to **establish a decentralised network of specialised institutions**, each targeting specific AI governance issues. This would allow for more focused and flexible regulation and governance tailored to particular aspects of AI.

8.3 (c) The Regime Complex Model for AI Governance:

A **non-hierarchical regime complex model** could involve a **network of international institutions and agreements** that work together to govern specific issue areas related to AI. This model would allow for cooperation in different forums, even when geopolitical or institutional conditions stall progress in others. The regime complex model also promotes **incremental progress** and **trust-building** from various state and non-state actors, producing mutually reinforcing changes over time (Keohane & Victor, 2011).

The regime complex model also offers adaptability in line with **technological advancements**, allowing for the inclusion of diverse governance stakeholders, including **big tech companies**. This is essential, given the **technical complexity and contextual nature** of AI.

Building a strong regime complex would involve:

- **Aligning targets** between different governance actors
- **Improving information-sharing** between institutions
- **Developing institutional partnerships**

- **Creating conflict resolution mechanisms**

There is precedent for this type of governance in other areas of international policy-making, notably in climate change governance (Galaz et al., 2012). Developing a **robust regime complex** for AI governance is not only possible but necessary, given the increasing global impact of AI technologies. Ultimately, this will require a **web of institutions** and **multiple stakeholders** working together toward overlapping and intersecting functions and outcomes.

8.2 (d) Examples of Current Efforts Towards Global AI Governance

Various efforts have been undertaken by states, international institutions, and private stakeholders to move toward global AI governance. For instance, the UN has engaged in discussions about governing lethal autonomous weapons systems (LAWS) under the Convention on Certain Conventional Weapons (UN, 2023). The OECD members adopted AI ethics principles in 2019, which were subsequently endorsed by G20 leaders. UNESCO's *Recommendation on the Ethics of Artificial Intelligence* (2021) aims to guide member states in developing legal frameworks for AI. The G7 launched the Hiroshima AI Process in 2023 to promote cooperation in AI governance, while BRICS countries agreed to form an AI study group. In December 2023, the Council of Europe (CoE) drafted a legally binding international convention on AI and human rights（Roberts et al., 2024）.

Efforts by states to establish international AI bodies include the *Global Partnership on AI* (GPAI), launched in 2020 by 15 founding countries to support the ethical adoption of AI; the *Trade and Technology Council*, established in 2021 to coordinate EU and US activities in AI; and the *UN High-Level Advisory Body on AI*, formed in 2023 to provide governance recommendations (Roberts et al., 2023). Additionally, the UK formed the *AI Safety Institute*（Samson, 2023）.

Private stakeholders have developed governance mechanisms, such as product and process standards for AI published by the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC), which are voluntary. The *Partnership on AI* (PAI) was established in 2016 by major tech companies, civil society organisations, and academic stakeholders to develop AI guidance and inform public policy. The *Frontier Model Forum*, founded by four big tech companies in 2023, aims to establish governance mechanisms for advanced AI systems（Roberts et al., 2024）. However, these efforts have mainly focused on regulatory consistency rather than governance itself.

8.2 (e) Barriers to Global AI Governance

The development of a consensus for a strong global AI governance framework faces challenges, primarily due to "first- and second-order cooperation problems"（Roberts et al., 2024）.

First-order cooperation problems stem from international anarchy, characterised by the absence of a common government in world politics. This creates uncertainty over the enforcement of agreements and other states' intentions（Lechner, 2022）. States' threat perceptions, levels of trust, and alignment of interests influence cooperative action in AI governance. For example, the Chinese policy of promoting military-civil fusion through AI has led the US, along with ideologically aligned countries, to oppose China. The US has enacted export controls on semiconductors to hinder China's AI development while promoting domestic production（Rajagopalan, 2024）. Conversely, China has promoted its competitiveness through policies like *Made in China 2025* and its Belt and Road Initiative（Cyrill, 2018）. The EU has pursued a policy of "digital sovereignty" to reduce reliance on foreign technologies（Madiega, 2020）.

Second-order cooperation problems arise from the dysfunction of international institutions required to address complex policy issues. Although new institutions have been created to manage emerging challenges, the resulting institutional fragmentation and overlapping mandates limit their effectiveness（Roberts et al., 2024）. Given that AI capabilities and regulation are concentrated in the US, China, and the EU, a multilateral agreement between these three powers could significantly advance global AI governance. However, each jurisdiction adopts distinct policy approaches, with the US favouring a laissez-faire approach, the EU enforcing legislative control, and China employing a hybrid model of self-discipline and targeted legislation（Samson, 2023）.

9. Conclusion and Way Forward

9.1 Summary of Key Findings

Our comparative analysis of **EU and Chinese AI regulations** reveals both convergences and divergences in their approaches to AI governance. Both frameworks prioritise **responsible AI development**, focusing on common themes like safety, security, fairness, transparency, and accountability. However, differences stem from their **political systems**—with the EU emphasising **fundamental rights** and **legislative control** while China integrates **socialistic values**, **national security**, and **industry self-discipline**.

We explored the **barriers to global AI governance**, identifying the "**first-order cooperation problems**" rooted in international anarchy and geopolitical rivalries,

particularly among the **USA, EU, and China**. Additionally, "**second-order cooperation problems**" arise from the dysfunction of international institutions, exacerbating the fragmentation of AI governance. Despite these challenges, there is scope for consensus-building and global governance through a **regime complex model**, which allows for **incremental progress** across multiple international forums.

9.2 Outlook on Future Global Governance of AGI

The future of **Artificial General Intelligence (AGI)** governance will be crucial in determining the trajectory of **global prosperity and innovation**. AGI, with its potential to surpass human intelligence, brings both immense opportunities and profound risks. It has the potential to revolutionise industries, boost productivity, and solve some of the world's most pressing challenges. However, it also poses risks to **national security**, **economic inequality**, and **societal cohesion**.

