Abstract

Reclaiming the Value of Deep Thinking

2026 Blue Book on the Development of AI for Social Sciences and Humanities

Editors-in-Chief

CHEN Zhimin

WU Libo

FUDAN UNIVERSITY 1905 復旦大學

Scan the QR code to access the full content of this book.

Preface:
Guiding the Transformation of Intelligence with the Depth of Thought

In his congratulatory letter marking the 120th anniversary of Fudan University, General Secretary Xi Jinping emphasized the need to advance knowledge innovation, theoretical innovation, and methodological innovation in philosophy and the social sciences. This important guidance points the way forward for the development of the New Liberal Arts in universities, and provides a fundamental compass for our exploration of the integration of intelligence with the social sciences and humanities.

The *Blue Book on the Development of AI for Social Sciences and Humanities 2026* (hereinafter the *Blue Book*) takes "knowledge innovation, theoretical innovation, and methodological innovation" as its overarching framework, and "Reclaiming the Value of Thinking" as its theme. It offers a profound analysis of the interactive tensions between the social sciences and humanities and artificial intelligence, vividly presents the latest explorations of Chinese universities in the paradigm transformation of intelligent social sciences and humanities, and provides important insights for promoting the high-quality development of philosophy and the social sciences by upholding fundamental principles while breaking new ground.

This is a work that responds to the questions of our time and is grounded in a deep concern for reality. As one turns its pages, the image of scholars in the social sciences and humanities—thinking deeply and working earnestly—comes vividly into view. They remain focused on frontier practice, directing their attention to major issues of human concern, such as national development, sustainable transformation, public governance, and social trust. They raise real questions, build real theories, and create real value, enabling research in philosophy and the social sciences to retain its penetrating capacity to answer the questions of the times and resolve the puzzles of development in the intelligent era.

This is also a work that moves beyond instrumental rationality and advocates human–AI symbiosis. A defining feature of the 2026 *Blue Book* is that it advances "AI

empowerment" toward "two-way empowerment and symbiosis." It discusses not only how technology drives the iteration of knowledge production, but also how the academic community can guide technology toward the good and enhance the value of technological empowerment. This mutually reinforcing and mutually enriching mode of integration unleashes technological potential while upholding human-centered boundaries. It also allows us to see the boundless possibilities of intelligence supporting the humanities and thought guiding innovation.

This is, above all, a work that returns to the essence of thought and calls for deep thinking. The book presents a group of high-value research achievements in the social sciences and humanities. What they rely on is not "faster generation," but "deeper cognition." Every argument has been questioned, refined, and tempered through long cycles, complete chains of inquiry, and strong contextual engagement. What ultimately emerges is a set of intellectual achievements that reach the essence of things, reveal underlying patterns, and illuminate the logic of reality.

Today, the wave of intelligence is surging forward. It has brought new tools, new scenarios, and new methods to philosophy and the social sciences, while also placing higher demands on the independence, depth, and creativity of thought. I believe that the publication of the new edition of the *Blue Book* will not only provide an important reference for the academic community, encouraging more scholars to devote themselves to the integration and innovation of the social sciences and humanities with artificial intelligence, and to continuously expand the breadth and depth with which we know and understand the world. It will also inspire more readers to safeguard thought and sharpen thinking in the intelligent era, preserving the composure and resolve needed for independent reflection, rational judgment, value inquiry, and discerning choice, so that the light of thought and the wings of intelligence may coexist, reinforce each other, and shine together.

Since its first release in 2025, the *Blue Book* has played an important role in advancing the paradigm transformation of intelligent social sciences and humanities in Chinese universities and in promoting the development of the New Liberal Arts. I hope that the *Blue Book* team will continue to work deeply on the core theme of integrating the social sciences and humanities with intelligence; serve as steadfast guardians of

deep thinking, pathfinders of intelligent empowerment, and pioneers in the development of the New Liberal Arts; and continue to produce achievements that combine intellectual depth with practical value. In doing so, the team will contribute greater wisdom to the construction of China's independent knowledge system in philosophy and the social sciences, and to the advancement of Chinese modernization.

I offer these words as a foreword.

Qiu Xin

Party Secretary of Fudan University

May 2026

Table of Contents

Abstract

Breakthroughs in generative large models, multimodal learning, and agent technologies are accelerating the evolution of artificial intelligence from a set of discrete tools into a general-purpose capability that can be embedded in research, education, industry, and governance workflows. At the same time, public and academic debates around "replacement or empowerment," "efficiency or responsibility," and "generation or thinking" continue to intensify. On the one hand, AI has significantly lowered the formal barriers to writing, programming, retrieval, annotation, modeling, and analysis, releasing unprecedented efficiency gains. On the other hand, capability diffusion, deep deployment, and competitive incentives are reshaping modes of knowledge production, organizational collaboration, public trust, and social governance. In response to this irreversible process, this Blue Book takes "Reclaiming the Value of Thinking" as its theme. It emphasizes the need to free human creativity, judgment, and value discernment from repetitive formal labor, and, through an integrated development–risk–governance lens, asks how AI can be embedded in the social sciences and humanities and in broader social operations in a healthy, robust, and safe manner. It discusses not only how AI empowers knowledge production, but also how the social sciences and humanities can provide direction, boundaries, and institutional engineering support for AI.

This edition continues the main thread of "empowerment" and "symbiosis." It builds an integrated narrative around the loop of "data–theory–mechanism–governance," seeking to turn fragmented technological advances into reusable methodological capabilities, and to elevate scattered university experiments into transferable organizational experience and institutional solutions. The book consists of three parts and eleven chapters. Part I focuses on the integration and reshaping of AI and the social sciences and humanities, emphasizing how intelligence as a public capability is acquired, distributed, used, and governed. Part II examines frontiers and trends in key domains of intelligent social sciences and humanities, systematically presenting paradigm shifts in complex systems and social simulation, social-science causal inference, computational social science, digital-intelligence humanities, and

public governance. Part III focuses on the exploration, action, and institutional development of AI4SSH in Chinese universities. Through index-based assessment, infrastructure review, and typical case analysis, it traces the staged pathway through which AI4SSH in Chinese universities is moving from project-based exploration toward platform building, ecosystem coordination, and institutionalized development.

What this Blue Book addresses is not only a leap in technical capability, but also a broader transformation in the conditions of knowledge production and social governance. As high-quality text generation, complex code implementation, image and video understanding, multimodal reasoning, and agent-based collaboration become callable capabilities, the bottleneck of research is shifting from "whether materials can be processed" to "whether good questions can be posed, real mechanisms constructed, and verifiable evidence chains formed." As simulation, optimization, and agents begin to participate in policy reasoning, organizational management, and public services, the governance challenge is also shifting from "whether AI should be used" to "how AI can be used responsibly under value conflicts, uncertainty, and accountability constraints." This is especially true in major issues such as sustainable development, climate change, public governance, cultural inheritance, and social trust. These issues are often cross-scale, cross-system, highly contextual, and deeply uncertain. They require not only the coupling of multimodal data and models, but also the incorporation of distributional effects, equity trade-offs, institutional constraints, and value judgments into a shared reasoning structure. In this sense, AI4SSH is not a marginal technical application. It is a reconstruction of knowledge and governance capabilities for addressing major societal challenges.

