

Exploration of Technology Applications for Intelligent, Intensive, and Efficient Converged Media Collection and Editing Systems: Postprint

Authors: Feng Enda

Date: 2023-10-08T00:00:00+00:00

Abstract

The “Opinions on Accelerating the In-depth Integrated Development of Media” provides guidance for the transition from converged media to intelligent media. This paper investigates the pain points of existing converged media collection and editing systems, and explores intelligent selection and application methods—including combined deployment of intelligent capabilities, efficient content production assistants, and efficient content production coordination—to meet the heightened demands that new communication formats impose on content production systems.

Full Text

Intelligent, Intensive, and Efficient Media Convergence Editing System Technology: An Exploratory Study

Xinhua Daily Media Group, Nanjing, Jiangsu 210000

Abstract: The *Opinions on Accelerating the In-Depth Development of Media Convergence* provides clear guidance for the evolution from converged media to intelligent media. This paper examines the pain points of existing media convergence editing systems and explores intelligent capability combination applications, efficient content production assistants, and coordinated content production mechanisms to meet the heightened demands of new communication formats on content production systems.

Keywords: media convergence editing; intelligent assistant; intensive efficiency; dialogue system

1. Overview and Pain Points of Media Convergence System Construction

1.1 Mainstream Media Editing System Development Media editing systems at traditional media organizations have generally evolved through three stages: traditional editing, “central kitchen” models, and media convergence platforms. Most media organizations at all levels have already established foundational information systems for content production, including content editing systems, issuance management systems, and media asset libraries, thereby initiating the convergence development process. In December 2019, Xinhua News Agency piloted the nation’s first intelligent editorial department, and in December 2020, *People’s Daily* launched its “Creative Brain” platform. These developments by China’s two major central media organizations inaugurated a new era of using artificial intelligence to reconstruct content production workflows.

1.2 Pain Points in Media Convergence System Construction Current projects transitioning from media convergence to intelligent media typically adopt a phased construction approach based on overall design. Constrained by budgets, construction timelines, and sub-project divisions, these initiatives often result in redundant investment in similar intelligent capabilities across different components, with limited cross-project capability sharing. Additionally, the flexibility of intelligent capabilities—specifically, their ability to be rapidly reconfigured and retrained according to media convergence business requirements—remains restricted, as does the resource consumption associated with optimization training. As convergence deepens, media convergence system construction extends beyond mere technological projects to encompass the integration of human resources, financial assets, and physical infrastructure. Achieving the overarching goals of comprehensive media convergence requires intelligent-driven approaches that foster proactive, open, and efficient talent cultivation and organizational methods, combined with new technology-enhanced content production models.

2. Intensive and Efficient Intelligent Capability Applications

2.1 Intelligent Capability Combination Applications Speech recognition represents a common intelligent capability in modern AI platforms. This technology converts speech signals into text by extracting acoustic features and intelligently matching them with acoustic and language models. Domestic technology companies such as iFlytek Open Platform, Baidu AI Open Platform, and Alibaba Cloud currently provide stable and reliable intelligent services in this domain. In media convergence systems, speech recognition is typically used for simple content creation assistance tasks such as transcribing interview recordings or converting video audio tracks to text. However, when integrated with media asset retrieval interfaces, this capability enables rapid search using voice input as the retrieval condition. Extending it with real-time processing capabil-

ities and combining it with mobile editing modules allows for voice-input-based drafting, significantly improving mobile content creation efficiency. When further enhanced with timeline attributes, this capability yields intelligent subtitling tools for video content. As the era of intelligent media approaches, the reuse of intelligent capabilities will become increasingly common; configuring each subsystem separately would waste resources. Encapsulating intelligent capabilities as microservices within an AI platform provides a flexible and efficient foundation for future media business transformation and development.

2.2 Efficient Intelligent Assistants The application of “big data + AI” capabilities has transformed the working methods and workloads of editorial staff. While new technologies offer unprecedented convenience, the exponential increase in available materials, information, formats, and options also presents new challenges. Task-oriented intelligent assistants can help editorial staff more efficiently acquire relevant services, materials, and information, thereby improving response speed and completion efficiency for editorial tasks and enabling them to meet emerging market demands for responsive content production.

