Management Studies, Nov.-Dec. 2023, Vol. 11, No. 6, 322-328
doi: 10.17265/2328-2185/2023.06.003

The Path to Reshaping the Structure and Chain System of the Publishing Industry in the Era of Big Data

DENG Liyan
Guangxi University of Foreign Languages, Nanning, China

Today, we are living in the era of "big data" where massive amounts of data are used for quantitative decisions and communication management. With the continuous penetration of big data-based intelligent technology in all fields of human life, the enormous commercial value inherent in the data industry has become a crucial force that drives the aggregation of new industries. For the publishing industry, the introduction of big data and relevant intelligent technologies, such as data intelligence analysis and scenario services, into the structure and value system of the publishing industry, has become an effective path to expanding and reshaping the demand space of publishing products, content decisions, workflow chain, and marketing direction. In the integration and reconstruction of big data, cloud computing, artificial intelligence, and other related technologies, it is expected that a generalized publishing industry pattern dominated by virtual interaction will be formed in the future.

Keywords: the era of big data, data analysis, scenario service, virtual publishing

On December 14, 2020, Beijing Publishing Group announced to initiate a deep-level strategic cooperation relationship with Urban Brain Co., Ltd., Beijing Zhongguancun Science City. The two parties would carry out multiple forms of cooperation in the publication of smart management systems, research on publishing big data, and the development and construction of a "publishing brain". On the part of Beijing Publishing Group, the incorporation of technological "hard power" this time would undoubtedly greatly facilitate the deep integration of the publishing industry chain and the information service technology chain with big data technology at its core, leaving a profound imprint of "big data technology" on the transformation of the structure and chain system of the publishing industry in Beijing. In the context of "the era of big data", the deep transformation of Beijing Publishing Group into the big data system is by no means an individual case. A large number of publishing companies have chosen to introduce big data and related intelligent technologies into their own product structure and business chain operation mode, as an important breakthrough to enhance their competitiveness and reconstruct the industrial value.

Status Quo of the Publishing Industry in the Era of Big Data

Definition and Application Background of "Big Data"

"Big data", also known as massive data, is a concept originated from the IT field, and refers to huge data group collected in multiple forms through multiple channels. It has such characteristics as huge amount of data, real and effective data sources, complex and diverse channels, and fast data analysis and processing speed and is

DENG Liyan, M.Sc., associate professor, Guangxi University of Foreign Languages, Nanning, China.

the outcome of the rapid development of information technology in the contemporary society. The concept of big data has been concerned about by the general public, largely due to the increasing popularity of Internet information technology. The Internet has evolved into a basic infrastructure for storing massive data and embodying the application value of big data and related intelligent technologies. In reality, in modern social life, any behavior and transaction of humans can be digitized and electronicized. Big data and related intelligent technologies have deeply explored and utilized the value of the data industry, on the basis of the collection, acquisition, screening, analysis, induction, and organization of various types of data generated in human life.

Today, people define and describe this information age that is supported and led by massive data and related innovative technologies by using the term "the era of big data". The technological background of big data also provides a driving force for the integration and upgrading of the industrial form for the publishing industry. Starting from 2014, the term "big data" has been included in the government work report successively, and the related concepts of big data have served as a guiding pointer for the adjustment of the layout of ideological and cultural industries. In 2015, with the issuance of the "Guiding Opinions on Promoting the Integrated Development of Traditional Publishing and Emerging Publishing" by the former State Administration of Press, Publication, Radio, Film, and Television and the Ministry of Finance as the landmark, publishing companies began to deeply explore the use of big data and related intelligent technologies for data analysis, scenario services, and other functions, with a view to tap the new value embedded in the publishing industry brought by "big data".

Transformation and Upgrading of Digital Publishing Driven by Big Data

Up to now, the publishing and communication industry in China has gone through the stage of digitized publishing, which involves the full-course digital processing of information content and industrial models with the help of computers or terminal devices. On this basis, the introduction of the concept of big data technology and conducting the transformation and upgrade of digital publishing has become an inevitable path to adapt to the times, embrace the changes of the thinking and behavior of the information industry. As a consequence, driven by big data and empowered by related intelligent technologies, a group of professional houses holding literature and knowledge resources have embarked on the transformation and upgrade of digital publishing, backed by financial support from government policies. They apply big data analysis, intelligent indexing, semantic search, and other related technologies to develop professional database publishing platforms that provide academic analysis, knowledge hotspot correlation, and other services for professional readers in specialized fields. For example, notable professional database platforms, such as Zhejiang University Digital Publishing Service Platform developed by Zhejiang University Press, Faxin Platform (www.faxin.cn) established by People's Court Press, and Architecture Digital Library of China Architecture & Building Press are all dedicated to offering one-stop academic services, encompassing information retrieval and research support, catering to users in fields such as education, law, architecture, and other specialized domains.

