摘要单行本

Reclaiming the Value of Thinking

重新发现“深度思考”的价值

2026 人文社会科学智能发展蓝皮书

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

主编：陈志敏　吴力波

復旦大學

牵头单位：

教育部哲学社会科学实验室——复旦大学国家发展与智能治理综合实验室

上海科学智能研究院

总 顾 问：金　力

主　　编：陈志敏　吴力波

副 主 编：文少卿　张计龙　黄　昊

编　　委：漆　远　李　昊　袁嘉灿　周　阳　陈　斌　施正昱

徐盈辉　黄　翔　肖　晓　杨庆峰　林俊宇　魏忠钰

刘　宝　陈思明　胡安宁　周葆华　郑　磊　赵　星

王译晗　伏安娜　程蕴涵　殷沈琴　汪东伟　许　闲

刘　闯　高奇琦

执行编辑：汤维祺　王凌翔　张以颖

注：按参与章节先后排序

扫描二维码获取本书完整内容：

序：以思想之深引领智能之变

习近平总书记在致复旦大学建校 120 周年的贺信中指出，要推动哲学社会科学知识创新、理论创新、方法创新。这为高校新文科建设指明了前进方向，也为我们探索智能与人文社科的融合之道提供了根本遵循。

《2026 人文社会科学智能发展蓝皮书》（下称《蓝皮书》）以“知识创新、理论创新、方法创新”为总纲，以“重新发现‘深度思考’的价值”为主题，深刻辨析了人文社科与人工智能之间的互动张力，生动展现了中国高校在人文社会科学智能范式变革的最新探索，为守正创新推动哲学社会科学高质量发展提供了重要启示。

这是一部回应时代命题、饱含现实关怀的著作。翻开书页，一位位人文社科领域学者沉潜思考、躬身探索的形象跃然眼前。他们始终聚焦前沿实践，将目光对准国家发展、可持续转型、公共治理、社会信任等人类重大议题，提出真问题、建构真理论、创造真价值，让哲学社会科学研究在智能时代保持回应时代之问、解答发展之惑的穿透力。

这是一部超越工具理性、倡导人机共生的著作。2026 年《蓝皮书》的一个鲜明特点，就是把“AI 赋能”升级为“双向赋能、融合共生”：既讨论技术如何推动知识生产方式迭代更新，也讨论学术共同体如何引导技术价值向善、为技术赋能增效。这种相互促进、相得益彰的融合范式，释放了技术潜力、守住了人本底线，更让我们看到智能助力人文、思考引领创新的无限可能。

这是一部回归思想本质、呼唤深度思考的著作。本书呈现了一批高价值的人文社会科学研究成果，它们依赖的不是“更快的生成”，而是“更深的认知”。每一个观点都经过长周期、全链条、强情境的叩问、淬炼和打磨，

最终沉淀出直击本质、洞察规律、通达事理的思想成果。

当前，智能浪潮奔涌而至，为哲学社会科学带来了新工具、新场景、新方法，也对思想的独立性、深刻性与创造性提出了更高要求。相信新一年《蓝皮书》的问世，不仅能为学界提供重要的研究参考，激励更多学人投身于人文社会科学与人工智能融合创新的事业，不断拓展我们认知世界、理解世界的广度与深度；也能感召更多读者在智能时代始终守护思想、砥砺思考，保留独立思索、理性判断、追问价值、明辨取舍的从容和定力，让思想之光与智能之翼融合共生、交相辉映。

自 2025 年首发以来，《蓝皮书》在推动中国高校人文社会科学智能范式变革、促进新文科建设等方面发挥了重要作用。希望《蓝皮书》团队始终深耕人文社会科学与智能深度融合的核心命题，做深度思考的坚守者、智能赋能的探路者、新文科发展的先行者，持续产出兼具思想深度与实践价值的成果，为服务中国哲学社会科学自主知识体系构建、支撑中国式现代化发展贡献更大的智慧力量！

是为序。

复旦大学党委书记　裘新

2026 年 5 月

摘 要

生成式大模型、多模态学习与智能体技术的突破，正在推动人工智能从单一工具集加速演化为可嵌入科研、教育、产业与治理流程的通用能力。与此同时，围绕“替代还是赋能”“效率还是责任”“生成还是思考”的讨论也不断升温。一方面，AI 显著降低了写作、编程、检索、标注、建模与分析的形式门槛，释放出前所未有的效率增益；另一方面，能力扩散、深度部署和竞赛激励也在重塑知识生产方式、组织协作模式、公共信任结构与社会治理条件。面对这一不可逆过程，本书以“重新发现‘深度思考’的价值”为主题，强调把人的创造力、判断力和价值辨析能力从重复性形式劳动中解放出来，并在“发展—风险—治理”一体化视角下，回答 AI 如何健康、稳健、安全地嵌入人文社会科学与社会运行：既讨论 AI 如何赋能知识生产，也讨论人文社会科学如何为 AI 提供方向、边界与制度工程支撑。

本期蓝皮书延续“赋能”与“共生”的主线，围绕“数据—理论—机制—治理”闭环构建整体叙事，力求把分散的技术进展沉淀为可复用的方法能力，把点状的高校探索提升为可推广的组织经验与制度方案。全书分为三篇十一章：第一篇聚焦人工智能与人文社会科学的融合与重塑，强调智能作为公共能力的获得方式、分配结构、使用规则与治理边界；第二篇聚焦人文社会科学智能关键领域的发展前沿与趋势，系统呈现复杂系统与社会仿真、社会科学因果推断、计算社会科学、数智人文和公共治理等方向的范式跃迁；第三篇聚焦中国高校 AI4SSH 的探索、行动与体系建制，通过指数评估、基础设施梳理和典型案例分析，刻画我国高校 AI4SSH 从项目探索走向平台建设、生态协同和制度化发展的阶段性路径。

