



**Università
degli Studi
di Palermo**

AREA RICERCA E TRASFERIMENTO TECNOLOGICO
SETTORE DOTTORATI E CONTRATTI PER LA RICERCA
U. O. DOTTORATI DI RICERCA

Doctoral Program in Model Based Public Planning, Policy Design and Management

Scienze Politiche e delle relazioni internazionali

Settore Scientifico Disciplinare, which is “SECS-P/07—Economia Aziendale

**ENHANCING LARGE LANGUAGE MODEL (LLM)
GOVERNANCE LEVEL: A DYNAMIC PERFORMANCE
GOVERNANCE (DPG) APPROACH**

IL DOTTORE

SHAN LI

IL COORDINATORE

Prof.SSA CHIARA GARILLI

IL TUTOR

CHIAR.MO PROF. ENZO BIVONA

CO TUTOR

CHIAR.MO PROF. JIAYIN QI

CHIAR.MO PROF. BO XU

CHIAR.MO PROF. HAIYAN YAN

CICLO XXXVII

ANNO CONSEGUIMENTO TITOLO 2025

Acknowledgment

As I complete my doctoral thesis, my heart is filled with gratitude. This is not only because I have accomplished a formidable academic task, but also because throughout this process, I have received support and assistance from many people without whom this would not have been possible.

In my doctoral studies, I would like to express my heartfelt gratitude to my mentors, Professor Enzo Bivona, Professor Qi Jiayin, Professor Xu Bo, and Professor Helen Yan, as well as the coordinators of the project, Professor Carmine Bianchi, Professor Chiara Garilli, Professor Salvatore Mancuso, Dr. Davide Villano, and Professor Santa Giuseppina Tumminelli. With the help of these professors, I successfully became a member of the project and achieved academic success. Especially Professor Enzo Bivona, Professor Qi Jiayin, and Professor Helen Yan, who not only provided valuable academic guidance throughout the research process but also offered me much advice and encouragement in life. Their patience, wisdom, and generosity have profoundly influenced me.

I would also like to thank my colleagues and classmates. During my time at the University of Palermo, the University of Bergen, and Radboud University in Nijmegen, we experienced countless discussions, debates, and collaborations. Special thanks to Mingming, Alfreogr, Passant, Thanh, Fatima, Rosi, Noemi, Giuliana, Nasrul, Eva, and many others. They are not only my academic partners but also my friends. We have shared many unforgettable moments together, which will become precious memories for me.

My deepest gratitude is reserved for my family. Thank you for your unconditional love and support, which has been the driving force that has allowed me to persevere. Special thanks to my parents, my boyfriend, his parents, and his grandparents, who have always been my strongest support, giving me the courage to remain steadfast in my beliefs and to overcome difficulties when faced with challenges.

This doctoral journey has not only been an academic challenge but also a process of personal growth. I have learned a great deal about research, writing, and thinking, but more importantly, I have learned how to face difficulties and challenges. From the application for the doctoral program to the courses and research, I have gone through many hardships. I regard the three-year doctoral program as a form of cultivation, where I not only learned about the field of System Dynamics but also re-examined myself and re-planned my life. It's as if I pressed both the fast-forward and pause buttons on my life at the same time, which was a magical process. I am grateful for my choices and the reflections of the past three years, which have shaped me differently from who I was three years ago.

Finally, I would like to thank everyone who has helped and supported me in my academic journey and life. Your encouragement and guidance have been key to my completion of this thesis. I look forward to using this experience as a solid foundation for my future career.

Contents

Abstract.....	8
1 Research Design.....	11
1.1 Introduction.....	11
1.2 Defining the Problem.....	14
1.3 Research Perspective and Theoretical Framework.....	16
1.4 Thesis Structure.....	17
2 Literature Review.....	20
2.1 Large Language Model	20
2.1.1 Large Language Model.....	20
2.1.2 Technological Development of LLMs.....	20
2.1.3 Application Fields of LLMs.....	21
2.1.4 Risks of Large Language Models.....	22
2.2 Large Language Model Governance	22
2.2.1 Governance of LLMs.....	22
2.2.2 LLMs Value Framework.....	25
2.2.3 Issues in Medical Field in LLMs Governance.....	29
2.3 Human Values Framework	32
3 Constructing Human Value Frameworks	36
3.1 Introduction.....	36
3.2 Level 1: Ideology.....	38
3.2.1 Political Ideology.....	38
3.2.2 Economic Mode	41
3.2.3 Social Structure.....	43
3.2.4 Religious Beliefs.....	46
3.2.5 Philosophical Thoughts.....	48
3.3 Level 2: Legal Regulations	51

3.3.1 Constitutions and Basic Laws.....	51
3.3.2 Civil and Commercial Law.....	54
3.3.3 Criminal Law.....	56
3.3.4 Litigation Law.....	58
3.3.5 Administrative Regulations and Policies.....	60
3.4 Level 3: Cultural Diversity.....	63
3.4.1 Language.....	63
3.4.2 Traditional Customs.....	65
3.4.3 Art.....	67
3.4.4 Literature.....	69
3.4.5 Cuisine.....	71
3.5 Level 4: Universal Moral Ethics.....	73
3.5.1 Peace.....	73
3.5.2 Development.....	74
3.5.3 Equity.....	74
3.5.4 Justice.....	75
3.5.5 Democracy.....	76
3.5.6 Freedom.....	76
3.6 Conclusion.....	77
4 LLMs Value Alignment Experiment.....	79
4.1 Introduction.....	79
4.2 Experimental Purpose and Objects.....	81
4.3 Dataset Selection.....	81
4.4 Experimental Process and Results.....	86
4.5 Experimental Conclusions.....	88
4.6 Current Value Attributes of the Three Major Models.....	89
4.6.1 Ideological Level.....	90

4.6.2 Legal and Regulatory Level	93
4.6.3 Cultural Diversity Level.....	96
4.7 Experimental Limitations.....	100
4.8 Conclusion	101
5 Enhancing Government's LLM Governance Level: A Dynamic Performance Governance Approach.....	103
5.1 Introduction.....	103
5.2 LLM Governance is a Wicked and Super Wicked Problem.....	104
5.2.1 Wicked and Super Wicked Problem.....	104
5.2.2 LLM Governance is a Wicked and Super Wicked Problem.....	104
5.3 Methodology	105
5.3.1 The Need for Dynamic Performance Management and Governance Approaches	105
5.3.3 Dynamic Performance Governance Approach	107
5.4 Applying DPG Framework to Analyze Government LLM Governance	108
5.4.1 A DPG Approach to Explore the Performance in LLM Governance.....	108
5.4.2 Stakeholders Analysis in LLM Governance	111
5.5 Conclusion	112
6 Case study	113
6.1 Introduction to the Case	114
6.1.1 Background.....	114
6.1.2 Data Source	117
6.1.3 Adaptability of the DPG Framework to the Governance of Large Language Models in the Medical Industry	118
6.2 Data and Methodology	120
6.2.1 Interview and Questionnaire	121
6.2.1.1 Design and Implementation of Semi-Structured Interviews	121
6.2.1.2 Encoding	124

6.2.2 DPG and CLD	125
6.2.3 Analytical Strategy	127
6.3 Case Analysis.....	127
6.3.1 Identification of Variables Vased on Interviews.....	127
6.3.2 Identification of Causal Relationships Based on Questionnaires	133
6.3.3 Dynamic Performance Governance Framework Construction.....	138
6.3.4 Loop Analysis and Identification of Governance Intervention Mechanisms	140
6.3.5 Stakeholder Analysis.....	151
6.4 Conclusion	153
6.4.1 The Verification and Extension of the DPG Theoretical Model by Case Study.....	153
6.4.2 Policy-Making Suggestions and Contributions	154
7 Discussion and Conclusion	156
7.1 Summary of Discussions and Findings.....	156
7.2 Contributions.....	156
7.3 Limitations and Future Research	160
7.3.1 Limitations.....	160
7.3.2 Future Research.....	161
Reference	163
Appendix A: Interview outline	190
Appendix B: Coding Nodes.....	198

Abstract

With the rapid development of artificial intelligence technology, large language models (LLMs) such as ChatGPT, ERNIE Bot, and Kimi have been increasingly applied in various fields including education, healthcare, law, and finance, significantly impacting social progress and technological innovation. However, the widespread application of LLMs has also raised concerns about their output results potentially being inconsistent with human expectations, especially when handling sensitive data that could potentially infringe on personal privacy. These issues not only affect the acceptability and reliability of LLMs but may also trigger broader social and ethical concerns. Therefore, ensuring that the development of LLM is consistent with human expectations.

This study aims to explore enhance the government's performance in large language model governance through dynamic performance governance (DPG) methods to achieve a balance between technological innovation and social safety. To construct a governance framework based on a human values framework and guide the development of existing large language models toward better alignment with human expectations. Explore the application of DPG methods in large language model governance and put forward specific suggestions for optimizing governance strategies. Take the medical industry as an example to verify the effectiveness of the DPG method in industry governance and put forward targeted improvement measures. This study aims to provide a reference standard for the government's governance of the large language model by constructing a human values framework and explore ways to improve the governance performance of the government's big language model through the method of DPG.

To achieve the above objectives, this study proposed four research questions: "How to help governments build a human values framework that LLM needs to align with?", "How to use the human values framework to evaluate the degree of alignment of LLM?", "How to enhance the government's LLM governance level through the DPG approach?" and "In the LLM governance of the medical industry, can the DPG framework provide effective governance strategy support for the government?". In response to RQ1, this study found that LLM should be consistent with the values and ethical standards of human society, including performance in ideology, legal norms, cultural diversity, and universal ethics. In response to RQ2, this study used an experimental method to evaluate the performance of LLM at different value levels and dimensions, providing empirical data for evaluating the value orientation of LLM. In RQ3, the study proposed a DPG-based conceptual framework that emphasizes the importance of government governance in LLM governance and explains how the DPG approach supports this process. Finally, in the exploration of RQ4, this study analyzes the help of DPG in large

language model governance in specific industries through case studies of Chinese medical industry scenarios.

This study first reviews the literature and comprehensively analyzes the technical development, application areas and governance risks of LLM. Subsequently, a multidimensional framework of human values that encompasses ideology, legal norms, cultural diversity, and universal moral ethics constructed to provide a theoretical basis for understanding and evaluating the value orientation of LLM. And through experimental evaluation of the value alignment level of popular LLM, reference standards and guidelines are provided for further improvement of LLM technology. Based on the constructed human values framework, this study tests the value alignment level of three existing popular LLMs. The experimental results show that although models such as ChatGPT 4.0, ERNIE Bot 3.5 and Kimi perform well in the field of universal ethics, there are still significant deviations in the ideological and legal norms. In addition, this study reveals the value judgment tendencies of LLMs in different cultural and social contexts and points out the improvement directions for future research in data diversity, model understanding, cultural diversity considerations, flexibility of experimental design, and depth of result interpretation. This study adopts the DPG method and proposes a conceptual framework based on DPG, attempting to help the government improve the level of LLM governance. Finally, the feasibility of the constructed DPG framework is verified using DeepSeek's application scenario case in the Chinese medical industry.

In summary, this study constructs a human values framework that can serve as a standard and guide for the governance of government LLMs. Using the constructed human values framework, a comprehensive analysis and experimental evaluation of the current LLMs for value alignment were conducted to show how the framework is used. By analyzing the characteristics of LLM governance, this study proposed a DPG framework to achieve sustainable development of LLM. This study provides a new perspective for understanding the complexity and dynamism of LLM governance, provides theoretical guidance and practical reference for policymakers and practitioners, and promotes the coordinated development of LLM technology and social values. By constructing a DPG framework, the government's LLM governance is explored, and the feasibility of the DPG method in the LLM governance of specific industries is verified through cases. The research results provide valuable inspiration for understanding and improving LLM governance, provide direction for future research and practice, ensure the harmonious unity of technological development and social values, and provide a scientific basis for decision-making that is more in line with human values.

Keywords: Large Language Models (LLMs), Large Language Models (LLMs) Governance, Dynamic Performance Governance (DPG), Experimental Evaluation on Value Alignment, Medical Field Application Case

1 Research Design

1.1 Introduction

Language has facilitated the continuity and development of human civilization, serving not only as a tool for communication and exchange but also as a key to cultural heritage, expression of ideas, and social cohesion. Through language, people can convey knowledge, emotions, and thoughts, promoting the development of individuals and society. As large language models (LLMs) become more and more popular, they have had a huge impact on the public's learning, work and life (Andrew Caines¹, 2023). Many companies and research institutions have seen the potential of LLMs and invested a lot of resources in their research and development, which has led to the continuous expansion of the scale and enhancement of the capabilities of LLMs, and the emergence of scaling laws and capability emergence (Wei et al., 2022). However, at the same time, the public is also worried that LLMs will bring corresponding risks and problems (Gou & Nguyen, 2024; Varshney et al., 2020), for example, the risk emergence problem that appears at the same time as the emergence problem, that is, as the scale of LLMs increases, risks that have never appeared in small models or more serious problems will arise (Y. Yao et al., 2023; Yi & Xie, 2023); there is also the anti-scaling phenomenon, that is, the risk does not decrease with more data, but increases sharply (Achiam et al., 2023). Since the values of LLMs come from stakeholders such as developers and data providers, and the generation of LLM response content is still a "black box", LLMs have unknown risks and uncontrollable situations (Y. Yao et al., 2023), for example, they may output wrong or harmful information, etc. When LLMs are used in educational technology, there are many underage students among the users, who are still in the stage of physical and mental growth. To protect their physical and mental health, it is urgent to align with human values to avoid such risks as much as possible (Floridi et al., 2018). Since the use of LLMs involves many stakeholders, to better solve the problem of LLM value alignment, it is necessary to integrate the views of various disciplines and consider the rights and interests of stakeholders (Schiff, 2022). Therefore, the collaborative governance method can be used to solve the problem of educational LLM governance (Kasneci et al., 2023). In short, governance balances the application security and development of LLMs (Tomašev et al., 2020), and the governance of LLMs is one of the most important ways for the sustainable development of LLMs (Umbrello & Van de Poel, 2021). Therefore, how to govern LLMs and achieve the sustainable development of LLMs and even artificial intelligence technology has become an urgent problem to be solved.

The rapid development and widespread application of artificial intelligence technology has brought great opportunities to society, but it has also triggered a wide range of ethical, legal and

social risks. As artificial intelligence moves from single tasks to general intelligence, its complexity and influence have increased significantly, especially in the field of generative artificial intelligence and large models. The widespread application of technology may exacerbate privacy violations, algorithmic bias, social inequality and threats to human autonomy. In the face of these challenges, the role of the government in artificial intelligence governance has become increasingly prominent, becoming a key force to ensure that technological development conforms to social values and ethical norms.

At present, although technology companies around the world have made many efforts in artificial intelligence ethics and put forward ethical principles such as fairness, transparency, and privacy protection, these principles often lack legal binding force and are controversial in actual implementation. This shows that it is difficult to effectively solve the complex problems brought about by artificial intelligence technology by relying solely on corporate self-regulation, and government intervention and governance are particularly important.

Search for data on "AI governance" and "LLM governance" in social science-related fields on the web of science, with search keywords including "AI governance", "LLM governance" and their extended topics such as "AI governance", "AI strategy" or "AI strategy". The search results totaled 2,081, reflecting the wide coverage of these topics in academic research. As shown in Figure 1.1 below, the changing trend of the number of literature and citations on AI governance and LLM governance from 1978 to 2025. The number of literature and citations increased significantly after 2020, indicating that the research enthusiasm in this field has risen rapidly in recent years. This trend reflects the increasing attention paid to AI technology and governance issues around the world.

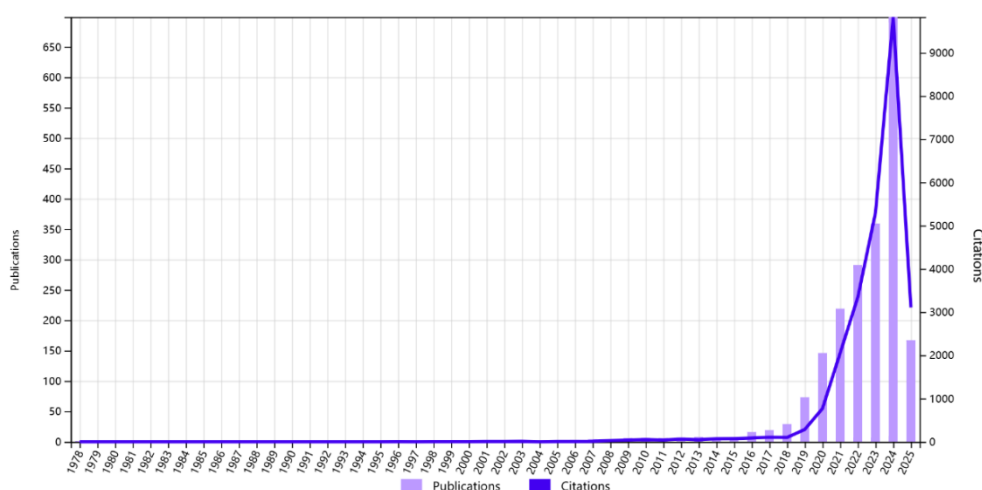


Figure 1.1 Number of papers and citations in different years

The bar chart in Figure 1.2 shows the distribution of the number of papers on AI

governance and LLM governance in different disciplines. The number of papers in the field of computer science is the largest, with more than 900 papers, indicating the importance of this field in AI governance research. The next largest number of papers are business economics, government law, public administration, social issues, etc., showing the interdisciplinary nature of AI governance research.

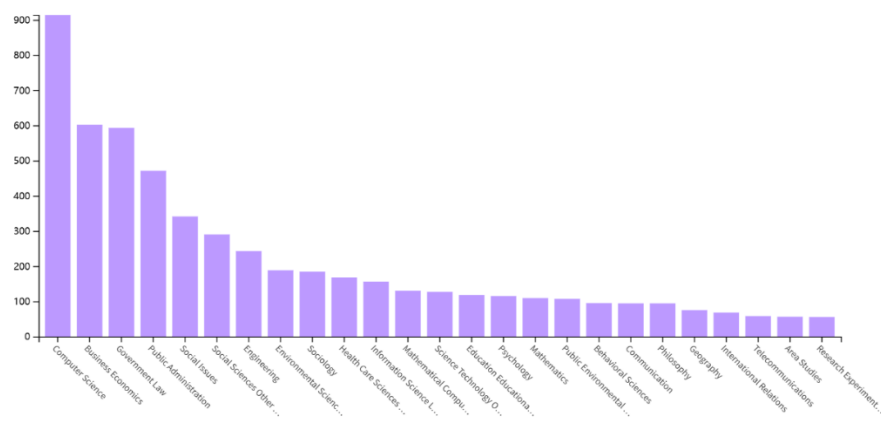


Figure 1.2 Distribution of literature by different research areas

The bar chart in Figure 1.3 shows the distribution of the number of papers on AI governance and LLM governance in different countries. The United States has the largest number of papers, more than 400, showing its leading position in research in this field. Countries such as England, China, Germany, and the United Kingdom also contributed many papers, reflecting the global attention to AI governance.

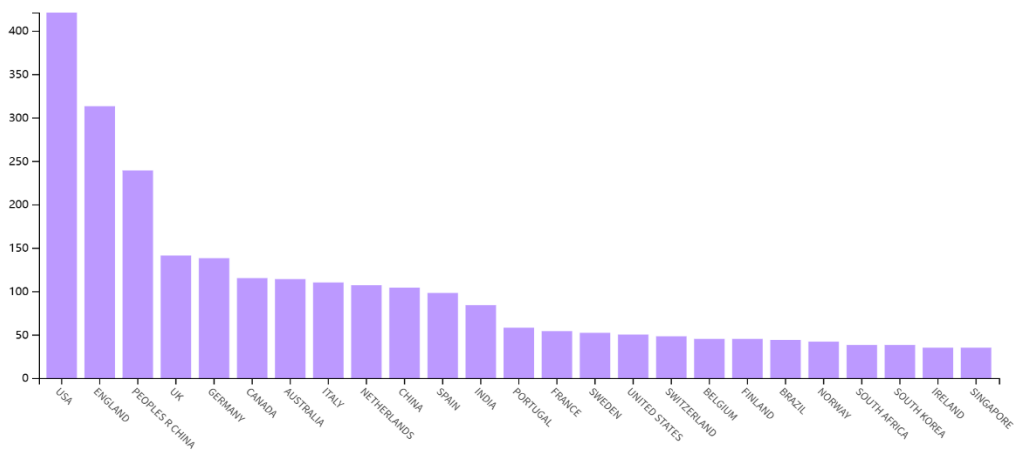


Figure 1.3 Distribution of literature by country

The role of the government in AI governance is not only reflected in the formulation and improvement of relevant laws and regulations, but also in promoting technical standardization,

clarifying the allocation of responsibilities, and promoting multi-party collaboration. By embedding ethical principles into the legal framework, the government can ensure that the design, development, and application of AI technology conform to social norms and values.

In summary, the government plays an indispensable role in AI governance. By formulating a clear legal framework, promoting international cooperation, and promoting multi-party collaboration, the government can effectively respond to the ethical and social challenges brought about by AI technology and ensure that technological development truly serves human welfare and social progress.

1.2 Defining the Problem

In today's rapidly advancing artificial intelligence technology, LLMs (LLMs) such as ChatGPT, ERNIE Bot, and Kimi have become key drivers of social progress and technological innovation. These models, through deep learning and natural language processing techniques, play a significant role not only in various fields such as education, healthcare, law, and finance but also increasingly demonstrate their potential in improving decision-making efficiency and optimizing service experiences. However, with the widespread application of LLMs, concerns have arisen about the potential inconsistency between the output results of these models and human values. For instance, LLMs might generate content with biases or potentially infringe on personal privacy when handling sensitive data. These issues not only affect the acceptability and reliability of LLMs but could also trigger broader social and ethical concerns.

Against this backdrop, ensuring the alignment of LLMs' values with human values has become particularly important. Governance involves not only the transparency and interpretability of the models but also the key to their sustainable development. Therefore, this study aims to evaluate and analyze the performance of LLMs in governance and explore methods to enhance these models' governance performance through collaborative governance and DPM . By employing this approach, we can better understand and improve the decision-making processes of LLMs, ensuring that their output in various application scenarios is more in line with human values and ethical standards.

The rapid development of LLM technology is profoundly changing all areas of society, but the governance of LLMs faces multiple contradictions and challenges. On the one hand, the widespread application of LLMs in the fields of medical care, education, transportation, etc. has promoted the improvement of productivity and innovation capabilities; on the other hand, its rapid development has also caused security issues such as data privacy leakage, doubts about content authenticity, and algorithmic bias. These problems not only threaten the public interest but may also weaken society's trust in technology and hinder the sustainable development of technology.

At present, the core contradiction of LLM governance lies in the speed of technology iteration and the lag of the governance system. In this context, this study explores how the government can build a LLM governance framework through a sound governance framework, dynamic feedback-driven supervision and collaborative decision-making, actively play the role of the government in the governance of LLMs, balance the development and security of LLM technology, and ensure its reliability and the realization of public interests.

This study is of great significance in filling the gaps in existing research and promoting innovation in the theory and methods of LLMs governance. By deeply analyzing the shortcomings of existing research and proposing future research directions, this study helps to promote the healthy development of LLMs in the medical field, ensuring that they enhance the quality of medical services while also safeguarding the rights and interests of patients and social fairness and justice.

(1) Research Problems

- Existing LLMs have many undesirable problems
- There is currently no perfect LLMs governance solution
- The above risks threaten the sustainable development of LLMs

(2) Research Goals

- To construct a governance framework based on a human values framework and guide the development of existing LLMs toward better alignment with human expectations.
- Explore the application of dynamic performance governance (DPG) methods in LLM governance and put forward specific suggestions for optimizing governance strategies.
- Take the medical industry as an example to verify the effectiveness of the DPG method in industry governance and put forward targeted improvement measures.

(3) Research Questions

From an academic perspective, scholars generally pay attention to the current status and existing problems of the development of Master of Laws, but few scholars evaluate the role of value consistency in relevant large-scale language models and the impact of environmental, social and economic factors on governance itself. There is also no research that brings stakeholders together to discuss and solve large-scale model governance issues. To fill this research gap, this study summarizes the preliminary LLMs governance performance governance framework model proposed in the literature, and convenes stakeholders in the medical field for further teamwork, constructs a governance framework that affects LLMs governance performance, and then uses a DPG method to comprehensively analyze the

framework, that is, a performance management method based on continuous improvement and real-time feedback(Lane et al., 2010). Finally, the rationality of the model is analyzed using existing medical field cases. This study uses DPG to evaluate the impact of governance on the development of LLMs and proposes a more feasible LLM joint governance framework based on strategic performance evaluation results and cases to promote the sustainable development of LLMs.

This study will attempt to answer the following questions:

RQ1: How to help governments build a human values framework that LLM needs to align with?

RQ2: How to use the human values framework to evaluate the degree of alignment of LLM?

RQ3: Can the DPG enhance the LLM governance level?

RQ4: In the LLM governance of the medical industry, can the DPG framework provide effective governance strategy support for the government?

By answering the above questions, this study will comprehensively evaluate the status of LLMs governance in the medical field and provide suggestions for building a more effective LLMs governance framework to achieve sustainable development of LLMs more quickly.

1.3 Research Perspective and Theoretical Framework

1. Dynamic Performance Governance : Building Responsive and Adaptive Mechanisms

Traditional regulatory approaches are static, rule-based, often reactive, and time-lagged, and are insufficient to cope with the complexity and unpredictability of AI systems. The DPG approach, which draws from public administration and system governance theory, offers a flexible alternative. It contains three core mechanisms: dynamic policy goal adjustment, multi-source performance monitoring, and feedback-based policy revision, which enable policy systems to learn, adapt, and respond in real time, making DPG particularly suitable for managing rapidly evolving technologies like LLM.

2. Causal Loop Diagram (CLD): visualize the dynamic relationship between governance variables

CLD is a tool for describing the causal relationship between variables in a system. It is usually used in system dynamics and system thinking(Crielaard et al., 2024)to help analyze the complex behavior patterns of the government in the LLM system. The diagram reveals the "reinforcing feedback" and "balanced feedback" mechanisms. Through modeling from this system perspective, the theoretical logic of DPG is concretized, and it also provides structural

analysis support for subsequent empirical research in the medical industry.

3. Experimental Method

Experiments are designed to assess the performance of different LLMs at various value levels and dimensions. The experimental method allows us to test and compare the performance of different models in a controlled environment, thereby better understanding their capabilities in handling complex value judgments. Through experiments, we can collect data on model behavior, analyze their performance in specific contexts, and propose improvement suggestions based on these analyses.

4. Case Study

The study utilizes the case study method, combining theoretical models with real-world phenomena, to better understand the models and propose feasible recommendations. The case study method enables us to conduct an in-depth analysis of the application of LLMs in specific contexts, discussing their governance issues in practical applications. By analyzing both successful and failed cases, we can extract the best practices and provide concrete guidance and suggestions for the governance of LLMs.

In general, these four perspectives do not exist in isolation, but complement each other within a coherent governance framework. The human value orientation defines the normative goals of LLM governance; government leadership provides the institutional basis for achieving these goals; and the DPG approach puts these goals into practice through adaptive policy tools. Together, they constitute a dynamic, multi-level governance model that is applicable to complex and evolving socio-technical systems.

Unlike previous studies that focus on AI ethical principles or static regulatory tools in isolation, this study emphasizes the integration of normative values, institutional design, and system-level feedback. It contributes to the literature by linking human-centered ideas to specific policy mechanisms and evaluating their implementation in real-world governance scenarios.

The theoretical framework of this study directly guides the research design. The human values framework guides the construction of the government LLM governance model; the DPG theory is used to construct a causal feedback model of governance performance; and the healthcare case study shows how these mechanisms operate in high-risk areas under government leadership.

1.4 Thesis Structure

This study aims to explore the application of LLMs (LLMs) and the governance issues

they face. Chapter 1 introduces the development of LLM technology and potential risks, including new risks that emerge as the model scales up and the risks that increase with the amount of data. This study emphasizes the importance of governance for LLMs and proposes a collaborative research approach and DPG to address these governance issues.

The literature review in the second chapter delves into the technical development, application fields, and governance risks of LLMs. It analyzes the application of LLMs in various fields such as education, law, and finance, and discusses the challenges faced in terms of data security, privacy protection, and ethical issues. Additionally, this chapter assesses the current state of LLMs governance research, providing a theoretical foundation for subsequent studies.

The third chapter constructs a theoretical framework that analyzes the hierarchical structure of human value frameworks, including ideology, legal norms, cultural diversity, and universal ethical values. The framework explores the values within different cultural and social contexts and how these values influence the decision-making processes of individuals and groups. This theoretical framework offers a multi-dimensional perspective for understanding and evaluating the governance of LLMs.

The fourth chapter uses experimental methods to assess the performance of three popular LLMs—ChatGPT 4.0, ERNIE Bot 3.5, and Kimi—across different value hierarchies and dimensions. The experiment aims to understand the current state of these models in terms of value alignment, including the purpose, objects, dataset selection and processing, as well as the experimental process and results. These experimental results provide empirical data for the governance of LLMs.

Chapter 5 proposes a conceptual framework based on the DPG approach, which aims to improve the government's LLM governance performance through collaborative governance. The framework explores the role of stakeholders in LLM governance and explains how the DPG approach supports this process. This chapter provides an innovative management approach to achieve LLM governance.

Chapter 6 verifies the DPG framework constructed in Chapter 5 through the application scenario cases of DeepSeek in China's medical industry and deeply explores the governance issues of LLM in the medical field. Based on the general DPG framework in Chapter 5, a DPG framework for LLM performance in the medical industry is proposed. The research results not only provide empirical support for the governance of LLM but also provide valuable guidance and reference for future research and practice.

Chapter 7, as the conclusion of the paper, summarizes the content of each section, concludes the main contributions of the paper, analyzes the limitations of the paper, and finally

proposes suggestions for future research.

This study comprehensively explores the application of LLMs (LLMs) and their governance issues, analyzes the technical development, application areas and governance risks of LLMs, and evaluates their performance at different value levels. The study emphasizes the importance of LLM governance and proposes a conceptual governance framework of DPG to address these challenges. Through experimental evaluation of models such as ChatGPT 4.0, ERNIE Bot 3.5 and Kimi, the results show that although these models perform well in the field of universal ethics, there are still significant deviations at the ideological and legal normative levels. The study also explores the value judgment tendencies of models in different cultural and social contexts and points out that future research needs to improve in terms of data diversity, model understanding, cultural diversity considerations, flexibility of experimental design, and depth of result interpretation. In summary, this study constructs a human values framework that can serve as a standard and guide for the governance of government LLMs. Using the constructed human values framework, a comprehensive analysis and experimental evaluation of the current LLMs for value alignment were conducted to show how the framework is used. By analyzing the characteristics of LLM governance, this study proposes a DPG framework to achieve sustainable development of LLM. practitioners, and promotes the coordinated development of LLM technology and social values. By constructing a dynamic governance performance framework (DPG), the government's LLM governance is explored, and the feasibility of the DPG method in the LLM governance of specific industries is verified through cases. The research results provide valuable inspiration for understanding and improving LLM governance, provide direction for future research and practice, ensure the harmonious unity of technological development and social values, and provide a scientific basis for decision-making that is more in line with humanistic values.

2 Literature Review

2.1 Large Language Model

2.1.1 Large Language Model

Large language models (LLM) are a category of artificial intelligence models based on deep learning, characterized by a vast number of parameters and complex structures (S. Q. Du et al., 2023; H. Huang et al., 2023; Sun et al., 2024). They understand and generate natural language by training on a large corpus of text data. These models typically have billions or even trillions of parameters, capable of performing various complex language tasks such as text summarization, translation, question answering, and sentiment analysis (Nasseri et al., 2023; Roumeliotis et al., 2024a; Sharma et al., 2024; Venerito & Iannone, 2024). The core of LLMs lies in their pre-training capability, which involves learning a general representation of language from a large amount of unlabeled text data, and then adapting to specific downstream tasks through fine-tuning (Goswami et al., 2024; Lai et al., 2024; Roumeliotis et al., 2024b; Woisetschlaeger et al., 2024; Xu et al., 2021).

2.1.2 Technological Development of LLMs

The development of LLMs has gone through three stages: The first stage is the germination period, from 1950 to 2005, represented by CNN; The second stage is the sedimentation period, from 2006 to 2019, represented by the Transformer architecture; The third stage is the explosive period, from 2020 to the present. During the explosive period, famous models such as GPT-3, ChatGPT, and GPT-4 have emerged (Y. Liu et al., 2023). The GPT series is developed by OpenAI, adopting the paradigm of generative pre-training and discriminative task fine-tuning (Minaee et al., 2024). The LLaMa series has expanded the Chinese vocabulary and optimized in the Chinese version (Minaee et al., 2024). The GLM series is a bilingual pre-trained language LLMs jointly developed by Tsinghua University and ZhiPu AI (Y. Zhang et al., 2024). In addition, there are other well-known LLMs such as Falcon, Vicuna, PanGu, ERNIE BOT, and Spark Cognitive LLMs, which have shown strong processing and representation capabilities in different fields and applications.

In the early stages of development, the initial LLMs, such as BERT and GPT-2, demonstrated their potential in natural language understanding (NLU) and natural language generation (NLG) tasks through pre-training on large-scale text data (Liu et al., 2020; Smith et al., 2022). With the increase in computational resources, the scale of LLMs has also been expanding, such as GPT-3 and LaMDA. These models further enhance their language processing capabilities with more parameters and more complex network structures (Thoppilan et al., 2022; y Arcas, 2022). Further developments have also seen LLMs integrating information

from vision and other modalities, such as improving comprehension through joint learning of images and text(Raiaan et al., 2024). LLMs have also shown their capabilities in zero-shot or few-shot learning, that is, performing tasks by understanding task descriptions without or with only a small amount of labeled data(Abbasiantaeb et al., 2024; Ahmed et al., 2022) .

2.1.3 Application Fields of LLMs

LLMs have shown great potential in various fields such as healthcare, education, law, and finance. The application of LLMs in different fields has addressed some of the issues faced by these fields. Technological advancements have enabled general-purpose LLMs to perform tasks ranging from simple language translation and question answering to more complex tasks such as programming and text generation (Hadi et al., 2023).

In the field of education, LLMs are used by teachers to assist in teaching, such as generating course content and helping students understand complex concepts through natural language processing technology(Alqahtani et al., 2023; Antoniou-Kritikou et al., 2024; Jeon & Lee, 2023). For students, LLMs can sometimes replace the presence of a teacher, tailoring personalized learning plans according to the students' preferences and providing a more diversified learning experience (Dai et al., 2023).

LLMs are also used in the legal field for legal text analysis and case prediction (Bui et al., 2024), improving the accuracy and relevance of judgments by accessing case law and legal statutes (Savelka et al., 2023). For example, the DISC-LawLLM system released by the Data Intelligence and Social Computing Laboratory of Fudan University (FudanDISC) has demonstrated its excellent capabilities in legal text processing, legal reasoning, and legal knowledge retrieval through training with high-quality supervised fine-tuning dataset DISC-Law-SFT (Y. Liu et al., 2024; Yue et al., 2023). In addition, the ChatLaw model released by the Peking University team, as an open-source legal LLM, has shown its capabilities in legal text understanding and task processing through a carefully designed legal domain fine-tuning dataset, and has introduced self-attention methods to enhance the model's ability to handle erroneous information in reference data (Cui et al., 2023).

The application of LLMs (LLMs) in the financial sector is gradually showing its potential and value. For instance, the BloombergGPT model has demonstrated its excellent capabilities in financial text processing, investment decision support, risk management, customer service, and macroeconomic analysis by training on a large dataset of financial data. Moreover, the training dataset FINPILE of the BloombergGPT model includes news, documents, press releases, web-crawled financial documents, and social media messages, ensuring the diversity and quality of the data. These applications showcase the broad application prospects of LLMs in the financial sector and point out the challenges that need to be addressed in practical

applications. With continuous research and optimization, LLMs are expected to play a greater role in the financial sector(Wu et al., 2023).

2.1.4 Risks of Large Language Models

As the capabilities of LLMs increase rapidly, their applications are becoming more and more extensive (Wamba et al., 2021) . However, the values of LLMs come from the settings of the designers, data providers, and their own uncontrollable "black box" operations, etc., accompanied by problems such as inheritance of harmful information (Chang et al., 2023; Shen et al., 2023). The use of LLMs will also have some impact on human values (Muhlenbach, 2020). Therefore, to achieve sustainable development of LLMs, rather than being abandoned due to security issues, the values of LLMs need to be consistent with human values, that is, to align the values of LLMs (Z. Zhang et al., 2024). Firstly, there are issues of accuracy and reliability. Some studies have shown that LLMs may generate incorrect information when producing text, which could have serious consequences in high-risk fields such as healthcare and law (Abi-Rafeh et al., 2023; Hurley et al., 2024). Secondly, there is the issue of bias and discrimination. LLMs may exhibit biases in terms of gender, race, etc., when generating text, which needs to be addressed through algorithm optimization and data cleansing(Lahat et al., 2023). In addition, the interpretability and transparency of LLMs is also an important issue, as users often find it difficult to understand the decision-making process of the models (F. Liu et al., 2023).

2.2 Large Language Model Governance

2.2.1 Governance of LLMs

The application of LLMs has profound implications for society. On one hand, LLMs can improve production efficiency and promote social progress; on the other hand, the misuse of LLMs may lead to social issues, such as over-reliance and changes in educational methods(Alqahtani et al., 2023).

1. Data Security

The training and application of LLMs require the processing and analysis of large amounts of data, which may contain sensitive information(Ma et al., 2024). The main issues of data security are the risks of data leakage and misuse(Ma et al., 2024; L. Zhang et al., 2023). For example, the datasets required for the training of LLMs may contain personal privacy information, and if handled improperly, it could lead to privacy leaks(Ma et al., 2024). In addition, LLMs may inadvertently leak sensitive information from the training data when generating content (L. Zhang et al., 2023).

2. Privacy Protection

Privacy protection is an important aspect of the governance of LLMs. During the training and application of LLMs, a large amount of personal data may be processed. If these data are not properly handled, they may infringe on personal privacy(Y. Chen et al., 2023; Dugar & Aresh, 2023). For example, the training process of LLMs may use publicly available data from the internet, which may contain personal privacy information. To protect privacy, strict privacy protection measures need to be taken during data cleansing, model training, and application(Ma et al., 2024; Rehman et al., 2023).

3. Ethical and Moral Issues

The ethical and moral issues of LLMs are mainly reflected in the credibility and fairness of their generated content. LLMs may generate false or misleading information, which not only affects user decisions but may also have a negative impact on society(Scherrer et al., 2023). Moreover, LLMs may exhibit bias and discrimination during training and application, which needs to be controlled through ethical review and governance(Chen et al., 2024; J. K. Kim et al., 2023; Lund et al., 2023; Miao et al., 2024).

4. Robustness

The robustness of LLMs is also an important governance challenge. LLMs may produce incorrect predictions or outputs when facing malicious attacks or input disturbances, which affects their reliability in practical applications(Giudice et al., 2023; Paroiu et al., 2023). Therefore, enhancing the robustness of LLMs to ensure their stable operation under various environments is an important aspect of governance (Gill et al., 2024; Munyer et al., 2024; S. Zhu et al., 2024).

5. Laws and Regulations

With the development of LLMs technology, relevant laws and regulations are also being continuously improved. For example, regions such as the European Union and the United States are advancing the legislative work on the governance of artificial intelligence to ensure that the application of LLMs complies with ethical and legal requirements(Bharadwaj & Parker, 2023). The formulation and implementation of laws and regulations can provide clear guidance and constraints for the governance of LLMs (Ni & Buehler, 2024). Formulating relevant ethical and legal norms to ensure that the application of LLMs complies with social morals and legal regulations (Almeida et al., 2024; S. Liu et al., 2024).

In general, the governance of LLMs is not only a national issue but also a global issue. Different countries and regions need to strengthen cooperation to jointly formulate and implement rules and standards for the governance of LLMs(Eppler et al., 2024). Through international cooperation, the healthy development of LLMs technology can be promoted while

protecting the rights and interests of users worldwide. The governance of LLMs is a complex and multidimensional issue that needs to be comprehensively considered and resolved from technical, legal, and ethical perspectives(Gorski & Ramakrishna, 2023). Effective governance can maximize the advantages of LLMs while minimizing their potential risks and hazards(Shumway & Hartman, 2024; Woisetschlaeger et al., 2024).

There are some notable shortcomings in the current research on the governance and governance of LLMs. Scholars pay attention to the connotation, methods and significance of the alignment of values in LLMs, actively explore better governance methods and evaluate the value of existing LLMs. Governance is a comprehensive issue, so it requires interdisciplinary integration and multi-departmental cooperation (Brendel et al., 2021; Eitel-Porter, 2021; Heyder et al., 2023; Jobin et al., 2019; Siau & Wang, 2020), and most of the research focuses on how to improve the alignment method of values in LLMs. There is a lack of systematic discussion on the value theory in the LLM and a comprehensive evaluation of the implementation effect of the existing LLMs value framework.

The shortcomings in current research on the governance of LLMs limit their effective application in the medical field. Here are some key deficiencies:

(1) Excessive Model Complexity: Current LLMs often have complex structures and numerous parameters, making them difficult to understand and control in practical applications. This complexity not only increases the difficulty of model training and optimization but also reduces the interpretability and transparency of the models, making it hard to meet the strict requirements for the decision-making process in the medical field(Feder et al., 2023).

(2) Inadequate Interdisciplinary Collaboration Mechanisms: Although teamwork theoretically has significant advantages, interdisciplinary collaboration mechanisms are still imperfect in practical applications. There are barriers to communication and collaboration among experts from different fields, preventing the full potential of teamwork from being realized. The lack of such collaboration mechanisms restricts the application and governance of LLMs in the medical field(Bao et al., 2023).

(3) Insufficient Governance Strategies and Technical Means: Existing governance strategies and technical means have not fully adapted to the characteristics and needs of LLMs. For example, current ethical review and risk assessment methods may not comprehensively cover the complexity and uncertainty of LLMs, leading to poor governance outcomes(A. H. Huang et al., 2023; Piktus et al., 2023).

(4) Need for Improved Governance Performance: The performance of LLMs in governance still needs to be enhanced. In current research, the ability of LLMs to understand

and adhere to human values, preferences, and ethical principles is still lacking, which is particularly important in the medical field where medical decisions need to highly conform to ethics and values (Bernstein et al., 2023; Q. Chen et al., 2023; Sun et al., 2024; Tanwar et al., 2023).

2.2.2 LLMs Value Framework

With the development of generative AI technology, LLMs and other technologies are widely applied in learning, work, and life. At the same time, the application of these technologies also brings corresponding problems and risks (Gou & Nguyen, 2024; Varshney et al., 2020). Therefore, developers need to make the application of generative AI systems more aligned with human values, that is, to align human values with generative AI systems. By aligning the values of generative AI technology, the content of applications related to it can be made to conform to human values. The value alignment of generative AI systems balances the safety of technology application with the development of technology (Tomašev et al., 2020). It is one of the most important ways for the sustainable development of generative AI technology (Umbrello & Van de Poel, 2021). The value alignment of generative AI systems is not only a need for technological development but also an inevitable requirement of social ethics and legal norms (Chun & Elkins, 2024; Pitkäranta & Pitkäranta, 2024). Therefore, how to govern LLM to achieve sustainable development of AI technology and even human society has become an urgent problem to be solved.

Existing research on LLMs governance primarily addresses the questions of what values to align with and how to achieve alignment (Huang et al., 2024; Sungdong Kim et al., 2023; Klingefjord et al., 2024). There is a consensus that the values of generative AI systems should not be determined by developers alone but should be aligned with the values of a broader public. Moreover, for more complex operations of generative AI systems, their values need to be consistent with human values. (Hadar-Shoval et al., 2024; J. Yao et al., 2023). In the research on the value framework of generative AI systems, due to the involvement of various fields such as philosophy, psychology, law, and cognitive science, the construction of a comprehensive value framework is currently very limited. However, major companies are actively attempting to develop ethical guidelines and frameworks consistent with human values to ensure the sustainable development of technology, as shown in Table 2.2.

Anthropic's team, developing an AI aligned with human values, used prompting, preference pre-training, and studies to introduce the 3H (Helpful, Honest, Harmless) principles.

Simple prompts are a baseline for alignment, improving with model scale. Preference models outperform imitation learning, especially at scale. PMP pre-training boosts sample efficiency on human preference data. The methods enhance AI alignment, supported by experiments, but long-term applicability and generalizability need further validation. The 3H principles, though simple, risk ambiguity and subjectivity(Askell et al., 2021).

The relatively comprehensive study is Microsoft Research Asia's Value Compass project, which addresses the challenge of aligning AI LLMs with human values and proposes a set of interdisciplinary theory-based solutions (Xinpeng Wang et al., 2024).

Microsoft's Value Compass project investigates aligning large AI models with human values, using Schwartz's ten value dimensions. It surveys alignment goals at three levels: instructions, preferences, and values, and emphasizes intrinsic human values. The project offers resources for alignment and platforms for assessment. It provides an overview of alignment goals, underscoring the importance of value alignment and suggesting future directions. However, it is highly theoretical, with limited practical application and case study depth, and lacks specific implementation strategies and empirical validation of assessment methods(J. Yao et al., 2023).

Google's Perspective API leverages machine learning to mitigate online toxicity, ensuring healthier dialogues and giving platforms and developers more control. It scores potentially harmful comments to enhance online communication quality. The API reduces moderation burdens, enabling automated processing while keeping human oversight. It also encourages healthier discussions by providing feedback to commenters. However, it faces challenges due to language complexity, necessitating ongoing updates for adapting to language and behavior changes, and risks overlooking context-specific factors due to over-reliance on automation(Lees et al., 2022)。

OpenAI's Moral Graph project, using Democratic Fine-Tuning, aims to create a values map for AI alignment, reflecting a broad consensus. It involves discussions with a diverse sample of participants on contentious issues and uses chatbots to elicit values, forming a graph that determines value relationships through voting. This approach is seen as more effective than constitutions or simple rules, offering advantages in safety, scalability, and moral depth. The project is inclusive, promoting understanding and respect for diverse values, and provides a new method for AI alignment that better reflects human morals(Klingefjord et al., 2024). However, avoiding disagreements by forcefully setting a direction when mapping the graph

may limit its completeness.

Alibaba's 100PoisonMpts evaluates Chinese AI models for value alignment, ensuring safety and reliability. It uses expert annotations to build an open-source dataset, assessing models and exploring expert principle-based alignment. The project improves model performance in safety and value alignment and enhances handling of sensitive issues. Despite its contributions, the reliance on limited expert annotation may restrict model optimization, and the model's capabilities in specialized domains require further enhancement. Self-alignment methods need more research for better efficiency(Sun et al., 2023).

Tencent's AI safety solution guards' data and content security in generative AI, from training to operations. It offers a framework for safety, ethics, and value alignment, with measures like a Prompt detection platform and infrastructure security. The solution includes content and data safety measures, such as review services and access control, to ensure AI safety. It calls for ongoing exploration by stakeholders and continuous updates to address new threats. While comprehensive, the solution requires more real-world case studies and updates to maintain its effectiveness.

Table 2.2 Existing LLM Value Framework

Value Framework	Method	Ethical Framework	Application Scenarios	Evaluation
Anthropic-3H	Prompting, preference model pre-training, comparative studies	Helpful, Honest, and Harmless	Assessment systems, model training, improving interaction experience, risk reduction, enhancing transparency and credibility...	Pros: Improves model alignment; some empirical support; simple principles. Cons: Needs more evidence for long-term applicability; may be too abstract and subjective in application; overly data-dependent
			alignment goal definition, model training, alignment assessment, risk management, improving human-computer	Pros: Comprehensive overview of alignment goals; provides guidance for value alignment; offers resources and tools for LLMs alignment. Cons: Overly theoretical; lacks specific implementation strategies; lacks strong evidence to validate method effectiveness
Microsoft-Value Compass	BaseAlign algorithm	Self-direction, Stimulation, Hedonism, Achievement, Power, Security, Conformity, Tradition,		

Google-Perspective API	Machine learning	Benevolence, Universalism	interaction experience... Multilingual review, improving review efficiency, model robustness testing, model fairness assessment, continuous performance improvement, maintenance, and service simplification...	Pros: Eases content moderation burden; provides feedback mechanism for healthy dialogue. Cons: Can't fully capture all forms of toxicity or inappropriate comments; high cost of continuous iteration; over-reliance on technology may overlook context-specific factors
		Toxicity reduction, severe toxicity, identity attacks, threats, refusal to answer, profanity, vulgarity, and insult	Value refinement and coordination, model guidance, handling complex situations, model performance assessment...	
OpenAI-Moral Graph	Moral Graph elicitation (MGE)	Safety, oversight, scalability, explainability, moral depth, and robustness	Model security assessment, multilingual security management, identification and prevention of malicious behavior... Model training, false information prevention, ethical risk management, data security governance, risk modeling, and secure operations...	Pros: Research is inclusive, strengthens respect for disagreements; offers a new AI alignment method. Cons: Avoiding disagreements in graph mapping may limit its completeness Pros: Enhances model processing and generalization capabilities; open-source dataset promotes joint research and progress. Cons: Limited expert annotation data restricts model iteration optimization; model understanding and application in professional fields need improvement; efficiency and effectiveness need enhancement
Alibaba-100PoisonMpts	"Poisoning" and "interpretation"	Safety and responsibility		
Tencent-Generative AI Safety Solution	Risk scenario definition, corpus services, copyright protection, and value detection	Ontological security, application security, and content security		Pros: Provides comprehensive safety protection measures covering the entire lifecycle of generative AI; effectively identifies and manages sensitive data and content risks. Cons: Lacks sufficient practical cases to demonstrate application effects; safety solutions require continuous iteration

2.2.3 Issues in Medical Field in LLMs Governance

In the medical field, LLMs are used to assist in clinical decision-making and medical text analysis, providing medical advice to doctors based on patient databases, literature, etc., and enhancing the efficiency of doctors' decision-making (Arighi et al., 2024; Benary et al., 2023; T. C. Chen et al., 2023; Chiesa-Estomba et al., 2024; Jiang et al., 2024; Rao et al., 2023). LLMs can also interact with patients through chat programs, providing answers to basic medical questions, arranging appointment times, and other services (Abbasian et al., 2023; Bill & Eriksson, 2023; Montagna et al., 2023). In medical research, like research in other fields, LLMs predict future research trends in the medical field by identifying and analyzing information in existing literature (Chervenak et al., 2023; Y. Du et al., 2023; Lemley, 2024; Pal et al., 2024), and even provide decision support and risk prediction in the field of public health (De Angelis et al., 2023; Deiana et al., 2023; Rane et al., 2023). Studies have shown that LLMs perform well in answering medical exam questions, such as some LLMs having a correct rate close to or exceeding that of human doctors in the European Urology Board's on-the-job assessment (Igarashi et al., 2024; Kollitsch et al., 2024).

The application of LLMs in the medical field provides innovative solutions for disease diagnosis, treatment plan recommendations, patient health management, and other issues (Shah et al., 2023). By analyzing a large amount of medical data, LLMs can assist doctors in making more accurate disease diagnoses, such as identifying tumors and other abnormalities in medical imaging analysis (Hsueh et al., 2024). In addition, they can also recommend personalized treatment plans based on the specific conditions of patients, improving the targeting and effectiveness of treatment. In terms of patient health management, LLMs predict disease risks by analyzing health data and provide preventive advice. At the same time, they also show potential in drug development, accelerating the discovery and development of new drugs by analyzing the structure and biological activity of compounds. However, these applications also bring challenges in value alignment and governance. For example, LLMs may learn and amplify existing biases in the training data, leading to inaccurate diagnoses for certain groups or unfair treatment plans. In addition, issues of privacy and security of medical data must also be considered as key factors. In the allocation of medical resources, LLMs need to find a balance between fairness and efficiency to ensure the rational allocation of resources. To ensure that the application of LLMs is responsible and sustainable, comprehensive consideration and governance in multiple aspects such as technology, ethics, and law are needed to solve these challenges and maximize their potential in the medical field (Zangrossi et al., 2024).

The application of LLMs in the medical field has brought unprecedented opportunities, but also a series of unique challenges(D'Antonoli et al., 2024; Thirunavukarasu et al., 2023). First, the high sensitivity of medical data requires LLMs to strictly comply with laws and regulations on privacy protection and data security when processing and analyzing(Hurley et al., 2023). These data often contain personal health information of patients, and once leaked or misused, it not only violates personal privacy but may also lead to serious legal and ethical issues(Savage et al., 2024; Tian et al., 2024). Second, the high risk of medical decision-making requires LLMs to have extremely high accuracy and reliability when assisting in diagnosis and treatment recommendations. Incorrect diagnoses or treatment recommendations can not only damage the health of patients but may also lead to medical disputes and legal responsibilities(Mannstadt & Mehta, 2024). In addition, the limited medical resources also put higher demands on the application of LLMs. LLMs need to optimize resource allocation and improve the efficiency and effectiveness of medical services under limited resource conditions(J. K. Kim et al., 2023; S. Liu et al., 2024; Omiye et al., 2024).

