Postprint of Practical Application of Data Technology in Omnimedia Communication

Authors: Yuan Ran

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

Abstract

In the era of all-media communication, data is becoming an important factor of production. By focusing on innovations in data products and data service models, the media industry is continuously implementing new data-driven transformations. This paper outlines the technical practices of data collection, analysis, and presentation by the Dazhong Newspaper Group in all-media communication, as well as application practices driven by scenarios such as data journalism, think tank analysis, data evaluation, and public opinion services. It also examines the issues arising in data security, acquisition, authenticity, and other aspects, and proposes solutions.

Full Text

Abstract

In the era of all-media communication, data has become a crucial production factor. By focusing on innovations in data products and service models, the media industry is continuously practicing new data-driven transformations. This paper outlines the technical practices of Qilu Media Group (Qilu Daily Press Group) in data collection, analysis, and presentation for all-media communication, as well as application-driven practices in scenarios such as data journalism, think tank analysis, data evaluation, and public opinion services. It also reflects on the issues of data security, acquisition, and authenticity that arise from these practices, and proposes solutions.

Keywords: All-media communication; Data factor; Data security; Data acquisition; Data authenticity

2.2 Using Data Intelligence to Drive Social Progress

Media data service provider Baifendian Technology emphasizes new technologies and has constructed three core business systems: enterprise-level, government-level, and SaaS services, forming a comprehensive portfolio of big data, artificial intelligence, and online service technology products. The company proposes a construction approach for an all-media central platform based on big data technology, which requires consideration of organizational, management, implementation, and technical systems during the development process. The all-media central platform comprises four components: a big data center asset platform, a big data center capability platform, a resource publishing and display platform, and a resource service sharing platform [**?**].

2.3 Building a New Intelligent Content Ecosystem for the Data Economy Era

Media data services such as Fanwen Technology focus on news content big data to create a new intelligent content ecosystem. The company proposes an innovative model of "Big Data + Artificial Intelligence + Cloud Services," providing comprehensive services including content big data technical services, central data center construction, all-media production process transformation, content data product monetization, and communication impact analysis [**?**].

2.4 Searching Information, Integrating Content, and Discovering Knowledge

Media data service provider Tuoersi (TRS) launched full-text retrieval and content management platform products more than 20 years ago. In the all-media era, TRS has newly built an "intelligent production and communication service platform with content assets at its core" to reconstruct news production processes, enhance data value mining and empowerment, and help media organizations achieve integrated communication and service monetization. The platform is dedicated to empowering media with data and industries with knowledge [**?**].

2. Flourishing Development of the All-Media Data Industry

With the continuous advancement of media convergence in the all-media era, the all-media data industry has flourished, attracting increasing participation from research institutions and private capital. These industry players focus on big data products and services, deeply cultivating basic technology research and service model innovations in media big data. In April 2020, the "Opinions of the Central Committee of the Communist Party of China and the State Council on Building a More Perfect Market-oriented Allocation System and Mechanism for Factors of Production" explicitly identified data as one of the five factors of

production and called for building a digital economy with data as the key element. The Fifth Plenary Session of the 19th CPC Central Committee adopted the "Proposals of the Central Committee of the Communist Party of China for Formulating the 14th Five-Year Plan for National Economic and Social Development and the Long-Range Objectives for 2035," proposing to establish data resource property rights and promote the development and utilization of data resources. The "Outline of the 14th Five-Year Plan for National Economic and Social Development and Long-Range Objectives for 2035 of the People's Republic of China" calls for "embracing the digital era and unlocking the potential of data elements." In the 13th Five-Year Plan, big data was still an emerging technology industry; in the 14th Five-Year Plan, big data is becoming a factor, resource, driver, and concept integrated into all areas of economic and social development.

2.1 Data as the Eye of Insight

In today's era of "product as service," media data service provider Qingbo Intelligence emphasizes service products based on data. In February 2020, Qingbo launched the public welfare product "Epidemic Check" for epidemic prevention, which played a positive role in rapid response and communication among government, research departments, and the public [**?**]. The Qingbo Index, as a big data evaluation platform, establishes influence indices for various vertical fields with a focus on communication capability assessment. Its "New Media Management and Assessment System" integrates the entire process of "strategy, collection, editing, distribution, and evaluation," achieving integrated management and assessment of new media matrices. Through comprehensive services of its big data platform, Qingbo uses data to build an intelligent and innovative industrial ecosystem.

3. Practices in All-Media Communication Data Processing

In all-media communication practice, data exists throughout its entire lifecycle, from collection, cleaning, and storage to mining, analysis, and visual presentation. Each stage involves numerous data algorithms and processing methods.

