https://doi.org/10.37420/j.cer.2024.008

Research on the Integrated Development Strategy of Future News Industry in the Smart Speaker Scene

Yuhan Gao, Cong Zhang, Ting He
School of Journalism and Communication, Beijing Institute of Graphic Communication, Beijing, China
***Corresponding author,E-mail: mail: 2214780882@qq.com**

Abstract:
The emergence of new media such as smart speakers presents both opportunities and challenges for the production and dissemination modes of news. Audio news is poised to become a pivotal component within the future ecosystem of news communication. This paper conducts an analysis of the development of smart speakers, their role in the integrated development of the news industry, and examines the current state of audio news. Based on these insights, it proposes strategies for the integrated development of the news industry within the context of smart speaker platforms.

Keywords:
Smart speakers, Journalism, Artificial intelligence, Audio News

1.Introduction

Lan Peng once pointed out that future modes of communication will transcend traditional instructional paradigms, with the emergence of interactive interfaces endowing users with greater audiovisual choices. [1] The advent of smart speakers has altered the conventional dissemination methods of broadcast news, introducing both opportunities and challenges characterized by its interactivity and personalization within the news industry. Since 2017, internet giants such as JD.com, Baidu, and Alibaba have launched their own smart speaker products, with smart voice technology now virtually ubiquitous across all smart devices. Data compiled by AVC and the China Industry Research Institute indicate that in 2017, the retail sales of smart speakers in China amounted to a mere 490 million yuan, soaring to 8.37 billion yuan by 2020. In 2021, the smart speaker market witnessed sales of 36.54 million units, generating sales revenue of 10.05 billion yuan, marking a 20.1% year-on-year growth. By 2022, market sales had reached 26.31 million units. With the continual integration and enhancement of artificial intelligence technology in smart speakers, the potential of China's smart speaker consumer market is poised for rapid expansion, with the industry anticipated to experience substantial growth. Presently, news media in Europe and America are adjusting their news content production methods and corresponding operational strategies to align with the direction of smart voice technology. For China's news industry, smart speakers may emerge as one of the most significant intelligent communication mediums following mobile phones. Thus, the expansion of smart speaker news content and the reform and integration of news production and dissemination methods pose essential questions necessitating profound contemplation within both academia and industry circles.

© 2024 by the author(s); licensee Mason Publish Group, this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (http://creativecommons.org/licenses/by/4.0/)

https://doi.org/10.37420/j.cer.2024.008

2.Development Opportunities of the News Industry in the Smart Speaker Scene

2.1. Smart Speakers Provide New Space for News Production

Echo, released by Amazon in November 2014, was the world's first smart speaker. Google launched its competitor, Google Home, two years later in November 2016. Apple introduced HomePod, which integrates the voice assistant Siri, in 2018. In the wave of technological advancement in smart speakers, the gap between China and the rest of the world is not significant. The domestic and international smart speaker markets started almost simultaneously. As Amazon and Google have not yet entered the domestic market, it has provided a favorable development space for smart speakers in China. Currently, China has the largest user base of smart speakers in the world. However, with a large population, the penetration rate is only around 10%, indicating significant room for growth. In 2020, the penetration rate of smart speakers in the United States exceeded one-quarter, far ahead of other countries. The penetration rate of smart speakers in the United Kingdom is approximately 20%, second only to the United States.

Since 2016, the Chinese smart speaker market has witnessed a rapid surge in sales, with growth rates consistently exceeding 110%. In 2018, the growth rate soared to over 400%, propelling sales from 288 million RMB in 2017 to 1.495 billion RMB, and further skyrocketing to 4.633 billion RMB in 2019. As of 2021, the Chinese smart speaker market recorded sales of 36.54 million units, amounting to 10.05 billion RMB, marking a 20.1% year-on-year increase. According to the "2021 China Smart Voice Market Analysis Report," the demand for smart voice applications in both daily consumer and enterprise-specific settings is steadily expanding. It is projected that by 2030, the consumer-grade application market will surpass 71 billion RMB, while the enterprise-grade market will reach 74 billion RMB. Following a general industry development model, the Chinese smart speaker market has established a trend of stable growth, with the pandemic-induced demand for online education also opening up new avenues for smart speaker market expansion.