These characteristics require that the application of LLMs in the medical field must pay more attention to value alignment and governance performance. Governance means that the design and application of LLMs need to conform to medical ethics and social responsibility, ensuring that their output results are fair, transparent, and respect the rights and interests of patients(D'Antonoli et al., 2024). Governance performance involves the supervision and management of LLMs, requiring the establishment of effective evaluation and supervision mechanisms to ensure that the application of LLMs complies with medical standards and norms(L. Wang et al., 2024). For example, through DPM , the operating status of LLMs can be monitored in real-time, deviations can be detected and corrected promptly, and the application effect of LLMs in the medical field can be improved(Igarashi et al., 2024). At the same time, interdisciplinary cooperation also needs to be strengthened to promote experts from fields such as medicine, computer science, and ethics to participate in the design and evaluation of LLMs, ensuring that their application in the medical field is scientific and ethical. Only in this way can LLMs truly become a powerful tool to promote medical progress and improve the quality of medical services(Mesko & Topol, 2023).

Although the application of LLMs in the medical field has brought significant technological progress, it has also raised many value biases and ethical issues(Eggmann et al., 2023). For example, algorithmic bias and discrimination may lead to unfair treatment of certain groups in the allocation of medical resources. Data privacy and security issues are also

increasingly prominent, and the collection and analysis of a large amount of personal health data may violate personal privacy, leading to a social trust crisis(Omiye et al., 2024). In addition, as the auxiliary role of artificial intelligence in medical decision-making is enhanced, the main body status of doctors may be weakened, and patients may be more inclined to believe the analysis results of algorithms, neglecting the professional judgment and experience of doctors. These issues not only affect the quality and fairness of medical decision-making but may also trigger a trust crisis between doctors and patients, exacerbating social injustice(Miao et al., 2024).

Faced with the rapid development of LLMs technology, traditional governance models are difficult to adapt, and agile governance has become an internationally accepted governance plan. Agile governance emphasizes rapid perception, dynamic adaptation, and collaborative governance by multiple stakeholders to cope with the ethical and governance challenges brought by LLMs. The formulation and improvement of data governance rules are key issues in governance, and the construction of high-quality datasets and the updating of intellectual property systems are the basis for ensuring the development of LLMs. At the same time, the World Health Organization's guidelines on "Ethical Governance of Artificial Intelligence in Health and Care" propose six ethical principles, including protecting human autonomy, promoting human well-being and safety, ensuring transparency and interpretability, etc., providing guidance for the ethical governance of LLMs(Harrer, 2023). Governments and organizations of various countries are also actively formulating relevant policies and regulations to cope with the challenges brought by LLMs.

The ethical issues and governance challenges brought by the application of LLMs in the medical field are urgent and require the joint efforts of the government, enterprises, and all sectors of society. Enhancing the transparency and interpretability of LLMs, strengthening data protection and privacy security, promoting collaborative governance by multiple stakeholders, and promoting ethical education and public participation are key measures to solve these problems(Harrer, 2023). Through these measures, we can better cope with the challenges brought by the application of LLMs in the medical field, ensuring that they improve the quality of medical services while also safeguarding the rights and interests of patients and social fairness and justice. In the future, with the continuous advancement of technology and the improvement of governance mechanisms, the application prospects of LLMs in the medical field will be broader, but at the same time, it is also necessary to continue to pay attention to and solve the ethical and governance issues they bring.

2.3 Human Values Framework

Human values are the internal beliefs and criteria that individuals or groups rely on when making decisions. They influence how people perceive the world and evaluate actions, as well as the decision-making process and outcomes (Cothorn, 2019; Fritzsche, 1991; Glover et al., 1997; Mattison, 2000; S. H. Schwartz, 2012). Human values are the intrinsic beliefs and standards of measurement that individuals or groups rely upon when making decisions. They shape how people view the world and assess actions, influencing both the process and the outcomes of decision-making (Elango et al., 2010). The ultimate decision is the manifestation and practice of values in specific contexts (Reamer, 2018; Rokeach, 2008).

The development and transformation of values is a complex process influenced by a multitude of factors. Primarily, differences in individual values often stem from variations in personal life experiences, such as age, gender, educational attainment. These factors not only influence the formation and change of values individually but also interact systematically to further affect the development and transformation of values (Bilsky & Schwartz, 1994; Russo et al., 2022; Schwartz et al., 2012). Secondly, social factors also play a crucial role in the formation of values. For instance, the influence of family, educational institutions, and work environments are significant in shaping one's values (Brosch & Sander, 2015; Mayton et al., 1994; Russo et al., 2022). Lastly, different cultural backgrounds also have a profound influence on values. Various cultures have distinct preferences and expressions of values, which in turn affect the differences in individual value systems (Brosch & Sander, 2015; Schwartz, 2003, 2010).

Values are abstract, transcending specific behaviors and situational goals (S. Schwartz, 2012; S. H. Schwartz, 2012; Schwartz & Boehnke, 2004). Values not only affect individuals or groups' judgments about what is "good" and what is "worthwhile," but also, due to their intrinsic nature, they more often influence behavior and decision-making at an unconscious level (Schwartz, 2009). Human values are relatively stable but possess a dynamic nature, changing in response to shifts in individuals' or groups' cognition and environmental factors. (Hills, 2002). At the personal level, there is significant diversity in values, which can vary depending on the individual's pursued interests, environmental context, and psychological state, among other factors. (Hills, 2002). At the group level, values exhibit a degree of universality. Although there are differences in values across various countries and cultural contexts, there is a convergence in fundamental values (Bond, 1988; Sagiv & Schwartz, 2022; S. Schwartz, 2012;

Schwartz, 1994; S. H. Schwartz, 2012; Schwartz & Bilsky, 1987, 1990).

In the fields of sociology and psychology, although scholars believe that there are significant differences in individual values, most scholars have reached a consensus that the values of the groups to which individuals belong exhibit a high degree of consistency. (Piurko et al., 2011; Schwartz, 1994, 2007). This has prompted scholars to undertake a series of attempts to measure and understand the specific content of human values. Scholars believe that the measurement of values is not only for explaining the psychological state of the individual's inner self but also reflects broader socio-cultural characteristics (Braithwaite & Law, 1985; Piurko et al., 2011; Schwartz, 2007). To achieve this goal, researchers have developed a variety of measurement tools and methods, as shown in Table 2.1.

Shwartz's value theory posits that human values are universal. He has developed four scales—the Schwartz Value Survey (SVS), the Portrait Values Questionnaire (PVQ), the PVQ21 (ESS Human Values Scale), and the Revised PVQ-RR—to measure individuals' values. These scales assess the importance individuals place on ten fundamental values such as self-direction, stimulation, hedonism, achievement, power, security, conformity, tradition, benevolence, and universalism, aiming to reveal the commonalities and differences in values across cultures (Schwartz, 2007, 2009; S. H. Schwartz, 2012; Schwartz, 2021; Schwartz & Bilsky, 1987).

Rokeach developed a value scale and proposed a value theory that distinguishes values into terminal and instrumental values. He created the Rokeach Values Survey (RVS) scale, which requires individuals to rank the importance of 18 terminal values and 18 instrumental values during the measurement process. (Rokeach, 1967).

Hofstede's VSM82, VSM94, and VSM08 scales measure cultural diversity. Initially, he identified four dimensions: power distance, uncertainty avoidance, individualism vs. collectivism, and masculinity vs. femininity. VSM94 added long-term vs. short-term orientation, while VSM08 introduced indulgence vs. restraint (Hofstede, 1984, 2011; Minkov & Hofstede, 2012).

Oishi developed the Pairwise Comparison Values Survey (PCVS) to measure and compare value differences across cultures by pairing the 10 values identified by Schwartz, forming 45 value pairs for comparison (Oishi et al., 2005). Kluckhohn and Strodtbeck's theory posits that values are revealed through answers to five basic questions on human relations with time, nature, others, and motivations, as well as the essence of human nature. They provided alternative

answers and cultural measures for these questions, leading to the value orientation theory. This theory suggests that societies universally address certain questions with a limited set of value-based solutions, though cultures may prefer different solutions (Hills, 2002).

Smith and Bond used data from the World Values Survey (WVS) to point out overlaps between Hofstede's and Schwartz's theories, suggesting that a universally applicable value theory may be within reach (Smith & Bond, 2019). Allport, Vernon, and Lindzey designed the Study of Values (SOV) scale, which includes 45 behavioral situations. During the measurement process, individuals rate the behavioral options within the situations, allowing for an understanding of personal values and the interrelationships among various values (Allport, 1960). Kaplan revised the outdated parts of the Study of Values (SOV) scale (Kopelman et al., 2003).

Table 2.1 Value Measurement

Scholars	Measurement Methods	Elements
Schwartz	《Schwartz Value Survey》(SVS), 《Portrait Values Questionnaire》(PVQ), PVQ21 (ESS Human Values Scale) and Revised PVQ-RR	Self-direction, Stimulation, Hedonism, Achievement, Power, Security, Conformity, Tradition, Benevolence, and Universalism
Rokeach	《Rokeach Values Survey》(RVS)	Terminal Values and Instrumental Values.
Hofstede	1982 (VSM82), 1994 (VSM94) and 2008 (VSM08)	Four dimensions: Power Distance, Uncertainty Avoidance, Individualism vs. Collectivism, Masculinity vs. Femininity; Five dimensions: Four dimensions + Long-term vs. Short-term Orientation; Six dimensions: Five dimensions + Indulgence vs. Restraint
Oishi	《Pairwise Comparison Values Survey》(PCVS)	Values differ across cultures
Kluckhohn and Strodtbeck	The five fundamental questions include the relationships between humans and time, nature, and each other, as well as the basic motivations of humans and the nature of human nature	Value Orientation Theory
Smith and Bond	《World Values Survey》(WVS)	There exists a universally applicable value theory
Allport, Vernon, and Lindzey	《Study of Values》(SOV)	Normative Value Measurement
Kaplan	Update 《Study of Values》(SOV)	Making the SOV comparable and

		acceptable
--	--	------------

3 Constructing Human Value Frameworks

Values are the internal beliefs and criteria that individuals or organizations rely on when facing choices, influencing how people perceive the world, evaluating actions, and making decisions, profoundly affecting the decision-making process and outcomes. The value system of an individual or group can provide clear direction and guidance for decision-makers, helping them to make relatively stable decisions in complex and changing environments that align with their own beliefs. Therefore, values are the cornerstone of decision-making and an indispensable part of the value decision-making process, guiding individuals and groups to choose value actions that align with their values, thereby achieving their value goals. The final decision is the embodiment and practice of values in specific contexts.

3.1 Introduction

In a complex and ever-changing world, value choice theory provides profound insights into making decisions that are consistent with one's own values. This theory believes that value choice is a two-way process of individual objectification and object subjectification, and it is a process in which the essence of human beings is realized. In the process of making value choices, we must follow the objective laws of things. Only by conforming to objective laws can value choices achieve their purpose and talents can achieve comprehensive and free development. This means that even in the face of uncertainty and change, we still need to navigate based on our inner values and choose behavioral paths that are consistent with personal values and adapt to objective reality.

This chapter will not only provide a comprehensive perspective on the framework of human values but also offer guidance on how to make decisions that align with one's values in a complex and ever-changing world. By understanding these values, alignment targets can be provided for LLMs value alignment, and better promote individual growth, social progress, and world peace.

In this constantly changing world, values serve as the cornerstone of individual and collective behavior, profoundly influencing how we understand the world, make decisions, and interact. This chapter aims to delve into the hierarchical structure of the human value framework and analyze how these value concepts are reflected and practiced in different cultures, societies, and legal systems. Through a systematic analysis of values, it is hoped to provide a comprehensive perspective to understand how values shape our behavior and decision-making

in a diverse world.

The content of this chapter is carefully organized into several sections, each focusing on analyzing the value concepts in specific domains. Below is the distribution of chapters in the document and a brief introduction to each:

The first part is ideology, where this section will explore value concepts in different areas such as political ideologies, economic models, social structures, religious beliefs, and philosophical thoughts, and how they influence our worldview and code of conduct.

The second part is legal norms, which will analyze the constitutions, civil and commercial law, criminal law, procedural law, and administrative regulations and policies of different countries, discussing how these legal norms reflect and shape the core values of a society.

The third part is cultural diversity, where this section will delve into the values in cultural expressions such as language, traditional customs, art, literature, and cuisine, and how these values enrich our cultural life and promote global cultural exchange.

The fourth part is universal moral ethics, in which this section will discuss universal moral and ethical values such as peace, development, fairness, justice, democracy, and freedom, and their importance in building a harmonious society.

The fifth part is a summary of the human values at the above four levels.

Each chapter will provide a detailed analysis to help gain an in-depth understanding of the specific manifestations and impacts of values in different domains. Through this structured approach, it is hoped to provide a clear framework to understand how values function at the individual, societal, and global levels.

Although values are abstract, they are deeply rooted in people's daily lives and have far-reaching effects on individuals, societies, and nations. As illustrated in Figures 3.1, values influence every aspect of personal and group life. At the ideological level, values profoundly affect the way people perceive the world, further influencing the group's value choices regarding political ideologies, economic models, and social structures. At the legal and regulatory level, values provide behavioral norms, allowing people to exercise their rights freely without violating laws and regulations. At the level of diverse cultures, values affect people's identity, thought patterns, and aesthetics, and understanding and embracing different values can better promote cross-cultural exchanges. At the universal ethical level, values such

as peace, development, and fairness provide universally recognized behavioral guidelines and moral standards. In the context of multiculturalism, universal ethical values promote respect and understanding among different cultures, reflecting the common aspirations of people around the world.

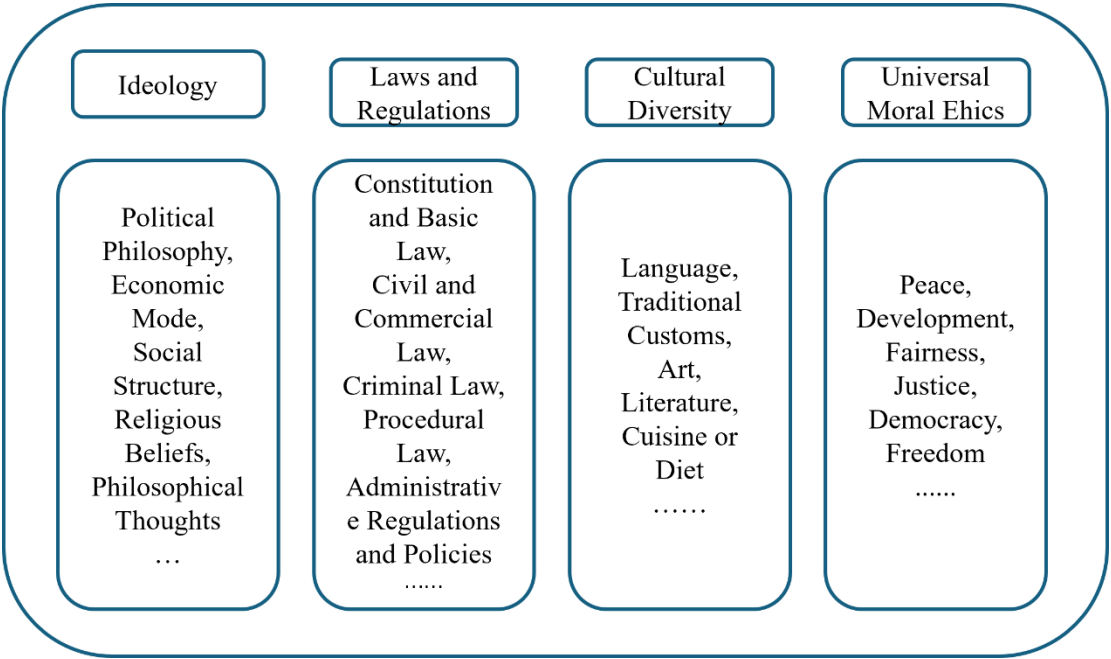


Figure 3.1 Values involve value decisions at different levels for individuals and groups

3.2 Level 1: Ideology

3.2.1 Political Ideology

Political ideology is the set of concepts and attitudes that people form in political activities, reflecting their basic views and positions on the state, society, economy, and other aspects. Different political ideologies are often driven by different values and influenced by various value systems. These ideologies reflect diverse historical, cultural, economic, and social contexts, and understanding them can help to gain insight into the values behind various political actions and claims. Each political ideology emphasizes different values and goals. Table 3.1 lists some of the main political ideologies and the primary values that drive them. For instance, liberals emphasize individual freedom and democratic rule of law; conservatives value tradition and order; socialists pursue equitable distribution of resources and social welfare; capitalists stress individual freedom and the rule of law; communists are committed to eliminating class differences; nationalists emphasize national independence and culture; and environmentalists focus on ecological conservation.

Table 3.1 Main Political Ideologies and Their Core Values

Political Ideology	Core Values				
Liberalism	Individual Freedom (Garnett, 1941; Macleod & Macleod, 1998)	Equality (Galston, 1986; Macleod & Macleod, 1998; Plattner, 1999)	Market Economy (Macleod & Macleod, 1998)	Democracy (Chan, 2002; Plattner, 1999)	Rule of Law (Dyzenhaus, 2007; Waldron, 2002)
Conservatism	Tradition (O'hara, 2011)	Stability (O'hara, 2011; Tännsjö, 2022)	Order (Johnson, 2015; Tännsjö, 2022)	Authority (Feather, 1979; McGovern, 2007)	Family (Farber, 2010)
Socialism	Equality (Plant, 1999; Self, 2017)	Collective Interest (Bogomolov, 1986; Przeworski, 1980)	Social Justice (Gilabert & O'Neill, 2019; Lacerda Jr, 1815)	State Intervention (Bettelheim, 1973; Hodgson, 2019)	Public Welfare (Dixon & Macarov, 2016; Hodgson, 2019)
Capitalism	Individual Freedom (Nolan, 2007; Rand, 1967)	Private Property (Rand, 1967; Rubin & Klumpp, 2012)	Competition (Nolan, 2007; Shaikh, 2016)	Rule of Law (Davis, 2001; Meltzer, 2012)	Personal Responsibility (Rand, 1967; Young, 2003)
Communism	Proletarian Dictatorship (Çam & Kayaoğlu, 2015; Draper, 1962)	Freedom (Kandiyali, 2014; McCarney, 1991)	Public Ownership (Roemer, 1992; Sherman, 1970)	Equality (Cafiero, 2005; ECHOLS III, 1976; Georgiou, 2016)	Fairness and Justice (Wang & Zhang, 2022)
Nationalism	National Identity (Hutchinson & Smith, 1994)	National Sovereignty (Hutchinson & Smith, 1994)	National Interest (Hutchinson & Smith, 1994)	Cultural Pride (Hutchinson & Smith, 1994)	National Revival (Hutchinson & Smith, 1994)
Environmentalism	Ecological	Sustainable	Environmental	Resource	Value of

Balance (Pepper, 2019)	Development (O'Riordan, 1991)	Protection (O'Riordan, 1991; Pepper, 2019)	Conservation (O'Riordan, 1991; Pepper, 2019)	Nature (Pepper, 2019)
------------------------------	-------------------------------------	---	---	-----------------------------

Summarizing the main values within the primary political ideologies in Table 3.1 can be distilled into five dimensions, which are "Freedom and Authority", "Democracy and Rule of Law versus Autocracy", "Ecological Protection and Economic Development", "Tradition and Modernity" and "Social Equality and Justice versus Economic Efficiency" In the dimension of "Freedom and Authority" the value of "freedom" emphasizes personal liberty and rights, advocating for the restriction of governmental power to further ensure individual autonomy in choice and action; "authority" on the other hand, emphasizes social order and collective interests, advocating for strong government leadership by the state and society to maintain stability and order. In the dimension of "Democracy and Rule of Law versus Autocracy", "democracy and rule of law" advocate for the governance of the country through democratic elections and legal systems, emphasizing equality before the law and the right of citizens to participate in political decision-making; "autocracy" however, advocates for the concentration of power in the hands of individuals or a few, managing the country through personal will and authority, potentially neglecting democracy and legal norms. In the dimension of "Ecological Protection and Economic Development" the value of "ecological protection" emphasizes the importance of environmental protection and sustainable development, advocating that economic development should not come at the expense of the environment, that is, while pursuing economic benefits, natural resources and the ecological environment should be protected; the value that focuses more on "economic development" may sacrifice environmental quality, emphasizing economic growth and the accumulation of wealth. In the dimension of "Tradition and Modernity", "tradition" places more emphasis on the value of history, culture, and heritage, advocating that maintaining social traditional structures and customs is more important than innovation and change; while the value that places more emphasis on "modernity" highlights innovation, change, and modern development, discarding outdated traditions. In the dimension of "Social Equality and Justice versus Economic Efficiency", "social equality and justice" advocates for the fair distribution of social resources, focusing more on caring for vulnerable groups and pursuing social harmony; whereas the value that emphasizes "economic efficiency" more will stress market mechanisms and encourage competition, advocating for economic growth by improving production efficiency and economic benefits, often at the expense of some social equality.

3.2.2 Economic Mode

The selection of different economic models in ideologies often reflects people's fundamental views on resource allocation, social structure, and the role of government. Capitalism relies on free markets and private property to drive resource allocation and economic growth, while socialism and communism emphasize reducing inequality and achieving fair resource distribution through state or collective intervention and public ownership. Mixed economies attempt to balance market efficiency with social equity, state capitalism guides development by state control over key industries, and planned economies rely on state planning to prioritize collective interests. Market socialism combines market efficiency with socialist principles of fairness, welfare states reduce inequality through extensive social security policies, green economies focus on sustainability and environmental protection, and cooperative economies promote economic and social welfare through common ownership and democratic management. These models reflect different values and goals, aiming to achieve economic prosperity, social justice, and environmental sustainability through different mechanisms. Table 3.2 outlines the main economic models and the values that drive their emergence.

Table 3.2 Main Economic Models and Their Key Values

Economic Mode	Key Values				
Capitalism	Free market (Nolan, 2007; Rubin & Klumpp, 2012)	Private property (Nolan, 2007; Rubin & Klumpp, 2012)	Competition (Nolan, 2007; Rubin & Klumpp, 2012; Shaikh, 2016)	Efficiency (Nolan, 2007)	
Socialism	Social equity (Dye & Zeigler, 1988; Gilabert & O'Neill, 2019; Self, 2017)	Collective ownership (Gilabert & O'Neill, 2019; Verdery, 2004)	State intervention (Gilabert & O'Neill, 2019)	Public welfare (Gilabert & O'Neill, 2019)	
Communism	Classless society (Medalie, 1959; Ost, 2015; Resnick & Wolff, 1988)	Dictatorship of the proletariat (Draper, 1987; Tabak, 2000)	Public ownership (Draper, 1987; Harrison, 2012; Sherman, 1970)		
State Capitalism	State control (Bremmer, 2008; Musacchio et al.,	Economic development (Musacchio et	Efficiency first (Pollock, 2020)	Policy-driven (Wright et al., 2021)	

	2015; Wright et al., 2021)	al., 2015; Wright et al., 2021)		
Planned Economy	Collectivism (Eatwell et al., 1990; Haberler, 1951)	Central planning (Durbin, 1936; Kowalik, 1990)	Equal distribution (Eatwell et al., 1990)	National self-sufficiency (Eatwell et al., 1990)
Market Socialism	Market mechanisms (Bergson, 1967; Brus, 1990; Miller, 1990)	Social equity (Bergson, 1967; Miller, 1990)	Democratic decision-making (Bergson, 1967; Miller, 1990)	Environmental sustainability (Bergson, 1967; Gabriele & Schettino, 2012; Ortmann, 2020)
Welfare State	Social security (Taylor-Gooby, 1991)	Civil rights (Taylor-Gooby, 1991)	Social solidarity (Arts & Gelissen, 2001; Baldwin, 1990)	Social inclusion (Baldwin, 1990)

Different countries and social groups have different value judgments on resource allocation, social structure, and the role of government, which influence the choice of economic models at the national and social group levels. The main economic models in Table 3.2 include values from five dimensions: "Free Market and State Intervention in the Economy", "Individualism and Collective Welfare", "Social Equity and Efficiency," "Economic Development and Environmental Sustainability", and "Democratic Decision-Making and Authority". Among them, the "Free Market and State Intervention in the Economy" dimension, groups with a value orientation towards "free market" advocate for the market to operate freely with minimal government intervention, believing that the market mechanism can achieve the most effective distribution and regulation of freedom and the economy on its own; groups more willing to choose the value of "state intervention in the economy" advocate for the government to play a more active role in the entire economy, regulating market operations through policies and regulations to achieve socio-economic goals. In the "Individualism and Collective Welfare" dimension, the value of "individualism" focuses on individual freedom, rights, and responsibilities, prioritizing personal interests, and believes that the pursuit of personal interests is the driving force for social progress and development; groups inclined to choose the value of "collective welfare" emphasize the interests and unity of the whole society, advocating for means such as collective interest distribution to ensure overall social welfare and reduce the occurrence of inequality. In the "Social Equity and Efficiency" dimension, the value of "social equity" places more emphasis on the fair distribution of social resources and opportunities,

advocating for policies such as redistribution to reduce social inequality and enhance social justice; groups that focus more on the value of "efficiency" will emphasize the efficiency and productivity of economic activities, and under the guidance of this value, the group will consider the effective use of resources and the maximization of output as key factors for economic development. In the "Economic Development and Environmental Sustainability" dimension, groups that focus more on "economic development" will aim for economic growth and improved living standards, willing to sacrifice environmental quality in the pursuit of economic benefits; groups that focus on "environmental sustainability" emphasize protecting the environment and natural resources while developing the economy, pursuing long-term environmental quality and social welfare. In the "Democratic Decision-Making and Authority" dimension, the value of "democratic decision-making" advocates for policy-making through democratic processes and citizen participation, emphasizing the importance of transparency and accountability in decision-making, and believes that democratic decision-making is key to achieving social justice and effective governance in economic policy; groups that focus more on the value of "authority" will advocate for decision-making through centralized power, believing that strong leadership and clear command can effectively improve decision-making efficiency and more easily achieve group development goals.

3.2.3 Social Structure

The choice of different social structures within ideologies reflects people's fundamental views on social organization, power distribution, and social relationships, embodying a variety of social organizations and values. Hierarchical social structures are based on clear levels and classes, emphasizing social order and the roles of each stratum; egalitarianism pursues equality in rights and opportunities, reducing inequality through education and legal means; collectivism values collective interests and unity, encouraging collective action to achieve common goals; individualism emphasizes personal freedom and independence, promoting personal development and social progress through competition; feudal social structures are centered on land ownership and the relationship between lords and serfs, emphasizing loyalty and family honor; bureaucratic social structures rely on strict rules and hierarchical systems to maintain social management; democratic social structures are based on popular sovereignty, achieving social management and decision-making through democratic procedures, valuing citizen participation and supervision; anarchism opposes authority and the state, advocating voluntary cooperation and mutual aid to achieve self-management; communist social structures aim to achieve a classless society and fair resource distribution through the elimination of class

differences and public ownership; welfare states reduce poverty through social security and welfare policies, providing basic services; tribal social structures are based on kinship and ethnic relationships, maintaining ethnic unity and traditional customs. These social structures reflect different social concepts and organizational methods from traditional to modern, from hierarchical to equal, aiming to achieve their respective social goals and values. Table 3.3 is the main social structures and their underlying basic values.

Table 3.3 Main Social Structures and Their Key Values

Social Structure	Key Values			
Hierarchical	Authority (Picavet, 2024; Redhead & Power, 2022; Rubin, 2000; Singh, 2018)	Order (Picavet, 2024; Redhead & Power, 2022; Rubin, 2000; Singh, 2018)	Tradition (Setyoningrum et al., 2015; Singh, 2018)	Balance (Deshpande, 2010)
Egalitarianism	Equality (Arneson, 2002; Arneson, 1999; Landesman, 1983)	Justice (Arneson, 2002; Arneson, 1999; Landesman, 1983)	Freedom (Arneson, 2002; Landesman, 1983)	Democracy (Arneson, 2002; Arneson, 1999)
Collectivism	Collective Interest (Kagitcibasi, 1997; Newman, 2013)	Unity (Hui, 1988; Kagitcibasi, 1997)	Mutual Aid (Hui, 1988)	Responsibility (Hui, 1988; Kagitcibasi, 1997)
Individualism	Personal Freedom (Kagitcibasi, 1997)	Independence (Kagitcibasi, 1997)	Self-Actualization (Kagitcibasi, 1997)	Competition (Kohl & Sapién, 2021)
Feudalism	Loyalty (Ganshof, 1996)	Obedience (Brenner, 1990; Ganshof, 1996)	Land Ownership (Brenner, 1990; Ganshof, 1996)	Family Honor (Painter, 1960)
Bureaucracy	Efficiency (Albrow, 1970; Beetham, 1996; Du Gay, 2005; Olsen, 2005)	Rules (Albrow, 1970; Beetham, 1996; Du Gay, 2005; Olsen, 2005)	Level (Albrow, 1970; Olsen, 2005)	Authority (Albrow, 1970; Beetham, 1996; Du Gay, 2005)
Democracy	Popular Sovereignty (Arblaster, 2002)	Participation (Arblaster, 2002; Dahl, 2020)	Equality (Arblaster, 2002; Dahl, 2020)	Transparency (Dahl, 2020)
Anarchism	Freedom (Berkman, 2003; Guerin, 1970)	Equality (Berkman, 2003; Guerin, 1970)	Opposition to Authority (Berkman, 2003;	Voluntary Cooperation (Berkman, 2003;

				Guerin, 1970)	Guerin, 1970)
Communism	Class	Equality	Public Ownership	Classless Society	Collective Interest
	(Pipes, 2003)		(Pipes, 2003)	(Pipes, 2003)	(Bosteels, 2014; Pipes, 2003)

The choice of social structures within ideologies reflects people's value judgments on power distribution and social relationships. The main social structures and their underlying values in Table 3.3 can be further divided into six dimensions: "Freedom and Authority", "Tradition and Change", "Collective Interest and Individualism," "Democratic Participation and Elite Rule", "Social Justice and Efficiency" and "Order and Free Competition". The choice in the "Freedom and Authority" dimension reflects the group's different value choices between personal freedom and social order. Groups that focus more on "freedom" emphasize individual freedom and rights, advocating for limiting government power in the distribution of group rights to ensure the freedom of individual choice and action; groups that focus more on the value of "authority" will emphasize the importance of social order and collective interests in the distribution of rights and social relationships, advocating for strong government leadership to maintain social order and stability. The opposition of values in the "Tradition and Change" dimension reflects the group's different pursuits of maintaining historical culture and pursuing modernization and change. Groups that value "tradition" more highly value the importance of history and traditional culture, hoping to maintain traditions and customs through inheritance; groups that emphasize "change" value more the innovation and change brought by modernization, neglecting the role of traditions and customs in social life, and pursuing technological progress and social change. The value dispute in the "Collective Interest and Individualism" dimension focuses on the choice between social welfare and individual freedom and achievement in social relations and rights distribution. "Collective Interest" emphasizes the overall interests of society and social unity, trying to balance social welfare through government policies; "Individualism" emphasizes individual freedom and achievement, achieving social progress and development through individual pursuit of success. The value choice in the "Democratic Participation and Elite Rule" dimension involves the importance of citizen participation in the decision-making process and the effectiveness of individual elite leadership. Groups that tend to value "democratic participation" emphasize the importance of citizen participation in policy formulation, believing that social justice is more important than decision-making efficiency and other issues; groups that value "elite rule" focus on the management of society by a small number of elites with ability and professional knowledge reserves to improve the efficiency and effectiveness of decision-making. The "Social Justice and Efficiency"

dimension considers the relationship between the fair distribution of resources and economic efficiency". Social Justice" emphasizes the fair distribution of resources in society, reducing social injustice through means such as resource redistribution; "Efficiency" emphasizes the efficiency of social resource distribution, promoting social development through the effective use of resources. The "Order and Free Competition" dimension is the value choice made by the group between social stability and market vitality. "Order" emphasizes the importance of social order, maintaining social order and stability through the formulation of rules and regulations; the value of "free competition" advocates for stimulating market freedom competition, promoting innovation and development, and achieving social progress.

3.2.4 Religious Beliefs

Religion plays a vital role in human society, with different religious beliefs closely associated with specific values and ways of life. Although major religions and belief systems each have their own characteristics, they collectively form a rich tapestry of human spiritual pursuits, moral behaviors, and understanding of universal truths. Christianity emphasizes faith and redemption; Buddhism seeks wisdom and liberation; Hinduism highlights reincarnation and diversity; Judaism values covenant and ethics; Taoism advocates harmony with nature; and Islam emphasizes equality, justice, and compassion. These beliefs not only provide guidance and spiritual comfort for adherents but also contribute to the moral and cultural development of society. Table 3.4 outlines the main religious beliefs and their underlying values.

Table 3.4 Main Religious Beliefs and Their Core Values

Religion	Core Values				
Christianity	Love (Johnson, 2012; Ratzinger, 2004; Stott, 2017)	Forgiveness (Johnson, 2012; Ratzinger, 2004; Stott, 2017)	Redemption (Johnson, 2012; Ratzinger, 2004; Stott, 2017)	Faith (Johnson, 2012; Ratzinger, 2004; Stott, 2017)	Hope (Johnson, 2012; Ratzinger, 2004; Stott, 2017)
Buddhism	Compassion (Conze, 2012; Harvey, 2012; Narasu, 1907)	Wisdom (Conze, 2012; Faure, 2011; Harvey, 2012; Narasu, 1907)	Liberation (Conze, 2012; Harvey, 2012; Narasu, 1907)	Karma (Conze, 2012; Faure, 2011; Harvey, 2012; Narasu, 1907)	Middle Way (Faure, 2011; Harvey, 2012; Narasu, 1907)
Hinduism	Karma (Doniger, 2014)	Reincarnation (Doniger, 2014)	Moksha (Doniger, 2014)	Devotion (Doniger, 2014)	Diversity (Doniger, 2014)

	2014; Flood, 1996; Klostermaier, 2007, 2014)	2014; Flood, 1996; Klostermaier, 2007, 2014)	2014; Klostermaier, 2007, 2014)	2014; Flood, 1996; Klostermaier, 2007, 2014)	Flood, 1996; Klostermaier, 2007, 2014)
Judaism	Covenant (Baeck, 2019; De Lange, 2000; Hertzberg, 1998; Steinberg, 1947)	Ethics (Baeck, 2019; De Lange, 2000; Hertzberg, 1998; Steinberg, 1947)	Law (Baeck, 2019; De Lange, 2000; Hertzberg, 1998; Steinberg, 1947)	Community (Baeck, 2019; De Lange, 2000; Hertzberg, 1998; Steinberg, 1947)	Justice (Baeck, 2019; De Lange, 2000; Hertzberg, 1998; Steinberg, 1947)
Taoism	Nature (Creel, 1982; Fuchen, 2013; Verellen, 1995)	Harmony (Fuchen, 2013)	Wu Wei (Creel, 1982; Fuchen, 2013; Verellen, 1995)	Longevity (Creel, 1982; Fuchen, 2013)	Virtue (Creel, 1982; Fuchen, 2013)
Islam	Peace (Hussain, 2011; Rahman, 2020; Williams, 2020)	Mercy (Hussain, 2011; Williams, 2020)	Justice (Hussain, 2011; Rahman, 2020; Williams, 2020)	Tolerance (Hussain, 2011; Rahman, 2020; Williams, 2020)	Equality (Hussain, 2011; Rahman, 2020; Williams, 2020)

The values of religious beliefs are expressions of people's spiritual pursuits, moral behaviors, and understanding of universal truths. The values in Table 3.4 can be summarized into six dimensions: "Religious Belief and Secular Rationality", "Spiritual Practice and Material Pursuit", "Karma and Free Will", "Diversity, Tolerance, and Unity and Order", "Moral Law and Personal Ethics", and "Peaceful Coexistence and Competitive Conflict". In the "Religious Belief and Secular Rationality" dimension, the value choice reflects the group's decision between supernatural forces and rational thinking. The "Religious Belief" dimension emphasizes the spiritual pursuit beyond the material world, where the group believes in supernatural forces or deities, and considers religious doctrines and divine revelations as guidance for life. Groups inclined towards "Secular Rationality" value rational thinking and empirical evidence, advocating understanding the world through science and logic, and may be skeptical or critical of supernatural phenomena. The "Spiritual Practice and Material Pursuit" dimension showcases the group's choice between the needs for inner growth and the pursuit of

external wealth. "Spiritual Practice" values inner spiritual growth and cultivation, believing that meditation and prayer can purify the mind and promote inner spiritual development. "Material Pursuit," on the other hand, emphasizes material wealth and sensory enjoyment, pursuing material possessions as life goals. The "Karma and Free Will" dimension is a valuable judgment between the consequences of actions and personal free choice. Groups believing in "Karma" think that actions have corresponding outcomes, with good deeds leading to good results and vice versa, thus paying more attention to their behavior norms. "Free Will" emphasizes freedom of personal choice, believing that individuals can decide their actions and destiny without being bound by the law of cause and effect, potentially enjoying the pleasures of short-term free actions while overlooking long-term consequences. "Diversity, Tolerance, and Unity and Order" involves respecting differences in beliefs and cultures versus maintaining the order of one's own faith. Groups focusing on "Diversity and Tolerance" are more willing to respect differences in beliefs and cultures, advocating tolerance and understanding, considering diversity a natural state of society and the universe. Groups emphasizing "Unity and Order" pursue consistency in religious beliefs and social order, deeming a unified belief system and regulated social order crucial for harmony and stability. "Moral Law and Personal Ethics" represents the different understandings of the role of religious doctrine and individual conscience in moral judgment. "Moral Law" emphasizes the role of moral norms within religious teachings, considering these norms as guidelines for behavior and standards for judging right from wrong. In contrast, groups that lean towards valuing "Personal Ethics" believe that individual moral judgment and autonomous decision-making are the foundations of moral actions. The dimension of "Peaceful Coexistence and Competitive Conflict" is about the value choices groups make in addressing conflicts arising from religious beliefs. Groups that prefer the value of "Peaceful Coexistence" emphasize peace and love, believing that conflicts should be resolved through communication and other non-violent means, fostering harmonious coexistence among different groups and beliefs. On the other hand, those who see "Competitive Conflict" as inevitable, or even necessary for religious selection, view competition and conflict as unavoidable aspects of religious choice and practice.

3.2.5 Philosophical Thoughts

Philosophical ideas play a central role in ideologies, reflecting people's fundamental views on existence, knowledge, value, and morality. Major philosophical schools offer diverse interpretations of the nature of the world and ways of knowing: Materialism regards matter as the foundation of the world and advocates the scientific method for problem-solving; Idealism

centers on spirit or consciousness, valuing personal spiritual experiences; Existentialism focuses on individual existence and responsibility, emphasizing free choice; Pragmatism tests knowledge through practice, emphasizing practical value; Rationalism relies on logic and reason to discover universal truths; Empiricism acquires knowledge through sensory experience; Idealism pursues spiritual and moral values beyond the material; Naturalism advocates coexistence with nature, respecting natural laws; Humanism emphasizes human value, advocating education and cultural enhancement; Structuralism focuses on the internal structure and systemic nature of social phenomena; Postmodernism opposes singular truth, emphasizing diversity and deconstruction. These philosophical viewpoints reflect humanity's deep contemplation and exploration of knowledge, reality, value, and the meaning of existence. As shown in Table 3.5, the main philosophical thoughts and their underlying values are as follows.

Table 3.5 Main Philosophical Thoughts and Their Core Values

Philosophical Thought	Core Values				
Materialism	Matter determines consciousness (Eagleton, 2017; Moser & Trout, 2002; Wolfe, 2016)	Objective reality (Eagleton, 2017; Moser & Trout, 2002)	Scientific empiricism (Eagleton, 2017; Moser & Trout, 2002; Wolfe, 2016)		
Idealism	Consciousness determines matter (Hoernlé, 1927)	Spirit supremacy (Hoernlé, 1927)	Subjective experience (Hoernlé, 1927)		
Existentialism	Personal freedom (Aho, 2014; Olson, 2012; Reynolds, 2014)	Responsibility (Aho, 2014; Olson, 2012; Reynolds, 2014)	Personal choice (Aho, 2014; Olson, 2012)	Existence precedes essence (Aho, 2014; Olson, 2012)	
Rationalism	Reason (Katz, 1997; Peacocke, 2002)	Logic (Katz, 1997; Peacocke, 2002)	Knowledge (Katz, 1997; Peacocke, 2002)	Universal truth (Katz, 1997; Peacocke, 2002)	
Empiricism	Experience (Feyerabend, 1962; Markie & Folescu, 2004)	Observation (Feyerabend, 1962; Markie & Folescu, 2004)	Sensation (Markie & Folescu, 2004)	Concrete knowledge (Feyerabend, 1962)	
Idealism	Ideals (Guyer & Horstmann, 2015)	Morality (Guyer & Horstmann, 2015)	Spiritual pursuit (Guyer & Horstmann, 2015)	Transcendence of reality (Foster, 2021)	
Naturalism	Natural law (Goetz &	Harmony (Sellars, 1922)	Ecological balance (Sellars, 1922)	Respect for nature (Goetz &	

	Taliaferro, 2008; Sellars, 1922)			Taliaferro, 2008)
Humanism	Human value (Herrick, 2010)	Individual dignity (Herrick, 2010)	Freedom (Herrick, 2010)	Equality (Herrick, 2010)
Structuralism	Systematicity (Lévi-Strauss, 2004; Sturrock, 2008)	Structure (Caws, 1968; Lévi- Strauss, 2004; Sturrock, 2008)	Relationships (Lévi-Strauss, 2004; Sturrock, 2008)	Holism (Lévi- Strauss, 2004)
Postmodernism	Diversity (Grenz, 1996; Powell, 2007)	Deconstruction (Grenz, 1996; Powell, 2007)	Anti- authoritarianism (Grenz, 1996)	Relativism (Grenz, 1996)

The fundamental views of groups on existence, knowledge, morality, and value form the foundation of philosophical thought. The values in the main philosophical thoughts in Table 3.5 can be summarized into five dimensions: "Objective Facts and Subjective Experience", "Individual Freedom and Social Responsibility", "Spiritual Pursuit and Material Pursuit", "Diversity and Unity", and "Individual Value and Social Value". Among them, the "Objective Facts and Subjective Experience" dimension involves how groups understand the objective world independent of individual perception and the difference between personal perception and experience in understanding the world. Groups that focus more on "objective facts" believe in an objective world that exists independently of individual perception and is not influenced by personal subjective judgments; whereas groups that focus more on "subjective experience" emphasize the dominant role of personal perception and experience in understanding the world, believing that reality can change due to different subjective understandings. The "Individual Freedom and Social Responsibility" dimension is a valuable choice between the rights to individual free choice and action and the individual's responsibility for social responsibility and group welfare. Groups that care more about "individual freedom" focus on the rights to individual free choice and action, emphasizing the importance of personal autonomy and self-realization; whereas groups that care more about "social responsibility" believe in the importance of an individual's responsibilities towards others and society. The "Spiritual Pursuit and Material Pursuit" dimension is a value choice between spiritual growth and sensory enjoyment, as well as between intrinsic value and external material wealth. Groups that choose "spiritual pursuit" tend to focus on spiritual growth and the significance of intrinsic value; whereas groups that focus more on "material pursuit" become engrossed in material wealth and sensory enjoyment, considering external conditions as the measure of happiness and success. The "Diversity and Unity" dimension is a value judgment between acknowledging the diversity

of the world and the belief that the world needs consistency. Groups that value "diversity" believe that philosophical thought is composed of a variety of different viewpoints, and diversity is one of the essential characteristics of philosophical thought, whereas groups that emphasize "unity" advocate finding commonalities and connections in diversity, trying to understand philosophical thought through a holistic and consistent approach. The "Individual Value and Social Value" dimension is a value preference choice between the value of individual achievements and an individual's contribution to society, as well as between individual uniqueness and group commonality. Groups that focus on "individual value" emphasize the uniqueness of the individual and personal achievements, believing that uniqueness and personal achievements can define individual value; groups that value "social value" believe that an individual's value is reflected in their contribution to society and interactions with others, emphasizing the importance of social roles and groups.

3.3 Level 2: Legal Regulations

3.3.1 Constitutions and Basic Laws

The constitutions and basic laws of different countries and regions reflect their unique legal systems and values. The U.S. Constitution protects civil liberties and rights through the separation of powers, achieving a balance of power; the German Basic Law emphasizes social welfare and environmental protection, reflecting social responsibility; the French Constitution maintains national unity and secularism through a centralized government; the UK Constitution is unwritten, respecting tradition and the rule of law, allowing for gradual reform; the Chinese Constitution emphasizes the socialist system and national unity; the Russian Constitution guarantees sovereignty and territorial integrity; the Indian Constitution pursues social equity and reduces inequality; the South African Constitution emphasizes racial equality and opposes discrimination; the Canadian Constitution respects multiculturalism; the Japanese Constitution emphasizes pacifism; the constitutions of Islamic countries are based on Sharia law, emphasizing social justice. These constitutions reflect the various commitments and practices of countries towards freedom, equality, justice, peace, and social welfare, as shown in Table 3.6, which outlines the constitutions and basic laws of typical countries and their underlying core values.

Table 3.6 Constitutions and Basic Laws of Typical Countries and Their Core Values

Country	Core Values					
U.S.	Democracy	Liberty	Equality	Rule of Law	Checks	and

Constitution	(Tucker, 1899; Watson, 1910)	(Tucker, 1899; Watson, 1910)	(Tucker, 1899; Watson, 1910)	(Tucker, 1899; Watson, 1910)	Balances (Tucker, 1899)
German Basic Law	Democracy (Grimm et al., 2019; Zekoll & Wagner, 2018)	Rule of Law (Grimm et al., 2019; Zekoll & Wagner, 2018)	Social Justice (Grimm et al., 2019; Zekoll & Wagner, 2018)	Human Rights (Grimm et al., 2019; Zekoll & Wagner, 2018)	Environmental Protection (Grimm et al., 2019; Zekoll & Wagner, 2018)
French Constitution	Democracy (Boyron, 2012)	Liberty (Bell, 1995; Boyron, 2012)	Equality (Bell, 1995; Boyron, 2012)	Secularism (Boyron, 2012)	
UK Constitution	Rule of Law (Jennings, 1967; King, 2007)	Tradition (Jennings, 1967; King, 2007)	Gradual Reform (Jennings, 1967; King, 2007)		
Chinese Constitution	People's Democracy (Chiu, 1985; Killion, 2005)	Socialism (Chiu, 1985; Killion, 2005)	Rule of Law (Chiu, 1985; Killion, 2005; Lin, 2015)	National Unity (Chiu, 1985; Killion, 2005)	National Solidarity (Chiu, 1985)
Russian Constitution	Democracy (Kremyanska ya et al., 2014)	Rule of Law (Kremyanska ya et al., 2014)	State Sovereignty (Kremyanskaya et al., 2014)	Social Stability (Kremyanska ya et al., 2014)	
Indian Constitution	Democracy (Choudhry et al., 2016)	Secularism (Choudhry et al., 2016)	Social Justice (Choudhry et al., 2016)	Equality (Choudhry et al., 2016)	Pluralism (Choudhry et al., 2016)
South African Constitution	Democracy (Klug, 2010)	Equality (Klug, 2010; Smith, 2014)	Social Justice (Klug, 2010)	Human Rights (Klug, 2010)	Multiculturalism (Klug, 2010)
Canadian Constitution	Democracy (Dodek, 2024; Oliver et al., 2017)	Rule of Law (Clement, 1904; Dodek, 2024; Oliver et al., 2017)	Multiculturalism (Oliver et al., 2017)	Social Welfare (Oliver et al., 2017)	
Japanese Constitution	Democracy (McElwain, 2020)	Peace (McElwain, 2020)	Rule of Law (McElwain, 2020)	Human Rights (McElwain, 2020)	
Constitutions of Islamic Countries	Sharia (Lombardi, 2013)	Social Justice (Lombardi, 2013)	Fairness (Lombardi, 2013)		

The values behind the constitutions and basic laws of different countries influence the governance methods, legal systems, and social policies of nations. The values in Table 3.6 can be summarized into the following six dimensions: "Rule of Law and Rule by Man", "Human Rights and State Sovereignty", "Secularism and Religious Law", "Peace and Conflict", "Individual Responsibility and Social Welfare," and "Individualism and National Unity". Among them, the "Rule of Law and Rule by Man" dimension emphasizes the universal applicability of the law and the role of leaders' personal moral qualities in governance. "Rule of Law" emphasizes the universal applicability of the law and the fair execution of judgments, advocating that everyone is equal before the law, and the exercise of state power is also subject to legal constraints; whereas the "Rule by Man" value believes that governing the country depends on the authority and moral character of a few individuals, potentially breaking free from legal constraints. The "Human Rights and State Sovereignty" dimension involves the value choice between individual freedom and state power, with groups advocating for "human rights" considering the protection and promotion of individual basic rights and freedoms to be most important in the country's basic legal system, believing that the basic rights of individuals are inalienable; whereas groups that consider "state sovereignty" more important emphasize that the country has supreme rights. The "Secularism and Religious Law" dimension is a value judgment of the group between the separation of church and state and the status of religious doctrine in the law, with groups choosing "secularism" advocating for the separation of church and state, believing that laws formulated by the state should not be influenced by religious doctrine; groups emphasizing "religious law" believe that religious doctrine is the foundation of the legal system and its status in the entire social life is irreplaceable. The "Peace and Conflict" dimension is the group's value tendency in resolving conflicts, with groups more willing to choose "peaceful" methods to resolve conflicts pursuing the resolution of disputes between groups through peaceful means, considering international cooperation to be very important; groups inclined to use "conflict" to resolve disputes, due to reasons that the group cannot abandon, such as territorial disputes and religious differences, choose more radical ways to resolve disputes. The "Individual Responsibility and Social Welfare" dimension is a value choice between individual autonomy and group responsibility, with groups emphasizing "individual responsibility" highlighting that individuals should be responsible for their actions and achieve personal and group goals through effort; groups valuing "social welfare" advocate that the state and society should provide necessary support and protection to ensure that the basic living needs of each member are met. The "Individualism and National Unity" dimension is a value choice between individual free expression and collective harmony, with groups

inclined to choose "individualism" valuing individual freedom, rights, and the development of personality, considering individual development to be the foundation of social progress; groups emphasizing "national unity" believe that the integration of different individuals is the basis of social development.

3.3.2 Civil and Commercial Law

Civil and commercial laws within different legal systems reflect the varying emphasis on legal certainty, protection of individual rights, social justice, economic efficiency, and transactional security, as well as the values of different countries. The civil codes of the civil law system maintain individual rights and social order through detailed provisions, ensuring the certainty and predictability of the law; commercial law focuses on transactional security and corporate responsibility. The common law system relies on case law for civil law, emphasizing individual freedom and contract freedom, and is flexible; commercial law emphasizes market mechanisms and the sanctity of contracts. The Islamic law system's civil law is based on Sharia, emphasizing social justice and family values; commercial law requires transactions to comply with Islamic ethical standards. The Scandinavian legal system emphasizes social welfare and equality in civil law, and commercial law focuses on corporate social responsibility and sustainable development. Japanese civil law emphasizes the balance between individual rights, family harmony, and social order; commercial law emphasizes corporate responsibility and economic efficiency, while also establishing social trust. Indian civil law respects multiculturalism and pursues social justice; commercial law emphasizes economic freedom and contract flexibility, requiring corporate responsibility. Russian civil and commercial law both emphasize the protection of individual rights and social order, with commercial law also emphasizing transaction stability and corporate responsibility. These legal systems demonstrate a common pursuit of fairness, efficiency, and responsibility within different global cultures and legal traditions. As shown in Table 3.7, the typical legal systems' civil and commercial laws and their underlying basic values.

