

生成式人工智能赋能下的新闻传播学研究方法革新探析

牟 怡 倪盈盈 蓝剑锋 马文斌

【内容摘要】本文探讨人工智能尤其是生成式 AI 对科学研究范式的深刻影响及其在社会科学中的应用潜力。随着 AI 技术的快速迭代，科学界对数据密集型研究范式的认可日益广泛。托马斯·库恩的范式理论表明科学研究模式会随科学本身发展和外部环境变化而转变。尽管自然科学领域已广泛采纳 AI 作为研究方法，但社会科学尤其是新闻传播学仍主要将 AI 视为研究对象而非方法工具。本研究聚焦新闻传播学定量实证研究的三大方法，分析其传统局限性，并探讨生成式 AI 如何为这些方法赋能，特别是在学术研究工具开发、数据共享及伦理治理方面的潜力。本研究旨在促进社会科学研究方法创新，推动学科融合发展，为社会科学研究范式转型提供新思路。

【关键词】人工智能；科学研究范式；生成式 AI；新闻传播学；数据密集型研究

Exploring the Innovation of Journalism and Communication Research Methods Empowered by Generative Artificial Intelligence

MOU Yi, NI Yingying, LAN Jianfeng, MA Wenbin

Abstract: This paper examines the profound impact of artificial intelligence, especially generative AI, on scientific research paradigms and its Application potential in social science research. With the rapid iteration of AI technology, the scientific community increasingly recognizes the data-intensive research paradigm. Thomas Kuhn's paradigm theory suggests that scientific research models evolve with both scientific development and changes in the external environment. Although natural sciences have widely adopted AI as a research method, social sciences, particularly journalism and communication studies, still primarily view AI as a research object rather than a methodological

tool. This study focuses on three major quantitative empirical research methods in journalism and communication, analyzing their traditional limitations and exploring how generative AI can empower these methods, especially in academic research tool development, data sharing, and ethical governance. The research aims to promote innovation in social science research methods, advance disciplinary integration, and provide new perspectives for the transformation of social science research paradigms.

Keywords: Artificial Intelligence; Scientific Research Paradigm; Generative AI; Journalism and Communication; Data-Intensive Research

一、引言

人工智能（以下简称 AI）的全面崛起对科学研究领域产生了重大影响，科学界对此也给予了正式认可。三位 AI 领域专家分别获得 2024 年诺贝尔物理学奖和化学奖，标志着人类科学研究基本模式的转变，AI 推动了科学研究第四范式的发展。

“范式”（paradigm）这一概念最初由美国著名科学哲学家托马斯·库恩（Thomas S. Kuhn）于 1962 年在《科学革命的结构》一书中提出，指的是常规科学所赖以运作的理论基础和实践规范，是从事某一科学的科学家群体所共同遵从的世界观和行为方式^[1]。新范式的产生，一方面是由于科学研究范式本身的发展，另一方面则是由于外部环境的推动。AI 的迅速崛起促使新问题不断产生，使得原有的科学研究范式受到各个方面的挑战。图灵奖得主詹姆斯·格雷（James Gray）提出科学研究的第四类范式“数据密集型科学研究”，即以数据为基础，融合实验、理论和计算机模拟为一体的数据密集型计算。这是继第一范式“实证科学”、第二范式“理论科学”和第三范式“计算科学”后的又一大研究范式^[2]。微软研究院副总裁托尼·海（Tony Hey）等人编撰的著作《第四范式：数据密集型科学发现》，邀请国际著名科学家对数据密集型科学发现的理念、应用和影响进行了全面分析，让这一理念产生了广泛影响。

然而社会科学领域，尤其是新闻传播学领域，目前过多将 AI 视为研究对象，而非研究方法。生成式人工智能（Generative AI）作为 AI 领域的重要分支，近年来在自然语言处理、图像生成、代码自动化等方面取得了革命性进展，其应用已逐步渗透到教育、医疗、媒体等社会关键领域。然而，在社会科学领域，对生成式 AI 的应用仍处于探索阶段，尤其是在辅助学术研究工具开发、数据共享以及伦理治理方面尚有待深入挖掘。因此，本文就新闻传播学定量实证研究的三大研究方法展开讨论，从传统方法的局限开始，进而探讨这些方法在生成式 AI 赋能之下如何焕发出新的生命力。

二、从问卷调查到计算模拟

问卷调查法作为社会科学研究的经典方法，在计算技术尚未成熟的历史阶段发挥了基础性作用。但随着研究问题复杂性的提升，其自身固有的局限性日益凸显，集中体现在获取数据的时间静态性与样本代表性挑战上，以及因果推断能力不足等问题。这些问题目前或可通过计算模拟方法进行一定程度的规避。

（一）传统问卷调查法的局限性

问卷调查法本质上是对特定时间社会现象的“静态截面”捕捉，缺乏对连续动态过程的追踪能力。比如在流行病学研究中，当研究者试图通过问卷调查了解公共疫情传播初期的公众风险认知时，数据收集与分析过程的时间延滞往往导致结果发布时公众认知已发生质变，严重影响研究发现的时效性和应用价值。又如西方国家大选之前展开的多轮民意测验（opinion poll），亦是横截面（cross-section）数据，反映的是调查进行之时民众的态度。

即使采用面板研究设计（longitudinal panel design，对同一研究样本进行多次重复测量的纵向研究方法），其时间分辨率仍受限于调查周期，无法精细刻画短时间内群体行为演变的微观过程。斯莱特（Slater）等人的媒介效果研究采用了两波面板设计，研究青少年对媒体暴力内容的选择性接触与攻击性态度之间的相互强化关系，仍然无法说明媒介接触与态度变化之间的即时互动关系，只能推断其长期趋势。这种静态的数据收集限制了研究者对动态传播过程的深入理解^[3]。