At the same time, some technology firms have also joined the fray with cutting-edge data analysis, intelligent indexing, and other technological advantages. On the one hand, they provide users with access to a vast amount of knowledge and data. On the other hand, their development focus has shifted to the integration and innovation of such knowledge and data for users. This shift has led to the establishment of well-known academic professional database platforms such as CNKI, qikan.com, CQVIP, and others, highlighting the innovative application trajectory of big data technology within the knowledge publishing industry.

Starting from 2016, Internet knowledge-paid service platforms have emerged, becoming a new vane for the integration of the digital publishing industry. Publishers are increasingly incorporating advanced technologies such as big data analysis, cloud computing, and artificial intelligence into the construction of comprehensive knowledge service platforms, to achieve the transition and upgrade from a professional knowledge digital content system to a comprehensive big data service platform. During this stage, knowledge-paid service platforms such as "Zhihu", "Dedao", "Ximalaya", and "Fandeng Reading" showcased high-quality content in a more refined manner, by leveraging big data resources on the Internet. These platforms offered integrated value-added services to users, such as information retrieval, reading information, entertainment and leisure, and knowledge elevation. At the same time, the publishing products were also expanded to encompass multiple fields such as online literature, audiobooks, animation, music, film, and television, thus forming a trend of cross-industry integration in the publishing industry. By the year 2020, the publishing industry has witnessed a convergence with state-of-the-art technologies such as big data analysis and decision-making, artificial intelligence scenario services, AR/VR virtual reality technology, and continued to optimize the product forms and industrial models, etc. This led to the emergence of publishing products in multiple formats, such as "natural resource big data platform", the AR book series of Suzhou Dreamers, and the science popularization products of HesiyiVR. The development path of the publishing industry, led by big data, is becoming increasingly clear.

Path to Adjusting the Structure of the Publishing Industry and Reshaping the Industry Chain With Big Data and Related Intelligent Technologies

At the current stage, technologies such as big data and, artificial intelligence have deeply penetrated all aspects of the publishing industry. These advancements allow for the precise identification of user demographics, collection and planning of publishing topics, market decision-making, and product marketing, thereby forming an effective path to reshaping the structure and chain system of the publishing industry.

Big Data Technology Promotes the Expansion of User Demand Space and Drives the Decision-Making of "Scenario-Based" Content Production

For the publishing industry, it is the choices and preferences of readers that ultimately determine the direction of the publishing market. However, in traditional publishing, it was impossible to define reader demands in a scientific manner. Editors and publishers can only rely on intuition, experience, and even luck to analyze reader demand space, based on limited samples of reader feedbacks. As a result, many best-selling books such as *Harry Potter* and *Sapiens: A Brief History of Humankind* almost failed to see the light of day due to misjudgments by publishers regarding reader demands. By contrast, in the era of big data, publishing companies are able to conduct market surveys using more scientifically efficient methods, for example, analyzing big data sample. This includes analyzing the real-time online search habits, reading duration, feedback, and purchase records of readers to determine their product preferences. What's more, data from positioning systems and sensors, etc. can also be adopted to perceive, record, and depict the current offline spatial location and actions of reader users. This is a new gift brought by big data and related intelligent technologies—"scenario data". Subsequently, there has been a three-dimensional description and all-round expansion of user demand space. To provide accurate and personalized "scenario-based services" has become the forefront of decision-making in the publishing industry in the era of big data.

In the era of big data, "scenario data" is a brand-new and important source of information. The "scenarios" include the content perceived by both "online virtual scenarios" and "offline real scenarios". The perception of online scenarios primarily involves sensing the virtual behaviors of users, including online browsing and social behaviors, etc., thereby determining relatively quantifiable data, such as the gender, age, hobbies, habits, and social relations of individual readers and users. The perception of offline scenarios tracks and analyzes the actual behaviors of readers and users and spatial scenarios that change at any time as the variable data, so as to enrich the existing user quantitative data module in a real-time manner, and thus achieve an "accurate portrayal" of immediate user demand in this particular time and location.

