Exploring Convergent Innovation Pathways between Frontier Science and Technology and Book Publishing (Postprint)

Authors: Liu Yazhen

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

Abstract

【Objective】 As a book planning editor in a large state-owned enterprise, this work integrates cutting-edge scientific and technological knowledge with book publishing to provide high-quality books for readers and expand the dissemination of frontier scientific and technological knowledge. 【Methods】 The article adopts an approach that treats books as products, with artificial intelligence as the thematic direction and intelligent speech as the breakthrough point, recognizing that speech is the most natural communication modality between humans and machines, and that natural language processing, speech signal analysis and recognition have always been important research hotspots in artificial intelligence. Through specific practical cases, it innovatively enhances the quality and speed of knowledge dissemination in the market, which can serve as a reference for new pathways in book publishing and planning. 【Results】 The cutting-edge technology books planned in the cases have all been published and distributed, and within a relatively short period, have already been reprinted multiple times, receiving positive feedback from both the market and readers. 【Conclusion】 The application of cutting-edge science and technology can not only improve people' s living standards, but also record important scientific research achievements and popularize knowledge through book publishing, thereby accelerating the speed of knowledge dissemination, enhancing cognitive levels, and enabling the mastery of scientific and technological knowledge.

Full Text

Abstract

[Purpose] As a book planning editor at a large state-owned enterprise, this article explores the integration of cutting-edge scientific and technological knowledge with book publishing to provide high-quality books for readers and expand

the dissemination of frontier scientific knowledge. [Method] Adopting a product-oriented approach with books as the medium, the article selects artificial intelligence as the thematic direction and intelligent speech as the breakthrough point, recognizing that voice represents the most natural form of communication between humans and machines. Natural language processing, speech signal analysis, and recognition have consistently been important research hotspots in AI. Through specific practical cases, this approach innovatively enhances the quality and speed of knowledge dissemination in the marketplace, serving as a reference for new book publishing and planning pathways. [Results] The cutting-edge technology books planned in these cases have all been published and distributed, undergoing multiple reprints within a short timeframe and receiving positive feedback from both the market and readers. [Conclusion] The application of frontier science and technology can not only improve people' s living standards but also record important scientific research achievements and popularize knowledge through book publishing, thereby accelerating knowledge dissemination, elevating cognitive levels, and enabling mastery of scientific and technological knowledge.

Keywords: Artificial Intelligence AI; Intelligent Speech; Book Publishing; Integrated Innovation

Classification Codes: G258.3; Document Code: A; Article ID: 1671-0134(2023)04-145-05; DOI: 10.19483/j.cnki.11-4653/n.2023.04.030

President Xi Jinping has emphasized that "core technologies are the nation' s most important tools" and that we must "make determined efforts, maintain perseverance, and identify the right focus to accelerate breakthroughs in core information technologies." According to estimates, China' s AI industry reached a scale of 404.1 billion yuan in 2021, representing a year-on-year growth of 40.4%. The integrated application of AI technology has propelled industrial transformation and upgrading to higher levels and efficiency, injecting strong momentum into high-quality economic and social development. Benefiting from the rapid development of 5G networks, big data technological support, and broad application scenario demands, China has experienced rapid growth in artificial intelligence, speech recognition, and related fields. As book planning editors, we should seize upon the topics of artificial intelligence and speech recognition to create a platform that consolidates forces from industry, academia, research, and application, bringing together a series of thematic publications on intelligent speech for the latest developments in frontier technology, innovative AI applications, reader enlightenment, knowledge sharing, and exchange.

1.2 Policy Guidance

Intelligent speech technology represents one of the earlier key technologies in AI development and constitutes a frontier technology sector receiving policy support from various countries. Since the release of policy documents such as the "New Generation Artificial Intelligence Development Plan," the government has

demonstrated high-level attention to new technologies like AI. In 2017, a New Generation AI Development Plan Promotion Office was established, comprising 15 departments including the Ministry of Science and Technology, National Development and Reform Commission, Ministry of Finance, Ministry of Education, Ministry of Industry and Information Technology, Chinese Academy of Sciences, Chinese Academy of Engineering, Central Military Commission Science and Technology Committee, and China Association for Science and Technology, focusing on coordinated planning of projects, bases, and talent deployment.