Given the **decentralised nature of AI**, future **global AGI governance** will need to strike a delicate balance between **cooperation** and **competition** among leading global powers. The **feasibility** of a unified **global governance framework for AGI** is contingent on overcoming geopolitical rivalries, particularly among China, the USA, and the EU. The **impact on global prosperity** could be significant if governance frameworks foster **open innovation ecosystems**, ensure **equitable access** to AI technologies, and mitigate the risk of **AGI monopolies** controlled by a handful of powerful nations or corporations. A **global AGI governance model** that encourages **inclusive development** can help bridge the **digital divide** and ensure that the benefits of AGI reach **developing countries** as well as **global AI hubs**. However, the risks of **fragmentation** due to differing political ideologies and regulatory approaches remain high.

Feasibility and global impact will largely depend on the cooperation of a few key actors, namely the **USA**, **EU**, and **China**. A **multilateral framework** agreed upon by these actors could lay the groundwork for broader global cooperation, especially if it addresses **shared concerns** like AGI safety, **bias mitigation**, and **responsible innovation**. Yet, the fragmented nature of global governance, coupled with the rapid technological advancements of AGI, makes this a challenging task. Nonetheless, the **long-term benefits** of **robust AGI governance**—from enhanced global prosperity to sustained innovation—underscore the necessity of concerted efforts on a global scale.

9.3 Policy Recommendations

Drawing from the analysis, several policy recommendations emerge as crucial for shaping future AI and AGI governance frameworks.

To begin with, **adopting a risk-based, multi-tiered regulatory approach** is fundamental. Modeled after the EU's AI Act, future AGI regulations should classify systems based on their risk profile, ensuring that regulatory efforts focus primarily on high-risk applications. These would include AGI technologies with implications for national security, biometric data, and critical decision-making processes (European Commission, 2021). Such an approach allows for efficient prioritization, which is vital in the dynamic field of AGI development (Cath, 2018).

Equally important is the **promotion of open standards and interoperability**. For global AI governance to remain coherent, it is imperative to encourage collaboration through international technical standards that ensure seamless regulatory integration across borders. Governments must actively support efforts to standardize AGI systems, creating mechanisms for interoperability (Floridi et al., 2018).

Next, **mandating human oversight and control** is essential for maintaining transparency and accountability in AGI systems. High-risk domains should have built-in human-in-the-loop safeguards to ensure critical decisions can be reviewed and acted upon by human operators, preventing system autonomy from leading to unintended outcomes (Jobin, Ienca, & Vayena, 2019).

Furthermore, **establishing international AGI governance forums** is necessary to foster global cooperation. These forums, potentially within the framework of existing organizations like the United Nations or as independent entities, should prioritize the development of ethical guidelines, safety protocols, and international monitoring mechanisms for AGI systems. Through collaboration, nations can create robust governance systems that reflect shared values and goals (Floridi et al., 2018).

Encouraging ethical AI research and innovation should also be a priority. Governments can incentivize ethical AI research by funding initiatives that promote fairness, transparency, and non-discrimination. This support can extend to universities, startups, and private sector companies that demonstrate a commitment to developing AI systems aligned with ethical principles (European Commission, 2019).

Additionally, **establishing regulatory sandboxes**—as seen in regulatory frameworks in both China and the EU—offers a controlled environment where AGI systems can be developed and tested safely. Sandboxes provide a means for fostering innovation while ensuring that ethical and safety concerns are appropriately managed (Campbell, 2019).

Lastly, **strengthening international cooperation on AGI safety** is imperative. Coordinated global efforts in safety research, information sharing, and collaborative testing infrastructure are necessary to mitigate the potential risks of AGI. The

international community must work together to ensure that AGI development remains aligned with global security standards and ethical principles (Bostrom, 2014; Yudkowsky, 2006).

9.4 Limitations

While this research provides valuable insights into the governance of AI and AGI, several limitations must be acknowledged.

One significant limitation is the **scope of the case studies** used to examine the regulatory approaches of the EU and China. By focusing primarily on these two regions, the analysis may have overlooked valuable insights from other nations with evolving AI regulatory frameworks. This limited geographic focus potentially reduces the generalizability of the findings to a broader, global context.

Another constraint lies in the **rapid pace of technological development** in the AGI field. AGI systems are progressing faster than regulatory bodies can adapt, creating a temporal gap between AGI advancements and the development of appropriate governance frameworks. This research, conducted within a specific timeframe, may not account for the most recent technological and regulatory developments, limiting the applicability of the findings over time.

Additionally, the research was limited by the **paucity of time** available for document analysis. Due to time constraints, the range of policy documents, case studies, and reports that could be thoroughly reviewed was restricted. This constraint meant that certain emerging or lesser-known AI governance frameworks may not have been fully captured in this analysis, potentially limiting the comprehensiveness of the findings.

The study also faced **data limitations**, particularly in accessing up-to-date and complete policy documents from emerging AI economies. This restricts a more nuanced understanding of global AI governance, especially in underrepresented regions where AI regulatory frameworks may be nascent or evolving.

Finally, **theoretical limitations** arise from the interdisciplinary nature of AI governance. While this study draws from multiple disciplines, including law, ethics, and technology studies, it may not fully address the complexity of integrating these fields into a cohesive governance model. Further interdisciplinary collaboration is necessary to refine these frameworks and provide a more comprehensive perspective on global AI governance.

9.5 Scope for Further Research

Global AI and AGI ethics require in-depth exploration, particularly with regard to how different cultural contexts approach the regulation and ethical considerations of AGI systems. Cross-cultural ethical research will help address divergent regional values, such as the prioritisation of fundamental rights in the EU versus national security in China (Floridi et al., 2018).

Moreover, the **role of AI governance in developing economies** remains underexplored. Research on how developing nations can both contribute to and benefit from global AI governance is essential, particularly in ensuring that AI does not exacerbate global inequalities (Cath, 2018).

