



FRANKFURTER
BUCHMESSE

■ AI and Audio: How — ARTIFICIAL INTELLIGENCE is Redefining the Audiobook Industry

September 2025



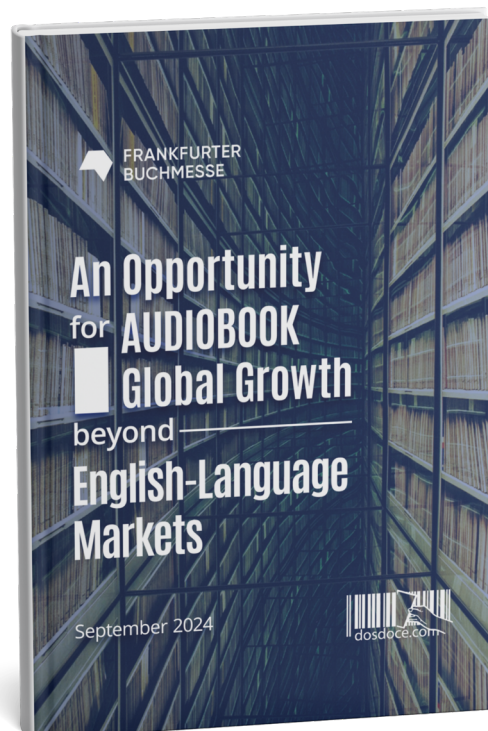
TABLE OF CONTENTS

	INTRODUCTION – EXECUTIVE SUMMARY	3
1	Section 1. Audio AI-Driven Tools & Services Map	8
	1.1 Segmentation of the AI tools & services 1.2 Global AI Race: origin of AI tools & services 1.3 Ranking of Business Activities in the AI landscape 1.4 Evolution of modern text-to-speech (TTS) 1.5 Learning curve, cost reductions and time to market 1.6 Ethical use of AI tools & services	
2	Section 2. Business Opportunities Driven by AI	24
	2.1 A more diverse and truly global audio industry 2.2 Enhanced Production Process 2.3 AI Translations - Seamless Language Listening 2.4 Distribution and Global Expansion 2.5 Avoiding a homogeneous global creative soundscape 2.6 Discoverability Amidst Content Abundance 2.7 Benchmarking AI initiatives in other creative industries	
3	Section 3. Challenges Derived from AI	42
	3.1 AI Voice Audiobooks will soon surpass Human production 3.2 A new mindset towards AI 3.3 Discoverability: AI in Marketing and Curation 3.4 Voice Authenticity 3.5 Quality, Trust and Transparency 3.6 Redefining Rights and Compensation 3.7 Innovation & Partnerships	
4	Sources	58
5	Credits	61
	About Dosdoce About Frankfurt Buchmesse / Frankfurt Audio Acknowledgements Imprint – Credits	

1. INTRODUCTION

As Artificial Intelligence (AI) tools have evolved from experimental curiosities to indispensable workflow components. In this transformational environment, the publishing industry is entering an era of unprecedented change, with AI technologies accelerating the pace and costs of audio production, increasing the creation of multilanguage catalogues, and making content accessible to broader audiences worldwide.

The majority of the publishing professionals interviewed last year in the **Global Evolution of the Audiobook Industry Report**,¹ published by the Frankfurt Book Fair (2024), declared that they had a “*dubious*” mindset about AI tools indicating that we’re tracking news and developments but not trusting the technology yet. In less than one year, the publishing community has moved to a “*cautiously progressive*” mindset which implies that most of them are already testing and evaluating different types of AI tools.



Artificial intelligence is by no means new; it is something that humans have been working on and studying for at least five decades. But over the last few years, the potentially unlimited access to large amounts of real-time data on human behavior – also known as “big data” – and a rapid increase in computational power have led to an explosion in the potential uses of artificial intelligence. As a result, it is no longer a mere scientific concept but a social reality that is already transforming everyday life, including the way we access and use all kinds of cultural content. Just like electricity in the late nineteenth century, artificial intelligence is set to become a structural commodity that affects most economic activities as well as the business models used by all kinds of companies. In other words, Artificial Intelligence is poised to revolutionize the publishing industry as the rest of the digital economy.

To help better understand the dynamics of this new scenario, the Frankfurter Buchmesse’s team encouraged Dosdoce.com to develop a new whitepaper aimed identifying the main AI driven tools and services used in the audio industry, as well as to cross-reference the main business opportunities and challenges derived from the irruption of AI across various dimensions, including audio production, distribution, marketing, and even content creation.