本期蓝皮书所面对的，不仅是技术能力的跃升，更是知识生产与社会治理条件的整体变化。当高质量文本生成、复杂代码实现、图像与视频理解、多模态推理和智能体协作逐渐成为可调用能力，研究瓶颈正在从“能否处理材料”转向“能否提出好问题、构建真机制、形成可检验的证据链”；当仿真、优化和智能体开始参与政策推演、组织管理和公共服务，治理挑战也从“是否使用 AI”转向“如何在价值冲突、不确定性和责任约束下负责任地使用 AI”。特别是在可持续发展、气候变化、公共治理、文化传承、社会信任等重大议题上，研究对象往往同时具有跨尺度、跨系统、强情境和高不确定性的特征，既需要跨模态数据与模型耦合，也需要把分布效应、公平权衡、制度约束和价值判断纳入同一推理结构。由此，AI4SSH 并非边缘化的技术应用，而是面向重大问题的知识与治理能力重构。

第一篇从宏观层面回答“智能如何成为公共能力”。第一章讨论 AI 发展与治理的协同，指出 AI 对社会的深层影响不止于生产力提升，而在于决策、预测、说服与组织能力的可规模复制，由此重写社会运行的底层参数。AI 健康发展的关键不只是模型是否“更聪明”，而是智能如何被获得、被分配、被使用和被约束。第二章面向 AGI 时代的人文社会科学智能融合与共生路径，聚焦复杂系统耦合建模、经济社会系统高通量仿真、因果推断和大模型智能体工作流，提炼可迁移的方法机制：跨学科接口从手工拼接走向可学习对接，政策—行为—风险在情景空间中被压力测试，“相关—因果—行动”的证据链被流程化与审计化，研究任务被组织为可分工、可复核、可追责的工作流。第三章进一步讨论 AI 伦理与治理的体系重构，围绕 AGI 的技术本质、概念歧义与人类位置，梳理风险谱系，分析决策自动化与信任重建，并提出面向伦理工程和制度设计的四维治理框架。

第二篇面向关键方法与治理议题，呈现人文社会科学智能从“解释世界”走向“模拟过程、评估干预、支撑决策”的研究跃迁。第四章以复杂系统与社会仿真为核心，指出在数字化、网络化和智能化深度融合的背景下，社会系统日益呈现复杂适应系统特征。大模型驱动的多智能体社会仿真能够在可控数字环境中模拟个体行为与群体互动，将复杂社会过程转化为可计算、可重复、可干预的“数字实验场”，为反事实推演、策略优化和政策评估提供实验化支撑。第五章聚焦深度学习赋能的社会科学因果推断，指出因果识别理论已相对成熟，但估计阶段的模型依赖性仍是实践瓶颈。深度学习可以通过表示学习、样本平衡和生成式分布建模增强反事实预测与机制分析能力，在坚守识别假设的前提下，提升因果估计的灵活性、稳健性和可复核性。第六章以“从文本到多模态”为主线，讨论大语言模型和多模态大模型如何重塑计算社会科学：一方面降低文本主题建模、分类、信息抽取和生成分析的成本，推动非结构化材料向可计算证据转化；另一方面把图像、视频、语音、轨迹等多模态事实流纳入研究对象，拓展计算社会科学的观测范围、实验空间和模拟能力。第七章从数智人文的新视角出发，讨论从“数字人文”到“数智人文”的范式跃迁，分析生成式 AI 对叙事、文化表达、真实性和作者性的冲击，并强调人文学科在智能时代仍需承担批判性重建、价值辨析和意义生产的核心功能。第八章聚焦人工智能嵌入公共治理的模式与责任，梳理 AI 在政务服务中的应用场景与潜在风险，区分“代理型”和“辅助型”两种嵌入模式，并从责任链重建、说明义务、审计监督、申诉救济和公共信任维护等方面，提出负责任使用 AI 的治理框架。

第三篇把视角落到中国高校 AI4SSH 的体系化行动。第九章构建“中国高校 AI4SSH 指数”，从研究核心能力、发展创新潜力和社会传播能力三个维度，系统观察我国高校在 AI 与人文社会科学深度融合中的总体进展与结

构特征。指数评估显示，我国高校 AI4SSH 发展已呈现“体系初构、梯次分明”的格局：研究产出规模和本土融合推进较快，若干高水平研究型大学已形成较为成熟的交叉研究体系；但国际学术影响力、源头创新能力、制度性支撑和社会服务转化仍存在短板，不同高校、不同区域之间的发展强度与能力结构差异显著。第十章聚焦 AI4SSH 基础设施和代表性项目建设，围绕数据底座与流通机制、算力融合与实验环境、仿真平台、模型/智能体与工具链等主线，强调 AI4SSH 基础设施不等同于硬件堆砌，而是可复用、可互操作、可评测、可审计的公共能力供给。其建设重点应从“资源规模”转向“规则、流程与能力的可复制供给”，以支撑研究全过程的智能化转型。第十一章汇编中国高校 AI4SSH 组织、推动与管理典型案例，展示高校如何通过场景牵引、平台底座和制度化治理，将协作意愿转化为协作能力。相关案例覆盖复旦大学 AI4SSH 点线面协同推进、文化遗产与早期中华文明智能化研究、数字政府复杂决策系统、教育战略智能决策、金融保险智能体行业技能基础设施等方向，呈现出从单点项目到平台生态、从技术应用到组织机制、从研究探索到公共服务的演进路径。

与上一年度相比，本期蓝皮书更加强调两类“从点到面”的转变。第一，研究方法从单点工具使用走向研究组织与工作流重构，更加重视数据治理、基准评测、证据链构建、过程记录和可审计机制对学术可靠性的意义。AI 不只是提升单个研究环节效率的工具，而正在成为组织复杂研究任务、连接多源证据、支持跨学科协作的新型科学语言。第二，高校实践从分散探索走向体系建制，围绕指数评估、基础设施、组织机制、人才培养和典型场景形成相互支撑的行动框架。AI4SSH 能力的建设，不再只是少数团队的技术试验，而是高校在学科重构、平台治理、公共服务和国际话语体系建设中的战略能力。