Further studies should also focus on **the impact of AGI on global labor markets**, especially in sectors vulnerable to automation. Research in this area could inform policy responses to job displacement and economic shifts caused by AGI advancements (Brynjolfsson & McAfee, 2014).

In addition, there is a growing need to study **public trust and acceptance of AGI systems** across various cultural, political, and economic settings. Public trust will significantly influence the effectiveness and legitimacy of global governance frameworks (Jobin, Ienca, & Vayena, 2019).

Lastly, the **environmental sustainability of AGI** demands more research. As AGI development is resource-intensive, studying the ecological impact of AGI technologies is critical for creating sustainable practices in the design, development, and deployment of these systems (Crawford & Calo, 2016).

*To conclude, as we stand at the precipice of AGI development, it is clear that a **collaborative, multi-stakeholder approach** will be necessary to ensure **responsible governance** that maximises the benefits of AGI while mitigating its risks. By learning from existing models and fostering **global cooperation**, we can create a **robust and adaptive governance framework** that supports innovation and ensures equitable, safe, and inclusive AGI development for future generations.*

Bibliography

1. Allen, G. C. (2019). Understanding China's AI Strategy: Clues to Chinese Strategic Thinking on Artificial Intelligence and National Security. Center for a New American Security. https://www.jstor.org/stable/resrep20446
2. Allnut, H., & Hardy, A. (2024). Open AI: Recent investigations and claims in the United States and the EU. *Lexology*. https://www.lexology.com/library/detail.aspx?g=014b6177-74f2-49df-8347-eb806b8f81ba
3. Bayamlioglu, E., Baraliuc, I., Janssens, L. A. W., & Hildebrandt, M. (2018). ETHICS AS AN ESCAPE FROM REGULATION: FROM "ETHICS-WASHING" TO ETHICS-SHOPPING? In BEING PROFILED (pp. 84–89). Amsterdam University Press. https://doi.org/10.1515/9789048550180-016
4. Beijing Academy of Artificial Intelligence. (2019). Beijing AI Principles. Datenschutz Datensich, 43, 656.
5. Binns, R. (2021). Fairness in Machine Learning: Lessons from Political Philosophy. https://arxiv.org/abs/1712.03586
6. Bostrom, N. (2014). Superintelligence: Paths, dangers, strategies. Oxford University Press, Incorporated. https://global.oup.com/academic/product/superintelligence-9780198739838?cc=us&lang=en& (Last accessed on: 16th May 2024)
7. Bradford, A. (2023, June 27). The Race to Regulate Artificial Intelligence. *Foreign Affairs*. https://www.foreignaffairs.com/articles/world/2023-06-27/race-regulate-artificial-intelligence
8. Brynjolfsson, E., & McAfee, A. (2014). The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies. W.W. Norton & Company. URL: https://wwnorton.com/books/The-Second-Machine-Age/ (Last accessed on: 16th May 2024)
9. Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research*, 81, 1-15. https://www.scirp.org/reference/referencespapers?referenceid=3614756
10. Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1), 1-12. https://www.scirp.org/reference/referencespapers?referenceid=2624207
11. Calo, R. (2017). Artificial Intelligence Policy: A Primer and Roadmap. U.C. Davis Law Review, 51, 399-435. URL: https://lawreview.sf.ucdavis.edu/sites/g/files/dgvnsk15026/files/media/documents/51-2_Calo.pdf (Last accessed on: 16th May 2024)

12. Campbell, T. A. (2019). *Artificial Intelligence: An Overview of State Initiatives*. Evergreen, CO: FutureGrasp, LLC. https://www.researchgate.net/publication/334731776
13. Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M., & Floridi, L. (2018). Artificial Intelligence and the 'Good Society': the US, EU, and UK approach. Science and Engineering Ethics, 24(2), 505–528. https://link.springer.com/article/10.1007/s11948-017-9901-7
14. Cath, C. (2018). Governing Artificial Intelligence: Ethical, Legal and Technical Opportunities and Challenges. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 376(2133), 20180080. URL: https://royalsocietypublishing.org/doi/10.1098/rsta.2018.0080 (Last accessed on: 16th May 2024)
15. China State Council. (2017). New Generation Artificial Intelligence Development Plan. URL: https://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm (Last accessed on: 16th May 2024)
16. Cook, T. D. (1985), "Postpositivist Critical Multiplism," in R. Shotland and M. Mark (eds), Social Science and Social Policy, (Beverly Hills, CA: Sage), pp. 21-62. Sage. https://link.springer.com/article/10.1023/A:1012749120909
17. Crawford, K., & Calo, R. (2016). There is a Blind Spot in AI Research. Nature, 538, 311-313. URL: https://www.nature.com/articles/538311a (Last accessed on: 16th May 2024)
18. Cyrill, M. (2018). What is Made in China 2025 and why has it made the world so nervous? *China Briefing*. https://www.china-briefing.com/news/made-in-china-2025-explained/
19. Daddow, O. (2011). New Labour and the European Union: Blair and Brown's logic of history (1st ed.). Manchester University Press. https://www.jstor.org/stable/j.ctvnb7rxc (Last accessed on: 16th May 2024)
20. Danks, D., & London, A. J. (2017). Algorithmic Bias in Autonomous Systems. Proceedings of the 26th International Joint Conference on Artificial Intelligence, 4691-4697. URL: https://www.ijcai.org/proceedings/2017/0654.pdf (Last accessed on: 16th May 2024)
21. Davenport, T. H., & Ronanki, R. (2018). Artificial Intelligence for the Real World. Harvard Business Review, 96(1), 108-116. URL: https://hbr.org/2018/01/artificial-intelligence-for-the-real-world (Last accessed on: 16th May 2024)
22. Diakopoulos, N. (2016). Accountability in Algorithmic Decision Making. Communications of the ACM, 59(2), 56-62. URL: https://dl.acm.org/doi/10.1145/2844110 (Last accessed on: 16th May 2024)
23. Ding, J. (2018). *Deciphering China's AI Dream*. Future of Humanity Institute, University of Oxford. Retrieved from https://www.fhi.ox.ac.uk/wp-content/uploads/Deciphering_Chinas_AI-Dream.pdf

