Publishing Research Quarterly (2025) 41:405–422
https://doi.org/10.1007/s12109-025-10049-1

ORIGINAL PAPER

Audiobooks and Artificial Intelligence: Tools for Synthetic Audiobook Creation and Implications for the Publishing Industry

Marta Magadán-Díaz[1] · **Álvaro Cabezas-Clavijo**[1] · **Pavel Sidorenko-Bautista**[2] · **Jesús I. Rivas-García**[1]

Accepted: 5 November 2025 / Published online: 14 November 2025

Abstract

This study explores the transformative impact of artificial intelligence (AI) on the production and consumption of audiobooks. As AI technologies advance, the publishing industry has increasingly adopted synthetic voices, which offer cost reductions and enable the creation of personalized, multilingual content. These developments, however, also raise significant ethical concerns related to authenticity, copyright, and voice cloning. To better understand the scope and implications of AI in the audiobook sector, this research conducted a review of 28 AI-based voice synthesis platforms. The analysis distinguished between platforms specifically tailored for audiobook production and those offering broader audio functionalities. Additionally, it examined the potential of these tools to enhance accessibility and influence user preferences. Findings reveal that only four out of the 28 platforms are explicitly designed for audiobook creation, while the majority serve more general audio applications. These platforms provide a wide array of synthetic voices and language options, thereby facilitating access to the audiobook market for smaller and independent publishers. Moreover, all platforms require paid subscriptions to unlock full functionality, with only limited features available in free versions. Most offer services such as voice cloning, personalization, and multilingual support. The study concludes that while AI significantly lowers entry barriers and operational costs in audiobook production, human narration remains a benchmark for quality and emotional engagement. To responsibly integrate AI technologies, the publishing sector must strike a balance between innovation and ethical accountability, adapting business models to ensure both sustainability and user trust.

Keywords Audiobooks · Artificial Intelligence · Platformization · Speech Synthesis Technology · Artificial Voices · Synthetic Voices

Abbreviations

AI Artificial Intelligence

Extended author information available on the last page of the article

TTS Text-to-Speech
LLMs Large Language Models

Introduction

Global digital transformation continues to profoundly reshape the publishing industry, revolutionizing how we produce and consume content (Nieborg et al. 2022; Spjeldnæs 2022). The rise of artificial intelligence (AI) accelerates this shift, fueling new formats (Shatzkin 2008; Magadán and Rivas 2018; 2019) and redefining traditional models of content interaction (Cabezas et al. 2024). A global digital communication model, defined by technological convergence, platform dominance, and algorithmic personalization, now imposes shared logics of visibility and distribution that deeply affect creative industries like publishing (Colbjørnsen 2021; Nieborg et al. 2022; Snelling 2021). This evolution presents both new opportunities for outreach and significant risks, including market concentration and the erosion of cultural diversity, making a critical examination of its intersection with regional publishing ecosystems crucial.

Within this interconnected landscape, new user demands drive the evolution of narratives and professional roles (Herranz et al. 2019). The integration of AI into audiobook production exemplifies this transformation (Saikaly 2023). Publishers who have traditionally relied on human narrators (Rivas and Magadán 2022; Bojda 2024) can now use AI to generate high-quality narrations at a much lower cost, forcing a reconsideration of business models (Ryzhko et al. 2024). While this automation could reduce prices and make audiobooks accessible to a broader audience (Kaur and Singh 2023), it also raises questions about the preservation of narrative creativity and emotional depth (Tattersall 2022a).

The audiobook market has expanded rapidly, driven by digitally native generations who integrate listening into their everyday routines. Millennials and Generation Z are the most active consumers, engaging with audiobooks while commuting, studying, or relaxing – often alongside music and podcasts (Viens 2019; Tattersall 2022b; Berglund 2024). Recent data from the UK, the USA, and Spain confirm this trend, with younger audiences leading consumption growth (Watson 2023; Curcic 2023; Ministerio de Cultura y Deporte 2023). As a result, audio has become the publishing industry's fastest-growing segment, with projected annual growth exceeding 10% – far outpacing ebooks and print books (Statista 2024).

The user's relationship with the format has also evolved. Initially a tool for accessibility, audiobooks now cater to a broad audience as a complementary leisure activity (Snelling 2021; Magadán and Rivas 2020; Rivas and Magadán 2022). Their 21st-century resurgence, fueled by streaming and subscription platforms (Colbjørnsen 2021), has pushed publishers to consistently adopt new technologies for distribution (Snelling 2021; Have and Pedersen 2020).

AI narration leverages text-to-speech (TTS) technology to offer personalization, reduce production costs, and generate content recommendations (Chen 2020; Dale 2022; Tattersall 2022a). However, the industry's approach suggests perceived audience resistance. Major platforms often use alternative labels like "self-narrated"

or "Narrated by Apple Books" instead of explicitly acknowledging AI-generated voices, implying that synthetic narration may not appeal to listeners (Kühne et al. 2020; Rodero and Lucas 2023; Erard 2023; Hui and Fielding 2023).

Given the rising popularity of audiobooks and the pivotal role AI tools play in expanding catalogues and global distribution, an in-depth analysis of these tools and their sector-wide implications is clearly warranted.

Objectives

This research primarily aims to explore the main AI tools with potential for audiobook production, analyzing their characteristics, functionalities, advantages, and limitations.

The specific objectives are as follows:

- to identify AI tools currently available that offer functionalities applicable to audiobook production, regardless of their current adoption in professional publishing workflows
- to determine whether these tools are designed for audiobook production or have a broader scope, enabling content generation across various formats (e.g., video, podcasts)
- to assess the economic costs of these tools and the distribution channels for content generated by these platforms.

Audiobooks and AI

Artificial intelligence is decisively transforming the audiobook industry, altering production, distribution, and consumption processes (Fernández 2024; Genelza 2024). Voice synthesis technology has evolved from rudimentary, robotic implementations to advanced models like Google WaveNet and Tacotron 2, which can replicate intonation, rhythm, and emotional nuance with near-human naturalness (Mahum et al. 2023; Kaur and Singh 2023). These improvements, powered by deep neural networks, enable the generation of fluid and expressive narrations from written text.

Initially designed for virtual assistants or navigation systems, this technology has expanded into cultural formats such as audiobooks, podcasts, and digital articles (Rodero and Lucas 2023; Kosch et al. 2024). Using text-to-speech (TTS), publishers can produce audiobooks more quickly and affordably, without relying on human narrators, and adapt the voice to different genres or narrative styles (Casian and Gabriela 2023). Users, in turn, benefit from personalization options – such as adjustments to tone, pace, accent, or emotion – that promote more immersive listening experiences tailored to their preferences (Tattersall 2022a).