Part I addresses, at the macro level, the question of how intelligence can become a public capability. Chapter 1 examines the coordination between AI development and governance, arguing that AI's profound social impact lies not only in productivity gains, but also in the scalable replication of capacities for decision-making, prediction, persuasion, and organization, thereby rewriting the underlying parameters of social operation. The key to the healthy development of AI is therefore not simply whether models become "smarter," but how intelligence is accessed, distributed, used, and constrained. Chapter 2 explores pathways toward the integration and symbiosis of AI

and the social sciences and humanities in the age of AGI. Focusing on complex-system coupling models, high-throughput simulation of socioeconomic systems, causal inference, and large-model agent workflows, it distills a set of transferable methodological mechanisms: interdisciplinary interfaces shift from manual assembly to learnable alignment; policies, behaviors, and risks are stress-tested across scenario spaces; the evidentiary chain from correlation to causality and action becomes proceduralized and auditable; and research tasks are organized into workflows that support division of labor, verification, and accountability. Chapter 3 turns to the reciprocal shaping of AI and the philosophy of cognitive science. It traces the evolution from symbolicism and connectionism to 4E cognition and predictive processing, identifies perception and attention, memory and cognitive maps, emotion and value, and cognitive control as foundational mechanisms of human intelligence, and examines how cognitive AI may move from object recognition to subject understanding, from content generation to meaning construction, and from labor substitution to cognitive collaboration. At the same time, AI provides new model-based, experimental, and comparative settings for revisiting classical philosophical questions concerning understanding, intentionality, embodiment, causality, and consciousness. Chapter 4 further examines the systemic reconstruction of AI ethics and governance. It analyzes the technological nature and conceptual ambiguities of AGI, reconsiders the place of human beings, maps emerging risk spectra, investigates automated decision-making and the rebuilding of trust, and proposes a four-dimensional governance framework for ethical engineering and institutional design.

Part II turns to key methods and governance issues, presenting the shift of intelligent social sciences and humanities from “explaining the world” toward “simulating processes, evaluating interventions, and supporting decisions.” Chapter 5 centers on complex systems and social simulation. It argues that, amid the deep convergence of digitization, networking, and intelligent systems, social systems increasingly display the features of complex adaptive systems. Large-model-driven multi-agent social simulation can model individual behavior and group interaction in controllable digital environments, translating complex social processes into computable, repeatable, and intervenable “digital experimental arenas.” This provides experimental

support for counterfactual reasoning, strategy optimization, and policy evaluation. Chapter 6 focuses on social-science causal inference empowered by deep learning. It notes that causal identification theory is relatively mature, but model dependence at the estimation stage remains a practical bottleneck. Through representation learning, sample balancing, and generative distribution modeling, deep learning can strengthen counterfactual prediction and mechanism analysis. Under maintained identification assumptions, it can improve the flexibility, robustness, and reproducibility of causal estimation. Chapter 7 follows the theme of expansion "from text to multimodality" and discusses how large language models and multimodal large models reshape computational social science. On the one hand, they reduce the cost of topic modeling, classification, information extraction, and generative analysis, transforming unstructured materials into computable evidence. On the other hand, they bring multimodal streams of facts—images, video, speech, trajectories, and other materials—into the scope of research, expanding the observational range, experimental space, and simulation capacity of computational social science. Chapter 8 approaches the issue from the perspective of digital-intelligence humanities. It discusses the paradigm shift from "digital humanities" to "digital-intelligence humanities," analyzes the impact of generative AI on narrative, cultural expression, authenticity, and authorship, and stresses that the humanities must continue to perform their core functions of critical reconstruction, value discernment, and meaning-making in the intelligent era. Chapter 9 focuses on the modes and responsibilities of AI embedded in public governance. It reviews the application scenarios and potential risks of AI in government services, distinguishes between "agentic" and "assistive" modes of embedding, and proposes a governance framework for the responsible use of AI from the perspectives of rebuilding responsibility chains, duties of explanation, audit and oversight, grievance and remedy mechanisms, and maintenance of public trust.

Part III shifts the focus to the systematic actions of Chinese universities in AI4SSH. Chapter 10 constructs the "China University AI4SSH Index." From the three dimensions of research core capability, development and innovation potential, and social communication capability, it systematically observes the overall progress and structural characteristics of Chinese universities in the deep integration of AI with the

social sciences and humanities. The index assessment shows that AI4SSH development in Chinese universities has begun to display a pattern of "initial system formation and clear tiering." Research output and domestic integration are advancing rapidly, and several high-level research universities have formed relatively mature interdisciplinary research systems. Yet shortcomings remain in international academic influence, original innovation capability, institutional support, and the transformation of research into social services. Development intensity and capability structure also differ significantly across universities and regions. Chapter 11 focuses on AI4SSH infrastructure and representative project development. Around data foundations and circulation mechanisms, compute integration and experimental environments, simulation platforms, models/agents, and toolchains, it emphasizes that AI4SSH infrastructure is not equivalent to the accumulation of hardware. Rather, it is a supply of public capabilities that are reusable, interoperable, assessable, and auditable. The priority should shift from "resource scale" to the replicable provision of rules, processes, and capabilities, thereby supporting the intelligent transformation of the entire research process. Chapter 12 compiles typical cases of AI4SSH organization, promotion, and management in Chinese universities. It shows how universities transform the willingness to collaborate into the capacity to collaborate through scenario traction, platform foundations, and institutionalized governance. The cases cover Fudan University's coordinated point-line-surface advancement of AI4SSH, intelligent research on cultural heritage and early Chinese civilization, complex decision systems for digital government, intelligent strategic decision-making in education, and industry skills infrastructure for financial and insurance agents. Together, these cases illustrate an evolutionary pathway from individual projects to platform ecosystems, from technical applications to organizational mechanisms, and from research exploration to public service.

Compared with the previous edition, this Blue Book places greater emphasis on two transitions "from point to system." First, research methods are moving from the use of single-point tools toward the reconstruction of research organization and workflows. Greater importance is attached to data governance, benchmark evaluation, evidence-chain construction, process records, and auditable mechanisms for academic

reliability. AI is not merely a tool for improving the efficiency of individual research steps. It is becoming a new scientific language for organizing complex research tasks, connecting multiple forms of evidence, and supporting interdisciplinary collaboration. Second, university practice is moving from scattered exploration toward institutionalized system building. Around index assessment, infrastructure, organizational mechanisms, talent training, and typical scenarios, universities are forming mutually reinforcing action frameworks. The development of AI4SSH capability is no longer a technical experiment by a small number of teams. It is becoming a strategic capability for universities in disciplinary reconstruction, platform governance, public service, and international discourse building.