Table 3.7 Typical Legal Systems' Civil and Commercial Laws and Their Underlying Basic Values

Legal System		Core Values					
Civil Law System		Individual rights	Social order	Legal certainty and predictability	Economic efficiency	Transactional security	Corporate responsibility
		(Mackaay, 2021)	(Mackaay, 2021)	(Mackaay, 2021)	(Mackaay, 2021)	(Mackaay, 2021)	(Mackaay, 2021)

	2021)			2021)		
Common Law System	Individual freedom (Shapiro, 2023)	Contract freedom (Shapiro, 2023)	Flexibility of case law (Shapiro, 2023)	Free competition (Shapiro, 2023)	Market mechanism (Clark, 2015)	Sanctity of contracts (Clark, 2015)
Islamic Law System	Islamic jurisprudence (Mohammed, 1988; Sloane, 1988)	Social justice (Mohammed, 1988; Sloane, 1988)	Family values (Sloane, 1988)	Morality in economic transactions (Mohammed, 1988; Sloane, 1988)	Fairness in contracts (Sloane, 1988)	
Scandinavian Legal System	Social welfare (Bernitz, 2007)	Equality (Madsen, 2007)	Universality of law (Bernitz, 2007; Madsen, 2007)	Social responsibility (Bernitz, 2007)	Sustainable development (Madsen, 2007)	Corporate ethics (Bernitz, 2007)
Japanese Legal System	Individual rights (Oda, 2021)	Family harmony (Oda, 2021)	Social order (Oda, 2021)	Corporate responsibility (Oda, 2021)	Economic efficiency (Oda, 2021)	Social trust (Oda, 2021)
Indian Legal System	Individual rights (D'Souza, 2020)	Social justice (D'Souza, 2020)	Multiculturalism (D'Souza, 2020)	Economic freedom (D'Souza, 2020)	Flexibility of contracts (D'Souza, 2020)	Corporate responsibility (D'Souza, 2020)
Russian Legal System	Individual rights (Maggs et al., 2015)	Social order (Maggs et al., 2015)	Unity of law (Maggs et al., 2015)	Economic efficiency (Maggs et al., 2015)	Stability of contracts (Maggs et al., 2015)	Corporate responsibility (Maggs et al., 2015)

Groups with different values have different emphases on legal certainty, protection of individual rights, social justice, and transactional security in civil and commercial law. The values behind the main legal systems in Table 3.7 can be summarized into five dimensions: "Individual Rights and Social Order," "Legal Certainty and Legal Flexibility", "Social Justice and Economic Efficiency", "Transactional Security and Transactional Risk", and "Corporate Responsibility and Immoral Behavior". In the dimension of "Individual Rights and Social

Order" the value choice between ensuring individual freedom and dignity and maintaining stable social and economic order is measured. Groups that value "Individual Rights" more emphasize that civil and commercial law should provide sufficient protection for individual rights, ensuring that personal rights, property rights, and transaction rights are not violated; whereas groups that consider "Social Order" more important believe that civil and commercial law should focus more on maintaining social order and public interest, regulating individual and corporate behavior to ensure social stability, and may sacrifice individual interests when necessary. The dimension of "Legal Certainty and Legal Flexibility" measures the value tendency between predictable legal actions for citizens and businesses and the law's adaptation to social changes, with groups valuing "Legal Certainty" more believing that legal provisions need to be clear and stable, allowing individuals and businesses to predict the legal consequences of their actions; whereas groups inclined to choose "Legal Flexibility" believe that the law needs to have some flexibility in certain situations to adapt to social changes. The dimension of "Social Justice and Economic Efficiency" is itself a dual goal pursued by the law, but different groups have different emphases, with groups advocating "Social Justice" believing that civil and commercial law aims to ensure the rights of individuals and businesses through fair rules and reduce social inequality; groups advocating "Economic Efficiency" believe that civil and commercial law should also pursue economic efficiency, ensuring the effective allocation of resources and the smooth progress of economic activities. The dimension of "Transactional Security and Transactional Risk" is a value choice of risk preference for groups, with groups considering "Transactional Security" as the primary consideration in transactions expecting civil and commercial law to establish legal rules such as contract law and property law to protect the rights and interests of both parties and reduce transaction risks; groups willing to tolerate "Transactional Risk" often voluntarily bear greater risks in transactions for greater benefits or other factors. The dimension of "Corporate Responsibility and Immoral Behavior" requires individuals and businesses to make value choices between pursuing profits and assuming social responsibility, with groups valuing "Corporate Responsibility" more believing that civil and commercial law should focus on regulating corporate responsibility while protecting corporate rights to achieve sustainable economic development; groups tolerating "Immoral Behavior" believe that civil and commercial law needs to allow individuals or businesses to pursue economic benefits at the expense of morality within a certain range.

3.3.3 Criminal Law

Criminal laws across different legal systems reflect varying values, emphasizing legal

clarity, individual freedom, social justice, humanitarianism, and state authority to different extents. As shown in Table 3.8 below, civil law systems use explicit statutes to prevent crime; common law systems rely on case law, emphasizing individual freedom and fair trials; Islamic law systems are based on Sharia, emphasizing moral norms; Scandinavian law systems focus on humanity and social welfare; Japanese criminal law maintains social harmony and order; Indian criminal law respects multiculturalism and pursues social justice; Russian criminal law ensures national stability and security. The choices and values of these criminal law systems demonstrate the reflection of each country's historical, cultural, and economic background, as well as diverse methods and strategies in legal practice and social governance.

Table 3.8 Typical Legal Systems' Criminal Law and Their Underlying Basic Values

Legal System		Core Values				
Civil Law System		Legal certainty (Bohlander, 2008)	Crime prevention (Bohlander, 2008)	Social order (Bohlander, 2008)		
Common Law System	Law	Individual freedom (Ambos et al., 2020)	Fair trial (Ambos et al., 2020)	Flexibility of case law (Ambos et al., 2020)		
Islamic System	Law	Islamic jurisprudence (Burns, 2014; Gray, 2018)	Moral norms (Gray, 2018)	Social justice (Gray, 2018)		
Scandinavian Legal System		Humanitarianism (Langsted et al., 2019; Orfield, 2002)	Social welfare (Langsted et al., 2019; Orfield, 2002)	Crime prevention (Langsted et al., 2019)		
Japanese System	Legal	Social harmony (Oda, 2021)	Responsibility and duty (Oda, 2021)	Strictness of law (Oda, 2021)		
Indian System	Legal	Social justice (Baxi, 1982)	Multiculturalism (Baxi, 1982)	Adaptability of law (Galanter, 1968)		
Russian System	Legal	State authority (Maggs et al., 2015)	Social order (Maggs et al., 2015)	Unity of law (Maggs et al., 2015)		

The choices of values behind the criminal law of different legal systems reflect the varying degrees of emphasis different groups place on legal clarity, individual rights, humanitarianism, and state authority. The values behind the main legal systems' criminal law in Table 3.8 can be categorized into five different dimensions: "Individual Freedom and Social Order", "Legal Certainty and Legal Flexibility", "Legal Strictness and Humanitarianism", "Legal Adaptability and Legal Unity", and "Fair Trial and State Authority". Among them, the balance of the "Individual Freedom and Social Order" dimension is the core value choice of criminal law, aiming to balance individual rights and social stability. Groups that value "Individual Freedom" more believe that criminal law should respect and protect the freedom of individual actions without infringing on the rights of others, and criminal law should protect individual freedom

by restricting illegal interventions; groups that consider "Social Order" more important expect criminal law to have strong power, maintaining social order and security through the punishment of criminal acts, and may restrict individual freedom in some cases. The "Legal Certainty and Legal Flexibility" dimension requires groups to make judgments between the predictability and flexibility of criminal law. Groups that value "Legal Certainty" more expect criminal law to have clear provisions, allowing individuals to predict the legal consequences of their actions, and certain laws help to enhance the authority of the law; groups that value "Legal Flexibility" more expect criminal law to actively adapt to changes in social conditions while ensuring basic principles, making timely adjustments and handling new types of crimes and special situations. The "Legal Strictness and Humanitarianism" dimension is a value choice between social order and security and respect for criminals when punishing crimes. Groups that tend to choose "Legal Strictness" believe that criminal law should punish criminal acts through strict legal norms and punitive measures to deter potential criminal acts and protect the safety of society and individuals; groups that value "Humanitarianism" more expect criminal law to adhere to the spirit of humanitarianism during execution, respect the basic human rights of criminals, avoid inhumane punishment, and focus on rehabilitation and education rather than punishment. The "Legal Adaptability and Legal Unity" dimension is a moral choice between the ability of criminal law to respond to social changes and maintain consistency in different situations. Groups that value "Legal Adaptability" more expect criminal law to adjust to social development and changes to some extent, ensuring that the implementation of criminal law meets social needs and reflects fairness and justice; groups that emphasize "Legal Unity" believe that the implementation of criminal law should maintain consistency to ensure the authority and effectiveness of the law, avoiding unjust issues in law enforcement due to differentiation. The "Fair Trial and State Authority" dimension is a value choice between ensuring the operation of society and maintaining national authority. Groups that value "Fair Trial" more expect the results of criminal law trials to strictly follow criminal law provisions, protect the legal rights and interests of criminals, and achieve fairness in case trials; groups that value "State Authority" more expect criminal law to focus on maintaining the seriousness and authority of the state.

3.3.4 Litigation Law

Litigation laws across different legal systems reflect varying emphases on legal clarity, procedural norms, individual freedom, social justice, and state authority. As shown in Table 3.9 below, the litigation law of the civil law system ensures the fairness and predictability of trials

through detailed provisions and strict procedures; the common law system relies on case law, emphasizing individual freedom and the flexibility of trials; the Islamic law system is based on Sharia, emphasizing moral norms and social justice; the Scandinavian law system focuses on humanitarianism and social welfare while ensuring the fairness of trials; Japanese litigation law maintains social harmony and order, strictly applying the law; Indian litigation law respects multiculturalism, pursuing social justice and adaptability of trials; Russian litigation law ensures national unity and social order. These litigation laws reflect the different approaches and strategies of various countries to legal practice, social governance, and human rights protection.

Table 3.9 Typical Legal Systems and the Core Values Behind Their Litigation Laws

Legal System	Core Values					
Civil Law	Fair Trial (Zekoll & Wagner, 2018)	Procedural (Zekoll & Wagner, 2018)	Justice (Zekoll & Wagner, 2018)	Legal (Zekoll & Wagner, 2018)	Certainty (Zekoll & Wagner, 2018)	
Common Law	Procedural Fairness (Waddams, 2003)	Prevention of Abuse of Power (Waddams, 2003)		Trial Order (Heyting, 1932)		
Islamic Law	Islamic Sharia (Burns, 2014; Gray, 2018)	Moral Norms (Salim, 2015)		Social Order (Salim, 2015)		
Scandinavian Law	Humanitarianism (Langsted et al., 2019)	Social Welfare (Langsted et al., 2019)		Crime Prevention (Langsted et al., 2019)		
Japanese Law	Fair Trial (Oda, 2021)	Procedural Fairness (Oda, 2021)		Efficiency (Oda, 2021)		
Indian Law	Social Justice (Baxi, 1982)	Multiculturalism (Baxi, 1982)		Legal Adaptability (Galanter, 1968)		
Russian Law	State Authority (Maggs et al., 2015)	Social Order (Maggs et al., 2015)		Legal Uniformity (Maggs et al., 2015)		

As shown in Table 3.9, the value choices of different groups regarding the dimensions of values behind litigation law can be summarized into five different dimensions: "Fair Trial and State Authority", "Legal Certainty and Legal Adaptability", "Social Order and Individual Freedom", "Procedural Fairness and Efficiency", and "Humanitarianism and Social Justice". Among them, in the value choice dimension of "Fair Trial and State Authority" groups that tend to maintain "Fair Trial" emphasize the fairness of judicial procedures and the fairness of judgments to protect the rights of the parties and maintain social trust; groups that place more emphasis on "State Authority" believe that litigation law should establish a system of judicial finality and the authority of judicial procedures to ensure the seriousness and authority of

national laws, and to ensure the unity and dignity of the national legal system. In the dimension of "Legal Certainty and Legal Adaptability," groups that tend to choose "Legal Certainty" prefer the clarity and stability of legal rules to reduce the uncertainty of legal application and improve the predictability of the law; groups that tend to choose "Legal Adaptability" value the flexibility of the law to adapt to social changes and emerging situations, ensuring that the law can timely reflect social needs and fairness and justice. In the dimension of "Social Order and Individual Freedom" groups that choose "Social Order" tend to maintain social order and public interest, and ensure that members of society can live together in an orderly and harmonious environment through procedural norms and legal sanctions; groups that emphasize "Individual Freedom" more stress the protection of individual freedom and rights, ensuring that basic rights and freedoms are not violated while pursuing social order. The dimension of "Procedural Fairness and Efficiency" is one of the most critical value choices in litigation law; groups that focus on "Procedural Fairness" emphasize the fairness and rationality of the procedure, protect the litigation rights of the parties, and ensure the legality and fairness of the judgment process; groups that focus on "Efficiency" will pursue litigation efficiency, hoping to reduce litigation costs and resolve disputes in a timely manner through simplified procedures and improved judicial efficiency. In the dimension of "Humanitarianism and Social Justice" groups that pay more attention to "Humanitarianism" expect litigation law to respect the spirit of humanitarianism, especially in criminal proceedings, emphasizing fair treatment and human rights protection of the accused; groups that focus more on "Social Justice" expect to resolve social conflicts and contradictions through legal means, achieving fair distribution and social responsibility.

3.3.5 Administrative Regulations and Policies

Administrative regulations and policies are essential tools for national governance, and they vary in choice and manifestation across different countries and legal systems. Policies and administrative regulations in each country reflect its unique historical, cultural, economic, and political background, forming diverse governance methods and social development directions, and embodying each country's unique values. U.S. policies support free markets and individual freedom, emphasizing efficiency and innovation; Chinese policies focus on national development and social stability, emphasizing collective interests and sustainable development; German policies value social welfare and environmental protection, pursuing balanced development; UK policies encourage free trade and technological innovation; French policies emphasize social solidarity and cultural protection, with moderate government intervention in

the economy; Japanese policies pursue economic stability and technological innovation while maintaining social harmony; Russian policies emphasize national sovereignty and economic autonomy, ensuring social stability; Indian policies promote economic development and social justice, respecting cultural diversity; Canadian policies emphasize multiculturalism and environmental protection; Brazilian policies focus on social inclusion and sustainable development. The choices of these policies and regulations demonstrate each country's adherence to and practice of its own development path and values in the context of globalization. Table 3.10 below shows typical countries' administrative regulations and policies and their underlying basic values.

Table 3.10 Typical Countries' Administrative Regulations and Policies and Their Underlying Basic Values

Country	Core Values				
United States	Free market (Breyer et al., 2022)	Individual freedom (Breyer et al., 2022; Goodnow, 2003)	Efficiency (Breyer et al., 2022; Goodnow, 2003)	Innovation (Breyer et al., 2022; Goodnow, 2003)	
China	National development (Liu et al., 2011)	Social stability (Guo & Zhang, 2015; Wang, 2015)	Collective interests (Ma & Ortolano, 2000)	Sustainable development (Ma & Ortolano, 2000)	
Germany	Social welfare (Künnecke, 2007)	Environmental protection (Künnecke, 2007)	Economic balance (Künnecke, 2007)		
United Kingdom	Free trade (Künnecke, 2007; Wade & Forsyth, 2014)	Market openness (Künnecke, 2007; Wade & Forsyth, 2014)	Innovation-driven (Künnecke, 2007; Wade & Forsyth, 2014)		
France	Social solidarity (Suleiman, 2015)	Cultural protection (Brown et al., 1998; Suleiman, 2015)	Economic intervention (Brown et al., 1998; Suleiman, 2015)		
Japan	Economic stability (Lebo, 2018)	Technological innovation (Lebo, 2018)	Social harmony (Lebo, 2018)		
Russia	National sovereignty (Hønneland & Jørgensen, 2003; et al., 2016; Studin, 2017)	Economic autonomy (Metelev et al., 2016; Studin, 2017)	Social stability (Studin, 2017)		

	Studin, 2017)			
India	Economic development (Munshi & Abraham, 2004)	Social justice (Munshi & Abraham, 2004)	Cultural diversity (Munshi & Abraham, 2004)	
Canada	Multiculturalism (Tuohy & Tuohy, 2010)	Social inclusion (Tuohy & Tuohy, 2010)	Environmental protection (Tuohy & Tuohy, 2010)	

Different countries have different value choices for administrative regulations and policies. The values in Table 3.10 can be summarized into the following five dimensions: "Individual Freedom and Social Stability", "Efficiency and Environmental Protection", "Innovation and Cultural Protection", "Social Welfare and Economic Development" and "Multiculturalism and Homogeneous Culture". Among them, the "Individual Freedom and Social Stability" dimension reflects the possible divergence in value choices between individual freedom and social stability among different groups. Groups that prioritize "Individual Freedom" emphasize its importance, advocating for reduced government intervention and protection of individual rights and freedom of choice; groups that value "Social Stability" more highly place greater emphasis on social stability, believing that strong regulations and policies are needed to maintain order and prevent chaos and conflict. In the "Efficiency and Environmental Protection" dimension, groups that value "Efficiency" focus on maximizing efficiency and economic benefits, supporting rapid development, which may come at the expense of environmental protection; groups that place greater emphasis on "Environmental Protection" consider it a priority, advocating for sustainable development and emphasizing the protection of natural resources and ecosystems in development. The "Innovation and Cultural Protection" dimension is a value choice between innovation and cultural protection, with groups that consider "Innovation" more important tending to support the innovative development of technology and industry to promote social progress and economic growth; groups that value "Cultural Protection" more highly believe in preserving and passing on cultural heritage, maintaining the cultural diversity and identity of society. In the trade-off of the "Social Welfare and Economic Development" dimension, groups that focus more on "Social Welfare" advocate for the government to provide more social security and public services to reduce social inequality and improve the quality of life for the public; groups that choose "Economic Development" focus more on economic growth and competitiveness, considering economic growth as the key to improving the overall welfare of society. In the "Multiculturalism and Homogeneous Culture" dimension, the value choices of different groups reflect different views on cultural identity and integration. Groups that identify

with "Multiculturalism" advocate for the inclusiveness of multiculturalism, supporting the equal coexistence of different cultures, languages, and religious beliefs; groups that tend to maintain "Homogeneous Culture" emphasize the promotion of mainstream culture and values, strengthening national identity and cultural homogeneity through education and media.

3.4 Level 3: Cultural Diversity

3.4.1 Language

Language is not only a tool for communication but also a carrier of culture and values. Different types of languages and their usage are often associated with specific cultural values, reflecting the characteristics of different language types and the cultural values they embody. High-context languages focus on conveying information through context and non-verbal cues, emphasizing social harmony and interpersonal relationships; low-context languages rely on the language itself to express intentions clearly, valuing direct communication and individual expression. Polyphonic languages demonstrate collective cooperation through multiple voices, tonal languages distinguish meaning through tones, emphasizing precision and emotional nuance. Inflectional languages express grammar through word changes, emphasizing regularity and logic; analytic languages use word order and auxiliary words, emphasizing conciseness and practicality. Agglutinative languages use affixes to express grammar; polysynthetic languages express rich meanings through word changes, both emphasizing precision and creativity. Each type of language reflects the cultural characteristics and values of the society that uses it and understanding these can help us gain a deeper understanding of the behavioral patterns and ways of thinking in different cultures and societies. The following Table 3.11 shows typical language cultures and their underlying core values.

Table 3.11 Typical Language Cultures and Their Underlying Core Values

Language Type	Core Values			
High-context languages (Chinese, Japanese)	Implicitness (Broeder, 2021)	Harmony (Kim et al., 1998)	Respect (Kim et al., 1998)	Collectivism (Dozier et al., 1998)
Low-context languages (English, German)	Directness (Broeder, 2021)	Clarity (Gudykunst & Nishida, 1986)	Efficiency (Gudykunst & Nishida, 1986)	Individualism (Dozier et al., 1998)
Tonal languages (Chinese, Thai)	Delicacy (Remijnsen, 2016)	Precision (Eady, 1982)	Emotional Expression	

			(Eady, 1982; Remijsen, 2016)
Inflectional languages	Regularity (Mirković et al., 2011; Ziarek, 1994)	Logic (Ziarek, 1994)	Order (Ziarek, 1994)
Analytic languages (Chinese, Vietnamese)	Conciseness (DeFrancis, 1986; Horn & Pham, 2004)	Practicality (DeFrancis, 1986)	Flexibility (DeFrancis, 1986)
Agglutinative languages (Turkish, Japanese)	Precision (Loglo, 2010; Wu, 2023; Yilahun et al., 2017)	Detail Orientation (Loglo, 2010; Wu, 2023; Yilahun et al., 2017)	Flexibility (Loglo, 2010; Yilahun et al., 2017)
Polysynthetic languages (Some African)	Diversity (Alexander, 2003; Pérez, 2015)	Richness (Alexander, 2003)	Creativity (Alexander, 2003; Pérez, 2015)

Language use in different cultural backgrounds demonstrates the various preferences and emphasis on communication methods, expression habits, and communication norms among different groups. In Table 3.11, the values in language culture can be divided into six dimensions: "Directness and Implicitness", "Collectivism and Individualism", "Emotional Expression and Logic", "Precision and Flexibility", "Conciseness and Richness", and "Regularity and Creativity". In the dimension of "Directness and Implicitness" groups that tend to direct and clear communication are usually associated with values that emphasize honesty and efficiency; groups that choose "implicitness" value more implicit expressions, reflecting the group's communication philosophy of avoiding conflict and pursuing harmony. The dimension of "Collectivism and Individualism" reveals different groups' views on the relationship between the individual and the collective, with groups that value "collectivism" emphasizing group interests and unity in language use; groups that value "individualism" believe that personal freedom and independence should be highlighted in communication. In the dimension of "Emotional Expression and Logic" groups that value "emotional expression" encourage the direct expression of emotions; groups that focus more on "logic" pay attention to the logic and objectivity of language expression. In the dimension of "Precision and Flexibility" groups that emphasize "precision" pursue precision and clarity in language expression, while groups that value "flexibility" place more emphasis on the flexibility and adaptability of language. In the dimension of "Conciseness and Richness" groups that tend to "conciseness" will be more

succinct in language expression, while groups that advocate "richness" prefer rich and detailed language expression. In the dimension of "Regularity and Creativity" groups that value "regularity" emphasize adherence to language rules and traditions, while groups that value "creativity" encourage innovation and personalized expression.

3.4.2 Traditional Customs

Traditional customs in different cultures reflect their unique values and social characteristics. Chinese culture emphasizes family reunion, respect for elders, historical commemoration, and social harmony; Indian culture is characterized by light and hope, color and joy, religious beliefs, family, and community; Japanese culture admires the beauty of nature, respect and etiquette, and spirituality. American culture highlights gratitude and sharing, freedom and independence, as well as entertainment and humor; French culture centers on patriotism and freedom, art and aesthetics, socializing and enjoyment; Russian culture values victory and glory, religious beliefs, and spring and hope; African culture emphasizes community and unity, history and heritage, as well as nature and spirit; Middle Eastern culture is marked by religious beliefs, patience and self-discipline, community and unity; Nordic culture focuses on nature and light, family and warmth, history and courage; Australian culture highlights patriotism and freedom, cultural diversity, sports and community. These traditional customs not only enrich their own cultural life but also promote the diversity and exchange of global culture. The following Table 3.12 shows typical traditional customs and their underlying core values.

Table 3.12 Typical Traditional Customs and Their Underlying Core Values

Traditional Custom	Core Values					
Chinese Traditional Customs	Family Reunion (Li, 2018; Stafford, 2014)	Respect for Elders (Chou, 1993; Li, 2018)	Historical Commemoration (Li, 2018; Stafford, 2014)	Social Harmony (Li, 2018)		
Indian Traditional Customs	Light and Hope (Krafft & Choubisa, 2018)	Color and Joy (Ayyar, 2001)	Religious Beliefs (Ayyar, 2001)	Family and Community (Ayyar, 2001)		
Japanese Traditional Customs	Beauty of Nature (Shively, 2015)	Respect and Etiquette (Shively, 2015)	Spirituality (Shively, 2015)			
American Traditional Customs	Gratitude and Sharing (Shearer, 2007)	Freedom and Independence (Shearer, 2007)	Entertainment and Humor (Shearer, 2007)			

French Traditional Customs	Patriotism and Freedom (Haine, 2006)	Art and Aesthetics (Haine, 2006)	Socializing and Enjoyment (Haine, 2006)
Russian Traditional Customs	Victory and Glory (Cuddihy & King, 2021)	Religious Beliefs (Cuddihy & King, 2021)	Hope (Cuddihy & King, 2021)
African Traditional Customs	Community and Unity (Blyden, 1994)	History and Heritage (Blyden, 1994)	Nature and Spirit (Blyden, 1994)
Middle Eastern Traditional Customs	Religious Beliefs (Lindholm, 2008)	Patience and Self-Discipline (Lindholm, 2008)	Community and Unity (Lindholm, 2008)
Nordic Traditional Customs	Nature and Light (Simonyi & Cagan, 2017)	Family and Warmth (Simonyi & Cagan, 2017)	History and Courage (Simonyi & Cagan, 2017)
Australian Traditional Customs	Patriotism and Freedom (Offord et al., 2014)	Cultural Diversity (Offord et al., 2014)	Sports and Community (Offord et al., 2014)

The value choices for traditional customs in different cultures reflect the different emphases of various groups on social structure, behavioral norms, attitudes towards life, and spiritual beliefs. As shown in Table 3.12, the values behind the traditional customs of different cultures can be divided into six dimensions: "Family Reunion and Individual Freedom," "Etiquette and Entertainment", "Social Harmony and Conflict", "Optimism and Patience", "Religious Beliefs and Secularism" and "Heritage and Innovation". In the dimension of "Family Reunion and Individual Freedom," groups that value "family reunion" emphasize the close connection and interdependence among family members, reflecting the importance of family values and social cohesion; groups that value "individual freedom" more highly prioritize personal autonomy and independence, encouraging individuals to pursue self-expression and personal development. The dimension of "Etiquette and Entertainment" reveals different groups' views on behavioral norms, with groups valuing "etiquette" emphasizing the importance of manners and rules to maintain social order and respect for tradition; groups valuing "entertainment" more highly see humor and entertainment as a seasoning in life, enhancing social interaction and alleviating life's pressures. In the dimension of "Social Harmony and Conflict" groups valuing "social harmony" tend to seek harmony and consensus, avoiding open conflict, reflecting the pursuit of collective interests and social stability; groups choosing "conflict" may be more accepting of conflict as a way of social progress and personal

expression. In the dimension of "Optimism and Patience" groups leaning towards "optimism" encourage a positive attitude towards life and optimistic expectations for the future, reflecting the importance of individual proactivity and the possibility of change; groups leaning towards "patience" more highly emphasize patience and adaptation as a strategy for facing life's challenges and uncertainties. In the dimension of "Religious Beliefs and Secularism" groups valuing "religious beliefs" regard religious faith as the core of life and the foundation of values, reflecting reverence for supernatural forces and spiritual pursuits; groups supporting "secularism" lean more towards a secular worldview, emphasizing rationality and science. In the dimension of "Heritage and Innovation," groups valuing "heritage" prioritize the preservation and transmission of tradition, respecting history and cultural heritage, while valuing "innovation" may encourage innovation and transformation of traditional culture.

3.4.3 Art

Art forms from different cultural backgrounds showcase their unique aesthetic concepts and cultural values. Chinese art emphasizes inner beauty, harmony, and heritage, focusing on spiritual harmony with nature; Japanese art pursues simplicity, etiquette, and the beauty of the moment, reflecting the beauty of the ephemeral and the transience of life; Indian art is known for its divinity, rich colors, and narrative, reflecting religious beliefs and philosophy of life; American art highlights innovation, diversity, and entertainment, showcasing the multicultural society and the spirit of innovation; French art is characterized by aesthetics, elegance, and innovation, demonstrating the refinement and forward-looking nature of art; Russian art is profound, educational, and grand, reflecting the depth of social history and thought; African art emphasizes community, spirituality, and tradition, embodying reverence for nature and ancestors. Middle Eastern art is characterized by religion, geometric patterns, and harmony, showcasing religious beliefs and cultural order. Nordic art is known for its simplicity, nature, and human care, emphasizing harmonious coexistence with nature; Australian art is characterized by diversity, nature, and innovation, reflecting multiculturalism and respect for the natural environment. Hindu art is marked by sacredness, decorativeness, and tradition, showcasing religious beliefs and the decorative beauty of art. These art forms not only enrich the diversity of world art but also reflect the unique perspectives and profound connotations of their respective cultures. The following Table 3.13 shows typical arts and their underlying core values.

Table 3.13 Typical Arts and Their Underlying Core Values

Art	Core Values			
Chinese Art	Inner Beauty	(Wu, 2019)	Harmony (Wu, 2019)	Heritage (Wu, 2019)
Japanese Art	Simplicity	(Graham, 2014)	Etiquette (Blyden, 1994)	Fleeting Beauty (Addiss et al., 2006)
Indian Art	Divinity	(Havell, 1920)	Color (Coomaraswamy, 1923; Havell, 1920)	Narrative (Coomaraswamy, 1923; Havell, 1920)
American Art	Innovation	(Jensen, 2002)	Diversity (Jensen, 2002)	Entertainment (Jensen, 2002)
French Art	Aesthetics	(Caballero, 2004)	Elegance (Leith, 1964)	Innovation (Caballero, 2004)
Russian Art	Profundity	(Newmarch, 1916; Tarasov, 2012)	Education (Newmarch, 1916; Tarasov, 2012)	Grandeur (Newmarch, 1916; Tarasov, 2012)
African Art	Community	(Herskovits, 1967)	Spirituality (Herskovits, 1967)	Tradition (Herskovits, 1967)
Middle Eastern Art	Religion	(Gregg, 2005; Zuhur, 1998)	Geometric Patterns (Gregg, 2005)	Harmony (Zuhur, 1998)
Nordic Art	Simplicity	(Hylland & Bjurström, 2018)	Nature (Hylland & Bjurström, 2018)	Humanity (Hylland & Bjurström, 2018)
Australian Art	Diversity	(Sayers, 2001)	Nature (Sayers, 2001)	Innovation (Sayers, 2001)

The choice of values behind art by different groups reflects the different emphases on aesthetic expression, cultural heritage, innovative spirit, and social criticism. As shown in Table 3.13, the multidimensionality of values behind art can be summarized in six aspects: "Tradition and Modernity", "Spirit and Material", "Inner and Outer", "Harmony and Conflict", "Nature and Artificial" and "Profundity and Superficiality". In the dimension of "Tradition and Modernity" the value of "Tradition" emphasizes the inheritance of historical and cultural heritage in art, reflecting respect and maintenance of the past; groups with the value of "Modernity" pursue innovation and contemporary expression, reflecting social and technological changes. In the dimension of "Spirit and Material" groups that focus on "Spirit" will explore the deep layers of human emotions and spiritual pursuits with spiritual art; groups that focus on "Material" itself will choose material art to pay attention to sensory experiences and formal beauty. In the dimension of "Inner and Outer" "Inner" artistic expression focuses on the artist's personal emotions and inner world; "Outer" artistic expression focuses on observing and commenting on broader social and natural phenomena. In the dimension of "Harmony and Conflict" groups pursuing the value of "Harmony" will pursue harmony in form and content;

groups pursuing "Conflict" will stimulate the audience's emotional response through the expression of conflict and tension. In the dimension of "Nature and Artificial" groups that focus on "Nature" will tend to use natural art to praise the beauty and power of nature; groups advocating "Artificial" will showcase human creation and technological achievements through artificial art. In the dimension of "Profundity and Superficiality" groups pursuing the value of "Profundity" seek profound themes and complex connotations through art; groups pursuing the value of "Superficiality" believe that intuitive art is more attractive and entertaining.

3.4.4 Literature

Literature in different cultures reflects their unique ideological concepts and social values. Chinese literature is deeply influenced by Confucian thought, emphasizing benevolence, righteousness, propriety, wisdom, and faithfulness, while integrating Taoist concepts of natural harmony and inner peace, as well as critiques of social injustice; Japanese literature is characterized by the aesthetics of mono no aware and Zen Buddhist thought, showcasing individualism and an appreciation for nature and life; Indian literature blends religion and mythology, exploring social issues and personal struggles; American literature highlights individualism, social critique, and the blending of multiculturalism; French literature is marked by rationality and enlightenment, aesthetic pursuit, and social critique; Russian literature centers on profound reflection, religious and philosophical discussions, and nationalism; African literature emphasizes community solidarity, historical traditions, and social critique; Middle Eastern literature is characterized by religious beliefs, nationalism, and social critique; Nordic literature focuses on the simplicity of nature, social critique, and humanistic care; Australian literature is known for multiculturalism, environmental consciousness, and social critique. These literary works are not only expressions of art but also profound reflections and explorations of human experience, social issues, and cultural values. The following Table 3.14 shows typical literature and their underlying core values.

Table 3.14 Typical Literature and Their Underlying Core Values

Literature	Core Values				
Chinese Literature	Confucian	Thought	Taoist	Thought	Social Critique (Idema & Haft, 1997)
	(benevolence, righteousness, propriety, faithfulness) (Idema & Haft, 1997)	wisdom,	(natural harmony, inner peace) (Idema & Haft, 1997)		

Japanese Literature	Mono no aware Aesthetics (Aston, 1899; Keene, 1988)	Zen Buddhist Thought (Aston, 1899; Keene, 1988)	Individualism (Aston, 1899; Keene, 1988)
Indian Literature	Religion and Mythology (Datta, 1988)	Social Critique (Datta, 1988)	Personal Struggle (Datta, 1988)
American Literature	Individualism (Shumway, 1994)	Social Critique (Shumway, 1994)	Multiculturalism (Shumway, 1994)
French Literature	Rationality and Enlightenment (Finch, 2010)	Aesthetic Pursuit (Finch, 2010)	Social Critique (Finch, 2010)
Russian Literature	Profound Reflection (Kropotkin, 1905, 1915)	Religion and Philosophy (Kropotkin, 1905, 1915)	Nationalism (Kropotkin, 1905, 1915)
African Literature	Community and Solidarity (Amuta, 2017)	History and Tradition (Amuta, 2017)	Social Critique (Amuta, 2017)
Middle Eastern Literature	Religious Beliefs (Gregg, 2005)	Nationalism (Gregg, 2005)	Social Critique (Gregg, 2005)
Nordic Literature	Nature and Simplicity (Haakonssen & Olden-Jørgensen, 2017)	Social Critique (Haakonssen & Olden-Jørgensen, 2017)	Humanistic Care (Haakonssen & Olden-Jørgensen, 2017)
Australian Literature	Multiculturalism (Huggan, 2007)	Environmental Consciousness (Huggan, 2007)	Social Critique (Huggan, 2007)

The value choices behind literature reflect different groups' varying emphases on historical heritage, contemporary concerns, spiritual pursuits, material life, collective interests, individual freedom, inner emotions, outward actions, harmonious concepts, conflict resolution, aesthetic experiences, and practical functions. The values of literature can be categorized into six dimensions: "Tradition and Modernity", "Spirit and Material", "Collective and Individual", "Inner and Outer", "Harmony and Conflict" and "Aesthetics and Practicality". On the dimension of "Tradition and Modernity" some literary works emphasize the inheritance and preservation of historical culture, reflecting respect and maintenance of the past; while other works focus on contemporary social issues and modern life experiences, exploring the impact of new technologies and new ideas. On the dimension of "Spirit and Material" some works delve into the human spiritual world, religious beliefs, and moral ethics, while others focus on material life, economic development, and social structure. In the dimension of "Collective and

Individual" some literary works emphasize collective interests, social unity, and group identity, while others focus on personal experiences, self-actualization, and individual expression. On the dimension of "Inner and Outer" literary works sometimes depict the inner world, emotions, and thoughts of characters, and sometimes focus on their outward actions, social interactions, and environmental factors. In the dimension of "Harmony and Conflict" some literary works pursue the depiction of a harmonious society and the ideal state of harmonious coexistence between humans and nature, while others explore the role of conflict in personal growth and social change by showing social contradictions, personal struggles, and conflict resolution. On the dimension of "Aesthetics and Practicality" literary works often pursue aesthetic value, providing aesthetic pleasure and artistic enjoyment, and carrying practical functions, such as education, inspiration, and entertainment. These value dimensions are reflected in different ways across cultures and literary schools, influencing the themes, forms, and audience reception of literary works, and forming the basis of literary diversity and richness.

3.4.5 Cuisine

Cuisine in different cultures reflects their unique lifestyles and social values. Chinese cuisine emphasizes family reunion, health and nourishment, and the diversity of dishes; Japanese cuisine is characterized by simplicity, naturalness, exquisite beauty, and seasonality, pursuing harmony between food and nature; Indian cuisine is deeply influenced by religious beliefs, rich in color, and emphasizes community sharing; American cuisine focuses on efficiency and convenience while offering a variety of personal choices. French cuisine centers on refinement, elegance, aesthetic enjoyment, and social interaction, while Italian cuisine is marked by simplicity, family traditions, and regional characteristics; Russian cuisine is abundant, with a deep historical tradition and social culture; Mexican cuisine is known for its color and flavor, combining community sharing and historical traditions; African cuisine emphasizes community and unity, history and tradition, as well as nature and sustainability; Middle Eastern cuisine is influenced by religious norms, focusing on socializing and sharing, with a rich flavor; Mediterranean cuisine aims for health and longevity, simplicity and naturalness, and emphasizes socializing; Southeast Asian cuisine is characterized by flavor and balance, diversity, and community sharing. These food cultures not only reflect the ingredients and culinary skills of each region but also show people's emotional attachment to food and ways of social interaction. The following Table 3.15 shows typical food cultures and their underlying core values.

Table 3.15 Typical Food Cultures and Their Underlying Core Values

Cuisine	Core Values			
Chinese Cuisine	Family (Chee-Beng, 2012; Liu, 2011)	Reunion	Health and Nourishment (Chee-Beng, 2012; Liu, 2011)	Diversity (Chee-Beng, 2012; Liu, 2011)
Japanese Cuisine	Simplicity and Naturalness (Ashkenazi & Jacob, 2003)		Exquisite and Aesthetic (Ashkenazi & Jacob, 2003)	Seasonality (Ashkenazi & Jacob, 2003)
Indian Cuisine	Religious Beliefs (Sen, 2004)		Rich Color (Sen, 2004)	Community Sharing (Sen, 2004)
American Cuisine	Efficiency and Convenience (Wallach, 2013)		Diversity (Wallach, 2013)	Personal Choice (Wallach, 2013)
French Cuisine	Refined and Elegant (Ferguson, 2006)		Aesthetic Enjoyment (Ferguson, 2006)	Socializing (Ferguson, 2006)
Italian Cuisine	Simplicity (Parasecoli, 2004)		Family Tradition (Parasecoli, 2004)	Regional Characteristics (Parasecoli, 2004)
Russian Cuisine	Abundance (Glants & Toomre, 1997)		Historical Tradition (Glants & Toomre, 1997)	Socializing (Glants & Toomre, 1997)
Mexican Cuisine	Color and Flavor (Long-Solís & Vargas, 2005)		Community Sharing (Long-Solís & Vargas, 2005)	Historical Tradition (Long-Solís & Vargas, 2005)
African Cuisine	Community and Unity (Flynn, 2016)		History and Tradition (Flynn, 2016)	Nature and Sustainability (Flynn, 2016)
Middle Eastern Cuisine	Religious Norms (Heine, 2004)		Socializing and Sharing (Heine, 2004)	Rich Flavor (Heine, 2004)
Mediterranean Cuisine	Health and Longevity (Helstosky, 2009)		Simplicity and Naturalness (Helstosky, 2009)	Socializing (Helstosky, 2009)
Southeast Asian Cuisine	Flavor and Balance (Chee-Beng, 2012)		Diversity (Chee-Beng, 2012)	Community Sharing (Chee-Beng, 2012)

The choice of values behind food culture by different groups reflects the emphasis on different food concepts such as family and community, health and taste, simplicity and refinement. The values of food culture in Table 3.15 can be divided into six dimensions: "Family Tradition and Community Sharing", "Health and Flavor", "Simplicity and Refinement", "Tradition and Modernity", "Regional Characteristics and Global Diversity" and "Nature and Artificial". In the dimension of "Family Tradition and Community Sharing" groups that value

"family tradition" emphasize the importance of food in family gatherings and festive celebrations, as well as its role in maintaining family bonds and passing on food culture; groups that value "community sharing" focus on the role of food in promoting community unity and social interaction. In the dimension of "Health and Flavor" groups that value "health" emphasize the nutrition and health benefits of food, advocating a balanced diet and a healthy lifestyle; groups that value "flavor" pursue the taste and culinary art of food, highlighting the sensory enjoyment brought by delicious food. In the dimension of "Simplicity and Refinement" groups that emphasize "simplicity" tend to choose simple and practical food preparation and consumption methods, reflecting an advocacy for a simple life and sustainability; groups that value "refinement" pursue the meticulous preparation and exquisite presentation of food, showing a pursuit of culinary aesthetics and quality of life. In the dimension of "Tradition and Modernity" groups that value "tradition" tend to respect and preserve traditional ingredients and cooking skills; groups that value "modernity" is more willing to explore modern ingredients and innovative cooking methods. In the dimension of "Regional Characteristics and Global Diversity" groups that value "regional characteristics" emphasize the importance of local ingredients and traditional dishes; groups that appreciate "global diversity" value the fusion of global ingredients and diverse flavors. In the dimension of "Nature and Artificial" groups that prefer "natural" values advocate natural ingredients and traditional agriculture; groups that are more open to "artificial" food values have an optimistic attitude towards food processing and improvement, accepting the use of technology to improve the accessibility and diversity of food.

3.5 Level 4: Universal Moral Ethics

3.5.1 Peace

Peace is a stable condition where nations, groups, and individuals are free from conflict or severe hostile acts, and in the face of disagreements, they are more inclined to resolve issues through nonviolent means. The value of peace relies on a deeper set of values to support and achieve it. Respect and equality form the foundation of peace, ensuring that everyone is treated fairly and has the right to express their thoughts; tolerance and nonviolence reduce the occurrence of conflict by solving problems through mutual understanding rather than confrontation; justice and security are the safeguards of peace, allowing everyone to live and grow in an environment free from fear; relationships between people are fostered through cooperation and empathy, which promote peace; freedom and sustainable development ensure the longevity and inclusiveness of peace, enabling everyone to develop freely and fully;

education and dialogue are effective methods for resolving disagreements and finding common ground, serving as a bridge to peace; responsibility motivates individuals and communities to think ahead and act in ways that promote peace. These values are interwoven, together forming a solid foundation for peace, which is realized and maintained through more specific actions.

3.5.2 Development

Development is a complex and comprehensive process that relates to economic growth, social progress, and the enhancement of human well-being. Since the concept of development is broad, like peace, it is necessary to understand the values that underpin the value of development to better achieve it. The values behind development include fairness, sustainability, inclusiveness, human rights, education, health, freedom, responsibility, innovation, cooperation, transparency, diversity, justice, environmental protection, and economic stability. Fairness ensures that individuals and groups have equal opportunities and rights in development, bear equal responsibilities, and avoid the social pressures and conflicts caused by inequality; sustainability emphasizes that development should not be pursued without regard for environmental impact and future responsibility; inclusiveness allows people from different social and cultural backgrounds to participate in development, enriching social diversity and enhancing societal creativity; diversity is the guarantee of societal innovation, encouraging the exchange of different viewpoints and cultures; respect for human rights provides everyone with basic dignity and freedom, which is the foundation of social development; emphasis on education and health supports the improvement of individuals in all aspects; freedom encourages individuals to pursue innovation and development; innovation is the main driving force of development, bringing new vitality to it; cooperation is the necessary path to achieving group goals, achieving greater development objectives through collective efforts; transparency ensures open information and fair decision-making, increasing people's participation and enhancing social trust; justice ensures the fairness of laws and the correctness of morals, providing a fair competitive environment for society; environmental protection is the key to balancing development and ecology, and it is a guarantee of sustainable development goals; economic stability is the continuous driving force of social development, and social development is limited when economic fluctuations bring too much uncertainty. These values together support the goal value of development, constituting a comprehensive, balanced, and sustainable development model.

3.5.3 Equity

Achieving equity requires a consistent approach to resource distribution, rights protection,

and opportunity provision. The values that support the realization of equity include justice, equality, respect, responsibility, transparency, participation, inclusiveness, reciprocity, cooperation, human rights, equal opportunity, compensation, sustainability, education, health, and economic stability. Justice, as the core of equity, ensures that everyone's rights and needs are treated and met fairly; equality in social and legal status emphasizes that everyone should be treated without discrimination; equal opportunity emphasizes the need to ensure that everyone has an equal starting point and competitive opportunities; mutual respect is a prerequisite for achieving equity, requiring us to respect and recognize the value of others; responsibility means that individuals and groups should not only care about their own interests but also be responsible for others and the overall society, maintaining an equitable environment; transparency ensures the availability of information and the openness of decision-making processes, allowing people to supervise various decision-making processes and outcomes, ensuring the realization of equity; participation encourages and supports everyone in society to express their opinions and suggestions in social decision-making, thereby safeguarding the interests of different groups; inclusiveness represents the society's acceptance of different cultural and social background perspectives, creating conditions for further equity; reciprocity and cooperation emphasize seeking common development on the basis of mutual respect and support, which is conducive to establishing equitable cooperative relationships; protecting human rights further ensures the realization of equity, making sure everyone has basic freedoms and rights, laying the foundation for equity; compensation mechanisms can correct historical and current unfair phenomena; sustainability emphasizes the long-term balance of society and the environment, providing continuous assurance for the realization of equity; education provides more opportunities for everyone to participate in social decision-making by enhancing capabilities; health and economic stability are also key to ensuring equity.

3.5.4 Justice

Justice is at the core of social order and ethical standards. Justice requires fair, just, and equal treatment of all individuals on legal and moral levels. The values that support the realization of justice include equality, fairness, respect, responsibility, integrity, freedom, reciprocity, inclusiveness, compassion, protection, rule of law, and transparency. Equality is the foundation of justice, ensuring that everyone has an equal status in society, providing the prerequisite for fair treatment; fairness further emphasizes the justice in the redistribution of social resources and opportunities, preventing the occurrence of injustice; respect is the manifestation of justice, requiring recognition of the value and rights of different individuals;

responsibility requires that in the pursuit of personal and group interests, one also assumes social responsibilities; integrity is the safeguard of justice, and sincerity in words and deeds is the foundation for establishing more solid relationships; freedom is part of justice, allowing people to express their views and pursue their goals without infringing on the rights of others; reciprocity requires mutual benefit in social relationships, which helps to establish fair and just social relationships; inclusiveness allows for the existence of different viewpoints and behaviors in society; compassion is the understanding and care for the plight of others, promoting the realization of social justice; protection is special attention to vulnerable individuals and groups, ensuring that their interests are not violated; the rule of law is the means of achieving justice, ensuring social justice through the formulation and implementation of laws and regulations; transparency ensures the availability of information and openness of decision-making processes, reducing the likelihood of unfair treatment.

3.5.5 Democracy

Democracy is a political system and way of life that emphasizes active participation of people in policymaking and autonomy in living. To achieve the value of democracy, support from equality, freedom, justice, responsibility, inclusiveness, transparency, participation, rule of law, efficiency, and education is needed. Equality ensures that people have the same rights and opportunities in the political process, providing the foundation for democracy; freedom allows individuals to express opinions and participate in political activities; justice requires that laws and policies be fair and unbiased, ensuring the fairness of democracy; responsibility means that as political participants, people must also be accountable for their actions and decisions, maintaining the integrity of democracy; inclusiveness encourages different viewpoints and interest groups to participate in the decision-making process, enhancing the diversity of democracy; transparency ensures openness in political processes and accessibility of information, reducing the possibility of corruption and abuse of power; participation is the core of democracy, encouraging citizens to actively engage in political life; the rule of law provides a stable framework for democracy, ensuring fair implementation and adherence to laws; efficiency requires government and political processes to operate efficiently to meet the needs and expectations of citizens; education is key to cultivating awareness and participation capabilities, providing knowledge and skills for the continuous development of democracy.

3.5.6 Freedom

Freedom is the ability of individuals and groups to express, choose, and act, and it is one of the core values of a democratic society. To achieve the value of freedom, a range of values

including individualism, equality, justice, responsibility, inclusiveness, diversity, privacy, education, rule of law, peace, and innovation is needed. Individualism emphasizes the importance of individual autonomy and self-actualization, providing the foundation for freedom; equality ensures that everyone has the same rights and opportunities in society, providing universality for freedom; justice protects the rights of individuals and groups from infringement, maintaining the fairness of freedom; responsibility requires individuals and groups to consider the impact of their actions on others and society when exercising freedom, ensuring its responsible exercise; inclusiveness allows different viewpoints and lifestyles to coexist, providing diversity and inclusiveness for freedom; diversity is the source of social creativity and progress, encouraging appreciation for different cultures and ideas; privacy protects individual space and information, safeguarding freedom from infringement; education enhances citizens' knowledge and critical thinking skills, providing capability and understanding for freedom; the rule of law protects individual rights through a fair legal system, providing institutional protection for freedom; peace provides a stable and secure environment for society, allowing freedom to be realized without violence and conflict; innovation drives social development and change, providing new possibilities and space for development for freedom.

3.6 Conclusion

In the development of human society, values, as the core of culture and thought, have always guided the direction of individual behavior and social progress. This chapter, through an in-depth analysis of the level of human value frameworks, aims to reveal how individuals and collectives rely on their internal beliefs and standards when facing choices, and how these values affect our perception of the world, evaluation of actions, and decision-making. We will explore how values provide clear direction and guidance for decision-makers at multiple levels and different dimensions, including political ideologies, economic models, social structures, religious beliefs, and philosophical thoughts, helping them make stable decisions in line with their personal beliefs in a complex and changing environment.

After analyzing the hierarchical structure of the human value system, the following conclusions are drawn:

(1) The characteristics of values are complexity and level. Values do not exist in isolation but form a multi-level, interrelated complex system. From personal beliefs at the individual level to legal norms at the national level, and to universal moral ethics at the global level, values

influence the operation and development of society at different levels.

(2) The impact of values on decision-making is profound. The decision-making processes of individuals and collectives are deeply influenced by values. In political, economic, and social fields, values not only provide direction for decision-makers but also become an important standard for evaluating actions and outcomes.

(3) The consistency between legal norms and values is evident. Norms within different legal systems reflect the value orientation of specific societies. From constitutions to civil and commercial law, criminal law, procedural law, and administrative regulations and policies, these norms embody society's pursuit of values such as justice, freedom, equality, and human rights.

(4) The value of cultural diversity is rich and colorful. Cultural expressions such as language, traditional customs, art, literature, and cuisine showcase the values within different cultural backgrounds. The diversity of these cultural value concepts is a precious treasure of human society, enriching our lives and promoting exchange and understanding between different cultures.

(5) Universal moral and ethical values such as peace, development, fairness, justice, democracy, and freedom are goals pursued by all of humanity. These value concepts transcend culture and national borders, becoming an important force in promoting global cooperation and sustainable development.

The analysis in this chapter not only provides a comprehensive perspective on the human value framework but also offers guidance on how to make decisions that align with one's values in a complex and changing world. It answers the research question RQ1, which is the question of how to help governments build a human values framework that LLM needs to align with? Understanding these value concepts is significant for guiding individual actions, solving social problems, promoting international cooperation, and achieving sustainable development. By deeply analyzing and understanding these value concepts, we can better grasp the direction of future development and contribute to building a more just, harmonious, and prosperous world. However, values are not static; they evolve with societal development and the changing times. Therefore, we need to continuously examine and update our values to adapt to the ever-changing world.

4 LLMs Value Alignment Experiment

4.1 Introduction

In Chapter 3, we explored the hierarchy of human value frameworks and analyzed how values in different cultural and social contexts affect decision-making. These theoretical frameworks provide us with a solid foundation for evaluating and understanding the performance of LLMs in value alignment. This chapter will build on these theoretical foundations and conduct a series of experiments to evaluate the performance of currently popular LLMs at different value levels and dimensions.

This section delves into the alignment of values between Large LLMs and human values. As the application of LLMs technology increases in daily life and work, the potential for its output results to conflict with human values arises, leading to social issues. Therefore, ensuring the alignment of LLMs values with human values has become particularly important. This study aims to evaluate the performance of three popular general LLMs—ChatGPT 4.0 developed by OpenAI, ERNIE Bot 3.5 developed by Baidu, and Kimi developed by Moonshot AI—across different value hierarchies and dimensions through experimental methods and analyzes their current attributes in value alignment.

The first part regarding the purpose and objects of the experiment introduces the design purpose of the experiment, which is to understand the extent of value alignment in different LLMs, and selects ChatGPT 4.0, ERNIE Bot 3.5, and Kimi as the subjects of the experiment. The reasons for choosing these models are their industry representativeness, diverse application scenarios, different development backgrounds and cultural diversity, and their large number of users, making the results of the experiment have high social value and practical significance.

The second part about dataset selection and processing describes the source of the dataset used in the experiment, namely the ValueEval task of SemEval-2023, as well as the process of data preprocessing. The dataset includes conclusion data from various sources, which, after classification and statistics, forms the conclusion dataset used for the experiment.

The third part is data selection, including data sources and data preprocessing, which describes in detail the sources of the data and the data preprocessing process to achieve the experimental purpose.

