Teaching Exploration on Cultivating News Communication Talents' Data Competence Empowered by Generative AI

Xu Wang

College of Communication, Liaoning Communication University, Shenyang, Liaoning, China

Abstract. This paper explores the application pathways of generative artificial intelligence (AI) in cultivating data competence among journalism and communication professionals. The study finds that AI technologies can effectively enhance abilities in data acquisition, analysis, and storytelling, yet they also pose risks such as insufficient data authenticity and algorithmic bias. The research proposes a "human-AI collaboration" pedagogical framework, constructing application strategies from three dimensions—course objectives, content, and teaching models—while emphasizing the central role of critical thinking and ethical judgment. This provides theoretical reference for reforming journalism and communication education.

Keywords: Generative Artificial Intelligence; Journalism and Communication Education; Data Competence Development.

1. Introduction

In recent years, the rapid advancement of generative artificial intelligence (AI) has profoundly reshaped the field of journalism and communication. From data-driven news production to intelligent public opinion analysis tools, AI is not only transforming traditional news production processes but also imposing new demands on the data competencies of journalism professionals. However, current journalism and communication education significantly lags in cultivating data skills: curricula are slow to update, and there is a serious lack of training in data literacy and the application of AI tools.

2. Alignment Between Generative AI and Data Competence Development in Journalism Education

2.1 Data Competence Requirements for Journalism and Communication Professionals

In the era of intelligent communication, the cultivation of data competence must transcend the single dimension of traditional statistical analysis, establishing a multidimensional framework that integrates technical operation, methodological thinking, and value judgment. This framework should be grounded in a deep understanding of evolving journalistic practices, rather than a mere adoption of technological tools [1].

From a technical perspective, data processing in journalism is undergoing two notable shifts: First, data forms are expanding from structured to unstructured formats—such as social media text, video metadata, and live-stream interaction logs—requiring interdisciplinary skills like natural language processing and computer vision. Second, analytical workflows are shifting from manual execution to automation, necessitating programming capabilities to support end-to-end processes including data scraping, cleaning, and visualization.

At the level of methodological thinking, the core of data competence lies in establishing a "question-data-method" mapping ability. In typical journalistic scenarios such as public opinion analysis, practitioners must select appropriate analytical methods based on specific problem characteristics: sentiment analysis may require combining lexicon-based and machine learning models, while tracking information diffusion paths may involve social network analysis tools. More importantly, students must develop critical data thinking—assessing data source reliability, questioning the validity of analytical results, and recognizing model limitations. Such cognitive training is far more crucial than memorizing software operations.

2.2 The Educational Potential of Generative AI

Generative AI offers transformative educational value by creating three types of learning environments that are difficult to achieve through traditional pedagogy, each uniquely beneficial for competence development.

First is the creation of controllable complex learning environments. Through carefully designed prompt engineering, instructors can generate datasets with specific pedagogical features—such as social media texts containing typical online linguistic phenomena or survey data embedded with intentional bias patterns. This controllability allows learners to progressively engage with various data challenges without being constrained by the unpredictability of real-world data. In practice, instructors can systematically increase data noise, missing values, or contradictions by adjusting generation parameters, thereby training students to handle real-world data complexity.

Second is interactive learning with immediate feedback. In traditional data analysis instruction, students often wait for instructor feedback, whereas AI tools can provide real-time, process-oriented guidance. For example, in coding exercises, AI can not only identify syntax errors but also explain the underlying logical flaws; in data interpretation training, AI can simulate "experts" with different professional backgrounds, offering critiques from perspectives such as news value or statistical rigor. This interaction creates a "one-on-one" tutoring effect rarely achievable in conventional classrooms, fostering flexible data thinking.

Third is cross-modal integrated competence training. The multimodal capabilities of generative AI open new possibilities for developing data storytelling skills. Learners can input analytical results into text generation models to produce draft reports in different styles, then compare and analyze how presentation formats affect communication effectiveness. They can also convert structured data into visual chart suggestions and evaluate the cognitive load of different visualization forms. This training bridges the traditional gap between technical operations and communication artistry, helping students build organic connections between data technology and journalistic practice.

3. Theoretical Application Pathways of Generative AI in Curriculum Design

3.1 Revising Course Objectives

The integration of generative AI fundamentally shifts the objectives of journalism and communication data courses across three dimensions: capability orientation, evaluation criteria, and educational focus. Traditional goals often center on mastering statistical software skills, whereas AI-era education should be redefined as cultivating "human-AI collaborative data decision-making ability" [2].

In terms of capability orientation, course objectives should evolve from isolated technical skills to a tripartite structure of "technical understanding—tool application—value judgment." Students must not only understand how generative AI works and its limitations but also master strategies for deploying AI tools appropriately within news production workflows. For instance, in public opinion monitoring, students should learn to identify which steps (e.g., data cleaning) can be automated by AI and which (e.g., conclusion validation) require human oversight.

Evaluation criteria must shift from "result correctness" to "process rationality." Traditional coding assignments focus on whether code runs and results are accurate. In contrast, AI-assisted assessment should emphasize how students select and use AI tools, verify AI outputs, and integrate AI-generated content into analytical workflows. This shift necessitates new rubrics focused on key decision points in the analytical process.

3.2 Restructuring Teaching Content

Curriculum content should be restructured according to three principles: authenticity, progression, and integration, incorporating generative AI organically while preserving journalistic identity. This is not a simple technological overlay but a systematic reorganization of pedagogical elements.