Overall, this book seeks to generate sustained impact at three levels. First, it provides the research community with an integrated framework and case base for "method–evidence–explanation–action," encouraging interdisciplinary collaboration to move from conceptual dialogue toward interface alignment, evidence co-construction, and mechanism testing. Second, it offers universities a systematic roadmap covering platform construction, data governance, compute and experimental environments, organizational mechanisms, evaluation systems, and talent training, thereby supporting the sustainable provision and scaled spillover of AI4SSH capabilities. Third, it provides public governance and social development with a toolbox and responsibility framework oriented toward real-world constraints, so that AI capabilities can enter decision-making processes under institutional conditions of transparency, verifiability, auditability, and accountability. We hope this annual Blue Book will become a shared language among academia, government, industry, and the public. It seeks both to understand the boundaries of AI capabilities through denser evidence and to transform technological gains into public capabilities through more robust institutions. In an intelligent era of accelerating change, it aims to preserve the value of deep thinking and to make thinking a key capability that connects technology, society, and the future.

Part I

Integration and Reshaping of AI and the Social Sciences and Humanities

Chapter1
Coordinating AI Development and Governance[①]

AI will continue to reshape how knowledge is produced and how society operates. In the face of this irreversible shift, the key is not to avoid the technology, but to calibrate capability boundaries through practice, reduce information asymmetry through transparency and verifiable mechanisms, and build platform-style infrastructure that supports cross-disciplinary innovation and the modernization of public governance. This chapter approaches the relationship between AI and the social sciences and humanities (SSH) through the lens of "public capability": it examines how AI is reshaping knowledge production and social operations, and how SSH can, in turn, provide sustainable boundaries and institutional support for AI development.

We have been trying to answer a question that seems paradoxical but is structurally the same: on the one hand, AI is empowering SSH at an unprecedented speed—expanding our observational scale, computational capacity, and explanatory toolkit; on the other hand, SSH must also support AI development and governance, putting an institutional steering wheel and seat belt on technological acceleration.

Across the narrative of development, risk, governance, constraints, and outlook, one message stands out: healthy AI development is not primarily about becoming "smarter," but about how intelligence is acquired, allocated, used, and constrained. Making intelligence a public capability means ensuring that technological gains can be understood, supervised, and jointly used within an institutional framework.

[①] This chapter was written under the leadership of Professor Wu Libo. Professor Wu Libo currently serves as a full-time mentor at Shanghai Innovation Institute, Assistant President of Fudan University, Vice Dean of Shanghai Innovation Institute, and Executive Director of the MOE Laboratory for National Development and Intelligent Governance at Fudan University. Her research focuses on energy and environmental economics, policy assessment modeling, applied statistical analysis of big data, and computational social science. She received the National Science Fund for Distinguished Young Scholars in 2019, was selected as a Changjiang Young Scholar by the Ministry of Education in 2016, and served as a lead author of Working Group III of the IPCC Sixth Assessment Report.

Chapter Highlights

- When decision-making, persuasion, prediction, and organizational capacity become software-defined and scalable, the “underlying parameters” of social systems can be rewritten; the key is how intelligence is acquired, allocated, used, and constrained.
- AI empowers SSH by lowering the barriers of scientific language and methods, restructuring research workflows, and reducing the coordination costs of cross-disciplinary collaboration. Its value is not only efficiency, but an expanded space for explanation and verification.
- SSH empowers AI in the opposite direction by structuring value conflicts into decision-ready trade-offs, translating social consequences into measurable indicators, and turning governance principles into accountability chains—thereby converting technological gains into public capability.
- Risk is not confined to the model itself; it also evolves through capability diffusion, deep embedding in critical systems, and competitive incentives. Irreversible diffusion and deep deployment can sharply compress the space for correction.
- A viable path forward is to map risk typologies to governance levers and to reduce information asymmetry through transparency, third-party evaluation, and public infrastructure—so that AI gradually becomes understandable, supervisable, and jointly usable as a public capability.

Chapter2

Pathways toward Human–AI Symbiosis in the Social Sciences and Humanities in the AGI Era[①]

This chapter focuses on the methodological "foundation" of two-way empowerment and symbiosis between AI and the social sciences and humanities (SSH). As AI rapidly enters research and governance workflows, knowledge integration is no longer a simple juxtaposition of disciplines. It requires both horizontal interoperability across fields and vertical embedding of AI into the research pipeline: horizontally, different disciplines must align on the same complex problem; vertically, new techniques must be integrated into how evidence is produced, checked, and used.

We use four research exemplars not to showcase isolated technical advances, but to extract transferable mechanisms. Coupled modeling illustrates how cross-disciplinary interfaces can move from manual stitching to learnable alignment. Social-system simulation shows how policy–behavior–risk can be stress-tested in an expanded scenario space. Causal inference shows how assumptions and robustness checks can be made explicit and auditable through workflow design. LLM-agent workflows show how long-horizon research tasks can be organized into collaborative, verifiable, and accountable processes.

Taken together, these pathways provide a general-purpose toolkit for downstream applications, and a research substrate for evidence-based governance and "computable social theory"—one that is verifiable, collaborative, and continuously improvable.

---

① This chapter was written under the leadership of Professor Wu Libo.

Chapter Highlights

- Across four exemplars, this chapter shows how AI can restructure SSH research workflows at multiple stages, with a shared goal of turning cross-disciplinary methodological capacity into reusable public infrastructure.
- In coupled modeling, the key challenge is interface alignment across scale, semantics, and data modalities. AI helps upgrade interfaces into learnable, testable, and iterative mechanism links through inference, alignment, adaptation, and optimization.
- In social-system simulation, the focus is a unified dynamic frame for policy–behavior–risk–resources. AI expands scenario coverage via data fusion, scenario generation, behavior/network modeling, and uncertainty quantification—enabling stress tests and option comparison.
- In causal inference, the move is from correlation to defensible causal evidence. AI supports structured research design and workflow-based validation so that assumptions, data constraints, and robustness checks become explicit and auditable.
- In agentic research, the shift is from point tools to organized research workflows. Task decomposition, role separation, tool use, and adversarial review make integrated research more verifiable and accountable.
- All four capabilities point to a human–AI collaborative evidence-production system: standardized schemas, auditable logs, task-oriented benchmarks, and human-in-the-loop gates reduce information asymmetry and coordination costs, aligning knowledge production with governance practice.