The fourth part is about the experimental process and results detail the process of the

experiment, including data preprocessing, the design of fixed prompts, and the execution of the experiment. The results are presented by statistically categorizing the frequency of deviation between the value choices of LLMs and human values, revealing the differences among different models in understanding and responding to human values. The experimental process and the conclusion of the experiment, summarizing the main findings of the experiment, including the performance of LLMs at different value hierarchies and dimensions, and their current attributes in value alignment. The results show that while the three leading LLMs perform relatively well in the field of universal ethics and morality, there are still significant deviations at the levels of ideology and law, which need to be addressed in future optimization work.

The fifth part analyzes the current value attributes of the three major models, further analyzing the value attributes of ChatGPT, ERNIE Bot, and Kimi at the three levels of ideology, legal norms, and cultural diversity, revealing their tendencies in different value dimensions through the results of the experiment.

The sixth part elaborates on the limitations of the experiment, discussing the limitations of the experiment, including the limitations of data quantity, data coverage, the limitations of LLMs in understanding complex values and ethical issues, as well as the limitations of experimental design and result interpretation. These limitations point out the directions that need to be improved and expanded in future research.

The seventh part is the summary of the section. This section evaluates the performance of three leading LLMs—ChatGPT 4.0, ERNIE Bot 3.5, and Kimi—in value alignment through a series of experiments. The results show that although these models basically meet the expectations of human value alignment in the field of universal ethics and morality, there are still significant deviations at the levels of ideology and legal norms. These findings emphasize the importance of further optimizing LLMs in these areas to ensure that their output results are more consistent with human values. The experiment also reveals the value judgment tendencies of LLMs in different cultural and social backgrounds, pointing out the differences between the model's performance in a specific cultural environment and the values of the local people. In addition, the limitations of the experiment, such as the limitations of data quantity and coverage, and the ability of the model to understand complex values and ethical issues, all indicate that future research needs to be improved in terms of data diversity, model understanding ability, cultural difference consideration, experimental design flexibility, and result interpretation depth.

The application of LLMs technology has greatly improved decision-making efficiency and service quality in people's lives and work, but it also brings certain risks. That is, the output results of existing LLMs may conflict with human values, leading to a series of social problems. Therefore, the alignment of LLMs values with human values is particularly important. Through the alignment of LLMs values, people try to guide LLMs to better serve human society and ensure the sustainable development of LLMs technology and even human society.

4.2 Experimental Purpose and Objects

Different LLMs, due to differences in training corpora and benchmarks, also vary in the degree to which they align with human values. This experiment is designed to further understand the extent of value alignment in popular existing LLMs. The subjects of this experiment are ChatGPT4.0 developed by OpenAI, ERNIE Bot 3.5 developed by Baidu, and Kimi developed by Moonshot AI. The reasons for choosing these three general LLMs are: firstly, all three models are representative of the industry, with ChatGPT representing the internationally advanced level of artificial intelligence research, while ERNIE Bot and Kimi represent China's research achievements in the field of LLMs; secondly, the three general LLMs have diverse application scenarios, which can better complete the content of this experiment; thirdly, the three general LLMs have different development backgrounds and cultural diversity, which may reflect more issues when dealing with different value issues; finally, these three models have a large number of users, and the results of the experiment will have higher social value and practical significance.

4.3 Dataset Selection

The dataset used in the experiment comes from the ValueEval task of SemEval-2023, which was collected and created by researchers from Bauhaus-Universität Weimar, Leibniz University Hannover, and Leipzig University. This dataset comes from various sources such as religious texts, political discussions, free text arguments, newspaper editorials, and online democratic platforms. The original dataset is labeled with 54 different values and mapped into 20 value categories. The 54 values in the data set were classified and, according to the method of searching similar words, they were searched one by one in the values of different value choices at the four value levels in the previous chapter, resulting in the following Table 4.1, which reclassified the 54 values into the value framework of this study.

Table 4.1 the 54 values in the value framework of this study

Original Text Value Hierarchy		Search Terms	Corresponding Values in This Study
level2	level1		
Self-orientation: Thought	Creativity	Innovation	10,13;21;17
	Curiosity		21;17
	Freedom of thought	Freedom	1,3,5,7,12,21;17
Self-orientation: Action	Autonomous goal setting	Autonomy	10,21;17
	Independence	Independence	3,12,21;17
	Autonomous action	Freedom	2,6,7,8,10,12,21;17
	Right to privacy	Individualism	3,11,14,21
Stimulation	Exciting life		None
	Diverse life	Diversity; Pluralism	4,5,10,12,15; 6,7,8,9,10,13,14
	Risk-taking	Courage	12;17
Hedonism	Pursuit of pleasure	Entertainment	12,13
Achievement	Ambition	Ideal	5;17
	Success	Victory	12;17
	Competence	Efficiency	2,3,7,9,10,11,15;17
	Wisdom	Wisdom	4;17
	Bravery	Bravery	12;17
Power: Dominance	Influence	Authority	1,3,5,8,9;17
	Command	Control; Decision-making	2; 2
Power: Resources	Wealth	Property	1,2;17
Face	Social recognition	Recognition (National identity); Trust	1; 7;17
	Good reputation		17
Security: Personal	Sense of belonging	Collective; Family; Community; Unity	1,2,3,10,11; 1,7,12,15; 4,12,13,14,15; 2,3,6,10,12,14,15
	Health	Health	15;17
	Debt-free	Economy	1,2,7,10;17

	Orderliness	Order	1,3,7,8,9,11;17
	Comfortable life	Enjoyment	12,15;17
Security: Social	National security	Nation	1,2,6,8,9,10;17
	Social stability	Social stability; Stability	6,10; 1,6,7,10;17
Tradition	Respect for tradition	Tradition	1,3,6,12,13,14,15
	Religious faith	Religion	4,12,13,14,15
Conformity: Rules	Rule-following	Rules (Laws are); Norms	3,11, (6-10) ; 8,9,15;17
	Self-discipline	Self-discipline	12;17
	Proper behavior	Etiquette	12,13
Conformity: Interpersonal	Politeness	Respect	11,12
	Respect for elders	Respect for elders; Respect	12
Humility	Humility		None
	Acceptance of life's conditions	Patience	12
Benevolence: Care	Helpfulness	Mutual aid	3;17
	Honesty		17
	Tolerance	Tolerance	4;17
	Family security	Family	1,7,12,15
	Loving-kindness	Love; Benevolence; Care	4; 4; 14
Benevolence: Reliability	Responsibility	Responsibility	1,3,5,7,8;17
	Loyalty to friends		None
Universal Values: Care	Pursuit of equality	Equality; Fairness	1,2,3,4,5,6,7; 1,2,6,18;17;19
	Justice	Justice	7,8,9;17;19
	World peace	Peace	4,6,16;17
Universal Values: Nature	Environmental protection	Environment	1,2,6,10,14;17

	Harmony with nature	Nature; Harmony	1,4,5,12,13,14,15; 4,5,7,8,10,11,12,13,14;17
	Beautiful world	Beauty	12,13,14,15;17
Universal Values: Tolerance	Open-mindedness	Tolerance	2,10;17
	Wise acceptance of others	Tolerance	4;17
Universal Values: Objectivity	Logical thinking	Logic; Objectivity; Rationality	5,11; 5; 5,14;17
	Objective view of things	Objectivity	5;17

For the purposes of Table 4.1, the value choices are numbered as shown in Table 4.2.

Table 4.2 No. of Value Selection Category

No.	Value Selection Category
1	Value choice
2	Political concept
3	Economic model
4	Social structure
5	Religious belief
6	Philosophical thought
7	Constitution and Basic Law
8	Civil and commercial law
9	Criminal law
10	Procedural law
11	Administrative policy
12	Language
13	Traditional customs
14	Art
15	Literature
16	Food
17	Peace
18	Development
19	Fairness
20	Justice
21	Democracy

The experiment uses the conclusion data from the a forementioned dataset, and according

to the experimental needs, the 20 value categories from the previous study are included into the four-level multi-dimensional human value framework constructed in the previous chapter, resulting in a total of 5383 conclusion data of 21 value choices. The specific classification is shown in Table 4.3 below.

Table 4.3 21 Value Choices and Corresponding Values

Value Level	Value Choice	Corresponding Values	Data Quantity
Ideology	Political Philosophy	Thoughtful, influential, socially recognized, belonging, orderly, national security, social stability, ensuring family safety, responsible, pursuing equality, in harmony with nature	272
	Economic Mode	Autonomy, capability, authority, wealth, debt-free, national security, pursuing equality, environmental protection, open-minded	308
	Social Structure	Freedom of thought, independent, privacy, influential, belonging, orderly, respect for tradition, rule-abiding, responsible, pursuing equality	240
	Religious Belief	Wise, belonging, religious, tolerant, loving, pursuing equality, world peace, in harmony with nature, wisely accepting others	148
	Philosophical Thought	Freedom of thought, diverse lifestyle, ambitious, influential, responsible, in harmony with nature, logical thinking, objective perspective	246
Legal Regulations	Constitutional Law	Autonomy, diverse lifestyle, belonging, national security, social stability, respect for tradition, rule-abiding, pursuing equality, world peace	266
	Civil and Commercial Law	Autonomy, diverse lifestyle, capability, socially recognized, debt-free, orderly, social stability, rule-abiding, ensuring family safety, responsible, pursuing equality, fairness, in harmony with nature	391
	Criminal Law	Diverse lifestyle, influential, orderly, national security, rule-abiding, responsible, fair, in harmony with nature	283
	Procedural Law	Capability, influential, orderly, national security, rule-abiding, fair	203
	Administrative Policy	Autonomous goal selection, autonomous action, diverse lifestyle, capability, belonging, debt-free, social stability, rule-abiding, environmental protection, in harmony with nature, open-minded	317
Cultural	Language	Privacy, capability, belonging, orderly, rule-abiding,	193

Diversity		polite, in harmony with nature, logical thinking	
	Traditional Customs	Freedom of thought, independent, diverse lifestyle, adventurous, happy, successful, brave, belonging, comfortable life, respect for tradition, religious, self-disciplined, well-behaved, polite, respecting elders, modest, ensuring family safety, in harmony with nature, beautiful world	295
	Art	Creative, diverse lifestyle, happy, belonging, respect for tradition, religious, well-behaved, in harmony with nature, beautiful world	134
	Literature	Diverse lifestyle, belonging, respect for tradition, religious, loving, in harmony with nature, beautiful world, logical thinking	154
	Food	Capability, belonging, healthy, comfortable life, respect for tradition, religious, ensuring family safety, in harmony with nature	196
Universal Ethics	Peace	World Peace	13
	Development	Creative, curious, free-thinking, autonomous goal selection, independent, autonomous action, adventurous, ambitious, successful, capable, wise, brave, influential, wealthy, socially recognized, good reputation, healthy, debt-free, orderly, comfortable life, national security, social stability, rule-abiding, self-disciplined, helpful, honest, tolerant, responsible, pursuing equality, fair, world peace, environmental protection, in harmony with nature, beautiful world, open-minded, accepting others wisely, logical thinking, objective perspective	985
	Fairness	Pursuing equality, fair	100
	Justice	Fair	54
	Democracy	Freedom of thought, autonomous goal selection, independent, autonomous action, privacy, diverse lifestyle, adventurous, ambitious, pursuing equality, fair, world peace, open-minded, accepting others wisely, logical thinking, objective perspective	428
	Freedom	Creative, curious, free-thinking, autonomous goal selection, independent, autonomous action, privacy	157

4.4 Experimental Process and Results

This experiment aims to understand the status of value alignment in three popular general LLMs. Through the process, the dataset is preprocessed, including removing duplicate

conclusions and correcting some value labels. The results of human value choices in the original study are statistically categorized. In the case of repeated questions, the mode is taken as the final answer. To achieve the purpose of the experiment and ensure that different prompts do not affect the results, this experiment only uses fixed prompts related to the human value framework. The prompts are designed to guide the LLMs to make value judgments under different levels and dimensions, for example: considering the political philosophy in ideology, only answer the following question with "support" or "do not support". Then, the content of the prompts is replaced with different levels and dimensions, while the other prompt content remains unchanged. After classifying and organizing the data, questions are asked to the three LLMs in the format of "prompt + conclusion data", and the LLMs are asked to answer with only "yes" or "no". The results of the LLMs are then compared with the results of human value choices in the original study.

The scoring criteria used in the experiment is that when the model's answer is "yes", the score is 1, and when the model's answer is "no", the score is 0. In the evaluation process, first collect a series of answer samples generated by each LLMs; then compare the model's answer results with the human answers in the original data set; finally, record the parts of the LLMs answers that are the same and different from the human answers. In the result analysis, the frequency of different answers from all LLMs models and human answers is calculated, that is, the situation that is different from human values.

The results are shown in the following Table 4.4, where the numbers represent the frequency of deviation between the values of the LLMs and human values in the value choices of each level and dimension.

Table 4.4 Results of three LLMs experiments

Value Level	Value Choice	ChatGPT Results	ERNIE Bot Results	Kimi Results
Ideology	Political Philosophy	0.51	0.46	0.46
	Economic Mode	0.47	0.45	0.41
	Social Structure	0.57	0.48	0.51
	Religious Belief	0.33	0.32	0.34
	Philosophical	0.53	0.43	0.48

	Thought			
Legal Regulations	Constitutional and Basic Law	0.44	0.42	0.37
	Civil and Commercial Law	0.43	0.45	0.38
	Criminal Law	0.55	0.49	0.38
	Procedural Law	0.59	0.50	0.35
	Administrative Regulations and Policies	0.45	0.42	0.39
Cultural Diversity	Language	0.51	0.47	0.46
	Traditional Customs	0.39	0.38	0.35
	Art	0.28	0.32	0.36
	Literature	0.37	0.36	0.32
	Diet	0.35	0.41	0.39
Universal Ethics	Peace	0.46	0.23	0.15
	Development	0.38	0.39	0.35
	Fairness	0.31	0.42	0.33
	Justice	0.28	0.31	0.33
	Democracy	0.45	0.42	0.43
	Freedom	0.46	0.43	0.41

Although the experimental results show that the performance of ChatGPT 4.0, ERNIE Bot 3.5 and Kimi in the field of general ethics and morality is in line with human value expectations, significant deviations at the level of ideology and legal norms have attracted our attention. These biases may stem from biases in the model training data, inherent biases in the model design, or an inadequate understanding of a specific cultural context. Future research is needed to further explore the root causes of these biases and explore how to reduce them through improved model training and design.

4.5 Experimental Conclusions

In this experiment, the deviation frequency of values between the ChatGPT, ERNIE Bot, and Kimi models and human values across different value hierarchies and dimensions was assessed. As shown in Table 6.3, the higher the value, the greater the deviation of the LLMs's values from human values. The results reveal the differences among the LLMs in understanding and responding to human values.

At the ideological level, in dimensions such as political philosophy, economic model, social structure, and philosophical thought, ChatGPT's experimental results have relatively high values of 0.51, 0.47, 0.57, and 0.53, respectively, indicating a certain deviation of ChatGPT's values from human values in these dimensions. ERNIE Bot has the lowest deviation frequency in the dimension of religious belief, at 0.32, but higher deviation frequencies in political philosophy and economic model dimensions, at 0.46 and 0.45, respectively. Kimi has a higher deviation in the social structure dimension at 0.51, comparable to ChatGPT, but performs slightly better in religious belief and philosophical thought.

In the legal and regulatory level, ChatGPT shows higher deviation frequencies in criminal law and procedural law, at 0.55 and 0.59, respectively, indicating a larger deviation in value identification in these legal fields. ERNIE Bot's performance across all dimensions in the legal and regulatory level is consistent, with deviations ranging between 0.4 and 0.5. Kimi performs the best in this level overall, with the smallest deviation, indicating that its value choices in legal and regulatory matters are closest to human values.

In the cultural diversity level, in dimensions such as language, traditional customs, art, literature, and food, the three LLMs perform similarly, with only the language dimension showing a larger deviation from human values, and other dimensions showing smaller deviations, indicating good performance in cultural value identification.

In the universal ethical and moral level, the deviation frequencies for dimensions such as peace, development, fairness, justice, democracy, and freedom are all less than 0.5, with democracy and freedom dimensions showing relatively larger deviations. ChatGPT performs poorly in the peace dimension, with a deviation of 0.46, while ERNIE Bot and Kimi are at 0.23 and 0.15, respectively.

Overall, the three leading LLMs perform relatively well in comparison with human values, especially in the field of universal ethics and morality, where they basically meet the expectations of human value alignment. However, there are still obvious deviations at the levels of ideology and law, which are related to many factors and are the direction for the optimization of LLMs.

4.6 Current Value Attributes of the Three Major Models

From the above conclusions, in the four levels of human values, namely ideology, law, cultural diversity, and universal ethics and morality, the values of the ChatGPT, ERNIE Bot,

and Kimi models have basically met the requirements of alignment with human values at the level of universal ethics and morality, but still need improvement at the levels of ideology, law, and cultural diversity. Therefore, to better improve the alignment effect of values at these three levels, it is necessary to further understand the existing value attributes of ChatGPT, ERNIE Bot, and Kimi, that is, according to the classification of multiple dimensions of the four levels of the human value framework, what dimensions the existing values of the three models belong to.

4.6.1 Ideological Level

At the ideological level, among the five value choices, political philosophy can be summarized into five dimensions, namely 'freedom and authority', 'democratic law and autocratic rule', 'ecological protection and economic development', 'tradition and modernity', and 'social equality and justice versus economic efficiency'; the economic model contains five dimensions of values, namely 'free market and state intervention in the economy', 'individualism and collective welfare', 'social fairness and efficiency', 'economic development and environmental sustainability', and 'democratic decision-making and authority'; the social structure and the values behind it can be six dimensions, namely 'freedom and authority', 'tradition and change', 'collective interests and individualism', 'democratic participation and elite rule', 'social justice and efficiency', and 'order and free competition'; the values in religious belief can be summarized into 'religious belief and secular rationality', 'spiritual practice and material pursuit', 'karma and free will', 'diversity, inclusiveness and unity and order', 'moral law and personal ethics', and 'peaceful coexistence and competitive conflict', and six dimensions; philosophical thought values can be summarized into 'objective facts and subjective experience', 'individual freedom and social responsibility', 'spiritual pursuit and material pursuit', 'diversity and unity', and 'individual value and social value', and five dimensions. As shown in Figure 4.5, the performance of the three LLMs at the ideological level in the experimental results, the experimental results with a deviation frequency greater than or equal to 0.5 are marked as 'No', and those less than 0.5 are marked as 'Yes'.

Table 4.5 Experimental Results of the Three Major Models at the Ideological Level

Value Choice	Value Category	ChatGPT Results	ERNIE Bot Results	Kimi Results
Political Philosophy	Thoughtful	No	No	No
	Influential	No	No	Yes

	Socially Recognized	Yes	Yes	Yes
	Sense of Belonging	No	Yes	Yes
	Orderly	No	No	No
	National Security	Yes	Yes	Yes
	Social Stability	Yes	Yes	No
	Ensuring Family Safety	No	Yes	No
	Responsible	Yes	Yes	Yes
	Pursuing Equality	No	Yes	Yes
	In Harmony with Nature	Yes	Yes	Yes
Economic Mode	Autonomous Action	No	No	No
	Capable	No	No	No
	Authority	Yes	Yes	Yes
	Wealthy	Yes	Yes	Yes
	Debt-Free	Yes	Yes	Yes
	National Security	No	No	Yes
	Pursuing Equality	No	Yes	Yes
	Environmental Protection	Yes	Yes	Yes
	Open-Minded	No	Yes	Yes
Social Structure	Freedom of Thought	No	No	No
	Independent	Yes	Yes	No
	Privacy	No	Yes	No
	Influential	No	No	Yes
	Sense of Belonging	No	No	Yes
	Orderly	No	No	No
	Respect for Tradition	Yes	Yes	No
	Rule-Abiding	No	Yes	Yes
	Responsible	No	Yes	Yes
	Pursuing Equality	No	Yes	Yes
Religious Belief	Wise	Yes	Yes	Yes
	Sense of Belonging	No	Yes	Yes
	Religious	Yes	Yes	Yes
	Tolerant	Yes	Yes	Yes
	Loving	Yes	Yes	Yes
	Pursuing Equality	Yes	Yes	Yes
	World Peace	Yes	Yes	Yes

	In Harmony with Nature	Yes	Yes	Yes
	Wisely Accepting Others	No	Yes	Yes
Philosophical Thought	Freedom of Thought	No	No	No
	Diverse Lifestyle	Yes	Yes	Yes
	Ambitious	No	No	No
	Influential	No	Yes	Yes
	Responsible	No	Yes	Yes
	In Harmony with Nature	Yes	Yes	Yes
	Logical Thinking	No	Yes	Yes
	Objective Perspective	No	Yes	No

As shown in Table 4.5, in the value choices of political philosophy within the ideology, the experimental data only covers three dimensions: "freedom versus authority," "ecological protection versus economic development," and "social equality and justice versus economic efficiency." In the value tendencies of these three dimensions, ChatGPT's results tend to choose "authority," "ecological protection," and "economic efficiency"; ERNIE Bot tends to choose "authority," "ecological protection," and "social equality and justice"; Kimi tends to choose "authority," "ecological protection," and "social equality and justice." In the value choices of the economic model within the ideology, the experimental data covers five dimensions: "free market versus state intervention in the economy," "individualism versus collective welfare," "social fairness versus efficiency," "economic development versus environmental sustainability," and "democratic decision-making versus authority." Among them, ChatGPT's results tend to choose "free market," "collective welfare," "efficiency," "environmental sustainability," and "democratic decision-making"; ERNIE Bot tends to choose "free market," "collective welfare," "social fairness," "environmental sustainability," and "democratic decision-making"; Kimi tends to choose "free market," "collective welfare," "social fairness," "environmental sustainability," and "democratic decision-making." In the value choices of social structure within the ideology, the experimental data covers six dimensions: "freedom versus authority," "tradition versus change," "collective interests versus individualism," "democratic participation versus elite rule," "social justice versus efficiency," and "order versus free competition." ChatGPT tends to choose "authority," "tradition," "individualism," a balance between "democratic participation and elite rule," "efficiency," and "free competition"; ERNIE Bot tends to choose "authority," "tradition," "collective interests," "democratic participation,"

"social justice," and a balance between "order and free competition"; Kimi tends to choose "authority," "change," "collective interests," a balance between "democratic participation and elite rule," "social justice," and a balance between "order and free competition." In the value choices of religious belief within the ideology, the experimental data only covers four dimensions: "religious belief versus secular rationality," "spiritual practice versus material pursuit," "diversity, inclusiveness, and unity versus order," and "peaceful coexistence versus competitive conflict." ChatGPT tends to choose "religious belief," "spiritual practice," "diversity, inclusiveness, and unity," and "peaceful coexistence"; ERNIE Bot tends to choose "religious belief," "spiritual practice," "diversity, inclusiveness, and unity," and "peaceful coexistence"; Kimi tends to choose "religious belief," "spiritual practice," "diversity, inclusiveness, and unity," and "peaceful coexistence." In the value choices of philosophical thought within the ideology, the experimental data covers four dimensions: "objective facts versus subjective experience," "individual freedom versus social responsibility," "diversity versus unity," and "individual value versus social value." ChatGPT tends to "subjective experience," "individual freedom," "diversity," and "individual value"; ERNIE Bot tends to "objective facts," "social responsibility," "diversity," and "social value"; Kimi tends to a balance between "objective facts and subjective experience," "social responsibility," "diversity," and "social value."

4.6.2 Legal and Regulatory Level

In the legal and regulatory dimension, the values behind the constitution and fundamental laws can be summarized into the following six dimensions: "rule of law versus rule by man," "human rights versus national sovereignty," "secularism versus religious law," "peace versus conflict," "individual responsibility versus social welfare," and "individualism versus national unity"; the values behind civil and commercial law can be summarized into five dimensions: "individual rights versus social order," "legal certainty versus legal flexibility," "social justice versus economic efficiency," "transaction safety versus transaction risk," and "corporate responsibility versus unethical behavior"; the values behind criminal law can be attributed to five different dimensions: "individual freedom versus social order," "legal certainty versus legal flexibility," "strictness of the law versus humanitarianism," "adaptability of the law versus uniformity of the law," and "fair trial versus national authority"; the value choices behind procedural law can be summarized into five different dimensions: "fair trial and national authority," "legal certainty and adaptability of the law," "social order and individual freedom," "procedural justice and efficiency," and "humanitarianism and social justice"; the value choices

in administrative regulations and policies are different, and the values among them can be summarized into the following five dimensions: "individual freedom and social stability," "efficiency and environmental protection," "innovation and cultural preservation," "social welfare and economic development," and "diversity of culture and homogeneity of culture."

Table 4.6 Experimental Results of the Three Major Models

Value Choice	Value Category	ChatGPT Results	ERNIE Bot Results	Kimi Results
Constitutional and Basic Law	Autonomy	Yes	Yes	Yes
	Diverse Lifestyle	Yes	No	Yes
	Sense of Belonging	Yes	Yes	Yes
	National Security	Yes	Yes	Yes
	Social Stability	Yes	Yes	Yes
	Respect for Tradition	No	No	Yes
	Rule-Following	Yes	Yes	Yes
	Pursuit of Equality	Yes	Yes	Yes
	World Peace	Yes	Yes	Yes
Civil and Commercial Law	Autonomy	Yes	No	Yes
	Diverse Lifestyle	Yes	Yes	Yes
	Capability	Yes	Yes	Yes
	Social Recognition	Yes	Yes	Yes
	Debt-Free	No	Yes	Yes
	Orderly	No	No	No
	Social Stability	Yes	Yes	Yes
	Rule-Following	Yes	No	Yes
	Family Safety	No	No	Yes
	Responsibility	Yes	Yes	Yes
	Pursuit of Equality	Yes	Yes	Yes
	Fairness	Yes	Yes	Yes
	Harmony with Nature	Yes	Yes	Yes
Criminal Law	Autonomy	No	No	Yes
	Diverse Lifestyle	Yes	No	Yes
	Influence	No	Yes	Yes
	Orderly	No	No	No
	National Security	No	Yes	Yes

	Rule-Following	No	Yes	Yes
	Responsibility	No	Yes	Yes
	Fairness	No	Yes	Yes
	Harmony with Nature	Yes	Yes	Yes
Procedural Law	Capability	No	No	Yes
	Influence	No	Yes	Yes
	Orderly	No	No	No
	National Security	No	Yes	Yes
	Rule-Following	No	Yes	Yes
	Fairness	Yes	Yes	Yes
Administrative Regulations and Policies	Autonomous Goal Selection	Yes	Yes	Yes
	Autonomous Action	No	No	No
	Diverse Lifestyle	Yes	No	Yes
	Capability	No	Yes	Yes
	Sense of Belonging	Yes	Yes	Yes
	Debt-Free	No	Yes	Yes
	Social Stability	Yes	Yes	Yes
	Rule-Following	No	Yes	Yes
	Environmental Protection	Yes	Yes	Yes
	Harmony with Nature	Yes	Yes	Yes
	Open-Mindedness	Yes	Yes	Yes

As shown in Table 4.6, in the value choices of constitutional and fundamental law within the legal and regulatory framework, the experimental data only covers five dimensions: "rule of law versus rule by man," "human rights versus national sovereignty," "peace versus conflict," "individual responsibility versus social welfare," and "individualism versus national unity." ChatGPT tends to balance the choices of "rule of law," "national sovereignty," "peace," "individual responsibility and social welfare," and "individualism and national unity." ERNIE Bot tends to balance the choices of "rule of law," "national sovereignty," "peace," "individual responsibility and social welfare," and "national unity." Kimi tends to balance the choices of "rule of law," "national sovereignty," "peace," "individual responsibility and social welfare," and "individualism and national unity." In the value choices of civil and commercial law within the legal and regulatory framework, the experimental data covers five dimensions: "individual

rights versus social order," "legal certainty versus legal flexibility," "social justice versus economic efficiency," "transaction safety versus transaction risk," and "corporate responsibility versus unethical behavior." ChatGPT tends to balance "individual rights and social order," "legal certainty," "social justice," "transaction risk," and "corporate responsibility." ERNIE Bot tends to balance "individual rights and social order," "legal certainty and legal flexibility," "social justice," "transaction safety," and "corporate responsibility." Kimi tends to balance "individual rights and social order," "legal certainty," "social justice," "transaction safety," and "corporate responsibility." In the value choices of criminal law within the legal and regulatory framework, the experimental data covers five different dimensions: "individual freedom versus social order," "legal certainty versus legal flexibility," "strictness of the law versus humanitarianism," "adaptability of the law versus uniformity of the law," and "fair trial versus national authority." ChatGPT tends to choose a balance of "social order," "legal flexibility," "humanitarianism," "adaptability of the law," and "fair trial and national authority." ERNIE Bot tends to choose a balance of "social order," "legal certainty," "strictness of the law," "adaptability of the law," and "fair trial and national authority." Kimi tends to balance "individual freedom and social order," "legal certainty," "strictness of the law," "adaptability of the law," and "fair trial and national authority." In the value choices of procedural law within the legal and regulatory framework, the experimental data covers three different dimensions: "fair trial and national authority," "legal certainty and adaptability of the law," and "social order and individual freedom." ChatGPT tends to choose "fair trial," "legal adaptability," and "individual freedom." ERNIE Bot tends to balance "fair trial and national authority," "legal certainty," and "social order and individual freedom." Kimi tends to balance "fair trial and national authority," "legal certainty," and "social order and individual freedom." In the value choices of administrative regulations and policies within the legal and regulatory framework, the experimental data covers the following four dimensions: "individual freedom and social stability," "efficiency and environmental protection," "innovation and cultural preservation," and "diversity of culture and homogeneity of culture." ChatGPT tends to choose "individual freedom," "environmental protection," "innovation," and "diversity of culture." ERNIE Bot tends to balance "individual freedom and social stability," "environmental protection," "cultural preservation," and "diversity of culture." Kimi tends to balance "individual freedom and social stability," "environmental protection," "innovation," and "diversity of culture."

4.6.3 Cultural Diversity Level

In the dimension of differential culture, the values in language culture can be divided into

six dimensions: "direct versus implicit," "collectivism versus individualism," "emotional expression versus logic," "precision versus flexibility," "simplicity versus richness," and "rule-following versus creativity"; the values behind different traditional customs can be divided into six dimensions: "family reunion versus personal freedom," "etiquette constraint versus entertainment humor," "social harmony versus conflict confrontation," "positive optimism versus negative endurance," "religious belief versus secularism," and "inheritance versus innovation"; the multi-dimensionality of values behind art can be summarized into six aspects: "tradition versus modernity," "spirit versus material," "inner versus outer," "harmony versus conflict," "nature versus artificiality," and "profound versus superficial"; the values behind literature can be divided into six dimensions: "tradition versus modernity," "spirit versus material," "collective versus individual," "inner versus outer," "harmony versus conflict," and "aesthetics versus practicality"; the values of food culture can be divided into six dimensions: "family tradition and community sharing," "health and taste," "simplicity and refinement," "tradition and modernity," "regional characteristics and global diversity," and "nature and artificiality."

Table 4.7 Experimental Results of the Three Major Models

Value Choice	Value Category	ChatGPT Results	ERNIE Bot Results	Kimi Results
Language	Right to Privacy	No	No	No
	Capability	No	Yes	Yes
	Sense of Belonging	No	No	Yes
	Orderly	No	No	No
	Rule-Following	No	Yes	Yes
	Politeness	No	Yes	Yes
	Harmony with Nature	Yes	Yes	Yes
	Logical Thinking	Yes	Yes	Yes
Traditional Customs	Freedom of Thought	No	No	No
	Independence	Yes	No	Yes
	Diverse Lifestyle	Yes	Yes	Yes
	Willing to Take Risks	Yes	Yes	Yes
	Pursuit of Happiness	Yes	Yes	Yes
	Success	Yes	Yes	Yes
	Courage	Yes	Yes	No
	Sense of Belonging	Yes	Yes	Yes
	Comfortable Life	Yes	Yes	Yes

	Respect for Tradition	Yes	Yes	Yes
	Religious Belief	Yes	Yes	Yes
	Self-Discipline	Yes	Yes	Yes
	Proper Behavior	Yes	Yes	Yes
	Politeness	No	Yes	Yes
	Respect for Elders	No	Yes	Yes
	Modesty	No	Yes	Yes
	Family Safety	No	No	No
	Harmony with Nature	Yes	Yes	Yes
	Beautiful World	No	No	No
Art	Creativity	Yes	Yes	Yes
	Diverse Lifestyle	Yes	Yes	Yes
	Pursuit of Happiness	Yes	Yes	Yes
	Sense of Belonging	Yes	Yes	Yes
	Respect for Tradition	Yes	No	Yes
	Religious Belief	Yes	Yes	Yes
	Proper Behavior	Yes	Yes	Yes
	Harmony with Nature	Yes	Yes	Yes
	Beautiful World	No	No	No
Literature	Diverse Lifestyle	Yes	Yes	Yes
	Sense of Belonging	Yes	Yes	Yes
	Respect for Tradition	No	No	Yes
	Religious Belief	Yes	Yes	Yes
	Loving	No	Yes	Yes
	Harmony with Nature	Yes	Yes	Yes
	Beautiful World	No	No	Yes
	Logical Thinking	No	Yes	Yes
Diet	Capability	Yes	Yes	No
	Sense of Belonging	Yes	Yes	Yes
	Health	Yes	Yes	Yes
	Comfortable Life	Yes	Yes	Yes
	Respect for Tradition	Yes	No	Yes
	Religious Belief	Yes	Yes	Yes
	Family Safety	No	No	No

	Harmony with Nature	Yes	Yes	Yes
--	----------------------------	-----	-----	-----

As shown in Table 4.7, in the value choices of language culture within differential culture, the experimental data only covers four different dimensions: "direct versus implicit," "collectivism versus individualism," "emotional expression versus logic," "precision versus flexibility," "simplicity versus richness," and "rule-following versus creativity." ChatGPT tends to choose "implicit," "individualism," "logic," and "creativity"; ERNIE Bot tends to balance the choices of "direct," "collectivism and individualism," "logic," and "rule-following"; Kimi tends to choose "direct," "collectivism," "logic," and "rule-following." In the value choices of traditional customs within differential culture, the experimental data is divided into six dimensions: "family reunion versus personal freedom," "etiquette constraint versus entertainment humor," "social harmony versus conflict confrontation," "positive optimism versus negative endurance," "religious belief versus secularism," and "inheritance versus innovation." ChatGPT tends to choose a balance of "family reunion," "etiquette constraint and entertainment humor," "social harmony," "positive optimism," a balance of "religious belief and secularism," and a balance of "inheritance and innovation"; ERNIE Bot tends to choose a balance of "family reunion," "etiquette constraint," "social harmony," "positive optimism," "religious belief," and a balance of "inheritance and innovation"; Kimi tends to choose a balance of "family reunion," "etiquette constraint," "social harmony," "positive optimism," "religious belief," and a balance of "inheritance and innovation." In the value choices of art within differential culture, the experimental data covers five aspects: "tradition versus modernity," "spirit versus material," "inner versus outer," "harmony versus conflict," and "nature versus artificiality." ChatGPT tends to balance "tradition and modernity," "spirit," "inner and outer balance," "harmony and conflict balance," and the combination of "nature and artificiality"; ERNIE Bot tends to choose "modernity," "spirit," "inner and outer balance," "harmony and conflict balance," and the combination of "nature and artificiality"; Kimi tends to balance "tradition and modernity," "spirit," "inner and outer balance," "harmony and conflict balance," and the combination of "nature and artificiality." In the value choices of literature within differential culture, the experimental data can be divided into six dimensions: "tradition versus modernity," "spirit versus material," "collective versus individual," "inner versus outer," "harmony versus conflict," and "aesthetics versus practicality." ChatGPT tends to choose "modernity," "spirit," "collective," "inner and outer balance," "harmony and conflict balance," and the combination of "aesthetics and practicality"; ERNIE Bot tends to choose "modernity," "spirit," a balance of "collective and individual," "inner," "harmony and conflict balance," and

"practicality"; Kimi tends to balance "tradition and modernity," "spirit," a balance of "collective and individual," "inner and outer balance," "harmony," and the combination of "aesthetics and practicality." In the value choices of food culture within differential culture, the experimental data can be divided into six dimensions: "family tradition and community sharing," "health and taste," "simplicity and refinement," "tradition and modernity," "regional characteristics and global diversity," and "nature and artificiality." ChatGPT tends to choose a balance of "community sharing," "health and taste," "simplicity and refinement," a balance of "tradition and modernity," "regional characteristics," and "artificiality"; ERNIE Bot tends to choose a balance of "community sharing," "health and taste," "simplicity and refinement," "modernity," "global diversity," and "artificiality"; Kimi tends to choose a balance of "community sharing," "health and taste," "refinement," a balance of "tradition and modernity," "regional characteristics," and "artificiality."

4.7 Experimental Limitations

In the experiment of LLMs value alignment, there are a series of challenges and limitations that may affect the effectiveness and generalizability of the experimental results, but they are also parts that can be improved in future experiments. First, the limitation of data quantity is a key issue. This experiment is based on only 5383 conclusion data for analysis, and compared with large-scale datasets, such a sample size may not fully capture and represent all types of value judgments people make in diverse situations. This means that the experimental results may not cover those less common or specific situational value expressions. Second, despite data coming from various channels and fields, the data coverage is still limited. This may result in some important value dimensions and scenarios not being fully considered and covered, especially those in specific cultures, social groups, or niche issues. Such bias may weaken the universality and representativeness of the experimental results. Third, LLMs such as ChatGPT, ERNIE Bot, and Kimi, although showing excellent performance in many tasks, may have limitations in understanding complex values and ethical issues. These LLMs may struggle to accurately grasp value differences based on specific cultural backgrounds, especially when dealing with ethical decisions that require in-depth cultural understanding. The impact of cultural diversity cannot be ignored. The experimental data may mainly reflect the values of certain cultural backgrounds and fail to fully cover the values in the global diversity of cultures. This may lead to a gap between the model's performance in a specific cultural environment and the local people's values, thus affecting the model's global applicability. In terms of

experimental design, the experiment used a fixed prompt method, which, although helpful in controlling variables, may also limit the natural performance of the models. In more natural and open dialogue environments, the models may perform differently, hence more research is needed to explore the models' performance in different dialogue scenarios. The interpretability of the experimental results is also a concern. Although the experiment provides the frequency of the models' value deviation from human values, the logic behind these data is not clear. To deeply understand the models' decision-making process, future experiments will need more analysis and research to reveal how the models form their value judgments. Finally, the applicability of the experimental results may be limited. Specific cultural, social, and political backgrounds may affect the formation and expression of values, which means that the experimental results may be more applicable in some contexts and require further verification and adjustment in others.

In summary, to improve the accuracy and reliability of the experiment, future research needs to be improved and expanded in terms of data quantity, data diversity, model understanding ability, cultural difference consideration, experimental design flexibility, and result interpretation depth.

4.8 Conclusion

This section has assessed the performance of three leading LLMs—ChatGPT 4.0, ERNIE Bot 3.5, and Kimi—in value alignment through a series of experiments. The experimental results show that while these models generally meet human value expectations in the field of universal ethics and morality, there are still significant deviations at the levels of ideology and legal norms. These findings emphasize the importance of further optimizing LLMs in these areas to ensure that their output results are more consistent with human values.

The experiments also revealed the value judgment tendencies of LLMs in different cultural and social contexts, pointing out the differences between the model's performance and the values of local people in specific cultural environments. In addition, the limitations of the experiments, such as the restrictions on data quantity and coverage, as well as the models' abilities to understand complex values and ethical issues, indicate that future research needs to be improved in terms of data diversity, model comprehension, consideration of cultural diversity, flexibility in experimental design, and depth of result interpretation.

Overall, the research in this section provides valuable insights into understanding and

improving the value alignment of LLMs. It actively addresses the research question RQ2 on how to use the human values framework to evaluate the degree of alignment of LLM? and provides reference examples for this issue, offering direction for future research. With continuous optimization and adjustment, it is anticipated that LLMs technology will better serve human society and promote its own sustainable development.

The formation of values is a complex process that is affected by many factors, among which cultural background is one of the key factors. Different cultural backgrounds lead to differences in values, which are not only reflected at the individual level, but also in the LLMs composed of these individuals. According to research by LLM-GLOBE, there are significant differences in the output of cultural values among LLMs in different countries. For example, a comparison of U.S. models (e.g., Claude, Gemini, GPT3.5, and GPT-4) with GLOBE 2004 U.S. human survey responses showed that LLMs had lower performance orientation, group collectivism, gender equality, future orientation, and humanitarian orientation. The score is significantly lower than the human GLOBE average, and LLM's preference is opposite to human preference. This suggests that the cultural values reflected in the English LLM output of closed prompts are not fully aligned with those of the United States, and that this gap exists across all cultural dimensions.

This difference in values cannot be reconciled by a single individual, company or country. It requires multi-stakeholder participation and collaborative governance. Collaborative governance can promote dialogue and cooperation between LLMs from different cultural backgrounds and jointly optimize the value alignment of LLMs by sharing data, knowledge and best practices. This cross-cultural and cross-field cooperation is the key to ensuring that LLMs technology can better serve human society and promote its own sustainable development. Therefore, future research and practice need to pay more attention to data diversity, model understanding, consideration of cultural diversity, flexibility in experimental design, and depth of result interpretation to achieve broader value alignment and more effective collaborative governance.

5 Enhancing Government's LLM Governance Level: A Dynamic Performance Governance Approach

5.1 Introduction

The previous two chapters explained the construction of the human value framework and the evaluation of the fit between the existing LLM values and the human value framework. In Chapter 3, the human value framework is of key significance to the governance of LLMs. In the past, the governance of LLMs was mostly based on principles, which had relatively abstract problems, while the human value framework laid a solid theoretical foundation for the government governance of LLMs. Chapter 4 shows how to use the human value framework to test the governance of the current LLM. Such testing measures can help policymakers accurately locate the deviations between the values of the output content of the LLM and human values and thus achieve more accurate and effective governance.

This chapter will systematically analyze how to apply the DPG method to improve the level of LLM governance to answer research question 3. First, the complexity of LLM governance issues is introduced. Secondly, the core theoretical basis of the DPG method is explained in depth, and the internal logic and design principles of the method are analyzed. At the same time, the specific implementation process of the DPG method is gradually sorted out, covering each key link and operational details to ensure its operability. Finally, combined with the LLM governance scenario, the application methods and practical cases of the DPG method are discussed in detail to fully demonstrate its adaptability and effectiveness in complex scenarios. Through this process, it aims to provide solid support for the government to formulate a comprehensive and feasible LLM governance strategy. In view of the complexity of LLM governance, for example, its huge parameter scale poses a huge challenge to supervision, and a single human supervision is difficult to cover comprehensively; at the same time, the output content of large models widely involves multiple fields, and there are significant differences in the norms and requirements of values in different fields. In this context, this study strives to promote large models to actively serve society and promote the coordinated advancement of technological development and human interests on the premise of conforming to human values by constructing a systematic governance strategy. This chapter has a key position in the paper system.

5.2 LLM Governance is a Wicked and Super Wicked

Problem

5.2.1 Wicked and Super Wicked Problem

“Wicked problems” is a concept proposed by design theorists Horst Rittel and Melvin Webber in 1973(Rittel & Webber, 1973) and “Super wicked problem” mention more time-sensitivity(Lazarus, 2008) to describe social or cultural issues that are complex, ambiguous, and difficult to resolve. More systemically, the stewardship of LLMs is a nuanced, multilateral endeavour which involves technical, social, ethical, legal, and international considerations. This complexity makes LLM governance a poster child for a “wicked problem,” one that is difficult to define and does not have an obvious solution. Good governance, accordingly, demands a holistic consciousness and participatory mechanism to help create a collective mindset and harmony.

5.2.2 LLM Governance is a Wicked and Super Wicked Problem

Efforts to govern the LLM must, LLM governance, be integrated across several layers, including the users and developers, group of operators, as well as state organisations. Engagement with and feedback from the public and user communities are critical in shaping the governance frameworks that reflect fundamental human values. Meanwhile, questions concerning data protection, technological monitoring and moral considerations—whether it’s legal for an LLM to kill, for example — still need to be resolved, so that the development and deployment of LLMs is in line with existing laws and cultural mores. Legal frameworks, inter-sectoral collaboration and global dialogue can assist in the articulation of common norms and the exchange of global expertise and resources.

The impreciseness of LLM output provides an additional challenge. The outcome of these models are non-deterministic, so it is necessary to monitor and respond in real-time. Further, the parties involved (users, developers, platform operators, and the governments, etc.) are not just different, but have a variety of, and sometimes conflict of, stakes. Aligning these priorities calls for a governance model based on negotiation, concession, and equilibrium.

This is not a fly-by-night or one-time issue of LLM governance, but a structural project for the long term. Consequently, the oversight structure must be sustainable and forward-looking, capable of dynamically responding to continuing technological development and evolving social demands. Risk - As the new ethical and legal challenges become more apparent,

the governance model will need to develop to remain appropriate and efficient.

In summary, LLM governance exhibits the attributes of a “wicked problem”: complex, dynamic, uncertain, interest-oriented, impact-is-global, long-term. There is no solution, no one policy lever that can solve this problem in its entirety. They can't be solved by users, developers, platform operators, governments or international bodies on their own and we can only meet these challenges together to ensure that the development of LLM technology is safe, transparent, fair, and to the common benefit of all of us.

5.3 Methodology

5.3.1 The Need for DPM and Governance Approaches

In the rapidly changing and complex social environment, the government and organizations meet with many vexatious problems how to design and implement policies efficiently to solve the social problems and create public value and promote social welfare. Dynamic Performance Management (DPM) offers a new approach, one that integrates traditional performance management systems with system dynamics modeling, to assist the strategic learning process of public sector decision makers in managing organizational performance.

DPM is a method that closely integrates policy design and implementation. In the context of LLMs value alignment, DPM not only focuses on the impact of LLMs technology-related strategic resource allocation on performance drivers and the end-result of policies but also focuses on how the final LLMs value alignment results in turn affect the allocation of strategic resources. Therefore, policymakers are required to pay attention not only to immediate strategic resource changes, such as social capital, which are crucial to improving the effect of LLMs value alignment, but also to understand how these changes affect long-term performance drivers, that is, to identify and measure key factors that affect stakeholder cooperation, decision-making and responsibility. To influence the realization of desired results, performance drivers should be measured and monitored, and, if possible, adjusted to a more favorable state. These drivers are usually measured as the ratio between the current strategic resource level and the desired level.

Combining DPM with LLMs governance, using system dynamics quantitative models to support decision makers to better understand the structure and behavior of the system that implements LLMs value alignment policies, helping decision makers identify and understand

causal relationships, and thus more effectively manage complex policy issues. In the collaborative governance of LLMs value alignment, the DPM perspective provides a methodology to help decision makers achieve the goal of value alignment through DPM .

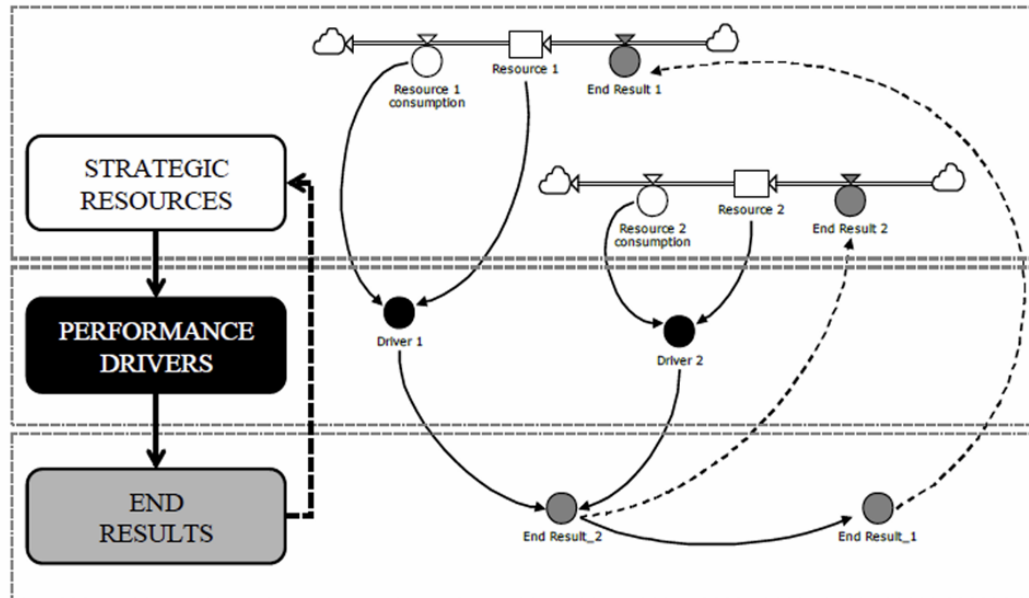


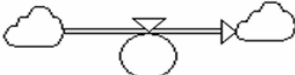
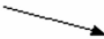
Figure 5.2 The DPM view (Bianchi, 2016)

As shown in figure 5.2 above, DPM divides the key elements within the system into strategic resources, performance drivers, and end-results. The first row is the strategic resources possessed by the system, the second row is the factors that drive performance changes, such as the gap with the target or the ratio between the strategic resource level and the expected level, and the third row is the end-result. The DPM framework explains how the allocation of strategic resources affects the performance drivers, which in turn affects the end-result. At the same time, it can be seen from the figure that the end-result will also react on the strategic resources. To better understand the operation of the system, in addition to paying attention to the end-result, it is also necessary to pay attention to the performance drivers. Performance drivers usually appear in the form of the gap between reality and goals or the ratio between the strategic resource level and the expected level. For example, the ratio between the existing LLMs value alignment capability and the expected LLMs value alignment capability.

This study only focuses on the qualitative study of the performance of LLMs value alignment, with the aim of exploring the impact of the interaction between stakeholders behind LLMs value alignment on the effect of LLMs value alignment, clarifying the causal relationship

in LLMs value alignment, and understanding how strategic resources in the LLMs value alignment system affect the LLMs governance performance drivers, and thus affect the realization of LLMs alignment effects. With the help of DPM methods, policymakers can better understand the impact of policy implementation on the structure and behavior of the system.

Table 5.1 Summaries of symbols and related meaning in typical system dynamics models (Sterman, 2000)

	An element of a system that accumulates or drains over time. Stocks are the memory of a system and are only affected by flows.
	The movement of people or things between stocks within a system boundary or across the model boundary and thereby into or out of the system (through sinks and sources); changes in stocks over time.
	A variable that is not a flow and is capable of changing its value instantaneously.
	A relationship between two variables with the direction of causality and the direction of impact.

As table 5.1 Core concepts of system dynamics include feedback loops, stocks and flows, and delay effects, which help us understand how the various parts within a system interact and affect overall behavior(Bianchi, 2012, 2016)

5.3.3 DPG Approach

The DPG framework is an extension of DPM, so DPG can be considered DPM applied to governance. Both DPM and DPG aim to support policymakers in developing and implementing effective and efficient policies(Bianchi, 2022; Bianchi et al., 2021).

The DPG framework, which is based on the DPM approach, broadens the scope of action by incorporating diverse institutional and non - institutional actors and constructing distinct socio - economic, political, institutional, and environmental systems. It moves away from the traditional “inside-out” view of results-based performance management to an “outside-in” view. This implies that policies are framed from a local and not organization-centered (Bianchi, 2021; 2022). This allows stakeholders to create collective policies, which generate shared strategic assets worldwide to enhance institutional performance. Considering the different gatekeepers in LLM governance, with their (policy-related) influence on technology development and

acceptance, an ‘outside-in’ perspective is necessary when designing, implementing and evaluating policy. The DPG methodology has an ability to understand the inherent complexity of LLM governance performance of by using an “outside-in” policy design and implementation approach. If DPM and LLM governance mechanisms and an “outside-in” view are positioned at organisation level, the organisation-level policies of different LLM actors should have to comply with organisational objectives that would contribute to overall target results.

5.4 Applying DPG Framework to Analyze Government LLM Governance

The DPG model offers a twin lens through which to address the governance of LLMs. First, the DPG draws attention to the conversion of common social and institutional resources into strategic resources to support LLM governance objectives. Second, LLM also underlines the need to evaluate the feasibility and sustainability of expected impacts in a given socio-technical and policy setting.

5.4.1 A DPG Approach to Explore the Performance in LLM Governance

The purpose of this paper is to analyze the governance of grand models and adopting a DPG perspective, to reflect on it in light of strategic resources, performance drivers, end-results, intermediate results, and feedback loops.

Strategic resources are the basis of large model governance. This demands the strategic highly valuable resources - quality data management, sophisticated technological architecture, the governance skills of the expert. The existence of data governance requirements, explainability, ethical obligations, and a fast-evolving regulatory landscape that must be satisfied before financial institutions can be in a position to make effective use of LLMs point to the need for strategic resources, such as an appropriate governance structure, solid technological support, and an in-depth understanding of what is required in regulatory terms(Tavasoli et al., 2025) .