总体而言，本书希望在三个层面形成持续影响：其一，为研究共同体提供“方法—证据—解释—行动”一体化的通用框架与案例库，推动跨学科协作从观点对话走向接口对齐、证据共建和机制检验；其二，为高校提供从平台建设、数据治理、算力与实验环境，到组织机制、评价制度和人才培养的系统路线图，促进 AI4SSH 能力的可持续供给与规模化外溢；其三，为公共治理和社会发展提供面向现实约束的工具箱与责任框架，使 AI 能力在透明、可验证、可审计和可问责的制度条件下进入决策流程。我们期待这份年度蓝皮书能够成为学界、政府、产业与公众之间的共同语言：既以更高密度的证据认识 AI 的能力边界，也以更稳健的制度把技术增益转化为公共能力，从而在不断加速的智能时代，守住“深度思考”的价值，并让思考成为连接技术、社会与未来的关键能力。

目　录

第一篇

人工智能与人文社会科学的融合与重塑

第一章　AI 发展与治理的协同*

AI 将持续改变知识生产与社会运行方式。面向这一不可逆过程，关键不在于回避技术，而在于以实践校准能力边界，以透明与可验证机制降低信息不对称，并以平台化基础设施支撑跨学科创新与公共治理现代化。本章从“公共能力”的视角讨论 AI 与人文社会科学的关系：既关注 AI 如何重塑知识生产与社会运行，也强调人文社会科学如何为 AI 发展提供可持续的边界与制度支撑。

我们一直试图回答一个看似矛盾、实则同构的问题：一方面，AI 正以前所未有的速度“赋能”人文社会科学，扩展我们的观测尺度、计算能力与解释工具；另一方面，人文社会科学也必须反过来为 AI 的发展与治理提供支撑，给技术的加速装上制度的方向盘与安全带。

从“发展—风险—治理—困境—展望”的叙事中，我们看到 AI 健康发展的关键不在于“更聪明”，而在于“智能如何被获得、被分配、被使用、被约束”。让智能成为公共能力，意味着让技术增益在制度框架内可被理解、可被监督、可被共同使用。

---

* 本章由吴力波教授牵头撰写。

吴力波教授现任上海创智学院全时导师，复旦大学校长助理，上海创智学院副院长，复旦大学国家发展与智能治理综合实验室执行主任。主要从事能源环境经济与政策评估建模研究、大数据应用统计分析与计算社会科学研究。2019 年国家杰出青年科基金获得者，2016 年入选教育部青年长江学者，IPCC 第六次评估报告第三工作组主要作者。

本章提要：

- 当决策、说服、预测与组织能力被软件化并可规模复制，社会运行的“底层参数”将被重写；AI 的关键在于智能如何被获得、分配、使用与约束。
- AI 对人文社会科学的赋能主要体现在科学语言与方法门槛下降、研究流程重构与跨学科协同成本降低；其价值不止于效率提升，更在于拓展可解释与可验证的研究空间。
- 人文社会科学对 AI 的反向赋能在于：将价值冲突结构化为可决策权衡，将社会后果转化为可测指标，将治理原则落实为责任链条与问责机制，从而把技术增益转化为公共能力。
- 风险并不只在模型内部，更在能力扩散、关键系统嵌入与竞赛激励结构中演化；“不可逆”的扩散与深度部署会显著压缩纠偏空间。
- 健康发展路径应以风险分型映射治理抓手，通过透明披露、第三方测评与公共基础设施建设，减少信息不对称，使 AI 逐步成为可被理解、可被监督、可被共同使用的公共能力。

第二章　AGI 时代的人文社会科学智能融合与共生路径*

本章聚焦 AI 赋能人文社会科学融合共生的“方法底座”。在 AI 加速进入研究与治理流程的背景下，知识融合不再只是学科之间的并列叠加，而是要求在同一研究链条中实现横向互通与纵向嵌入：横向是不同学科对同一复杂问题的共同建模与共同证据语言；纵向是大模型、多模态处理、因果推断与仿真工具对研究流程的深度嵌入，使问题界定、证据组织与方案评估形成可迭代闭环。

本章以四个研究样本为支点，但目的并非呈现单一领域的技术进展，而是提炼可迁移的共性机制。复杂系统耦合建模展示跨学科接口如何从手工拼接升级为可学习对接；社会系统模拟展示政策—行为—风险如何在情景空间中被压力测试；因果推断展示‘相关—因果—行动’的证据链如何被流程化与审计化；大模型智能体展示研究任务如何被组织为可分工、可复核、可追责的工作流。四者共同指向一个方向：把方法论能力建设成可复用的公共基础设施。

本章内容为后续关键领域应用提供通用工具箱，也为“证据型治理”与“可计算的社会理论”奠定可验证、可协作、可持续迭代的研究底盘。

---

* 本章由吴力波教授牵头撰写。张人禾、张小曳院士，漆远、袁嘉灿教授，李昊、徐盈辉研究员周阳、陈斌副教授、汤维祺副研究员，施正昱、史少杰、陈子健、仲晓辉博士合作撰写。

本章提要：

- 本章以四个研究样本展示 AI 如何在不同环节重构人文社会科学的研究流程，其共同目标是将跨学科方法能力沉淀为可复用的公共基础设施。
- 复杂系统耦合建模强调跨尺度、跨语义、跨数据形态的接口对齐；AI 通过推断、对齐、适应与优化把接口升级为可学习、可校验、可迭代的机制连接。
- 社会系统模拟强调把政策—行为—风险—资源置于同一动态框架；AI 通过数据融合、情景生成、行为/网络建模与不确定性量化扩展情景覆盖，支持压力测试与方案比较。
- 因果推断强调从相关走向可辩护的因果证据；AI 通过结构化研究设计与工作流化验证让假设、数据约束与稳健性检验显性化，提高透明度与可审计性。
- 大模型智能体强调从单点工具走向研究工作流的组织化跃迁；任务分解、角色分工、工具调用与对抗式审查使综合研究更可复核、更可追责。
- 四类能力共同指向‘人机协作的证据生产体系’：以标准化 schema、可审计日志、面向任务的评测基准与人在回路机制，降低信息不对称与协作成本，推动知识生产与治理实践同向迭代。