Advances in TTS also benefit major platforms such as Google, Apple, and Amazon, enabling them to convert extensive catalogs into audio and expand multilingual

offerings (Franganillo 2023). Furthermore, automatic translation and localization allow for the adjustment of accents and vocal inflections, thereby increasing international accessibility (Ji et al. 2024).

Although the perceptual gap between human and synthetic voices has narrowed, emotional resonance remains a benchmark of quality. AI-driven tools enable the selection of distinct narrative styles and the generation of more accurate transcriptions, which supports the inclusion of individuals with hearing loss. However, this evolution raises ethical and legal dilemmas related to voice attribution, consumer transparency, and intellectual property rights (Chuks-Okeke et al. 2024; Avilés and Vernier 2023).

In summary, AI offers significant opportunities to reduce costs and democratize audiobook production, but it also requires the publishing sector to reflect on narrative quality, the preservation of professional work, and the ethical management of these emerging tools.

Producing an Audiobook with AI

The production process of an audiobook begins with a printed or manuscript work, whether previously published or not, which then enters the pre-production phase. The publisher may carry out this phase or outsource it to another party. During this stage, the decision is made whether the recording will strictly follow the text or, alternatively, require a script adaptation for the narration. Producing books with AI involves several steps (see Fig. 1), from converting the written text into synthesized speech to enhancing the result.

This model is not intended to encompass the full range of audiobook production possibilities enabled by AI. In fact, there is a growing number of works conceived directly as audio productions (born audio or audio originals), which are not adapted from existing print materials (Have and Pedersen 2020; Linkis 2021; Tattersall 2022b; K. Berglund 2024). Additionally, the term "recording" might be misleading in this context, as TTS processes do not involve recording in the traditional sense, but rather algorithmic text-to-voice conversion (Kaur and Singh 2023). While many current productions use a single synthetic voice to preserve narrative consistency, the use of multiple voices in a single work is not only technically feasible but increasingly common (Kimberg and Chin 2024). Therefore, this model should be understood as a simplified representation of a complex and evolving process, shaped by technological innovation, audience practices, and editorial experimentation (Ryzhko et al. 2024).

Before narration begins, the audiobook text undergoes preprocessing (Berglund and Dahllöf 2021). This stage may involve editing the content to ensure a smooth reading, including correcting ambiguities or simplifying technical terms that the AI system might mispronounce (Kimberg and Chin 2024). The text must be well structured to ensure the narration sounds coherent and natural. The choice and configuration of the synthetic voice are essential. AI platforms offer various options for personalizing the voice and adjusting features such as tone, pace, and emotion (Casian and Gabriela 2023). For example, a dramatic fiction audiobook may require a more

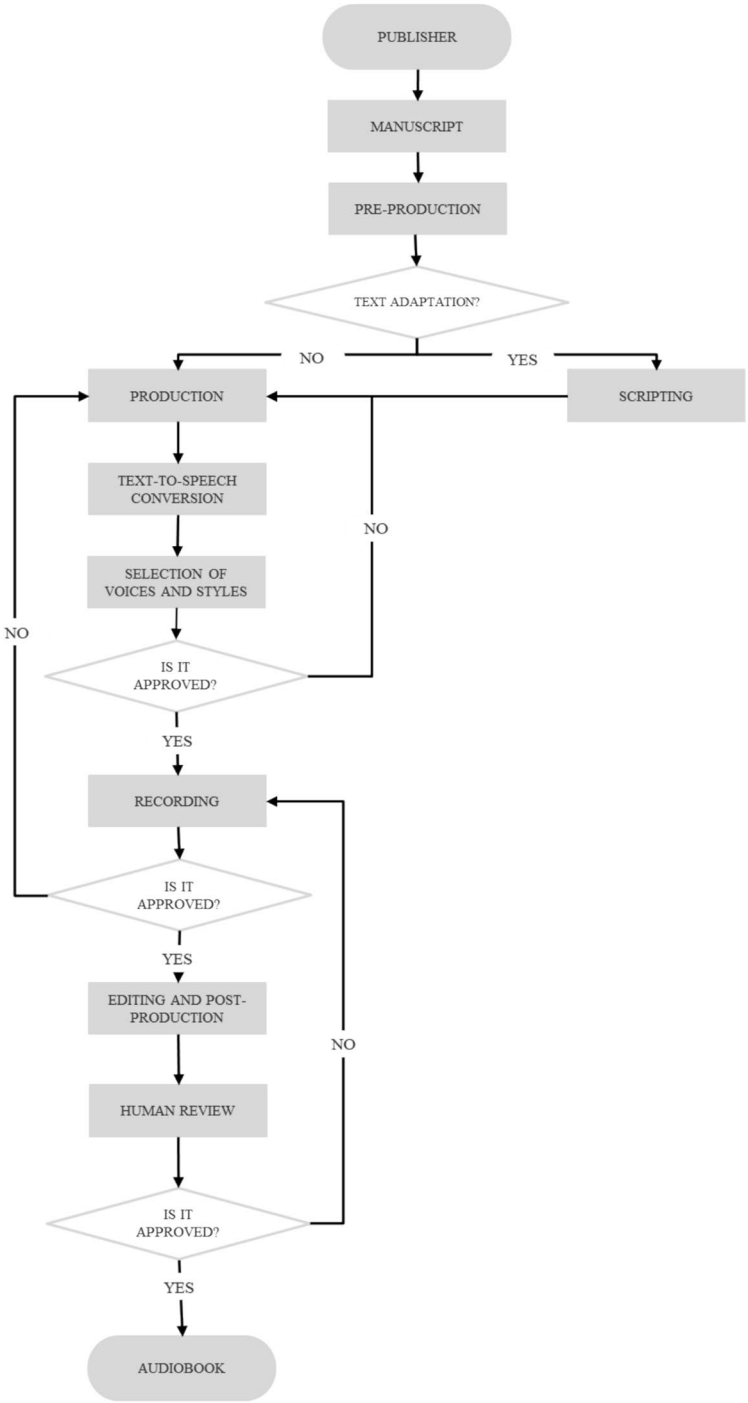


Fig. 1 Flowchart of audiobook production with AI

emotional and measured voice, while a technical text might call for a more neutral and steady tone. Proper voice configuration ensures that the narration matches the genre and style of the book. Recent advances in TTS have allowed for finer control over intonation and rhythm, generating more natural and expressive voices thanks to neural networks and deep learning models. Some AI systems can detect the context of a sentence and automatically adjust intonation, such as emphasizing questions or conveying surprise, which significantly enhances the listener's experience (Kimberg and Chin 2024).