同样样本代表性问题始终是问卷调查法难以避免的方法论挑战，特别是在跨文化、跨区域比较研究中。亨里希（Henrich）等人在《行为与脑科学》上发表的研究指出，大量心理学研究过度依赖西方、受教育、工业化、富裕和民主社会（WEIRD）样本，其结论的普适性特别是在发展中国家应用时值得质疑^[4]。以世界价值观调查（WVS）为例^[5]，该项目自1981年由密歇根大学政治学家罗纳德·英格尔哈特（Ronald Inglehart）发起以来，已完成七波调查，覆盖全球近100个国家和地区，调查主题涉及政治态度、宗教信仰、性别观念、幸福感、社会容忍度等广泛价值观领域。尽管提供了不同文化背景下价值观变迁的纵向和横向数据，但样本在不同国家间的抽取范围、概念的文化等值性难以完全保证，不同受访者对问卷题项的理解与反应模式存在文化差异，如社会期许效应（social desirability bias）在集体主义文化中可能更为显著，这些问题都直接影响了研究结果的可靠性。

此外，问卷调查的样本规模受到人力、物力等资源约束，对于需要小概率事件或特殊人群的研究，往往面临样本获取困难的问题。以政治传播学研究为例，对政治极化现

象进行实证分析时经常面临采样困境。持有极端政治立场的人群在总体人群中比例较低，且参与调查的意愿有限，这使得相关研究难以获取足够样本量，从而无法进行合理的统计推断。在新媒体研究领域，这一问题表现得尤为突出。如巴克希（Bakshy）等人2015年在《科学》杂志发表的关于Facebook用户信息获取行为的研究虽具开创性^[6]，但存在明显的样本代表性问题。他们研究分析了1000多万名美国Facebook活跃用户的新闻浏览习惯，认为用户的信息过滤气泡主要源于个人选择而非算法推荐。然而，这一研究样本仅限于每日登录且经常与政治内容互动的用户，完全忽略了占美国人口近30%的低频或非Facebook用户群体，包括老年人、低收入群体和农村居民。这些被排除在外的群体恰恰是数字鸿沟研究中的关键对象，他们的信息获取行为可能与样本中的活跃用户有本质区别。此外，研究仅关注单一平台Facebook，无法捕捉用户跨平台的信息获取行为，使其结论难以推广至整体社交媒体生态或更广泛的信息环境。这一例子清晰展示了新媒体研究中常见的样本偏差问题，即以平台活跃用户行为为基础的发现难以代表整体社会中多元化的媒体使用情况。

（二）计算模拟方法的理论优势与实践价值

计算模拟方法为突破传统调查研究的限制提供了创新路径，在数据时效性、样本规模和理论建模等方面展现出显著优势，实现了从“毫秒”到“年”的时间单位都可捕捉的连续观察能力。

2007年，沃茨（Watts）和多兹（Dodds）在《消费者研究杂志》（JCR）发表了突破性研究，他们构建了一个包含10000个节点（代表个体）的大规模网络模型，精细模拟了信息在社会网络中的传播过程^[7]。与传统研究仅能观测信息传播的起点和终点不同，该模型精确记录每一条信息从发出到扩散的完整时间轨迹，包括信息通过哪些节点传播、每个节点接收和转发信息的时间点，以及信息传播过程中的衰减和放大效应。这种精细化的数据使沃茨和多兹能够观察到传统“两级传播”理论难以观测的“级联效应”（cascade effects）动态过程。传统“两级传播”理论强调意见领袖在信息传播中的中介作用，认为大众媒体信息首先影响意见领袖，再由意见领袖影响普通公众。然而沃茨和多兹的模拟表明，在大多数情况下，信息的广泛传播并非主要依赖少数高影响力的意见领袖，而是通过大量普通个体形成的“关键人群”（critical mass）共同作用产生的级联效应。他们成功发现信息传播通常表现为一种临界现象——当接收到信息的易感个体达到一定比例时，信息传播会呈现出爆发式增长。这种非线性的传播动态难以通过问卷调查捕捉，因为问卷调查通常只能提供静态“快照”，无法记录完整的传播链条。

此外，问卷调查研究通常依赖统计相关性分析，结论推断主要建立在变量间相关关系的基础之上。然而，正如统计学的基本原则所指出的——相关不等于因果。即使采用

结构方程模型等高级统计技术，问卷数据分析仍难以确证因果关系，特别是在存在内生性问题（如遗漏变量）的情形下。

计算模拟通过“生成式”的建模方法提供了因果推断的新思路。不同于统计模型对现有数据的拟合，计算模拟从理论假设出发，构建微观行为与交互规则，再生成宏观的系统行为，进而检验理论假设的合理性。阿克塞尔罗德（Axelrod）和哈蒙德（Hammond）的文化传播与种族主义研究是这一方法论创新的典型案例。该研究针对“为何种族主义在人类社会中普遍存在”这一社会科学核心问题，构建了基于行为体的计算模型（agent-based model）^[8]。阿克塞尔罗德等人通过设计“标记—协作”的计算模型，即个体行为体倾向于与具有相似标记（如种族特征）的个体互动，将协作经验反馈至标记偏好中。在无须假设先天种族偏见的情况下，通过调整模型参数——如种族标记可辨识度、协作收益比率、群体流动性等，探索了不同社会条件下种族主义形成的临界阈值。模型预测在高度流动的社会中，即使存在微弱的“标记—协作”偏好，也能导致显著的种族分离。而在稳定社区结构中，邻里效应会部分缓解这一趋势。