第三章　AI 与认知科学哲学的双向奔赴*

人工智能的持续演进，不仅扩展了机器的信息处理、内容生成和行动能力，也重新打开了“智能如何可能”这一认知科学的根本问题。大模型和智能体展现出的推理、规划与交互能力，正在迫使哲学重新审视理解、意向性、具身性、因果和意识等核心概念；与此同时，脑科学关于感知、记忆、情绪、价值和认知控制的研究，也为 AI 突破模式匹配局限、走向情境理解、意义生成和反思判断提供了机制参照。由此，AI 与认知科学哲学形成了真正的“双向奔赴”：认知科学哲学通过澄清智能的结构、边界和价值方向，推动 AI 走向可信、可理解、可校准；AI 则把长期依赖概念分析和思想实验的心智问题转化为可建模、可实验、可比较的研究对象，推动哲学进入与工程、数据和实验持续对话的新阶段。本章从历史、机制、能力与未来四个层面展开：梳理 AI 与认知科学哲学相互塑造的理论轨迹，揭示人类智能的认知底座，讨论认知型 AI 在主体理解、意义生成和深度思考方面的能力要求，并围绕理解、意向性、具身认知、因果与意识等前沿议题，展望二者协同演进的未来方向。

* 本章由复旦大学哲学学院黄翔副教授和类脑智能科学与技术研究院肖晓研究员合作撰写。

本章提要：

- 从符号主义、联结主义、4E 认知到预测加工，AI 与认知科学哲学始终相互塑造：认知理论为机器智能提供模型，AI 则成为检验和重构心智理论的重要实验场。
- 人类智能并非线性的“输入—输出”过程，而是感知与注意、记忆与认知地图、情绪与价值、认知控制在身体、环境、经验和目标之间协同建构意义的动态系统。
- 脑科学和认知科学启示 AI 实现四重跃迁：从对象识别走向情境理解，从信息存储走向经验组织，从情绪分类走向价值理解，从流畅生成走向自我监控、错误修正和反思判断。
- 面向人文社会科学的认知型 AI，应发展主体理解、意义生成和深度思考三类能力；其目标不是替代人的判断，而是帮助人比较证据、辨析价值并保持认知主动性。
- 大模型与智能体正在重新激活“中文屋”、意向性、4E 认知、因果与反事实、意识等哲学问题，并将其转化为可建模、可实验、可比较的研究对象，为可信 AI 和人机共生提供新的理论基础。

第四章　AGI 伦理与治理的体系重构*

在以大模型、智能体与具身智能为代表的新一轮人工智能跃迁中，“通用性”不再只是能力指标，而逐步演化为影响社会运行方式的制度变量。AGI 相关讨论之所以需要纳入社会伦理治理框架，是因为其潜在影响不止于技术应用的得失，更将触及权利实现、程序正当、责任分配、公共价值与文化心态等治理核心环节。

本章以“风险谱系—信任与责任—治理机制”的主线组织论述：首先澄清 AGI 概念与技术谱系，继而从数据偏见、算法价值嵌入、隐私保护、虚假与幻觉信息、认知操控与技术非中立等维度构建伦理风险图谱；在此基础上，聚焦决策自动化所引发的责任真空与信任困境，强调可解释、可追溯与救济机制的制度化意义；最后提出面向 AGI 时代的四维治理框架，推动治理从原则宣示走向可操作的文化、制度、技术、伦理的协同。

AGI 时代的社会伦理治理是一项跨层级、跨主体、跨文化的系统工程。有效治理既要在技术层面提升透明、可解释与可追溯能力，也要在制度层面明确责任链条与救济通道，更要在伦理与文化层面守住公共价值与社会信任。唯有在“原则—机制—实践”之间形成可迭代闭环，通用智能的技术增益才可能转化为可持续的公共能力。

---

* 本章由复旦大学杨庆峰教授牵头，林俊宇、张瑞瑗，合作撰写。

杨庆峰教授为复旦大学科技伦理与人类未来研究院教授、研究员，复旦大学哲学学院博士生导师。研究领域包括技术哲学、记忆哲学、数据伦理与人工智能伦理等。2013—2014 年美国达特茅斯学院访问学者，2017 年澳大利亚斯威本科技大学访问学者。全国应用伦理教指委秘书长、上海市自然辩证法研究会副理事长。国家社会科学基金重大项目“当代新兴增强技术前沿的人文主义哲学研究”首席专家。

林俊宇为复旦大学科技伦理与人类未来研究院高级工程师；张瑞瑗为复旦大学哲学学院科学技术哲学专业博士研究生。

本章提要：

- AGI 讨论存在概念与现实的双重歧义：既需区分 AGI 与 ASI 等层级关系，也需从“路径”而非单点事件理解通用能力的演进，从而避免治理议题被夸大或被弱化。
- AGI 与 AIGC、Agent、具身智能构成“部分—整体、当下—未来”的协同谱系：内容生成、自治执行与物理交互共同推动通用能力从表达走向行动与存在。
- 伦理风险不仅来自数据偏见与隐私侵犯，更在生成式能力加持下扩展为虚假与幻觉信息、认知操控与价值观重塑，并因技术非中立性而呈现结构性不平等与自我强化特征。
- 决策自动化将责任、信任与权力推入重定义阶段：算法黑箱、主体分散与自动化偏差可能造成责任真空与信任崩塌，必须以可解释、可追溯与救济机制守住程序正当底线。
- 面向 AGI 时代的数据治理需从“合规”转向“信任构建”，以多样性、透明性、可解释性与可追溯机制贯穿全生命周期，形成可审计的责任链。
- 治理需要从原则走向机制：通过文化—制度—技术—伦理（CITE）四维协同，将公平、对齐与多元参与嵌入模型研发与部署流程，并以公众参与和弱势群体赋权提升治理的可接受性与可持续性。