24. Dutton, W. H. (2023). The Fifth Estate: The Power Shift of the Digital Age (1st ed.). Oxford University Press. https://academic.oup.com/book/46084
25. EDPB. (2020). European Data Protection Board: Annual Report 2020. URL: https://www.edpb.europa.eu/our-work-tools/our-documents/annual-report/edpb-annual-report-2020_en (Last accessed on: 16th May 2024)
26. Esteva, A., Robicquet, A., Ramsundar, B., et al. (2019). A Guide to Deep Learning in Healthcare. Nature Medicine, 25, 24-29. URL: https://www.nature.com/articles/s41591-018-0316-z (Last accessed on: 16th May 2024)
27. European Commission. (2020). Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonized Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts. URL: https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence (Last accessed on: 16th May 2024)
28. European Commission. (2024). *The Artificial Intelligence Act: Regulating AI for Trustworthiness and Innovation*. European Commission Publications. https://ec.europa.eu/commission/presscorner/detail/en/ip_24_4123
29. European Union. (2024, June). *The EU Artificial Intelligence Act*. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202401689
30. European Union. (2022, October). *Digital Services Act*. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32022R2065
31. European Union. (2022, September). *Digital Markets Act*. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32022R1925
32. European Commission. (2019, April). *Ethics guidelines for trustworthy AI*. https://dcoe.org/uploads/2019/06/EC-190408-AI-HLEG-Guidelines.pdf
33. European Union. (2016, April). *General Data Protection Regulation*. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679
34. Fisher, E., & Rip, A. (2013). Responsible Innovation: Multi-Level Dynamics and Soft Intervention Practices. In Responsible Innovation (pp. 165–183). John Wiley & Sons, Ltd. https://library.oapen.org/bitstream/id/2431f06c-0034-4457-8b76-e09e34b9332b/9781000292749.pdf
35. Fjeld, J., Achten, N., Hilligoss, H., Nagy, A., & Srikumar, M. (2020). *Principled artificial intelligence: Mapping consensus in ethical and rights-based approaches to principles for AI*. Berkman Klein Center for Internet & Society. https://cyber.harvard.edu/publication/2020/principled-ai
36. Floridi, L. (2021). *The European legislation on AI: A brief analysis of its philosophical approach*. *Philosophy and Technology, 34*(2), 215–222. https://doi.org/10.1007/s13347-021-00460-9
37. Floridi, L. (Ed.). (2021). *Ethics, governance, and policies in artificial intelligence*. Springer Verlag. https://doi.org/10.1007/978-3-030-81907-1

38. Floridi, L., Cowls, J., Beltrametti, M., et al. (2018). AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. Minds and Machines, 28(4), 689-707. URL: https://link.springer.com/article/10.1007/s11023-018-9482-5 (Last accessed on: 16th May 2024)
39. Galaz, V., Crona, B., Dauriach, A., Jouffray, J.-B., Österblom, H., & Troell, M. (2012). Polycentric systems and interacting planetary boundaries—emerging governance of climate change–ocean acidification–marine biodiversity. *Ecological Economics, 81*, 21–32. https://ideas.repec.org/a/eee/ecolec/v81y2012icp21-32.html
40. Goertzel, B., & Pennachin, C. (2007). Artificial General Intelligence. Springer. URL: https://link.springer.com/book/10.1007/978-3-540-68677-4 (Last accessed on: 16th May 2024)
41. Haenlein, M., & Kaplan, A. (2019). A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. California Management Review, 61(4), 5–14. https://journals.sagepub.com/doi/10.1177/0008125619864925
42. Hagendorff, T. (2020). *The ethics of AI ethics: An evaluation of guidelines*. *Minds and Machines, 30*(1), 99–120. https://doi.org/10.1007/s11023-020-09517-8
43. Hale, E. (2023). China wants AI to rival ChatGPT. Censorship makes that tricky. *Al Jazeera*. https://www.aljazeera.com/economy/2023/3/2/china-wants-to-copy-chatgpts-success-censorship-makes-it-tricky
44. Harwell, D., & Duo, E. (2020). Huawei tested AI software that could recognize Uighur minorities and alert police, report says. *The Washington Post*. https://www.washingtonpost.com/technology/2020/12/08/huawei-tested-ai-software-that-could-recognize-uighur-minorities-alert-police-report-says/
45. Jasanoff, S. (2016). The ethics of invention: Technology and the human future. W. W. Norton & Company. https://wwnorton.com/books/The-Ethics-of-Invention/
46. Jobin, A., Ienca, M., & Vayena, E. (2019). The Global Landscape of AI Ethics Guidelines. Nature Machine Intelligence, 1, 389-399. URL: https://www.nature.com/articles/s42256-019-0088-2 (Last accessed on: 16th May 2024)
47. Jonas Tallberg, Eva Erman, Markus Furendal, Johannes Geith, Mark Klamberg, Magnus Lundgren, The Global Governance of Artificial Intelligence: Next Steps for Empirical and Normative Research, *International Studies Review*, Volume 25, Issue 3, September 2023, viad040, https://academic-oup-com.libproxy.ucl.ac.uk/isr/article/25/3/viad040/7259354 (Last accessed on: 14th June 2024)