这个理论预测与历史社会学数据高度契合，与欧洲城市化进程中种族隔离现象的演变轨迹基本吻合。可以说，计算模拟使阿克塞尔罗德和哈蒙德能够清晰观察微观个体行为到宏观种族现象之间的因果链。无独有偶，谢林（Schelling）^[9]的居住隔离模型也揭示了微观个体偏好与群体结构之间的因果机制——个体仅持有 20%—30% 的同类偏好，仍可在群体层面形成 80% 以上的居住隔离效应，这项发现超越了传统相关分析的解释深度。

其次，计算模拟方法支持反事实推理，通过改变参数可以模拟“如果 A 不发生，B 会怎样”的假设情境，为政策评估提供更科学的依据。如冈萨雷斯—拜隆（González-Bailón）等人^[10]在社会运动研究中通过操控网络结构参数，发现当早期采纳者比例超过总体的 5%—7% 时，社会抗议行为通常会触发系统性扩散，这项研究结果为预测社会稳定的临界点、评估候选人政治提案的传播效能提供了精确参考。

另外，社会系统的一个核心特征是“涌现性”（emergence），即系统性的宏观现象并非简单地由个体行为线性叠加产生，而是通过个体间的非线性交互作用涌现出来。问卷调查建立在线性归约思维的基础上，通常将复杂现象分解为可测量的变量，再通过统计模型探究变量间的线性关系。这种方法在研究简单社会现象时有效，但面对复杂系统中的涌现现象时，解释力度相对不足。在传播学领域，特别是社交媒体上的信息爆炸、舆论极化、谣言传播等社会现象中，传统问卷调查能发现个体的信息接收偏好与其社会网络位置存在关联，但难以解释为何小范围的偏好差异会导致整体大规模社会结果出现，如“信息茧房”。计算模拟则能够通过设计个体的信息选择偏好和社交网络结构，模拟出信息茧房的形成过程，揭示出个体行为到宏观现象的涌现机制。

米勒（Miller）和佩吉（Page）^[11]在《复杂适应系统》一书中指出，计算模拟使研究者能够超越传统社会科学的均衡思维，探索社会系统中的非线性动态过程、突现行为和路径依赖现象。该范式转向的学术价值在森托拉（Centola）^[12]发表于《科学》的研究中得到进一步验证。森托拉通过精心设计的网络实验，系统比较了不同网络拓扑结构（小世界网络、无标度网络等）对健康行为传播的影响机制，发现与流行病学模型预测相反，个体行为采纳的社会传染需要“多源强化”，因此在高聚类（即个体的朋友之间也相互认识的紧密社交圈）的网络中传播更为有效。森托拉的研究发现挑战了传统流行病学模型的预测逻辑：虽然病毒传播只需要一次接触就能感染，但社会行为的采纳需要个体从多个社会联系那里接收到一致的信号才会发生改变，这被称为“多源强化”机制。因此，在高聚类网络中，当某种行为开始传播时，个体更容易同时从几个相互关联的朋友那里听到相同的建议或看到相同的行为示范，这种社会压力和验证效应使得行为采纳比在连接分散的随机网络中更加有效和稳定，这也解释了为什么现实中很多健康行为推广在紧密社区中比通过广泛但松散的社交网络更成功。

三、从内容分析到数据挖掘

（一）传统内容分析法的局限

内容分析（content analysis）是新闻传播学科中的经典研究方法。传播学先驱贝雷尔森（Berelson）将其定义为“一种对传播内容进行客观、系统和定量描述的研究技术”^[13]。传统内容分析的精髓在于，通过一套严谨的操作规程，将纷繁复杂的非结构化媒体文本（如新闻报道、访谈记录）转化为可进行统计分析的结构化数据，进而揭示传播内容的内在特征、演变趋势与潜在规律。其操作核心高度依赖于研究者的精心设计与人工执行：研究者须预先制定详尽的编码手册（coding scheme），对分析单位、分析类目和判断规则进行清晰界定，以确保分析过程的标准化与可重复性；为控制主观偏差，通常需要组织多名编码员进行独立的平行编码，并通过计算编码员间信度系数来检验结果的客观与可靠程度。最后再通过语义判断、频次统计、主题归纳等步骤，将质性的文本内容系统性地转化为量化的数据变量，为后续运用统计学进行假设检验和规律探索奠定基础^[14]。

然而，随着互联网以及移动社交媒体的普及，我们进入了一个信息爆炸性增长的时代。面对海量、瞬息万变且形态多样的传播内容，严谨但缓慢的传统内容分析方法暴露出其固有的局限性。其瓶颈主要体现在：人工编码耗时巨大、成本高昂，其结果易受到研究者个人理解与主观判断的影响；面对社交媒体上难以计量的帖子和评论，传统方法在处理规模上心有余而力不足；经典内容分析多聚焦于纯文本，对于图文、音视频等多

模态内容，以及用户间的互动行为模式，难以进行有效的整合分析，无法全面还原复杂的媒体生态。

（二）数据挖掘：自动化、规模化与多维度的突破

为应对上述挑战，数据挖掘（data mining）技术应运而生，为新闻传播研究带来了革命性的突破。数据挖掘是从海量数据中自动萃取有价值的、先前未知的模式和知识的技术集合。相较于传统内容分析，数据挖掘的核心优势在于其强大的自动化处理能力与规模化分析潜力。借助机器学习算法，研究者能够对数百万乃至数千万规模的媒体内容进行快速、深入的分析，从而发现仅凭人工观察难以察觉的深层模式与潜在关联^[15]。