In essence, in the era of big data, efforts are made to precisely target the needs of each individual "person" as every "small data", and expand the space of early generalized and categorized reader and user needs in an all-round way. With strong support from big data analysis and other technologies, publishing companies can accurately grasp the personalized "scenario demands" formed by the behavioral habits, hobbies, and living conditions of users. In this way, they can make scientific and precise decisions on content production, including the determination of publishing topics, the control of the content to be pushed, as well as marketing decisions, such as promotion channels and interfacing strategies. This has brought equal publishing and interfacing opportunities for almost all publishing topics and works. As a consequence, we can witness the promotion of "Zhaofeng's Economics Lecture Notes" on the "Dedao" platform, which became a bestseller after getting the recognition of users, and also the exposure of a niche work entitled "Anonymous Zone". This work was a collection of emotionally-driven stories written by more than 1,000 "anonymous users" with high readership compiled by the "Zhihu" platform. In the process of making decisions on content, publishers use targeted big data analysis, AI interaction, and other technologies to provide content that suits different types and forms of publications, such as "learning", "sports", "leisure", "emotion", etc., based on the contextual dimensions of each user. This approach satisfies and serves the users' scenario states and immediate needs, thereby deeply improving their product experience, and realizing the function of scenario services under the big data technology. In the future, the publishing industry will increasingly move towards personalized and intelligent scenario-based services driven by big data and related technologies.

Empowering and Optimizing the Processes of Publishing Industry With Big Data Technology

In the traditional publishing activities, the business processes of the publishing industry mainly revolved around three elements: editing, printing, and distribution. However, as the publishing industry enters the stage of digital publishing, publishing information elements are all presented in the form of data. Based on precise positioning of user scenarios, big data and related technologies such as artificial intelligence are deeply integrated into the specific business processes of the publishing industry, such as editing, proofreading, printing, transportation, and sales. Through the intelligent monitoring, collection, and analysis of publishing information data at various steps, they facilitate the coordinated production of all elements, leading to the restructuring of the whole industrial process and the improvement of the overall efficiency of the publishing sector. For example, in the process of editing and proofreading, big data and artificial intelligence technologies, etc. can be adopted to typeset, edit, and enter data intelligently, and even conduct typesetting and editing in different formats simultaneously, such as electronic books, audio and video production, etc., so as to produce publications in different formats. In the processes of printing, transportation and sales, unified analysis across platforms can be achieved through big data analysis, and multi-dimensional analysis, including the classification and ranking of

publications, sales information, and logistics and transportation, can be carried out to determine the optimal number of copies to be printed and sales channels, thereby maximizing the benefits of publishing.

In addition to specific business processes, the constituent elements of the publishing industry chain are also becoming more and more complex, and publishing entities, such as content providers, technology service providers, content publishers, channel distributors, and equipment manufacturers cross and permeate each other, actively promoting the transformation from data production to data integration platforms. The specific approach is to connect the content production, digital operation and promotion, physical and digital distribution platforms, and other constituent elements of the industry of publishing products in series, control the big data information of publishing and distribution, and build a functionally-improved publishing and distribution big data platform with a closed loop for supply and demand. This serves as an effective path to achieve the intelligent upgrade of the chain of publishing industry. For example, Inner Mongolia Education Publishing House has developed a comprehensive content database and application platform that can store and manage its own books, materials, and resource in a digital way, edit and produce intelligently, and sets up a multi-purpose database through one-time production and multiple releases, so as to offer intelligent paper and electronic books for readers and users. China Publishing Group Digital Media Co., Ltd. has developed "Dajia Reading Platform" targeted at high-quality reading. It worked with professional and authoritative publishers, such as Zhonghua Book Company and People's Literature Publishing House, to attract numerous readers by intelligently recommending high-quality content after big data analysis. As a matter of fact, traditional publishing organizations with limited technological capabilities can also ride and wave and benefit from the empowerment of big data technology through cooperation and win-win and other strategies, by taking advantages of their content ownership. For example, in March 2020, the Baidu Intelligent Cloud joined hands with Zhiduiliu to create a platform called "Digital Intelligent Publishing Cloud". They made use of the powerful big data and cloud computing technologies of Baidu Intelligent Cloud, as well as the software development capabilities of Zhiduiliu to help traditional publishing organizations produce and distribute digital publishing products in a targeted manner. Xiaoe Tech, a technology service provider, offered UI design, mini-program development, and other technical services for many publishing companies with online business needs. CITIC Publishing Group also came up with a publishing platform plan in 2020, with an aim to leverage its strong financial and technological edges to cooperate with small and medium-sized publishing enterprises in multiple ways, and help them complete transformation and upgrading towards digitalization. Through collaborative efforts among different publishing entities, the publishing industry is able to aggregate big data information resources of the publishing industry, and develop a comprehensive and interconnected chain that combines various publishing elements in an open, multidimensional, and networked manner.