While AI can facilitate audiobook production by automating certain aspects of the process (Kimberg and Chin 2024), human intervention is still required to ensure the quality and consistency of the result. For English, TTS offers an acceptable level of quality with good sound output, leading technological platforms to market TTS-produced audiobooks. In the case of TTS in other languages, such as Spanish, quality is also reaching acceptable levels, although some post-production work is still necessary to achieve a satisfactory standard (Bookwire 2023).

After generating the narration, the production team adjusts and edits to improve the audio quality, using specific techniques such as removing noise, adjusting the volume, and adding sound effects to enrich the listening experience. Editing tools help correct AI pronunciation errors and adjust pauses to make the narration sound more natural. This process is vital to ensure a high-quality final product. Before distribution, quality checks are essential, involving listening to samples and adjusting based on feedback to ensure the audiobook meets the established quality standards. This stage ensures any defects are corrected before the product reaches the public. Once the production process is complete, the audiobook can be distributed and published on audiobook platforms, online stores, or other distribution services. Some platforms allow direct uploads to services such as Audible or Apple Books, facilitating access to a broader audience (Berglund 2021; Have and Pedersen 2020; Rivas and Magadán 2022; Tattersall, 2022a). Furthermore, metadata to optimize search (Shetty et al. 2015) and categorization of audiobooks help improve content visibility and discovery by listeners.

Publishing Implications of AI Narration

While this work does explore certain technical aspects of AI audiobook production – such as voice configuration and automated text-to-speech conversion – this is not out of a purely technical or engineering interest. Rather, these elements are examined from an expanded editorial perspective. In today's digital landscape, technical decisions have become inseparable from editorial ones, as they influence accessibility, narrative experience, audience targeting, and algorithmic content distribution. Understanding how these tools function is therefore essential not only to describe the production process, but also to reconsider the role of the editor within an ecosystem increasingly shaped by automated platforms.

The idea of selecting a single "appropriate" voice for each audiobook reflects a more linear and limited conception of voice configuration. AI-driven platforms are

rapidly moving toward more flexible models, where users can choose among multiple voices, accents, tones, and even narrative styles – allowing the listening experience to be tailored to personal preferences, contextual situations, or accessibility needs (Casian and Gabriela 2023; Kaur and Singh 2023; Ji et al. 2024; Kimberg and Chin 2024). Rather than a fixed one-to-one assignment of voice to text, we are entering a scenario in which a single work may come to life in various ways, depending on who listens and how they choose to engage with it (Tattersall 2022a; Ji et al. 2024). This shift entails not only a technical transformation but also a redefinition of editorial roles and listener expectations in relation to narrative experience (Magadán and Rivas 2020; Rivas and Magadán 2022; Franganillo 2023).

Methodology

This study follows an exploratory and descriptive methodological approach rooted in well-established qualitative research traditions (Yin 2017). The research design follows a scoping review approach, which provides a robust method for analyzing emerging or underexamined phenomena and identifying patterns within defined samples. This method is particularly suitable for analyzing under-theorized technological developments, such as the implementation of AI tools in audiobook production. Following the methodological guidelines set out by Levac et al. (2010), this review enables the systematic identification, selection, and comparative mapping of digital tools. Researchers have applied this approach in similar technology-driven contexts, including studies on educational technologies, digital platforms, and AI applications in media (Franganillo 2023).

The research design combined documentary review techniques and comparative platform analysis, starting with a purposive selection based on targeted keyword searches through specialized search engines (Google, Google Scholar) and sector-relevant websites (e.g., PublishDrive). The research applied clearly defined inclusion and exclusion criteria, focusing on platforms that remained active and accessible in 2024 and offered functionalities related to audiobook production or voice synthesis. These criteria align with qualitative sampling principles commonly used in platform research (Yin 2017). The team verified each selected platform through direct access and, when available, conducted hands-on testing of demo versions to assess technical features and overall usability.

This approach builds on methodological frameworks previously employed in studies on digital media and technology (Hine 2020), where online content and platform functionalities serve as valid data sources for analyzing processes of digital transformation. It enables researchers to map a rapidly evolving technological ecosystem and provides a valuable dataset for future, more in-depth investigations.

This work searched the web to identify tools dedicated to creating synthetic voices for audiobooks and conducted various searches using the following keywords: "Audiobooks + AI," "Audiobooks making with AI," and "AI audiobook maker," along with recommendations from PublishDrive (Mădălina 2024). The researchers conducted the study in English and at two different times: an initial

review in March 2024, followed by an update in September 2024, which resulted in a sample of 28 generative AI tools.

Several prior studies have addressed emerging phenomena in the publishing and media fields using qualitative or documentary methodological approaches, which supports the suitability of the framework adopted in this study. For instance, Colbjørnsen (2021) analyzes editorial platformization from a theoretical and descriptive perspective, while Rodero and Lucas (2023) employ an experimental design to examine the perception of synthetic voices, and Avilés and Vernier (2023) conduct a systematic characterization of AI applications in digital radio. Although these studies do not use exactly the same methodological design as this research, they share an interest in mapping emerging technologies through analytical strategies tailored to dynamic digital contexts.

The collected data were organized into an analysis table, recording the following variables:

- tool name
- release year
- responsible company
- country of the company
- tool type: 1. Tool specifically for audiobook creation; 2. Generic digital audio tool.
- services: 1. Audiobook generation; 2. Audio editing; 3. Text-to-speech; 4. Customisable voices; 5. Multilingual; 6. Voice cloning; 7. Voiceover; 8. Voice synthesis; 9. Voice effects
- number of voices
- number of languages
- cost: The lowest and highest subscription fees available on the platform
- content distribution: 1. Email; 2. Social media; 3. Direct link; 4. Directly to audiobook platforms
- supported audiobook platforms: 1. Apple Books; 2. Audible; 3. Storytel; 4. Other.

Finally, the researchers accessed the respective websites to verify the different variables for each tool and, where possible, tested a demo or free version of the application.

The choice of this methodology responds to the exploratory nature of the topic, providing a structured yet flexible framework to map a rapidly evolving sector. Although this approach presents certain limitations – such as its reliance on publicly available information – it enables a scalable analysis of trends, functionalities, and market positioning.