3.1 Data Collection and Tagging in All-Media Communication

The massive generation of information data poses challenges to data collection, where the breadth and precision of collection determine the depth of subsequent data mining and analysis. Different needs and data types require different collection methods and tools. While data is typically categorized as structured, semi-structured, or unstructured, during the collection process it can be further subdivided into finer granularity and preliminarily tagged to facilitate later data mining and analysis.

3.1.1 Self-Developed Media Data Collection Tools Using web request mechanisms, self-developed web crawler software can obtain public data from the internet in real time. This data mostly comprises interaction information between content and users, representing real-time updated massive data in both structured and unstructured forms. By designing relational and non-relational data structures for storage, Qilu Media Group has independently developed software crawler tools using Python, deployed on distributed server clusters, to collect content data from authoritative sources into the group' s data resource pool for subsequent development and utilization.

3.1.2 Data Interface-Based Collection This approach obtains data by directly interfacing with business systems through data APIs. It requires extensive coordination to clarify the other party' s business processes and interface specifications, with significant time spent on initial integration and joint debugging, but yields high-precision, real-time data. This method is primarily limited to internal enterprise data. Qilu Media Group obtains real-time data on original manuscripts, hot news, and trending content by interfacing with partners such as Shandong Radio and Television' s "Shandian Cloud" and commercial platforms.

3.2 Data Analysis and Presentation in All-Media Communication

The trends and patterns formed through data analysis and mining represent the value of data. Presenting this value through data visualization is also a crucial part of data practice. Qilu Media Group continuously promotes the application of data in all-media production to enrich news content forms. The group' s award-winning work "Walking on the Yellow River Beach: My Relocation Story," which won the first prize in the China News Award, uses data materials as the main axis and fully integrates graphic illustrations, VR, and aerial photography technologies to create an innovative major thematic reporting work. The algorithms and new technologies used in this long-scroll page include the following:

3.2.1 Horizontal Scrolling Display Method for Long-Form Content By setting a fixed page height to establish a specific display area and configuring a large background image, the horizontal axis of the display area is set to scroll while the vertical axis hides portions exceeding the display area. Multiple sub-display areas are created within the main display area, with properties such as inline-table set for these sub-areas.

3.2.2 Positioning of News, Video, VR, and Aerial Footage on Long-Form Content Different arrays are defined to record different events and icons, with relative position values recorded within the arrays. When calculating positions, the total width of the background image and the current screen width are combined to form a new coordinate system to precisely calculate the absolute position of each event and icon in the display 画面 at the current time.

3.2.3 Multi-Terminal Compatible Display Method CSS style codes control compatible display effects, scaling each image proportionally to adapt to mobile phone screen widths and heights. When calculating relative positions, the total width of the background image serves as the coordinate system, with specific position data dynamically obtained and calculated while ensuring compatibility with different sizes of mobile and PC terminals.

4. Data Application Practices in All-Media Communication

As a modern media group producing news and information content, Qilu Media Group continuously experiments and explores all-media communication data applications.

4.1 Data Journalism

Data journalism has become a standard feature of daily media operations, focusing on "data analysis," "content planning," and "visual presentation," with the fundamental goal of telling good news stories. Data serves as the foundation of these stories and requires structured and knowledge-based organization. Qilu Media Group actively explores interactive and visual data journalism integrated into major thematic reporting planning, continuously launching data news products such as "Glorious 13th Five-Year Plan," "Remembering the Entrustment, Taking the Lead, and Pioneering Comprehensively," "Epidemic Prevention and Control," "2020 Two Sessions," and "2020 Shandong Provincial Flood Emergency Command Big Data," forming a unique production method that integrates technology and content in converged media product development [**?**].

4.2 Think Tank Analysis Services

By employing big data analysis technology combined with senior editorial staff for in-depth analysis, Qilu Media Group produces monthly data think tank analysis reports such as "Selected Compilation of Central Media Reports on Shandong," submitted to the Shandong Provincial Propaganda Department and provincial leaders. Through automatic monitoring of central media reports on Shandong, weighted calculations automatically capture high-attention articles and videos for database integration. After algorithmic matching and analysis by the platform, the reports form big data think tank analysis products containing heavy focus points, key manuscript dissemination quantities and paths, netizen heat, emotional tendencies, and regional distribution, all from the perspective of central media coverage. Based on this foundation, various thematic templates defined by the platform can quickly generate different data think tank products.