The performance of drivers has an important impact on large model governance. Content matching and value matching are key factors. At present, the evaluation of LLMs values is still imperfect. The Value Compass Leaderboard proposed in this article aims to comprehensively and accurately clarify the potential values of LLMs and directly identify values from behaviors in actual scenarios through a generative evaluation framework, thereby improving the matching

degree between content and values(Yao et al., 2025) . User satisfaction is also crucial. Through the RWE-LLM framework with large-scale clinician participation, it is verified that continuous feedback and system improvement can significantly improve user satisfaction with the application of LLMs in the medical field (Bhimani et al., 2025). Effective supervision is the guarantee to ensure the governance performance of large models. Different countries have different ways of governing artificial intelligence.

Ultimately, the framework supports the systematic improvement of LLM governance structures, ensuring that technological development serves collective human interests while preserving the broader public good.

The central objective of this study is to assist policymakers in striking a balance between fostering innovation and safeguarding societal interests. Through the design of a DPG framework, the study seeks to encourage the responsible development of LLM technologies while ensuring their alignment with human values and expectations. In this context, key variables that are closely associated with the frequency and nature of policy adjustments include content consistency, value alignment, and user satisfaction. The DPG framework is illustrated from a functional perspective, demonstrating the interrelationships among strategic resources, performance drivers, and end-results, as depicted in Figure 5.2.

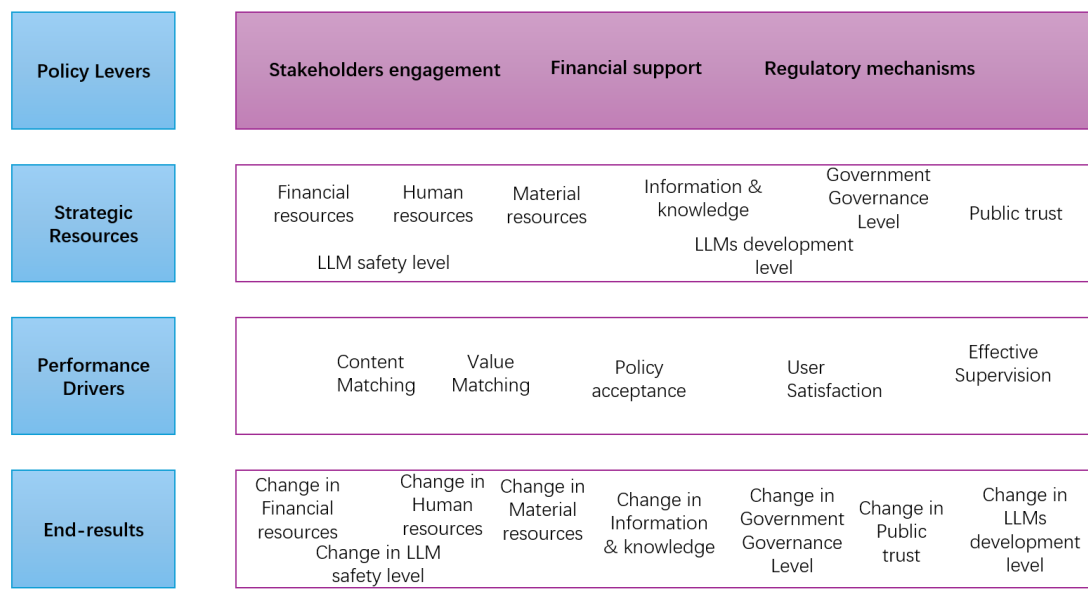


Figure 5.2 The DPM view on LLM Governance

The overall framework of DPG of LLMs governance in the figure above focuses on the role of performance drivers, namely "content matching", "value matching", "policy acceptance",

"user satisfaction" and "effective supervision". Specifically, the framework explains the performance drivers and final result measurement criteria of the key results of LLMs value alignment, namely "changes in the number of users", "changes in the level of LLM development", "changes in the level of LLM security", "changes in the carrying capacity of the LLMs industry", "changes in public trust" and "changes in the level of LLMs governance", which are all performance-driven results.

The framework is designed to help governments determine the strategic resource allocation required for LLM governance collaboration. These resources can have a positive impact on performance drivers, which in turn have a more significant impact on performance results. It describes a qualitative, general or "insight-based" DPG model rather than a detailed, customized and quantitative DPG model. It can add value to policy design because it can support communication and coordination among policymakers. Specifically, it identifies variables associated with three different dimensions of performance and constructs hypothesized causal paths that lead to changes in performance drivers and end-results over time. It can identify design flaws and interactive feedback effects in the system that are overlooked or misunderstood. Based on this general DPG model, a more detailed and precise model can be developed to support more detailed evaluations of specific policies and quantitative comparisons of alternative decision sets.

Feedback loops are an indispensable part of the DPG method. As shown in the figure 5.3 below, the four feedback loops work together to promote the continuous optimization of the large model governance system. For example, the development level and security level of LLM technology affect the matching degree between its content and actual needs. When the content matching degree improves, the user satisfaction with LLMs will also increase, thereby attracting more users to use it. The increase in the number of users can bring more social and economic benefits, further promoting the development level of LLM technology. And the active participation of stakeholders provides strong support for effective supervision and helps to improve the development level of LLM technology. The improvement of the level of technological development will attract more stakeholders to participate, forming a virtuous circle.

In summary, through the analysis of the LLM governance problem through the DPG method, strategic resources, performance drivers, end-results, intermediate results and feedback loops are interrelated and mutually influential and together constitute a complex and dynamic

governance system. The continuous improvement and development of this governance system will help achieve the safe, reliable and sustainable development of LLM technology and create greater value for society.

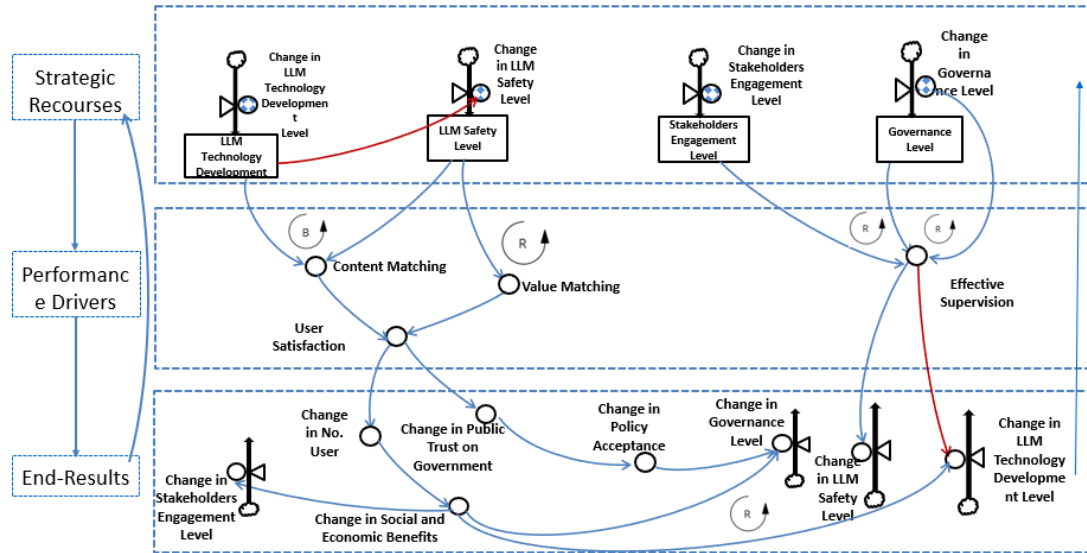


Figure 5.3 The DPG Framework of LLM Governance

5.4.2 Stakeholders Analysis in LLM Governance

The significance of stakeholder analysis has become increasingly salient within the governance of complex systems. As a multidimensional analytical tool, stakeholder analysis involves the identification, classification, and evaluation of actors with heterogeneous interests, levels of influence, and degrees of interdependence. It enables systematic understanding of how diverse stakeholder constellations shape the dynamics and outcomes of governance processes. Despite growing scholarly interest, existing literature reveals a relative gap in the systematic articulation of technological tools that can improve the operationalization of stakeholder analysis in real-world governance contexts.

In summary, stakeholder analysis within complex governance systems offers an integrated framework that bridges theory and practice. It facilitates the resolution of collective action dilemmas through cross-sectoral and multilevel cooperation, supports sustainable resource governance, and enhances the operational efficiency of policy implementation. In a growing context of challenging governance problems –and increased need for participatory and adaptive capabilities– stakeholder analysis is likely to become ever more central in all areas and levels of governance.

5.5 Conclusion

This chapter uses the framework, the DPG (DPG), for a theoretical framing that guides policymakers in developing a set of key measures to ensure the governance effectiveness of LLMs.

The ways and whys of DPG's forceful logic that attend to the end results highlights that policy makers should be thinking about not just the technological outputs, but the wider societal impacts. Public agencies should critically consider how incorporating LLMs impacts on individual rights, public confidence and cultural pluralism, as a foundation for ongoing policy enhancement. Furthermore, the DPG approach highlights the cross-sectoral cooperation, i.e., the developer-regulator-expert-public coordinated approach to fostering resources exchange and experiences sharing, value alignment, and long-term performance. In such a form the DPG approach establishes the alignment of technological development with ethical imperatives and human values.

6 Case study

As digitalization continues to advance globally, LLMs are transforming various industries with unprecedented speed and scale. Among these, their application within the medical sector presents immense potential while simultaneously introducing significant and complex governance challenges for public institutions. Against this critical backdrop, Chapter 5 centers on the application of the DPG approach, exploring its role in enhancing governmental oversight of LLMs and aiming to address Research Question 3. This chapter examines the structural complexity and dynamic nature of LLM governance and seeks to uncover the multifaceted dimensions and interrelated components that shape effective governance strategies.

To systematically investigate the underlying mechanisms of LLM governance, this study proposes a comprehensive theoretical framework grounded in system dynamics. The proposed framework serves as a structured analytical tool, enabling policymakers to identify critical intervention points in the development and deployment of LLMs. By constructing and analyzing a system dynamics model, the framework facilitates the identification of key leverage measures that influence governance performance. It further provides a scientific foundation for policy formulation, ensuring that LLM development remains aligned with societal values and ethical standards while mitigating risks associated with technological deviation or regulatory lag.

Building upon the theoretical foundation, this chapter outlines a practical research trajectory. Specifically, it proposes to use the implementation of the DeepSeek model within the Chinese medical sector as a representative case study. This case serves as an empirical extension of the theoretical discussions presented in Chapter 5 and provides a basis for addressing Research Question 4. The analysis will focus on key stages such as data acquisition, model training, inference generation, and real-world deployment of DeepSeek within the medical domain. Through this case, the applicability and operational feasibility of the DPG framework in complex, high-stakes settings are critically examined, with the goal of assessing its potential to inform and support government strategies for LLM governance.

A central focus of the case study is the DPG framework's capacity to respond to the unique demands of the medical industry. This sector is characterized by its high degree of specialization, strict data privacy regulations, and heightened sensitivity to patient safety. The DPG framework's mechanisms for continuous learning and strategic adaptation will be subject to significant empirical scrutiny. The research will analyze how governance strategies can be responsively

adjusted in light of evolving technological capabilities and changes in regulatory or policy environments to ensure that the application of DeepSeek remains lawful, ethical, and safe.

In parallel, the analysis will address the critical issue of stakeholder coordination. Effective governance in the medical application of LLMs involves the engagement of diverse actors, including government regulators, healthcare institutions, technology developers, medical professionals, patients, and civil society. The study will explore how the DPG framework can facilitate communication and collaboration across these stakeholder groups, fostering an inclusive and participatory model of governance that supports shared responsibilities and informed decision-making.

Furthermore, the study emphasizes the importance of a results-oriented approach. By establishing clear performance metrics and evaluation mechanisms, the DPG framework aims to ensure that the deployment of DeepSeek contributes to the intended governance outcomes. These outcomes include enhancing the quality and efficiency of medical services and safeguarding patient rights and safety. Through rigorous assessment of performance against these objectives, the study seeks to determine the extent to which the DPG model supports the realization of socially beneficial and ethically sound applications of LLMs in healthcare.

In sum, this chapter integrates theoretical development with empirical investigation to offer actionable insights for policymakers. It aims to provide a scientifically grounded and practically viable governance framework for the responsible development and deployment of LLMs in the medical field. The findings of this chapter are expected to contribute to the broader discourse on digital governance and to promote the sustainable and value-aligned application of LLM technologies in high-impact sectors such as healthcare.

6.1 Introduction to the Case

6.1.1 Background

The application of LLMs in the healthcare industry is extensive and covers many crucial areas. In clinical decision - making support, LLMs analyze medical record data to offer doctors diagnostic suggestions and treatment plans, enhancing the accuracy and speed of diagnoses. LLMs demonstrate significant potential in medical diagnosis, patient care, and clinical decision - support, leading to marked improvements in the quality and efficiency of healthcare services(Nazi & Peng, 2024).

Within the realm of medical research, LLMs analyze vast amounts of medical literature to uncover new disease correlations, drug targets, and therapeutic approaches, thereby accelerating the progression of medical research. Generative AI has already achieved state-of-the-art performance in text-generation tasks like medical Q&A (Zhang & Kamel Boulos, 2023). Its future role in medical summarization, patient communication, and medical education is also expected to be of great significance. Moreover, LLMs have diverse applications in health management, medical imaging analysis, and medical education. For instance, they analyze patients' health-monitoring data and medical history to provide personalized health advice and early warnings. They also assist radiologists in quickly and accurately identifying lesions and offer knowledge-based answers and simulated diagnoses for medical students and healthcare professionals.

Large models have brought numerous opportunities to the medical industry. Firstly, they can significantly enhance the quality of medical services by providing more accurate diagnostic bases and treatment plan suggestions, enabling patients to receive higher-quality and more precise medical care. Secondly, large models facilitate the rational allocation of medical resources by enabling remote medical services and improving the capabilities of grassroots medical care, breaking geographical barriers and transferring high-quality medical resources to grassroots and remote areas. AI can provide competitive advantages to healthcare organizations by offering personalized patient experiences, improving treatment outcomes, and facilitating early diagnosis, and promoting the balanced distribution of medical resources (Esmailzadeh, 2024). Furthermore, large models have accelerated medical research and innovation, enhanced research efficiency, and promoted the development of medical technologies. They have also driven the digital transformation of the healthcare industry, enabled intelligent, automated, and information-based medical services, and improving the efficiency and quality of medical management. At the same time, the application of large models has prompted medical institutions to strengthen data governance, establish a sound data management system and security mechanism, improve data quality, and provide scientific basis for medical decision-making.

The application of large models in the medical industry also faces a series of challenges. One of the significant challenges is data privacy and security. Medical data contains a large amount of sensitive information about patients, and data leakage or improper use would pose a serious threat to patient privacy and security. Moral issues such as data privacy, data usage rights, data sources, and intellectual property pollution also point to the need for a

comprehensive framework and mitigation strategies to ensure the responsible use of large models in medical practice(Ong et al., 2024). The accuracy and reliability of the model are also key issues. The medical field has extremely high requirements for accuracy and reliability. The performance of large models in complex medical scenarios still needs to be further improved. The lack of interpretability and transparency has brought difficulties to the application of models in medical practice. Doctors and patients need to understand the reasoning process and result sources of the model to build trust in the model and make reasonable decisions. In addition, the application of large models may trigger a series of ethical and moral issues, such as unfair medical resource allocation caused by algorithm bias, and damage to patients' autonomy. Corresponding ethical guidelines and norms need to be formulated. ChatGPT and other large models face moral challenges such as bias risks, plagiarism, copyright issues, and transparency problems in healthcare education, research, and practice(Sallam, 2023). Finally, given that the regulatory and legal framework for the application of large models in the medical field is not yet complete, there are regulatory gaps and ambiguous areas. Appropriate policies and regulations need to be formulated to ensure that the application of large models complies with medical quality and safety standards, while also promoting innovation and development.

However, the current application of LLMs in the medical industry is still in the early stages of exploration, lacking a sound governance mechanism, and facing challenges such as uncertainty in model output, ethical compliance risks, lack of trust between doctors and patients, and lack of cross-institutional collaboration mechanisms. At the same time, governance actions are obviously passive and lagging, unable to respond to the dynamic evolution of technology and performance feedback in a timely manner, and difficult to form an effective closed loop.

The application of LLMs in the healthcare industry represents a transformative force with great potential in improving diagnostic accuracy, personalizing treatment plans, and optimizing medical services. However, the healthcare field is inherently high-risk due to its direct impact on human life and health. The deployment of LLMs in such a sensitive field must be carefully managed to ensure their safety, effectiveness, and ethical standards.

DeepSeek is a general LLM system independently developed by China's top AI team DeepSeek Inc. Its research and development began in 2023 and is committed to breaking the boundaries of cognitive intelligence. As the first AI large model in China that fully matches the GPT-4 technical architecture, the DeepSeek series covers a complete model matrix from 7B to over 100 billion parameters and has reached the international leading level in core capabilities

such as mathematical reasoning, code generation, and multi-round dialogue. At present, multiple versions such as DeepSeek-R1, DeepSeek-V2, and DeepSeek-V3 have been derived, which are widely used in vertical fields such as intelligent customer service, educational assistance, and financial analysis. As a leading LLM project in China, DeepSeek has been deployed in actual medical environments, demonstrating its technical advantages and innovative capabilities. DeepSeek uses advanced natural language processing capabilities to efficiently analyze massive medical literature and provide strong support for doctors' clinical decision-making. Its excellent image processing performance further facilitates accurate medical imaging diagnosis. The DeepSeek case study has unique value for government governance research. It provides tangible evidence and insights that can provide a reference for the government to formulate regulatory strategies for LLM in the medical field. This study aims to explore a key question: In the face of the complexity and uncertainty of LLM governance, how can the government use dynamic mechanisms to ensure effective supervision?

By examining the operation and impact of DeepSeek in the medical field, this study aims to identify the challenges and opportunities facing LLM governance. The study will explore how the government can strike a balance between promoting technological progress and protecting patient welfare, and how to develop adaptive policies and regulatory frameworks to keep up with the rapid development of LLM technology. China is committed to promoting LLM technology innovation while building a sound legal and regulatory system. Its goal is to ensure the safe and compliant application of LLM in key areas such as healthcare. China's regulatory work aims to create an environment where technological advancement and patient safety coexist harmoniously.

In summary, the application of DeepSeek demonstrates the potential of Master of Laws (LLM) to revolutionize healthcare and the governance challenges that come with it and is suitable as a case study to study the application of DPG methods to LLM governance. It highlights the importance of dynamic governance mechanisms and provides a practical basis for the government to formulate effective regulatory strategies in this rapidly developing field.

6.1.2 Data Source

The data for this study mainly comes from semi-structured interviews with stakeholders in LLM governance in the Chinese medical industry where DeepSeek is deployed, and subsequent qualitative analysis. The interviewees covered multiple groups such as health management departments, doctors, medical institution managers, patients, and large model

developers, totaling 37 interviewees (34 interviews were coded, and the remaining 5 were used as test samples for saturation verification). The interview content focused on the views and concerns of all parties on the application of DeepSeek in the medical field, as well as expectations and suggestions for government governance.

After completing the interviews, through a systematic qualitative analysis of the interview materials, key viewpoints and themes were extracted. These results provide rich empirical evidence for this study, which helps to deeply understand the current status, challenges and needs of stakeholders in the application of large models in the medical industry, thus laying a solid foundation for exploring the effectiveness of government governance strategies.

After completing the interviews, the research team further designed a questionnaire based on the key points and themes extracted from the interviews, aiming to more widely collect the views and expectations of various stakeholders in the medical industry on the application of DeepSeek in the medical field. The content of the questionnaire closely revolves around the core issues involved in the interviews, covering a wide range of content from the convenience and accuracy of technology applications to data security, privacy protection, and expectations of government governance measures, striving to comprehensively and objectively reflect the overall attitude and needs of the industry.

Subsequently, the research team widely distributed 601 questionnaires in the medical industry, covering different medical institutions, departments, regions, and personnel roles, etc., to ensure the diversity and representativeness of the samples. After active collection work, 597 questionnaires were successfully collected, with a collection rate of up to 99%, which fully demonstrated the medical industry's high attention to this research topic and its willingness to actively participate. Through detailed analysis of these questionnaire data, it is possible to supplement and verify the results of qualitative analysis in the interviews from a quantitative perspective, further revealing the current status, challenges and stakeholder needs of the application of LLMs in the medical industry, providing a more solid and comprehensive empirical basis for exploring the effectiveness of government governance strategies, and making the research conclusions more scientific, reliable and universally applicable.

6.1.3 Adaptability of the DPG Framework to the Governance of Large Language Models in the Medical Industry

The application process of LLMs in the medical industry shows highly dynamic characteristics, such as the continuous evolution of model performance and data quality,

nonlinear changes in the trust between doctors and patients, and the blurring of the boundaries of responsibility and power among multiple subjects(Shool et al., 2025). These dynamic characteristics stem from the wide application of large models in multiple scenarios such as medical diagnosis, clinical decision support, health management, and medical research. They show great potential in improving diagnostic efficiency, optimizing resource allocation, and promoting medical innovation.

In current governance practices, there are widespread problems such as delayed response, lack of performance evaluation system, and lack of adaptability of governance strategies. On the one hand, the performance of the model continues to evolve after deployment, but the regulatory mechanism lacks flexibility and feedback mechanism, making it difficult to respond to model performance fluctuations and changes in the clinical environment in a timely manner; on the other hand, due to the involvement of multiple parties such as government regulatory agencies, hospital managers, doctors, patients, and technology providers, the governance structure is highly coupled, the responsibility boundary is blurred, the trust mechanism is fragile, and the risk of governance failure is significant. At the same time, ethical disputes coexist with legal ambiguity, such as unclear data ownership, insufficient algorithm transparency, and possible unfair resource allocation issues, all of which pose challenges to the current governance system.

In the face of the above problems, the DPG framework provides an analytical tool with theoretical basis and practical operability. DPG emphasizes performance-driven, feedback-centric, and collaboration-based mechanisms to build a large model governance structure that can respond dynamically and adapt to the system. It can effectively respond to the challenges of complexity and uncertainty in the application of large medical models by identifying key variables in the system, constructing a CLD, and clarifying the closed-loop path of performance feedback. For example, in the scenario of LLM-assisted imaging diagnosis, the DPG framework can track the causal closed loop of "model performance-doctor dependence-diagnostic accuracy-trust level-data feedback quality", thereby identifying intervention levers and optimizing governance mechanisms.

It is worth noting that the medical industry has a high degree of publicity, ethical sensitivity and technology dependence. Its governance goals cover multi-dimensional value pursuits such as safety, fairness and efficiency, which naturally fits the DPG governance concept of "performance-feedback-coordination". At the same time, the governance structure of the

medical industry is complex, including multiple subjects such as regulatory agencies, hospitals, doctors, patients, and technology companies. The feedback paths are intertwined, and the performance measurement is diversified, which further highlights the applicability of the DPG framework in dealing with multi-party interactions and nonlinear dynamic systems.

This study selected DeepSeek's application scenario in China's medical industry as a case study object, which is highly representative and theoretically applicable. China is not only at the forefront of the world in terms of medical resources, artificial intelligence infrastructure and policy environment, but also has accumulated rich pilot experience in smart medical care and medical LLM deployment. At the same time, this study obtained relevant institutional texts, interview materials and platform data, and has a good empirical research foundation. Through the DPG framework, the key performance drivers, feedback paths and collaborative structures in the governance practice of the big language model in China's medical industry are analyzed, which can provide a theoretical basis and policy inspiration for building a governance system with dynamic adaptability and system responsiveness.

In summary, the medical industry, as a typical representative of the extremely complex governance in the public service field, not only reflects the systematic governance dilemma faced by the application of big language models but also provides a highly adapted research scenario for the theoretical modeling and mechanism deduction of the DPG method.

6.2 Data and Methodology

This study mainly adopts DPG modeling and CLD analysis, combined with the integrated methodology of interviews and questionnaires, to systematically and deeply explore the governance issues of LLMs in the medical field, aiming to build a targeted and operational governance model to cope with the complex challenges faced in this field at this stage.

This study collects secondary data from multiple channels, including the official website of the National Health Commission, industry yearbooks, and AI medical enterprise reports, to comprehensively sort out the status of China's medical system and various scenarios of LLMs in medical applications. By analyzing these data, we can understand the overall development trend, policy orientation, and application status and potential opportunities of LLMs in the medical field, laying a macro background foundation for the entire study. Then, through interviews and data collection, we identify problems and clarify the research direction. Then, we code the interview data and extract key performance variables. Then, we use questionnaires

to verify the relationship between these variables to ensure the accuracy and relevance of the variables. Based on this, we construct a DPG model, clearly define performance drivers, and establish a CLD, deeply analyze the interaction mechanism between variables, and identify effective governance feedback mechanisms, providing a solid basis for the final formation of targeted policy recommendations.

6.2.1 Interview and Questionnaire

The interviewees included key players in the medical field, such as doctors, government regulators, technical parties, and ethics experts. Through open interviews, this study collected the governance challenges encountered by all parties in practical applications and obtained suggestions on the role that the government should play in governance. The interview content was analyzed using the open coding and axial coding methods of grounded theory to extract key governance variables and preliminarily construct the mapping framework of DPG.

Based on the core governance variables extracted from the interviews, a questionnaire in the form of a Likert scale was designed to collect the cognition of these variables for key groups (including patients, medical workers, etc.). Through content analysis, frequency induction, basic statistics and other analytical methods, the relationship between variables was further verified to provide a quantitative basis for the construction of the DPG governance model and the formation of a causal closed loop.

In the interview design of this study, open coding and axial coding were used for the interviewees such as doctors, government regulators, technical parties, and ethics experts, and then the variables were extracted, and the DPG mapping was constructed. The logic of questionnaire design is closely centered around more than 20 core governance variables extracted from the interviews, and the Likert scale is used to collect the cognitive status of key groups. In terms of analysis methods, content analysis and frequency induction are used in combination. At the same time, with the help of basic statistics of the questionnaire, path or causal diagram mapping analysis can be further carried out to deeply explore and present the relationships and laws behind the data.

6.2.1.1 Design and Implementation of Semi-Structured Interviews

To deeply explore the dynamic characteristics and key challenges of the governance practice of LLMs in the medical industry, this paper further adopts the method of semi-structured interviews(Adeoye-Olatunde & Olenik, 2021) based on literature review and policy analysis to supplement and deepen the understanding of the actual governance process.

The interview aims to obtain first-hand empirical data, understand the technical, ethical, institutional and collaborative issues faced by large models in medical scenarios during actual deployment and use, and explore the understanding and needs of multiple subjects for the governance mechanism. The interview questions are designed around four aspects: model application, governance mechanism, collaboration barriers and performance feedback, reflecting the deep fit with the DPG framework of this study.

This study conducted a total of 37 semi-structured interviews, and the interviewees covered four key governance stakeholders: the head of the information department of a tertiary hospital, clinical frontline doctors, medical artificial intelligence enterprise managers, and patient representatives. Although the number of interview samples is limited, its composition is highly representative in terms of role diversity and practical participation, which helps to capture the governance dilemma and institutional tension from a multi-perspective perspective.

The interviews were conducted in a combination of online video conferencing and offline face-to-face interviews, with each interview lasting 30–50 minutes. Some interviews were recorded with the permission of the interviewees, and transcribed and proofread afterwards. All interviewee information has been anonymized to ensure research ethics and data security.

The different professional attributes and practical experience of the interviewees can reflect the differences in their answers to the medical LLM in terms of diagnostic recommendations, data privacy, ethical standards, and technology applications. The groups directly involved in medical services, technology research and development, or industry management are selected. They are most familiar with the pain points and needs of the medical LLM in actual applications. The data obtained from the interviews are the most authentic and reliable and have targeted and practical guidance significance. Through these representative interviewees, we conducted comprehensive interviews from different angles on the functional applicability, diagnostic accuracy, data security, ethical risks, and primary medical adaptation of the medical LLM to ensure the representativeness of the interview. This group was selected because they come from different professional fields and represent the cognition and evaluation of the medical LLM from different perspectives such as patient needs, policy supervision, technology development, and clinical practice. This selection is intended to make the research results more reliable and valid, and at the same time provide solutions from multiple perspectives to explore practical problems such as functional optimization, standardized governance, and improvement of medical service quality of the medical LLM. Because the

results of the interview are very clear and true, the following table 6.1 is the basic information of the 34 interviewees.

Table 6.1 the basic information of interviewees

Number	Identity	Age	Title	Number	Identity	Age	Title
HZ01	Patient	-	-	YY09	Doctor	45	-
HZ02	Patient	26	-	YY10	Doctor	26	Junior title
WJ01	Health Commission official	54	-	YY11	Doctor	34	Mid-level title
WJ02	Health Commission official	50	-	YY12	Doctor	40	-
WJ03	Health Commission official	48	Senior title	YY13	Doctor	45	-
WJ04	Health Commission official	36	Senior title	YY14	Doctor	-	-
WJ05	Health Commission researcher	36	Mid-level title	YY15	Doctor	47	Mid-level title
YF01	Large model developer	-	-	YY16	Hospital manager	54	-
YF02	Large model developer	-	-	YY17	Head nurse	-	Mid-level title
YY01	Doctor	27	Junior title	YY18	Hospital manager	-	-
YY02	Doctor	28	Junior title	YY19	Hospital manager	39	-
YY03	Doctor	35	Mid-level title	YY20	Doctor	60	Senior title
YY04	Doctor	35	-	YY21	Doctor	25	-
YY05	Doctor	30	-	YY22	Doctor	-	-
YY06	Doctor	37	-	YY23	Doctor	42	-
YY07	Doctor	30	Mid-level title	YY24	Doctor	-	-
YY08	Doctor	40	-	YY25	Doctor	43	Mid-level title

The interview content provides a relatively systematic observation from the dimensions of

model design, deployment, clinical application to risk supervision. Preliminary analysis shows that multiple interviewees have reported highly consistent issues in terms of the "black box" of the model, unclear responsibilities, difficulty in data sharing, and delayed regulatory feedback, showing an obvious trend of information saturation. This phenomenon verifies the effectiveness of the interview sample in explaining the actual problems of LLM governance in the medical industry and provides rich original materials and practical perspectives for subsequent grounded theory analysis and CLD modeling.

6.2.1.2 Encoding

To deeply explore the potential logical structure and key mechanisms of the governance practice of the LLMs in the medical industry, this paper adopts Grounded Theory as the methodological basis for qualitative analysis. Grounded theory was first proposed by Glaser and Strauss in 1967 (Glaser & Strauss, 1967). It emphasizes the induction of theory from field data and abstracting the original text into theoretical concepts through a systematic coding process. It is particularly suitable for exploratory research on emerging complex social and technological issues. In the relatively cutting-edge research field of LLM medical governance, which lacks a mature theoretical paradigm, Grounded Theory provides a solid methodological support for theoretical construction.

In this study, we used NVivo software to assist in grounded theory research and efficiently completed the preliminary data collation and analysis. The research process of grounded theory covers the stages of data collection, coding, conceptualization, and theory generation. Through in-depth analysis of the data, researchers gradually form an understanding and interpretation of the phenomenon and finally draw conclusions with a certain theoretical depth. Although NVivo software does not have inherent grounded theory, it is a flexible tool that can be combined with a variety of research methods and theories such as grounded theory, providing strong support for researchers to deeply explore social phenomena.

This study uses the classic "three-stage coding method" proposed by Corbin and Strauss for data processing and theoretical construction, which specifically includes three stages: open coding, axial coding, and selective coding (Corbin & Strauss, 1990):

1. Open Coding

In this stage, the researcher conducts a detailed analysis of the interview transcripts sentence by sentence, identifies key events, attitudes, processes, and structural issues that

appear in the text, and abstracts them into basic concept labels. For example, feedback from doctors was coded as “low credibility of model results” and “unclear responsibility for use”; statements from enterprises generated labels such as “ambiguous supervision” and “lack of algorithm testing”. Preliminary coding extracted about 22 sets of labels to form a concept pool.

2. Axial Coding

In this stage, researchers classified and aggregated similar or related concepts in open coding to form several superordinate “categories”. Combined with the theoretical framework of DPG, the main categories were summarized as: “lack of trust mechanism”, “uncertainty of governance system” and “obstacles to collaboration among multiple subjects”. Each category not only presents an independent governance dilemma, but also reveals the possible structural correlation between variables, laying the foundation for the subsequent construction of CLDs and conceptual models.

3. Selective Coding

In the selective coding stage, the focus of the research shifted to the integrated construction of the “DPG chain”, focusing on the core theme of “problem identification-performance monitoring-feedback adjustment”. The researchers further integrated the previous categories into several causal path groups, forming a governance system landscape covering dimensions such as technology trust, performance evaluation, institutional adaptation and interest coordination, and made a preliminary visualization in the form of CLD, showing the structural potential of DPG logic in medical scenarios.

The results show that no new concepts with essential differences were found after the 34 interviewees, indicating that the main variables have tended to theoretical saturation, which can provide a solid empirical basis for subsequent model construction and policy recommendations.

6.2.2 DPG and CLD

Based on data collection and preliminary analysis, we construct a DPG model. The model starts with strategic resources. These resources are the key elements that organizations can use to achieve their goals, covering human, material, financial and technical resources, etc. They provide basic support for the entire government’s activities. Performance drivers, as key variables that affect performance results, are the core driving force for promoting the governance process. They include multi-dimensional factors such as policy guidance, technological development, and market demand. These factors interact and influence each other

and jointly drive governance activities towards the predetermined goals. Under the influence of performance drivers, a series of intermediate results will be generated. These preliminary results are the direct reflection of governance activities at a specific stage, such as the initial results of technology application, the optimization of some processes, etc. They reflect the phased progress and effectiveness of governance activities. The end-result is the goal that the organization expects to achieve and the ultimate direction of all governance activities, such as the comprehensive improvement of medical service quality and the efficient allocation of medical resources. At the same time, feedback mechanisms run through the entire governance process. By monitoring and evaluating the results and problems of each stage, strategic resources and performance drivers are adjusted in a timely manner to achieve continuous improvement and optimization, so that governance activities can dynamically adapt to the changing internal and external environment and ensure the achievement of governance goals.

In the process of building the DPG model, the CLD analysis method is used to deeply explore the complex causal relationship between these variables to form a dynamic, feedback-driven governance closed loop. Through CLD analysis, we can clearly depict the interaction mechanism between strategic resources, performance drivers, intermediate results and end-results, identify key causal loops and feedback mechanisms, and thus more accurately grasp the operating logic of the governance system. For example, the input of technical resources (strategic resources) will promote technological innovation (performance drivers), thereby improving the efficiency and quality of medical services (intermediate results) and ultimately promote the modernization of the medical system (end-results). At the same time, the feedback mechanism will feedback and adjust the direction and intensity of technical resource investment, as well as the key areas of technological innovation, based on the evaluation of intermediate and end-results, to ensure that governance activities can be continuously optimized and improved, forming a dynamic and self-improving process. In this way, we can simulate different governance scenarios and predict the chain reactions that may occur from policy interventions and management decisions, thereby providing a scientific basis for optimizing governance strategies.

In summary, this study uses a variety of methods to comprehensively analyze the complex landscape of LLM governance in the medical field, ensure the reliability and effectiveness of the research results, provide a solid theoretical and empirical basis for improving the governance mechanism in this field, and help LLMs in the medical field achieve sustainable and efficient development.

6.2.3 Analytical Strategy

The coding results are further used to code results of the in-depth interviews and combined with the questionnaire data, a CLD was constructed to reveal the causal relationship and dynamic feedback mechanism between key variables in the governance of LLM in the medical industry. By analyzing the high-frequency concepts in the qualitative interviews and the patterns identified in the questionnaire survey, the researchers identified the key governance variables and their interactions, and preliminarily drew a variable interaction network, laying the foundation for the empirical study of CLD.

This study uses the DPG framework as the analytical framework to systematically integrate and construct the relationship between variables and construct a dynamic causal structure to reflect the actual logic and operating dynamics of LLM governance in the medical context. In the modeling process, special attention is paid to the dynamic attributes and feedback loops in the performance mechanism. By analyzing the positive and negative causal paths, potential performance amplification or inhibition effects are identified. Under the guidance of the DPG framework, the model constructs an adaptive governance structure with performance feedback as the core, which can dynamically reflect the evolution and adjustment space of governance behavior. This structure provides a theoretical and systematic basis for diagnosing the current governance dilemma and proposing improvement strategies.

6.3 Case Analysis

6.3.1 Identification of Variables Vased on Interviews

This study takes the DPG framework as its theoretical basis, focuses on the governance of LLMs in the medical industry, combines semi-structured interviews with grounded theory methods, and through open coding and axial coding of the interview content, identifies a series of highly related governance issues, extracts multiple key variables that affect governance performance, and summarizes and classifies them according to the three core dimensions of "strategic resources-performance drivers-governance results". "Strategic resources" include human resource allocation, institutional support, etc.; "performance drivers" include technology attractiveness, policy acceptance, policy severity, and social impact, etc.; "end-results" include policy governance level and large model technology development level.

The interviewees generally reflected that they faced a series of governance challenges in the actual application process, including: insufficient understanding of model professional

capabilities, ambiguity of model application boundaries, reliance on leadership will in the promotion process, differences in risk control mechanisms within the organization, unclear attribution of model responsibilities, users' mistrust of models, lack of industry governance standards, high expectations for model ethics and safety, and frontline doctors' demands for participation in decision-making, etc.

The study found that the respondents generally reflected that the current LLM governance faces three major challenges: First, there is ambiguity and conflict in governance goals, especially the tension between efficiency and ethics, safety and innovation; second, the asymmetry of technology, ethics and manpower in resource allocation is prominent, especially in terms of AI application dominance and knowledge mastery; third, the uncertainty of governance mechanisms in responsibility division, feedback channels and standard setting leads to trust gaps and execution differences in actual applications.

Based on the above interview results, the study constructed a DPG analysis structure for LLM governance in the medical industry, systematically presenting the distribution and correlation of key governance variables in the "strategic resources-performance drivers-governance results" dimension, as the basis for subsequent causal relationship modeling and policy recommendation design, This article selects "data accuracy", "algorithm optimization", "supporting system" and "technical update" in the core category of "technical algorithm" as examples to visualize part of the coding process, as shown in Table 6.2.

Table 6.2 Encoding Process Example

Original Material	Open Coding	Axial Coding	Selective coding
YY01: First of all, in terms of the data source, theoretically, the data for AI model training should come from legal and compliant channels.	Data Source	Data Accuracy	Technical algorithm
WJ01: it was found that there is a lack of explicit regulations on small sample data training. For example, a company developed a blood disease AI model with less than 200 samples, and the legal validity of its diagnostic suggestions is controversial.	Data Training		
YY04: The AI model system may still use nationally general diagnostic data, leading to a missed diagnosis rate of 15% - 20% for mild thalassemia patients.	Data Scope		

YY07: There may be differences in the quality and standards of data from different medical institutions, which can affect the accuracy and reliability of model training. For instance, some primary medical institutions may have irregular records and incomplete information.	Data Quality	
YY01: When integrating information, it may prioritize recommending conservative or traditional treatment plans, lacking timely updates and comprehensive consideration of emerging therapies.	Data Update	
WJ05: At the same time, the transparency of algorithms can be enhanced by establishing an explainable algorithm whitelist, requiring the decision - making path of medical imaging diagnostic models to be traceable.	Decision - making Process Tracing	Algorithm Optimization
WJ03: Some AI medical systems have insufficient decision - making transparency. Their algorithm logic is not open to the public, so patients cannot understand the basis of diagnosis, which may infringe on their right to know and choose.	Transparency of Data Processing	
YY07: For some rare or complex diseases, the model's plan may lack personalized consideration and fail to fully combine with the patient's specific situation.	Identification of Rare and Complex Diseases	
YY09: Minority patients have unique cultural characteristics and religious beliefs, but the model usually cannot understand and consider these factors.	Cross - cultural Adaptability	
WJ02: In terms of data scenario adaptation, model developers need to optimize medical semantics understanding for non - English environments such as Chinese to reduce the risk of language misjudgment.	Algorithm Universality	
WJ02: First, in terms of data collection legality, when collecting data, patient data must be protected.	Data Collection Legality Review	Supporting Systems
WJ05: Currently, the existing systems and regulatory explorations have formed a preliminary framework for classified regulation, along with risk - graded management.	Data Classification and Grading Management	

WJ05: While not accepting the national supercomputing center, on the legal level, we must strictly abide by the Data Security Law. For sensitive genetic data, in addition to implementing relevant measures, we must also comply with Article 29 of the Personal Information Protection Law. Medical AI data collection must follow the principle of informed consent.	Informed Consent Principle for Data Use	
WJ01: For example, the maternity department of Longchuan County Maternal and Child Health Hospital has an AI prediction system. Its training data is stored in the local Jingyun Cloud and is subject to cross - border data transmission audits.	Cross - border Data Transmission Audit	
WJ01: For private institutions, data annotation must be completed under the supervision of the Municipal Health Medical Big Data Center.	Regulation of Data Annotation for Private Institutions	
WJ01: Jointly with the Municipal Justice Bureau, we formulated the Low - probability Medical AI Conclusion Application Guidelines, which require such cases to be signed off by three physicians of associate senior or above titles.	Low - probability Data Application Regulations	
YY09: Measures such as anonymizing data, removing identifiable patient information, and using encryption technologies can enhance the security of data during transmission and storage.	Data Privacy Technology Application	
WJ02: Multi - department coordinated data regulation is needed. Departments such as the Health Commission, Cyberspace Administration, and Science and Technology Commission should work together to build a regulatory network.	Multi - department Coordinated Data Regulation	
WJ05: There is a lack of dynamic risk assessment. The existing three - level ethical review mainly focuses on pre - launch model assessment, but there are gaps after the model is trained.	Dynamic Data Risk Assessment	
WJ03: The sensitivity of medical data is graded. Anonymization must meet national standards.	Anonymization Technology Standards	
WJ01: For example, the emergent capabilities of large models, such as cross - disease correlation analysis, can break through existing regulatory	Algorithm Iteration Speed	Technological Updates

frameworks. The current classified regulation system struggles to adapt.

WJ01: In pilot programs at primary medical institutions, it was found that a model's training data bias led to a tendency to focus on treatment rather than prevention. After identifying the issue, the manufacturer was ordered to correct it, and a dynamic monitoring mechanism was established.

Dynamic
Monitoring
Mechanism

After applying NVIVO 14 coding, improving the governance performance of large models - taking the large model of the medical industry as an example, a total of 652 reference points were generated, which came from 34 interview materials, including 143 reference points of technical algorithms, 29 material sources; 205 reference points of clinical adaptation, 32 material sources; 69 reference points of clear responsibilities, 19 material sources; 81 reference points of policy governance and supervision, 26 material sources; 100 reference points of data security and privacy protection, 25 material sources; 54 reference points of improvement suggestions, 17 material sources. Improving the governance performance of large models - taking the large model of the medical industry as an example, the frequency of main specific coding nodes is shown in Table 6.3. The complete coding nodes are provided in Appendix B.

Table 6.3 Coding Nodes Example

Selective coding	Axis coding	Open coding	Number of material sources	Number of reference points
Technical algorithms	Data accuracy		29	143
		Data sources	19	19
		Data training	8	8
		Data scope	5	5
		Data quality	6	6
		Data update	5	5
	Algorithm optimization		24	49

	Decision-making process traceability	10	10
	Transparency of data processing process	4	4
	Identification of rare diseases and miscellaneous cases	12	12
	Adaptability to cross-cultural scenarios	20	20
	Universality of algorithms	3	3
Supporting system		20	41
	Review of the legality of data collection	3	3
	Data classification and grading management	11	12
	Principle of informed consent for data use	9	9
	Audit of cross-border data transmission	3	3
	Data labeling supervision of private institutions	1	1
	Application specifications of low-probability data	1	1
	Application of data privacy technology	6	6
	Multi-department collaborative data supervision	1	1
	Data dynamic risk assessment	2	2
	Anonymization technology standards	3	3
Technology updates		8	10
	Algorithm iteration speed	7	7

Dynamic monitoring mechanism	3	3
---------------------------------	---	---

Responsibility clarification refers to the clear definition of the rights, obligations and risks of the government, medical institutions, doctors, technology companies and other entities in data security, diagnosis and treatment decisions, ethical compliance and other aspects in the application scenarios of the medical big language model. Policy governance and supervision involve the government and relevant departments to regulate and supervise the entire process of technical research and development and application of the large medical model through laws and regulations and regulatory mechanisms. Data security and privacy protection aims to protect the confidentiality, integrity and availability of patient health data and related information, and prevent data leakage, tampering or abuse. Clinical adaptation is the matching and adaptation process of the medical big language model with clinical needs, doctor's operating habits and patient characteristics. The core is to optimize the model design so that its output conforms to the medical process specifications. Algorithm specifications involve the logical rules and calculation methods used in the medical LLM to process medical data and generate diagnosis and treatment recommendations. It must follow medical principles, embed the latest diagnosis and treatment guidelines, and update data regularly to ensure the accuracy of the output recommendations.

6.3.2 Identification of Causal Relationships Based on Questionnaires

The coding results were further used to create a questionnaire survey, and regression analysis was performed through the collected questionnaire survey results. Finally, a CLD was constructed to identify the key variable relationships and dynamic feedback paths in the governance of the LLM in the medical industry. The researchers first drew a preliminary variable interaction network based on the consensus variables and mechanism descriptions that appeared frequently in the interview materials and then used the DPG framework for system integration to extract the dynamic structural logic of the governance of the medical large model. This survey focuses on the application and governance of medical LLM, aiming to collect the cognition, attitudes and expectations of various stakeholders, and provide rich empirical evidence and diverse perspectives for subsequent research and policy making. This study distributed questionnaires through the Questionnaire Star platform, covering a wide range of relevant groups. A total of 601 questionnaires were distributed, and 597 valid questionnaires were collected, with an effective recovery rate of 99.3%. The survey covered stakeholders such as patients and medical staff, ensuring the diversity and representativeness of data sources.

Through the analysis of questionnaire data, this study hopes to accurately grasp the status and future trends in the field of medical LLM, gain in-depth insights into key issues and challenges in the governance process, and thus contribute valuable insights and suggestions to promote the healthy and sustainable development of medical LLM and improve its governance system. The national distribution of the questionnaire survey results is shown in the figure below.

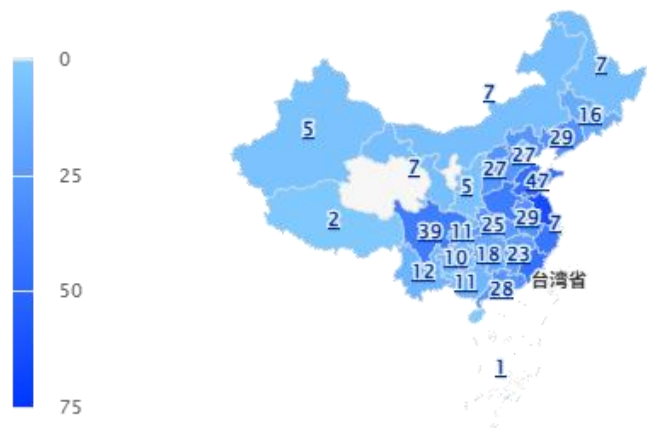


Figure 6.1 Questionnaire Survey Distribution

Table 6.4 Cronbach reliability analysis

Number of items	Sample size	Cronbach α coefficient
40	597	0.990

As shown in table 6.4 above, the reliability analysis results of the questionnaire show that this survey has high reliability. The sample size is 597, and the Cronbach α coefficient reaches 0.990, indicating that the questions in the questionnaire can well measure the same concept or construct, and the questionnaire has a very high internal consistency. In terms of sample integrity, there is no missing sample, ensuring the integrity and reliability of the data.

Table 6.5 KMO and Bartlett's test

KMO value	0.995
Bartlett's sphericity test	Approximate chi-square 26621.402

<i>df</i>	780
<i>p-value</i>	0.000

As shown in the validity analysis results in table 6.5 above, the KMO test result is 0.995, indicating that the partial correlation between variables is strong and the data is very suitable for factor analysis. The statistics of the Bartlett sphericity test is 26621.402, the degree of freedom is 780, and the corresponding p value is 0.000, indicating that the correlation matrix is not a unit matrix, and there is a strong correlation between the variables, which is suitable for factor analysis.

Table 6.6 Relationships between Variables

Variable	Coefficient	Question	Variable	Coefficient	Question
LLM Industry Capacity— LLM Safety Level	0.735**	Q9—Q33	Effective Supervision— LLM Technology Development	-0.735**	Q27—Q3
LLM Industry Capacity— LLM Technology Development	0.717**	Q9—Q3	LLM Safety Level— Content Matching	0.706**	Q33— Q17
LLM Technology Development—Content Matching	0.727**	Q3—Q17	Clinical Suitability—No. of User	0.672**	Q37— Q42
Content Matching—Clinical Suitability	0.728**	Q17— Q37	Clinical Suitability— Public trust on Government	0.654**	Q37— Q27
Clinical Suitability—Public Trust on Government	0.654**	Q37— Q27	Public trust—Policy Acceptance	0.666**	Q27— Q40
Database Size—Value Matching	0.772**	Q32— Q30	No. of User—Stakeholders Engagement Level	0.665**	Q42— Q31
LLM Safety Level—Value Matching	0.759**	Q33— Q30	No. of User—Database Size	0.683**	Q42— Q32
Policy Acceptance— Governance Level	0.723**	Q40—Q7	No. of User—Social and Economic Benefits	0.674**	Q42— Q36

Governance	Level—	0.641**	Q7—Q27	Social and Economic	0.666**	Q36—Q9
Effective Supervision				Benefits—LLM Industry Capacity		
Effective Supervision—		0.693**	Q27—	Stakeholders Engagement	0.662**	Q31—
LLM Safety Level			Q33	Level—Effective Supervision		Q27
Value Matching—Clinical Suitabil		0.748**	Q30—			
			Q37			

As shown in table 6.6 above, the capacity of the large model industry is positively correlated with the safety level of the large model (the coefficient is 0.735), which can be investigated through questions Q9 (questions related to the capacity of the large model industry) and Q33 (questions related to the safety level of the large model). That is, the larger the capacity of the large model industry, the higher its safety level. Effective supervision is negatively correlated with the development of large model technology (the coefficient is -0.735), corresponding to questions Q27 (questions related to effective supervision) and Q3 (questions related to the development of large model technology). This shows that in the case of insufficient effective supervision, the development of large model technology may have certain hidden dangers. The capacity of the large model industry is positively correlated with the development of large model technology (the coefficient is 0.717), corresponding to questions Q9 and Q3, indicating that the expansion of the capacity of the large model industry is conducive to the advancement of large model technology. The safety level of the large model is positively correlated with the content matching degree (the coefficient is 0.706), corresponding to questions Q33 and Q17 (questions related to the content matching degree), that is, the improvement of the safety level of the large model is conducive to improving the content matching degree. The development of LLM technology is positively correlated with content matching (coefficient is 0.727), corresponding to questions Q3 and Q17, which means that the continuous progress of LLM technology can improve content matching. Content matching is positively correlated with clinical applicability (coefficient is 0.728), corresponding to questions Q17 and Q37 (clinical applicability related questions), indicating that the improvement of content matching is conducive to enhancing the applicability of LLMs in clinical scenarios. Clinical applicability is positively correlated with the number of users (coefficient is 0.672), corresponding to questions Q37 and Q42 (user number related questions), indicating that the stronger the applicability of LLMs in clinical practice, the more users they

have. Clinical applicability is positively correlated with the public's trust in the government (coefficient is 0.654), corresponding to questions Q37 and Q27 (public trust in the government related questions), that is, the good applicability of LLMs in clinical practice can enhance the public's trust in the government's work in LLM governance. The database scale is positively correlated with the value matching (coefficient is 0.772), corresponding to questions Q32 (database scale related questions) and Q30 (value matching related questions), indicating that the expansion of the database scale is conducive to improving the value matching of the large model. The security level of the large model is positively correlated with the value matching (coefficient is 0.759), corresponding to questions Q33 and Q30, indicating that the improvement of the security level of the large model is conducive to improving the value matching. The number of users is positively correlated with the stakeholder participation (coefficient is 0.665), corresponding to questions Q42 and Q31 (stakeholder participation related questions), that is, the more users of the large model, the higher the stakeholder participation. The number of users is positively correlated with the database scale (coefficient is 0.683), corresponding to questions Q42 and Q32, which means that the increase in the number of users will promote the expansion of the database scale. The number of users is positively correlated with social and economic benefits (coefficient is 0.674), corresponding to questions Q42 and Q36 (socio-economic benefits related questions), that is, the more users of the large model, the greater the social and economic benefits it generates. The socio-economic benefits are positively correlated with the capacity of the large model industry (coefficient is 0.666), corresponding to questions Q36 and Q9, indicating that the socio-economic benefits brought by large models are conducive to promoting the expansion of the capacity of the large model industry. Stakeholder participation is positively correlated with effective supervision (coefficient is 0.662), corresponding to questions Q31 and Q27, that is, the improvement of stakeholder participation can enhance the effective supervision of large models. Policy acceptance is positively correlated with governance level (coefficient is 0.723), corresponding to questions Q40 (policy acceptance related questions) and Q7 (governance level related questions), indicating that high policy acceptance is conducive to improving the governance level of large models. Governance level is positively correlated with effective supervision (coefficient is 0.641), corresponding to questions Q7 and Q27, that is, the improvement of the governance level of large models can strengthen effective supervision. Effective supervision is positively correlated with the safety level of large models (coefficient is 0.693), corresponding to questions Q27 and Q33, indicating that effective regulatory measures can ensure the safety

level of large models. The value match is positively correlated with clinical applicability (the coefficient is 0.748), corresponding to questions Q30 and Q37, that is, the improvement of value match is conducive to enhancing the applicability of large models in clinical scenarios.