Results

Of the 28 generative AI tools and resources in the selected sample, all require paid versions to access their full creative capabilities, although some provide limited free options. These free versions have allowed for a sufficient understanding of their offerings and functionalities for this study.

The United States was the country of origin for 17 of the 28 analyzed platforms (60.7%), followed by India and the United Kingdom, each with two tools. Seven other countries have developed an AI tool for audiobook creation. A key factor contributing to this could be the size of the US market, its early adoption of digital technologies, or the presence of a well-established and advanced technological infrastructure.

Although variables such as the year of release may seem merely descriptive, they are essential for contextualizing technological evolution in publishing. The concentration of new platforms in 2022–23 reflects the rapid diffusion of generative AI and marks a structural shift in audiobook production and business models.

Of the 28 tools analyzed, only four (14.3%) are specifically designed for audiobook creation (Apple Books, Audie.ai, AuthorVoices, and Pozotron), while the remainder are generic digital audio tools, applicable to a wide range of content formats (Table 1).

The period between 2017 and 2023 reveals a substantial surge in tools and resources leveraging AI for digital audio editing and creation, including audiobook production. This growth may stem from advances in technology, as well as the growing popularity of audiobooks and audio content consumption overall. Notably, the sharp rise in 2023 registrations indicates that the sector continues to attract new platforms eager to meet the rising demand for audio content (see Fig. 2).

Audiobook and digital audio platforms provide a well-rounded set of services, each backed by a comparable number of AI tools and resources. This approach suggests that platforms focus on offering comprehensive features to meet user needs and expectations rather than relying on unique elements to stand out in the market (see Fig. 3).

Services such as audio editing, audiobook generation, and voice synthesis are central, with 24 to 26 of the platforms analyzed offering these features. This suggests that these functions are considered essential and play a key role in audio content creation and editing.

Building on this, advanced features such as voice cloning and personalization also appear extensively. Voice cloning, for example, enables platforms to provide a more immersive and tailored experience (though it raises ethical concerns), as they can replicate specific voices, which proves valuable for personalized or educational content. Similarly, multilingual support through voice synthesis in various languages is crucial for enabling tools to reach a diverse global audience. The availability of multiple languages is of paramount relevance due to the global reach of audiobooks. As shown in Fig. 4, there is considerable variation in both the number of voices and languages offered by the digital audio editing tools.

Table 1 Origin information of AI tools for audiobook creation, editing, and digital audio projects

Platform	Launch year	Company	Country	Tool type
Apple Books	2010	Apple	USA	Specific
Artlist.io	2016	Artlist Ltd	Israel	General
Audie.ai	2023	NO INFO	USA	Specific
Audiosonic	2023	Writesonic	USA	General
Audyo	2023	NO INFO	USA	General
AuthorVoices	2023	Archieboy Holdings, LLC	USA	Specific
Descript	2017	Descript, Inc	USA	General
Dubverse.ai	2021	Deterministic Algorithms LLP	India	General
ElevenLabs	2022	ElevenLabs, Inc	Poland	General
Fliki	2021	Nine Thirty Five	India	General
Lovo	2019	Lovo, Inc	USA	General
Murff TTS	2020	Murf AI, Inc	USA	General
Podcastle	2020	Podcastle, Inc	Armenia	General
Pozotron	2017	Pozotron, Inc	USA	Specific
Resemble.ai	2019	Resemble, Inc	Canada	General
Respeecher	2018	Respeecher, Inc	Ukraine	General
Revoicer	2021	Revoicer	USA	General
Scribe Audio	2020	Scribe Audio, Inc	USA	General
Speechify	2017	Speechify, Inc	USA	General
Speechki	2019	Speechki, Inc	USA	General
StoryBook	2022	StoryBook AI, Inc	USA	General
Typecast	2019	Neosapience	South Korea	General
Veed.io	2018	VEED Ltd	United Kingdom	General
Vidnoz.ai	2018	Wise Reward Limited	USA	General
WeelSaid Labs	2018	WellSAid Labs	USA	General
Voxificar	NO INFO	Voxify	Mexico	General
Writesonic	2020	Writesonic, Inc	USA	General
Wondercraft	2023	Wondercraft, Inc	United Kingdom	General

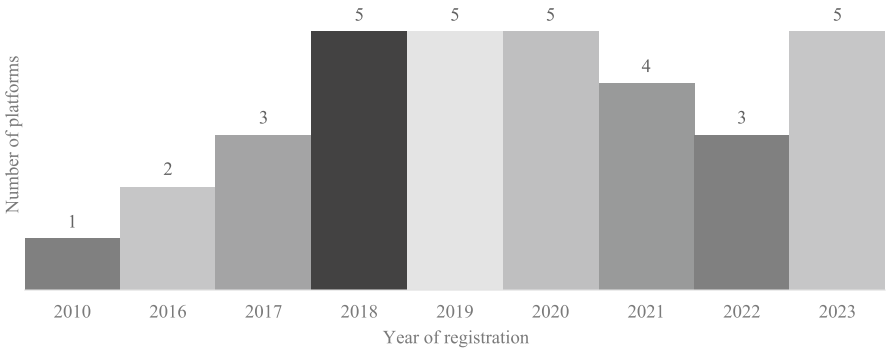


Fig. 2 Distribution of AI tools and resources for digital audio editing and audiobooks by registration year

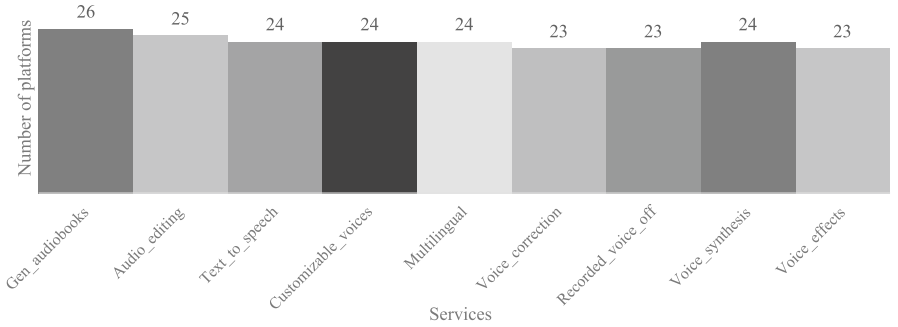


Fig. 3 Comparison of services offered by AI tools and resources for digital audio editing and audiobooks