数据挖掘的引入，标志着新闻传播研究正从经验观察范式向数据驱动范式发生根本性转变^[16-17]。数据挖掘通过采集、清洗、分析及可视化等一系列流程，对非结构化或半结构化的数据进行深度处理与模式识别，其核心特征包括：自动化、多维度和整合性等。数据挖掘能够高效整合多种维度的数据，构建起全面、立体的传播分析框架，并最大限度减少人为干预，使得研究结论更具客观性^[18]。

总的来说，数据挖掘的引入推动了新闻传播研究的“计算转向”^[19]。它使研究者能够借助算法模型，对信息传播路径、用户接受模型、舆论影响力演化等复杂过程进行系统性分析，并在此基础上催生方法论的创新，为探索新的理论视角和研究路径提供坚实的技术支撑。

（三）生成式AI技术赋能：重塑分析的深度与广度

内容分析与数据挖掘的演进并非相互割裂，而是在人工智能技术的催化下走向深度融合。如果说内容分析侧重于语义的精确理解，数据挖掘侧重于数据的系统探索，那么人工智能则成为连接二者的桥梁，创造出一种更为系统、高效和智能的研究路径——基于人工智能的数据驱动传播研究^[20]。

1. 自动化编码：从“模型微调”到“零样本”

自动化编码是AI赋能内容分析最直观的体现。早期，研究者多依赖主题模型（如LDA）和情感词典，但其效果受限于对语境的把握不足^[21]。随后，以BERT（Bidirectional Encoder Representation from Transformers）为代表的预训练模型带来了第一次飞跃，它通过在大规模语料上的预训练，能够理解复杂的语义关系，研究者只须针对特定任务（如新闻主题分类、价值观识别）进行“微调”（fine-tuning），即可达到甚至超越人工编码的准确率^[22]。

自2020年OpenAI发布拥有1750亿参数的GPT-3模型以来，我们进入了大型语言模型（large language model, LLM）的纪元^[23]。GPT-3.5、GPT-4等后续版本在长文本

理解、逻辑推理和“少样本学习”（Few-shot Learning）能力上实现了惊人的提升。与需要大量标注数据进行微调的 BERT 不同，大语言模型凭借其强大的零样本能力（zero-shot learning），可以直接根据研究者在提示词（Prompt）中给出的自然语言指令和定义完成分类任务。这意味着研究者无须掌握复杂的编程或机器学习技术，即可运用最前沿的 AI 能力，极大地降低了技术门槛^[24]。从微调到零样本，人工智能帮助内容分析的编码完成了质的飞跃。

2. 智能语义理解：从“表层识别”到“深层洞察”

人工智能技术将内容分析从表层的词频统计推向了深层的语义挖掘。传统的“词袋模型（bag-of-words model）”无法分辨词序和语境，而现代神经语言模型，尤其是 LLMs，能够精准捕捉词汇在特定语境下的微妙含义。

例如，在一项对意大利医疗健康主题的社交媒体帖子进行分类的研究中，研究者试图区分讨论不同疾病（如糖尿病、结肠癌、胶质母细胞瘤）的用户群体。研究发现，传统的词袋模型完全无法有效表征用户，其分类结果几乎等同于随机猜测。然而，LLM 则能够为同一个词语（如帖子中出现的“colons”一词，在不同语境下可能指代结肠镜检查或结肠本身）根据其所在句子的不同而生成独特的、情境化的向量表征，从而深刻理解文本的真实意图。最终，像 Mistral 这样的大语言模型达到了 99.4% 的分类准确率^[25]。这种能够有效区分不同癌症讨论组用户的能力，证明了现代 LLMs 在理解语言复杂性和相似主题上的强大能力，也为理解复杂的社会话语提供了前所未有的深度和精度。

3. 多模态内容处理：跨越符号的边界

随着图像和视频在新闻传播中地位的日益凸显，多模态内容分析成为新的前沿。研究范围得以从纯文本扩展到图像、音频、视频等多种媒体形式。以 OpenAI 的 CLIP（contrastive language-image pre-training）模型为代表的技术，通过大规模图文对比训练，成功实现了文本与图像在高维向量空间的对齐^[26]。这意味着研究者可以“用语言检索视觉”。

此外，以 DALL·E、Sora、Imagen 等为代表的多模态生成模型，则为符号学和叙事结构研究打开了全新的维度。研究者可以通过生成式 AI 来测试不同图文组合的传播效果，这在以往是令人难以想象的。这些技术的整合，使研究者能对新媒体内容进行全方位的深度分析，从而捕捉其完整的传播意义。

四、从传统实验到计算实验

传统实验法在新闻传播学研究中占据重要地位，根本原因在于其在高度可控的环境下揭示媒介信息对认知^[27]、情感与行为因果机制的能力^[28]。从班杜拉（Bandura）经典的“波波玩偶”观察学习实验^[29]到近年虚假信息传播效应的因果检验^[30]，实验法历

来被视为验证传播理论与模型的“金标准”。然而，当下媒介使用场景由单一封闭转向开放、多屏、多模态的分布式空间，用户行为呈现高度情境化、瞬时化，且深受算法塑造^[31]。因此，新闻传播实验研究面临双重挑战：一是如何在复杂、多变的环境中有效控制变量，二是如何真实再现信息传播的生态场景。

