

生成式人工智能背景下数据新闻的转型路径与实践逻辑 ——基于三家主流媒体的质性研究

袁雯倩, 杨学永*

南京航空航天大学, 江苏 南京 211100

DOI:10.61369/SE.2025120001

摘 要 : 随着生成式人工智能 (AIGC) 加速介入新闻生产, 数据新闻作为技术依赖度极高的新闻类型, 正面临生产流程、协作机制与新闻理念的深度重塑。本文选取澎湃新闻、人民日报客户端与央视新闻三家主流媒体为案例, 结合 2020—2021 年与 2023—2024 年两个阶段的数据新闻作品, 探讨 AIGC 介入后数据新闻在选题策略、可视化手段等方面的演变路径。研究发现, AIGC 不仅在图表生成、文本润色等环节显著提升效率, 更促使编辑、记者与算法之间的协作逻辑发生结构性转变。不同类型媒体在技术采纳路径上呈现出差异化特征: 党媒趋于稳健与规范, 市场化媒体更重交互性与技术探索。

关 键 词 : 生成式人工智能; 数据新闻; 新闻生产; 人机协作; 质性研究

Transformation Path and Practical Logic of Data Journalism in the Context of Generative Artificial Intelligence — A Qualitative Study Based on Three Mainstream Media Outlets

Yuan Wenqian, Yang Xueyong*

Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu 211100

Abstract : With the accelerated integration of generative artificial intelligence (AIGC) into news production, data journalism, as a highly technology-dependent form of journalism, is undergoing profound transformations in its production processes, collaborative mechanisms, and journalistic principles. This paper selects three mainstream media outlets—The Paper, People's Daily app, and CCTV News—as case studies and analyzes data journalism works produced during two periods: 2020–2021 and 2023–2024. It explores the evolutionary paths of data journalism in terms of topic selection strategies and visualization methods following the introduction of AIGC. The study finds that AIGC not only significantly enhances efficiency in areas such as chart generation and text refinement but also drives structural changes in the collaborative logic among editors, journalists, and algorithms. Different types of media exhibit distinct characteristics in their technology adoption paths: party-affiliated media tend to be more stable and standardized, while market-oriented media prioritize interactivity and technological exploration.

Keywords : generative artificial intelligence; data journalism; news production; human-machine collaboration; qualitative research

引言

新闻从来不是在真空中完成生产的。从机械排版、卫星传输到数字可视化、智能算法, 每一次技术浪潮的出现都深刻地介入了新闻的呈现形态与实践逻辑。自 2022 年底 ChatGPT 引发全球关注以来, 生成式人工智能 (AIGC) 加速进入新闻业, 以其在文本撰写、视觉生成和结构建构方面的能力, 对新闻生产的各个环节产生实质性影响。

数据新闻作为融合技术性与专业性最紧密的报道样态, 理应成为观察这场变革的重要窗口。一方面, 它长期依赖数据采集、清洗、分析与可视化等流程, 具备天然的自动化潜力; 另一方面, 面对 AIGC 的大规模应用, 数据新闻亦面临由“结构化协作”向“模型化介

项目信息: 本文系中央高校基本科研项目“智能传播时代下媒体与受众互动模式的变革方向与模式”(项目编号: ND2025003)和教育部产学研合作协同育人项目(项目编号: 241204554182113)“无人机飞行教学训练虚拟仿真”的阶段性成果。

作者简介: 袁雯倩, 学生, 研究方向: 媒介技术、传播理论、媒介治理。

通讯作者: 杨学永, 助理研究员, 研究方向: 媒介技术、智能传播、媒介治理。

入”的模式转变。在此背景下，本文以三家中国主流媒体为样本，选取2020—2021年与2023—2024年两个阶段具有代表性的数据新闻作品为分析对象，探讨以下问题：（1）在AIGC普及前后，主流媒体的数据新闻在选题类型与数据来源方面有何结构性变化？（2）不同属性的媒体在技术采纳路径上表现出怎样的差异？