48. Kania, E. B. (2021). Artificial intelligence in China's revolution in military affairs. Journal of Strategic Studies, 44(4), 515–542. https://www.tandfonline.com/doi/full/10.1080/01402390.2021.1894136 (Last accessed on: 16th May 2024)
49. Keohane, R. O., & Victor, D. G. (2011). The regime complex for climate change. *Perspectives on Politics,* *9*(1), 7–23. https://www.cambridge.org/core/journals/perspectives-on-politics/article/abs/regime-complex-for-climate-change/F5C4F620A4723D5DA5E0ACDC48D860C0
50. Lechner, S. (2022). Anarchy in international relations. *Oxford Research Encyclopedia of International Studies*. https://oxfordre.com/internationalstudies/view/10.1093/acrefore/9780190846626.001.0001/acrefore-9780190846626-e-79
51. Lee, K. (2018). AI Superpowers: China, Silicon Valley, and the New World Order. Germany: HarperCollins. https://dl.acm.org/doi/10.5555/3299596
52. Liang, F., Das, V., Kostyuk, N., & Hussain, M. M. (2018). Constructing a Data-Driven Society: China's Social Credit System as a State Surveillance Infrastructure. Policy and Internet, 10(4), 415–453. https://onlinelibrary.wiley.com/doi/10.1002/poi3.183
53. Macnaghten, P., Kearnes, M. B., & Wynne, B. (2005). Nanotechnology, Governance, and Public Deliberation: What Role for the Social Sciences? Science Communication, 27(2), 268–291. https://journals.sagepub.com/doi/10.1177/1075547005281531
54. Madiega, T. (2020). Digital sovereignty for Europe. *European Parliamentary Research Service*. https://www.europarl.europa.eu/RegData/etudes/BRIE/2020/651992/EPRS_BRI(2020)651992_EN.pdf
55. Mozur, P. (2018). Inside China's Dystopian Dreams: A.I., Shame and Lots of Cameras. The New York Times. URL: https://www.nytimes.com/2018/07/08/business/china-surveillance-technology.html (Last accessed on: 16th May 2024)
56. Müller, V. C., & Bostrom, N. (2016). Future Progress in Artificial Intelligence: A Survey of Expert Opinion. In V. C. Müller (Ed.), Fundamental Issues of Artificial Intelligence (pp. 555-572). https://research.tue.nl/en/publications/future-progress-in-artificial-intelligence-a-survey-of-expert-opi (Last accessed on: 16th May 2024)
57. National Technical Committee 260. (2024). *Basic safety requirements for generative artificial intelligence services*. Cybersecurity of Standardization Administration of China. https://cset.georgetown.edu/publication/china-safety-requirements-for-generative-ai-final/

58. Natlareview.com. (2023). China Releases Draft Measures for the Management of Generative Artificial Intelligence Services. https://natlawreview.com/article/china-releases-draft-measures-management-generative-artificial-intelligence-services#:~:text=On%20April%2011%2C%202023%2C%20the%20Cyberspace%20Administration%20of,the%20territory%20of%20the%20People%E2%80%99s%20Republic%20of%20China.
59. NSCAI Final Report. (2021). *Final Report of the National Security Commission on Artificial Intelligence*. Washington, DC: The National Security Commission on Artificial Intelligence. https://www.nscai.gov/wp-content/uploads/2021/03/Full-Report-Digital-1.pdf
60. O'Neil, C. (2016). Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. Crown. URL: https://dl.acm.org/doi/10.5555/3002861 (Last accessed on: 16th May 2024)
61. Owen, R., Macnaghten, P., & Stilgoe, J. (2012). Responsible research and innovation: From science in society to science for society, with society. Science & Public Policy, 39(6), 751–760. https://academic.oup.com/spp/article-abstract/39/6/751/1620724?redirectedFrom=fulltext
62. People's Republic of China. (2024, April). *Model Artificial Intelligence Law (MAIL) v.2.0*. https://www.scribd.com/document/726703792/1713270587983
63. People's Republic of China. (2024, May). *Artificial Intelligence Law of the People's Republic of China (Draft for Suggestions from Scholars)*. https://cset.georgetown.edu/publication/china-ai-law-draft/
64. People's Republic of China. (2024, April). *Basic Safety Requirements for Generative Artificial Intelligence Services*. https://cset.georgetown.edu/publication/china-safety-requirements-for-generative-ai-final/
65. People's Republic of China. (2024, January). *Guidelines for the Construction of a Comprehensive Standardization System for the National Artificial Intelligence Industry (Draft for Feedback)*. https://cset.georgetown.edu/publication/china-ai-standards-system-guidelines-draft/
66. People's Republic of China. (2022, November). *Provisions on the Administration of Deep Synthesis Internet Information Services*. https://www.chinalawtranslate.com/en/deep-synthesis/
67. People's Republic of China. (2022, March). *Opinions on Strengthening the Management of Science and Technology Ethics*. https://www.gov.cn/gongbao/content/2022/content_5683838.htm
68. People's Republic of China. (2021, December). *Provisions on the Management of Algorithmic Recommendations in Internet Information Services*. https://www.chinalawtranslate.com/en/algorithms/