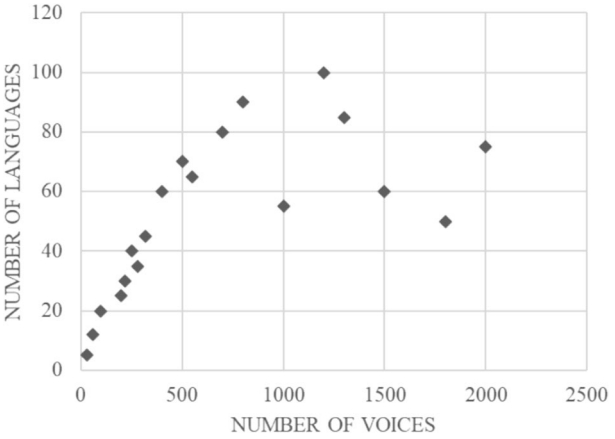
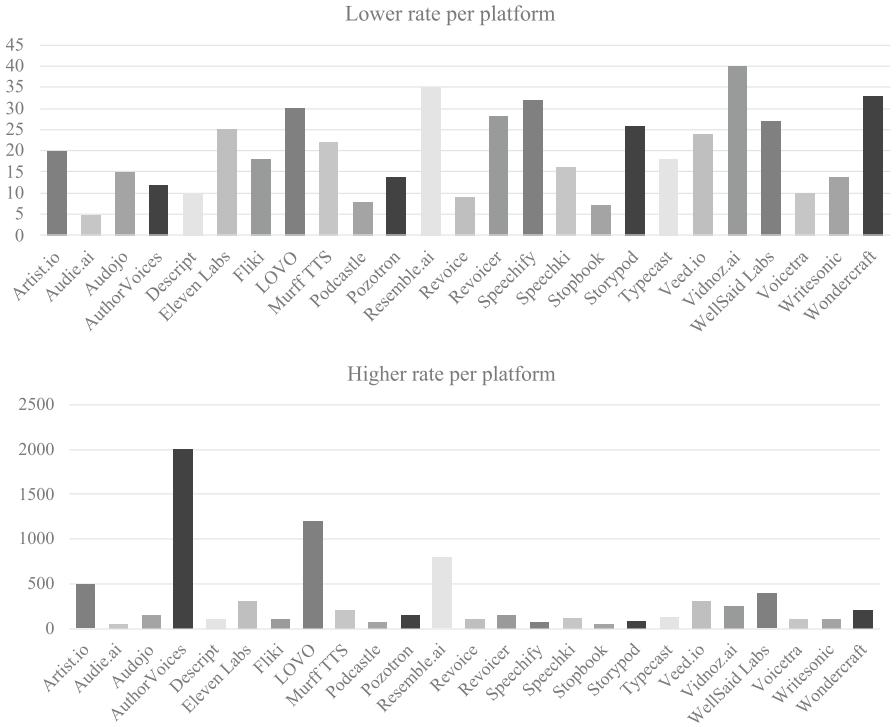


Fig. 4 Relationship between the number of voices and languages offered by AI tools and resources for digital audio and audiobook editing

Some platforms prioritize offering a wide variety of voices in a limited number of languages, while others focus on multilingual support with fewer voice options. This difference highlights varying approaches to accessibility and personalization, which could influence user selection criteria.

Most platforms focus on the lower range (fewer than 250 voices and fewer than 50 languages), suggesting that many tools offer a limited selection of voices and languages. However, some platforms stand out by providing over 1,000 voices, such as Fliki (2,000), Vidnoz (1,240), and Speechki (1,100), highlighting an emphasis on personalizing the user experience and providing differentiated content with a broader range of results.

Another notable point is that several tools support a high number of languages (between 60 and 140) but offer a moderate number of voices. Voxificar (140), Lovo (100), and Audie.ai (90) are the platforms providing the largest selection of languages (see Figs. 5 and 6).



Figs. 5 and 6 Costs of AI tools and resources for digital audio editing and audiobooks

Figures 4 and 5 reveal significant heterogeneity in the subscription or hiring costs for the digital audio editing and audiobook tools included in the study sample. Regarding the most affordable platform rates (Fig. 6), Resemble.ai ($1) and Audyo ($3) offer the lowest fees, while WeelSaid Labs ($44.08) and Revoicer ($37) have the highest rates. However, it is significant to note that pricing systems vary considerably from platform to platform, with companies adopting different strategies for customer acquisition. For example, Resemble.AI's $1 fee applies only to the first month (after which it increases to $29) and provides limited services.

The platforms are divided into two categories: those offering affordable prices across lower and higher-tier options and those with a broader pricing range, reflecting higher features and catering to a more professional audience. Platforms with significantly higher fees, such as AuthorVoices, provide professional services that cover audiobook creation and include distribution and promotion of the content.

Figure 7 examines the publication channels for various tools, showing that most platforms (89.3%) allow content distribution via email, social media, or direct links. However, only three platforms (10.7%) support direct uploads to audiobook platforms: Apple Books, Audie.ai, and AuthorVoices.

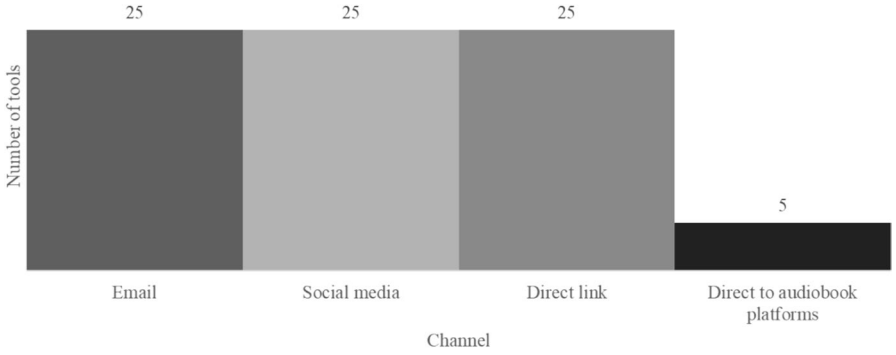


Fig. 7 Publication and distribution channels for AI tools and resources for digital audio editing and audiobooks

Discussion and Conclusions

This study has reviewed the primary platforms for generating audiobooks using generative AI. The analysis found that most of the 28 platforms examined offer similar, generic services to create audio content, with only a few tools dedicated to audiobook creation. Most of these tools have been launched recently (since 2017) and offer various pricing models, enabling users to choose the most suitable service based on features and costs.