一、文献综述与理论基础

数据新闻作为融合新闻专业能力与技术工具的创新报道形态，自进入中国媒体实践以来便受到学界广泛关注。其依托结构化数据与可视化技术重构新闻叙事路径，被认为是传统新闻业在数字化转型中的突破口（马玉霞，2015）。陈虹与秦静（2016）指出，数据新闻打破了以人为中心的经验判断路径，通过大数据挖掘、图表呈现与互动模块拓展新闻边界，为新闻的“客观性”提供了技术支持。学界在早期研究中主要聚焦于数据新闻的定义演化、生产流程、可视化风格与伦理规范等方面（林溪声，2018；秦静，潘玉，2015），但整体上仍偏重描述与经验总结，缺乏对技术逻辑深层次作用机制的系统探讨。

随着人工智能特别是生成式AI（AIGC）的迅猛发展，新闻业正面临新一轮技术范式的剧烈重构。2022年底ChatGPT的面世被广泛视为“智能新闻”变革的重要拐点（陈昌凤，2023）。彭兰（2023）指出，在生成式技术广泛介入新闻采集、写作与分发的背景下，新闻工作已从“工具增强”阶段迈入“人机共创”新阶段，原有的人力编辑逻辑开始让位于算法协作模式。这一趋势在数据新闻领域尤为显著。从数据获取、清洗到初步分析与图表构建，AIGC工具已具备相当程度的自动化处理能力。一些主流媒体如澎湃新闻、人民日报客户端等已在数据新闻中引入AI图表生成与摘要模型，在部分页面标注“AI辅助生成”或“智能推荐”等字样（郎劲松，杨海，2014）。

学界对AIGC在新闻流程中的介入，主要有三类回应路径。一种是技术工具主义视角，强调其在提升效率、扩展报道类型方面的优势，尤其在突发事件、专题数据追踪等高频场景中表现突出（陈虹，秦静，2016）。另一种则持批判立场，指出算法在信息筛选与表征中的“黑箱效应”，可能放大数据偏误、削弱记者主体性与专业判断（邵婉霞，阴晴霞，2024）。第三种路径尝试将人机关系置于组织传播语境中，探讨AIGC如何通过“流程嵌入”重塑编辑部内部的角色划分与组织结构（肖鳕桐，方洁，2020）。在此背景下，新闻生产正由传统的“记者—编辑—排版”线性流程转向“人机协作—多工协同—动态迭代”的网状流程形态。

值得注意的是，这一过程不仅是技术替代，更是新闻理念与制度逻辑的重组。塔克曼提出的“新闻常规”理论认为，新闻并非孤立的创作活动，而是制度性安排下组织成员对不确定性的协商结果（Tuchman，1978 引用自林溪声，2018）。AIGC的引入在改变个体操作方式的同时，也重构了组织如何设定“标准化流程”的基础逻辑。例如，传统数据新闻强调数据清洗、分析与叙事之间的逻辑一致性，而生成式AI则可能提供快速但难以验证的

初步文稿，导致把关机制压力增大（邱淑露，2025）。

新闻职业角色也在此过程中经历转型。记者正从传统的“信息发掘者”向“技术驾驭者”与“价值监督者”角色升级。在智能化传播语境中，记者必须具备“算法素养”与“跨域理解力”，方能有效引导模型生成内容并保持对新闻价值的把握（虞鑫，王金鹏，2022）。这一论断回应了彭兰（2023）所言的“智能传播三态关系”：人机协同、人机交流与人机共生。在数据新闻场景中，这种协同关系尤其明显，记者与AI不再是主—辅关系，而是在流程中不断切换角色、共享决策权。^[1-5]