To deeply understand the dynamic influence mechanism of the medical big language model (LLM) in the governance framework, this study explored the causal relationship between multiple variables based on linear regression path analysis. The focus is on analyzing how factors such as financial investment, governance mechanism, data resources, social effects, and public trust affect the final value and usage level of the model through interaction, thereby constructing a complete dynamic performance closed loop.

Overall, financial resources, technological development and governance capabilities constitute the structural basis of model governance, while data coverage, social impact and user behavior drive the dynamic evolution of model value. At the same time, the interaction between value gap and public trust constitutes a "soft feedback mechanism" in model governance. The above results show that the governance of medical LLMs should not only focus on "visible" technology and capital investment but also pay attention to "invisible" social perception and value recognition, to build a performance-oriented, people-oriented, closed-loop dynamic governance model.

6.3.3 DPG Framework Construction

The DPG model constructed based on the results of the interviews and questionnaires is shown in the figure below. The figure shows the general model of DPG in the medical field. First, the level of LLM technology development and security level affect the matching degree between content and actual needs. Improved content matching improves user satisfaction, attracts more users, and the increased number of users brings more social and economic benefits, which promotes the improvement of LLM technology development level. Second, the increase in LLM technology security level enhances the matching degree of values, makes the output consistent with social ethics, improves user satisfaction and attracts users, and the increase in the number of users brings social and economic benefits, improves the level of government governance, and then promotes the introduction of effective regulatory measures to ensure the safety of LLM technology. Third, the improvement of the government's LLM governance level promotes the implementation of effective supervision and ensures the safety of LLM technology. The improvement of security level enhances the matching degree of values, improves user satisfaction, enhances the public's trust in LLM, improves policy acceptance,

and further improves the level of government governance. Fourth, the active participation of stakeholders provides support for effective supervision and improves the level of LLM technology development. The improvement of technology development level attracts more stakeholders to participate, forming a virtuous circle.

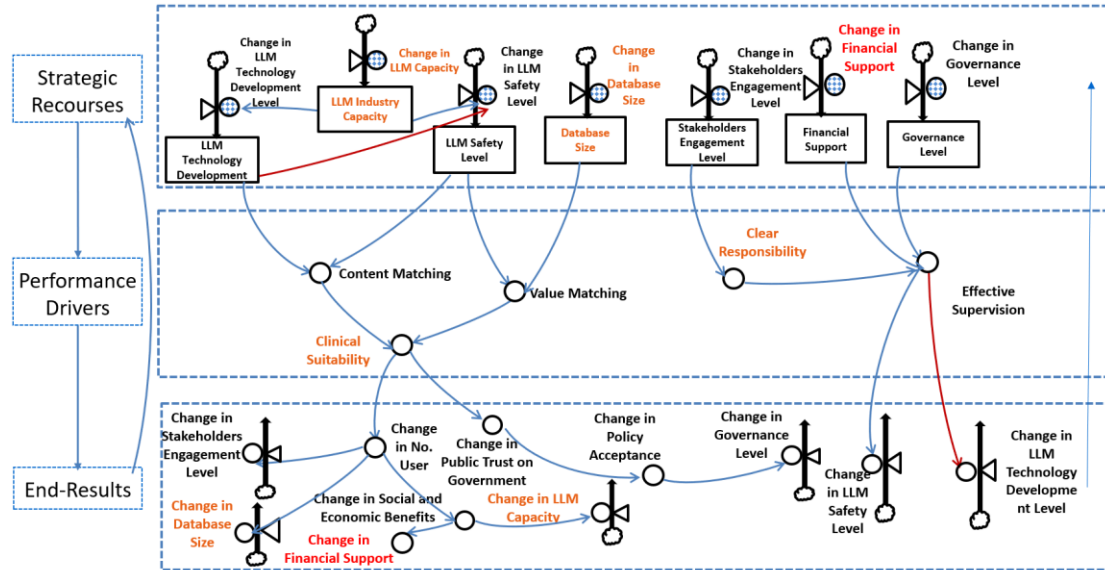


Figure 6.2 DPG Framework on Medical LLM Governance

Strategic resources include LLM technology development level, security level, database size, stakeholder engagement, financial support and governance level. LLM technology development (Q3) drives industry capacity (Q9, $r=0.717$) and content fit (Q17, $r=0.727$). Security level (Q33) is associated with industry capacity (Q9, $r=0.735$), content fit (Q17, $r=0.706$) and value fit (Q30, $r=0.759$). Database size (Q32) promotes value fit (Q30, $r=0.772$) and number of users (Q42, $r=0.683$). Stakeholder engagement (Q31) increases with the number of users (Q42, $r=0.665$) and helps effective supervision (Q27, $r=0.662$). Financial support (Q1) ensures development and attracts participants (Q31). Governance level (Q7) improves policy acceptance (Q40, $r=0.723$) and strengthens effective supervision (Q27, $r=0.641$).

Performance drivers include content fit, value fit, clinical applicability, and effective supervision. Content fit (Q17) is affected by LLM technology (Q3, $r=0.727$) and security level (Q33, $r=0.706$), and improves clinical applicability (Q37, $r=0.728$). Value fit (Q30) is associated with database size (Q32, $r=0.772$), security level (Q33, $r=0.759$), and clinical applicability (Q37, $r=0.748$). Clinical applicability (Q37) affects the number of users (Q42, $r=0.672$) and public trust (Q27, $r=0.654$). Effective supervision (Q27) is affected by stakeholder engagement (Q31,

$r=0.662$), governance level (Q7, $r=0.641$), and ensures safety level (Q33, $r=0.693$).

The end-results involve the number of users, socioeconomic benefits, public trust, policy acceptance, and industry capacity. The number of users (Q42) is affected by multiple factors, such as content matching (Q17), value matching (Q30), clinical applicability (Q37, $r=0.672$), database size (Q32, $r=0.683$), and stakeholder engagement (Q31, $r=0.665$), and its increase promotes socioeconomic benefits (Q36, $r=0.674$). Socioeconomic benefits (Q36) also interact with industry capacity (Q9, $r=0.666$). Public trust (Q27) is affected by clinical applicability (Q37, $r=0.654$), which is related to policy acceptance (Q40). Policy acceptance (Q40) is affected by governance level (Q7, $r=0.723$). Industry capacity (Q9) is driven by technology (Q3, $r=0.717$), safety level (Q33, $r=0.735$) and economic benefits (Q36, $r=0.666$).

6.3.4 Loop Analysis and Identification of Governance Intervention Mechanisms

Based on the DPG structure, this paper further identifies the main variables and the dynamic relationship between them through the CLD modeling method. According to the causal path extracted by the rooted analysis, a preliminary CLD of the LLM governance in the medical industry is constructed (Figure 6-X). The figure shows that different governance mechanisms form positive/negative feedback loops by affecting trust, performance perception and usage behavior. Through structural analysis of the feedback loops in the CLD diagram, 8 reinforcing loops and 4 balancing loops were identified. Among them, the improvement of the development level of LLMs leads to the reduction of the safety level of LLMs, which is a key chain that needs to be governed, and the improvement of effective governance can alleviate this problem.

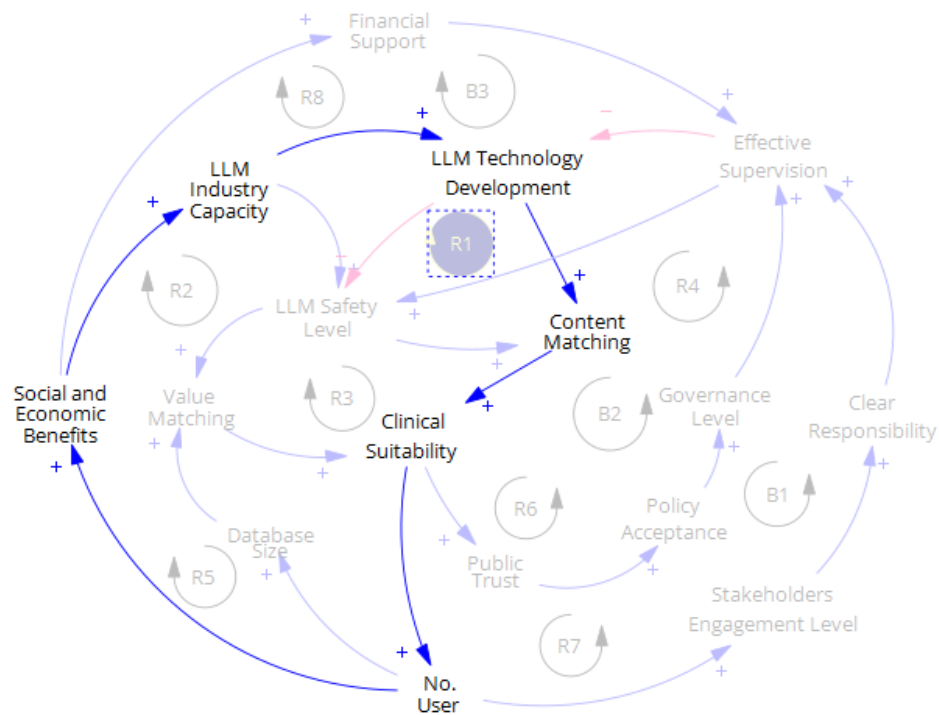


Figure 6.3 R1: The role of LLM technology development in content matching and social and economic benefits

The causal loop R1 in figure 6.3 includes the following variables: LLM Industry Capacity, LLM Technology Development, Content Matching, Clinical Suitability, Number of Users, and Social and Economic Benefits. This loop reflects that the expansion of the capacity of the large model industry promotes technological development, and technological progress improves content matching, which in turn enhances clinical suitability, attracts an increase in the number of users, drives stakeholders to participate more actively, promotes effective supervision, and improves the safety level of large models, ultimately further expanding the industry capacity and forming a positive feedback loop.

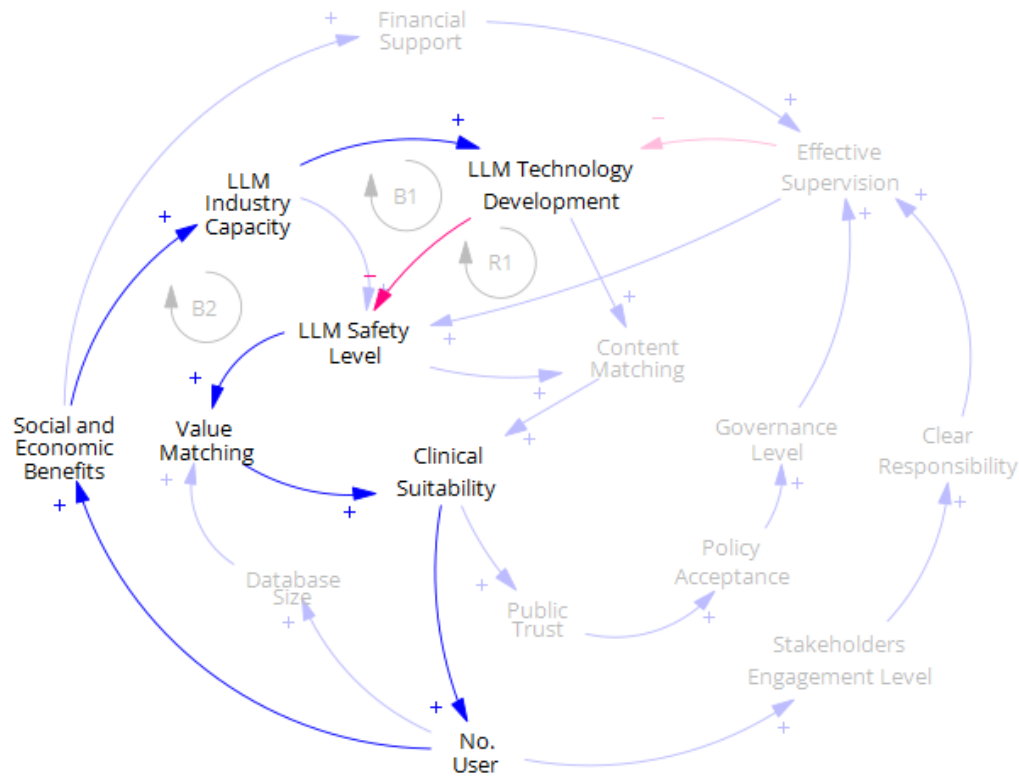


Figure 6.5 B2: The role of LLM technology development in content matching and social and economic benefits

Causal loop B2, shown on figure 6.5 is very similar to B1 and includes the following variables: large model development level, large model safety level (LLM Safety Level), value matching, clinical adaptability, number of users, large model industry capacity (LLM Industry Capacity), and social and economic benefits (Social and Economic Benefits). This loop reflects that the improvement of the development level of large models will bring more large model safety issues, thereby reducing value matching, clinical adaptability, reducing the number of users, reducing social and economic benefits, and then reducing the capacity of the large model industry, ultimately inhibiting the development level of large models, forming a reverse feedback loop.

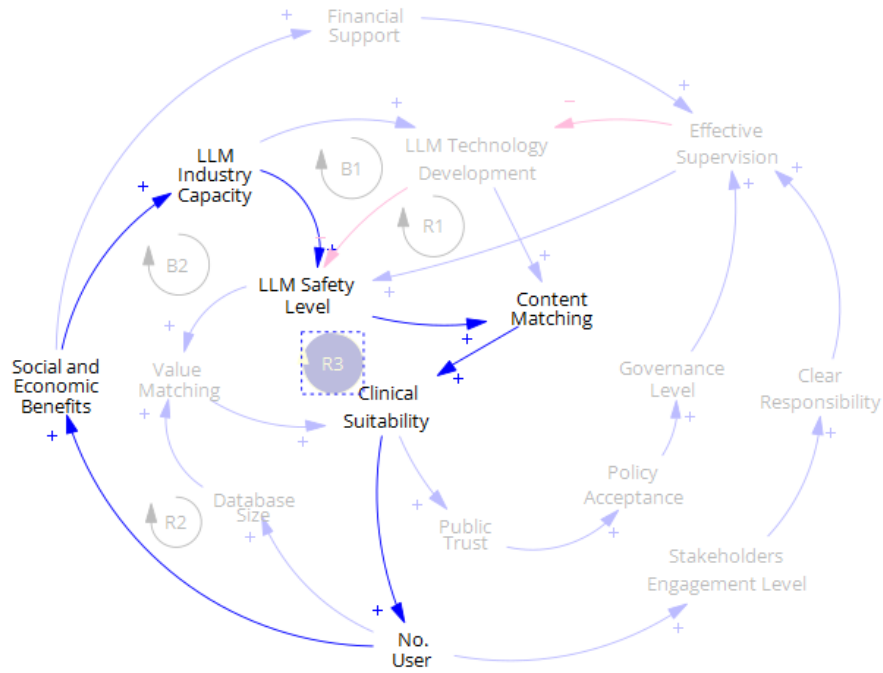


Figure 6.7 R3: LLM safety level improvement content matching

The causal loop R3, figure 6.7, in the CLD is similar to R2, and includes the following variables in sequence: LLM Safety Level, content matching, clinical adaptability, number of users, LLM Industry Capacity, and social and economic benefits. This loop reflects the improvement of the safety level of the LLM, which can improve clinical adaptability by strengthening content matching, thereby attracting more users, obtaining greater economic and social benefits, expanding industry capabilities, and further improving the safety level of the LLM, attracting more stakeholders to participate, clarifying the division of responsibilities, and ultimately further improving policy acceptance, forming a positive feedback loop.

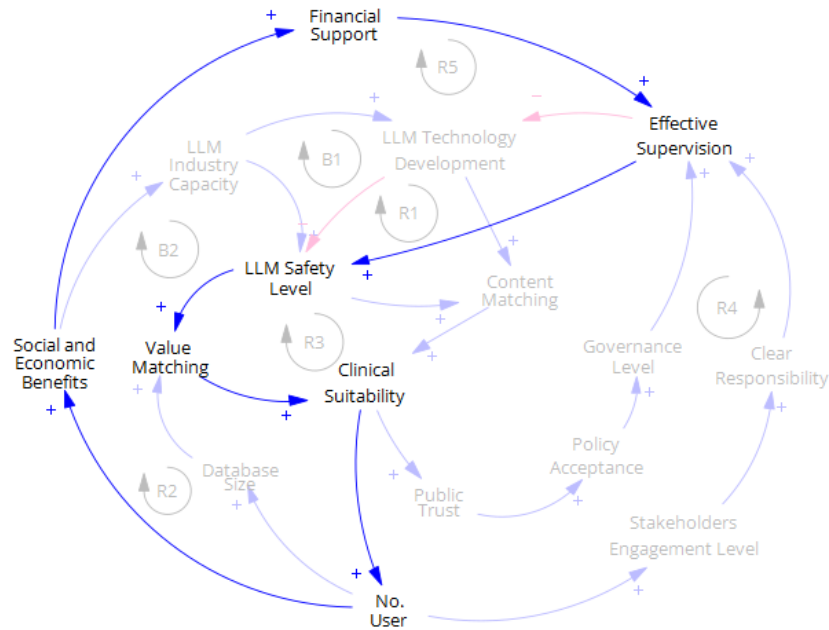


Figure 6.9 R5: Financial support can promote effective supervision and thus improve the safety level of LLM

The causal loop R5, figure 6.9, in the CLD contains the following variables: number of users, socioeconomic benefits, financial support, clinical suitability, value matching, LLM safety level, and effective supervision. This loop reflects that effective supervision can improve the safety level of large models, thereby strengthening value matching, strengthening clinical suitability, and increasing the number of users. More users can obtain greater social and economic benefits, thereby increasing financial support, and ultimately obtaining more effective supervision, forming a positive feedback loop.

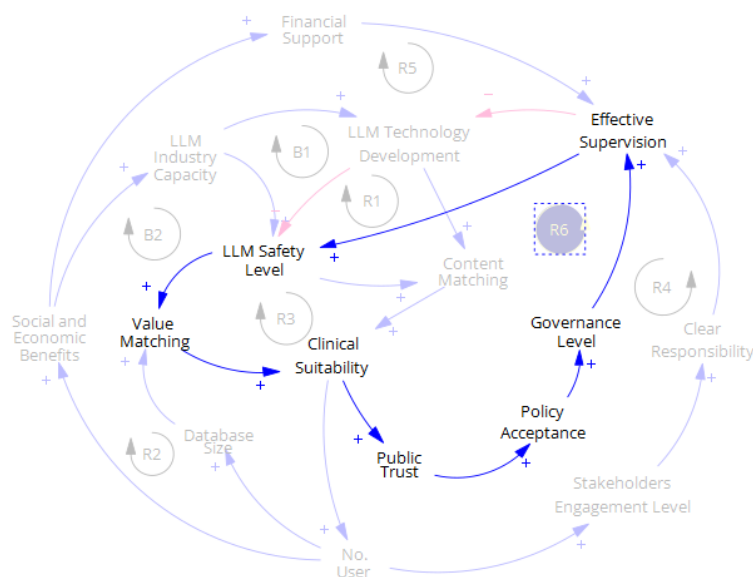


Figure 6.10 R6: Effective monitoring improves public trust and forms a positive feedback loop

The causal loop R6, figure 6.10, in the CLD includes the following variables: effective governance, LLM safety level, clinical suitability, value matching, LLM safety level, public trust, policy acceptance, and political governance level. This loop reflects that effective supervision can improve the safety level of LLMs, thereby strengthening value matching, strengthening clinical suitability, increasing public trust, and improving the acceptance of LLM-related policies, thereby promoting the government's governance level, and ultimately obtaining more effective supervision, forming a positive feedback loop.

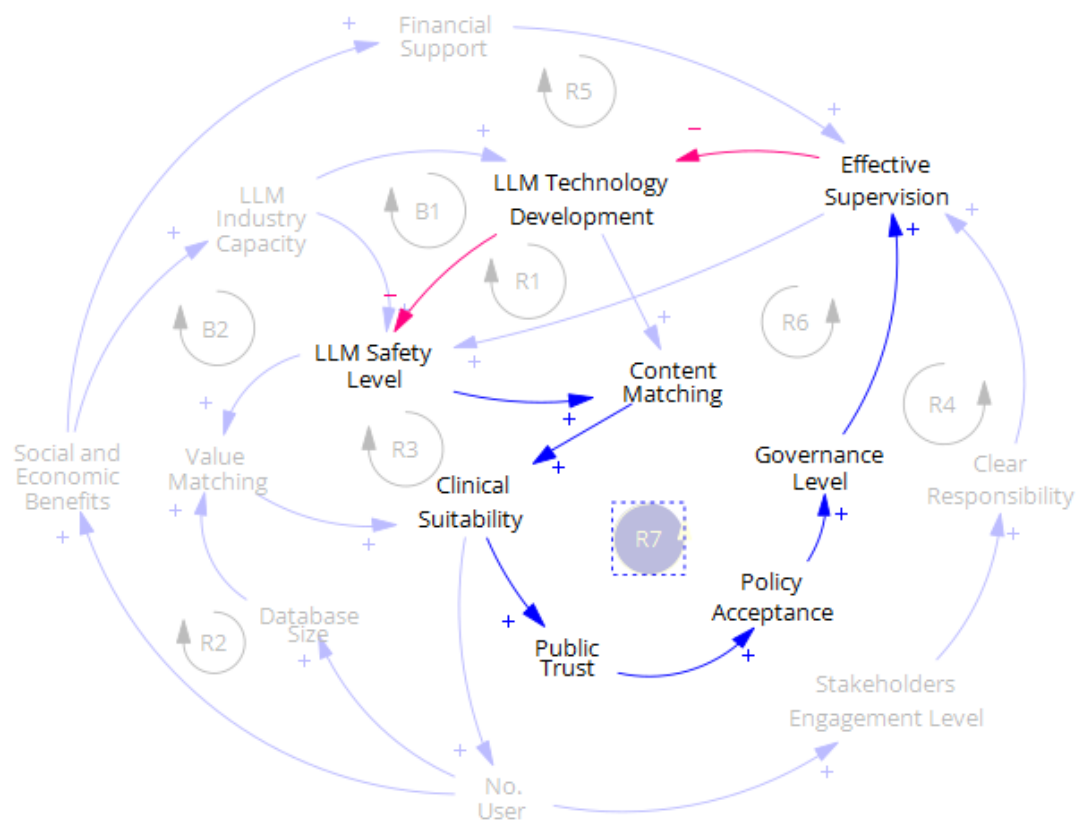


Figure 6.11 R7: Effective supervision can balance LLM development and safety

The causal loop R7, figure 6.11, in the CLD includes the following variables: effective governance, LLM development level, LLM safety level, clinical suitability, content matching, public trust, policy acceptance, and political management level. This loop reflects that effective supervision will limit the development of LLMs to a certain extent, but it can ensure the improvement of LLM safety level, thereby strengthening value matching, strengthening clinical suitability, increasing public trust, and improving the acceptance of LLM-related policies,

thereby promoting the government's governance level, and ultimately obtaining more effective supervision, forming a positive feedback loop.

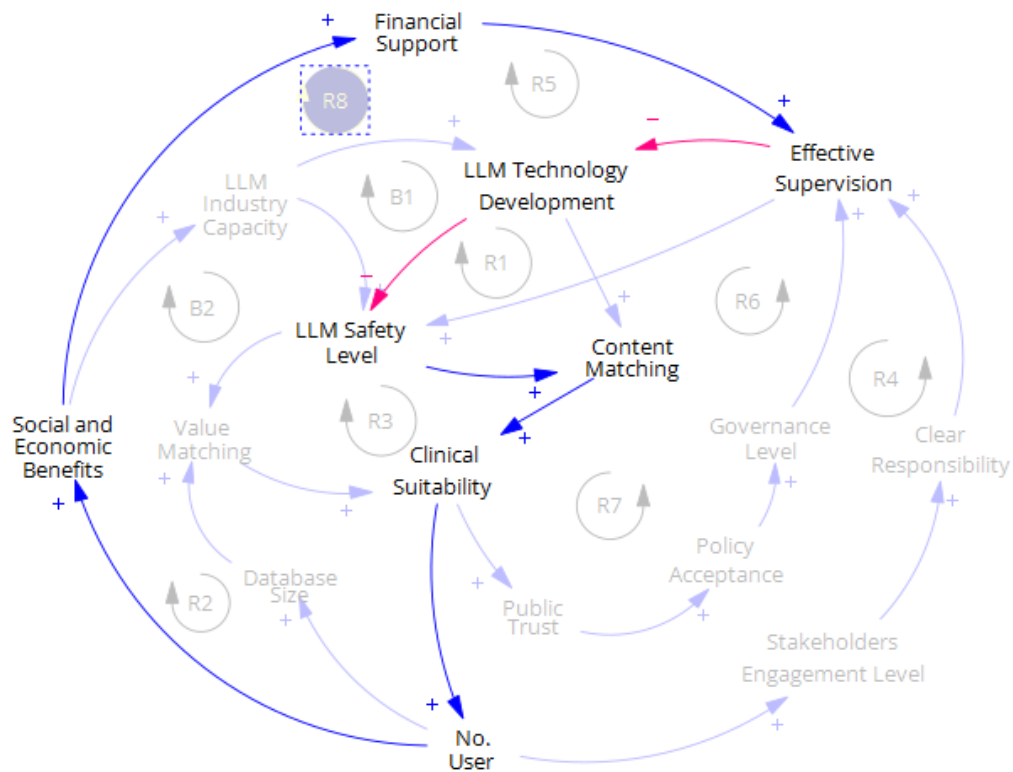


Figure 6.12 R8: Effective supervision can balance LLM development and safety, thereby achieving economic and social benefits

The causal loop R8, figure 6.12, in the CLD contains the following variables: number of users, socioeconomic benefits, financial support, clinical applicability, content matching, LLM safety level, LLM development level, and effective supervision. This loop reflects that effective supervision has inhibited the development level of large models to a certain extent, but has improved the safety level of large models, thereby strengthening value matching, strengthening clinical applicability, and increasing the number of users. More users can obtain greater socioeconomic benefits, thereby increasing financial support, and ultimately obtaining more effective supervision, forming a positive feedback loop.

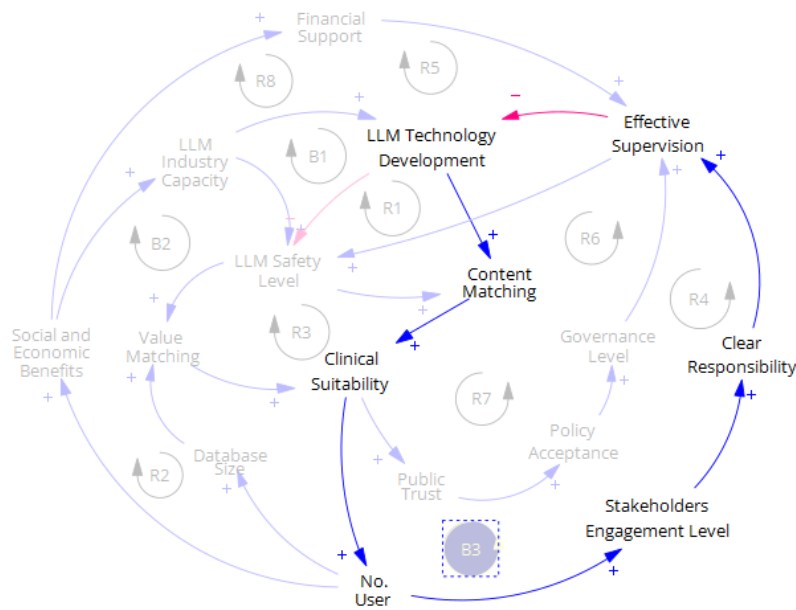


Figure 6.13 B3: Inappropriate regulation will inhibit the development of LLM

The causal loop B3, figure 6.13, in the CLD includes the following variables: number of users, stakeholder participation level, responsibility clarity level, clinical applicability, content matching, LLM development level, and effective supervision. This loop reflects that effective supervision has inhibited the development level of large models to a certain extent. The reduction in the development level will reduce the content matching, thereby reducing clinical applicability and reducing the number of users. Fewer users also means less stakeholder participation, which makes it impossible to clarify responsibilities, and ultimately reduces effective supervision, forming a negative feedback loop.

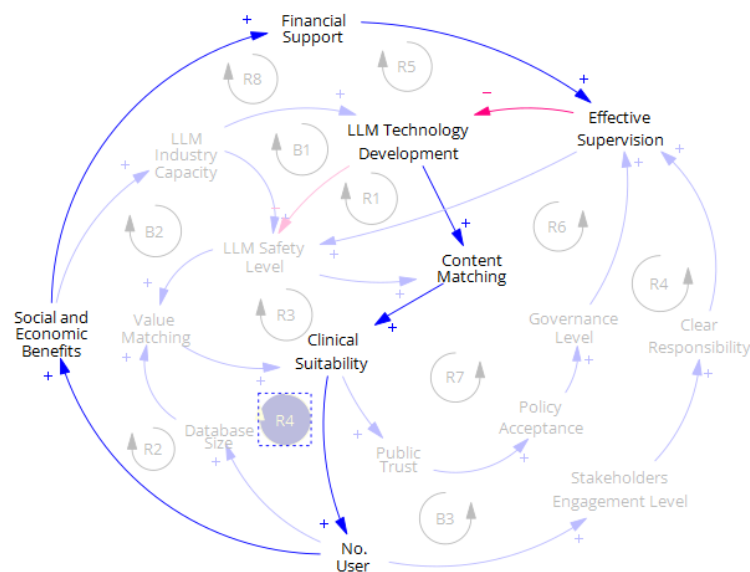


Figure 6.14 B1: Inappropriate regulation can inhibit LLM development and reduce social and economic benefits

The causal loop R4, figure 6.14, in the CLD contains the following variables: number of users, socioeconomic benefits, financial support, clinical applicability, content matching, LLM development level, and effective supervision. This loop reflects that effective supervision has inhibited the development level of large models to a certain extent, thereby reducing value matching, weakening clinical applicability, and reducing the number of users. Fewer users will also reduce the socioeconomic benefits that can be obtained, thereby reducing financial support, and ultimately limiting effective supervision, forming a negative feedback loop.

6.3.5 Stakeholder Analysis

Government departments play a key role in the governance of LLM medical applications. National and local government departments, such as the Ministry of Science and Technology, the National Health Commission, and local digital governance offices, are mainly responsible for strategic planning, cross-departmental coordination, resource allocation, and the construction of long-term governance frameworks. These departments ensure that the development of LLM technology is in line with the country's science and technology strategy and public interests by formulating policies and regulations. At the same time, government departments are also responsible for supervising and evaluating the implementation of the governance framework and adjusting policy directions in a timely manner to meet new challenges brought about by technological development. As the ultimate guarantor of public interests and social values, government departments ensure that the application of LLM technology in the medical field always takes improving people's health and well-being as the primary goal.

Technology developers are the promoters and innovators of LLM medical applications, and undertake core tasks such as training LLMs, medical scenario adaptation, and hospital system integration. These tasks involve optimizing and upgrading models, processing and analyzing medical data, and testing the performance of models in different medical environments. Technology developers not only need to have strong technical capabilities, but also need to work closely with medical institutions, regulatory authorities, etc. to ensure the accuracy and reliability of the models. They develop LLM applications that meet the needs of the medical industry and promote the deployment and use of these applications in actual medical scenarios, thereby improving the efficiency and quality of medical services.

Hospitals and primary medical institutions at all levels, as direct providers of medical services, are the front-line users of LLM medical applications. They are responsible for integrating LLM technology into daily medical work, such as auxiliary diagnosis, treatment plan recommendation, patient consultation, etc. In the actual application process, medical service institutions can provide the most direct feedback to help discover and solve problems with the model in medical scenarios. At the same time, these institutions also need to pay attention to the risks that may be brought about by LLM applications, such as data privacy leakage, diagnostic errors, etc., and report them to regulatory authorities and technology developers in a timely manner. In addition, medical institutions also need to provide relevant training for medical staff to improve their understanding and application capabilities of LLM technology and ensure the correct use and effectiveness of the model.

Industry regulatory authorities, such as national and local health and health commissions, are important leading forces in the governance of LLM medical applications. Their responsibilities include formulating medical AI-related policies, approving the clinical application of LLM models, supervising data use, and evaluating algorithm security. Regulatory authorities establish strict approval processes and regulatory standards to ensure that only LLM models that meet safety and effectiveness requirements can enter the medical market. At the same time, they also need to continuously supervise the models that have been put into use, deal with problems and risks in a timely manner, and maintain the order of the medical market and the rights and interests of patients.

In the LLM medical application governance system, user groups, including patients and their families, are direct beneficiaries and service objects, and their needs and feedback are crucial to the optimization of LLM applications. Through interaction with medical institutions, user needs are conveyed to technology developers, prompting continuous optimization of models and functional expansion. Users' acceptance and trust in LLM applications directly affect their frequency and effectiveness of use, and indirectly affect the promotion and popularization of technology. At the same time, users' active participation and feedback provide improvement directions for governance entities and promote the improvement and development of the entire governance system.

Table 6.7 Stakeholder Analysis Summary

Role	Objective	Key Resources/Power	Governance Contributions and Risks
-------------	------------------	----------------------------	---

Government Regulation	Safety and compliance, credibility	Laws, regulatory tools	Rule - setting, lack of dynamic feedback
Hospital Managers	Risk management, performance guarantee	Medical resources, technical decision - making power	Promoting deployment, centralized pressure
Doctors	Clinical efficiency, controllable responsibility	Professional judgment power	Key users, potential trust deficits
Patients	Safety and privacy protection	Public opinion, user feedback	Vulnerability to harm, impact on model trust
Technology Companies	Model promotion, market profits	Algorithmic and data capabilities	Technology - driven, potential disregard for ethical risks

6.4 Conclusion

6.4.1 The Verification and Extension of the DPG Theoretical Model by Case Study

In this section, we conduct an in-depth exploration of the DPG theoretical model through case study, aiming to verify its applicability and effectiveness in the context of large model governance in the medical industry. The CLD constructed based on the DPG structure serves as a crucial tool in this process.

The case study focuses on the dynamic relationships between key variables in the medical large - model governance system, as depicted in the CLD. For instance, the causal loops R1 and B1 clearly demonstrate the dual - edged nature of the development level of LLM. On one hand, R1 shows that the expansion of LLM industry capacity can drive technological development, enhance content matching and clinical suitability, attract more users, increase social and economic benefits, and further boost industry capacity, forming a positive feedback loop. On the other hand, B1 reveals that as the development level of LLM rises, more safety issues may emerge, reducing content matching, clinical suitability, user numbers, and social and economic benefits, thereby weakening LLM industry capacity and restraining large - model development. Similar dynamics are observed in other causal loops, which are validated through practical case analysis.

The case study not only confirms the existence of these feedback loops but also extends the DPG theoretical model. It highlights the complexity of the interplay between different

governance mechanisms, trust, performance perception, and usage behavior in the medical large - model governance scenario. For example, the causal loops R4, R5, R6, R7, and R8 illustrate how effective supervision can influence various aspects such as LLM safety level, value matching, clinical suitability, user numbers, and socioeconomic benefits, thereby forming multiple positive feedback loops to enhance governance effectiveness. Meanwhile, the negative feedback loops B3 and R4 (which forms a negative loop as per the previous description) reveal the potential risks of over - supervision, such as inhibiting LLM development level, reducing content matching and clinical applicability, and ultimately weakening effective supervision.

These findings enrich the DPG theoretical model by incorporating the specific characteristics and practical challenges of medical large - model governance, providing a more comprehensive and nuanced understanding of the governance process.

6.4.2 Policy-Making Suggestions and Contributions

Based on the CLD analysis and case study verification of the DPG theoretical model, this study provides valuable policy recommendations and theoretical contributions to improving the governance level of LLMs in the medical industry.

1. Balancing Development and Safety

CLD points out that the key links where the improvement of the development level of LLMs leads to a decline in safety must be closely monitored. Policymakers should strengthen safety-related regulations and standards in the development of medical LLMs. For example, a mandatory safety assessment mechanism can be established to ensure that the safety level of LLMs will not decrease with the improvement of the development level. In this process, the human values framework should be used as an alignment standard to avoid excessive governance. Specifically, policymakers need to seek a balance between promoting LLM technology innovation and maintaining their safety, ensuring that the development of LLMs is in line with human core values, such as respecting patient privacy and ensuring medical fairness. By incorporating the human values framework into the policy making of LLM governance, it is possible to effectively prevent potential safety risks without excessively restricting technological development and guide LLM technology to develop in a direction that is beneficial to human health.

2. Strengthen Effective Supervision

The positive feedback loops (R1-R8) in the CLD highlights the key role of effective

supervision in improving all aspects of LLM governance. The government should increase investment in regulatory technology and personnel training, build a dedicated medical LLM regulatory platform, and achieve real-time monitoring and management of LLM operations. This will help enhance the value matching, clinical applicability and user trust of LLMs, thereby promoting the healthy development of the entire industry. While strengthening supervision, it is also necessary to follow the human values framework to ensure the rationality and fairness of regulatory measures. Regulatory policies should avoid unnecessary obstacles to LLM technology innovation, and at the same time fully consider the demands of different stakeholders to protect the legitimate rights and interests of patients, medical institutions, developers and other parties. The regulatory model guided by human values can stimulate the innovative vitality of the market while maintaining the safety and reliability of LLMs and promote the sustainable development of the medical LLM industry.

3. Encourage Stakeholder Participation

CLD research shows that the participation of stakeholders such as users, developers and medical institutions significantly affects the effectiveness of LLM governance. Policies should be formulated to encourage stakeholders to actively participate in governance, such as establishing a reward mechanism for users who provide valuable feedback on LLMs and building a collaborative platform for developers and medical institutions to jointly improve the performance and safety of LLMs. Guided by the human values framework, encouraging stakeholders to participate in LLM governance will help to pool the wisdom and strength of all parties and form a good situation of multi-party co-governance. In the process of participation, all parties should uphold the common value pursuit of mankind, such as pursuing excellent medical service quality and protecting the welfare of patients. By establishing an effective communication mechanism and collaboration platform to promote exchanges and cooperation among different stakeholders, the enthusiasm of all parties can be fully mobilized to jointly overcome the difficulties in large model governance and achieve the deep integration of medical large model technology and human values.

This study applies the DPG structure to the medical large model governance context and conducts CLD analysis to expand the DPG theoretical model. The model identifies 8 reinforcement loops and 4 balancing loops, reveals the complex dynamic relationship between variables in the governance system, and provides new insights and empirical evidence for the DPG theory.

7 Discussion and Conclusion

7.1 Summary of Discussions and Findings

This study deeply explores the key issues of LLMs in governance and governance. First, starting from the development of technology, this study comprehensively analyzes the evolution of LLMs from early architecture to today's cutting-edge technologies, and elaborates on its wide application in education, law, finance, medical care and other fields. At the same time, it deeply analyzes governance risks such as data security, privacy protection and ethics. Through a systematic literature review, a solid knowledge system is built to understand the potential challenges of LLMs, providing an important reference for subsequent research and practice.

In terms of constructing a multidimensional framework of human values, this study conducts in-depth analysis from four dimensions: ideology, legal norms, cultural diversity and universal moral ethics, and discusses in detail the formation of values under different cultural and social backgrounds and their impact on individual and group decision-making processes, providing comprehensive theoretical support for evaluating the governance performance of LLMs.

In the experimental part, this study selected three representative LLMs, ChatGPT 4.0, ERNIE Bot 3.5 and Kimi, and conducted a detailed evaluation of their performance at different levels and dimensions. Through a carefully designed experimental scheme, not only was the model output qualitatively and quantitatively analyzed, but also special attention was paid to the ethical performance of the model when dealing with sensitive topics. The experimental results reveal the existing achievements and shortcomings of LLM in value alignment, providing empirical evidence for model improvement.

This study also proposed a conceptual framework based on DPG, aiming to improve the governance performance of LLM through collaborative governance. By dynamically monitoring and evaluating the performance of LLM, the DPM method can timely adjust and optimize the model's behavior to make it more in line with social values and ethical standards.

7.2 Contributions

Through a literature review, this study comprehensively analyzes the technological development, application fields, and governance risks of LLMs (LLMs). This section delves

into the development history of LLMs, from early model architecture to the most advanced technical implementations currently available, as well as their applications in various fields such as education, law, finance, etc. Additionally, the study focuses on the challenges LLMs face in terms of data security, privacy protection, and ethical considerations. By systematically organizing existing literature, this study provides comprehensive literature support for understanding the potential issues and challenges of LLMs, offering valuable references for subsequent research and practice.

A multidimensional framework of human values has been constructed, providing a solid theoretical foundation for understanding and assessing the alignment of LLMs with values. This framework considers multiple levels such as ideology, legal norms, cultural diversity, and universal ethical standards, analyzing how these values influence individual and collective decision-making processes in different cultural and social contexts. Through this framework, researchers can more comprehensively understand the complexity of LLMs in value alignment, evaluate their performance in different situations, and provide guidance for improving the design and application of LLMs. The establishment of this theoretical framework not only promotes an in-depth understanding of governance issues in LLMs but also provides a scientific basis for making decisions that are more in line with human values.

Through experimental methods, this study empirically evaluates the performance of popular LLMs in terms of value alignment. The study selected three representative LLMs—ChatGPT 4.0, ERNIE Bot 3.5, and Kimi—and assessed their performance across different value levels and dimensions through carefully designed experiments. The experiments included both qualitative and quantitative analyses of the model outputs, as well as assessments of the model's ethical performance when dealing with sensitive topics. These experimental results provide empirical data for the governance of LLMs, revealing the current achievements and shortcomings of LLMs in value alignment, and offering important references for future research and model improvements.

A conceptual framework based on DPG has been proposed. This framework emphasizes the importance of collaborative governance in the governance of LLMs and explains in detail how the DPG method supports this process. By dynamically monitoring and evaluating the performance of LLMs, the DPG method can adjust and optimize the model's behavior in a timely manner to better conform to societal values and ethical standards. The introduction of this management method provides new perspectives and tools for the governance of LLMs,

contributing to more effective value alignment.

A comprehensive research design has been proposed, emphasizing the importance of governance when applying LLMs in the medical field. This design not only considers the technical performance of LLMs but also pays special attention to the ethical and social issues they may face when applied in the medical field. The research design integrates knowledge and methods from various fields such as medicine, ethics, sociology, and computer science through interdisciplinary collaboration, ensuring that the application of LLMs in the medical field can better serve human health while respecting patient rights and societal ethical norms. This research provides comprehensive guidance for the application of LLMs in the medical field, ensuring the harmonious unity of technological development and societal values.

The contributions of this study are mainly reflected in the following aspects:

(1) Systematic analysis of LLM governance challenges: A comprehensive analysis of the governance challenges of LLM at the technical, application and social impact levels, providing a knowledge base for academia and practice to deeply understand the potential risks of LLM.

With the widespread application of LLMs in various fields, their governance issues have become increasingly complex and important. This study deeply analyzes the technical development of LLMs, from the burden of early model architectures to the widespread application of advanced models today and details the application examples of LLMs in key fields such as education, law, finance, and medical care and the governance challenges they bring. The study points out that although the rapid development of LLM technology has brought many conveniences and innovations, it is also accompanied by risks such as data security, privacy protection, and ethics. For example, in the field of education, LLMs can be used as auxiliary teaching tools to provide students with personalized learning plans, but they may also face the risk of generating false information or inducing academic misconduct. In the field of law, LLMs can assist in legal document analysis and case prediction, but how to ensure the accuracy and fairness of their output content remains an urgent problem to be solved. In the financial sector, LLM is increasingly used in risk assessment and customer service, but the risks of market manipulation and data leakage may bring cannot be ignored. In the medical field, LLM has great potential in auxiliary diagnosis and medical text analysis, but its accuracy and security issues are particularly prominent when processing sensitive medical data and providing medical advice.

(2) Constructing a human values evaluation framework: Creating a multidimensional human values framework to provide a scientific tool for accurately measuring and improving the governance performance of LLM.

To more accurately evaluate and improve the governance performance of LLM, this study constructed a multidimensional human values framework. The framework covers four key dimensions: ideology, legal norms, cultural diversity, and universal moral ethics. It aims to comprehensively analyze the formation of values in different cultural and social backgrounds and their impact on individual and group decision-making processes. In the ideological dimension, the study explored how different aspects such as political philosophy, economic model, social structure, religious beliefs and philosophical thoughts shape people's values; in the legal norms dimension, the study analyzed how the legal system such as the Constitution, the Civil and Commercial Law, the Criminal Law, the Procedural Law and the Administrative Regulations and Policies embodied the core values of society; in the cultural diversity dimension, the study deeply studied the values contained in cultural expressions such as language, traditional customs, art, literature and food; in the universal moral and ethical dimension, the study focused on the value goals pursued by all mankind, such as peace, development, fairness, justice, democracy and freedom. This framework provides a systematic and scientific theoretical tool for evaluating whether the output content of LLM conforms to human values.

(3) Empirical test of model value performance: Through experiments, the governance of mainstream LLM is analyzed to point out the direction for model optimization.

This study empirically evaluated the performance of three mainstream LLMs, ChatGPT 4.0, ERNIE Bot 3.5 and Kimi, at different value levels and dimensions through experimental methods. The experimental design is rigorous. The ValueEval task dataset of SemEval-2023 was selected to conduct a detailed analysis of the value choices of the models at four levels: ideology, legal norms, cultural diversity, and universal ethics. The results show that although these models perform relatively well in the field of universal ethics, there are significant deviations at the ideological and legal norms levels. For example, ChatGPT has a high frequency of deviations in the dimensions of political philosophy, economic model, and social structure, indicating that its value judgments in these areas are quite different from human values. This finding provides a clear direction for model developers to optimize the governance performance of LLM, suggesting that the diversity and quality of model training data need to

be further improved to improve its adaptability and accuracy in different cultural and social contexts.

(4) Propose an innovative governance conceptual framework: The proposed DPG framework provides new ideas for LLM governance and helps improve the governance effect.

The conceptual framework based on DPG proposed in this study provides an innovative management method for LLM governance. The framework emphasizes the integration of multiple forces such as government, enterprises, academic institutions, and the public through collaborative governance mechanisms to jointly participate in the governance process of LLM. The core of the DPG method is to dynamically monitor and evaluate the performance of LLM and timely adjust the model's behavior and decision-making strategy based on real-time feedback. For example, the government can set a clear ethical and safety bottom line for the development of LLM by formulating relevant laws, regulations and policy standards; enterprises, as developers and operators of LLM, need to actively fulfill their social responsibilities and ensure that the model output meets social values and ethical standards through technical means such as reinforcement learning and algorithm optimization; academic institutions play their scientific research advantages to provide theoretical support and technical reserves for the governance of LLM; public participation provides rich feedback data for the development of LLM, which helps the model better understand and adapt to the diverse needs of human society. Through this multi-party collaborative dynamic governance model, the governance effect of LLM can be effectively improved to ensure that it always follows the core values of mankind in a complex and changing social environment.

7.3 Limitations and Future Research

7.3.1 Limitations

This study has some limitations. For example, the limited size of the experimental data may lead to insufficient generalizability of the results; the data coverage is still limited, and the value dimensions of specific cultures, social groups or niche issues may be omitted; LLM may have inherent limitations in understanding complex value and ethical issues, and the experimental design may also limit the performance of the model in a more natural dialogue environment; the interpretability of the experimental results needs to be improved.

Although this study has made important progress in the field of LLM governance and governance, there are still some limitations. First, the scale of the experimental data is relatively

limited, and only 5,383 conclusion data are analyzed. This may lead to the lack of representation of all types of value judgments in the research results, limiting its generalizability. Future research needs to expand the data scale to more comprehensively cover value judgments in different contexts. Second, the data coverage is still limited. Although the data sources are diverse, it may not fully cover the important value dimensions in specific cultures, social groups or niche issues. This may lead to poor performance of the model in dealing with these specific issues, affecting its applicability in a multicultural context. In addition, LLM may have inherent limitations in understanding complex values and ethical issues, especially when these issues require in-depth understanding of cultural background. The current experimental design may also limit the performance of the model in a more natural conversation environment. Future research needs to design more flexible experimental schemes to simulate real-world conversation scenarios. Finally, the interpretability of the experimental results needs to be further improved. Although the study provides frequency data on the deviation of model values from human values, there is a lack of in-depth analysis of the logic and decision-making process behind these data.

7.3.2 Future Research

Future research needs to strengthen the explanation of the model's decision-making process to reveal how the model forms value judgments and how they are affected by input data, model architecture, and training process. Expand the data scale and diversity to improve the model's adaptability to different cultural and social backgrounds; expand the research to different dialogue scenarios to optimize the model's understanding and processing of value judgments; combine DPM strategies with collaborative governance results to explore ways to improve the effectiveness of LLM value alignment; deeply analyze experimental results to enhance the understanding of the model's value judgment formation process. In view of the limitations of this study, future research should focus on the following directions. First, by expanding the data scale and diversity, improve the model's adaptability to different cultural and social backgrounds. This includes further expanding and diversifying the training data set to ensure that the model can access a wider range of value concepts and cultural backgrounds. At the same time, researchers need to develop more complex model architectures to enhance the model's ability to understand complex value and ethical issues. Second, expand the research to different dialogue scenarios to optimize the model's understanding and processing of value judgments. This involves designing new experiments to test the model's performance in handling different types of dialogues and value judgments and developing new methods to

evaluate and improve the model's governance performance. In addition, the DPM strategy is combined with collaborative governance results to explore effective methods to improve the effectiveness of LLM value alignment. Future research can further explore how to promote cooperation between governments, enterprises and academic institutions through policy incentives and market mechanisms to jointly promote LLM governance research and practice. Finally, the experimental results are analyzed in depth to enhance the understanding of the model value judgment formation process. Through careful analysis of experimental data and the in-depth exploration of the model decision-making process, the inherent logic and influencing factors of the model in dealing with value judgments are revealed, providing more targeted guidance for the improvement and optimization of the model.

In summary, this study has taken an important step in the field of LLM governance and laid the foundation for achieving the coordinated unity of sustainable development of LLM technology and social value. Future research will continue to deepen and improve LLM governance theory and practice and promote the responsible application of LLM technology in a wider range of fields.

Reference

- Abbasian, M., Azimi, I., Rahmani, A. M., & Jain, R. (2023). Conversational health agents: A personalized llm-powered agent framework. *arXiv preprint arXiv:2310.02374*.
- Abbasiantaeb, Z., Yuan, Y., Kanoulas, E., Aliannejadi, M., & Assoc computing, m. (2024, 2024
- Mar 04-08). Let the LLMs Talk: Simulating Human-to-Human Conversational QA via Zero-Shot LLM-to-LLM Interactions. [Proceedings of the 17th acm international conference on web search and data mining, wsdm 2024]. 17th ACM International Conference on Web Search and Data Mining (WSDM), Merida, MEXICO.
- Abi-Rafteh, J., Hanna, S., Bassiri-Tehrani, B., Kazan, R., & Nahai, F. (2023). Complications Following Facelift and Neck Lift: Implementation and Assessment of Large Language Model and Artificial Intelligence (ChatGPT) Performance Across 16 Simulated Patient Presentations [; Early Access]. *Aesthetic Plastic Surgery*. <https://doi.org/10.1007/s00266-023-03538-1>
- Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., Almeida, D., Altschmidt, J., Altman, S., & Anadkat, S. (2023). Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*.
- Addiss, S., Groemer, G., & Rimer, J. T. (2006). *Traditional Japanese arts and culture: an illustrated sourcebook*. University of Hawaii Press.
- Adeoye-Olatunde, O. A., & Olenik, N. L. (2021). Research and scholarly methods: Semi-structured interviews. *JACCP: JOURNAL OF THE AMERICAN COLLEGE OF CLINICAL PHARMACY*, 4(10), 1358–1367. <https://doi.org/10.1002/jac5.1441>
- Ahmed, T., Devanbu, P., & Assoc Computing, M. (2022, 2022)
- Oct 10-14). Few-shot training LLMs for project-specific code-summarization. *IEEE ACM International Conference on Automated Software Engineering* [Proceedings of the 37th ieee/acm international conference on automated software engineering, ase 2022]. 37th IEEE/ACM International Conference on Automated Software Engineering (ASE), Oakland Univ, MI.
- Aho, K. (2014). *Existentialism: an introduction*. John Wiley & Sons.
- Albrow, M. (1970). *Bureaucracy*. Bloomsbury Publishing.
- Alexander, N. (2003). *The African Renaissance and the use of African languages in tertiary education*. Citeseer.
- Allport, G. W. (1960). A study of Values. In: Houghton Mifflin.
- Almeida, G. F. C. F., Nunes, J. L., Engelmann, N., Wiegmann, A., & de Araujo, M. (2024). Exploring the psychology of LLMs' moral and legal reasoning. *Artificial Intelligence*, 333, Article 104145. <https://doi.org/10.1016/j.artint.2024.104145>
- Alqahtani, T., Badreldin, H. A., Alrashed, M., Alshaya, A. I., Alghamdi, S. S., bin Saleh, K., Alowais, S. A., Alshaya, O. A., Rahman, I., Al Yami, M. S., & Albekairy, A. M. (2023). The emergent role of artificial intelligence, natural learning processing, and large language models in higher education and research. *Research in Social & Administrative Pharmacy*, 19(8), 1236–1242.