Due to the rapid proliferation of smart speakers, foreign media outlets are actively strategizing around "sound" and delving into the realm of audio news. The Washington Post in the United States is exploring the transformation of its daily news into audio format; Germany's Der Spiegel has been providing daily brief news updates via smart voice devices for over a year now; The Wall Street Journal has established news briefing services on both Amazon and Google platforms, maintaining updates with content 2 to 3 times daily; The Economist, leveraging its extensive exploration in the audio domain over the years, has amassed a sizable and stable listener base.

In China, major news outlets such as People's Daily, CCTV News, and Xinhua News Agency have all launched audio news columns. CCTV News has registered official accounts and established ten audio albums, including "Anchor's Daily Broadcast," "Good Morning! News Is Here," and "News 1+1." On WeChat public accounts, CCTV News typically sends out the first post of "Good Morning! News Is Here" around 5:50 a.m. daily. Moreover, on the CCTV News app, most news articles feature a "headphones" icon for users to listen to audio broadcasts. Around 7:00 a.m. daily, the "Good Morning! News Is Here" is pushed through the "Timeline" section of the app, and users can also listen to TV broadcasts by selecting "TV+" feature. Furthermore, prominent regional media outlets such as Cover News, Shenyang Evening News, and Southern Metropolis Daily have collaborated with smart speaker Tmall Genie, launching audio news products tailored for smart

© 2024 by the author(s); licensee Mason Publish Group, this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (http://creativecommons.org/licenses/by/4.0/)

https://doi.org/10.37420/j.cer.2024.008

speaker platforms.

2.2. Smart speakers enhance the user experience in news consumption

In 2019, the Reuters Institute for the Study of Journalism at the University of Oxford released a survey report on smart speakers. The majority of users expressed satisfaction with their smart speakers, deeming them convenient, enjoyable, and socially engaging. For the average user, the adoption of smart speakers has led to a reduction in time spent on television and other screen devices.[2]

The screenless interaction of smart speakers fulfills users' desire to minimize screen exposure. In a world where individuals spend significant time glued to screens or smartphones, interruptions from pop-up messages or floating advertisements frequently disrupt sustained attention. Voice-enabled devices, however, offer a distinct advantage by eliminating screens, thereby reducing distractions. As information is conveyed primarily through sound, the concentrated delivery process minimizes opportunities for additional auditory distractions. For heavy users, smart speakers have become integral devices, accompanying them from sleep to wakefulness, gradually supplanting smartphones and radios. Voice input is perceived as an opportunity to streamline and simplify life, rendering human-machine interaction more natural and intuitive.

Due to its ease of operation and low requirement for fine motor skills, smart speakers have gained immense popularity among the elderly and people with disabilities. During an interview, an elderly individual who had never mastered computers or any other smart devices remarked that smart speakers, compared to other smart products, can be quickly and easily utilized. Within a matter of days, he learned how to interact with Tmall Genie. For individuals with disabilities, this convenience is particularly evident.

2.3. Smart speakers provide opportunities for the integration, transformation, and upgrading of the media industry

In 2015, JD teamed up with iFLYTEK to launch the Dingdong Smart Speaker, pioneering China's smart speaker industry. Currently, domestic enterprises entering this field include not only major tech giants focusing on the internet like Baidu and Alibaba, but also technology manufacturing companies represented by Huawei and Xiaomi. Traditional speaker manufacturers such as Harman Kardon and Xiaoniao Sound have also joined the market competition.