与此同时，学界也日益重视生成式AI带来的伦理挑战与制度空白。生成内容缺乏可追溯性、真实性难以验证、算法歧视问题等，引发了对于新闻公信力的忧虑（邵婉霞，阴晴霞，2024）。目前，多数媒体尚未建立系统的AI内容审核机制，对“谁负责”“如何审定”“是否标注”缺乏统一规范，这一监管滞后显然与其快速扩张的应用趋势不匹配（陈昌凤，2023）。因此，部分学者主张构建“人机双轨”的审核制度，将记者与算法各自的优势结合，以保障内容的质量与价值边界（邵婉霞，阴晴霞，2024）。

整体来看，当前关于AIGC与数据新闻融合的研究呈现出三个明显演进趋势：从工具层面的“可用性”探讨向机制层面的“协作逻辑”转向；从内容生成环节的技术替代关注扩展至流程与组织层级的深度协同；从单一的效率逻辑回归至新闻专业主义与伦理价值的守护。这些趋势共同构成了数据新闻在AIGC时代走向制度化创新的研究基础。

二、研究设计

（一）样本抽样策略

本研究对象主要涵盖澎湃新闻新闻客户端“美数课”专栏、人民日报客户端“图解中国”和“数读”专栏，以及央视新闻客户端、央视网的“数说××”和“数据可视化”等相关专题。样本时间段划分为2020—2021年（AIGC广泛应用之前）与2023—2024年（AIGC广泛应用之后）两段。为保证样本的可比性，本研究采用目的抽样法，每家媒体每一阶段选取10篇相关数据新闻作品，共计60篇。“数据新闻”被界定为具有明确数据基础并通过可视化形式呈现信息的新闻页面。澎湃新闻和人民日报客户端均以固定专栏发布此类内容，而央视新闻多以专题形式呈现，因此本研究对央视样本以专题标签而非栏目命名进行识别。

（二）编码维度与执行流程

研究团队首先编制统一编码表，并在部分样本上进行了预编码试验，以检验各项编码指标的可操作性。编码维度主要包括

四类：

- 1. 选题属性：议题类型（如经济、社会、科技等）和报道节奏；
- 2. 数据来源与处理：数据来源（官方、第三方或自采）及是否说明数据获取和处理方法；
- 3. 可视化与交互：采用的可视化类型（如静态图表、动态图表、地图等）和互动功能设计（如点击筛选、动态交互等）；

4.AIGC 介入线索：报道中显性提示 AIGC 使用的要素，如 AI 标注（“由 XXX 智能系统生成”之类提示）、自动生成提示、机器人署名等。

需要指出的是，由于研究资源与公开信息所限，本研究仅针对新闻页面上可观察到的显性要素进行定量编码，不涉及隐性技术流程的溯源验证。每篇样本统一填写编码表，并进行基础频次统计与类型分布分析，识别三家平台数据新闻的呈现特征差异。

表 1 内容分析编码框架与变量设置

编码维度	变量名称	操作性定义	编码说明与取值示例
选题属性	议题类型	报道聚焦的核心议题领域	1= 宏观政策与治理；2= 社会民生；3= 科技创新与数字经济；4= 经济金融；5= 国际与地区事务；6= 其他
	报道节奏 / 类型	报道在时间性与深度上的定位	1= 时讯性 / 快报；2= 解释性专稿 / 背景解读；3= 系列策划 / 专题；4= 评论性数据报道
数据来源与处理	数据来源类型	报道主要依托的数据来源类别	1= 国家统计局等官方统计部门；2= 党政机关 / 事业单位；3= 科研院所 / 智库；4= 行业组织；5= 商业平台 / 互联网公司；6= 记者自采；7= 多源混合
	是否说明方法	报道中是否对数据采集、清洗、分析方法作出说明	0= 完全未说明；1= 只笼统说明来源（如“根据公开数据整理”）；2= 较为详细说明（包括口径、处理方式等）
可视化与交互	可视化主类型	报道中承担信息呈现主体功能的可视化形态	1= 单一静态图表；2= 多图组合（图组 / 图册）；3= 信息长图 / 长页面图解；4= 动态图 / 动画图表；5= 地图可视化；6= 图文混排为主
	可视化复杂度	图表组合的丰富程度与信息承载量	1= 简单（1 - 2 幅基础图表）；2= 中等（多幅图表，结构清晰）；3= 复杂（多模块、多层级信息）
	交互设计程度	页面是否提供互动操作及其丰富程度	0= 无互动（纯静态浏览）；1= 有限互动（如切换标签、悬浮提示）；2= 较强互动（如自定义筛选、可操作时间轴等）
AIGC 介入线索	AIGC 显性标注	页面中是否出现表明使用 AI/AIGC 的文字或图标	0= 无任何相关标注；1= 弱提示（如“智能推荐”“由系统生成部分内容”）；2= 明确标注（如“本页面由 × × 大模型生成 / 辅助生成”）
	自动生成署名	是否出现“机器人”“系统”“算法”等非人类署名，或强调自动生成流程	0= 未出现；1= 出现非人类署名或自动生成声明