Chapter3
The Two–Way Convergence of AI and the Philosophy of Cognitive Science[①]

The continuing evolution of artificial intelligence has not only expanded machines' capacities for information processing, content generation, and action; it has also reopened the foundational question of cognitive science: how is intelligence possible? The reasoning, planning, and interactive capabilities displayed by large models and AI agents are compelling philosophy to revisit such core concepts as understanding, intentionality, embodiment, causation, and consciousness. At the same time, research in brain science on perception, memory, emotion, value, and cognitive control offers mechanistic guidance for moving AI beyond pattern matching toward contextual understanding, meaning-making, and reflective judgment. AI and the philosophy of cognitive science have thus entered a genuine two-way convergence. By clarifying the structure, boundaries, and value orientation of intelligence, the philosophy of cognitive science can help make AI more trustworthy, intelligible, and calibratable. AI, in turn, is transforming questions of mind that once relied primarily on conceptual analysis and thought experiments into objects that can be modeled, experimentally investigated, and systematically compared, bringing philosophy into sustained dialogue with engineering, data, and empirical research. This chapter unfolds across four dimensions—history, mechanisms, capabilities, and future directions. It traces the theoretical trajectory through which AI and the philosophy of cognitive science have mutually shaped one another; identifies the cognitive foundations of human intelligence; examines the capacities that cognitive AI requires for understanding subjects, generating meaning, and supporting deep thinking; and explores the future co-evolution of the two fields

[①] This chapter was jointly written by Associate Professor Huang Xiang of the School of Philosophy, Fudan University, and Research Professor Xiao Xiao of the Institute of Science and Technology for Brain-inspired Intelligence, Fudan University.

through frontier questions concerning understanding, intentionality, embodied cognition, causation, and consciousness.

Chapter Highlights

- From symbolism and connectionism to 4E cognition and predictive processing, AI and the philosophy of cognitive science have continually shaped one another: cognitive theories provide models for machine intelligence, while AI serves as an important experimental arena for testing and reconstructing theories of mind.
- Human intelligence is not a linear input-output process. It is a dynamic system in which perception and attention, memory and cognitive maps, emotion and value, and cognitive control jointly construct meaning across the body, environment, experience, and goals.
- Brain science and cognitive science point to four transitions for AI: from object recognition to contextual understanding, from information storage to the organization of experience, from emotion classification to value-sensitive understanding, and from fluent generation to self-monitoring, error correction, and reflective judgment.
- Cognitive AI for the social sciences and humanities should develop three classes of capability: understanding human subjects, generating meaning, and supporting deep thinking. Its purpose is not to replace human judgment, but to help people compare evidence, deliberate over values, and retain cognitive agency.
- Large models and AI agents are reactivating philosophical questions surrounding the Chinese Room, intentionality, 4E cognition, causation and counterfactuals, and consciousness. By turning these questions into objects that can be modeled, tested, and compared, they provide new theoretical foundations for trustworthy AI and human-AI symbiosis.

Chapter4
Reconstructing the Ethical and Governance Architecture for AGI[①]

In the new wave of AI advances—represented by foundation models, AI agents, and embodied intelligence—"generality" is no longer merely a capability metric. It is increasingly becoming an institutional variable that can reshape how society operates. Discussions of AGI therefore must be situated within a social-ethics and governance framework: the stakes go far beyond application-level gains and losses, reaching core governance domains such as the realization of rights, procedural fairness, responsibility allocation, public values, and cultural attitudes.

This chapter is organized around three threads—(i) a risk spectrum, (ii) trust and responsibility, and (iii) governance mechanisms. We first clarify the concept of AGI and its technological landscape. We then build an ethical-risk map across data bias, value embedding in algorithms, privacy protection, misinformation and hallucinations, cognitive manipulation, and the non-neutrality of technology. Third, we focus on the accountability vacuum and trust dilemmas triggered by decision automation, emphasizing the institutionalization of explainability, traceability, and redress mechanisms. Finally, we propose a four-dimensional governance framework to move governance from principle statements to operational coordination across culture, institutions, technology, and ethics.

Social-ethics governance in the AGI era is a cross-level, multi-actor, and cross-cultural systems project. Effective governance must enhance transparency, explainability, and traceability at the technical level; clarify accountability chains and

---

[①] This chapter was led by Prof. Yang Qingfeng (Fudan University), co-authored with Lin Junyu and Zhang Ruiyuan. Prof. Yang is a professor and researcher at the Institute for Ethics and the Future of Humanity, and a PhD supervisor at the School of Philosophy, Fudan University. His research spans philosophy of technology, philosophy of memory, data ethics, and AI ethics. He was a visiting scholar at Dartmouth College (2013–2014) and Swinburne University of Technology (2017). He serves as Secretary-General of the National Teaching Steering Committee for Applied Ethics and Vice President of the Shanghai Society for Dialectics of Nature. He is PI of the National Social Science Fund Major Project "Humanistic Philosophy at the Frontier of Emerging Enhancement Technologies." Lin Junyu is a senior engineer at the Institute for Ethics and the Future of Humanity, Fudan University; Zhang Ruiyuan is a PhD candidate in Philosophy of Science and Technology at the School of Philosophy, Fudan University.

pathways for remediation and redress at the institutional level; and protect public values and social trust at the ethical and cultural level. Only by forming an iterative loop among principles, mechanisms, and practice can the technological gains of general intelligence be transformed into sustainable public capability.

Chapter Highlights

- AGI debates face dual ambiguities in concept and reality. Governance should distinguish AGI from ASI and understand general capability as an evolving pathway rather than a single event—avoiding both exaggeration and complacency.
- AGI co-evolves with AIGC, AI agents, and embodied intelligence along a "part–whole, present–future" spectrum: content generation, autonomous execution, and physical interaction jointly push general capability from expression to action and presence.
- Ethical risks extend beyond data bias and privacy. With generative power, risks expand into misinformation and hallucinations, cognitive manipulation, and value reshaping—and can become structural inequality with self-reinforcing dynamics due to technology's non-neutrality.
- Decision automation pushes responsibility, trust, and power into a phase of redefinition. Black-box systems, fragmented actors, and automation bias may create accountability vacuums and trust collapse; procedural safeguards require institutionalized explainability, traceability, and redress.
- Data governance for the AGI era must shift from compliance to trust-building, embedding diversity, transparency, explainability, and traceability across the full lifecycle to form auditable accountability chains.
- Governance must move from principles to mechanisms. Through Culture–Institutions–Technology–Ethics (CITE) coordination, fairness, alignment, and pluralistic participation can be embedded into development and deployment workflows, and participation of vulnerable groups can improve legitimacy and sustainability.

Part II

Frontiers and Trends in Key Domains of Intelligent Social Sciences and Humanities

Chapter5
From Observation to Decision: Paradigm Transformation in Complex Systems and Social Simulation[①]

In the context of deep convergence among digitization, networking, and intelligent systems, human society is rapidly evolving into a prototypical complex adaptive system. Whether in public opinion diffusion, market dynamics, or macro collective behavior, these processes exhibit high uncertainty and complexity. Traditional social-science paradigms face structural methodological pressure. For a long time, approaches centered on surveys and statistical inference have been strong at describing patterns after the fact, but weaker at representing the dynamic evolution driven by multi-agent interaction.