In addition to internet technology companies and speaker manufacturers, an increasing number of media outlets have also entered the smart speaker industry. Audio platform Ximalaya has introduced the "Xiao Ya Speaker," primarily embedding Ximalaya's audio content. Knowledge service provider "Get" has also launched a custom speaker developed in collaboration with Xiaomi, facilitating users to listen to "Get" content.

Among various types of smart speakers, there are customized versions of "news columns." In a simple survey conducted by the author, when news is invoked on the "Xiao Ai Speaker," it is a news product compiled by the content service team of Xiaomi, the internet company behind it. When news is invoked on the "Xiao Du Speaker," sometimes it is content from Ximalaya, and sometimes it is news content from mainstream media, without careful editing and processing. Platforms like Ximalaya have been collaborating with major media outlets to deeply customize audio programs and push them on many speakers. However, from a user experience perspective, the deep integration of the news industry with the internet has not yet reached the

© 2024 by the author(s); licensee Mason Publish Group, this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (http://creativecommons.org/licenses/by/4.0/)

https://doi.org/10.37420/j.cer.2024.008

depth expected by users. In the future, regardless of the type of speaker product, media industry involvement should be present in the creation and editing of news content, through deep integration, as well as big data and algorithm recommendations, to provide different column content for different audiences, thereby improving program quality and expanding the influence of media content.

3.The Challenges Faced by the News Industry in the Development of Smart Speaker Scenarios

3.1. Low Timeliness in News, Inaccurate Intelligent Recommendations

Currently, smart speakers often deliver outdated news, which undermines the timeliness and accuracy of the news content. Users have no opportunity to skip or freely select news content while using smart speakers, which adversely affects their user experience.

Moreover, the level of news recommendations by smart speakers is generally low, with unclear information retrieval logic and chaotic content recommendations. According to a survey conducted for the Tmall Genie smart speaker news listening experience report, most smart speaker users express a desire for personalized news recommendations. For instance, if a user requests Japanese news, news from Fujian, or news from New Zealand every day, then it would be beneficial to recommend more of such news to that user, rather than solely recommending local news. Users' interactions with smart speakers remain largely at a basic level of "commands and control," such as accessing music, inquiring about the weather, or setting timers. Voice assistants excel at handling specific queries, such as "Xiao Ai, what's the weather like today?" However, most users find it challenging to retrieve the specific hot news topics they want to hear through the smart speaker accurately. Unlike traditional radio where users listen to news sequentially, users are more selective in their listening habits. The development of audio news is influenced by the current state of media convergence, with limitations in smartness level and evident shortcomings in deep-level experiences, such as intelligent interaction between media and users.

3.2. Monotonous Content Format with Low Degree of Specialization

Currently, many media outlets still directly convert text into speech, often derived from broadcasts or print materials, resulting in content that lacks attractiveness and appears rigid and mundane. The human voice in news conveys corresponding information through its tone and timbre, which cannot yet be replicated by robotic "news readers," underscoring the irreplaceable role of genuine human voices in news reporting.

Moreover, the content of news briefings is typically singular, sourced mainly from broadcasts or other media, with durations often lengthy, making it challenging for listeners to discern key points. According to a report by the Reuters Institute for the Study of Journalism, many users express dissatisfaction with news delivery via smart speakers, with 46% of UK users regularly listening to news updates, yet only 1% considering it the most useful function of smart speakers. Users frequently complain about the lengthy programs and cumbersome operation, often resorting to alternative functionalities such as information search.

Different types of media may release identical information within a close timeframe, potentially leading to repetition or inconsistency in the content heard by users, resulting in an unsatisfactory user experience.