On the whole, big data technology has empowered the process of the publishing industry. This not only enables accurate docking between supply and demand sides and effectively improving the efficiency of the publishing industry, but also helps publishing companies prolong their industry chain, connect the content creators, producers, and processors, technical service providers, and end-reader and users through big data publishing platforms, ultimately improving production and operation management, as well as the profitability of the whole publishing industry.

Application of Big Data Technology Drives the Expansion of the Boundaries of Publishing and the Direction of Virtual Publishing

Nowadays, the analysis and application, data carriage and scenario service capabilities of technologies, such as big data, cloud computing, and artificial intelligence have been continuously improved, which makes the

publishing industry gradually form a new content ecosystem. The creation methods, presentation styles, and communication forms of publishing works have also become more and more diversified, with text, pictures, voice, video, and other elements serving as tangible carriers of publishing content. The boundaries and domain of the publishing products have been expanded with each passing day. At the same time, the knowledge service attribute of the publishing industry is increasingly prominent, and the publishing knowledge service system centered on users' scenario demands is evolving into a more sophisticated state. Based on the continuous integration of publishing information resources into the big data publishing platforms and the development towards a diversified, personalized, and service-oriented direction, comprehensive knowledge-paid service platforms such as Ximalaya and Zhihu have become new publishing paradigms and formed a one-stop platform for value aggregation and monetization. These platforms capture real-time needs of users in different time and space scenarios as their data foundation, and match and push knowledge services to users in an intelligent way. In the process of providing knowledge services, publishing platforms focus on reading, integrating, refining, and transforming massive fragmented data by using technologies such as big data and artificial intelligence, and then offer filtered product content in different forms that can match the scenarios to the users. This has also become a new growth point for the publishing industry in terms of market and marketing in the era of big data. With the deep fusion of technologies such as big data processing, cloud computing, VR/AR, and intelligent man-machine interaction, the forms of publishing products provided by the knowledge service platforms are increasingly diversified and aligned with the scenario experiences and emotional needs of readers. The publishing industry is undergoing a transition from physical publishing and digital publishing to the stage of virtual publishing in broad sense, that is, "intelligent technology + publishing". For example, there are audio publications that can free users from visual constraints, fill in and piece together the fragmented time of users, and even be controlled by voice. Also there are VR/AR publications that can provide immersive sensory experience, construct reading scenes that intertwine between virtual and reality, and offer users new and exciting reading experiences through their eyes and mind, etc. All of these have embodied the trend of "virtualization" in the publishing industry. In the future, publications will be capable of dynamically adjusting to users' needs based on their time and location, and offering readers with formats of knowledge product and service that meet their desires. Big data and related intelligent technologies have offered essential technical elements for the virtualization of the publishing industry layout.

Conclusion

In the environment of big data, big data and related intelligent technologies have penetrated and improved the efficiency of the publishing industry, reshaped the structure and chain system of the publishing industry, and also participate in the influence on the future development of the publishing industry. Moving forward, big data and related intelligent technologies will further strengthen their core competitiveness in terms of the sharing and openness of data resources, as well as the interactivity in the era of intelligence. They will drive the industry with technology, empower the industrial efficiency, service capabilities, and chain system of publishing companies, and ultimately create a virtualized and intelligent industrial ecology that can deliver a perfect personalized reading experience to users.

References

Gong, M. Y. (2018). The development trend of the digital publishing industry in the era of big data. *Research on Transmission Competence*, *2*(31), 147.

Sun, Y. (2019). Analysis of the development trend of the digital publishing industry in the era of big data. *Research on Transmission Competence, 3*(36), 210.

Sun, Y. L. (2017). Prospect of the development of the digital publishing industry in the era of intelligent technology. *Information on Publication, 30*(9), 13-15.

Wan, A. L., & Wang, J. F. (2019). Virtual publishing: A duet from the technological revolution to reading scenes. *Journal of Hebei University (Philosophy and Social Science), 44*(1), 146-153.

Wang, H. (2019). The future trends of the digital publishing industry in the age of big data. *Media, 21*(18), 61-63.

Zhang, H. D. (2019). Empowering the development of the digital publishing industry with big data technology. *Journalism & Communication, 26*(1), 57-58.