In the data acquisition phase, the use of AI-generated synthetic data must follow strict instructional design guidelines. Instructors should provide structured prompt templates to guide students in generating pedagogically targeted datasets. For example, when training data cleaning skills, prompts can instruct AI to embed specific types of missing values, outliers, or inconsistencies. Crucially, the relationship between synthetic and real data must be clarified—synthetic data for skill-specific training, real data for comprehensive assessment—forming a progressive learning sequence [3].

The focus of data analysis instruction should shift from writing code to reviewing and optimizing AI-generated code. When AI can automatically produce basic code, teaching should advance to evaluating the efficiency and security of AI-generated code, adjusting parameters for specific journalistic contexts, and integrating multiple AI-generated modules into a cohesive workflow. For instance, in text analysis tasks, students could compare different AI-generated sentiment analysis scripts in terms of accuracy, runtime efficiency, and interpretability.

Data storytelling training can be significantly enhanced through AI's text generation capabilities. A "generate-deconstruct-optimize" three-step method can be implemented: first, AI generates a basic analytical report; second, students critically analyze its logical flaws; third, students guide the AI to refine the output. This process helps students understand the translation mechanism between data analysis and narrative journalism, cultivating the ability to convert technical language into public discourse.

3.3 Innovating Teaching Models

Pedagogical innovation must be grounded in the principle of "augmentation rather than replacement," leveraging the complementary strengths of AI and human instructors to form a new educational community.

Blended learning design should clearly delineate roles between AI and human instructors. AI is better suited for standardized tasks such as knowledge delivery (e.g., via intelligent Q&A systems) and skill drills (e.g., automated coding assessment), while classroom time should focus on high-order thinking activities—such as debating AI-generated results or simulating data interpretation scenarios. This requires reengineering traditional workflows into a closed loop: "AI-assisted pre-learning → in-depth classroom discussion → AI-supported reinforcement."

Task-based learning can be personalized through AI. Instructors can design parameterized task templates, allowing AI to generate customized scenarios based on individual student profiles. For example, in a data visualization assignment, AI could generate datasets with different emphases (e.g., temporal trends vs. regional comparisons) for different students, all requiring the same design principles. This approach ensures pedagogical consistency while respecting learner diversity [4].

4. Potential Challenges and Mitigation Strategies

4.1 Technical Limitations

The application of generative AI in journalism education faces several inherent technical constraints, requiring educators to maintain critical awareness. A primary issue is the "authenticity paradox" of generated data—while AI-generated data can be highly realistic, it fundamentally represents a statistical fitting of existing data distributions and cannot fully replicate the spontaneity and novelty of real-world social dynamics. For instance, in simulating online public opinion, AI struggles to generate truly disruptive or innovative expressions of opinion, potentially leading students to develop path dependency on conventional patterns during training.

4.2 Educational Risks

The most concerning educational risk is "cognitive atrophy due to technological dependence." When students routinely rely on AI for end-to-end data processing, they may lose substantive understanding of key stages in the data pipeline—such as rule-setting in data cleaning or feature engineering. This "knowing how without knowing why" state severely limits their ability to solve

non-standardized problems. For example, when encountering edge cases that AI cannot handle, students lacking foundational training may be completely at a loss.

The risk of algorithmic bias transmission is even more insidious. When instructors use AI-generated content as teaching material, they may inadvertently introduce cultural biases or group stereotypes from the AI's training data into the classroom. Particularly in cross-cultural communication analysis, mainstream AI models trained primarily on English corpora may embed Western-centric perspectives, influencing students' interpretations of local communication phenomena. More troublingly, such biases often appear in the guise of statistical "scientific validity," making them more deceptive.

4.3 Mitigation Strategies

To address technical limitations, institutions should establish "AI usage technical guidelines" that clearly define the scope of synthetic data use—e.g., permitting synthetic data only for basic skill training, while requiring real datasets for final assessments. Additionally, students should be required to submit a "technical roadmap" when using AI tools, explaining their rationale for tool selection and acknowledging anticipated limitations.

To mitigate educational risks, a "phased de-AI" strategy should be implemented: allowing moderate AI use in early stages, then gradually increasing requirements for independent work as learning progresses. For algorithmic bias, instructors can design "bias (bias means 'bias tracing') exercises" where students reverse-engineer AI outputs to identify potential value orientations. Most importantly, human instructors must retain supervisory authority at critical junctures—especially in areas involving ethical judgment and creative thinking.

5. Conclusion

This study systematically examines the theoretical pathways for integrating generative AI into the cultivation of data competence in journalism and communication education. Findings indicate that AI offers significant advantages in expanding educational resources, improving learning efficiency, and deepening cognitive engagement. However, realizing its educational value requires rigorous instructional design. The core insight is that AI should serve educational goals—the deployment of technology must be deeply integrated with the professional characteristics of journalism education. In critical areas such as data ethics, critical thinking, and creative expression, clear human-AI division of labor must be established to prevent technological applications from undermining the core values of journalism. Educators must strike the right balance—leveraging technological empowerment to enhance teaching effectiveness while safeguarding the essential nature of journalism education.

References

[1] Peng, L. (2021). The New Content Revolution in the Age of Intelligence. Journalism & Writing, (1), 5–12.

[2] Yu, G., & Yang, Y. (2020). The Impact of Algorithmic Recommendation on News Communication and Its Governance Path. Xinwen Jie, (1), 4–11.

[3] Hu, Y., & Li, J. (2022). Research on the Impact of Generative AI on News Production. Modern Communication, 44(3), 1–8.

[4] Shi, A., & Wang, P. (2021). Journalistic Ethics in the Era of Intelligent Communication: Challenges and Countermeasures. Journalist, (3), 4–12.