With rapid advances in AI, social science is accelerating toward a complex-systems paradigm—shifting from "ex post explanation" to "process modeling and mechanism-based reasoning." By building LLM-driven multi-agent systems and simulating individual behaviors and group interactions in controllable digital environments, complex social processes can be translated into a computable, repeatable, and intervenable "digital experimental arena." This enables counterfactual reasoning and strategy optimization, providing experimental-style support for policy analysis and intelligent decision-making.

Against this backdrop, this chapter reviews AI-driven social simulation methods and key application pathways from a complex-systems perspective. We first introduce a general-purpose social simulation framework oriented toward real-world alignment, using SocioVerse as an exemplar, and explain its unified modeling mechanism across environment, users, interaction structure, and behavior generation. We then present an

[①] This chapter was led by Associate Professor Wei Zhongyu (Fudan University), and with collaboration of Associate Professor Liu Bao and Professor Chen Siming. Professor Wei Zhongyu is an Associate Professor and PhD supervisor at the School of Data Science, Fudan University, and leads the Data Intelligence and Social Computing Group. His research focuses on large-model–driven social computing and multimodal foundation models.

institution-constrained multi-agent platform for strategic interaction, using ProcureGym as an exemplar. Finally, we discuss social media simulation and visual analytics, highlighting an emerging human–AI collaborative loop from generation to analysis to steering.

Chapter Highlights

- In a digitized, networked, and intelligent society, social systems increasingly exhibit complex adaptive dynamics. Traditional paradigms centered on "ex post explanation" struggle to capture the dynamic evolution driven by multi-agent interaction; social science is moving from observational induction toward process modeling and mechanism-based reasoning.
- LLM-driven multi-agent social simulation provides a computable, repeatable, and intervenable "digital experimental arena" for complex social processes. It supports counterfactual reasoning and strategy optimization, offering experimental-style support for policy analysis and intelligent decision-making.
- In general-purpose social simulation frameworks (e.g., SocioVerse), the key is real-world alignment: ensuring consistency across environment context, population distribution, interaction structure, and behavior generation is foundational to credibility, interpretability, and scalability.
- In institution- and game-constrained settings (e.g., ProcureGym), policy rules, agent heterogeneity, and market interactions can be formalized as a computable Markov game. This supports sensitivity analysis of mechanism parameters and counterfactual evaluation, shifting policy assessment from experience-based judgment to verifiable mechanism experiments.
- In open social media settings, multi-agent simulation and visual analytics are converging into a human–AI collaborative loop of "generation–analysis–steering–re-simulation." Simulation provides controllable synthetic data and process logs; visual analytics strengthens explanations of emergent structure and diffusion pathways; together they improve understanding and intervention capacity.
- As social simulation moves from method exploration to system building and deployment, credibility floors must be protected: multi-source data and behavior alignment, interpretability and uncertainty disclosure, ethics and privacy compliance, and auditable versioning/logging mechanisms determine whether simulation can become governance-ready infrastructure.

Chapter6
From Identification to Estimation: Deep Learning – Enabled Causal Inference in Social Science[①]

The core concern of social science is often not whether variables are correlated, but what mechanisms cause what outcomes. This makes causal inference a foundational method for understanding social stratification, explaining behavioral differences, and evaluating the effects of public policy. In recent years, theoretical frameworks for causal identification have become increasingly mature. Yet in practice, a more prominent and often overlooked bottleneck lies in model dependence at the estimation stage. Many methods require parametric modeling of key quantities, such as propensity scores or conditional expectations. Once model specification becomes rigid or misspecified, systematic bias may arise under high-dimensional covariates, nonlinear relationships, and complex interaction structures. The resulting conclusions may then reflect "model artifacts" more than information contained in the data.

Deep learning, as a highly flexible nonparametric approximator, offers a new pathway for easing this bottleneck. On the one hand, representation learning and sample-balancing ideas move treatment-effect estimation beyond "control variables" or fitted equations toward more data-adaptive counterfactual prediction. On the other hand, generative distribution modeling and counterfactual simulation provide an operational computational framework for characterizing mediation mechanisms and estimating "cross-world" nested counterfactuals. Following two lines—effect estimation and mechanism analysis—this chapter reviews the logic through which deep learning and causal inference can be integrated. It emphasizes that, while identification assumptions and causal logic must be preserved, social-science causal analysis can move from

[①] This chapter was led by Professor Anning Hu, School of Social Development and Public Policy, Fudan University. Professor Hu is Dean of the School of Social Development and Public Policy at Fudan University, Chair of the Computational Sociology Section of the Chinese Sociological Association, a leading talent in philosophy and social sciences under the Ten Thousand Talents Program, a Young Chang Jiang Scholar, and a Shanghai Leading Talent.

parametric specification toward distribution-level modeling and verifiable evidence production.

Chapter Highlights

- The identification framework for social-science causal inference is now relatively mature, but practical bottlenecks increasingly concentrate in model dependence at the estimation stage. Parametric specifications of key quantities, such as propensity scores and conditional expectations, can introduce systematic bias under high dimensionality, nonlinearity, and complex interactions.
- As a data-adaptive nonparametric approximator, deep learning provides a technical pathway for reducing estimation bias caused by rigid specification. Its value lies in improving the stability of counterfactual prediction, not in replacing causal identification theory.
- For treatment-effect estimation, methods such as TARNet and CFR reframe causal inference as a problem of representation learning and sample balance. Shared representation spaces and distribution-matching constraints improve comparability between treated and control groups, reduce extrapolation risk, and strengthen causal generalization.
- For mechanism analysis, generative frameworks such as normalizing flows can model the conditional distributions of mediators and outcomes at the distribution level. They support simulated estimation of nested counterfactual quantities, including natural indirect effects, allowing mediation analysis to move from linear coefficient decomposition toward intervention, generation, and comparison under invertible transformations.
- Method integration requires three baseline disciplines: model complexity cannot compensate for unobserved confounding; highly flexible models require more data and face finite-sample risks; and the interpretability and auditability of deep models must be reinforced through explicit assumptions, robustness checks, and process records.
- Looking ahead, the key is not to replace traditional causal methods with deep learning. It is to build an integrated framework that combines rigorous identification with flexible estimation. Under the constraints of counterfactual logic and identification assumptions, representation learning and generative modeling can

support a more robust, more granular, and more verifiable chain of causal evidence production.