In December 2017, the Ministry of Industry and Information Technology released the "Three-Year Action Plan for Promoting Development of the New Generation Artificial Intelligence Industry (2018-2020)," which explicitly supported innovative applications of new-generation speech recognition, intelligent dialogue, audio-video fusion, and speech synthesis technologies. Additionally, in May 2018, the Ministry of Industry and Information Technology and the Anhui Provincial Government signed the "Cooperation Agreement on Further Jointly Promoting the Development of Anhui' s Intelligent Speech Industry," proposing to build "China Sound Valley" into a national highland for intelligent speech industry development within approximately five years. AI and traditional industries are accelerating their deep integration.

1.3 Industry Environment

Benefiting from the rapid development of AI technology, the global intelligent speech industry is experiencing vigorous growth. Traditional products and applications based on intelligent speech have seen further optimization in experience and service, while new business scenarios continue to emerge and fusion applications of speech interaction technology constantly update, with smart healthcare, autonomous vehicles, smart communities, smart homes, smart education, smart airports, and smart ports appearing in large numbers. As one of the next generation' s most important human-computer interaction interfaces, enterprises are joining the intelligent speech field to seize this technological high ground.

Currently, the world attaches great importance to AI and provides policy support for related technologies such as intelligent speech technology R&D and industrialization. With continuous development in AI technologies like big data analysis, cloud computing, and deep learning, the widespread application of intelligent speech technology has gained strong support. The global speech market is showing explosive growth, particularly with OpenAI's launch of the ChatGPT conversational AI software. This language algorithm model generates responses through one-on-one, question-and-answer style dialogue between users and the platform. The model can provide intelligent answers in real-time based on user input, following the system' s language algorithm model and existing databases, greatly enhancing AI' s ability to converse with customers and revealing the dawn of large-scale AI implementation.

Currently, in the intelligent speech domain, foreign technology giants have seized

the initiative, accelerating the expansion of AI-centered ecosystems within their existing products and businesses. Meanwhile, domestic internet companies are also seizing the opportunity, increasing investment, and building fusion innovation in the intelligent speech field.

2. Book Topic Positioning

The positioning of book topics aims to help readers establish the connection between human speech production and machine speech production, starting with voice as the starting point. The content revolves around three core elements: technology, industry, and application, introducing basic knowledge of intelligent speech technology in a popular science format, and illustrating current industry development status and trends through practical application cases. This enables readers to comprehensively understand AI speech technology, uncover the essence of technological development, its history, and methods for application transformation. Simultaneously, it summarizes new models and cases of industrial application, providing market planning and strategic layout for enterprises, and further driving market enthusiasm for the development of the intelligent speech industry.

The book "AI Intelligent Speech Technology and Industrial Innovation Practice" takes three dimensions—technology, application, and industry—as its entry point to research hotspots and trends in the intelligent speech and semantics field. The planning editor first identified experts from the China Academy of Information and Communications Technology in the intelligent speech domain, using this theme as a topic for in-depth discussion to explain intentions and identify the first author for the AI intelligent speech book. This author discusses current development status and research situations in intelligent speech from a government perspective. After quickly agreeing on the book theme with the first author, the planning editor discussed the composition of the writing team and sought participation from other industry experts.

The book begins with "Dialogue Between Humans and Machines," describing the process of human speech generation, transmission, and perception, triggering reflection on machine intelligent speech listening and speaking. It then elaborates on the history of technological exploration. Centered on this core content, the editor visited the Institute of Linguistics at the Chinese Academy of Social Sciences and found researchers in phonetics, grammar, natural language processing, semantics, and corpora. After explaining the purpose of the visit and the book publishing plan, the researcher immediately agreed to participate. He has chaired and participated in projects funded by the National Social Science Fund, National Natural Science Fund, National 973 Program, and U.S. National Science Foundation, as well as joint research projects with companies such as iFlytek, Nokia, and Motorola. He expressed great interest in the AI intelligent speech thematic book publishing plan and agreed to participate as an editorial board member, writing content for relevant chapters.