2.2.1 Dialogue Optimization for Intelligent Assistants Intelligent assistants typically improve efficiency by converting user input into effective service or content output. Current intelligent assistants are most commonly deployed as online customer service agents, which predominantly use task-oriented dialogue systems to automatically respond to user needs, recommend products, and invoke services. Research indicates that retrieval-based dialogue systems are currently the most prevalent in the market. However, such systems often inadequately understand user questions and rely heavily on manually predefined response templates that match keywords against databases to select reply statements or trigger services. While this approach standardizes answers, its limitation becomes apparent after several dialogue rounds, frequently requiring users to select “this answer was not helpful” or “human customer service” options. Consequently, this solution proves unsuitable for media convergence scenarios. This paper therefore focuses on generative human-computer dialogue systems that fully leverage advances in deep learning.

2.2.2 Methods for Enhancing Intelligent Assistant Response Capabilities 2.2.2.1 Incorporating Classified Intent Recognition to Optimize Interaction Experience and Improve Service Delivery Efficiency

Due to the specialized terminology inherent in news content production systems, classifying user input within vertical domains during human-computer dialogue can enhance intent recognition accuracy. Model training can employ conditional sequence adversarial networks for intent recognition [2]. The method involves four preprocessing steps for user input: (1) tokenization and extraction of features such as word vectors, part-of-speech tags, word frequency, and topics; (2) concatenation and vectorization of feature values; (3) calculation of vector distances using RBF functions to group proximate feature vectors into the same

category; and (4) minimal manual review and verification. The dialogue generation process employs an unsupervised adversarial network structure (Figure 1 [Figure 1: see original paper]), which reduces model training costs compared to supervised approaches. The generator in this architecture can produce sufficient training materials from limited corpora and annotation data, while the discriminator's binary classifier distinguishes between real and fake corpora, and the category classifier categorizes samples based on probability distribution. Given that input lengths in human-computer dialogue scenarios are typically short, TextCNN convolutional neural networks with hyperparameter tuning can rapidly accomplish short-text intent classification. The generator's ability to learn statement features and mass-produce training inputs helps the classifier quickly achieve optimal performance.

2.2.2.2 Leveraging Media Knowledge Graphs to Support Content Decision-Making and Improve Response Generation Quality

In the specific context of media convergence centers, human-computer dialogues typically involve system resources, functions, and services. Media organizations have accumulated substantial industry knowledge over the years, which can be processed from existing structured, semi-structured, and unstructured data within production systems to form knowledge graphs that serve dialogue systems. Structured data such as personnel information (name, gender, age, position, responsibilities, editorial focus, published articles, etc.) can be converted into RDF data using D2R technology [4], transforming relational database query results into triple format to complete knowledge extraction.

Since data sources in media convergence systems are typically heterogeneous, knowledge fusion requires data cleaning, entity alignment, attribute value decision-making, and relationship construction under a unified architecture following knowledge extraction. Based on the fused knowledge, entity similarity calculations categorize similar entities, and reasoning using ANN or PRA constructs new entity relationships. Quality assessment then filters high-credibility components to form high-quality knowledge graphs stored in Neo4j databases. Semi-structured and unstructured data require more complex intelligent model-based knowledge extraction processes before integration into the knowledge graph.

As the intelligent assistants discussed here primarily serve content production, the dialogue tasks inherently carry scope constraints that reduce implementation complexity to some degree. The process begins by filtering a smaller subgraph collection from the complete knowledge graph using fuzzy retrieval methods. Reinforcement learning based on policy gradient algorithms then selects optimal knowledge within this narrowed subgraph. The overall structure is illustrated in Figure 2 [Figure 2: see original paper]. The knowledge selection module takes triple knowledge, dialogue topic information, and contextual information as simultaneous inputs. Through each round of human-computer dialogue, the system uses the initial topic as a starting point to search for optimal input-target fitting paths within the subgraph collection by adjusting network parameters via

gradient optimization algorithms, ultimately guiding users to entities identified through intent recognition. Using known initial and target topic sequences along with other dialogue data, reinforcement learning identifies which knowledge from the knowledge graph should be used to construct optimal response topics.

Most prior human-computer dialogue processes treat each interaction as a new conversation, with machine-generated responses failing to leverage historical dialogue information. In contrast, human-like dialogue typically incorporates conversation history when formulating replies [5]. Therefore, the response generation module employs Transformer networks for encoding and decoding input text, utilizing the network's self-attention layers [6] to examine context and relatively distant words in the input sequence during encoding. Compared to popular RNN networks, Transformer eliminates recurrent structures and temporal dependencies, enabling parallel processing of all words and symbols in a sequence, resulting in significantly faster training speeds and higher-quality generated responses. Rather than training the knowledge selection and response generation modules independently, joint training with reinforcement learning helps the model better utilize user input and knowledge graph outputs.