Chapter7
From Text to Multimodality: AI–Enabled Development of Computational Social Science[①]

Computational social science (CSS) has carried a "methodological ambition" from the outset: to describe social structures, behavioral interactions, and institutional operations in computable ways, so that social research can not only explain the past but also evaluate policy, reason about risk, and identify mechanisms under stricter evidentiary constraints. In the era of large models, this ambition has gained new technical support across the full chain of "data–model–reasoning–validation." Text is no longer merely research material. It can be structured, modeled, and aligned with behavioral and institutional processes as computable evidence. At the same time, evidence about the social world is no longer limited to text and statistical tables. It now extends to graph networks, spatiotemporal trajectories, images and videos, audio, and multimodal fact streams composed of multi-source sensor data. As a result, computational social science is moving from empirical analysis centered on text and variables toward mechanism modeling and system evaluation centered on multimodal evidence and complex systems.

This chapter shows how AI is reshaping three key capacities of computational social science. The first is large-scale evidence extraction and knowledge organization. Large models substantially reduce the cost of extracting facts, concepts, and relationships from unstructured materials, allowing policy texts, historical archives, public-opinion content, and scholarly literature to be encoded more systematically as searchable and verifiable structured evidence. The second is the computable representation of structure and interaction. Graph neural networks, representation learning, and multi-agent frameworks make it possible to describe the dynamic evolution of social networks, organizational relationships, and collective behavior at a finer grain, providing new modeling languages for structural causality and mechanism-

[①] This chapter was co-authored by Professor Baohua Zhou and doctoral students Yuqing Wu and Lingyun Zhang, School of Journalism, Fudan University.

based explanation. The third is a closed loop from explanation to intervention. Reinforcement learning, simulation, and digital-twin-style frameworks allow policy evaluation to move beyond ex post attribution and toward stress testing, distributional-effect identification, and option comparison in controllable scenario spaces, pushing research toward verifiable intervention reasoning.

It is important to emphasize that expanded capability does not automatically mean knowledge progress. Richer multimodal evidence can also introduce more noise, bias, and invisible selection mechanisms. Stronger model reasoning can also generate narratives that appear plausible but are difficult to test. Therefore, while presenting technical pathways, this chapter uses the "data–theory–mechanism" loop as its basic standard. It emphasizes the importance of reproducible evidence chains, interpretable assumptions, and auditable processes, and it discusses governance requirements in data compliance, algorithmic bias, and research responsibility. Ultimately, the value of AI-enabled computational social science does not lie in replacing social-science judgment. It lies in enabling social science to build more reliable explanations and more robust action-oriented knowledge within evidence systems that are denser and more transparent.

Chapter Highlights

- Large language models and multimodal large models are reshaping the content-analysis paradigm of computational social science. From text-based topic modeling, classification, and generation to the understanding, generation, and simulation of multimodal social data such as images, videos, sounds, and digital traces, AI is increasingly becoming a general-purpose "method engine" connecting social data with research reasoning.
- At the level of text analysis, large language models substantially ease the dual difficulty of high-cost manual coding and traditional algorithms' limited capacity to handle complex contexts. They show advantages in efficiency and interpretability in tasks such as topic labeling, sentiment and stance detection, and information disorder detection. Yet cultural context, emotional resonance, and implicit meaning still require human verification and theoretical constraints.
- For classification and identification tasks, large models may produce systematic biases in settings involving implicit emotions, polarized topics, and value alignment. This means that computational-social-science applications must incorporate bias diagnosis, uncertainty disclosure, and tiered validation into standard workflows.
- In text generation, large models can support explanation, summarization, and information extraction. They significantly reduce the workload of processing massive materials and improve the supply of structured information. At the same time, researchers must guard against emotion loss and minority-view loss caused by compression, as well as misleading risks from generated explanations and extrapolated conclusions. A human–AI evidence chain and review mechanism is therefore essential.
- Multimodal large models expand the data boundary of computational social science. They provide zero-shot or few-shot coding capabilities for media-content coding, fake-content and complex-pragmatic recognition, digital-behavior trace analysis, urban perception, and spatial governance. Cross-modal reasoning also improves the ability to identify and explain complex social contexts.
- Multimodal generation and multimodal agents provide new tools for "controlled social-situation generation–silicon-sample experiments–multi-agent simulation."

They significantly strengthen the experimentalization and reproducibility of social interaction. External validity, deep reasoning across modalities, and safety and controllability remain key bottlenecks. Standardized deployment requires benchmark evaluation, ethical compliance, and interpretable frameworks.

Chapter8
From Technological Embedding to Ethical Value Reflection: New Perspectives and Trends in Digital–Intelligent Humanities[①]

Digital-intelligent humanities is not simply the "digitization" of traditional materials. Under deep technological embedding, it reshapes research objects, methodological logic, and modes of knowledge production. Following the thread of "technological embedding—paradigm reconstruction—ethical reflection," this chapter traces the mechanisms through which digital humanities is moving toward digital-intelligent humanities. It also shows how digital-intelligent practices in fields such as literature, history, and philosophy generate transferable methodological experience.

The intervention of generative AI brings not only efficiency gains, but also systemic challenges to narrative structure, cultural expression, and academic ethics. Behind these challenges lie ontological questions about digital existence and epistemological tensions between computational rationality and humanistic understanding. Healthy development in digital-intelligent humanities must therefore rest on critique and reflexivity. In a new human–AI collaborative academic ecology, it must rebuild data governance, algorithmic transparency, and academic norms, so that technological progress and humanistic value move in the same direction. This chapter proceeds in five dimensions. It first examines the systemic reconstruction of tools, objects, and methods. It then analyzes digital-intelligent transformation in traditional humanities fields, including literature, history, and philosophy. It further discusses the impact and challenges of generative AI for narrative structure, cultural expression, and creative ethics. It then offers a deeper theoretical examination at the levels of ontology and epistemology. Finally, it proposes pathways for rebuilding the critical and reflexive functions of the humanities. Through this multi-dimensional examination, the chapter

---

[①] This chapter was led by Associate Professor Shaoqing Wen, Institute of Science and Technology Archaeology, Fudan University, and co-authored by Yiying Zhang, MOE Laboratory for National Development and Intelligent Governance, Fudan University.

offers an analytic framework that combines theoretical depth with practical concern for understanding new perspectives and trends in digital-intelligent humanities.

Chapter Highlights

- Digital humanities is moving from “document digitization–computational analysis” toward “intelligent generation–paradigm reconstruction.” The core of digital-intelligent humanities is not tool upgrading alone, but the systemic reconstruction of research objects, methodological logic, and knowledge-production structures.
- At the tool level, technologies such as pretrained language models and knowledge graphs are shifting humanities research from shallow statistics to deep semantic computation. Computers are moving from “counters” toward analytical agents with semantic-reasoning capabilities.
- At the object level, texts, archives, and cultural heritage are being transformed into computable data structures and semantic networks. Research questions are expanding from linear narrative to cross-spatiotemporal structural association and dynamic generation.
- At the method level, “computable interpretation” emphasizes interaction between computational discovery and meaning interpretation: distant reading provides macro-level navigation, close reading provides deep explanation, and the two form a complementary rather than substitutive relationship.
- Generative AI directly enters narrative and creative processes. It raises challenges involving authorship, cultural homogenization, algorithmic bias, copyright, and authenticity or hallucination, prompting the humanities to reassess the boundaries of technology at the ontological and epistemological levels.
- The strategic value of digital-intelligent humanities lies in rebuilding the critical and reflexive functions of the humanities. Through data critique, algorithm critique, and application critique, it can build a human–AI collaborative academic ecology and interdisciplinary talent system, creating institutional support that combines technical competence with humanistic concern.