（一）传统实验法的局限

实验法在新闻传播学的扎根可追溯至 20 世纪中叶，社会心理学与大众传播研究的交叉奠定了其实证基础^[32]。在“有限效果”学派盛行的阶段，实验成为检验媒介影响因果机制的关键方法。无论是麦奎尔（McGuire）提出的说服信息加工模型^[33]，还是班杜拉针对观察学习的经典实验，均在高度控制的实验室环境下展示了媒介暴露对认知态度与行为的直接影响。进入 80 至 90 年代，实验法在新闻框架、议程设置、启动效应等领域广泛应用^[34]。研究细化到刺激材料的微观元素，如标题措辞、图像构图、报道顺序，多模态信息刺激（视觉、听觉、文本）被引入，以更加精准地测量信息处理路径^[35]。在中国，21 世纪初新闻传播领域开始引进实验法。研究主题涵盖虚假信息识别与抵制^[36]、媒介暴力对青少年行为影响^[37]、算法推荐对信息选择偏好的干预^[38]。这些研究多数采用高控制的实验室环境，强调因果推断的稳健性。

在方法论上，传统实验法具有以下三大核心环节：首先，刺激材料设计（stimulus design）。研究者根据理论假设，人工构建新闻文本、视频、社交媒体帖子等实验刺激，确保在自变量操控的同时，保持其他信息特征一致。例如，在虚假新闻研究中，新闻真实性为自变量，版式、字数、视觉风格均保持相同，以排除混杂变量。其次，反应变量测量（response measurement）。其涵盖行为指标（点击率、停留时间）、生理指标（眼动追踪、心率、皮肤电反应）以及态度量表（问卷、语义差异法）。近年来，实时大脑成像技术（如 FMRI、EEG）被引入，揭示认知加工的神经机制^[39]。最后，组间对比与统计推断（Between-/Within-subjects Comparison）。通过随机分组与对照设计，比较不同实验条件下因变量的变化，采用方差分析（ANOVA）、协方差分析（ANCOVA）等统计方法，得出因果效应。

实验法的最大优势在于能够通过排除外部干扰，实现因果推断的严密性。然而，尽管传统实验法在因果推断方面具备无可替代的优势，但其局限性也日益显著。这包括：（1）外部效度不足。实验室环境与真实媒介使用情境存在显著差异^[40]。实验受试者的注意力状态与信息处理方式，往往与其日常媒介行为不符，影响结果的可推广性。（2）样本代表性受限。传统实验多依赖高校学生便利样本，该群体在年龄、教育背景及媒介习惯上同质性强^[41]，难以反映广泛受众的多样性。（3）情境动态性缺失。现实媒介环境中，用户行为动态且受算法和社交网络结构深刻影响。传统实验多为单次静态刺激，无

法模拟这种复杂的互动过程。(4)成本与时间消耗高。多条件、多轮次设计下,受试者招募与管理成本显著,限制了实验规模与频次。(5)伦理挑战日益严峻。部分实验涉及欺骗(deception)或隐私数据采集,如未告知的浏览行为跟踪^[42],在数据伦理日益严格的当下引发争议。这些局限性在数字化、平台化的媒介生态中尤为明显,推动了网络实验、自然实验及混合方法的兴起,也为AI驱动实验提供了现实需求和技术条件。

(二)网络实验的崛起与转型逻辑

互联网技术的兴起,为新闻传播学实验研究带来首次重大“空间规模”跳跃。自21世纪初,随着宽带普及和网页交互功能提升,实验研究突破实验室物理局限,进入线上实施阶段。网络实验依托在线调查系统、交互网页及社交平台插件,实现了跨地域、多样化样本的快速招募与实验控制。美国计算社会学家马修·萨尔加尼克(Matthew J. Salganik)则在《计算社会学:数据时代的社会研究》一书中对网络实验方法进行深入阐释^[43]。

首先,此转变体现在样本来源扩展上。传统实验常依赖校园便利样本,而网络实验则通过众包平台(Amazon Mechanical Turk、Prolific)、在线调查工具(Qualtrics、问卷星)和社交媒体嵌入(Facebook A/B测试)等多渠道获取更广泛、多元的受众。这不仅扩大了样本规模,还提升了外部效度。

其次,网络实验突破了刺激材料单一化限制。多媒体技术允许实验呈现文字、图像、音频、视频,甚至沉浸式虚拟场景,更贴近真实媒介环境。例如,研究虚假新闻传播时,可构建模拟社交媒体界面,嵌入虚构新闻帖,并通过算法动态调整信息流顺序,考察用户选择与认知变化^[44]。

网络实验虽然在规模和情境模拟上实现突破,但局限依旧明显。首先,实验控制力下降。网络环境中硬件差异、网络延迟、屏幕尺寸不一以及外部干扰,均削弱刺激一致性,增加噪声。其次,参与者注意力与数据真实性难以得到保障。受试者可能并行多任务处理,存在挂机、刷单行为^[45],且网络匿名性导致身份虚报,增加样本质量筛选难度。

更深层的挑战则源自信息生态的动态性与算法驱动性。数字传播高度依赖个性化推荐与实时更新,而网络实验的静态设计难以捕捉由算法引发的动态交互及其对认知、情绪、行为的连锁影响。Meta(原Facebook)2014年的“情绪感染”实验揭示了算法对用户情绪潜移默化的影响,同时暴露出重大伦理隐忧:大规模采集用户数字足迹和情绪反应,将面临知情同意和数据保护难题^[46]。