To achieve this objective, this report is divided into three main sections. The first section analyzes the growing ecosystem of artificial intelligence tools and services most commonly used today by audio industry professionals worldwide. Dosdoce has conducted extensive research identifying more than 160 Audio AI-driven entities from the creation of synthetic voices engines to simultaneous translations of audiobooks and podcasts into multiple languages, as well as technological developments aimed at marketing and advertising audio content, to applications for automating all kinds of processes in the pre- and post-production of audio content with the support of artificial intelligence.

One of the key takeaways of this report is that AI tools will empower professionals by freeing them from repetitive and tedious tasks in production processes, allowing them to channel their time into more creative and complex aspects of their work. "The 160+ AI solutions and services identified in the Global Audio AI-Driven Tools & Services Map will become an extension of the work of audio professionals to improve their productivity and creativity, not to replace it. When all audio professionals use the same AI tools, human inspiration and creativity will be the differentiating factor," indicated **Javier Celaya, founding partner of Dosdoce.com and author of Global Audio AI-Driven Tools & Services Map.**

Another key finding of the Report is that AI tools are offering invaluable automatization of processes, as well as new levels of efficiency and cost savings. Without a doubt, AI can generate complete audiobooks in days rather than weeks, supposedly reducing production costs by up to 50% compared to production processes using human voices, but the supposed time and cost savings are not immediate, as we will deep-dive later on in the Report, given that there is a learning curve and process adaptation that can take up to at least two or three years. The second and third sections of this report aggregate the opinions of more than 40 industry experts across the globe from leading publishing houses, to audio production studios, as well as distributors and streaming platforms, that have generously contributed with their insights about the main business opportunities and challenges derived from the irruption of AI across various dimensions, including audio production, distribution, promotion, and even content creation.

Readers will see that another of the key takeaways shared by the experts interviewed is that AI tools are now capable of managing complex processes that previously demanded extensive human coordination such as metadata enrichment, generating alternative text for images or assembling indexes. These tools are also very effective in time consuming editorial tasks such

as assessing manuscripts, performing copyediting, and proof-reading translation services, but the mentioned industry leaders affirm that most AI tools still need a human review to carefully validate and approve the final versions of these processes.

As a derivative of all these changes, an increase in jobs postings in the publishing industry requiring knowledge of generative AI in just two years is no coincidence; it is a clear sign of a profound transformation taking place in the job market. The demand for knowledge in generative AI tools has skyrocketed because companies are witnessing their direct impact on productivity. The most in-demand AI skills include AI-enhanced content creation, data-driven marketing, AI audio strategy, and project management into business processes—all requiring strong analytical and strategic capabilities.

One of the most critical decisions early in any AI project involves defining a skills training strategy specifically for publishing houses internal audio teams. The ultimate success of your AI strategy will hinge on gaining strong commitment and engagement from your knowledgeable and experienced publishing professionals. We are no longer talking about an emerging technology, but about a set of tools and services that are transforming the way we work, as we will see in later on this report.

Another recurring theme discussed among the industry leaders that contributed with their insights for this report is that Artificial Intelligence is likely to radically change the way we access and consume all kinds of cultural content in the next decade challenging long-standing norms in copyright and remuneration. There is real concern among the publishing sector that translators, voice talent and production studios may see their economic foundations erode. To adapt, the publishing industry, together with the tech sector, must create innovative models to manage derivative works and compensation, including fair royalty structures for content created or enhanced by AI, as we will see later on in this report.

In the conversations we have had with them, they pointed out that AI is one of the sector's biggest challenges, but at the same time our biggest business opportunity. Readers of Sections 2 and 3 will see that the insights and quotes collected bring to light the remarkable potential for scale, innovation and accessibility, as well as to issues related to cultural diversity, quality control, and above all creative rights and consumer trust.

We hope that this report will provide the publishing sector with new ideas on the possible uses of AI tools to grow their audio businesses, as well as enough insights from industry peers to encourage them to embrace the upcoming industry transformation.



Javier Celaya

Founder of Dosdoce.com

SECTION 1- GLOBAL AUDIO AI-DRIVEN TOOLS & SERVICES MAP

We seem to be approaching a pivotal moment in which AI-driven publishing solutions can help publishers scale audiobook production, expand distribution opportunities, and create new revenue sources. At the same time, AI tools are helping creators and producers writing their manuscripts, generating voices, editing audio, designing covers, drafting show notes, translating content into dozens of languages, and even recommending what content to produce next. The walls between human and machine-made audio are rapidly disappearing.