第二篇

人文社会科学智能关键领域发展前沿与趋势

第五章　从观测到决策的跃迁：复杂系统与社会仿真范式变革*

在数字化、网络化与智能化深度融合的背景下，人类社会正加速演化为典型的复杂适应系统。无论是舆论传播、市场运行，还是宏观社会行为，均表现出高度不确定性与复杂性，对传统社会科学研究范式带来结构性挑战。长期以来，以问卷调查、统计分析与案例研究为代表的方法，主要依赖对历史数据的观测与归纳，难以刻画多主体交互驱动下的动态演化机制。

随着计算能力与 AI 技术的快速发展，社会科学研究正加速向复杂系统范式转型，由“事后解释”转向“过程建模与机制推演”。通过构建由大语言模型驱动的多智能体系统，并在可控的数字环境中模拟个体行为与群体互动，复杂社会过程得以转化为可计算、可重复的“数字实验场”。在这一框架下，不仅可以复现既有运行规律，还能够开展反事实推演与策略优化，推动社会治理从经验驱动走向数据驱动与模型驱动的协同决策模式，从而为政策分析与智能决策提供实验化支撑。

本章从复杂系统视角出发，梳理 AI 驱动的社会仿真方法体系及其关键应用路径。具体而言，首先介绍面向真实世界对齐的通用社会仿真框架，以 SocioVerse 为代表，阐述其在环境、用户、交互结构与行为生成等多维度上的统一建模机制；其次，分析面向制度约束与策略博弈的多智能体仿真平台，以 ProcureGym 为例，探讨 AI 在政策评估与经济决策中的应用；最后，聚焦社交媒体场景，分析多智能体仿真与可视分析融合的方法体系，揭示其在复杂社会行为理解与干预中的作用。通过上述展开，进一步呈现社会仿真从方法探索走向系统构建、从现象复现走向决策支持的发展趋势。

---

* 本节由复旦大学魏忠钰副教授牵头，刘宝副教授和陈思明研究员合作撰写。

魏忠钰教授现为复旦大学大数据学院副教授，博士生导师，数据智能与社会计算课题组负责人，主要从事大模型驱动的社会计算和多模态大模型的研究。

本章提要

- 在数字化、网络化与智能化深度融合背景下，社会系统呈现典型的复杂适应系统特征，传统以“事后解释”为主的范式难以刻画多主体交互驱动的动态演化机制，社会科学正加速由观测归纳走向过程建模与机制推演。
- 以 LLM 驱动的多智能体系统为代表的社会仿真方法，为复杂社会过程提供可计算、可重复、可干预的“数字实验场”，支持反事实推演与策略优化，为政策分析与智能决策提供实验化支撑。
- 通用社会仿真框架（以 SocioVerse 为例）的关键在于“真实世界对齐”（alignment）：在环境语境、用户分布、交互结构与行为生成四个层面建立一致性机制，是提升仿真可信性、可解释性与可扩展性的基础。
- 面向制度约束与策略博弈的多智能体平台（以 ProcureGym 为例）将政策规则、主体异质性与市场互动形式化为可计算的马尔可夫博弈，支持对机制参数的敏感性分析与反事实评估，推动政策评估从经验判断走向可验证的机制实验。
- 在开放社交媒体场景中，多智能体仿真与可视分析正在融合为“生成—分析—调控—再仿真”的人机协同闭环：仿真提供可控合成数据与过程日志，可视分析强化对结构涌现与传播路径的解释能力，共同提升对复杂社会行为的理解与干预能力。
- 社会仿真从方法探索走向系统构建与应用落地，仍需守住可信度底线：多源数据与行为对齐、可解释性与不确定性披露、伦理与隐私合规，以及可审计的日志机制，决定其能否成为面向治理的基础设施。

第六章　从识别到估计的革命：深度学习赋能的社会科学因果推断*

社会科学的核心关切往往不是“变量是否相关”，而是“何种机制导致了何种结果”，这使因果推断成为理解社会分层、解释行为差异与评估公共政策效果的基础方法。近年来，因果识别理论框架日益成熟，但在实践层面，一个更突出且常被忽视的瓶颈来自估计阶段的模型依赖性：大量方法需要对倾向得分或条件期望等关键量进行参数化建模，一旦模型设定刚性或错设，在高维协变量、非线性关系与复杂交互结构下就可能产生系统偏误，甚至使结论更多反映“模型产物”而非数据所蕴含的信息。深度学习作为高度灵活的非参数逼近器，为缓解这一瓶颈提供了新路径：一方面，通过表示学习与样本平衡思想，将处理效应估计从“控制变量/拟合方程”推进到更具数据自适应性的反事实预测；另一方面，通过生成式分布建模与反事实模拟，为中介机制刻画与“跨世界”嵌套反事实的估计提供了可操作的计算框架。本章围绕“效应估计—机制分析”两条主线，梳理深度学习与因果推断的融合逻辑，强调在坚守识别假设与因果逻辑前提下，推动社会科学因果分析由参数化设定走向分布级建模与可复核的证据生产。

---

* 本章由复旦大学社会发展与公共政策学院胡安宁教授牵头编写。

胡安宁教授现任复旦大学社会发展与公共政策学院院长，中国社会学会计算社会学分会理事长，万人计划哲学社会科学领军人才、青年长江学者、上海市领军人才。

本章提要

- 社会科学因果推断的识别框架已较成熟，但实践瓶颈日益集中于估计阶段的模型依赖性：倾向得分与条件期望等关键量的参数设定在高维、非线性与复杂交互条件下易引入系统偏误。
- 深度学习作为数据自适应的非参数逼近器，为缓解“刚性设定”带来的估计偏差提供技术路径，其价值在于提高反事实预测稳定性，而非替代因果识别理论。
- 在处理效应估计上，TARNet/CFR 等方法将因果问题重构为表示学习与样本平衡问题：通过共享表征空间与分布匹配约束提升处理组与对照组的可比性，降低外推风险并增强因果泛化能力。
- 在机制刻画上，基于归一化流等生成式框架可对中介与结果的条件分布进行分布级建模，支持对自然间接效应等“嵌套反事实”量的模拟估计，使中介分析从线性系数分解走向可逆变换下的干预、生成与比较。
- 方法融合需坚持三条底线：未观测混杂无法被模型复杂度弥补；高度灵活模型带来更高数据需求与有限样本风险；深度模型的可解释性与可审计性必须通过明确假设、稳健性检验与过程记录加以强化。
- 面向未来，关键不在“用深度学习取代传统因果方法”，而在构建兼具识别严谨性与估计灵活性的整合框架：在反事实逻辑与识别假设约束下，充分利用表示学习与生成建模能力，形成更稳健、更精细、可复核的因果证据生产链条。