The proliferation of AI tools parallels the expansion of the audiobook sector, driven by demand for flexible digital consumption and mobile access. Audiobooks have evolved from an accessibility-oriented format into a mainstream cultural medium embraced by younger audiences (Linkis 2021; Rivas and Magadán 2022; Tattersall 2022a; Berglund 2024). By offering increasingly diverse synthetic voices and multilingual options, these platforms allow small publishers and independent authors to reach global audiences at lower costs while preserving emotional engagement through tailored narration (Ji et al. 2024).

Nevertheless, audiobook listeners prefer human voices over synthetic ones, as the former facilitates greater overall enjoyment of the content. Human voices promote the creation of more vivid and numerous mental images, generating a more emotional response and increasing attention levels (Rodero and Lucas, 2023). Apart from the content itself, the most crucial factor in capturing the audiobook listener's attention is the quality of the narrator's voice (warmth, closeness, naturalness, intonations, emotions), which is further enhanced by adding sound effects to the story. Listeners who experience telepresence (the sensation of being immersed in the story) and emotional connection tend to have a more favorable attitude towards audiobooks, with these psychological factors fostering greater emotional engagement and enjoyment (Ji et al., 2024). Varied narrative styles, such as multiple voices or dramatized readings, along with music and sound effects, create this emotional connection by enhancing the emotional implications of the content (Cahill and Moore 2017). Audio stories that include non-verbal sound elements (e.g., sound effects, reverb filters, and volume variations) generate more vivid mental imagery than

verbal stories, enriching the listening experience (Rodero and Romero 2022). In this regard, most platforms studied offer features like voice and sound effects, which can support immersion in the content, fostering emotional connections that enhance the enjoyment of the reading.

On the other hand, most available tools lack the functionality for direct publication on audiobook platforms, such as Audible, Storytel, and Apple Books, a significant limitation in the current market landscape. Global platforms like Audible and Apple Books dominate the audiobook industry, establishing a dominant presence through subscription models and extensive libraries (Berglund, 2021). These platform strategies encompass more than just books; they also include other forms of audio content, such as exclusive narrative podcasts, talk shows, and original programming. Within this context, users act as consumers and active participants in the ecosystem, generating data and providing feedback that directly contributes to service improvements. Consequently, audiobook platformization enables companies like Audible and Apple Books to gather data on user listening habits and apply these insights to tailor content to their preferences, enhancing user retention and engagement. The fact that AI-based audiobook creation tools do not offer direct access to these main distribution networks highlights a critical gap – not in content generation but in distribution and audience reach (Fernández, 2024; Genelza, 2024).

The emergence of AI narration tools has significantly lowered production barriers, enabling content creation without studios or professional narrators. While this democratization fosters innovation and inclusion, it also raises concerns about quality control and copyright. As observed in the self-publishing ebook market, low entry costs may lead to an excess of AI-generated works that do not always meet professional standards (Cabezas et al. 2024).

The analysis reveals, in general terms, that most tools in the study sample offer extensive and efficient creative and editing resources aimed at producing high-quality content, making them viable options for creating professional-grade audiobooks. Besides, many of these tools provide customized packages without human mediation, significantly lowering costs. This shift could pose challenges for publishing houses that fail to adapt efficiently to the new model of audio content production and distribution (Fernández, 2024).

The AI incorporation into audiobook production is profoundly transforming the publishing industry (Saikaly, 2023). Publishers who have traditionally relied on human narrators face a significant challenge due to AI's capacity to generate high-quality narrations at considerably lower costs (Bojda, 2024; Ji et al., 2024; Rivas and Magadán, 2022; Tattersall, 2022a). This shift compels companies in the sector to reconsider their business models (Magadán and Rivas, 2022) and evaluate the integration of AI technologies into their operations (Ryzhko et al., 2024). Automation can substantially lower production costs, potentially leading to lower prices and making audiobooks more accessible to a broader audience (Kaur and Singh 2023). However, this transformation also presents challenges, especially in maintaining quality and creativity in narration areas where human narrators continue to excel.

Moreover, AI use in audiobook production has significant implications for distribution and marketing (Fernández, 2024). With the ability to generate content quickly and at scale, publishers can expand their catalogues and reach new markets.

However, the increasing content volume requires distribution platforms to adapt, as visibility for specific titles may suffer in an increasingly saturated environment. In this context, strategic use of metadata and recommendation algorithms becomes essential to help listeners discover new titles (Shetty et al. 2015). Publishers that fail to adopt effective digital marketing and content optimization strategies risk losing relevance in a rapidly evolving market.

Regarding the limitations of this study, the market for AI tools used in audio content generation is highly dynamic, with new platforms emerging and others disappearing quickly. This evanescence and dynamism within the sector make it challenging to analyze the phenomenon at a specific moment, as this study has done. Therefore, continuous sector monitoring is necessary to identify new services or platforms that may join the catalogue presented here. Additionally, one must consider the growing capabilities of large language models (LLMs) that underpin the main AI tools, which could rapidly generate new services or functions potentially disruptive to the audiobook value chain, which would open the door to new consumption models for this format, offering more personalization options and even lower costs than those currently available.

Although many studies highlight listeners' preference for human voices due to their emotional depth (Kühne et al. 2020; Rodero and Lucas 2023), recent research suggests that this distinction is narrowing. Advanced synthetic voices can increasingly evoke similar levels of engagement and authenticity (Casian and Gabriela 2023; Kosch et al. 2024), indicating that perception depends more on the listening experience than on the voice's origin. The intersection between human and AI narration thus opens new possibilities for storytelling and audience connection. Ultimately, AI represents a powerful tool to democratize audiobook production and expand global access, but its responsible integration will depend on balancing technological efficiency with human creativity and ethical transparency.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. Not applicable for that section.

Data Availability Not applicable for that section.

Declarations

Competing interest All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

References

Avilés, Claudio, and Emiliano Vernier. 2023. "Radio e Inteligencia Artificial: Sistematización y Caracterización de Aplicaciones y Prestaciones." *Revista Latinoamericana de Ciencias de La Comunicación* 22 (44): 41–53. https://doi.org/10.55738/alaic.v22i44.1042.

Berglund, Karl. 2024. *Reading Audio Readers: Book Consumption in the Streaming Age*. Bloomsbury Publishing.

Berglund, Karl. 2021. "Introducing the Beststreamer: Mapping Nuances in Digital Book Consumption at Scale." *Publishing Research Quarterly* 37 (2): 135–51. https://doi.org/10.1007/s12109-021-09801-0.