3. Intelligent-Driven Editing and Resource Integration

3.1 Intelligent-Driven Editing The supply-demand relationship in the media market is undergoing subtle changes driven by technology. Previously, what users consumed largely depended on what media organizations produced. However, as the “everything is media” era approaches, user-generated content has gradually gained market share, with the self-media industry proliferating rapidly—though not without issues such as fake news and lack of journalistic professionalism. Behind this phenomenon lies a direct hit on the pain points of traditional media content production inertia: insufficient editing and distribution speed, disconnect between content focus and public interests, and inability to capture audience preferences in content presentation formats. The *Opinions on Accelerating the In-Depth Development of Media Convergence* calls for the main force to fully advance onto the main battlefield and occupy new communication high ground. As mainstream media, the first priority must be synchronizing with user needs. In the new communication environment, user demands for “novelty” and “speed” have intensified. This cannot be solved simply by building manuscript editing and distribution platforms; rather, it requires leveraging intelligent capabilities to enable user participation in content production. Responsive news products precisely address this user need, generating greater resonance with audiences.

3.1.1 User Indirect Participation in Content Production During manuscript editing, semantic understanding of user-input content combined with big data aggregation capabilities can analyze historical and current user access data (reading, forwarding, commenting, liking, dwell time, etc.) for similar topics and feed insights back to editors. This approach can also

predict future content dissemination trends based on historical communication effectiveness and user preference data. Users participate invisibly in guiding content production, helping editorial staff create content that better aligns with audience tastes and generates greater interest.

3.1.2 User Direct Participation in Content Production Research indicates that WeChat, Douyin, and Toutiao rank as the most frequently used news channels among netizens, each characterized by social attributes and strong connectivity. Users increasingly value participation in news dissemination processes—the popularity of live streaming stems precisely from the sense of participation users gain through interacting with hosts, where dialogue can even change the thematic content of the broadcast, comments receive on-screen feedback, and users can communicate with fellow viewers through bullet comments. Responsive news represents a content production model featuring user participation, as illustrated in Figure 3 [Figure 3: see original paper]. Users submit explicit requirements to the intelligent assistant on news client applications (via voice, text, or video), which are transmitted back to the content production platform. Through intelligent capabilities such as natural language understanding and video theme capture, user needs are converted into content production requirements. The system then orchestrates intelligent media asset libraries, writing robots, intelligent editing tools, and other AI platform capabilities to rapidly produce text, images, videos, and other news content on corresponding topics, finally completing initial review through intelligent auditing before entering the pending publication repository.

3.2 Resource Integration and Convergence Novel and rapid content production cannot rely solely on editing systems alone. A series of subsystems related to content production must be coordinated to form a responsive content production system that adapts to new communication formats and better addresses market challenges. The golden era of profit-driven operations through print layouts has passed for the newspaper industry. The new media operations era demands higher standards for event organization, planning, and promotion. This requires integrating editing business systems, omnimedia assessment systems, and human resource systems to generate personnel profiles. For different thematic reporting activities, intelligent system capabilities can select appropriate personnel to form working groups, enabling collaborative resonance among management teams, content teams, and technical teams. This achieves comprehensive cross-media integration of human, financial, and material resources, converging media resources and production elements.

Amid rapid developments in 5G, AI, and big data technologies, the evolution from media convergence to intelligent media represents an inevitable trend. Media organizations at all levels have varying conditions for deploying intelligent capabilities, making the identification of suitable intelligent application strategies based on their positioning and resource allocation a key focus for each organization. New communication formats will continue to challenge newspaper

content production systems, which will face increasing demands for intensive, efficient, and intelligent capability integration and architectural flexibility.

References

- [1] http://www.gov.cn/zhengce/2020-09/26/content_{5547310}.htm
- [2] Zhou Xinyu. Chinese ASR Recognition Dialogue Short Text Intent Matching Based on Neural Networks [D]. Beijing: Beijing University of Posts and Telecommunications, 2020.
- [3] M. Mirza, and S. Osindero, “Conditional generative adversarial nets” , arXiv preprint arXiv:1411.1784, 2014.
- [4] Sun Guozi. Multi-Turn Question Answering System Based on Knowledge Graph [C]. 2020.4.
- [5] Xu Cong. Research on Dialogue Models Based on Deep Learning and Reinforcement Learning [D]. Beijing: University of Science and Technology Beijing, 2019.
- [6] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need [C]. Advances in neural information processing systems. Long Beach, 2017: 5998-6008.

Author Profile: Feng Enda (1985-), male, from Wuxi, Jiangsu, holds intermediate professional title in electronic information engineering. Research interests: backend development of information systems, intelligent capability applications.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv – Machine translation. Verify with original.