第七章　从文本到多模态的拓展：计算社会科学智能发展前沿与趋势*

计算社会科学（Computational Social Science，CSS）旨在以可计算方式刻画社会结构、行为互动与制度运行，使社会研究从解释过去走向评估政策、推演风险和识别机制。进入大模型时代，文本不再只是研究材料，而成为可结构化、可建模、可复核的“计算证据”；社会证据也从文本和统计数据，拓展到图网络、时空轨迹、图像、视频、语音和传感数据构成的多模态事实流。计算社会科学由此迈向以多模态证据和复杂系统为基础的机制建模与系统评估。

本章围绕 AI 重塑计算社会科学的三类能力展开：一是规模化证据提取与知识组织，大模型降低了从政策文本、历史档案、舆情内容和学术文献中抽取事实与关系的成本；二是结构与互动的可计算表达，图神经网络、表示学习和多智能体框架使社会网络、组织关系与群体行为得到细粒度刻画；三是从解释到干预的闭环，强化学习、仿真和数字孪生式框架推动政策评估走向压力测试、分布效应识别和方案比较。

但能力扩张并不自动等于知识进步。多模态证据越丰富，越需要警惕噪声、偏差和选择机制；模型推理越强，越需要防止看似合理但不可检验的叙事。本章以“数据—理论—机制”闭环为标准，讨论 AI 赋能计算社会科学的技术路径及治理要求。

---

* 本章由复旦大学新闻学院周葆华教授，博士生吴雨晴、张凌赟合作撰写。周葆华教授为国家万人计划哲学社会科学领军人才（2021），教育部首批青年长江学者（2015）。

本章提要

- 大语言模型与多模态大模型正在重塑计算社会科学的内容分析范式：从文本主题建模、分类与生成，到图像/视频/声音/数字痕迹等多模态社会数据的理解、生成与模拟，AI 逐步成为连接社会数据与研究推理的通用“方法引擎”。

- 在文本分析层面，大语言模型显著缓解了“人工编码高成本”与“传统算法难解复杂语境”的双重困境，在主题标注、情感与立场识别、信息失序检测等任务中展现出效率与可解释性优势，但在文化语境、情感共鸣与隐性语义等仍需人工核验与理论约束。

- 面向分类与识别任务，大模型在情绪隐晦、话题极化与价值对齐情境下易出现系统性偏差，提示计算社科应用必须将偏差诊断、不确定性披露与分层校验纳入标准流程。

- 在文本生成环节，大模型可用于解释、摘要与信息提取，显著降低海量材料处理负荷并提升结构化信息供给能力，但需警惕“压缩导致的情绪与少数观点损失”、以及生成解释与结论外推带来的误导风险，强调人机协作的证据链与审阅机制。

- 多模态大模型扩展了计算社会科学的数据边界：在媒体内容编码、虚假内容与复杂语用识别、数字行为痕迹分析、城市感知与空间治理等方向提供零/少样本编码能力，并通过跨模态推理提升对复杂社会语境的识别与解释潜力。

- 多模态生成与多模态智能体为“可控社会情境生成—硅样本实验—多智能体仿真”提供新工具，使社会互动的实验化与可重复性显著增强；外部效度、模态整合深度推理与安全可控仍是关键瓶颈，需要以基准评测、伦理合规与可解释框架推动规范化落地。

第八章　从技术嵌入到伦理价值的反思：数智人文的新视角与新趋势*

数智人文不只是把传统材料“数字化”，更在技术深度嵌入之下重塑研究对象、方法逻辑与知识生产方式。本章以“技术嵌入—范式重构—伦理反思”为主线，梳理数字人文向数智人文跃迁的机制，展示文学、历史与哲学等典型领域的数智化实践如何生成可迁移的方法论经验。

生成式 AI 的介入不仅带来效率提升，也带来叙事结构、文化表达与学术伦理的系统性挑战；其背后牵涉到数字存在的本体论追问与计算理性与人文理解之间的认识论张力。因而，数智人文的健康发展必须以批判性与反思性为底座，在人机协同的新型学术生态中重建数据治理、算法透明与学术规范，推动技术进步与人文价值同向而行。本章拟从五个维度展开论述：首先审视工具、对象与方法的系统再造，继而分析文学、历史学与哲学等传统人文领域的数智化转型实践，进一步探讨生成式 AI 对叙事结构、文化表达与创作伦理的冲击与挑战，在本体论与认识论层面进行深层理论审视，最终提出重塑人文学科批判性与反思性功能的路径建议。通过多维度的系统考察，本研究试图为理解数智人文的新视角与新趋势提供一个兼具理论深度与实践关怀的分析框架。

---

* 本章由复旦大学科技考古研究院文少卿副教授牵头，复旦大学国家发展与智能治理综合实验室张以颖合作撰写。

文少卿副教授现任国家发展与智能治理综合实验室副主任，“中欧人才”（国家自然基金委，2021 年）、“兰台青年学者”（中国历史研究院，2023 年），专注分子考古、量化考古与法医考古，利用古基因组学研究中华民族谱系，探讨文明起源与民族交融的历史进程。