69. People's Republic of China. (2021, September). *Ethical Norms for New Generation Artificial Intelligence Released*. https://cset.georgetown.edu/wp-content/uploads/t0400_AI_ethical_norms_EN.pdf
70. People's Republic of China. (2021, July). *Artificial Intelligence Standardization White Paper (2021 Edition)*. https://cset.georgetown.edu/publication/artificial-intelligence-standardization-white-paper-2021-edition/
71. People's Republic of China. (2021, November). *The PRC Personal Information Protection Law*. https://www.china-briefing.com/news/the-prc-personal-information-protection-law-final-a-full-translation/
72. People's Republic of China. (2019, June). *Governance Principles for a New Generation of Artificial Intelligence: Develop Responsible Artificial Intelligence*. https://digichina.stanford.edu/work/translation-chinese-expert-group-offers-governance-principles-for-responsible-ai/
73. Rajagopalan, R. (2024). AI chips for China face additional US restrictions. *The Diplomat*. https://thediplomat.com/2024/04/ai-chips-for-china-face-additional-us-restrictions/
74. Roberts, H., Hine, E., Taddeo, M., & Floridi, L. (2024). Global AI governance: Barriers and pathways forward. *International Affairs*. https://ssrn.com/abstract=4588040 or http://dx.doi.org/10.2139/ssrn.4588040
75. Roberts, H., Hine, E., Floridi, L. (2023). Digital Sovereignty, Digital Expansionism, and the Prospects for Global AI Governance. In: Timoteo, M., Verri, B., Nanni, R. (eds) Quo Vadis, Sovereignty? . Philosophical Studies Series, vol 154. Springer, Cham. https://doi.org/10.1007/978-3-031-41566-1_4
76. Russell, S. (2019). Human-Compatible Artificial Intelligence. Computer Science Division, University of California, Berkeley. URL: https://people.eecs.berkeley.edu/~russell/papers/mi19book-hcai.pdf (Last accessed on: 16th May 2024)
77. Russell, S., & Norvig, P. (2016). Artificial Intelligence: A Modern Approach (3rd ed.). Pearson. URL: https://aima.cs.berkeley.edu/global-index.html (Last accessed on: 16th May 2024)
78. Sahin, K. (2020). The West, China, and AI surveillance. *Atlantic Council GeoTech Cues*. https://www.atlanticcouncil.org/blogs/geotech-cues/the-west-china-and-ai-surveillance/
79. Samson, P. (2023). On advancing global AI governance. *Centre for International Governance Innovation*. https://www.cigionline.org/articles/on-advancing-global-ai-governance
80. Smith, B., & Crampton, N. (2024). *Global Governance: Goals and Lessons for AI*. Microsoft. https://dcgdigitaldelivery.blob.core.windows.net/global-governance/Global_Governance_Goals_and_Lessons_for_AI-2024_Whitepaper_AC.pdf

81. Stanford University. (2023). *Artificial Intelligence Index Report 2023: Chapter 6: Policy and Governance*. https://aiindex.stanford.edu/report/
82. State Council of the People's Republic of China. (2017). *New Generation Artificial Intelligence Development Plan*. https://english.www.gov.cn/policies/latest_releases/2017/07/20/content_281475742458322.htm
83. Stilgoe, J., Owen, R., & Macnaghten, P. (2013). Developing a framework for responsible innovation. Research Policy, 42(9), 1568–1580. https://www.sciencedirect.com/science/article/pii/S0048733313000930?via%3Dihub
84. Todd J (2016). The UK's relationship with Europe. Springer. URL https://link.springer.com/book/10.1007/978-3-319-33669-5 (Last accessed on: 16th May 2024)
85. United Nations. (2023). First Committee approves draft resolution on lethal autonomous weapons systems. United Nations. Retrieved from https://press.un.org/en/2023/gadis3731.doc.htm
86. Veale, M., & Binns, R. (2017). Fairer Machine Learning in the Real World: Mitigating Discrimination Without Collecting Sensitive Data. Big Data & Society, 4(2), 1-17. URL: https://journals.sagepub.com/doi/full/10.1177/2053951717743530 (Last accessed on: 16th May 2024)
87. Voigt, P., & Von dem Bussche, A. (2017). The EU General Data Protection Regulation (GDPR): A Practical Guide. Springer. URL: https://link.springer.com/book/10.1007/978-3-319-57959-7 (Last accessed on: 16th May 2024)
88. Weiss, T. G. (2000). Governance, good governance and global governance: Conceptual and actual challenges. *Third World Quarterly, 21*(5), 795–814. https://www.scirp.org/reference/referencespapers?referenceid=1255325
89. White House. (2022). *Blueprint for an AI Bill of Rights*. https://www.whitehouse.gov/ostp/ai-bill-of-rights/
90. White House. (2023). FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence. *The White House*. https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/
91. Whittaker, M., Crawford, K., Dobbe, R., et al. (2018). AI Now Report 2018. AI Now Institute. URL: https://ec.europa.eu/futurium/en/system/files/ged/ai_now_2018_report.pdf
92. Whyman PB, Petrescu AI (2017). The economics of Brexit: a cost-benefit analysis of the UK's economic relationship with the EU. Palgrave Macmillan,

London. URL: https://link.springer.com/book/10.1007/978-3-319-58283-2 (Last accessed on: 16th May 2024)
93. Yudkowsky, E. (2006). Artificial Intelligence as a Positive and Negative Factor in Global Risk. https://academic.oup.com/book/40615/chapter/348239228
94. Zappettini F (2019). The official vision for 'global Britain': Brexit as rupture and continuity between free trade, liberal internationalism and 'values'. In:Koller K, Kopf S, Miglbauer M (eds) Discourses of Brexit. Routledge, New York, pp. 140–154.URL: https://www.taylorfrancis.com/chapters/edit/10.4324/9781351041867-9/official-vision-global-britain-franco-zappettini (Last accessed on: 16th May 2024)
95. Zhang, A. H. (2024). The promise and perils of China's regulation of artificial intelligence. *Columbia Journal of Transnational Law*. https://ssrn.com/abstract=4708676

Appendix

Appendix Table: List of Codes and Themes Created for Core Documents

Theme	**Codes**	**EU AI Act 2024**	**MAIL v 2.0 224**	**Deep Synthesis Law 2022**	**Algorithmic Recommendations Law 2021**
Legislative basis/Subject matter	Legislative basis/Subject matter	Chapter 1: Article 1: Subject matter	Chapter 1: Article 1:Legislative basis	Chapter 1(GENERAL) Article 1	Chapter 1(GENERAL) Article 1
Scope of application	Scope of application (Where this law applies and where does not)	Chapter 1: Article 2: Scope Will not apply in specific scenarios: e.g if exclusively for military, defence or	Chapter 1: Article 2: Scope of application ARTICLE 77 – MILITARY ARTIFICIAL INTELLIGENCE The regulations governing the R&D, provision, and use of	Chapter 1(GENERAL) Article 2	

		national security purposes Use for the sole purpose of scientific research and development. Use for purely personal non-professional activity AI systems are released under free and open-source licences, unless they are placed on the market or put into service as high-risk AI systems or as an AI system that falls under Article 5 or 50	artificial intelligence by the Chinese People's Liberation Army and the Chinese People's Armed Police Force shall be separately stipulated by the Central Military Commission in accordance with the principles prescribed in this Law.		