Berglund, Karl, and Mats Dahllöf. 2021. "Audiobook Stylistics: Comparing Print and Audio in the Bestselling Segment." *Journal of Cultural Analytics* 4 (1): 1–30. https://doi.org/10.22148/001c.29802.

Bojda, Tomáš. 2024. "Radio and Audiobook Acting. Comparative Analysis of Two Audio Productions of The Master and Margarita." *Theatralia* 27 (1): 167–93. https://doi.org/10.5817/TY2024-1-8.

Bookwire. 2023. "Evolución Del Mercado Digital (Ebooks y Audiolibros) En España y América Latina." Informe Bookwire 2023. https://www.bookwire.es/fileadmin/media/_footer/unternehmen/knowledge_hub/BOOKWIRE_INFORME_2023.pdf.

Cabezas, Álvaro, Marta Magadán, Pavel Sidorenko, and Jesús I. Rivas. 2024. "This Book Is Written by ChatGPT: A Quantitative Analysis of ChatGPT Authorships through Amazon.Com." *Publishing Research Quarterly* 40 (2): 147–63. https://doi.org/10.1007/s12109-024-09998-w.

Cahill, Maria, and Jennifer Moore. 2017. "A Sound History: Audiobooks Are Music to Children's Ears." *Children and Libraries* 15 (1): 22–29. https://doi.org/10.5860/cal.15n1.22.

Casian, Marincaş, and Czibula Gabriela. 2023. "A Generative Modelling-Based Approach for Text-to-Speech Synthesis in Romanian Language." In *2023 IEEE 19th International Conference on Intelligent Computer Communication and Processing* (ICCP), 323–30. Cluj-Napoca, Romania: IEEE. https://doi.org/10.1109/ICCP60212.2023.10398699.

Chen, Yuan. 2020. "The Reform of Publishing Format and the Reconstruction of Editing Power in the Era of Intelligence." *Open Access Library Journal* 7 (12): 1–10. https://doi.org/10.4236/oalib.1107052.

Chuks-Okeke, Ekene, Natalie Linero, and Brenda Leong. 2024. "Generative AI and Intellectual Property: Copyright Implications for AI Inputs, Outputs." IAPP, August 7. https://iapp.org/news/a/generative-ai-and-intellectual-property-copyright-implications-for-ai-inputs-outputs.

Colbjørnsen, Terje. 2021. "The Streaming Network: Conceptualizing Distribution Economy, Technology, and Power in Streaming Media Services." *Convergence* 27 (5): 1264–87. https://doi.org/10.1177/1354856520966911.

Curcic, Dimitrije. 2023. "Book Sales Statistics." Wordsrated. https://wordsrated.com/amazon-publishing-statistics/.

Dale, Robert. 2022. "The Voice Synthesis Business: 2022 Update." *Natural Language Engineering* 28 (3): 401–8. https://doi.org/10.1017/S1351324922000146.

Erard, Gabriel. 2023. "Apple Books Ahora Ofrece Audiolibros Narrados Por Inteligencia Artificial." *Hipertextual*, 2023. https://hipertextual.com/2023/01/apple-books-audiolibros-narrados-inteligencia-artificial.

Fernández, Peter. 2024. "How AI Might Transform the Book." *Library Hi Tech* 41 (8): 15–17. https://doi.org/10.1108/LHTN-08-2024-0130.

Franganillo, Jorge. 2023. "Los Grandes Modelos de Lenguaje: Una Oportunidad Para La Profesión Bibliotecaria." Anuario ThinkEPI 17: 1–5. https://doi.org/10.3145/thinkepi.2023.e17a28.

Genelza, Genesis Gregorious. 2024. "A Systematic Literature Review on AI Voice Cloning Generator: A Game-Changer or a Threat?" *Journal of Emerging Technologies (JET)* 4 (2): 51–61. https://doi.org/10.57040/ag587791.

Have, Iben, and Birgitte-Stougaard Pedersen. 2020. "The Audiobook Circuit in Digital Publishing: Voicing the Silent Revolution." *New Media & Society* 22 (3): 409–28. https://doi.org/10.1177/1461444819863407.

Herranz, José María, Raquel Caerols, and Pavel Sidorenko. 2019. "La Realidad Virtual y El Vídeo 360o En La Comunicación Empresarial e Institucional." *Revista de Comunicación* 18 (2): 177–79. https://doi.org/10.26441/RC18.2-2019-A9.

Hine, Christine. 2020. *Ethnography for the Internet. Embedded, Embodied and Everyday*. London (UK): Routledge. https://doi.org/10.4324/9781003085348.

Hui, Alan, and Kate Fielding. 2023. "Friend, Foe or Frenemy: Foreseeable Impacts of AI on Arts, Culture and Creativity." A New Approach (ANA), October. https://newapproach.org.au/wp-content/uploads/2023/10/ANA_Analysis-Paper_Friend-foe-or-frenemy.pdf.

Ji, Dan, Boquan Liu, Jinghong Xu, and Jiankun Gong. 2024. "Why Do We Listen to Audiobooks? The Role of Narrator Performance, BGM, Telepresence, and Emotional Connectedness." *Sage Open* 14 (2): 1–13. https://doi.org/10.1177/21582440241257357.

Kaur, Navdeep, and Parminder Singh. 2023. "Conventional and Contemporary Approaches Used in Text to Speech Synthesis: A Review." *Artificial Intelligence Review* 56: 5837–80. https://doi.org/10.1007/s10462-022-10315-0.

Kimberg, Dan, and Spencer Chin. 2024. "Text-to-Audio Conversion Using LLMs to Preserve Character Identification." *Technical Disclosure Commons*. https://www.tdcommons.org/dpubs_series/7267.

Kosch, Lukas, Annika Schwabe, Hajo Boomgaarden, and Günther Stocker. 2024. "Experiencing Literary Audiobooks: A Framework for Theoretical and Empirical Investigations of the Auditory Reception of Literature." *Journal of Literary Theory* 18 (1). https://doi.org/10.1515/jlt-2024-2005.

Kühne, Katharina, Martin H Fisher, and Yuefang Zhou. 2020. "The Human Takes It All: Humanlike Synthesized Voices Are Perceived as Less Eerie and More Likable. Evidence From a Subjective Ratings Study." *Front Neurorobot* 16 (14): 593732. https://doi.org/10.3389/fnbot.2020.593732.

Levac, Danielle, Heather Colquhoun, and Kelly O'Brien. 2010. "Scoping Studies: Advancing the Methodology." *Implementation Science* 5 (69): 1–9. https://doi.org/10.1186/1748-5908-5-69.