2.2 Using Book Market Analysis Data as Basis for Topic Planning

2.2.1 Market Capital Investment Analysis Major domestic and international manufacturers are actively deploying intelligent speech technology and industry. Technology giants such as Google, Apple, and Microsoft are committed to building speech ecosystems, while internet enterprises like Baidu, Alibaba, and Tencent are transforming their AI foundational capabilities. Traditional speech companies like iFlytek and AISpeech are focusing on industry application integration. In recent years, startups such as "Chumen Wenwen"and "Orion Star" have also emerged, creating a diversified intelligent speech ecosystem. As intelligent speech technology gradually matures, industry and application demands for intelligent speech will continue to expand, providing momentum for large-scale AI implementation from the interaction endpoint.

2.2.2 Analysis of AI-related Books in Market On JD.com' s book best-seller list, "Artificial Intelligence: From Beginner to Master" published by China Water & Power Press in January 2021 has over 20,000 reviews with a 99% positive rating. "The Age of Intelligent Speech: Business Competition, Technological Innovation, and Virtual Immortality" published by Electronic Industry Press in May 2019 has over 20,000 reviews with a 99% positive rating. "Speech Recognition: Principles and Applications" published by Electronic Industry Press in July 2020 has over 30,000 reviews with a 98% positive rating.

The reasons are as follows: First, books on AI intelligent speech industrial applications are relatively scarce in the book market, allowing us to leverage this scarcity to enter the market. Second, the author team consists of technical experts from authoritative domestic research institutes, operators, and higher education institutions, ensuring original, high-quality content. Third, from the perspective of content demand analysis, the books dissect the internal reasons for technology research and analyze its development direction, providing high reference value. Fourth, they fully utilize the innovative theories and practices of the author team, incorporating the latest thinking and summaries from multi-party "industry-academia-research-application" collaborative work. Fifth, the content is rigorous, referencing and citing foreign monographs and literature to facilitate further extended learning for readers.

Based on the above market analysis and current industrial applications of intelligent speech, China Industry and Information Technology Publishing and Media Group planned the book "AI Intelligent Speech Technology and Industrial Application Innovation Practice."

The book was published in December 2021. As of January 2023, within one year, it has sold nearly 3,000 copies, providing a platform for knowledge sharing through the medium of books for engineering technicians and managers engaged in AI technology R&D, product application, and market planning, as well as for university faculty and students in AI-related majors and individuals interested in intelligent speech technology. The book has received recommendations,

endorsements, and positive feedback from experts in AI-related fields.

Among them, the Director of the Cloud Computing and Big Data Research Institute at the China Academy of Information and Communications Technology recommended: "In the process of AI engineering moving from academia to industry, 'integration' is particularly important. To make AI a scientific engineering practice, it needs to integrate and develop with related surrounding technologies. Taking intelligent speech as the entry point, the book emphasizes the importance of ecological platforms, standard construction, and scenario integration from three dimensions—technology, application, and industry—combining multiple hot topics such as big data, cloud computing, security, and multimodality, making AI more 'down-to-earth' through innovative practice." The Director of the Auditory Intelligence Research Center at Tsinghua University' s AI Research Institute, founder of Deli Yintong, and Dean of the Deli Yintong Research Institute commented: "Combining intelligent speech technology with industrial application practice, starting from basic theory, introducing several different branches of speech processing technology, explaining national policies, standard formulation, and industry implementation practices in speech technology, while also providing prospects for future integration with other technologies, this meets the current needs of academia and industry to further integrate technology into industry. The tide is rising and the wind is favorable for the industrial development of intelligent speech. Universities are continuously cultivating talent entering the field. I hope all personnel dedicated to the theoretical exploration and industrial development of intelligent speech processing technology will ride the wave on this broad road and create a new era of intelligent speech!"

2.3 Developing Series Books Based on Market Industry Environment —Case 2

2.3.1 Technical Hotspot Background Analysis Intelligent speech is one of the important core foundational technologies of artificial intelligence. Governments, related industries, and capital sectors worldwide continue to pay attention to the R&D and industrialization of intelligent speech technology. All countries are striving for breakthroughs in key technologies in related fields, hoping to establish their own technological development as the standard, seize market opportunities, and thereby gain leadership and discourse power in the "government-industry-academia-research-application" ecosystem of intelligent speech.