Theme	Codes	EU AI Act 2024	MAIL v 2.0 224	Deep Synthesis Law 2022	Algorithmic Recommendations Law 2021
THEMES RELATED TO RISK MANAGEMENT					
AI oversight system / Regulatory approach	Categorised AI oversight system	CHAPTER II (PROHIBITED AI PRACTICES) Article 5; CHAPTER III (HIGH-RISK AI SYSTEMS) Article 6; Article 7; Article 8; Article 16(a) CHAPTER V (GENERAL-PURPOSE AI MODELS): Article 51.	Chapter III: Article 25;: Article 26 to 33.		

Safety and Security	Risk management system; Internal management systems; Record keeping of automatically generated logs; technical documentation; education and training of employees; checking accuracy, robustness and cybersecurity Auditing; Post market surveillance, monitoring and	CHAPTER III (HIGH-RISK AI SYSTEMS) Article 9; Article 14: Article 15: Article 18; Article 16 (d); Article 19; Article 16 (e); Article 11; Article 12; Article 26 (6). Chapter I: Article 4. CHAPTER V(GENERAL-PURPOSE AI MODELS) Article 53; Article 55. CHAPTER IX (POST-MARKET MONITORING, INFORMATION SHARING AND MARKET	Chapter 1: Article 5. Chapter IV: Article 34; Article 35; Article 41; Article 46 (1); Article 49 (a); Article 34; Article 45 (a); Article 49;Article 52.	Chapter II: Ordinary Provisions; Article 7., 8, 13 Chapter III (Data and Technical Management Specifications): Article 15,16 Chapter IV: Oversight Inspections and Legal Responsibility: Article 20	Chapter II (Regulation of Information Services) Article 7.8 Chapter IV: (Oversight and Managemen)t Article 27.28

	information sharing	SURVEILLANCE): SECTION 1(Post-market monitoring): Article 72; SECTION 2 (Sharing of information on serious incidents :Article 73; SECTION 3(Enforcement): Article 74 to 84.			
Fairness, equality and Non-discrimination	Fairness, equality and Non-discrimination	CHAPTER X CODES OF CONDUCT AND GUIDELINES: Article 95- 2 (e): Codes of conduct for voluntary application of specific requirements:	Chapter 1; Article 8: Chapter IV: Article 40		

		assessing and preventing the negative impact of AI systems on vulnerable persons or groups of vulnerable persons, including as regards accessibility for persons with a disability, as well as on gender equality. CHAPTER X CODES OF CONDUCT AND GUIDELINES Article 95- 2 (e): Codes of conduct for voluntary application of			

		specific requirements: assessing and preventing the negative impact of AI systems on vulnerable persons or groups of vulnerable persons, including as regards accessibility for persons with a disability, as well as on gender equality. Article 95- 2 (e): Codes of conduct for voluntary application of specific requirements:			

		assessing and preventing the negative impact of AI systems on vulnerable persons or groups of vulnerable persons, including as regards accessibility for persons with a disability, as well as on gender equality.			
	Model and data stewardship	CHAPTER III (HIGH-RISK AI SYSTEMS): Article 26 (2): Obligations of deployers of high-risk AI systems- deployer shall	Chapter IV: Article 46 (2): Special obligations for developers of foundation models- Establish and maintain	Chapter III (Data and Technical Management Specifications): Article 14:	

		ensure that input data is relevant and sufficiently representative in view of the intended purpose of the high-risk AI system.	a comprehensive model and data stewardship system for foundational models in accordance with national regulations.		
	Data governance and management	Article 10: Data and data governance for high risk systems	Chapter II: Article 18: Supply of data production factors		
	Inclusiveness in Design/ Multi-stake collaboration/Public oversight	CHAPTER X CODES OF CONDUCT AND GUIDELINES Article 95- 2 (d): Codes of conduct for voluntary	Chapter IV: (obligations of AI developers and providers/deployers) Article 46 (7)	10	

		application of specific requirements: facilitating an inclusive and diverse design of AI systems, including through the establishment of inclusive and diverse development teams and the promotion of stakeholders' participation in that process			
Human-centric principle/Promotion of Human Values	Human Review of Automated Decision; Ability to Opt out of	CHAPTER III (HIGH-RISK AI SYSTEMS) Article 27:;	Chapter 1: Article 4;Article 14 (a);(b) (c)	Chapter 1(GENERAL) Article 1; Article 4	Chapter 1(GENERAL) Article 1, Article 4

	Automated Decisions Human Values and Human Flourishing	Article 95- 2 (e):		Chapter II: Ordinary Provisions; Article 6:	Chapter II: Regulation of Information Services; Article 6, 14,15,17 Chapter III (Protection of User's Rights and Interests) Article 18 to 21
	Content moderation			Chapter II: Ordinary Provisions; Article 10	Chapter II: Regulation of Information Services; Article 9,10,11, 13
Openness, Transparency and Explainability	Openness and transparency Information that user is interacting with AI/synthesised content and inform that the content is AI generated	CHAPTER III (HIGH-RISK AI SYSTEMS): Article 13; Article 26 (6):; Article 26 (11). CHAPTER IV (TRANSPARENCY OBLIGATIONS FOR PROVIDERS AND DEPLOYERS OF CERTAIN AI	Chapter 1: Article 6:; Chapter IV: Article 38; Article 46 (1);: Article 38 (a to d): Article 46 (1).	Chapter III (Data and Technical Management Specifications): Article 17,18	Chapter II: Regulation of Information Services; Article 12 Chapter III (Protection of User's Rights and Interests) Article 16