Chapter9

From Efficiency to Responsibility: Application Modes and Responsibility Reconstruction in AI–Embedded Public Governance[①]

With the rapid development of big data, algorithmic models, and compute infrastructure, artificial intelligence—especially generative AI—is being embedded at accelerating speed across the full process of public governance. From intelligent Q&A, intelligent service guidance, and intelligent approval to assisted decision-making and proactive service push, AI not only improves the responsiveness, precision, and coordination efficiency of government services, but also drives public governance from passive response toward proactive service, from experience-based judgment toward data-driven operations, and from process optimization toward capability reconstruction. At the same time, AI is no longer merely a back-office tool. It is increasingly entering core governance links such as rights realization, procedural operation, responsibility allocation, and public-value formation, thereby reshaping the relationships among public departments and citizens, technology and institutions, and efficiency and legitimacy. Yet AI-enabled public governance is not a linear story of technological progress. Risks such as algorithmic "black boxes," model hallucinations, intelligence divides, organizational capability erosion, blurred boundaries of power and responsibility, and shifts in public value are becoming increasingly visible. In particular, between agency and assistance, and between human–AI collaboration and technological substitution, public governance must carefully define the boundaries of AI application.

This chapter discusses three questions: the application scenarios of AI in government services, the two modes through which AI is embedded in public

---

[①] This chapter was led by Prof. Lei Zheng, School of International Relations and Public Affairs, Fudan University. Prof. Zheng is Director of the Fudan Digital and Mobile Governance Lab, Vice Chair of the China Information Association's Digital Governance Committee, Co-Chair of the China Chapter of the International Digital Government Society, a member of the UN E-Government Survey Expert Group, and a member of the Shanghai Public Data Open Expert Committee.

governance, and the reconstruction of responsibility that follows. It seeks to identify a path toward responsible AI governance that balances efficiency gains with value commitments.

Chapter Highlights

- AI is being widely applied in intelligent Q&A, intelligent service guidance, intelligent approval, assisted decision-making, and proactive service push. It is moving government services from passive response toward proactive, precise, and intelligent provision.
- While AI improves administrative efficiency, optimizes service processes, and strengthens data-based assessment, it also introduces new problems, including intelligence divides, algorithmic bias, data security risks, and unequal access to services.
- AI embedded in public governance mainly takes two forms: an agentic mode and an assistive mode. They differ fundamentally in the allocation of decision-making power, responsibility attribution, and governance legitimacy.
- Different governance matters should determine the depth of AI intervention based on rule clarity, rights impact, complexity of value conflicts, and the level of social trust. High-risk scenarios should adhere to human–AI collaboration and human-in-the-loop safeguards.
- Assistive systems may drift in practice toward de facto agentic governance because of automation bias, organizational dependence, and performance-driven incentives. This can lead to responsibility offloading and power alienation.
- Responsible AI governance should rebuild full-lifecycle responsibility chains around data, models, deployment, use, correction, and remedies, ensuring that public responsibility cannot be replaced by technology or vendors.
- AI use in public governance should be accompanied by duties to explain, human review, complaint and appeal channels, audit oversight, and dynamic evaluation, embedding fairness, transparency, privacy, security, and public trust throughout technological governance.

Part III

Exploration, Action, and Institutional Development of AI4SSH in Chinese Universities

This is an abstract.

Please look forward to the official release of the complete version during the 2026 World Artificial Intelligence Conference (WAIC).

Chapter10
Indexed Observation of AI4SSH Development in Chinese Universities[①]

AI technology is penetrating the social sciences and humanities with unprecedented breadth and depth, driving a shift in research paradigms from tool-based application to methodological reconstruction. In this process, universities are both the main sites of frontier exploration and key organizers of interdisciplinary ecosystems. Building a sustainable path between diffuse application and institutionalized development requires an evaluation tool that can provide both a panoramic profile and a directional benchmark.

This chapter introduces the China University AI4SSH Index. Built around three dimensions—research core capacity, development and innovation potential, and social communication capacity—and supported by a multi-source evidence chain, the index offers an indexed observation of AI4SSH development in Chinese universities. Its purpose is to identify structural features, reveal capability gaps and growth drivers, and support the transition of integrative research from spontaneous exploration to organized, assessable, and iterative institutional development.

Following a structure of "method—profile—conclusions and outlook," this chapter addresses three key questions: What gradients and clusters characterize AI4SSH development in Chinese universities? Which capability dimensions are accumulating rapidly, and where do obvious weaknesses remain? In the next stage, how can institutional supply, platform support, and talent mechanisms help move the field from scale accumulation toward quality leadership?

[①] This chapter was led by Professor Zhao Xing of the Institute for Big Data, Fudan University, and co-authored with Dr. Wang Yihan of the National Experimental Base for Intelligent Evaluation and Governance, Fudan University, and others. Professor Zhao Xing is currently Deputy Director of the National Experimental Base for Intelligent Evaluation and Governance, Fudan University; a Shanghai "Shuguang Scholar"; a member of the Ministry of Education Teaching Guidance Committee; an editorial board member of Journal of Informetrics, a leading international journal in informetrics and science evaluation; and an editorial board member of five major Chinese journals.

Chapter Highlights

- Using the China University AI4SSH Index as its analytical tool, this chapter offers a panoramic view of the overall progress and structural features of Chinese universities in the deep integration of AI with the social sciences and humanities. The index serves both as an evaluation benchmark and a development guide, providing quantitative reference points for disciplinary planning, resource allocation, and ecosystem building.
- The index system covers three dimensions—research core capacity, development and innovation potential, and social communication capacity—and breaks them down into a multi-level indicator structure. It emphasizes identifying universities' comprehensive AI4SSH capacity and sustainable development potential from the perspectives of paradigm integration, organizational coordination, and ecosystem building.
- The overall pattern shows a structured development landscape of an emerging system with clear tiers. Research output and domestic integration are advancing quickly, generating visible clustering effects and a regional coordination framework. However, substantial differences remain across universities and regions in development intensity and capability structure.
- Ecosystem building is entering an accelerated stage of entity-based and institutionalized development. The establishment of interdisciplinary research institutions and platform-based layouts has become a key variable in moving AI4SSH from scattered exploration toward systematic development, although institutional support and long-term mechanisms are still at an early stage in most universities.
- Social communication and public influence show a stratified distribution. The conversion of academic outputs into public agendas and social services still needs improvement. How to transform research capacity into policy support and social empowerment is a missing link that must be strengthened in the next stage.
- Looking ahead, it is urgent to move paradigm integration from application following to innovation leadership; build a sustainable ecosystem with internal–external linkage and institutional support; strengthen demand-driven social service

and discourse systems; and cultivate interdisciplinary talent with a strong disciplinary foundation, deep cross-disciplinary capacity, and strong collaborative ability.