To help the publishing community better understand the potential benefits of using all types of AI-driven solutions and services, Dosdoce has conducted extensive research to identify the main technological solutions and services related to the world of audio, from the creation of synthetic voices, to solutions for simultaneous translations of audiobooks and podcasts into multiple languages, as well as technological developments aimed at marketing and advertising audio content, to applications for automating all kinds of processes in the pre- and post-production of audio content with the support of artificial intelligence.

At the end of a three-month research process (June-August 2025), we selected a sample of more than 160 tools that we believe can be very useful for professionals across the industry, whether they are publishing groups, audio production companies, streaming platforms, or even creators and narrators themselves.

GLOBAL AUDIO AI-DRIVEN TOOLS & SERVICES MAP

VOICE ENGINE MODELS & AGGREGATORS



SOUND DESIGN TOOLS



STREAMING PLATFORMS & SALES CHANNELS



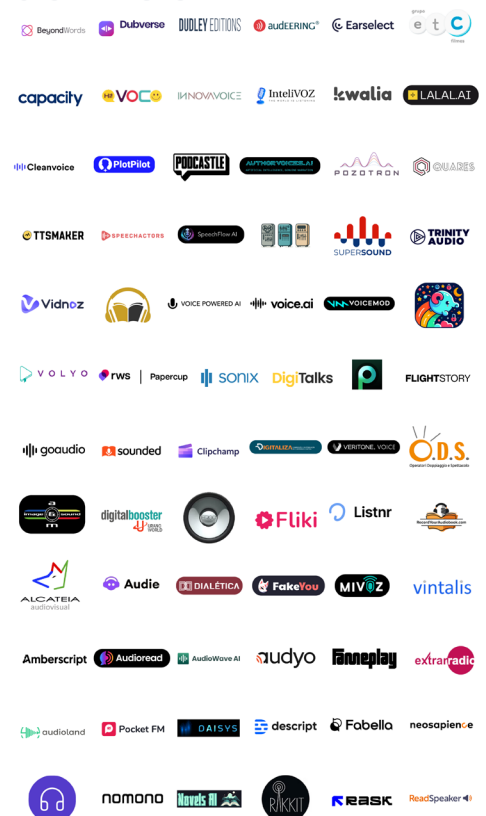
MARKETING SERVICES



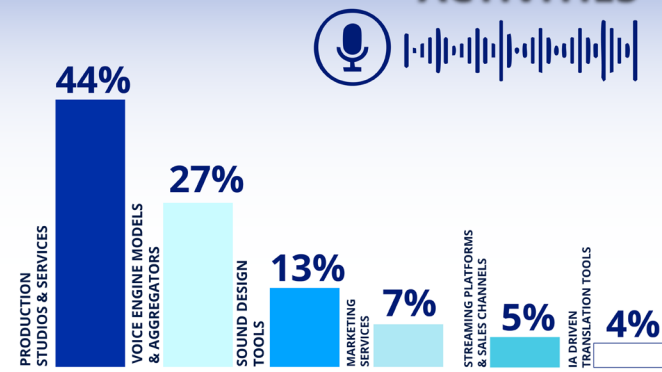
AI DRIVEN TRANSLATION TOOLS



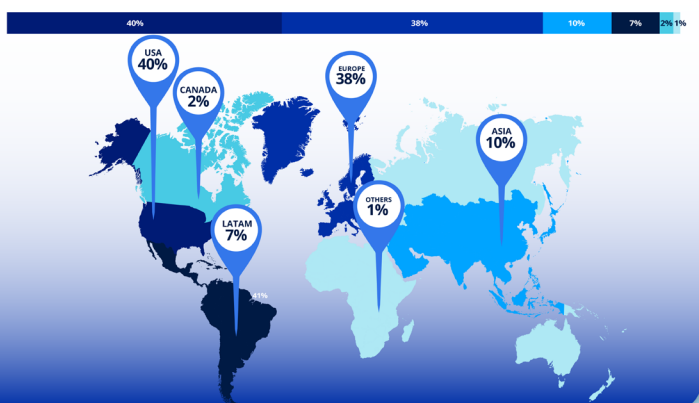
PRODUCTION STUDIOS & SERVICES



RANKING BUSINESS ACTIVITIES



GLOBAL AI RACE



1.1 SEGMENTATION OF THE AI TOOLS & SERVICES

The AI solutions and services aggregated in the Global Audio AI-Driven Tools & Services Map have been divided into six groups, representing each of the key business activities in the audio industry:

- AI voice models: Entities that have developed their own voice synthesis technology engine, as well as companies that aggregate several of these third-party models.
- Audio production studio and services: Entities that use several AI based tools to offer value-added services around their production processes.
- Translation services: Multilingual translation tools that use AI to help create international catalogs for distribution around the world.
- Distribution/sales channels for audio content: Distribution platforms that are committed to using AI to offer better service in their sales channels.
- Sound design: Tools that use AI to enrich the sound design processes of audiobooks, podcasts, audio fiction, etc.
- Marketing and advertising: AI-based tools to improve all audio content promotion processes in order to increase their visibility in the market.

Important Note:

In several cases, it has been difficult to assign a single category to some entities, since many of them carry out various complementary business activities related to the use of AI. For example, WonderCraft is a tool for creating AI-generated ads, but it is also a podcast producer. To carry out these tasks, it aggregates up to six Voice Engine models. Since their main business activity is AI-powered advertising, we have decided to assign them to the Marketing category.

The same applies to Ingram, which is one of the leading distribution platforms in the United States. In this case, we have also chosen to assign them to the Marketing category, since they have developed a wide range of AI-supported marketing services for publishers, such as AI Generated Ad Copy & Image, which automatically creates optimized ad visuals and headlines based on your book's metadata.

And this is true for many other entities, showing that AI tools and services are increasingly spreading into multiple processes and tasks across all organizations in the book world.

1.2 GLOBAL AI RACE: ORIGIN OF AI TOOLS & SERVICES

AI is transforming the balance of power worldwide, igniting a high-stakes technological contest among leading nations. The United States, the European Union, and China are each pursuing their own regulatory strategies, reflecting their particular priorities, objectives, and underlying principles.

Unlike earlier forms of AI that focus mainly on analyzing data and detecting patterns, generative AI has the ability to produce entirely new material—ranging from text, voice and images to software code—making it a game-changing catalyst for innovation across multiple sectors, including publishing and media. In the United States, a market-oriented approach prevails, with private enterprises driving progress and government intervention kept to a minimum. The European Union takes a different path, prioritizing ethically responsible AI through comprehensive legislation such as the AI Act, designed to promote transparency, ensure accountability, and manage risks—a subject we will examine in greater detail later in this report. China, by contrast, advances its vision of digital sovereignty through strict state control, ensuring AI developments serve national objectives including economic expansion, social harmony, and strategic military goals².

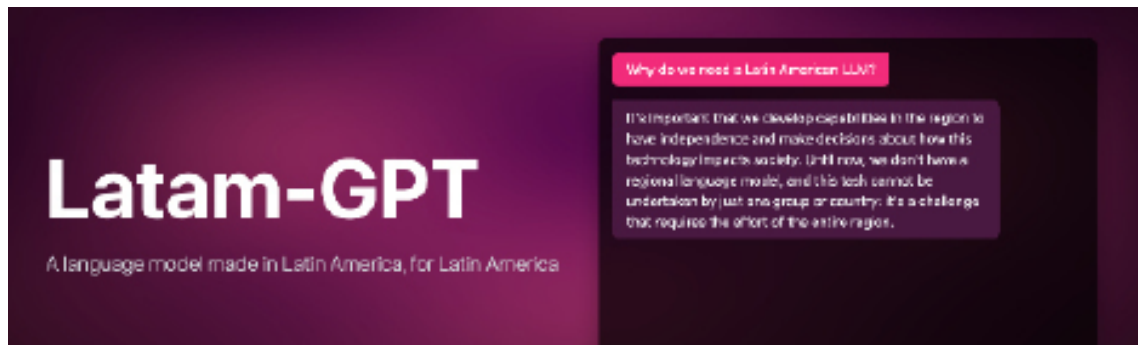
Taking into consideration this geopolitical scenario, we decided to analyze the geographical origin of the AI tools because it has implications regarding privacy, data sovereignty and algorithm bias, among other issues. As the table indicates, the geographical origin of the tools aggregated in the Global Audio AI-Driven Tools & Services Map shows a notable concentration in markets that have been investing in the development of new audio tools based on various AI technologies for the last couple of years.