© 2024 by the author(s); licensee Mason Publish Group, this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (http://creativecommons.org/licenses/by/4.0/)

https://doi.org/10.37420/j.cer.2024.008

3.3. Challenges in News Content Curation: Potential Copyright Disputes

The scrutiny of news content on smart speakers faces significant challenges. Some smart speakers push news sourced from non-mainstream media outlets or automated readings of online articles, which may compromise the authenticity and credibility of the news. For instance, Malanshan Broadcasting Station, with its main anchor Ma Ben, entered the Ximalaya Headline channel in 2013, boasting a fan base of 636,000. The station's team, including anchors such as Tang Meili, Zhizhi, Kevin Teacher, and Xian gu, updates audio content such as radio dramas and bedtime stories under the banner of "Gifts of Sound" on platforms like Penguin FM and WeChat public accounts. It is imperative to acknowledge the substantial regulatory challenges and difficulties in ensuring the authenticity of both commercial and self-media content. Self-media outlets, particularly prominent in audio news presentation, often capitalize on users' psychological needs. However, such news sources frequently lack the rigorous fact-checking mechanisms employed by traditional media, thus heightening the risk of copyright disputes.

4.Smart Speakers Pave New Paths for Media Convergence

4.1. Increased Investment in News in the Smart Speaker Scene

The potent traffic-driving capability of smart speakers holds significant commercial value for news media. Mainstream media outlets should increase their investment in the field of audio news by focusing on building professional personnel teams and reforming news production processes. For instance, Swedish Radio demonstrated foresight by reimagining the potential of linear broadcasting before the advent of smart speakers in Sweden. [3] It established a dedicated team for producing intelligent audio content, swiftly capturing the initial listener market. Similarly, BBC swiftly adapted to the era by establishing numerous teams dedicated to the research and development of new products, particularly in the realms of news, broadcasting, and other forms of self-media. Concurrently, they independently developed various technologies to explore the potential evolution of audio landscapes. Domestically, Southern Metropolis Daily established a specialized team for operating audio content, equipped with a comprehensive workflow encompassing content selection, production, and final quality control. The integration of the entire process of audio content production and operation with the editorial team connects thematic selection, production, and quality control, forming the core team of the column in the form of "operation + host".

4.2. Reframing Narrative Frameworks in Journalism

Taiwanese scholar Guoren Zang has concretized the concept of news framing, proposing a three-tier framework consisting of the upper, middle, and lower levels. At the highest level, the structure involves qualitative assessments of news events and the criteria for factual selection. The middle level encompasses the selection of event plots, text structures, event attribution, and outcome analysis. The lowest level concerns news language styles and narrative techniques. Integrating frame theory with the Technology Acceptance Model (TRA), the news industry should adopt unique narrative frameworks and language styles tailored to the smart speaker environment.

Distinct news framing is required for morning, noon, and evening news broadcasts, along with personalized news delivery to meet the usage habits and lifestyle scenarios of smart speaker users. This necessitates a combination of algorithmic analysis and narrative frameworks to enhance user news experiences. For instance,

© 2024 by the author(s); licensee Mason Publish Group, this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (http://creativecommons.org/licenses/by/4.0/)

https://doi.org/10.37420/j.cer.2024.008

the "Tmall Genie Voice Headlines" program by the Shenyang Evening News offers morning and evening editions. In the morning, smart speaker users typically engage in information search, seeking the latest news, weather updates, and traffic conditions, thus emphasizing political news, major events, and public service information. In the evening, users seek relaxation and entertainment, leading to a focus on lighter news topics such as culture, entertainment, technology, and sports. News organizations need to continually experiment to identify suitable news production frameworks, reshaping the integration of news content. Additionally, due to the characteristics of smart voice technology, traditional broadcast news content needs to be streamlined, with news segments generally kept under 1 minute to meet user preferences for simple, clear, and rapid information delivery. Smart voice platforms serve not only as crucial entry points for accessing content but also as pivotal steps in retaining users' engagement and sustaining their listening experiences.

4.3. Enhancing Assisted and Interactive News Experience

The news industry can leverage the backend data of smart speakers to identify interactive points for user engagement and incorporate interactive elements into future audio news content. For instance, inserting questions within key information segments enables users to interactively engage with the smart speaker by providing answers. As comprehensive advancements in voice technology unfold, the adaptation of content tailored for smart voice playback terminals becomes increasingly crucial.