Linkis, Sara Tanderup. 2021. "Reading Spaces: Original Audiobooks and Mobile Listening." *Sound Effects* 10 (1): 43–55. https://doi.org/10.7146/se.v10i1.124197.

Mădălina, Amalia. 2024. "AI Audiobook Narration: How to Turn a Book into an Audiobook Using AI." *PublishDrive*, July 3. https://publishdrive.com/ai-audiobook-narration-for-ebooks.html.

Magadán, Marta, and Jesús I. Rivas. 2018. "Digitization and Business Models in the Spanish Publishing Industry." *Publishing Research Quarterly* 34: 333–46. https://doi.org/10.1007/s12109-018-9593-0.

Magadán, Marta, and Jesús I. Rivas. 2019. "Digitalización y Modelos de Negocio En La Industria Editorial Española." *Journal of Technology Management and Innovation* 14 (1): 63–72. https://doi.org/10.4067/S0718-27242019000100063.

Magadán, Marta, and Jesús I. Rivas. 2020. "El Audiolibro En España: ¿industria o Modelo de Negocio?" *Profesional de La Información* 29 (6): e290625. https://doi.org/10.3145/epi.2020.nov.25.

Mahum, Rabbia, Aun Irtaza, and Ali Javed. 2023. "Text to Speech Synthesis Using Deep Learning." In *Intelligent Multimedia Signal Processing for Smart Ecosystems*, edited by Shabir A. Parah, Nasir N. Hurrah and Ekram, 289–305. Springer. https://doi.org/10.1007/978-3-031-34873-0_12.

Ministerio de Cultura y Deporte. 2023. "Barómetro de Hábitos de Lectura y Compra de Libros, 2022." Madrid: Ministerio de Cultura y Deporte. Gobierno de España. https://www.cultura.gob.es/actualidad/2023/02/230227-barometro-habitos-lectura.html.

Nieborg, David, Thomas Poell, and J. van Dijck. 2022. "Platforms and Platformization." In *The SAGE Handbook of the Digital Media Economy*, edited by Terry Flew, Jennifer Holt and Julian Thomas, 29–49. Sage Publications. https://doi.org/10.4135/9781529757170.n4.

Rivas, Jesús I., and Marta Magadán. 2022. "An Overview of the Audiobook Marketplace in Spain." *Publishing Research Quarterly* 38 (1): 168–79. https://doi.org/10.1007/s12109-022-09864-7.

Rodero, Emma, and Ignacio Lucas. 2023. "Synthetic versus Human Voices in Audiobooks: The Human Emotional Intimacy Effect." *New Media & Society* 25 (7): 1746–64. https://doi.org/10.1177/14614448211024142.

Rodero, Emma, and Laura Romero. 2022. "Let Me Listen to Where You Are. Spatial Dimension Resources in Audio Stories Can Increase Imagery, Transportation, Attention, and Recall." *Media Psychology* 25 (1): 155–79. https://doi.org/10.1080/15213269.2021.1880439.

Ryzhko, Olena, Tetiana Krainikova, Svitlana Vodolazka, and Kateryna Sokolova. 2024. "Generative AI Changes the Book Publishing Industry: Reengineering of Business Processes." *Communication & Society* 37 (3): 255–71. https://doi.org/10.15581/003.37.3.255-271.

Saikaly, Tony. 2023. "The Transformative Role of Artificial Intelligence in the Publishing Industry." *FADEL*, October 18. https://fadel.com/whitepaper-the-transformative-role-of-artificial-intelligence-in-the-publishing-industry/.

Shatzkin, Mike. 2008. "Digital Publishing in the US: Driving the Industry to Vertical Niches?" *Logos* 19 (2): 56–60. https://doi.org/10.2959/logo.2008.19.2.56.

Shetty, Sonal, Akash Sabarad, Harish Hebballi, Moula Husain, S. M. Meena, and Shiddu Nagaralli. 2015. "Content Based Audiobooks Indexing Using Apache Hadoop Framework." In *Proceedings of the*

Third International Symposium on Women in Computing and Informatics, 496–501. Kochi, India: ACM. https://doi.org/10.1145/2791405.2791485.

Snelling, Maria. 2021. "The Audiobook Market and Its Adaptation to Cultural Changes." *Publishing Research Quarterly* 37 (4): 642–56. https://doi.org/10.1007/s12109-021-09838-1.

Spjeldnæs, Kari. 2022. "Platformization and Publishing: Changes in Literary Publishing." *Publishing Research Quarterly* 38 (4): 782–94. https://doi.org/10.1007/s12109-022-09912-2.

Statista. 2024. "Audiobooks–Worldwide." Statista Market Forecast. Statista. https://www.statista.com/outlook/amo/media/books/audiobooks/worldwide.

Tattersall, Elisa. 2022a. "Audiobook Apps: Exploring Read-Ing Practices and Technical Affordances in the Player Features." *Information Research* 27 (4). https://doi.org/10.47989/irpaper943.

Tattersall, Elisa. 2022b. "Audiobook Routines: Identifying Everyday Reading by Listening Practices amongst Young Adults." *Journal of Documentation* 77 (6): 1–16. https://doi.org/10.1108/JD-06-2021-0116.

Viens, Ashley. 2019. "This Graph Tells Us Who´s Using Social Media the Most." World Economic Forum, 2019. https://www.weforum.org/stories/2019/10/social-media-use-by-generation/.

Watson, Amy. 2023. "Frequency of Audiobook Download or Access through the Internet in the United Kingdom (UK) as of May 2022, by Age Group." Statista. https://www.statista.com/statistics/1302492/frequency-distribution-of-audiobook-downloading-and-accessing-uk/.

Yin, Robert K. 2017. *Applications of Case Study Research*. 6th ed. London, UK: Sage Publications.

Authors and Affiliations

Marta Magadán-Díaz[1] · Álvaro Cabezas-Clavijo[1] · Pavel Sidorenko-Bautista[2] · Jesús I. Rivas-García[1]

✉ Marta Magadán-Díaz
marta.magadan@unir.net

Álvaro Cabezas-Clavijo
alvaro.cabezas@unir.net

Pavel Sidorenko-Bautista
pavel.sidorenkobautista@unir.net

Jesús I. Rivas-García
jesus.rivas@unir.net

[1] Faculty of Economics and Business, Universidad Internacional de La Rioja, La Rioja, Spain

[2] Faculty of Education, and PhD Program in Humanities and Digital Society, Universidad Internacional de La Rioja, La Rioja, Spain