三、研究发现

（一）议题的轻微转向：稳定中的内在重塑

无论是在前 AIGC 阶段，还是在 AIGC 快速进入新闻机构之后，宏观政策与社会民生始终构成数据新闻最稳定的支撑面。人民日报客户端的数据新闻以清晰的政策主线为骨架，通过结构化数据对国家治理图景进行可视化呈现；央视新闻的作品则更多围绕重大主题节点，以国家叙事为中心展开；澎湃新闻“美数课”在保持对公共议题关注的同时，议题更贴近城市脉动和社会现实。^[6-10]

真正发生变化的是“重心的微调”。进入 AIGC 普及阶段后，与科技创新、数字经济、新质生产力相关的主题出现得更频繁，也更“自然”地融入日常报道结构。这样的增长并非简单迎合舆论热度，而是与国家战略叙事、社会技术氛围及受众兴趣共同作用的结果。数据新闻在这一时期不再只是政策图解的工具，而是逐渐承担起解释科技系统、建构数字社会想象的角色。

此外，一类新的议题形态开始变得重要——跨领域的、结构性的、由多源数据支撑的“问题链条”。例如，用人口趋势解释教育资源调整，用消费结构变化观察青年就业状态，用产业布局和资本流向解读地方城市转型。相比于前阶段“围绕单一问题展开”的报道方式，后期的数据新闻更倾向于把多个变量放在同一个叙事框架里讲述，形成一种“用结构解释现象”的新倾向。

（二）数据结构的调整：从单源依赖到多源整合

三家媒体的数据来源在两个阶段的差异，也呈现出一种“从

单到多、从稳到活”的变化。前 AIGC 阶段的数据新闻大量依靠官方统计与政府数据库，呈现出典型的“单源型报道”特征。数据新闻在这一时期的任务，是把厚重而严肃的权威数据转译成易读、易懂的图表形态。

但随着 AIGC 工具和多元数据接口进入生产流程，记者与编辑在后阶段明显更主动地使用多源数据。官方统计仍是底座，但科研机构、行业平台、互联网企业的数据开始被自然地纳入分析链条。例如，涉及城市生活与消费趋势的报道，会并置官方消费统计与平台大数据；涉及就业结构的报道，会用教育部门数据与招聘平台数据做对照；讨论 AI 应用与产业链时，则常同时引用政策文件、市场调研数据与行业预测模型。

值得注意的是，这种多源拼接并未完全改变数据新闻的方法透明度。虽然部分报道开始简要交代处理逻辑，但“大部分作品对口径、处理方式、模型原理的说明仍相对有限”。这种谨慎并非技术能力不足，而更多源于主流媒体对“稳定性”和“责任界定”的要求。在这种制度环境下，方法透明度的提升往往滞后于工具能力的提升。