Chapter11
AI4SSH Infrastructure and Representative Projects: Current Status and Outlook[①]

The impact of artificial intelligence on research activity is expanding from isolated tool use to a systemic reshaping of research infrastructure, research workflows, and the organization of innovation. Against this backdrop, building AI4SSH infrastructure has become a key lever for universities seeking to advance the digital-intelligent transformation of the social sciences and humanities. It determines whether data, computing power, models, and tools can become a sustainable on-campus public provision, and whether interdisciplinary collaboration can move from project-based cooperation toward long-term institutionalization.

This chapter follows three main lines: data foundations and circulation mechanisms; integrated computing power and research apparatuses; and models, agents, and toolchains. On the basis of domestic and international policy trends, conceptual evolution, and public practices at Chinese universities, the chapter shows the staged transition of AI4SSH infrastructure from "resource aggregation" to "computable organization," from "computational support" to "experimental research environments," and from "model access" to a capability system that covers the full research process.

AI4SSH infrastructure is not a pile-up of hardware. For the social sciences and humanities, its core is to form reusable, interoperable, and auditable public capabilities: computable and circulable data; usable and sustainable computing environments; controllable and scalable models and tools; and organizational mechanisms and governance rules that support human–AI collaborative knowledge production.

---

① This chapter was led by Zhang Jilong, research librarian at the School of Data Science, Fudan University. Zhang Jilong currently serves as a member of a Ministry of Education teaching steering committee, Deputy Director of the MOE Laboratory for National Development and Intelligent Governance at Fudan University, Director of the Shanghai Joint Innovation Laboratory for Big Data in Scientific Research, and Executive Deputy Director of the Fudan–Aifadi Joint Research Center for Smart Library Studies. His research areas include smart libraries and scientific data management.

Chapter Highlights

- AI4SSH infrastructure has become an important component of global and Chinese science and technology strategy. A coordinated pattern is emerging across international strategic guidance, national top-level planning, local policy responses, and university platform implementation.
- Conceptually, research infrastructure is evolving toward digital-intelligence scholarly infrastructure. Its core breakthrough lies in system integration across data, computing power, models, and toolchains, which turns platform-based research, human–AI collaboration, and knowledge symbiosis into infrastructural capabilities guided by publicness, reusability, and sustainability.
- Three trends are visible in the construction of data foundations: a shift from dispersed resource repositories toward university-level foundations and integrated platform groups; an expansion of data objects from structured statistics to multi-source heterogeneous and multimodal knowledge objects; and a transition from static data sharing toward workflow-embedded circulation mechanisms and AI-ready data.
- Computing infrastructure is moving from single high-performance computing (HPC) platforms toward integrated foundations that combine supercomputing, AI computing, and general-purpose computing. Social simulation and multi-agent experimental platforms are also emerging, providing "experimental" research spaces for AI4SSH.
- Models, agents, and toolchains are becoming a capability system for the full research process. Local recomposition of general-purpose models, discipline-specific deepening of domain models, and platform-based toolchain provision are jointly moving AI capabilities from being merely usable toward being embeddable, reproducible, and governable.
- Platform homogenization, inconsistent accounting of computing capacity, insufficient resource accessibility, unclear ethical and compliance boundaries, and the absence of evaluation for platform contributions still constrain the large-scale spillover of infrastructure. Future development should use interoperability standards, evaluation and auditing mechanisms, service-oriented teams, and sustainable governance rules to move AI4SSH infrastructure from scale expansion toward high-quality public provision.

Chapter12
Typical Cases of Organization, Promotion, and Management of AI4SSH in Chinese Universities[①]

As AI for Social Sciences and Humanities (AI4SSH) moves from point-based exploration toward institutional system building, the key question is not whether any single technology is advanced. It is whether organizational arrangements, platform capabilities, incentive mechanisms, and governance rules can match the complexity of interdisciplinary integration. Universities are not only the main sites of knowledge production. They are also organizers of interdisciplinary ecosystems. They need to translate computing power and data foundations, cross-disciplinary coordination mechanisms, engineering-oriented support teams, and quality-control rules into sustainable public capabilities.

This chapter selects four university cases: Fudan University, Wuhan University, Renmin University of China, and Beijing Normal University. These cases cover different pathways, including platform-based infrastructure construction, digital-intelligence humanities engineering driven by real-world scenarios, complex decision systems for governance, and evidence-based teaching-research loops for education reform. Together, they point to a transferable experience chain: real public tasks create demand; shared platforms and engineering teams provide capacity; institutionalized academic governance guarantees quality; and alliances, training, and open cooperation support diffusion.

The purpose of this chapter is to present replicable elements of AI4SSH organizational innovation through case analysis. It offers reference points for universities seeking, in the context of the 15th Five-Year Plan, to advance the intelligent transformation of disciplines, build humanities and social-science laboratories and cross-disciplinary platforms, and improve research management and incentive systems.

[①] This chapter was written by Tang Weiqi and Wang Lingxiang of Fudan University's MOE Laboratory for National Development and Intelligent Governance. Source materials were provided by Fudan University's Office of Liberal Arts and the Alliance of University Philosophy and Social Sciences Laboratories.

Chapter Highlights

- The systematic advancement of AI4SSH requires organizational-capacity building, including stable talent pipelines, cross-school and cross-department coordination mechanisms, engineering support teams, and sustainable platform rules.
- Scenario-driven organization is one of the most effective adhesives for interdisciplinary collaboration. Complex public tasks such as education reform, cultural heritage protection, digital government, and national governance can embed academic explanation, engineering implementation, and social service into the same loop.
- Platform foundations determine replicability. Shared mechanisms and compliance rules for computing power, data, models, and toolchains can turn the willingness to collaborate into the capability to collaborate, lower cross-team costs, and improve the reproducibility of outcomes.
- Institutionalized academic governance and quality control are preconditions for scale. Academic committees, sub-laboratory matrices, versioned data governance, auditable logs, and third-party evaluation help turn the research process itself into a manageable evidence chain.
- Diffusion and ecosystem building depend on alliance networks and talent systems. Summer schools, annual conferences, training programs, and co-built shared platforms can create a common language and workflow templates, supporting a shift from individual cases to scaled applications.
- The experiences of leading universities point to a shared logic of "infrastructure–norms–outputs." Institutionally supplied public capabilities can promote innovation while keeping risks controllable, and can improve the interpretability of research outcomes as they enter public decision-making and social service.