总体而言,网络实验是传统实验的技术延展和范式渐进转变。它保留因果推断逻辑,扩展样本与情境维度,但面对复杂动态的信息生态时,控制假设与测量手段仍显不足。

当研究对象从“人—媒介”向“人—机—算法”系统转变时，传统实验框架的适用性受到限制。

这正为人工智能技术的深度介入创造契机。AI可利用机器学习从大数据中筛选高质量样本，避免无关或虚假参与者；在实验中实时生成个性化刺激（如自动生成不同立场的新闻、图像、视频）；利用自然语言处理和情感计算实时监测受试者反应；后期运用深度学习分析多模态数据（点击流、面部表情、语音、文本），揭示传统统计难以捕捉的复杂机制；更具未来感的是，AI可作为“虚拟受试者”或“交互主体”与人类共同构成实验生态，研究人机协同与冲突^[47]。贝尔（Bail）的研究指出，生成式AI对问卷研究、在线实验、自动内容分析和agent-based模型等方法有显著提升潜力，促进社会科学研究方法的现代化^[48]。

然而，AI驱动实验绝非传统实验或网络实验的简单替代，而是一场可能引发范式跃迁的革新，研究者须在技术创新与伦理规范间找到平衡。网络实验解决了“实验能否走出实验室”的问题，AI驱动实验则挑战“实验能否真实还原乃至预测人—机—算法交互的复杂现实”。

（三）基于智能体的算法测试方法创新

数字媒体时代，算法系统已成为信息传播的关键中介，从搜索引擎的结果排序到社交媒体的内容推荐，算法深刻重塑着信息流动的路径与格局。然而，这些算法系统通常以“黑箱”形式存在，其内部运作机制对外界研究者不透明，传统研究方法面临无法穿透算法黑箱的挑战。

基于智能体的测试（agent-based testing, ABT）方法为算法黑箱的可解释探索提供了创新路径，实现了从被动观察到主动测试的方法论转变。其原理在于构建模拟用户代理（user agents），通过有计划的交互行为探测算法系统的响应模式，进而推断其内部运作机制。这种方法将算法视为一个可被系统性测试的“研究对象”，而非不可接近的黑箱。2014年，桑德维格（Sandvig）等人开发了基于智能体的算法审计方法，系统测试了Google搜索算法中的种族和性别偏见^[49]。研究团队创建了具有不同人口统计特征的虚拟用户画像，包括不同的姓名、地理位置、搜索历史等属性，用户通过这些智能体进行相同关键词的搜索查询。结果发现搜索算法存在显著的性别与种族偏见，将高收入标签更多关联白种人男性用户，而黑种人用户则更多关联犯罪、低收入标签，这项研究为算法偏见问题提供了科学证据，推动了学术界、政策制定者和科技公司对算法公平性问题的重视，促进了相关技术标准和监管框架的发展。

在具体操作中，ABT通常遵循“输入—输出—推断”的研究范式：首先，设计具有特定属性和行为模式的智能体用户群；其次，让这些智能体按照预设策略与目标算法系

统进行交互（如浏览、点击、搜索等行为）；再次，记录算法系统对不同智能体的差异化响应；最后，通过统计分析和模式识别，推断算法可能采用的决策规则和偏好因素。比如研究者可以部署具有不同政治倾向的智能体群体，观察社交媒体的推荐算法如何根据政治偏好差异化分发内容，从而揭示算法中可能存在的政治极化强化机制。贝尔等人通过 Twitter 机器人实验研究了算法推荐对政治态度的影响机制^[50]。研究团队部署了 1,652 个具有不同政治倾向的智能体账户，让其在为期一个月的时间内接收来自对立政治阵营的信息推送。通过对智能体前后测量的态度变化分析，研究发现，算法推荐的跨党派信息接触不仅未能缓解政治极化，反而显著加剧了保守派用户的极化倾向，自由派用户的态度变化则相对较小。这一发现挑战了“多元接触假说（diverse exposure hypothesis）”，表明算法中介的信息接触可能产生与预期相反的极化效应。格施克（Geschke）等人构建了包含 10,000 个智能体的 Facebook 模拟环境，每个智能体具有不同的政治倾向、兴趣偏好和社交网络结构^[51]。模拟结果显示，即使在算法推荐机制相对中性的条件下，用户的同质化社交倾向（homophily）也会导致信息茧房的自然形成。当算法推荐引入个性化偏好匹配机制后，信息茧房效应被显著放大，用户接触到的信息多样性平均下降 40%，政治态度的极化程度相应增强。

这种动态网络模拟在传播学研究中展现出独特价值。它能够模拟信息传播中的网络动态适应性变化，如意见极化如何导致网络分裂成封闭的信息“同温层”社区，可系统探究节点行为与网络结构的协同演化如何影响信息传播的广度与深度，揭示传播内容特性（如情感强度、争议性、可信度）如何与网络动力学相互作用，形成不同的传播路径和结果。

目前基于智能体的动态网络模拟已在多个传播学研究领域取得突破。在谣言传播研究中，研究者通过模拟不同信息验证机制下的网络动力学，发现了谣言传播的临界阈值和最优干预时机。在舆论形成研究中，动态网络模拟揭示了少数派意见如何在特定网络结构条件下形成局部共识，并最终影响整体舆论走向。在社交媒体极化研究中，模拟结果展示了算法推荐、同质社交倾向和内容特性如何共同作用，导致网络社区的分化与隔离。