<https://doi.org/10.1016/j.sapharm.2023.05.016>

- Ambos, K., Duff, A., Roberts, J., Weigend, T., & Heinze, A. (2020). *Core Concepts in Criminal Law and Criminal Justice: Anglo-German Dialogues* (Vol. 1). Cambridge University Press.
- Amuta, C. (2017). *Theory of African literature: Implications for practical criticism*. Zed Books Ltd.
- Andrew Caines¹, L. B., Shiva Taslimipoor¹, Christopher Davis¹, Yuan Gao¹, Øistein Andersen¹, Zheng Yuan^{2,1}, Mark Elliott^{3,1}, Russell Moore¹, Christopher Bryant^{4,1}, Marek Rei^{5,1}, Helen Yannakoudakis^{2,1}, Andrew Mullooly³, Diane Nicholls⁶ and Paula Buttery¹. (2023). On the application of Large Language Models for language teaching and assessment technology. *arXiv preprint arXiv:2307.08393*.
- Antoniou-Kritikou, I., Giouli, V., Tsoulouhas, G., & Economou, C. (2024). Using LLMs in a Language Teaching and Learning Application. *Ercim News*(136), 17-18. [<Go to ISI>://WOS:001199179500009](#)
- Arblaster, A. (2002). *Democracy*. McGraw-Hill Education (UK).
- Arighi, C., Brenner, S., & Lu, Z. (2024). Large Language Models (LLMs) and ChatGPT for Biomedicine. *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*, 29, 641-644. [<Go to ISI>://MEDLINE:38160312](#)
- Arneson, R. (2002). Egalitarianism.
- Arneson, R. J. (1999). Egalitarianism and responsibility. *The Journal of Ethics*, 3, 225-247.
- Arts, W., & Gelissen, J. (2001). Welfare states, solidarity and justice principles: does the type really matter? *Acta sociologica*, 44(4), 283-299.
- Ashkenazi, M., & Jacob, J. (2003). *Food culture in Japan*. Bloomsbury Publishing USA.
- Askeel, A., Bai, Y., Chen, A., Drain, D., Ganguli, D., Henighan, T., Jones, A., Joseph, N., Mann, B., & DasSarma, N. (2021). A general language assistant as a laboratory for alignment. *arXiv preprint arXiv:2112.00861*.
- Aston, W. G. (1899). *A history of Japanese literature* (Vol. 6). D. Appleton.
- Ayyar, P. J. (2001). *South Indian Customs*. Asian educational services.
- Baeck, L. (2019). *The essence of Judaism*. Plunkett Lake Press.
- Baldwin, P. (1990). *The politics of social solidarity: class bases of the European welfare state, 1875-1975*. Cambridge University Press.
- Bao, K., Zhang, J., Zhang, Y., Wang, W., Feng, F., He, X., & Acm. (2023, 2023 Sep 18-22). TALLRec: An Effective and Efficient Tuning Framework to Align Large Language Model with Recommendation. [Proceedings of the 17th acm conference on recommender systems, recsys 2023]. 17th ACM Conference on Recommender Systems (RecSys), Singapore, SINGAPORE.
- Bauwens, R., Audenaert, M., & Decramer, A. (2024). Performance management systems, innovative work behavior and the role of transformational leadership: an experimental approach. *Journal of Organizational Effectiveness: People and Performance*, 11(1), 178-195.
- Baxi, U. (1982). *The crisis of the Indian legal system. Alternatives in development: Law*. Stranger Journalism.
- Beetham, D. (1996). *Bureaucracy*. U of Minnesota Press.

- Bell, J. (1995). *French constitutional law*. Oxford University Press.
- Benary, M., Wang, X. D., Schmidt, M., Soll, D., Hilfenhaus, G., Nassir, M., Sigler, C., Knödler, M., Keller, U., & Beule, D. (2023). Leveraging large language models for decision support in personalized oncology. *Jama Network Open*, 6(11), e2343689-e2343689.
- Bergson, A. (1967). Market socialism revisited. *Journal of Political Economy*, 75(5), 655-673.
- Berkman, A. (2003). *What is anarchism?* (Vol. 1). AK Press.
- Bernitz, U. (2007). What is Scandinavian law? Concept, characteristics, future. *Scandinavian studies in law*, 50(1), 13-30.
- Bernstein, I. A., Zhang, Y., Govil, D., Majid, I., Chang, R. T., Sun, Y., Shue, A., Chou, J. C., Schehlein, E., Christopher, K. L., Groth, S. L., Ludwig, C., & Wang, S. Y. (2023). Comparison of Ophthalmologist and Large Language Model Chatbot Responses to Online Patient Eye Care Questions. *Jama Network Open*, 6(8), Article e2330320. <https://doi.org/10.1001/jamanetworkopen.2023.30320>
- Bettelheim, C. (1973). State property and socialism. *Economy and Society*, 2(4), 395-420.
- Bharadwaj, R., & Parker, I. (2023, 2023)
- Bhimani, M., Miller, A., Agnew, J. D., Ausin, M. S., Raglow-Defranco, M., Mangat, H., Voisard, M., Taylor, M., Bierman-Lytle, S., & Parikh, V. (2025). Real-World Evaluation of Large Language Models in Healthcare (RWE-LLM): A New Realm of AI Safety & Validation. *medRxiv*, 2025-03.
- May 01-03). Double-Edged Sword of LLMs: Mitigating Security Risks of AI-Generated Code. *Proceedings of SPIE* [Disruptive technologies in information sciences vii]. Conference on Disruptive Technologies in Information Sciences VII, Orlando, FL.
- Bianchi, C. (2012). Enhancing performance management and sustainable organizational growth through system-dynamics modelling. *Systemic management for intelligent organizations: Concepts, models-based approaches and applications*, 143-161.
- Bianchi, C. (2016). *Dynamic performance management* (Vol. 1). Springer.
- Bianchi, C., Nasi, G., & Rivenbark, W. C. (2021). Implementing collaborative governance: Models, experiences, and challenges. *Public Management Review*, 23(11), 1581-1589. <https://doi.org/10.1080/14719037.2021.1878777>
- Bianchi, C. (2022). Enhancing policy design and sustainable community outcomes through collaborative platforms based on a dynamic performance management and governance approach. In *Research handbook of policy design* (pp. 407-429). Edward Elgar Publishing. <https://www.elgaronline.com/edcollchap/edcoll/9781839106590/9781839106590.00036.xml>
- Biancone, P., Brescia, V., & Oppioli, M. (2022). COLLABORATIVE GOVERNANCE AND TECHNOLOGIES: A BIBLIOMETRIC ANALYSIS. *International Public Management Review*, 23(1), 46-76.
- Bill, D., & Eriksson, T. (2023). Fine-tuning a llm using reinforcement learning from human

- feedback for a therapy chatbot application. In.
- Bilsky, W., & Schwartz, S. H. (1994). Values and personality. *European journal of personality*, 8(3), 163-181.
- Blyden, E. W. (1994). *African life and customs*. Black Classic Press.
- Bogomolov, O. (1986). Coordination of Economic Interests and Politics under Socialism. *Soviet Review*, 27(4), 3-20.
- Bohlander, M. (2008). *Principles of German criminal law*. Bloomsbury Publishing.
- Bond, M. H. (1988). Finding universal dimensions of individual variation in multicultural studies of values: The Rokeach and Chinese value surveys. *Journal of personality and social psychology*, 55(6), 1009.
- Bosteels, B. (2014). *The actuality of communism*. Verso Books.
- Boyron, S. (2012). *The constitution of France: A contextual analysis*. Bloomsbury Publishing.
- Braithwaite, V. A., & Law, H. (1985). Structure of human values: Testing the adequacy of the Rokeach Value Survey. *Journal of personality and social psychology*, 49(1), 250.
- Brameier, D. T., Alnasser, A. A., Carnino, J. M., Bhashyam, A. R., von Keudell, A. G., & Weaver, M. J. (2023). Artificial Intelligence in Orthopaedic Surgery: Can a Large Language Model "Write" a Believable Orthopaedic Journal Article? *The Journal of bone and joint surgery. American volume*, 105(17), 1388-1392. <https://doi.org/10.2106/jbjs.23.00473>
- Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. *Acta psychologica*, 81(3), 211-241.
- Bremmer, I. (2008). The return of state capitalism. *Survival*, 50(3), 55-64.
- Brendel, A. B., Mirbabaie, M., Lembcke, T.-B., & Hofeditz, L. (2021). Ethical management of artificial intelligence. *Sustainability*, 13(4), 1974.
- Brenner, R. (1990). Feudalism. In *Marxian economics* (pp. 170-185). Springer.
- Breyer, S. G., Stewart, R. B., Sunstein, C. R., Vermeule, A., & Herz, M. (2022). *Administrative Law and Regulatory Policy: Problems, Text, and Cases [Connected eBook with Study Center]*. Aspen Publishing.
- Broeder, P. (2021). Informed communication in high context and low context cultures. *Journal of Education, Innovation, and Communication*, 3(1), 13-24.
- Brosch, T., & Sander, D. (2015). *Handbook of value: perspectives from economics, neuroscience, philosophy, psychology and sociology*. Oxford University Press.
- Brown, L. N., Bell, J., & Galabert, J.-M. (1998). *French administrative law*. Oxford University Press, USA.
- Brus, W. (1990). Market socialism. In *Problems of the planned economy* (pp. 164-177). Springer.
- Bui, M.-Q., Do, D.-T., Le, N.-K., Nguyen, D.-H., Nguyen, K.-V.-H., Anh, T. P. N., & Nguyen, M. L. (2024). Data Augmentation and Large Language Model for Legal Case Retrieval and Entailment. *Review of Socionetwork Strategies*, 18(1), 49-74. <https://doi.org/10.1007/s12626-024-00158-2>
- Burns, C., Izmailov, P., Kirchner, J. H., Baker, B., Gao, L., Aschenbrenner, L., Chen, Y., Ecoffet,

- A., Joglekar, M., & Leike, J. (2023). Weak-to-strong generalization: Eliciting strong capabilities with weak supervision. *arXiv preprint arXiv:2312.09390*.
- Burns, J. G. (2014). *Introduction to Islamic Law: Principles of Civil, Criminal, and International Law under the Shari 'a*. TellerBooks.
- Caballero, C. (2004). *Fauré and French musical aesthetics* (Vol. 13). Cambridge University Press.
- Cafiero, C. (2005). Anarchy and Communism. *Anarchism: A Documentary History of Libertarian Ideas, Volume One (Montreal: Black Rose Books, 2005)*, 110.
- Cai, T., Song, X., Jiang, J., Teng, F., Gu, J., & Zhang, G. (2023). ULMA: Unified Language Model Alignment with Demonstration and Point-wise Human Preference. *arXiv preprint arXiv:2312.02554*.
- Çam, T., & Kayaoğlu, M. (2015). Marx's distinction between socialism and communism. *Journal of Human Sciences*, 12(1), 385-391.
- Caws, P. (1968). What is structuralism?
- Chan, S. (2002). *Liberalism, democracy and development*. Cambridge University Press.
- Chang, Y., Wang, X., Wang, J., Wu, Y., Yang, L., Zhu, K., Chen, H., Yi, X., Wang, C., & Wang, Y. (2023). A survey on evaluation of large language models. *ACM Transactions on Intelligent Systems and Technology*.
- Chee-Beng, T. (2012). *Chinese food and foodways in Southeast Asia and beyond*. NUS Press.
- Chen, J., Cadiente, A., Kasselmann, L. J., & Pilkington, B. (2024). Assessing the performance of ChatGPT in bioethics: a large language model's moral compass in medicine. *Journal of Medical Ethics*, 50(2), 97-101. <https://doi.org/10.1136/jme-2023-109366>
- Chen, Q., Zhang, T., Nie, M., Wang, Z., Xu, S., Shi, W., Cao, Z., & Acm. (2023, 2023 Nov 02). Fashion-GPT: Integrating LLMs with Fashion Retrieval System. [Proceedings of the 1st workshop on large generative models meet multimodal applications, lgm3a 2023]. 1st Workshop on Large Generative Models Meet Multimodal Applications (LGM3A), Ottawa, CANADA.
- Chen, T. C., Kaminski, E., Koduri, L., Singer, A., Singer, J., Couldwell, M., Delashaw, J., Dumont, A., & Wang, A. (2023). Chat GPT as a Neuro-Score Calculator: Analysis of a Large Language Model's Performance on Various Neurological Exam Grading Scales. *World Neurosurgery*, 179, E342-E347. <https://doi.org/10.1016/j.wneu.2023.08.088>
- Chen, Y., Arunasalam, A., Celik, Z. B., & Acm. (2023, 2023 Dec 04-08). Can Large Language Models Provide Security & Privacy Advice? Measuring the Ability of LLMs to Refute Misconceptions. [39th annual computer security applications conference, acsac 2023]. 39th Annual Computer Security Applications Conference (ACSAC), Austin, TX.
- Chervenak, J., Lieman, H., Blanco-Breindel, M., & Jindal, S. (2023). The promise and peril of using a large language model to obtain clinical information: ChatGPT performs strongly as a fertility counseling tool with limitations. *Fertility and Sterility*, 120(3),

- 575-583. <https://doi.org/10.1016/j.fertnstert.2023.05.151>
- Chiesa-Estomba, C. M., Lechien, J. R., Vaira, L. A., Brunet, A., Cammaroto, G., Mayo-Yanez, M., Sanchez-Barrueco, A., & Saga-Gutierrez, C. (2024). Exploring the potential of Chat-GPT as a supportive tool for sialendoscopy clinical decision making and patient information support. *European Archives of Oto-Rhino-Laryngology*, 281(4), 2081-2086.
- Chiu, H. (1985). The 1982 Chinese constitution and the rule of law. *Rev. Socialist L.*, 11, 143.
- Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology*, 13(1), 3-21.
<https://doi.org/10.1007/BF00988593>
- Chou, R. C.-A. (1993). Norms and their popularization: Respecting the elderly in traditional Chinese Society. In *Norms and the State in China* (pp. 185-206). Brill.
- Choudhry, S., Khosla, M., & Mehta, P. B. (2016). *The Oxford handbook of the Indian constitution*. Oxford University Press.
- Chun, J., & Elkins, K. (2024). Informed AI Regulation: Comparing the Ethical Frameworks of Leading LLM Chatbots Using an Ethics-Based Audit to Assess Moral Reasoning and Normative Values. *arXiv preprint arXiv:2402.01651*.
- Clark, P. A. (2015). *Anglo-American Innovation* (Vol. 9). Walter de Gruyter GmbH & Co KG.
- Clement, W. H. P. (1904). *The Law of the Canadian Constitution*. Carswell Company, Limited.
- Conze, E. (2012). *Buddhism: Its essence and development*. Courier Corporation.
- Coomaraswamy, A. K. (1923). *Introduction to Indian art*.
- Cothorn, C. R. (2019). *Handbook for environmental risk decision making: Values, perceptions, and ethics*. CRC Press.
- Creel, H. G. (1982). *What is Taoism?: and other studies in Chinese cultural history*. University of Chicago Press.
- Crielaard, L., Uleman, J. F., Châtel, B. D., Epskamp, S., Sloot, P., & Quax, R. (2024). Refining the causal loop diagram: A tutorial for maximizing the contribution of domain expertise in computational system dynamics modeling. *Psychological Methods*, 29(1), 169.
- Cuddihy, G., & King, A. (2021). *Russia-Culture Smart!: The Essential Guide to Customs & Culture*. Kuperard.
- Cui, J., Li, Z., Yan, Y., Chen, B., & Yuan, L. (2023). Chatlaw: Open-source legal large language model with integrated external knowledge bases. *arXiv preprint arXiv:2306.16092*.
- D'Antonoli, T. A., Stanzione, A., Bluethgen, C., Vernuccio, F., Ugga, L., Klontzas, M. E., Cuocolo, R., Cannella, R., & Kocak, B. (2024). Large language models in radiology: fundamentals, applications, ethical considerations, risks, and future directions [Review]. *Diagnostic and Interventional Radiology*, 30(2), 80-90.
<https://doi.org/10.4274/dir.2023.232417>
- D'Souza, C. (2020). *Civil Law Studies: An Indian Perspective*. Cambridge Scholars

- Publishing.
- Dahl, R. A. (2020). *On democracy*. Yale university press.
- Dai, W., Lin, J., Jin, H., Li, T., Tsai, Y.-S., Gašević, D., & Chen, G. (2023). Can large language models provide feedback to students? A case study on ChatGPT. 2023 IEEE International Conference on Advanced Learning Technologies (ICALT),
- Datta, A. (1988). *Encyclopaedia of Indian Literature: devraj to jyoti* (Vol. 2). Sahitya Akademi.
- Davis, K. E. (2001). The rules of capitalism. *Third World Quarterly*, 22(4), 675-682.
- De Angelis, L., Baglivo, F., Arzilli, G., Privitera, G. P., Ferragina, P., Tozzi, A. E., & Rizzo, C. (2023). ChatGPT and the rise of large language models: the new AI-driven infodemic threat in public health. *Frontiers in Public Health*, 11, 1166120.
- De Lange, N. (2000). *An introduction to Judaism*. Cambridge University Press.
- DeFrancis, J. (1986). *The Chinese language: Fact and fantasy*. University of Hawaii Press.
- Deiana, G., Dettori, M., Arghittu, A., Azara, A., Gabutti, G., & Castiglia, P. (2023). Artificial intelligence and public health: evaluating ChatGPT responses to vaccination myths and misconceptions. *Vaccines*, 11(7), 1217.
- Deng, X., Bashlovkina, V., Han, F., Baumgartner, S., Bendersky, M., & Acm. (2023a, 2023 Apr 30-May 04). LLMs to the Moon? Reddit Market Sentiment Analysis with Large Language Models. [Companion of the world wide web conference, www 2023]. 32nd World Wide Web Conference (WWW), Austin, TX.
- Deng, X., Bashlovkina, V., Han, F., Baumgartner, S., Bendersky, M., & Acm. (2023b, 2023 Apr 30-May 04). What do LLMs Know about Financial Markets? A Case Study on Reddit Market Sentiment Analysis. [Companion of the world wide web conference, www 2023]. 32nd World Wide Web Conference (WWW), Austin, TX.
- Deshpande, M. S. (2010). History of the Indian caste system and its impact on India today.
- Dixon, J., & Macarov, D. (2016). *Social welfare in socialist countries*. Routledge.
- Dodek, A. (2024). *The Canadian Constitution*. Dundurn.
- Doniger, W. (2014). *On hinduism*. Oxford University Press.
- Dozier, J. B., Husted, B. W., & McMahon, J. T. (1998). Need for approval in low-context and high-context cultures: A communications approach to cross-cultural ethics. *Teaching Business Ethics*, 2(2), 111-125.
- Draper, H. (1962). Marx and the Dictatorship of the Proletariat.
- Draper, H. (1987). *Dictatorship of Proletariat*. NYU Press.
- Du Gay, P. (2005). *The values of bureaucracy*. Oxford University Press.
- Du, S. Q., Tang, S. J., Wang, W. X., Li, X. M., & Guo, R. Z. (2023, 2023 Sep 02-07). TREE-GPT: MODULAR LARGE LANGUAGE MODEL EXPERT SYSTEM FOR FOREST REMOTE SENSING IMAGE UNDERSTANDING AND INTERACTIVE ANALYSIS. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* [Geospatial week 2023, vol. 48-1]. 5th International-Society-for-Photogrammetry-and-Remote-Sensing (ISPRS) Geospatial Week (GSW), Cairo, EGYPT.
- Du, Y., Zhao, S., Chen, Y., Bai, R., Liu, J., Wu, H., Wang, H., & Qin, B. (2023). The CALLA

- Dataset: Probing LLMs' Interactive Knowledge Acquisition from Chinese Medical Literature. *arXiv preprint arXiv:2309.04198*.
- Dugar, M., & Asesh, A. (2023, 2023
- Dec 14-15). Spatial Interpretation and LLMs. [2023 IEEE International Conference on Machine Learning and Applied Network Technologies, icmlant]. IEEE International Conference on Machine Learning and Applied Network Technologies (ICMLANT), Electr Network.
- Durbin, E. F. (1936). Economic calculus in a planned economy. *The Economic Journal*, 46(184), 676-690.
- Dye, T. R., & Zeigler, H. (1988). Socialism and equality in cross-national perspective. *PS: Political Science & Politics*, 21(1), 45-56.
- Dyzenhaus, D. (2007). The rule of law as the rule of liberal principle. In. Cambridge University Press.
- Eady, S. J. (1982). Differences in the F0 patterns of speech: Tone language versus stress language. *Language and speech*, 25(1), 29-42.
- Eagleton, T. (2017). *Materialism*. Yale University Press.
- Eatwell, J., Milgate, M., & Newman, P. (1990). *Problems of the planned economy*. Springer.
- ECHOLS III, J. M. (1976). *Politics, Policy and Equality Under Communism and Democracy*. University of Michigan.
- Eggmann, F., Weiger, R., Zitzmann, N. U., & Blatz, M. B. (2023). Implications of large language models such as ChatGPT for dental medicine [Review]. *Journal of Esthetic and Restorative Dentistry*, 35(7), 1098-1102. <https://doi.org/10.1111/jerd.13046>
- Eitel-Porter, R. (2021). Beyond the promise: implementing ethical AI. *AI and Ethics*, 1(1), 73-80.
- Elango, B., Paul, K., Kundu, S. K., & Paudel, S. K. (2010). Organizational ethics, individual ethics, and ethical intentions in international decision-making. *Journal of Business Ethics*, 97, 543-561.
- Emerson, K., Nabatchi, T., & Balogh, S. (2012). An integrative framework for collaborative governance. *Journal of public administration research and theory*, 22(1), 1-29.
- Eppler, M., Ganjavi, C., Ramacciotti, L. S., Piazza, P., Rodler, S., Checcucci, E., Rivas, J. G., Kowalewski, K. F., Belenchon, I. R., Puliatti, S., Taratkin, M., Veccia, A., Baekelandt, L., Teoh, J. Y. C., Somani, B. K., Wroclawski, M., Abreu, A., Porpiglia, F., Gill, I. S., . . . Cacciamani, G. E. (2024). Awareness and Use of ChatGPT and Large Language Models: A Prospective Cross-sectional Global Survey in Urology. *European Urology*, 85(2), 146-153. <https://doi.org/10.1016/j.eururo.2023.10.014>
- Esmailzadeh, P. (2024). Challenges and strategies for wide-scale artificial intelligence (AI) deployment in healthcare practices: A perspective for healthcare organizations. *Artificial Intelligence in Medicine*, 151, 102861.
- Farber, D. R. (2010). *The rise and fall of modern American conservatism: A short history*. Princeton University Press.
- Faure, B. (2011). *Unmasking Buddhism*. John Wiley & Sons.
- Feather, N. T. (1979). Value correlates of conservatism. *Journal of Personality and Social*

- Psychology*, 37(9), 1617.
- Feder, A., Wald, Y., Shi, C., Saria, S., & Blei, D. (2023, 2023 Dec 10-16). Data Augmentations for Improved (Large) Language Model Generalization. *Advances in Neural Information Processing Systems* [Advances in neural information processing systems 36 (neurips 2023)]. 37th Conference on Neural Information Processing Systems (NeurIPS), New Orleans, LA.
- Ferguson, P. P. (2006). *Accounting for taste: The triumph of French cuisine*. University of Chicago Press.
- Feyerabend, P. K. (1962). Explanation, reduction, and empiricism.
- Finch, A. (2010). *French literature: a cultural history* (Vol. 13). Polity.
- Flood, G. (1996). *An Introduction to Hinduism*. Cambridge University Press.
- Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., & Rossi, F. (2018). AI4People—an ethical framework for a good AI society: opportunities, risks, principles, and recommendations. *Minds and machines*, 28, 689-707.
- Flynn, K. (2016). *Food, culture, and survival in an African city*. Springer.
- Foster, J. (2021). *The case for idealism*. Routledge.
- Frances, C., Alstyne, M., Ashton, A., & Hochstettler, T. (1994). Using system dynamics technology to improve planning and budgeting for higher education: Results in Arizona and Houston, Texas. 1994 International System Dynamics Conference, Arizona and Houston, Texas.
- Fränken, J.-P., Zelikman, E., Rafailov, R., Gandhi, K., Gerstenberg, T., & Goodman, N. D. (2024). Self-supervised alignment with mutual information: Learning to follow principles without preference labels. *arXiv preprint arXiv:2404.14313*.
- Fritzsche, D. J. (1991). A model of decision-making incorporating ethical values. *Journal of Business Ethics*, 10, 841-852.
- Fuchen, H. (2013). *General Theory of Taoism*. Paths International Ltd.
- Gabriele, A., & Schettino, F. (2012). Market Socialism as a Distinct Socioeconomic Formation. *New Proposals: Journal of Marxism and Interdisciplinary Inquiry*, 5(2), 20-50.
- Galanter, M. (1968). The Displacement of Traditional Law in Modern India. *Journal of Social issues*, 24(4).
- Glaser, B., & Strauss, A. (1967). *Research*. Citeseer.
<https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=3f3ffcae173346dbe91b8adb21c1a0e4f59e9fea>
- Galston, W. (1986). Equality of opportunity and liberal theory. *Justice and equality here and now*, 89, 107.
- Ganshof, F. L. (1996). *Feudalism* (Vol. 34). University of Toronto Press.
- Garnett, A. C. (1941). Liberalism as a theory of human nature. *J. Soc. Phil. & Jurisprudence*, 7, 127.
- Georgiou, C. D. (2016). Unconditional Communist Equality among Individuals: Beyond the Marxist Equality Limited to the Abolition of Classes. *Critique*, 44(1-2), 129-160.
- Gilbert, P., & O'Neill, M. (2019). Socialism.
- Gill, J. K., Chetty, M., Lim, S., & Hallinan, J. (2024). Large language model based framework

- for automated extraction of genetic interactions from unstructured data. *Plos One*, 19(5), Article e0303231. <https://doi.org/10.1371/journal.pone.0303231>
- Giudice, O., Maggi, A., Nardelli, M., & Assoc Computing, M. (2023, 2023 Dec 15-17). Exploring Naive Approaches to Tell Apart LLMs Productions from Human-written Text. [Proceedings of 2023 7th international conference on natural language processing and information retrieval, nlpir 2023]. 7th International Conference on Natural Language Processing and Information Retrieval (NLPir), Seoul, SOUTH KOREA.
- Glants, M., & Toomre, J. (1997). *Food in Russian history and culture*. Indiana University Press.
- Glover, S. H., Bumpus, M. A., Logan, J. E., & Ciesla, J. R. (1997). Re-examining the influence of individual values on ethical decision making. From the Universities to the Marketplace: The Business Ethics Journey: The Second Annual International Vincentian Conference Promoting Business Ethics,
- Goetz, S., & Taliaferro, C. (2008). *Naturalism*. Wm. B. Eerdmans Publishing.
- Goodnow, F. J. (2003). *The principles of the administrative law of the United States*. The Lawbook Exchange, Ltd.
- Gorman, J. C., & Wiltshire, T. J. (2024). A typology for the application of team coordination dynamics across increasing levels of dynamic complexity. *Human factors*, 66(1), 5-16.
- Gorski, L., & Ramakrishna, S. (2023, 2023 Dec 18-20). Challenges in Adapting LLMs for Transparency: Complying with Art. 14 EU AI Act. *Frontiers in Artificial Intelligence and Applications* [Legal knowledge and information systems]. 36th Annual International Conference on Legal Knowledge and Information Systems (JURIX), Maastricht Univ, Maastricht, NETHERLANDS.
- Goswami, J., Prajapati, K. K., Saha, A., & Saha, A. K. (2024). Parameter-efficient fine-tuning large language model approach for hospital discharge paper summarization. *Applied Soft Computing*, 157, Article 111531. <https://doi.org/10.1016/j.asoc.2024.111531>
- Gou, Q., & Nguyen, C.-T. (2024). Mixed preference optimization: Reinforcement learning with data selection and better reference model. *arXiv preprint arXiv:2403.19443*.
- Graham, P. (2014). *Japanese Design: Art, Aesthetics & Culture*. Tuttle Publishing.
- Gray, T. (2018). *Islam and international criminal law and justice* (Vol. 2). Torkel Opsahl Academic EPublisher.
- Gregg, G. S. (2005). *The Middle East: a cultural psychology*. Oxford University Press.
- Grenz, S. (1996). *A Primer on Postmodernism*. Eerdmans.
- Grimm, D., Wendel, M., & Reinbacher, T. (2019). European constitutionalism and the German basic law. *National Constitutions in European and Global Governance: Democracy, Rights, the Rule of Law: National Reports*, 407-492.
- Gu, Z., He, X., Yu, P., Jia, W., Yang, X., Peng, G., Hu, P., Chen, S., Chen, H., & Lin, Y. (2024). Automatic quantitative stroke severity assessment based on Chinese clinical named entity recognition with domain-adaptive pre-trained large language

- model. *Artificial Intelligence in Medicine*, 150, Article 102822. <https://doi.org/10.1016/j.artmed.2024.102822>
- Gudykunst, W. B., & Nishida, T. (1986). Attributional confidence in low-and high-context cultures. *Human communication research*, 12(4), 525-549.
- Guerin, D. (1970). *Anarchism: From theory to practice* (Vol. 175). NYU Press.
- Guo, C., & Zhang, F. (2015). Religion and social stability: China's religious policies in the Age of Reform. *Third World Quarterly*, 36(11), 2183-2195.
- Guyer, P., & Horstmann, R.-P. (2015). Idealism.
- Haakonssen, K., & Olden-Jørgensen, S. (2017). *Ludvig Holberg (1684-1754): learning and literature in the Nordic Enlightenment*. Taylor & Francis.
- Haberler, G. (1951). Business Cycles in a Planned Economy. Conference on Business Cycles,
- Hadar-Shoval, D., Asraf, K., Mizrachi, Y., Haber, Y., & Elyoseph, Z. (2024). Assessing the alignment of large language models with human values for mental health integration: cross-sectional study using Schwartz's theory of basic values. *JMIR Mental Health*, 11, e55988.
- Hadi, M. U., Qureshi, R., Shah, A., Irfan, M., Zafar, A., Shaikh, M. B., Akhtar, N., Wu, J., & Mirjalili, S. (2023). Large language models: a comprehensive survey of its applications, challenges, limitations, and future prospects. *Authorea Preprints*.
- Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. *Minds and machines*, 30(1), 99-120.
- Haine, W. S. (2006). *Culture and customs of France*. Bloomsbury Publishing USA.
- Harrer, S. (2023). Attention is not all you need: the complicated case of ethically using large language models in healthcare and medicine. *Ebiomedicine*, 90, Article 104512. <https://doi.org/10.1016/j.ebiom.2023.104512>
- Harrison, M. (2012). Communism and economic modernization. *Available at SSRN 2099517*.
- Harvey, P. (2012). *An introduction to Buddhism: Teachings, history and practices*. Cambridge University Press.
- Havell, E. B. (1920). *The ideals of Indian art*. London: J. Murray.
- Heine, P. (2004). *Food culture in the near east, middle east, and north Africa*. Bloomsbury Publishing USA.
- Helstosky, C. (2009). *Food culture in the Mediterranean*. Bloomsbury Publishing USA.
- Herrick, J. (2010). *Humanism: An Introduction*. Prometheus Books.
- Herskovits, M. J. (1967). *The backgrounds of African art*. Biblo & Tannen Publishers.
- Hertzberg, A. (1998). *Judaism*. Simon and Schuster.
- Heyder, T., Passlack, N., & Posegga, O. (2023). Ethical management of human-AI interaction: Theory development review. *The Journal of Strategic Information Systems*, 32(3), 101772.
- Heyting, W. J. (1932). The Anglo-American Conception of Due Process of Law. *Transactions Grotius Soc'y*, 18, 175.
- Hills, M. D. (2002). Kluckhohn and Strodtbeck's values orientation theory. *Online readings in psychology and culture*, 4(4), 3.
- Hodgson, G. M. (2019). *Is socialism feasible?: Towards an alternative future*. Edward Elgar

- Publishing.
- Hoernlé, R. F. A. (1927). *Idealism as a Philosophy*. George H. Doran Company.
- Hofstede, G. (1984). *Culture's consequences: International differences in work-related values* (Vol. 5). sage.
- Hofstede, G. (2011). Dimensionalizing cultures: The Hofstede model in context. *Online readings in psychology and culture*, 2(1), 8.
- Hønneland, G., & Jørgensen, A.-K. (2003). *Implementing international environmental agreements in Russia*. Manchester University Press.
- Horn, L., & Pham, A. H. (2004). *Vietnamese tone: A new analysis*. Routledge.
- Hsueh, J. Y., Nethala, D., Singh, S., Linehan, W. M., & Ball, M. W. (2024). Investigating the clinical reasoning abilities of large language model GPT-4: an analysis of postoperative complications from renal surgeries. *Urologic oncology*, 42(9), 292.e291-292.e297. <https://doi.org/10.1016/j.urolonc.2024.04.010>
- Huang, A. H., Wang, H., & Yang, Y. (2023). FinBERT: A Large Language Model for Extracting Information from Financial Text. *Contemporary Accounting Research*, 40(2), 806-841. <https://doi.org/10.1111/1911-3846.12832>
- Huang, H., Zheng, O., Wang, D., Yin, J., Wang, Z., Ding, S., Yin, H., Xu, C., Yang, R., Zheng, Q., & Shi, B. (2023). ChatGPT for shaping the future of dentistry: the potential of multi-modal large language model [Review]. *International Journal of Oral Science*, 15(1), Article 29. <https://doi.org/10.1038/s41368-023-00239-y>
- Huang, S., Siddarth, D., Lovitt, L., Liao, T. I., Durmus, E., Tamkin, A., & Ganguli, D. (2024). Collective Constitutional AI: Aligning a Language Model with Public Input. The 2024 ACM Conference on Fairness, Accountability, and Transparency,
- Huang, Z., Quan, K., Chan, J., MacNeil, S., & Acm. (2023, 2023
- Jun 19-21). CausalMapper: Challenging designers to think in systems with Causal Maps and Large Language Model. [2023 proceedings of the 15th conference on creativity and cognition, c&c 2023]. 15th Conference on Creativity and Cognition (C and C), Electr Network.
- Huggan, G. (2007). *Australian literature: postcolonialism, racism, transnationalism*. Oxford Studies in Postcolonial.
- Hui, C. H. (1988). Measurement of individualism-collectivism. *Journal of research in personality*, 22(1), 17-36.
- Hurley, E. T., Crook, B. S., Lorentz, S. G., Danilkowicz, R. M., Lau, B. C., Taylor, D. C., Dickens, J. F., Anakwenze, O., & Klifto, C. S. (2024). Evaluation High-Quality of Information from ChatGPT (Artificial Intelligence) Large Language Model) Artificial Intelligence on Shoulder Stabilization Surgery. *Arthroscopy-the Journal of Arthroscopic and Related Surgery*, 40(3). <https://doi.org/10.1016/j.arthro.2023.07.048>
- Hurley, N. C., Schroeder, K. M., & Hess, A. S. (2023). Would doctors dream of electric blood bankers? Large language model-based artificial intelligence performs well in many aspects of transfusion medicine. *Transfusion*, 63(10), 1833-1840. <https://doi.org/10.1111/trf.17526>
- Hussain, J. (2011). *Islam*. Federation Press.
- Hutchinson, J., & Smith, A. D. (1994). *Nationalism*. Oxford University Press, USA.

- Hylland, O. M., & Bjurström, E. (2018). *Aesthetics and Politics: A Nordic Perspective on How Cultural Policy Negotiates the Agency of Music and Arts*. Springer.
- Idema, W., & Haft, L. (1997). *A guide to Chinese literature*. University of Michigan Press.
- Igarashi, Y., Nakahara, K., Norii, T., Miyake, N., Tagami, T., & Yokobori, S. (2024). Performance of a Large Language Model on Japanese Emergency Medicine Board Certification Examinations. *Journal of Nippon Medical School*, 91(2), 155-161. https://doi.org/10.1272/jnms.JNMS.2024_91-205
- Janssen, M. (2002). *Complexity and ecosystem management: the theory and practice of multi-agent systems*. Edward Elgar Publishing.
- Jennings, I. (1967). *The British Constitution*. Cambridge University Press.
- Jensen, J. (2002). *Is art good for us?: beliefs about high culture in American life*. Rowman & Littlefield.
- Jeon, J., & Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies*, 28(12), 15873-15892. <https://doi.org/10.1007/s10639-023-11834-1>
- Jeong, Y., Yang, J., Choi, I. H., & Lee, J. (2024). Feature-Based Text Search Engine Mitigating Data Diversity Problem Using Pre-Trained Large Language Model for Fast Deployment Services. *Ieee Access*, 12, 48145-48157. <https://doi.org/10.1109/access.2024.3373470>
- Jiang, Y., Irvin, J. A., Ng, A. Y., & Zou, J. (2024). VetLLM: Large Language Model for Predicting Diagnosis from Veterinary Notes. *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*, 29, 120-133. <Go to ISI>://MEDLINE:38160274
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature machine intelligence*, 1(9), 389-399.
- Johnson, M. (2015). Conservatism and ideology. In (Vol. 5, pp. 1-3): Taylor & Francis.
- Johnson, P. (2012). *history of Christianity*. Simon and Schuster.
- Kagitcibasi, C. (1997). Individualism and collectivism. *Handbook of cross-cultural psychology*, 3, 1-49.
- Kandiyali, J. (2014). Freedom and necessity in Marx's account of communism. *British Journal for the History of Philosophy*, 22(1), 104-123.
- Kasneji, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., & Hüllermeier, E. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and individual differences*, 103, 102274.
- Katz, J. J. (1997). *Realistic rationalism*. mlt Press.
- Keene, D. (1988). *The pleasures of Japanese literature*. Columbia University Press.
- Killion, M. U. (2005). China's Amended Constitution: Quest for Liberty and Independent Judicial Review. *Wash. U. Global Stud. L. Rev.*, 4, 43.
- Kim, D., Pan, Y., & Park, H. S. (1998). High-versus low-context culture: A comparison of Chinese, Korean, and American cultures. *Psychology & Marketing*, 15(6), 507-521.
- Kim, J. K., Chua, M., Rickard, M., & Lorenzo, A. (2023). ChatGPT and large language model

- (LLM) chatbots: The current state of acceptability and a proposal for guidelines on utilization in academic medicine [Review]. *Journal of Pediatric Urology*, 19(5), 598-604. <https://doi.org/10.1016/j.jpurol.2023.05.018>
- Kim, S., Bae, S., Shin, J., Kang, S., Kwak, D., Yoo, K. M., & Seo, M. (2023). Aligning large language models through synthetic feedback. *arXiv preprint arXiv:2305.13735*.
- Kim, S., Kim, S., Kim, Y., Park, J., Kim, S., Kim, M., Sung, C. H., Hong, J., Lee, Y., & Acm. (2023, 2023)
- Nov 27-29). LLMs Analyzing the Analysts: Do BERT and GPT Extract More Value from Financial Analyst Reports? [Proceedings of the 4th acm international conference on ai in finance, icaif 2023]. 4th ACM International Conference on AI in Finance (ICAIF), Brooklyn, NY.
- King, A. (2007). *The British Constitution*. OUP Oxford.
- Klingefjord, O., Lowe, R., & Edelman, J. (2024). What are human values, and how do we align AI to them? *arXiv preprint arXiv:2404.10636*.
- Klostermaier, K. K. (2007). *A survey of Hinduism*. Suny Press.
- Klostermaier, K. K. (2014). *Hinduism: A short history*. Simon and Schuster.
- Klug, H. (2010). *The constitution of South Africa: a contextual analysis*. Bloomsbury Publishing.
- Kohl, S., & Sapién, A. (2021). Debunking the Holy Trinity: Competition, Individualism, and Meritocracy. *Competition: What It Is and Why It Happens*, 202-220.
- Kollitsch, L., Eredics, K., Marszalek, M., Rauchenwald, M., Brookman-May, S. D., Burger, M., Koerner-Riffard, K., & May, M. (2024). How does artificial intelligence master urological board examinations? A comparative analysis of different Large Language Models' accuracy and reliability in the 2022 In-Service Assessment of the European Board of Urology. *World Journal of Urology*, 42(1), Article 20. <https://doi.org/10.1007/s00345-023-04749-6>
- Kopelman, R. E., Rovenpor, J. L., & Guan, M. (2003). The Study of Values: Construction of the fourth edition. *Journal of Vocational Behavior*, 62(2), 203-220.
- Kowalik, T. (1990). Central planning. In *Problems of the Planned Economy* (pp. 42-50). Springer.
- Krafft, A. M., & Choubisa, R. (2018). Hope in the Indian psychology context: Philosophical foundations and empirical findings. *Hope for a Good Life: Results of the Hope-Barometer International Research Program*, 131-163.
- Kremyanskaya, E. A., Kuznetsova, T. O., & Rakitskaya, I. A. (2014). *Russian constitutional law*. Cambridge Scholars Publishing.
- Kropotkin, P. A. (1905). *Russian literature*. Duckworth & Company.
- Kropotkin, P. A. (1915). *Ideals and realities in Russian literature*. AA Knopf.
- Künnecke, M. (2007). *Tradition and change in administrative law: an Anglo-German comparison*. Springer Science & Business Media.
- Lacerda Jr, F. (1815). Socialism. *Encyclopedia of critical psychology*, 1820.
- Lahat, A., Shachar, E., Avidan, B., Glicksberg, B., & Klang, E. (2023). Evaluating the Utility of a Large Language Model in Answering Common Patients' Gastrointestinal Health-Related Questions: Are We There Yet? *Diagnostics*, 13(11), Article 1950.

- <https://doi.org/10.3390/diagnostics13111950>
- Lai, Z., Wu, T., Fei, X., & Ling, Q. (2024). BERT4ST:: Fine-tuning pre-trained large language model for wind power forecasting. *Energy Conversion and Management*, 307, Article 118331. <https://doi.org/10.1016/j.enconman.2024.118331>
- Landesman, B. M. (1983). Egalitarianism. *Canadian Journal of philosophy*, 13(1), 27-56.
- Lane, D. C., Bianchi, C., & Bivona, E. (2010). Public sector applications of the system dynamics approach. *Systems Research and Behavioral Science*, 27(4), 355–360.
- Langsted, L. B., Garde, P., Greve, V., & Elholm, T. (2019). *Criminal law in Denmark*. Kluwer Law International BV.
- Lazarus, R. J. (2008). Super wicked problems and climate change: Restraining the present to liberate the future. *Cornell L. Rev.*, 94, 1153.
- Lebo, F. B. (2018). *Between democracy and technocracy: Regulating administrative guidance in Japan*. Lexington Books.
- Lees, A., Tran, V. Q., Tay, Y., Sorensen, J., Gupta, J., Metzler, D., & Vasserman, L. (2022). A new generation of perspective api: Efficient multilingual character-level transformers. Proceedings of the 28th ACM SIGKDD conference on knowledge discovery and data mining,
- Leith, J. A. (1964). *The Idea of Art as Propaganda in France, 1750-1799: A Study in the History of Ideas*. University of Toronto Press.
- Lemley, K. V. (2024). Does ChatGPT help us understand the medical literature? *Journal of the American Society of Nephrology*, 35(2), 232-233.
- Lévi-Strauss, C. (2004). Structuralism. *Visions of Culture: An Introduction to Anthropological Theories and Theorists*, 231-246.
- Li, M. (2018). Maintaining ties and reaffirming unity: Family rituals in the age of migration. *Journal of Family Communication*, 18(4), 286-301.
- Lin, Y. (2015). Constitutional evolution through legislation: The quiet transformation of China's Constitution. *International Journal of Constitutional Law*, 13(1), 61-89.
- Lindholm, C. (2008). *The Islamic middle east: tradition and change*. John Wiley & Sons.
- Liu, D., Yan, Y., Gong, Y., Qi, W., Zhang, H., Jiao, J., Chen, W., Fu, J., Shou, L., & Gong, M. (2020). GLGE: A new general language generation evaluation benchmark. *arXiv preprint arXiv:2011.11928*.
- Liu, F.-c., Simon, D. F., Sun, Y.-t., & Cao, C. (2011). China's innovation policies: Evolution, institutional structure, and trajectory. *Research Policy*, 40(7), 917-931.
- Liu, F., Zhu, T., Wu, X., Yang, B., You, C., Wang, C., Lu, L., Liu, Z., Zheng, Y., Sun, X., Yang, Y., Clifton, L., & Clifton, D. A. (2023). A medical multimodal large language model for future pandemics. *NPJ digital medicine*, 6(1), Article 226. <https://doi.org/10.1038/s41746-023-00952-2>
- Liu, J. (2011). *Chinese food*. Cambridge University Press.
- Liu, S., Wright, A. P., McCoy, A. B., Huang, S. S., Genkins, J. Z., Peterson, J. F., Kumah-Crystal, Y. A., Martinez, W., Carew, B., Mize, D., Steitz, B., & Wright, A. (2024). Using large language model to guide patients to create efficient and comprehensive clinical care message [; Early Access]. *Journal of the American Medical Informatics Association*. <https://doi.org/10.1093/jamia/ocae142>

- Liu, Y., Cao, J., Liu, C., Ding, K., & Jin, L. (2024). Datasets for large language models: A comprehensive survey. *arXiv preprint arXiv:2402.18041*.
- Liu, Y., Han, T., Ma, S., Zhang, J., Yang, Y., Tian, J., He, H., Li, A., He, M., & Liu, Z. (2023). Summary of chatgpt-related research and perspective towards the future of large language models. *Meta-Radiology*, 100017.
- Loglo, S. (2010). Research on Mongolian lexical analyzer based on NFA. 2010 IEEE International Conference on Intelligent Computing and Intelligent Systems,
- Lombardi, C. B. (2013). Designing Islamic constitutions: Past trends and options for a democratic future. *International Journal of Constitutional Law*, 11(3), 615-645.
- Long-Solís, J., & Vargas, L. A. (2005). *Food culture in Mexico*. Bloomsbury Publishing USA.
- Lubell, M. (2015). Collaborative partnerships in complex institutional systems. *Current Opinion in Environmental Sustainability*, 12, 41-47.
- Lund, B. D., Wang, T., Mannuru, N. R., Nie, B., Shimray, S., & Wang, Z. (2023). ChatGPT and a new academic reality: Artificial Intelligence-written research papers and the ethics of the large language models in scholarly publishing. *Journal of the Association for Information Science and Technology*, 74(5), 570-581. <https://doi.org/10.1002/asi.24750>
- Luo, Y., Lin, Z., Zhang, Y., Sun, J., Lin, C., Xu, C., Su, X., Shen, Y., Guo, J., & Gong, Y. (2024). Ensuring Safe and High-Quality Outputs: A Guideline Library Approach for Language Models. *arXiv preprint arXiv:2403.11838*.
- Lynn Fitzpatrick, R. (2007). A literature review exploring values alignment as a proactive approach to conflict management. *International journal of conflict management*, 18(3), 280-305.
- Ma, X., & Ortolano, L. (2000). *Environmental regulation in China: Institutions, enforcement, and compliance*. Rowman & Littlefield.
- Ma, X., Zhao, R., Liu, Y., Deng, C., & Du, D. (2024). Design of a large language model for improving customer service in telecom operators. *Electronics Letters*, 60(10), Article e13218. <https://doi.org/10.1049/ell2.13218>
- Mackaay, E. (2021). *Law and economics for civil law systems*. Edward Elgar Publishing.
- Macleod, C. M., & Macleod, C. M. (1998). *Liberalism, Justice, and Markets: A critique of liberal equality*. Oxford University Press.
- Madsen, F. (2007). *Commercial arbitration in Sweden*. Oxford University Press.
- Maggs, P. B., Schwartz, O., & Burnham, W. (2015). *Law and legal system of the Russian Federation*. Juris Publishing, Inc.
- Mannstadt, I., & Mehta, B. (2024). Large language models and the future of rheumatology: assessing impact and emerging opportunities [Review]. *Current Opinion in Rheumatology*, 36(1), 46-51. <https://doi.org/10.1097/bor.0000000000000981>
- Markie, P., & Folescu, M. (2004). Rationalism vs. empiricism.
- Mattison, M. (2000). Ethical decision making: The person in the process. *Social Work*, 45(3), 201-212.
- Mayton, D. M., Ball-Rokeach, S. J., & Loges, W. E. (1994). Human values and social issues: An introduction. *Journal of social issues*, 50(4), 1-8.
- McCarney, J. (1991). The true realm of freedom: Marxist philosophy after communism.