本章提要

- 数字人文正经历从“文献数字化—计算分析”走向“智能生成—范式重构”的跃迁，数智人文的核心不在工具升级，而在研究对象、方法逻辑与知识生产结构的系统再造。
- 在工具层面，预训练语言模型与知识图谱等技术推动人文研究由浅层统计转向深层语义计算，使计算机从“计数器”走向具备语义推理能力的分析代理。
- 在对象层面，文本、档案与文化遗产被转化为可运算的数据结构与语义网络，研究问题由线性叙事扩展到跨时空的结构关联与动态生成。
- 在方法层面，“可计算阐释”强调计算发现与意义阐释的互动：远距离阅读提供宏观导航，精读提供深度解释，二者构成互补而非替代关系。
- 生成式 AI 直接介入叙事与创作环节，带来作者性、文化同质化、算法偏见、著作权与真实性（幻觉）等挑战，促使人文学科在本体论与认识论层面重估技术的边界。
- 数智人文的战略价值在于重塑人文学科的批判性与反思性：以数据批判、算法批判与应用批判为抓手，建设人机协同的学术生态与交叉人才体系，形成兼具技术能力与人文关怀的制度化支撑。

第九章　从效率到责任的重建：AI 嵌入公共治理的应用模式与责任重构*

随着大数据、算法模型和算力基础设施的快速发展，人工智能特别是生成式人工智能正在加速嵌入公共治理全过程。从智能问答、智能导办、智能审批到辅助决策、智能推送，AI 不仅提升了政务服务的响应速度、供给精度和协同效率，也推动政府治理从被动响应走向主动服务、从经验判断走向数据驱动、从流程优化走向能力重塑。与此同时，AI 已不再只是后台工具，而是逐步进入权利实现、程序运行、责任分配和公共价值塑造等治理核心环节，深刻重构公共部门与公民、技术与制度、效率与正当性之间的关系。然而，AI 赋能公共治理并不只是技术进步的线性展开。算法“黑箱”、模型“幻觉”、智能鸿沟、组织能力退化、权责边界模糊以及公共价值偏移等风险，正在不断显现。尤其是在代理与辅助、人机协同与技术替代之间，公共治理必须审慎划定人工智能的应用边界。

本章围绕 AI 在政务服务中的应用场景、嵌入公共治理的两种模式及其责任重构问题展开讨论，力图在效率提升与价值坚守之间，探寻负责任的人工智能治理路径。

---

* 本章由复旦大学国际关系与公共事务学院郑磊教授牵头，复旦大学国际关系与公共事务学院博士生张宏、陈奕醇和杨家西合作撰写。

郑磊教授现任复旦大学数字与移动治理实验室主任、中国信息协会数字治理专业委员会副会长、国际数字政府学会中国分部联合主席、《联合国电子政务调查报告》专家组成员、上海市公共数据开放专家委员会委员。

本章提要

- 人工智能正广泛应用于智能问答、智能导办、智能审批、辅助决策和智能推送等场景，推动政务服务由被动响应转向主动、精准和智能化供给。
- AI 在提升行政效率、优化服务流程和强化数据研判能力的同时，也带来了智能鸿沟、算法偏见、数据安全和服务不平等等新问题。
- 人工智能嵌入公共治理主要呈现代理型与辅助型两种模式，二者在决策权配置、责任归属和治理正当性方面存在根本差异。
- 不同治理事项应根据规则明确性、权益影响程度、价值冲突复杂性和社会信任水平，审慎选择 AI 的介入深度，高风险场景应坚持“人机协同、人在回路”。
- 需警惕辅助型系统在实践中因自动化偏差、组织依赖和绩效导向而滑向事实上的代理型治理，进而引发责任卸载和权力异化。
- 负责任的人工智能治理，应围绕数据、模型、部署、使用、纠错和救济等环节重建全生命周期责任链，确保公共责任不能被技术或供应商替代。
- 推进公共治理中的 AI 应用，应同步完善说明义务、人工复核、申诉救济、审计监督和动态评估机制，把公平、透明、隐私、安全与公众信任嵌入技术治理全过程。

第三篇

中国高校人文社会科学智能（AI4SSH）的探索、行动与体系建制

第十章　中国高校 AI4SSH 发展的指数化观察*

AI 技术正以前所未有的广度与深度渗透至人文社会科学领域，推动研究范式由“工具性应用”走向“方法论重构”。在这一进程中，高校既是前沿探索的主阵地，也是交叉学科生态的关键组织者。如何在扩散式应用与建制化发展之间形成可持续路径，需要兼具“全景画像”与“导向性参照”的评价工具。

本章引入“中国高校 AI4SSH 指数”，以研究核心能力、发展创新潜力与社会传播能力三大维度为主轴，结合多源数据证据链，对我国高校 AI4SSH 的发展态势进行指数化观察。指数的目的在于识别结构特征、揭示能力短板与增长动能，推动融合研究从自发探索走向有组织、可评估、可迭代的体系建制。

通过“方法—画像—结论与展望”的结构，本章尝试回答三类关键问题：我国高校 AI4SSH 整体呈现何种梯度与集聚特征；哪些能力维度正在加速积累、哪些环节仍是明显短板；下一阶段应如何以制度供给、平台支撑与人才机制推进从规模积累到质量引领的跃迁。

---

* 本章由复旦大学大数据研究院赵星教授牵头，复旦大学国家智能评价与治理实验基地王译晗博士等合作撰写。

赵星教授现为复旦大学国家智能评价与治理实验基地副主任，上海市“曙光学者”、教育部教指委委员、国际信息计量学和科技评价领域权威期刊 *Journal of Informetrics* 编委，五种中文重要期刊编委。

本章提要

- 本章以“中国高校 AI4SSH 指数”为工具，面向全景式把握我国高校在 AI 与人文社会科学深度融合中的总体进展与结构特征，兼具“评估基准”与“发展导向”双重功能，为学科布局、资源配置与生态构建提供量化参照。
- 指数体系覆盖“研究核心能力—发展创新潜力—社会传播能力”三大维度，并细化为多层指标结构，强调从范式融合、组织协同与生态构建等视角识别高校 AI4SSH 综合能力与可持续发展潜力。
- 总体态势呈现“体系初构、梯次分明”的结构化发展格局：研究产出规模与本土融合推进较快，形成明显的集聚效应与区域协同骨架，但不同高校、不同区域之间的发展强度与能力结构差异显著。
- 生态建设正进入“实体化、建制化”加速阶段：交叉研究机构的设立与平台化布局成为推动 AI4SSH 从零散探索走向体系化发展的关键变量，但制度性支撑与长效机制在多数高校仍处起步阶段。
- 社会传播与公共影响呈现“圈层化”分布：学术成果向公共议程与社会服务的转化效率仍有提升空间，如何将研究能力转化为政策支撑与社会赋能，是下一阶段需要补齐的链条。
- 面向未来，亟待推动范式融合由“应用跟随”迈向“创新引领”；构建“内外联动、制度保障”的可持续生态；强化需求驱动的社会服务与话语体系；加快培养“厚基础、强交叉、善协同”的复合型人才队伍。