五、结语

进入 21 世纪第三个十年，全球信息传播体系正经历以 AI 为核心的第四次技术革命浪潮^[52]。这一波技术变革不仅重塑了信息的生产与分发机制，更深刻影响了新闻传播学的研究对象、范式与方法论^[53]。AI 赋能的新闻传播实验范式变革，远非简单的技术升级，而是涵盖方法论、理论框架与伦理治理的整体性和系统性转型。人工智能的快速进步正为新闻传播研究创造前所未有的可能性。

然而，AI 赋能研究在方法上带来便利的同时，也引发了伦理、法律与学术诚信的复杂问题，这些问题既需要技术层面的对策，也需要制度化的规范升级。这包括知情同意与去欺骗化风险。AI 代理极具拟人性，使受试者难以区分人机边界。典型风险包括隐蔽的实验参与、未充分告知 AI 交互性质或实验目的。研究人员应当建立更严格的知情同意流程（如显著标注“您正在与 AI 代理互动”），完善事后说明与去欺骗程序，并在高拟真场景中慎重评估是否允许任何形式的行为诱导。而隐私与多模态数据风险亦是不容小觑的风险。面部表情、语音与行为数据属于高度敏感信息，应严格遵循数据最小化原则，优先实施设备端数据预处理，落实加密存储与严格的访问控制。必要时引入差分隐私、联邦学习等技术以降低泄露风险，并在伦理审查中开展风险—收益评估。

面对 AI 浪潮，新闻传播学研究者不仅应成为新技术的应用者，更应成为技术意义上的阐释者与范式革新的推动者，以保障学科理论的前瞻性。

[1] 库恩. 科学革命的结构（第四版）[M]. 金吾伦，胡新和，译. 北京：北京大学出版社，2012.

[2] 海，等. 第四范式：数据密集型科学发现 [M]. 潘教峰，等译. 北京：科学出版社，2012.

[3] SLATER M D. Reinforcing spirals model: Conceptualizing the relationship between media content exposure and the development and maintenance of attitudes [J]. *Media Psychology*, 2015, 18 (3): 370–395.

[4] HENRICH J, HEINE S J, NORENZAYAN A. The weirdest people in the world? [J]. *Behavioral and Brain Sciences*, 2010, 33 (2–3): 61–83.

[5] INGLEHART R, HAERPFER C, MORENO A, et al. World values survey: Round two – country-pooled datafile version [EB/OL]. (2014-12-01) [2024-12-01]. <http://www.worldvaluessurvey.org/WVSDocumentationWV2.jsp>.

[6] BAKSHY E, MESSING S, ADAMIC L A. Exposure to ideologically diverse news and opinion on Facebook [J]. *Science*, 2015, 348 (6239): 1130–1132.

[7] WATTS D J, DODDS P S. Influentials, Networks, and public opinion formation [J]. *Journal of Consumer Research*, 2007, 34 (4): 441–458.

[8] HAMMOND R A, AXELROD R. The Evolution of ethnocentrism [J]. *Journal of Conflict Resolution*, 2006, 50 (6): 926–936.

[9] SCHELLING T C. Dynamic models of segregation [J]. *Journal of Mathematical Sociology*, 1971, 1 (2): 143–186.

- [10] GONZÁLEZ-BAILÓN S, BORGE-HOLTHOEFER J, RIVERO A, et al. The dynamics of protest recruitment through an online network[J]. Scientific Reports, 2011, 1(1): 197.
- [11] MILLER J H, PAGE S E. Complex Adaptive Systems: An Introduction to Computational Models of Social Life [M]. Princeton: Princeton University Press, 2007.
- [12] CENTOLA D. The spread of behavior in an online social network experiment [J]. Science, 2010, 329 (5996): 1194–1197.
- [13] BERELSON B. Content Analysis in Communication [M]. New York: Free Press, 1952.
- [14] FRANZOSI R. Content Analysis: Objective, systematic, and quantitative description of content [J]. Content Analysis, 2008, 1 (1): 21–49.
- [15] ATTEWELL P, MONAGHAN D. Data Mining for the Social Sciences: An Introduction [M]. Oakland: University of California Press, 2015.
- [16] 张志安. 大数据与新闻传播研究的学术想象 [J]. 国际新闻界, 2018, 40 (2): 133–139.
- [17] 陈昌凤, 石泽. 价值嵌入与算法思维: 智能时代如何做新闻 [J]. 新闻与写作, 2021 (1): 54–59.
- [18] BARAM-TSABARI A, SEGEV E, SHARON A J. What’ s new? The Applications of data mining and big data in the social sciences [M] //The SAGE Handbook of Online Research Methods, 2017: 92–106.
- [19] 韩少卿, 巢乃鹏. 复杂性转向: 计算传播学的发生逻辑 [J]. 南京社会科学, 2024 (3): 109–122.
- [20] HANCOCK J T, NAAMAN M, LEVY K. AI-Mediated communication: definition, research agenda, and ethical considerations [J]. Journal of Computer-Mediated Communication, 2020, 25 (1): 89–100.
- [21] VENUGOPALAN M, GUPTA D. An enhanced guided LDA model augmented with BERT based semantic strength for aspect term extraction in sentiment analysis [J]. Knowledge-Based Systems, 2022, 246: 108668.
- [22] DEVLIN J, CHANG M W, LEE K, et al. BERT: Pre-training of deep bidirectional transformers for language understanding [C] //Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Minneapolis: ACL, 2019: 4171–4186.
- [23] BROWN T, MANN B, RYDER N, et al. Language models are few-shot learners

[C] //Advances in Neural Information Processing Systems, 33. Red Hook: Curran Associates, 2020: 1877–1901.