Audio platforms and smart voice products boast substantial user bases and offer high-quality auditory experiences, thereby presenting compelling opportunities for collaboration across various industries. As articulated by personnel affiliated with Tmall Genie, the expression of news through smart speakers differs significantly from any traditional news dissemination method, as it allows for seamless scene transitions based on user preferences, thereby enhancing audience engagement. For instance, "Nandu Audio Breakfast" is a collaborative venture between Southern Metropolis Daily and Tmall Genie, presenting a localized news talk show dubbed "Voice Headlines." Hosted by Zhang Pei, the program, in conjunction with Tmall Genie, delivers a nutritious "news breakfast" to listeners each morning, featuring discussions on policy impacts, local hot topics in Guangzhou and Shenzhen, and noteworthy updates. In "Nandu Audio Breakfast," listeners gain access to these exclusive news highlights. Traditional media actively employ voice interaction technology to reconstruct user relationships and reinforce their existing resources and advantages.[4]

5.Conclusions

In the current era, which is still predominantly visual-oriented, the inexorable trend toward an auditory shift has emerged. With the widespread utilization of voice-based media, there is burgeoning development and consumption of intelligent voice-based news products. This phenomenon is driving a transformational upgrade in media's news content production processes and providing a novel avenue for convergence and transformation in the news media industry. [5] Despite the existence of numerous challenges surrounding intelligent voice-based news products and services, in the era of intelligent media, news products are poised to emphasize contextualized, interactive, and segmented operational modes. Moreover, there will be a continual emphasis on algorithmic optimization, leveraging of data resources, and cultivation of demand-driven scenarios. [6] Through a process of push-and-pull interaction, real-time user experiences will be catered to. As articulated by Mukul Devichand, a senior figure at the BBC, "The evolving advancements in intelligent devices are undoubtedly heralding a transformative epoch, akin to the advent of telephony and the birth of the

© 2024 by the author(s); licensee Mason Publish Group, this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (http://creativecommons.org/licenses/by/4.0/)

https://doi.org/10.37420/j.cer.2024.008

Internet."

Funding

This project is funded by the National Radio and Television Administration-level social science research project "Research on Intellectual Property Protection Strategies for Audiovisual Programs" (Project No.: GD2324)

References

[1]Peng Lan. Smart Media: Future Media Trends - Report on the Development Trends of New Media (2016) [J]. International Journalism, 2016, 38(11): 6-24. DOI: 10.13495/j.cnki.cjjc.2016.11.001.Author 1, A.; Author 2, B. Book Title, 3rd ed.; Publisher: Publisher Location, Country, Year; pp. 154–196; ISBN.

[2]Reuters Institute:"The Future of Voice and the Implications for News",https://reutersinstitute.politics.ox.ac.uk/our-research/future-voice-and-implications-news

[3]"Amazing! Shen Wan Leads the Nation in 'Voice Headlines' Activity! Northeastern-Style Talk Show, Chatting 365 Days a Year!" Sohu Website, http://www.sohu.com/a/305471694_349248, April 2, 2019.

[4]Yin Jialin, Wu Jingyan. Exploration and Practice of Southern Metropolis Daily's 7 Premium Audio Programs[J]. Youth Journalist, 2019(21):18-19.

[5]Shi, Anbin, & Hu, Yu. (2019). Voice-Controlled Intelligent Media and News Communication: Current Situation and Prospects. Youth Journalist, 2019(01), 77-80.

[6]Wang, Haoyi. (2018). Exploration of the Current Development and Trends of Intelligent Speaker Based on Artificial Intelligence Technology. Communication World, 2018, 25(12), 225-226.

© 2024 by the author(s); licensee Mason Publish Group, this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (http://creativecommons.org/licenses/by/4.0/)