（三）AIGC 的后台介入：生产机制的隐性重组

最值得强调的是，AIGC 并没有在页面上大举“露脸”。显性标注极少，机器人署名几乎没有出现；技术团队的参与常以“支持单位”“数据平台”的方式隐身在署名下方。这一方面源于主流媒体对内容权威性与责任主体的高度敏感，另一方面也说明：AIGC 融入新闻生产并未以一种炫技式的姿态出现。

但在文本与页面的细部观察中，AIGC 的痕迹十分明显。许多

解释性段落呈现出高度模板化的语言结构；若干作品的“导语—三个层级的小标题—结语”模式高度一致；某些类型的图表风格甚至出现了“跨媒体趋同”。这些现象无法完全用“人工模仿”解释，背后反映的是自动摘要、自动出图、智能可视化模板等工具已经嵌入各媒体的后台流程。

更深的变化发生在“角色重新定位”上：记者的工作从“完整撰写一篇报道”转向“提出问题、挑选数据、校对模型输出”；编辑从“改文章”转向“设计结构与判断价值风险”；技术团队从后台支持变成了“流程的共同设计者”。AIGC 在这里不是一个“生产者”，而是一种“结构力量”，改变了新闻机构内部“谁做什么、谁先做什么、什么可以自动做”的基本秩序。

在这样的背景下，数据新闻看似保持了原有的专业性与稳健风格，但其生产逻辑已经发生了深刻的重组。

四、结论与讨论

技术并非以激进姿态强行介入，而是在现有制度框架与专业

惯习中，逐步导向一种新的行业秩序。新闻机构应用生成式技术，亟需超越效率工具的单一视角，转向对组织变革与专业影响的深度审视：将技术引入作为制度化建设的核心环节，系统性提升记者的数据素养与模型能力；以公开透明的技术参与方式，重建公众对新闻生产的信任。诚然，本研究存在样本量有限、分析依赖前端呈现、未深入受众理解等局限，但未来研究可通过扩大媒体类型覆盖、采用民族志方法深入编辑部场域、引入受众评估等路径，进一步拓展对生成式技术影响的认知边界。

生成式技术并非一场骤然降临的剧烈革命，而是一场持续的秩序重构过程。它深刻重塑了新闻组织的形态与记者的工作方式，在带来效率便利的同时，也划定了新的专业边界。新闻业的未来，正取决于我们如何在这一由技术驱动的新节奏中，重新锚定自身的专业定位与价值坐标。

参考文献

- [1] 马玉霞. 数据新闻的兴起与发展文献综述 [J]. 新闻世界, 2015(6): 210-212.
- [2] 陈虹, 秦静. 数据新闻的历史、现状与发展趋势 [J]. 编辑之友, 2016(1): 69-75.
- [3] 林溪声. 数据新闻生产的常态化趋势及其限度 [J]. 新闻爱好者, 2018(2): 35-38.
- [4] 郎劲松, 杨海. 数据新闻：大数据时代新闻可视化传播的创新路径 [J]. 现代传播, 2014(3): 32-36.
- [5] 陈昌凤. 生成式人工智能与新闻传播：实务赋能、理念挑战与角色重塑 [J]. 新闻界, 2023(6): 4-12.
- [6] 彭兰. 从 ChatGPT 透视智能传播与人机关系的全景及前景 [J]. 新闻大学, 2023(4): 1-16.
- [7] 邵婉霞, 阴晴霞. 生成式人工智能时代的新闻生产：关系再造、实践转型与价值坚守 [J]. 传播创新研究, 2024(1): 203-216.
- [8] 肖懿桐, 方洁. 内容与技术如何协作？——行动者网络理论视角下的新闻生产创新研究 [J]. 国际新闻界, 2020, 42(11): 99-118.
- [9] 邱淑露. 数据新闻叙事策略研究——以“网易数读”为例 [J]. 社会科学前沿, 2025, 14(8): 426-434.
- [10] 虞鑫, 王金鹏. 重新认识“信息茧房”——智媒时代工具理性与价值理性的共生机制研究 [J]. 新闻与写作, 2022, (03): 65-78.