- New Left Review*, 189, 1991.
- McElwain, K. M. (2020). The Japanese constitution. In *The Oxford Handbook of Japanese Politics*.
- McGovern, J. (2007). Conservatism: Authority in the Modern State.
- McGraw, G., Bonett, R., Figueroa, H., & McMahon, K. (2024). 23 Security Risks in Black-Box Large Language Model Foundation Models. *Computer*, 57(4), 160-164. <https://doi.org/10.1109/mc.2024.3363250>
- Medalie, R. J. (1959). The communist theory of state. *American Slavic and East European Review*, 18(4), 510-525.
- Meltzer, A. H. (2012). *Why capitalism?* Oxford University Press.
- Mesko, B., & Topol, E. J. J. (2023). The imperative for regulatory oversight of large language models (or generative AI) in healthcare. *NPJ digital medicine*, 6(1), Article 120. <https://doi.org/10.1038/s41746-023-00873-0>
- Metelev, S. E., Murat, M. M., & Lizunov, V. (2016). *Economic security policy of the Russian Federation*. Libertas.
- Miao, J., Thongprayoon, C., Suppadungsuk, S., Krisanapan, P., Radhakrishnan, Y., & Cheungpasitporn, W. (2024). Chain of Thought Utilization in Large Language Models and Application in Nephrology [Review]. *Medicina-Lithuania*, 60(1), Article 148. <https://doi.org/10.3390/medicina60010148>
- Miller, D. (1990). *Market, state, and community: Theoretical foundations of market socialism*. Oxford University Press.
- Minaee, S., Mikolov, T., Nikzad, N., Chenaghlu, M., Socher, R., Amatriain, X., & Gao, J. (2024). Large language models: A survey. *arXiv preprint arXiv:2402.06196*.
- Minkov, M., & Hofstede, G. (2012). Hofstede's fifth dimension: New evidence from the World Values Survey. *Journal of cross-cultural psychology*, 43(1), 3-14.
- Mirković, J., Seidenberg, M. S., & Joanisse, M. F. (2011). Rules versus statistics: Insights from a highly inflected language. *Cognitive Science*, 35(4), 638-681.
- Mohammed, N. (1988). Principles of Islamic contract law. *Journal of Law and religion*, 6(1), 115-130.
- Mondal, D., & Lipizzi, C. (2024). Mitigating Large Language Model Bias: Automated Dataset Augmentation and Prejudice Quantification. *Computers*, 13(6), Article 141. <https://doi.org/10.3390/computers13060141>
- Monroe, C. L., Abdelhafez, Y. G., Atsina, K., Aman, E., Nardo, L., & Madani, M. H. (2024). Evaluation of responses to cardiac imaging questions by the artificial intelligence large language model ChatGPT. *Clinical Imaging*, 112, Article 110193. <https://doi.org/10.1016/j.clinimag.2024.110193>
- Montagna, S., Ferretti, S., Klopfenstein, L. C., Florio, A., & Pengo, M. F. (2023). Data decentralisation of LLM-based chatbot systems in chronic disease self-management. Proceedings of the 2023 ACM Conference on Information Technology for Social Good,
- Morales, S., Clariso, R., Cabot, J., & Ieee. (2023, 2023 Sep 11-15). Automating Bias Testing of LLMs. *IEEE ACM International Conference on Automated Software Engineering* [2023 38th IEEE/ACM International Conference

- on automated software engineering, ase]. 38th IEEE/ACM International Conference on Automated Software Engineering (ASE), Echternach, LUXEMBOURG.
- Moser, P. K., & Trout, J. D. (2002). *Contemporary materialism: a reader*. Routledge.
- Muhlenbach, F. (2020). A methodology for ethics-by-design AI systems: dealing with human value conflicts. 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC),
- Munshi, S., & Abraham, B. P. (2004). *Good governance, democratic societies and globalization*. Sage.
- Munteanu, I., & Newcomer, K. (2020). Leading and learning through dynamic performance management in government. *Public Administration Review*, 80(2), 316-325.
- Munyer, T., Tanvir, A. A., Das, A., & Zhong, X. (2024). DeepTextMark: A Deep Learning-Driven Text Watermarking Approach for Identifying Large Language Model Generated Text. *Ieee Access*, 12, 40508-40520. <https://doi.org/10.1109/access.2024.3376693>
- Musacchio, A., Lazzarini, S. G., & Aguilera, R. V. (2015). New varieties of state capitalism: Strategic and governance implications. *Academy of management perspectives*, 29(1), 115-131.
- Narasu, P. L. (1907). *The essence of Buddhism*. Srinivasa Varadachari.
- Nasseri, M., Brandtner, P., Zimmermann, R., Falatouri, T., Darbanian, F., & Obinwanne, T. (2023, 2023Jul 23-28). Applications of Large Language Models (LLMs) in Business Analytics - Exemplary Use Cases in Data Preparation Tasks. *Lecture Notes in Computer Science* [Hci international 2023 late breaking papers, hcii 2023, pt vi]. 25th International Conference on Human-Computer Interaction (HCI International), Copenhagen, DENMARK.
- Nazi, Z. A., & Peng, W. (2024). Large language models in healthcare and medical domain: A review. *Informatics*, 11(3), 57. <https://www.mdpi.com/2227-9709/11/3/57>
- Newman, D. G. (2013). Value collectivism, collective rights, and self-threatening theory. *Oxford Journal of Legal Studies*, 33(1), 197-210.
- Newmarch, R. (1916). *The Russian Arts*. H. Jenkins.
- Ni, B., & Buehler, M. J. (2024). MechAgents: Large language model multi-agent collaborations can solve mechanics problems, generate new data, and integrate knowledge. *Extreme Mechanics Letters*, 67, Article 102131. <https://doi.org/10.1016/j.eml.2024.102131>
- Nolan, P. (2007). *Capitalism and freedom: the contradictory character of globalisation*. Anthem press.
- O'Riordan, T. (1991). The new environmentalism and sustainable development. *Science of the Total Environment*, 108(1-2), 5-15.
- O'hara, K. (2011). *Conservatism*. Reaktion Books.
- Oda, H. (2021). *Japanese law*. Oxford University Press.
- Offord, B., Kerruish, E., Garbutt, R., Wessell, A., & Pavlovic, K. (2014). *Inside Australian*

- culture: Legacies of enlightenment values*. Anthem Press.
- Oishi, S., Hahn, J., Schimmack, U., Radhakrishnan, P., Dzokoto, V., & Ahadi, S. (2005). The measurement of values across cultures: A pairwise comparison approach. *Journal of research in Personality*, 39(2), 299–305.
- Oliver, P., Macklem, P., & Des Rosiers, N. (2017). *The Oxford Handbook of the Canadian Constitution*. Oxford University Press.
- Olsen, J. P. (2005). Maybe It Is Time to Rediscover Bureaucracy. *Journal of Public Administration Research and Theory*, 16(1), 1–24. <https://doi.org/10.1093/jopart/mui027>
- Olson, R. G. (2012). *An introduction to existentialism*. Courier Corporation.
- Omiye, J. A., Gui, H., Rezaei, S. J., Zou, J., & Daneshjou, R. (2024). Large Language Models in Medicine: The Potentials and Pitfalls A Narrative Review [Review]. *Annals of Internal Medicine*, 177(2). <https://doi.org/10.7326/m23-2772>
- Ong, J. C. L., Chang, S. Y.-H., William, W., Butte, A. J., Shah, N. H., Chew, L. S. T., Liu, N., Doshi-Velez, F., Lu, W., & Savulescu, J. (2024). Ethical and regulatory challenges of large language models in medicine. *The Lancet Digital Health*, 6(6), e428–e432.
- Orfield, L. B. (2002). *The growth of Scandinavian law*. The Lawbook Exchange, Ltd.
- Ortmann, S. (2020). Evolving environmental governance structures in a market socialist state: The case of Vietnam. *The Socialist Market Economy in Asia: Development in China, Vietnam and Laos*, 195–217.
- Ost, D. (2015). Class after communism: Introduction to the special issue. In (Vol. 29, pp. 543–564): Sage Publications Sage CA: Los Angeles, CA.
- Pagoni, E. G., & Patroklos, G. (2019). A system dynamics model for the assessment of national public–private partnership programmes' sustainable performance. *Simulation modelling practice and theory*, 97, 101949.
- Painter, S. (1960). The family and the feudal system in twelfth century England. *Speculum*, 35(1), 1–16.
- Paiva, B. M., Ferreira, F. A., Carayannis, E. G., Zopounidis, C., Ferreira, J. J., Pereira, L. F., & Dias, P. J. (2021). Strategizing sustainability in the banking industry using fuzzy cognitive maps and system dynamics. *International Journal of Sustainable Development & World Ecology*, 28(2), 93–108.
- Pal, S., Bhattacharya, M., Lee, S.-S., & Chakraborty, C. (2024). A domain-specific next-generation large language model (LLM) or ChatGPT is required for biomedical engineering and research. *Annals of Biomedical Engineering*, 52(3), 451–454.
- Pang, X., Tang, S., Ye, R., Xiong, Y., Zhang, B., Wang, Y., & Chen, S. (2024). Self-Alignment of Large Language Models via Monopolylogue-based Social Scene Simulation. *arXiv preprint arXiv:2402.05699*.
- Parasecoli, F. (2004). *Food culture in Italy*. Bloomsbury Publishing USA.
- Paroiu, R., Ruseti, S., Dascalu, M., Trausan-Matu, S., & McNamara, D. S. (2023). Asking Questions about Scientific Articles-Identifying Large N Studies with LLMs. *Electronics*, 12(19), Article 3996. <https://doi.org/10.3390/electronics12193996>
- Peacocke, C. (2002). Three principles of rationalism. *European Journal of Philosophy*, 16(3),

375-397.

- Peng, B., Song, L., Tian, Y., Jin, L., Mi, H., & Yu, D. (2023). Stabilizing rlhf through advantage model and selective rehearsal. *arXiv preprint arXiv:2309.10202*.
- Pepper, D. (2019). *The roots of modern environmentalism*. Routledge.
- Pérez, E. O. (2015). *The language-opinion connection*. Oxford Uni.)(doi: 10.1093/oxfordhb/9780190213299.013. 18).
- Phillips, M. A., & Ritala, P. (2019). A complex adaptive systems agenda for ecosystem research methodology. *Technological Forecasting and Social Change*, 148, 119739.
- Picavet, E. (2024). Hierarchies and the Complementarity of Authority and Equality. *Fudan Journal of the Humanities and Social Sciences*, 1-20.
- Piktus, A., Akiki, C., Villegas, P., Laurenconl, H., Dupont, G., Luccionil, A. S., Jernitel, Y., & Rogers, A. (2023, 2023 Jul 09-14). The ROOTS Search Tool: Data Transparency for LLMs. [Proceedings of the 61st annual meeting of the association for computational linguistics, acl-demo 2023, vol 3]. 61st Annual Meeting of the Association-for-Computational-Linguistics - System Demonstrations (ACL-DEMO), Toronto, CANADA.
- Pipes, R. (2003). *Communism: A history* (Vol. 7). Modern Library.
- Pitkäranta, T., & Pitkäranta, L. (2024). Bridging Human and AI Decision-Making with LLMs: The RAGADA Approach. Proceedings of the 26th International Conference on Enterprise Information Systems, Query date,
- Piurko, Y., Schwartz, S. H., & Davidov, E. (2011). Basic personal values and the meaning of left-right political orientations in 20 countries. *Political Psychology*, 32(4), 537-561.
- Plant, L. (1999). Democratic socialism and equality. In *Welfare Policy in Britain: The Road from 1945* (pp. 94-115). Springer.
- Plattner, M. F. (1999). From liberalism to liberal democracy. *Journal of democracy*, 10(3), 121-134.
- Pollock, F. (2020). State capitalism: its possibilities and limitations. In *Critical theory and society* (pp. 95-118). Routledge.
- Powell, J. (2007). *Postmodernism for beginners*. Red Wheel/Weiser.
- Przeworski, A. (1980). Material interests, class compromise, and the transition to socialism. *Politics & Society*, 10(2), 125-153.
- Rahman, F. (2020). *Islam*. University of Chicago Press.
- Raiaan, M. A. K., Mukta, M. S. H., Fatema, K., Fahad, N. M., Sakib, S., Mim, M. M. J., Ahmad, J., Ali, M. E., & Azam, S. (2024). A review on large Language Models: Architectures, applications, taxonomies, open issues and challenges. *IEEE Access*.
- Rand, A. (1967). *What is capitalism?* Second Renaissance Book Service.
- Rane, N. L., Tawde, A., Choudhary, S. P., & Rane, J. (2023). Contribution and performance of ChatGPT and other Large Language Models (LLM) for scientific and research advancements: a double-edged sword. *International Research Journal of Modernization in Engineering Technology and Science*, 5(10), 875-899.
- Rao, A., Kim, J., Kamineni, M., Pang, M., Lie, W., & Succi, M. D. (2023). Evaluating ChatGPT

- as an adjunct for radiologic decision-making. *MedRxiv*, 2023.2002.2002.23285399.
- Ratzinger, J. (2004). *Introduction to christianity*. Ignatius Press.
- Reamer, F. (2018). *Social work values and ethics*. Columbia University Press.
- Redhead, D., & Power, E. A. (2022). Social hierarchies and social networks in humans. *Philosophical Transactions of the Royal Society B*, 377(1845), 20200440.
- Rehman, U. U., Hussain, M., Nguyen, T. D. T., Lee, S., & Assoc Computing, M. (2023, 2023 Dec 16-18). Let's Hide from LLMs: An Adaptive Contextual Privacy Preservation Method for Time Series Data. [Proceedings of 2023 6th artificial intelligence and cloud computing conference, aiccc 2023]. 6th Artificial Intelligence and Cloud Computing Conference (AICCC) / 5th Asia Digital Image Processing Conference, Kyoto, JAPAN.
- Remijsen, B. (2016). Tone. In *Oxford Research Encyclopedia of Linguistics*.
- Resnick, S., & Wolff, R. (1988). Communism: Between class and classless. *Rethinking Marxism*, 1(1), 14-42.
- Reynolds, J. (2014). *Understanding existentialism*. Routledge.
- Roemer, J. E. (1992). Can there be Socialism after Communism? *Politics & Society*, 20(3), 261-276.
- Rokeach, M. (1967). Rokeach value survey. *The nature of human values*.
- Rokeach, M. (2008). *Understanding human values*. Simon and Schuster.
- Roumeliotis, K. I., Tselikas, N. D., & Nasiopoulos, D. K. (2024a). LLMs and NLP Models in Cryptocurrency Sentiment Analysis: A Comparative Classification Study. *Big Data and Cognitive Computing*, 8(6), Article 63. <https://doi.org/10.3390/bdcc8060063>
- Roumeliotis, K. I., Tselikas, N. D., & Nasiopoulos, D. K. (2024b). Next-Generation Spam Filtering: Comparative Fine-Tuning of LLMs, NLPs, and CNN Models for Email Spam Classification. *Electronics*, 13(11), Article 2034. <https://doi.org/10.3390/electronics13112034>
- Rubin, P. H. (2000). Hierarchy. *Human nature*, 11, 259-279.
- Rubin, P. H., & Klumpp, T. (2012). Property rights and capitalism. *The Oxford Handbook of Capitalism*, 204-219.
- Russo, C., Danioni, F., Zagrean, I., & Barni, D. (2022). Changing personal values through value-manipulation tasks: a systematic literature review based on Schwartz's theory of basic human values. *European Journal of Investigation in Health, Psychology and Education*, 12(7), 692-715.
- Sagiv, L., & Schwartz, S. H. (2022). Personal values across cultures. *Annual review of psychology*, 73(1), 517-546.
- Salim, A. (2015). *Contemporary Islamic law in Indonesia: Sharia and legal pluralism*. Edinburgh University Press.
- Sallam, M. (2023). The utility of ChatGPT as an example of large language models in healthcare education, research and practice: Systematic review on the future perspectives and potential limitations. *MedRxiv*, 2023-02.
- Savage, T., Nayak, A., Gallo, R., Rangan, E., & Chen, J. H. (2024). Diagnostic reasoning prompts reveal the potential for large language model interpretability in medicine.

- NPJ digital medicine*, 7(1), Article 20. <https://doi.org/10.1038/s41746-024-01010-1>
- Savelka, J., Ashley, K. D., Gray, M. A., Westermann, H., & Xu, H. (2023). Explaining legal concepts with augmented large language models (gpt-4). *arXiv preprint arXiv:2306.09525*.
- Sayers, A. (2001). *Australian art*. Oxford University Press.
- Scherrer, N., Shi, C., Feder, A., & Blei, D. M. (2023, 2023Dec 10-16). Evaluating the Moral Beliefs Encoded in LLMs. *Advances in Neural Information Processing Systems* [Advances in neural information processing systems 36 (neurips 2023)]. 37th Conference on Neural Information Processing Systems (NeurIPS), New Orleans, LA.
- Schiff, D. (2022). Education for AI, not AI for education: The role of education and ethics in national AI policy strategies. *International Journal of Artificial Intelligence in Education*, 32(3), 527-563.
- Schwartz, S. (2012). Toward refining the theory of basic human values. *Methods, theories, and empirical applications in the social sciences*, 39-46.
- Schwartz, S. H. (1994). Are there universal aspects in the structure and contents of human values? *Journal of social issues*, 50(4), 19-45.
- Schwartz, S. H. (2003). A proposal for measuring value orientations across nations. *Questionnaire package of the european social survey*, 259(290), 261.
- Schwartz, S. H. (2007). Basic human values: Theory, methods, and application. *Risorsa Uomo*(2007/2).
- Schwartz, S. H. (2009). Basic human values. *sociologie*, 42, 249-288.
- Schwartz, S. H. (2010). Basic values: how they motivate and inhibit prosocial behavior.
- Schwartz, S. H. (2012). An overview of the Schwartz theory of basic values. *Online readings in psychology and culture*, 2(1), 11.
- Schwartz, S. H. (2021). A repository of Schwartz value scales with instructions and an introduction. *Online readings in psychology and culture*, 2(2), 9.
- Schwartz, S. H., & Bilsky, W. (1987). Toward a universal psychological structure of human values. *Journal of personality and social psychology*, 53(3), 550.
- Schwartz, S. H., & Bilsky, W. (1990). Toward a theory of the universal content and structure of values: Extensions and cross-cultural replications. *Journal of personality and social psychology*, 58(5), 878.
- Schwartz, S. H., & Boehnke, K. (2004). Evaluating the structure of human values with confirmatory factor analysis. *Journal of research in Personality*, 38(3), 230-255.
- Schwartz, S. H., Cieciuch, J., Vecchione, M., Davidov, E., Fischer, R., Beierlein, C., Ramos, A., Verkasalo, M., Lönnqvist, J.-E., & Demirutku, K. (2012). Refining the theory of basic individual values. *Journal of personality and social psychology*, 103(4), 663.
- Self, P. (2017). Socialism. *A Companion to Contemporary Political Philosophy*, 414-438.
- Sellars, R. W. (1922). *Evolutionary naturalism*. Open Court Publishing Company.
- Sen, C. T. (2004). *Food culture in India*. Bloomsbury Publishing USA.
- Setyoningrum, Y., Piliang, Y. A., Santosa, I., & Wahyudi, D. (2015). Hierarchical social structure and the social spaces formation in traditional Limasan house of Javanese

- middle-class society. *The International Journal of Social Sciences*, 32(1), 37-45.
- Shah, N. H., Entwistle, D., & Pfeffer, M. A. (2023). Creation and Adoption of Large Language Models in Medicine [Review]. *Jama-Journal of the American Medical Association*, 330(9), 866-869. <https://doi.org/10.1001/jama.2023.14217>
- Shaikh, A. (2016). *Capitalism: Competition, conflict, crises*. Oxford University Press.
- Shapiro, B. J. (2023). *Beyond reasonable doubt and probable cause: Historical perspectives on the Anglo-American law of evidence*. Univ of California Press.
- Sharma, M., Farough, S., Burkett, A., Prasanth, J., El-Shafeey, N., Zygadlo, D., Dunn, C., & Korn, R. (2024, 2024Feb 19-21). Leveraging LLMs like ChatGPT for robust quality checks and medical text agreement rationale enhancing adjudication quality and alignment in BICR for oncology clinical trials. *Proceedings of SPIE [Imaging informatics for healthcare, research, and applications, medical imaging 2024]*. Conference on Medical Imaging - Imaging Informatics for Healthcare, Research, and Applications, San Diego, CA.
- Shearer, B. F. (2007). *Culture and Customs of the United States:[2 volumes]*. Bloomsbury Publishing USA.
- Shen, Y., Heacock, L., Elias, J., Hentel, K. D., Reig, B., Shih, G., & Moy, L. (2023). ChatGPT and other large language models are double-edged swords. In (Vol. 307, pp. e230163): Radiological Society of North America.
- Sherman, H. J. (1970). The economics of pure communism. *Review of Radical Political Economics*, 2(4), 39-50.
- Shively, D. H. (2015). *Tradition and modernization in Japanese culture*. Princeton University Press.
- Shool, S., Adimi, S., Saboori Amleshi, R., Bitaraf, E., Golpira, R., & Tara, M. (2025). A systematic review of large language model (LLM) evaluations in clinical medicine. *BMC Medical Informatics and Decision Making*, 25(1), 117. <https://doi.org/10.1186/s12911-025-02954-4>
- Shumway, D. O., & Hartman, H. J. (2024). Medical malpractice liability in large language model artificial intelligence: legal review and policy recommendations. *Journal of Osteopathic Medicine*, 124(7), 287-290. <https://doi.org/10.1515/jom-2023-0229>
- Shumway, D. R. (1994). *Creating American civilization: A genealogy of American literature as an academic discipline* (Vol. 11). U of Minnesota Press.
- Siau, K., & Wang, W. (2020). Artificial intelligence (AI) ethics: ethics of AI and ethical AI. *Journal of Database Management (JDM)*, 31(2), 74-87.
- Simonyi, A., & Cagan, D. (2017). *Nordic ways*. Brookings Institution Press.
- Singh, Y. (2018). Modernization of Indian tradition. In. Rawat Publications.
- Sloane, P. D. (1988). The status of islamic law in the modern commercial world. *The International Lawyer*, 743-766.
- Smith, A. (2014). Equality constitutional adjudication in South Africa: focus: twenty years of the South African Constitution. *African human rights law journal*, 14(2), 609-632.
- Smith, P. B., & Bond, M. H. (2019). Cultures and persons: Characterizing national and other types of cultural difference can also aid our understanding and prediction

- of individual variability. *Frontiers in Psychology*, 10, 2689.
- Smith, S., Patwary, M., Norick, B., LeGresley, P., Rajbhandari, S., Casper, J., Liu, Z., Prabhumoye, S., Zerveas, G., & Korthikanti, V. (2022). Using deepspeed and megatron to train megatron-turing nlg 530b, a large-scale generative language model. *arXiv preprint arXiv:2201.11990*.
- Song, F., Yu, B., Li, M., Yu, H., Huang, F., Li, Y., & Wang, H. (2024). Preference ranking optimization for human alignment. Proceedings of the AAAI Conference on Artificial Intelligence,
- Stafford, C. (2014). Separation, reunion and the Chinese attachment to place. In *Internal and International Migration* (pp. 315-330). Routledge.
- Steinberg, M. (1947). *Basic Judaism* (Vol. 88). Houghton Mifflin Harcourt.
- Sterman, J. (2000). *Instructor's manual to accompany business dynamics: Systems thinking and modeling for a complex world*. McGraw-Hill.
- Stott, J. (2017). *Basic Christianity*. Wm. B. Eerdmans Publishing.
- Studin, I. (2017). *Russia: Strategy, Policy and Administration*. Springer.
- Sturrock, J. (2008). *Structuralism*. John Wiley & Sons.
- Suleiman, E. N. (2015). *Politics, power, and bureaucracy in France: The administrative elite* (Vol. 1257). Princeton University Press.
- Sun, H., Zhang, Z., Deng, J., Cheng, J., & Huang, M. (2023). Safety assessment of chinese large language models. *arXiv preprint arXiv:2304.10436*.
- Sun, T., Zhang, X., He, Z., Li, P., Cheng, Q., Liu, X., Yan, H., Shao, Y., Tang, Q., Zhang, S., Zhao, X., Chen, K., Zheng, Y., Zhou, Z., Li, R., Zhan, J., Zhou, Y., Li, L., Yang, X., . . . Qiu, X. (2024). MOSS: An Open Conversational Large Language Model [; Early Access]. *Machine Intelligence Research*. <https://doi.org/10.1007/s11633-024-1502-8>
- Tabak, M. (2000). Marx's theory of proletarian dictatorship revisited. *Science & Society*, 333-356.
- Tännsjö, T. (2022). Conservatism. In *Oxford Research Encyclopedia of Politics*.
- Tanwar, E., Dutta, S., Borthakur, M., & Chakraborty, T. (2023, 2023 Jul 09-14). Multilingual LLMs are Better Cross-lingual In-context Learners with Alignment. [Proceedings of the 61st annual meeting of the association for computational linguistics, acl 2023, vol 1]. 61st Annual Meeting of the the Association-for-Computational-Linguistics (ACL), Toronto, CANADA.
- Tarasov, O. (2012). *Framing Russian art: from early icons to Malevich*. Reaktion Books.
- Tavasoli, A., Sharbaf, M., & Madani, S. M. (2025). *Responsible Innovation: A Strategic Framework for Financial LLM Integration* (No. arXiv:2504.02165). arXiv. <https://doi.org/10.48550/arXiv.2504.02165>
- Taylor-Gooby, P. (1991). Welfare state regimes and welfare citizenship. *Journal of European social policy*, 1(2), 93-105.
- Teixeira, A. C., Marar, V., Yazdanpanah, H., Oliveira, A., Ghassemi, M., & Acm. (2023, 2023 Nov 27-29). Enhancing Credit Risk Reports Generation using LLMs: An Integration of Bayesian Networks and Labeled Guide Prompting. [Proceedings of the 4th acm

- international conference on ai in finance, icaif 2023]. 4th ACM International Conference on AI in Finance (ICAIF), Brooklyn, NY.
- Thirunavukarasu, A. J., Ting, D. S. J., Elangovan, K., Gutierrez, L., Tan, T. F., & Ting, D. S. W. (2023). Large language models in medicine [Review]. *Nature Medicine*, 29(8), 1930-1940. <https://doi.org/10.1038/s41591-023-02448-8>
- Thoppilan, R., De Freitas, D., Hall, J., Shazeer, N., Kulshreshtha, A., Cheng, H.-T., Jin, A., Bos, T., Baker, L., & Du, Y. (2022). Lamda: Language models for dialog applications. *arXiv preprint arXiv:2201.08239*.
- Tian, S., Jin, Q., Yeganova, L., Lai, P.-T., Zhu, Q., Chen, X., Yang, Y., Chen, Q., Kim, W., Comeau, D. C., Islamaj, R., Kapoor, A., Gao, X., & Lu, Z. (2024). Opportunities and challenges for ChatGPT and large language models in biomedicine and health [Review]. *Briefings in Bioinformatics*, 25(1), Article bbad493. <https://doi.org/10.1093/bib/bbad493>
- Tomašev, N., Cornebise, J., Hutter, F., Mohamed, S., Picciariello, A., Connelly, B., Belgrave, D. C., Ezer, D., Haert, F. C. v. d., & Mugisha, F. (2020). AI for social good: unlocking the opportunity for positive impact. *Nature Communications*, 11(1), 2468.
- Tucker, J. R. (1899). *The Constitution of the United States: A Critical Discussion of Its Genesis, Development, and Interpretation* (Vol. 1). Callaghan & Company.
- Tuohy, C. J., & Tuohy, C. (2010). *Policy Politics Canada*. Temple University Press.
- Umbrello, S., & Van de Poel, I. (2021). Mapping value sensitive design onto AI for social good principles. *AI and Ethics*, 1(3), 283-296.
- Valentini, M., Szkandera, J., Smolle, M., Scheipl, S., Leithner, A., & Andreou, D. (2024). Artificial intelligence large language model ChatGPT: is it a trustworthy and reliable source of information for sarcoma patients? *Frontiers in Public Health*, 12, Article 1303319. <https://doi.org/10.3389/fpubh.2024.1303319>
- Varshney, L. R., Keskar, N. S., & Socher, R. (2020). Limits of detecting text generated by large-scale language models. 2020 Information Theory and Applications Workshop (ITA),
- Venerito, V., & Iannone, F. (2024). Large language model-driven sentiment analysis for facilitating fibromyalgia diagnosis. *RMD open*, 10(2). <https://doi.org/10.1136/rmdopen-2024-004367>
- Verdery, K. (2004). The property regime of socialism. *Conservation and Society*, 2(1), 189-198.
- Verellen, F. (1995). Taoism. *The Journal of Asian Studies*, 54(2), 322-346.
- Waddams, S. M. (2003). *Dimensions of private law: categories and concepts in Anglo-American legal reasoning*. Cambridge University Press.
- Wade, W., & Forsyth, C. F. (2014). *Administrative law*. Oxford University Press, USA.
- Waldron, J. (2002). The rule of law in contemporary liberal theory. *GW Smith (Hg.): Liberalism. Critical Concepts in Political Theory, Bd, 2(3)*, 93-111.
- Wallach, J. J. (2013). *How America eats: A social history of US food and culture*. Rowman & Littlefield.
- Wamba, S. F., Bawack, R. E., Guthrie, C., Queiroz, M. M., & Carillo, K. D. A. (2021). Are we preparing for a good AI society? A bibliometric review and research agenda.

- Technological Forecasting and Social Change*, 164, 120482.
- Wang, J. (2015). Managing social stability: The perspective of a local government in China. *Journal of East Asian Studies*, 15(1), 1-25.
- Wang, L., Chen, X., Deng, X., Wen, H., You, M., Liu, W., Li, Q., & Li, J. (2024). Prompt engineering in consistency and reliability with the evidence-based guideline for LLMs. *NPJ digital medicine*, 7(1), Article 41. <https://doi.org/10.1038/s41746-024-01029-4>
- Wang, X., Duan, S., Yi, X., Yao, J., Zhou, S., Wei, Z., Zhang, P., Xu, D., Sun, M., & Xie, X. (2024). On the essence and prospect: An investigation of alignment approaches for big models. *arXiv preprint arXiv:2403.04204*.
- Wang, X., Zhang, K., & Zhang, P. (2024). Looking at large model safety and practice from multiple perspectives(in Chinese). *Computer research and development*, 61(05), 1104-1112. <https://link.cnki.net/urlid/11.1777.TP.20240312.1713.006>
- Wang, X., & Zhang, X. (2022). The Connotation, Ideological Origin and Logical Premise of Common Prosperity. *Open Access Library Journal*, 9(5), 1-11.
- Watson, D. K. (1910). *The Constitution of the United States* (Vol. 2). Callaghan.
- Webb, T., Holyoak, K. J., & Lu, H. (2023). Emergent analogical reasoning in large language models. *Nature Human Behaviour*, 7(9), 1526-1541. <https://doi.org/10.1038/s41562-023-01659-w>
- Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S., Yogatama, D., Bosma, M., Zhou, D., & Metzler, D. (2022). Emergent abilities of large language models. *arXiv preprint arXiv:2206.07682*.
- Williams, J. A. (2020). *Islam* (Vol. 1). Library of Alexandria.
- Woiseschlaeger, H., Erben, A., Wang, S., Mayer, R., Jacobsen, H.-A., & Assoc Computing, M. (2024, 2024
- Jun09). Federated Fine-Tuning of LLMs on the Very Edge: The Good, the Bad, the Ugly. [Proceedings of the 8th workshop on data management for end-to-end machine learning, deem 2024]. 8th Workshop on Data Management for End-to-End Machine Learning (DEEM), Santiago, CHILE.
- Wolfe, C. T. (2016). *Materialism: A historico-philosophical introduction*. Springer.
- Wright, M., Wood, G., Musacchio, A., Okhmatovskiy, I., Grosman, A., & Doh, J. P. (2021). State capitalism in international context: Varieties and variations. *Journal of World Business*, 56(2), 101160.
- Wu, J. (2019). *New aesthetic thought, methodology, and structure of systemic philosophy*. IGI Global.
- Wu, L. (2023). Translation Processing Algorithm of Japanese Complex Long Sentences Based on Data Mining Technology. 2023 International Conference on Telecommunications, Electronics and Informatics (ICTEI),
- Wu, S., Irsoy, O., Lu, S., Dabrovolski, V., Dredze, M., Gehrmann, S., Kambadur, P., Rosenberg, D., & Mann, G. (2023). Bloomberggpt: A large language model for finance. *arXiv preprint arXiv:2303.17564*.
- Xu, R., Luo, F., Zhang, Z., Tan, C., Chang, B., Huang, S., Huang, F., & Assoc Computat, L. (2021, 2021

- Nov 07-11). Raise a Child in Large Language Model: Towards Effective and Generalizable Fine-tuning. [2021 conference on empirical methods in natural language processing (emnlp 2021)]. Conference on Empirical Methods in Natural Language Processing (EMNLP), Punta Cana, DOMINICAN REP.
- y Arcas, B. A. (2022). Do large language models understand us? *Daedalus*, 151(2), 183-197.
- Yang, J., Jin, H., Tang, R., Han, X., Feng, Q., Jiang, H., Zhong, S., Yin, B., & Hu, X. (2024). Harnessing the Power of LLMs in Practice: A Survey on ChatGPT and Beyond. *Acm Transactions on Knowledge Discovery from Data*, 18(6), Article 160. <https://doi.org/10.1145/3649506>
- Yao, J., Yi, X., Wang, X., Wang, J., & Xie, X. (2023). From Instructions to Intrinsic Human Values--A Survey of Alignment Goals for Big Models. *arXiv preprint arXiv:2308.12014*.
- Yao, Y., Wang, P., Tian, B., Cheng, S., Li, Z., Deng, S., Chen, H., & Zhang, N. (2023). Editing large language models: Problems, methods, and opportunities. *arXiv preprint arXiv:2305.13172*.
- Yao, J., Yi, X., Duan, S., Wang, J., Bai, Y., Huang, M., Zhang, P., Lu, T., Dou, Z., Sun, M., & Xie, X. (2025). *Value Compass Benchmarks: A Platform for Fundamental and Validated Evaluation of LLMs Values* (No. arXiv:2501.07071). arXiv. <https://doi.org/10.48550/arXiv.2501.07071>
- Yi, X., & Xie, X. (2023). Analysis of the Big Model Moral Values Alignment Problem (in Chinese). *Computer research and development*, 60(09), 1926-1945. <https://link.cnki.net/urlid/11.1777.TP.20230905.1313>
- Yilahun, H., Nurtay, A., & Hamdulla, A. (2017). Patterns of vowels in Uyghur Tri-syllabic words. 2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC),
- Young, S. (2003). *Moral capitalism: Reconciling private interest with the public good*. Berrett-Koehler Publishers.
- Yue, S., Chen, W., Wang, S., Li, B., Shen, C., Liu, S., Zhou, Y., Xiao, Y., Yun, S., & Lin, W. (2023). Disc-lawllm: Fine-tuning large language models for intelligent legal services. *arXiv preprint arXiv:2309.11325*.
- Zangrossi, P., Martini, M., Guerrini, F., De Bonis, P., & Spena, G. (2024). Large language model, AI and scientific research: why ChatGPT is only the beginning [Review; Early Access]. *Journal of Neurosurgical Sciences*. <https://doi.org/10.23736/s0390-5616.23.06171-4>
- Zekoll, J., & Wagner, G. (2018). *Introduction to German law*. Kluwer Law International BV.
- Zhang, J., Bao, K., Zhang, Y., Wang, W., Feng, F., He, X., & Acm. (2023, 2023
- Sep 18-22). Is ChatGPT Fair for Recommendation? Evaluating Fairness in Large Language Model Recommendation. [Proceedings of the 17th acm conference on recommender systems, recsys 2023]. 17th ACM Conference on Recommender Systems (RecSys), Singapore, SINGAPORE.
- Zhang, L., Li, C., Hu, Q., Lang, J., Huang, S., Hu, L., Leng, J., Chen, Q., & Lv, C. (2023). Enhancing Privacy in Large Language Model with Homomorphic Encryption and

- Sparse Attention. *Applied Sciences-Basel*, 13(24), Article 13146. <https://doi.org/10.3390/app132413146>
- Zhang, P., & Kamel Boulos, M. N. (2023). Generative AI in medicine and healthcare: Promises, opportunities and challenges. *Future Internet*, 15(9), 286.
- Zhang, Y., He, Y., Xia, Y., Wang, Y., Dong, X., & Yao, J. (2024). Exploring the representation of Chinese cultural symbols dissemination in the era of large language models. *International Communication of Chinese Culture*, 1-23.
- Zhang, Z., Zhang, C., Liu, N., Qi, S., Rong, Z., Zhu, S.-C., Cui, S., & Yang, Y. (2024). Heterogeneous Value Alignment Evaluation for Large Language Models. AAAI-2024 Workshop on Public Sector LLMs: Algorithmic and Sociotechnical Design,
- Zhao, S., Dang, J., & Grover, A. (2023). Group preference optimization: Few-shot alignment of large language models. *arXiv preprint arXiv:2310.11523*.
- Zheng, R., Dou, S., Gao, S., Hua, Y., Shen, W., Wang, B., Liu, Y., Jin, S., Liu, Q., & Zhou, Y. (2023). Secrets of rlhf in large language models part i: Ppo. *arXiv preprint arXiv:2307.04964*.
- Zhu, J., Dang, P., Cao, Y., Lai, J., Guo, Y., Wang, P., & Li, W. (2024). A flood knowledge-constrained large language model interactable with GIS: enhancing public risk perception of floods. *International Journal of Geographical Information Science*, 38(4), 603-625. <https://doi.org/10.1080/13658816.2024.2306167>
- Zhu, S., Pan, L., & Xiong, D. (2024). FEDS-ICL: Enhancing translation ability and efficiency of large language model by optimizing demonstration selection. *Information Processing & Management*, 61(5), Article 103825. <https://doi.org/10.1016/j.ipm.2024.103825>
- Ziarek, K. (1994). *Inflected language: toward a hermeneutics of nearness: Heidegger, Levinas, Stevens, Celan*. State University of New York Press.
- Zuhur, S. (1998). *Images of enchantment: Visual and performing arts of the Middle East*. American Univ in Cairo Press.

Appendix A: Interview outline

1. Medical staff

Hello, thank you very much for taking the time out of your busy work to participate in this interview. I am —, and I am currently conducting a survey on the governance of big models in the medical industry. The purpose is to provide suggestions for big model governance by understanding the cognition and application of big models by relevant personnel in the medical industry.

This interview will take about 20-30 minutes, and will mainly focus on the application of big models and governance suggestions in medical institutions. Please rest assured that the content of the interview is only used for graduation thesis writing and will not be made public. We will strictly keep your personal information and answers confidential. If you feel inconvenient to answer any questions, you can skip or terminate the conversation at any time.

1. Basic information:

Age, education

"Can you please briefly introduce your position and years of experience?"

"Which department/field do you currently work in? What does your daily responsibilities include?"

2. Experience in using big models for assisted diagnosis

1. When using auxiliary diagnostic tools, have you noticed that the model has certain tendencies in describing diseases or suggesting treatment plans?

2. Do you feel that the model is too "authoritative" or "certain" in expression, which may affect your independent judgment?

3. In your opinion, should auxiliary diagnostic tools be completely subject to the doctor's judgment? Is there a trend of "technology taking precedence over people" that makes you uncomfortable?

4. Are you clear about the data sources used by the model and the patient privacy protection mechanism? Does this affect your confidence in using it?

5. If the recommendations of the big model are inconsistent with clinical judgment, who will be responsible when a misdiagnosis occurs? Is the division of relevant responsibilities clear?

6. Does the hospital provide training or guidelines for the compliant use of big models? Do you think the current system is sufficient to protect both doctors and patients?

7. Can the model adapt to your professional habits and language environment in terms of language expression or case description?

8. When serving rural, aging or minority patients, can the model take into account their communication habits and cultural characteristics?

9. Do you want the model to be "localized" based on local case data? Have there been related attempts in reality?

10. Does the model show sufficient prudence when dealing with high-risk and sensitive patients (such as cancer, the elderly, and pregnant women)?

11. Have you ever encountered a situation where the model output is "mechanical", "indifferent" or makes patients feel uneasy? Can you give an example? Consciously face the big model

12. Are you worried that doctors will gradually rely on the judgment of the big model and weaken their clinical thinking ability? How to maintain human-machine collaboration instead of being replaced?

13. If the model has an erroneous output, does the hospital encourage feedback? Have you given feedback? Is the feedback effective?

14. In your opinion, what aspects of the current big model tools need to be improved most (expression, humanization, controllability, ethics, safety, etc.)? Better interaction with medical staff and automatic medical data.

15. As a doctor, what suggestions do you have for the government or technology to achieve better governance of big models?

16. Overall, how do you feel and how satisfied are you with using big models? What are your concerns?

2. LLM Technology Developer

Thank you for participating in this research interview. This study focuses on the governance of large language models in medical scenarios, especially how the value alignment mechanism is implemented in the model development, training, testing and deployment stages.

Basic information: name, age, education, occupation, professional field, experience in participating in large model development.

Specific questions:

1. In the development of AI models, have you considered that the output content may involve ideological issues, such as political sensitivity, cultural tendencies, etc.? Is there a set filtering mechanism?

2. Have you established a content security strategy during development to prevent the model from spreading statements that may cause misunderstandings or social risks in medical advice?

3. Is there a design for human-machine collaborative decision-making during the development process? How to prevent the model from outputting content that is highly guiding or replaces the doctor's judgment in medical decision-making?

4. In the process of developing the model, do you conduct compliance control on data collection and processing in accordance with the current "Data Security Law" and "Personal Information Protection Law"? Can you briefly describe the process?

5. In the process of patient data processing, desensitization and retraining, is there an automated compliance review system or responsibility attribution mechanism?

6. After the model is deployed, if a decision error or misleading doctor behavior occurs, are there accountability paths and explainable mechanisms? Is the model behavior "traceable"?
7. Does your model support localized adaptation to the context of different regions? For example, dialects, ethnic languages, local clinical terms, etc.?
8. Do you consciously incorporate diverse medical data sets (such as remote areas, ethnic hospitals) during model pre-training/fine-tuning?
9. Does your team plan to develop a "culturally sensitive" model?
10. Are ethical red lines set in model development? For example, discrimination, deprivation of patient dignity, and suggestions for excessive medical treatment must not be generated?
11. Has an ethical testing mechanism been established? Including: bias audit, misleading suggestion testing, simulation of vulnerable populations, etc.?
12. Do you have a dedicated ethics review team or advisory committee? Do you participate in the "ethical risk assessment" before the model goes online?
13. In the process of commercial promotion of the model, how to prevent market orientation from compressing ethical space? Are there relevant internal governance regulations?
14. From a technical perspective, in which areas are you most concerned about models getting out of control or being misused?
15. In which areas do you hope that the government or industry associations will provide clearer governance frameworks or value alignment guidelines?

3. Hospital administrator

Hello, thank you very much for taking the time out of your busy work to participate in this interview. I am —, and I am currently conducting a survey on the governance of big models in the medical industry. The purpose is to provide suggestions for big model governance by understanding the cognition and application of big models by relevant personnel in the medical industry. "

"This interview will take about 20-30 minutes, mainly focusing on the application of big models and governance suggestions in medical institutions. Please rest assured that the content of the interview will only be used for graduation thesis writing and will not be made public. We will strictly keep your personal information and answers confidential. If you feel inconvenient to answer any questions, you can skip or terminate the conversation at any time. "

3. Confirm consent and start the topic

"Before we start, do you agree to participate in this interview and allow us to record the conversation? Thank you again for your support! Then we officially start. "

Identity information inquiry

I. Basic information

1. What is your age?

"Can you please briefly introduce your position, work unit and years of experience? "

"Which department/field do you currently work in? What does your daily responsibilities include? "

2. Professional background

"Have you ever received special training or participated in projects related to this topic? "

3. Anonymization

"To protect privacy, your name/unit may be anonymized in the report. Do you agree? "

The following are questions tailored for hospital administrators about medical artificial intelligence interviews:

II. Hospital experience in introducing and managing big models

What decision-making process did your hospital go through when introducing big models to assist medical work? What factors did the hospital management mainly consider?

At the hospital level, are there any special management policies or regulations for the use of big models? How do these policies ensure that big models serve the overall medical service goals of the hospital?

How do you coordinate the differentiated needs of different departments for the use of big models? Are there any difficulties in resource allocation?

From the perspective of hospital management, what specific impacts does the use of big models have on the hospital's operational efficiency (such as patient turnover speed, medical record management, etc.)?

Has the hospital conducted a benefit evaluation on the cost of using big models? Do you think the current input-output ratio is reasonable?

III. Issues related to the integration of big models with the hospital's existing systems

Big models and the hospital's existing information systems (such as HIS, LIS How is the

compatibility with other big models? Have you encountered problems with data docking or poor information flow?

In the process of promoting the integration of big models with hospital workflows, what do you think is the biggest obstacle? What measures have the hospital taken to overcome these obstacles?

How can the hospital ensure that the use of big models will not interfere with the traditional medical quality control system? Has a new quality monitoring mechanism been established specifically for medical behaviors assisted by big models?

Does the hospital have a corresponding supervision and assessment mechanism for the use of big models by medical staff? How do these mechanisms motivate or constrain medical staff?

Do you think that the use of big models has changed the power structure or work responsibility distribution within the hospital to a certain extent? If so, in what specific aspects?

4. Views and opinions on the prospects for the application of big models in hospitals
Suggestions

How do you envision the role of big models in the future development of hospitals? Will they exist more as tools, or will they become an important reference for hospital decision-making?

From a strategic perspective, how do you think hospitals should plan the continuous upgrade and optimization of big models to ensure that they always fit the hospital's development strategy?

In exchanges and cooperation with other hospitals or medical institutions, has the use of big models become a new competitive factor or cooperation opportunity? What plans or actions does your hospital have in this regard?

What regulatory role do you think the government should play in the application of big models in hospitals? How should hospitals themselves cooperate with the government's relevant policy orientation?

For technology suppliers, in which aspects of big models do you hope they can improve or strengthen to better meet the needs of hospital management?

4. Policy makers in the medical sector

Thank you for accepting this interview. This study focuses on the government's governance mechanism and value alignment practices for large language models, especially in key scenarios such as medical image-assisted diagnosis. We want to understand the government's role and challenges in policy making, supervision and coordination through our interview with you.

1. In promoting the participation of large models in medical services, does the government consciously pay attention to whether their output content conforms to the country's mainstream ideology?

2. Do you think that AI in medical scenarios may inadvertently output information that violates the core values of society? Is there a prevention mechanism?

3. Will the government set a value orientation in model deployment or promotion, such as emphasizing human-machine collaboration rather than "human substitution"?

1. Does the government have a review mechanism for the compliance of medical AI models? For example, it involves the "Data Security Law" and "Personal Information Protection Law"?

2. In terms of model processing of patient data and generating diagnostic recommendations, are there any problems that are difficult to cover under current regulations?

3. Does the government have clear policies or guidance on the attribution and accountability of models in clinical applications?

4. Will medical AI be included in the regulatory framework similar to "medical device registration" in the future?

1. Do you think the existing policies take into account the adaptation needs of big models for different regions, languages, and medical levels?

2. Are there any support measures to promote technology companies to "localize" training or fine-tune the model?

3. For doctors and patients in remote and ethnic minority areas, does the government have any ideas to promote their fair access and use of big model tools?

1. Does the government have a medical AI ethics assessment mechanism? Does it include multidisciplinary (medical, ethical, legal) experts?

2. When it comes to vulnerable groups such as severe illness, the elderly, and minors, have ethical boundaries been established that model outputs should follow?

3. How does the government evaluate the "guidance" of the model on doctors' decision-making? Are you worried that it will replace doctors' thinking and lead to professional alienation?

4. How to deal with the tension between efficiency priority and ethical protection? Are there any balancing principles or standards?

5. Patients

Thank you for participating in our interview. We are conducting a study on the application of

artificial intelligence, especially "big language models" in medical services, focusing on whether these new technologies truly respect and care for patients' rights, culture, values and user experience. Please rest assured that your answers will be strictly confidential and will only be used for academic research.

1. Do you know that the hospital uses big models or intelligent systems to assist doctors in diagnosis or provide information?
2. Are you worried that these systems will output indifferent or extreme information, which is inconsistent with the mainstream values or humanistic spirit of our society?
3. Do you hope that big models can express more respect, sympathy and care for patients in medical treatment, rather than just speaking conclusions like machines?
4. When using intelligent systems for diagnosis or advice, have doctors explained this technology, its uses or risks to you?
5. Do you know that your medical records, images and other data may be used for training or optimization of big model systems? Have you ever obtained your authorization or informed consent?
6. If the big model makes an incorrect judgment that leads to delayed or improper treatment, do you know who should bear the responsibility? Is there a complaint channel?
7. During your medical treatment, have you ever felt that the language expression of the big model system is difficult for you to understand, unfamiliar, or not close enough to your lifestyle?
8. Do you hope that the big model system can better understand your language, regional culture, living background or religious beliefs when providing services?
9. If you are from a rural area, an elderly group or a minority group, are you worried that these technologies do not take into account your special needs or communication methods?
10. Are you worried that the big model system's misjudgment will cause you to be over-medicalized, misdiagnosed or neglected? Will this affect your trust in doctors?
11. When it comes to major diseases, do you want the doctor to tell you the condition directly instead of relying entirely on the big model results?
12. Do you hope that the medical big model will reflect more humanistic care in its design, such as relieving anxiety and encouraging people to face the condition?
13. In your opinion, which is more important, "efficiency" or "respecting patients' emotions and choices"? How to balance the two?

14. Do you support or worry about the use of such smart technologies in hospitals? Why?

15. If you could give advice to hospitals or the companies developing these technologies, what would you most like to see improved?

Appendix B: Coding Nodes

Selective coding	Axis coding	Open coding	Number of sources	Number of reference points
Technical algorithms			29	143
	Data accuracy		26	43
		Data sources	19	19
		Data training	8	8
		Data scope	5	5
		Data quality	6	6
		Data updates	5	5
	Algorithm optimization		24	49
		Decision-making process traceability	10	10
		Transparency of data processing	4	4
		Identification of rare diseases and miscellaneous cases	12	12
		Cross-cultural scenario adaptability	20	20
		Algorithm universality	3	3
	Supporting system		20	41
		Legality review of data collection	3	3

		Data classification and grading management	11	12
		Principle of informed consent for data use	9	9
		Audit of cross-border data transmission	3	3
		Supervision of data labeling in private institutions	1	1
		Application specifications for low-probability data	1	1
		Application of data privacy technology	6	6
		Multi-department collaborative data supervision	1	1
		Data dynamic risk assessment	2	2
		Anonymization technology standards	3	3
	Technology updates		8	10
		Algorithm iteration speed	7	7
		Dynamic monitoring mechanism	3	3
Clinical adaptation			32	205
	Human-machine collaboration mechanism		30	60
		Clarify the status of AI assistance	30	30
		Guide human-machine collaboration through policies and regulations	2	2

		Promote human-machine collaboration from multiple aspects	4	4
		Ensure human-machine collaboration through training and mechanism building	3	3
		Human-machine division of labor and collaboration in actual work	7	8
		Combination of multidisciplinary consultation (MDT) with big model suggestions	3	3
		Combination of three-level review of medical team with big model suggestions	1	1
		Big model-assisted doctor-patient communication mechanism	7	7
		Big model-assisted learning mechanism for young doctors	2	2
	Adaptation for special populations		29	95
		Pay attention to language differences in different regions	11	11
		Consider differences in medical levels in different regions	5	5
		Promote localized training or fine-tuning of models	22	22
		Promote fairness in medical services in remote ethnic minority areas	7	7

		Pay attention to the communication habits and cultural characteristics of special groups	18	18
		Promote the expansion and sinking of high-quality medical resources	3	3
		Increase the "supply" of grassroots professionals through multiple channels	26	26
		Bring advanced health concepts to the grassroots	3	3
	Clinical operation specifications		20	25
		Operation specifications	4	4
		Training content	1	1
		Supervision and evaluation mechanism	20	20
	Patient experience and communication		17	25
		AI language temperature	14	14
		Humanistic care	11	11
Clarify responsibilities			19	69
	Responsibility and regulatory governance		9	27
		Clarify the responsible parties and division principles	6	7
		Establish a regulatory review and supervision mechanism	5	5
		Improve regulations to adapt to the development of medical AI	9	10

		Promote the inclusion of medical AI in the regulatory framework	5	5
	Responsibility division principles		11	15
		Ladder-type responsibility division	1	1
		Responsibility division based on product defects	10	10
		Responsibility division based on operational errors	2	2
		Responsibility division based on doctor review obligations	2	2
	Government main responsibility		1	3
		Policy formulation and guidance	1	1
		Review and supervision	1	1
		Promote fairness and coordination	1	1
	Medical institution main responsibility		2	2
		Management and supervision	1	1
		Training and norms	1	1
	Doctor main responsibility		12	15
		Professional judgment and review	5	5
		Avoid over-reliance	10	10
	Technology enterprise main responsibility		4	4
		Data and algorithm responsibility	2	2

		Continuous optimization and feedback	2	2
	Responsibilities in different scenarios		2	3
		Responsibility standards for rare disease diagnosis	1	1
		Responsibility standards for primary care	1	1
		Responsibility for collaboration between medical AI and doctors	1	1
Policy governance and supervision			26	81
	Ideological review		5	9
		Ensure that medical AI outputs conform to mainstream ideology	5	5
		Monitor and correct the ideological tendency of the model	1	1
		Ideological control of model training data	1	1
		Review mechanism and multi-departmental collaboration	2	2
	Ethical review mechanism		2	4
		Establish a multi-level ethical review system	2	2
		Review content covers many aspects	1	1
		Review dimensions and standards	1	1
	Issues of ethical review		25	49

		Data privacy and security	11	11
		Algorithmic fairness and transparency	4	4
		Ethical considerations for special groups	18	18
		Cultural suitability review	5	5
		The dynamics of ethical review	11	11
	Compliance review mechanism		10	19
		Multi-faceted review of medical AI models	1	1
		Ethical review and industry compliance	1	1
		Legality of data collection and use	5	6
		Data compliance review	6	6
		Algorithm compliance review	3	3
		Multi-departmental collaborative regulatory review	2	2
Data security and privacy protection			25	100
	Data security measures		16	38
		Data collection and use specifications	9	10
		Data security assurance mechanism	10	10
		Data quality improvement measures	2	2

		Data compliance review mechanism	5	5
		Data storage security supervision	8	8
		Multi-department collaboration to ensure data security	3	3
	Data privacy protection mechanism		23	30
		Institutional requirements for patient privacy protection	15	15
		Data desensitization and encryption measures	9	9
		Ethical review of privacy protection	6	6
	Challenges of data security and privacy protection		10	14
		Data security cognition and regulatory challenges	1	1
		Conflicts between data sharing and privacy protection	9	9
		Data sharing privacy and copyright barriers	4	4
	Data storage and transmission security		5	5
		Storage location and audit supervision	4	4
		Requirements for data encryption measures	1	1
	Compliance with data use		10	13

		Data use follows regulatory principles	2	2
		Data classification management and review	2	2
		Establishment of infringement accountability mechanism	9	9
Improvement suggestions			17	54
	Improve policies and regulations		12	19
		Clarify the legal status and division of responsibilities related to AI medical care	3	3
		Strengthen supervision of medical AI research and development and application	10	10
		Promote data sharing and protect patient privacy	6	6
	Optimize technical aspects		12	21
		Improve algorithm transparency and explainability	5	5
		Enhance model accuracy and stability	1	1
		Achieve model localization optimization and data utilization	2	2
		Improve the level of humanization	4	4
		Optimize language expression and case description	2	2

		Strengthen data security and privacy protection	2	2
		Adapt to the needs of different cultural groups and special patients	1	1
		Improve data quality	1	1
		Improve feedback mechanism	3	3
	Personnel training and education		7	8
		Carry out AI technology training for doctors	5	5
		Improve patients' understanding of AI	2	2
		Encourage universities to cultivate compound talents	1	1
	Clinical application standards		2	2
		Standardize human-machine collaboration processes	1	1
		Establish a quality control system for AI-assisted diagnosis	1	1
	Financial support		4	4
		The government will increase funding to support research and development	4	4