Based on machine learning, AI technology development mainly focuses on three directions: speech recognition, natural language understanding, and visual recognition. From the perspective of the intelligentization process, it is divided into three development stages: computational intelligence, perceptual intelligence, and cognitive intelligence. The first stage is computational intelligence, the initial stage of AI development, referring to computers' fast computing and massive storage capabilities. With the development of grid computing, distributed computing, and other technologies, machine computing

power has far surpassed humans, laying a solid foundation for the subsequent two stages. The second stage is perceptual intelligence, which has now reached relative maturity. Perceptual intelligence uses auditory, visual, tactile, and other sensory systems to understand the external world, enabling machines to acquire "listening, speaking, and seeing" capabilities—currently achieved through various technologies that imitate human organs. Its main scope includes facial recognition (seeing), natural language understanding (listening), speech synthesis (speaking), speech recognition (listening), image recognition (seeing), etc. Currently, machine perceptual intelligence technology has become very close to human capabilities, as exemplified by iFlytek' s "Xunfei Tingjian," Xiaomi' s voice assistant, and Microsoft' s facial recognition system. The third stage is cognitive intelligence, the highest form of AI development. The two cores of cognitive intelligence are "understanding" and "explanation," enabling machines with the ability to "comprehend semantics, perform logical reasoning, and learn to make judgments." In February 2023, the globally popular ChatGPT is a brand-new intelligent speech product with cognitive intelligence, allowing machines to think like humans! This thinking ability is manifested in machines' capacity to understand language, interpret real-world problems, read data, and explain phenomena. Therefore, realizing cognitive intelligence requires knowledge as the driving force, combined with natural language understanding technologies, to enable machines to possess more human-like emotional understanding, semantic comprehension, associative reasoning, affective computing, decision-making planning, and other core technologies.

Against this technical background, the author visited Beijing Language and Culture University and had the privilege of meeting Professor Xun Endong, Dean of the School of Information Science and Director of the Institute of Language Intelligence. His research focuses on natural language processing and language education technology, and he has chaired multiple "National High-Tech Research and Development Program," "National Social Science Fund," and "National Natural Science Fund" projects, as well as undertaken numerous corporate collaborative projects, primarily researching Chinese semantic computing and language education technology. During communication and exchanges with Dean Xun, the author discussed the planned thematic book on AI intelligent speech, and they reached a consensus to develop three books in the "Natural Language Structure Computation" series around language structure from three aspects: analysis technology algorithm theory, language knowledge database, and knowledge extraction (database construction): *Natural Language Structure Computation—GPF Language Analysis Framework*, *Natural Language Structure Computation—BCC Corpus*, and *Natural Language Structure Computation—Theory and Technology of Yihe Graph*. This series mainly elaborates on the theoretical knowledge and engineering construction ideas regarding the algorithm layer and data layer during the perceptual and cognitive stages of natural language processing.

Benefiting from the rapid development of computer technology and text recogni-

tion technology, the collection and organization of corpora have become faster and more convenient, making the construction of large-scale Chinese corpora possible. Since 2016, corpus construction has risen to the national strategic level, with related policies entering a comprehensive explosive period. The release of policies such as the "National Long-term Language and Script Career Reform and Development Plan Outline (2012-2020)" demonstrates the government' s high-level attention to new technologies like corpus construction. Against this backdrop, the Beijing Language and Culture University Corpus Center (BCC) emerged. Since its construction in 2015, the BCC corpus has played an important role in computational linguistics and linguistic ontology research. The author planned and published the book *Natural Language Structure Computation—BCC Corpus*, which provides a detailed introduction to the BCC corpus' s content, construction process, retrieval functions, query languages, programming languages, and how to use custom BCC corpora, hoping that the BCC corpus can be better understood by more people and provide better services for language ontology and teaching research.

Based on comprehensive market analysis and the current state of corpus construction and application in academia, these books primarily focus on artificial intelligence. There are not many book categories describing speech recognition, natural language processing, and corpora, and books on speech recognition and corpora were published earlier with fewer new titles, leaving readers with limited choices. The three books in the "Natural Language Structure Computation" series have good timing and value for publication.