第十一章 AI4SSH 基础设施和代表性项目建设现状与展望*

人工智能对科研活动的影响，正从单点工具应用扩展为对研究基础设施、研究流程、创新组织方式的系统性重塑。在这一背景下，AI4SSH 基础设施的建设已成为高校推进人文社会科学数智化转型的关键抓手。它既决定数据、算力、模型与工具能否形成可持续的校内公共供给，也决定跨学科协作能否从项目合作走向长期建制。

本章围绕“数据底座与流通机制—融合算力与研究装置—模型/智能体与工具链”三条主线，在系统梳理国内外政策导向与概念演进的基础上，结合国内高校公开实践，揭示出 AI4SSH 基础设施从“资源汇聚”走向“可计算组织”、从“计算支撑”走向“可实验研究环境”、从“模型接入”走向“研究全过程能力体系”的阶段性跃迁。

AI4SSH 基础设施并非硬件堆砌。对人文社会科学而言，其核心在于形成可复用、可互操作、可审计的公共能力：数据的可计算与可流通、算力与环境的可用与可持续、模型与工具的可控与可扩展，以及支撑人机协同知识生产的组织机制与治理规则。

---

* 本章由复旦大学大数据学院张计龙研究馆员牵头编写。

张计龙研究馆员目前担任教育部教指委委员、复旦大学国家发展与智能治理综合实验室副主任、上海市科研领域大数据联合创新实验室主任、复旦—阿法迪共建智慧图书馆学研究中心常务副主任等，研究领域包括智慧图书馆、科学数据管理等。

本章提要

- AI4SSH 基础设施建设已成为全球与我国科技战略的重要组成部分，国际战略引领、国家顶层统筹、地方政策响应与高校平台落地形成协同联动的推进格局。
- 概念层面，研究基础设施正在向“数智学术基础设施”演进。其核心突破在于以“数据—算力—模型和工具链”的系统集成为基础，将平台科研、人机协同与知识共生转化为基础设施能力，并以公共性、可复用性与可持续性为发展导向。
- 数据底座建设三大趋势：由分散资源库走向校级底座与平台群整合；数据对象由结构化统计扩展到多源异构与多模态知识对象；数据流通由静态共享走向嵌入全流程的机制化流转和 AI-Ready。
- 算力基础设施由单一 HPC 平台走向超算、智算与通算融合的综合底座；社会仿真与多智能体实验平台兴起，为相关研究提供了“可实验”的研究空间。
- 模型、智能体与工具链正构成面向研究全过程的能力体系：通用模型校内化重组、垂域模型学科化深入、工具链平台化供给，共同推动 AI 能力从“可用”向“可嵌入、可复现、可治理”转变。
- 平台同质化、算力统计口径不一、资源可及性不足、伦理与合规边界不清、平台贡献评价缺位等问题仍制约基础设施的规模化外溢。未来应以互操作标准、评测与审计机制、服务化团队与可持续治理规则为抓手，推动 AI4SSH 基础设施建设从规模扩张走向高质量公共供给。

第十二章 中国高校 AI4SSH 组织、推动与管理典型案例*

AI 赋能人文社会科学从“点状探索”走向“体系建制”，关键不在单项技术是否先进，而在组织方式、平台能力、激励机制与治理规则能否匹配交叉融合的复杂性。高校既是知识生产主阵地，也是交叉生态组织者：需要把算力与数据底座、跨学科协同机制、工程化交付能力与社会服务转化链条同时搭建起来，才能把技术增益转化为可持续的公共能力。

本章选取复旦大学、武汉大学、中国人民大学、北京师范大学四个案例，覆盖“平台型基础设施建设”“场景牵引的数智人文工程”“面向治理的复杂决策系统”“面向教育改革的循证教研闭环”等不同路径。四校实践共同指向一条可迁移的经验链条：以真实公共任务牵引跨学科协同，以平台化底座沉淀数据与模型资产，以制度化机制保障质量、透明与问责，并通过联盟网络、培训体系与开放合作实现扩散。

本章旨在以案例方式呈现 AI4SSH 组织创新的可复制要素，为高校在“十五五”背景下推进学科智能化转型、建设文科实验室与交叉平台、完善科研管理与激励制度提供参考。

* 本章由复旦大学国家发展与智能治理综合实验室汤维祺、王凌翔合作编写。素材由复旦大学文科处、高校哲学社会科学实验室联盟等提供。

本章提要：

- AI4SSH 的体系化推进需“组织能力建设”，包括稳定的人才梯队、跨院系协同机制、工程化支撑队伍以及可持续运行的平台规则。
- “场景牵引”是跨学科协作的最有效粘合剂：以教育改革、文化遗产保护、数字政府与国家治理等复杂公共任务为抓手，可将学术解释、工程实现与社会服务嵌入同一闭环。
- 平台化底座决定可复制性：算力、数据、模型与工具链的共享机制及合规规则，能够把协作意愿转化为协作能力，降低跨团队成本并提升成果可复核性。
- 制度化学术治理与质量控制是“规模化”前提：学术委员会、子实验室矩阵、版本化数据治理、可审计日志与第三方评测等机制有助于把研究过程前置为可管理的证据链。
- 扩散与生态建设依赖联盟网络与人才体系：通过暑期学校、学术年会、培训课程与共建共享平台，形成共同语言与工作流模板，推动从个案到规模化应用。
- 高校经验共同指向“基础设施—规范—产出”的联动逻辑：以公共能力为导向的制度供给能够同时促进创新活力与风险可控，并提升成果进入公共决策与社会服务的可解释性。