[24] KOJIMA T, GU S S, REID M, et al. Large language models are zero-shot reasoners [C] //Advances in Neural Information Processing Systems, 35. Red Hook: Curran Associates, 2022: 22199–22213.

[25] DE SANTIS E, MARTINO A, RONCI F, et al. From bag-of-words to transformers: A comparative study for text classification in healthcare discussions in social media [J] . IEEE Transactions on Emerging Topics in Computational Intelligence, 2024.

[26] RADFORD A, KIM J W, HALLACY C, et al. Learning transferable visual models from natural language supervision [C] //Proceedings of the 38th International Conference on Machine Learning, 2021: 8748–8763.

[27] VOGEL T, WANKE M. Attitudes and Attitude Change [M] . Psychology Press, 2016.

[28] IYENGAR S. News that matters: television and American opinion [J] . The American Journal of Sociology, 1989, 94 (4): 882–884.

[29] BANDURA A. Social Learning Theory [M] . Prentice-Hall, 1977.

[30] VOSOUGHI S, ROY D, ARAL S. The spread of true and false news online [J] . Science, 2018, 359 (6380): 1146–1151.

[31] THORSON K, WELLS C. Curated flows: A framework for mapping media exposure in the digital age [J] . Communication Theory, 2016, 26 (3): 309–328.

[32] HOVLAND C I, JANIS I L, KELLEY H H. Communication and Persuasion [M] . Yale University Press, 1952.

[33] MCGUIRE W J. Toward social psychology's second century [M] //KOCH S, LEARY D E. A Century of Psychology as Science. American Psychological Association, 1985: 558–590.

[34] SCHEUFELE D A, TEWKSBURY D. Framing, agenda setting, and priming: The evolution of three media effects models [J] . Journal of Communication, 2007, 57 (1): 9–20.

[35] LANG A. The limited capacity model of mediated message processing [J] . Journal of Communication, 2000, 50 (1): 46–70.

[36] 邓胜利, 汪奋奋. 互联网治理视角下网络虚假评论信息识别的研究进展 [J] . 信息资源管理学报, 2019, 9 (3): 73–81.

[37] 马福全. 媒介暴力对青少年的负面影响及学校干预途径 [J] . 教育理论与实践, 2018, 38 (14): 26–28.

- [38] 顾晨昱. 算法推荐内容对老年人健康信息回避行为的影响机制研究——数字反哺的干预实验 [J]. 国际新闻界, 2025, 47 (2): 27–48.
- [39] FALK E B, BERKMAN E T, MANN T, et al. Predicting persuasion-induced behavior change from the brain [J]. *Journal of Neuroscience*, 2012, 32 (25): 8421–8426.
- [40] MOOK D G. In defense of external invalidity [J]. *American Psychologist*, 1983, 38 (4): 379–387.
- [41] PETERSON R A, MERUNKA D R. Convenience samples of college students and research reproducibility [J]. *Journal of Business Research*, 2014, 67 (5): 1035–1041.
- [42] 吴国芳, 张万礼. 一种基于用户浏览行为的数据采集方法 [J]. 洛阳师范学院学报, 2014, 33 (8): 76–79.
- [43] 萨尔加尼克. 计算社会学: 数据时代的社会研究 [M]. 赵红梅, 赵婷, 译. 北京: 中信出版集团, 2019.
- [44] PENNYCOOK G, RAND D G. Fighting misinformation on social media using crowdsourced judgments of news source quality [J]. *Proceedings of the National Academy of Sciences*, 2019, 116 (7): 2521–2526.
- [45] HAUSER D J, SCHWARZ N. Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants [J]. *Behavior Research Methods*, 2016, 48 (1): 400–407.
- [46] ZIMMER M. “But the data is already public”: On the ethics of research in Facebook [J]. *Ethics and Information Technology*, 2010, 12 (4): 313–325.
- [47] 牟怡. 人机传播的内涵、外延及研究议程 [J]. 青年记者, 2023 (2): 9–11.
- [48] BAIL C A. Can generative AI improve social science? [J]. *Proceedings of the National Academy of Sciences*, 2024, 121 (21): e2314021121.
- [49] SANDVIG C, HAMILTON K, KARAHALIOS K, et al. Auditing algorithms: Research methods for detecting discrimination on internet platforms [R/OL]. 2014 [2024–12–01]. <https://api.semanticscholar.org/CorpusID: 15686114>.
- [50] BAIL C A, ARGYLE L P, BROWN T W, et al. Exposure to opposing views on social media can increase political polarization [J]. *Proceedings of the National Academy of Sciences*, 2018, 115 (37): 9216–9221.
- [51] GESCHKE D, LORENZ J, HOLTZ P. The triple-filter bubble: Using agent-based modelling to test a meta-theoretical framework for the emergence of filter bubbles and echo chambers [J]. *British Journal of Social Psychology*, 2019, 58 (1): 129–149.
- [52] BRYNJOLFSSON E, MCAFEE A. Machine, Platform, Crowd: Harnessing Our

Digital Future [M] . New York: W. W. Norton&Company, 2017.

[53] FLORIDI L. The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities [M] . Oxford: Oxford University Press, 2023.

（牟怡系上海交通大学媒体与传播学院教授、博士生导师；倪盈盈系上海交通大学媒体与传播学院在读博士；蓝剑锋系上海交通大学媒体与传播学院在读博士；马文斌系上海交通大学媒体与传播学院在读博士）