Natural Language Structure Computation—GPF Language Analysis Framework was published in November 2022, selling over 1,300 copies in just over three months. *Natural Language Structure Computation—BCC Corpus* was published in January 2023, with over 1,000 copies sold. *Natural Language Structure Computation—Theory and Technology of Yihe Graph* was published in March 2023, with over 800 copies sold. The series has received enthusiastic preface recommendations from Liu Ting, Council Member of the China Computer Federation, Vice Chairman of the Chinese Information Processing Society of China, and Professor at Harbin Institute of Technology: "The author has been deeply engaged in natural language processing for over 30 years. His innovative spirit embodied in these three series books also fills me with admiration. Facing China' s language and script, his monographs demonstrate the confidence and character of Chinese scholars, rather than forcing things to fit or following the crowd. The academic ideas and professional knowledge contributed can bring unique inspiration to scholars, engineers, and students in the natural language processing field."

2.3.2 Analysis of Market Economic Development Status Related to Natural Language Processing

2.3.2.1 International Market Governments in the United States, the United Kingdom, and other countries have established special projects with substantial funding to support the construction of corpora for their respective languages. For example, the "Lancaster Corpus of Mandarin Chinese (LCMC)" is a research project undertaken by the Department of Linguistics at Lancaster University and funded by the UK' s Economic and Social Research Council.

2.3.2.2 Domestic Industry Market Major national social science projects and the National Social Science Fund provide project approval support for related corpus construction. For instance, major National Social Science Fund projects such as "Corpus Construction and Comprehensive Comparative Study of Hakka Dialects at Home and Abroad" and "Corpus Construction Research on Unearthed Ancient Documents."

The Beijing Language and Culture University Corpus Center (BCC) is a globally large-scale and functionally strong online Chinese corpus system. BCC has greatly facilitated the computational turn in humanities and social sciences. The system has accumulated over 100 million visits and been cited in more than 4,000 papers both domestically and internationally, with influence extending across multiple fields and dozens of countries.

2.3.2.3 Book Open-Book Data Through the Kaijuan Smart program, a total of 1,971 books related to artificial intelligence, natural language processing, speech recognition, and corpora were found. Among them, 1,303 are science and technology books, mainly in computer science and engineering technology, such as *Natural Language Processing Methods: Mining Text Data with Machine Learning and Deep Learning Using Python*, *Python Natural Language Processing and Development/AI and Big Data Series*, and *Illustrated Speech Recognition*. Searching with "speech recognition"as the keyword yields 22 titles, all science and technology books, such as *Speech Recognition: Principles and Applications* and *Fundamentals of Speech Recognition: Kaldi Practice and Exploration*. Searching with "natural language processing" as the keyword yields 87 titles, all science and technology books, such as *Practical Natural Language Processing Based on BERT Model* and *Natural Language Processing Based on Deep Learning*. Searching with "corpus" as the keyword yields 374 titles, but only 12 are science and technology books, such as *Corpus Data Processing Based on Python* and *Construction and Application of Computer Corpora*.

Analysis of corpus-related books in JD.com' s book market shows that most books on corpora introduce general corpus construction processes or corpus-based ontological research. Currently, there are no specialized books on the construction and application of a specific Chinese corpus.

The global AI industry is burgeoning, and speech is one of the most natural ways of human communication. For AI to achieve true intelligence, speech interaction capabilities are essential. Intelligent speech technology, as a key technology for AI development from perceptual to cognitive intelligence, has promoted

rapid human-computer interaction development and offers rich industrial application scenarios. Sound technology, represented by speech, is beginning to be more widely applied across various industries, with related emerging enterprises rising rapidly. Intelligent speech technology continues to penetrate major application scenarios such as education, healthcare, home, and office, gradually enabling industrial transformation and upgrading, and possesses enormous development potential and commercial value. Currently, intelligent products are actively introducing intelligent speech technology to enhance human-computer interaction experience and efficiency. As a book editor, I hope to guide more industry professionals and readers to deeply explore frontier science and technology, master the latest knowledge in AI-related fields, and on the basis of databases constructed by experts, enable more readers, industry professionals, and people interested in frontier science and technology to jointly build and share an innovative knowledge platform. Through books as a carrier, we can disseminate knowledge, share experiences, inspire more authors and readers to broaden their perspectives, and contribute to a new round of scientific and technological development.

References

[1] Feng Zhiwei. The Pros and Cons of Rationalism and Empiricism in Natural Language Processing [J]. Yangtze River Academic, 2007(2): 79-85.

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