4.3 Data Evaluation to Improve Assessment Mechanisms

Performance assessment in all-media communication is more complex yet more humanized. Qilu Media Group' s editorial and reporter assessment regulations

have been implemented for many years. Under the integrated media convergence mechanism, the main force must enter the main battlefield, making the application of new technologies and new media the new direction. To reflect the guiding role of the command baton and introduce internet evaluation data, the group uses multi-dimensional indicators including original manuscript quantity, quality, communication effect, and influence to 统计 the workload and effectiveness of editorial, operational staff, and their departments. Assessment formulas are pre-established to calculate and generate royalty 统计 lists, making performance indicators clear and quantifiable, assessment results fair and transparent, and providing big data support for quantitative management.

4.4 Data Public Opinion Services

In all-media communication, the formation and development of public opinion can be clearly observed through public opinion service system monitoring. Qilu Media Group has always emphasized research and development in this area. After years of continuous iteration and updates, the current public opinion system serves various media within the group, and its public opinion services have expanded to cover major Shandong provincial departments and bureaus. Through the public opinion big data platform, real-time hot topics, regional distribution, and communication paths across the entire network can be clearly and quickly understood. Through continuous monitoring, multiple data types including text, links, images, videos, and audio can be collected to obtain public opinion information from massive data, filter out large amounts of invalid data, and provide timely warnings and 定向 tracking through manual analysis and judgment.

5. Reflections on Data Issues

5.1 Reflections on Data Security

On June 10, 2021, the Data Security Law was passed at the 29th session of the Standing Committee of the 13th National People' s Congress and took effect on September 1, 2021. This is a foundational law in the data field and an important law in the national security domain. The development of big data and cloud computing relies more on openness and sharing, while simultaneously facing security challenges. This is not merely a technical issue of data technology but also an issue of security awareness. The future of media is a contest of data capabilities, and data security is the foundation of data capabilities. In addition to technical security in infrastructure such as cloud, network, platforms, and data, legal security concerning data sensitivity, privacy, confidentiality, and copyright must also be considered, requiring a comprehensive data security protection system formed by law, regulations, standards, and technology.

5.2 Reflections on Data Acquisition

Obtaining valuable data from publicly available online information is challenging due to non-standardized data formats and varying degrees of data openness.

Regarding format issues, on one hand, crawled data needs to be processed, cleaned, and unified, while on the other hand, data standards can be designed and promoted to gradually regulate online data. Concerning data openness, the "Proposals of the Central Committee of the Communist Party of China for Formulating the 14th Five-Year Plan for National Economic and Social Development and the Long-Range Objectives for 2035" adopted at the Fifth Plenary Session of the 19th CPC Central Committee proposed expanding the orderly opening of basic public information data and building a national unified data sharing and opening platform. With top-level promotion by the Party and the state, data openness is expected to increase. For information with data property rights, needed data can be more efficiently obtained through purchase or cooperation.

5.3 Reflections on Data Authenticity

In the all-media era, where everyone has a microphone, the massive amount of data released by self-media has led to low credibility of online data. Although authoritative sources can significantly improve information accuracy, data timeliness is greatly compromised, and complete correctness cannot be guaranteed. Using multi-source comparison to improve data accuracy is a feasible method. If blockchain technology is introduced to establish a trusted data mutual trust mechanism, certified data would be trustworthy, greatly reducing the cost of obtaining correct data.

The new technologies, models, and business forms brought by data in all-media communication are gradually promoting in-depth integrated media innovation and development. With support and guidance from relevant policies, an increasing number of media organizations are entering the all-media communication data field, and more research institutions are investing R&D efforts, incubating data service models and entity companies. Although there are still growing pains in technology and application, as all-media communication practices deepen, more and more new applications, products, and services are being developed, becoming new drivers for the in-depth integrated development of media.

References

[1] Xi Jinping. Accelerating Media Convergence to Build an All-Media Communication Landscape [J]. Qiushi, 2019(6): 4-8.

[2] Miao Zhiru. Analysis of Key Points in Big Data Application Development in Major Public Health Events—Taking Qingbo Big Data as an Example [J]. China Media Technology, 2020(3).

[3] Yang Dong, Teng Jian. Practice of All-Media Central Platform Based on Big Data Technology [J]. China Media Technology, 2020(9): 19-20.

[4] Wang Ming. Big Data + Artificial Intelligence + Cloud Services: Technology Innovation Driving the Construction of County-Level Integrated Media Centers

[J]. Media, 2019(2): 17-18.

[5] Lin Songtao. AI Empowering Scenario Innovation in Media [J]. China Media Technology, 2020(4): 10-12.

[6] Tang Dailu. Exploration and Mechanism Thinking of 5G Converged Media Technology [J]. China Media Technology, 2020(8): 7-10.

Author Information: Yuan Ran (1972-), from Shouguang, Shandong, Senior Engineer, Information Technology Department of Qilu Media Group (Qilu Daily Press Group).

(Corresponding Editor: Chen Xuguan)

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