		SYSTEMS) Article 50 .			
	Explainability	Annex IV	Chapter IV: Article 39:.		
Theme	**Codes**	**EU AI Act 2024**	**MAIL v 2.0 224**	**Deep Synthesis Law 2022**	**Algorithmic Recommendations Law 2021**
Privacy	Privacy Consent Control over the use of data Ability to restrict data processing Right to rectification	Article 10-5 (b); Point 69 Article 2(7)	Chapter 1: Article 14 (c)	Chapter II: Ordinary Provisions; Article 9. Chapter III (Data and Technical Management Specifications): Article 14	Chapter IV: (Oversight and Management) Article 29

	Right to erasure/revocation				
Professional Responsibility	Responsible Design Impact Assessments	Article 27: Article 95- 2 (a): Article 95- 2 (b): Article 95- 2 (e):	Chapter 1: Article 9: Green Principle (of Sustainability)	Chapter 1(GENERAL) Article 5 Chapter II: Ordinary Provisions; Article 6:	
Accountability	Quality management system Codes of practice Human control/oversight of AI	CHAPTER III (HIGH-RISK AI SYSTEMS) Article 17; Article 16 (c); Article 14; Article 26 (2). CHAPTER V(GENERAL-PURPOSE AI MODELS) Article 56	Chapter 1; Article 7.y Chapter IV: Article 36; Article 45 (a); Article 52 (c);		

	Remedy and notify/Corrective actions and duty of information	CHAPTER III (HIGH-RISK AI SYSTEMS) Article 20; Article 16 (j);	Chapter IV: Article 37. Chapter IV: Article 50 to 51.		
	Designate authorised representatives for AI developers and providers located outside the State.	CHAPTER III (HIGH-RISK AI SYSTEMS) Article 22 (1).	Chapter IV: Article 44.		
	Other entities in AI value chain [besides developer, provider, and authorised representataives]	CHAPTER III (HIGH-RISK AI SYSTEMS): Article 23 to 25. CHAPTER V(GENERAL-PURPOSE AI			

		MODELS) Article 54			
	Liabilities/ Penalties	CHAPTER XII(PENALTIES): Article 99 to 101	Chapter VI: Article 66 to 68; Article 69; Article 71; Article 72	Chapter IV (Oversight Inspections and Legal Responsibility) Article 21	Chapter IV: (Oversight and Management) Article 30 Chapter V (Legal Responsibility) Article 31 to 33.
	Remedies	CHAPTER IX (POST-MARKET MONITORING, INFORMATION SHARING AND MARKET SURVEILLANCE): SECTION 4 (Remedies) Article 85 to 87	Chapter VI: Article 70; Article 73.	Chapter II: Ordinary Provisions; Article 12 Chapter IV (Oversight Inspections and Legal Responsibility) Article 21	Chapter III (Protection of User's Rights and Interests) Article 22
THEMES RELATED TO SUPPORT INNOVATION					

State support/measures to promote AI development and innovations	Govt. makes and implements plans for development of AI Construction of computing infrastructure Support innovation in algorithms, open source AI and foundation models Support construction of foundational and specialised databases, develop national integrated big	CHAPTER VI (MEASURES IN SUPPORT OF INNOVATION) Article 57 to 62	Chapter 1: Article 10; Chapter II:: Article 15 to 24. Chapter V: Article 59-60. Chapter VI: Article 71.		

	data centre system Promote AI industrial development, and integration/application of AI in various industries Support professional talent cultivation institutions and mechanisms Provide fiscal and procurement support , allocate special budgets foray support and development				

	Tax credit incentives Pilot projects for AI application in govt and public management Establish Regulatory sandboxes				
Governance mechanisms	Organograms Advisory forums Panel of experts)/ Creation of monitoring body/ notifying authorities, and conformity assessment bodies	CHAPTER VII (GOVERNANCE) Article 64: Article 65: Article 66: Article 67: Article 68: Article 69: Article 70: CHAPTER III (HIGH-RISK AI SYSTEMS): Article 28:es of high risk systems	Chapter 1: Article 3: Article 12: Article 13. Chapter V: Article 54; Article 55; Article 56;. Article 57: Article 58.: Article 59: Article 61: Article 62:Article 63:	Chapter 1(GENERAL) Article 3 Chapter IV (Oversight Inspections and Legal Responsibility) Article 21	Chapter 1(GENERAL) Article 3 Chapter IV: (Oversight and Managemen)t Article 23

	Data base	CHAPTER III (HIGH-RISK AI SYSTEMS):Article 29 to 38. CHAPTER VIII (EU DATABASE FOR HIGH-RISK AI SYSTEMS): Article 71			
	Detailed Procedures. Detailed Guidelines for implementation of regulations	CHAPTER III (HIGH-RISK AI SYSTEMS): Article 40 to 49. CHAPTER IX (POST-MARKET MONITORING, INFORMATION SHARING AND MARKET SURVEILLANCE): SECTION 5 : (Article 88 to 94)	Chapter IV: Article 47: Registry obligations for providers not listed in negative list Chapter IV: Article 48: Registry process: for providers not listed in negative list	Chapter IV (Oversight Inspections and Legal Responsibility) Article 19	Chapter IV: (Oversight and Managemen)t Article 24

		CHAPTER V (GENERAL-PURPOSE AI MODELS): Article 52. CHAPTER X (CODES OF CONDUCT AND GUIDELINES): Article 96. CHAPTER XI DELEGATION OF POWER AND COMMITTEE PROCEDURE: Article 97, 98			

	Liability fixation for government bodies	CHAPTER XII (PENALTIES): Article 100	Chapter VI(LIABILITIES): Article 76		
	International cooperation	NA	Chapter 1: Article 11 Chapter V: Article 64: Article 65:		