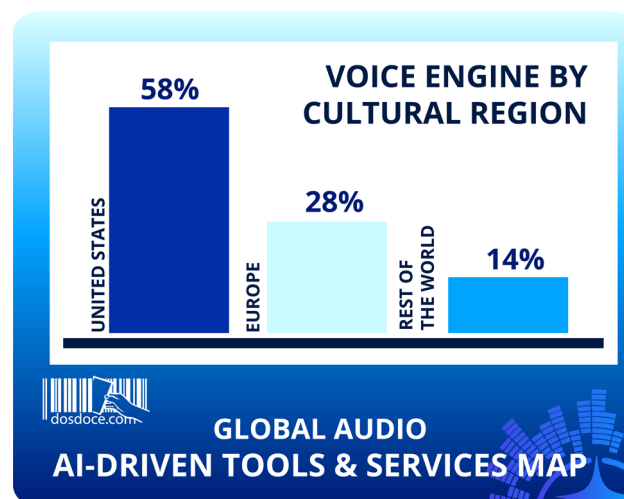
As the graph indicates, the United States leads the ranking as the headquarters of the tools and services analyzed, representing 40% of our sample, followed very closely by Europe with 38%. In third place, but somewhat distant from the previous two but growing very fast, we find the Asian continent, representing 10% of the entities, and then Latin America, which accounts for 7% of the entities analyzed.

Africa is merely represented in the Map, as well as other emerging regions and markets, but this situation could

change very rapidly as new AI tools and services come into the market with greater language and cultural sensibilities such as LATAM-GPT, a Latin American initiative to develop a language model built entirely in the region, among other initiatives like Masakhane in Africa.

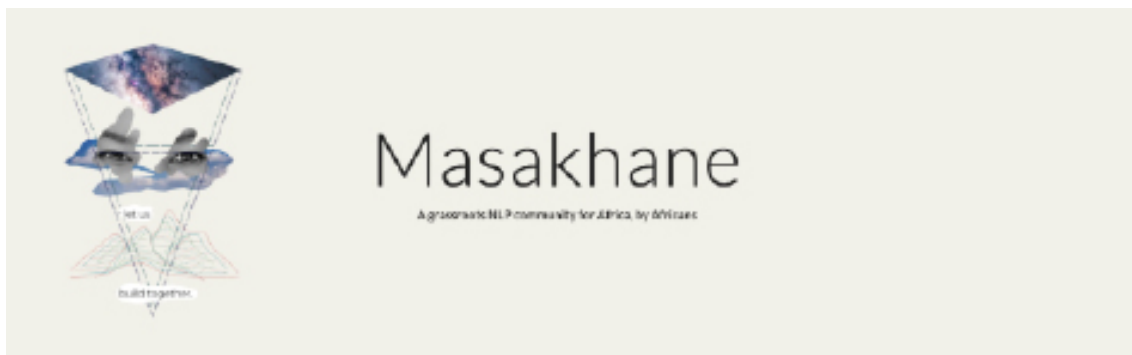


As we will analyze in the next section of the Report, the United States leads as the birthplace of the main Voice Engines, representing 58% of the entities analyzed, followed by Europe with 28% (mainly through the UK). Control over the ownership and management of voice and language models will be crucial in shaping the future of global culture, especially among emerging markets.



For example. "In just a few years, Swahili has gone from being digitally invisible to algorithmically indispensable. According to recent data, Swahili is now among the top 10 most spoken languages integrated into AI systems globally, reaching over 100 million native and second-language speakers across East and Central Africa. Tech giants—from Google and Meta to Huawei—

have incorporated this language into their multilingual models, signaling a recognition of Africa's rising digital market. Yet behind this celebratory narrative of inclusion lies a more sobering reality: Swahili-language AI, though localized, remains externally controlled. Estimates suggest that less than 5% of AI training data involving African languages are sourced or governed by African institutions themselves. This structural asymmetry risks perpetuating Africa's digital dependence."³



Supporting this viewpoint, **Ama Dadson, Founder & CEO AkooBooks and Frankfurt Audio Ambassador**, asserts that "AI can democratize storytelling giving African voices global reach through language, discovery, and new creative formats. Or it can deepen inequalities if we are excluded from owning the infrastructure. The next three years will decide which future we build." Ensuring emerging audio market languages and stories aren't left out of AI infrastructure and training data is essential to building truly inclusive audio industry. As we will see later on this Report, if only dominant languages and narratives are prioritized, we risk preventing diverse voices from shaping the digital landscape and deepening global inequities. This is closely tied to preventing a widening digital divide where the Global South becomes only a consumer, not a co-creator, of AI. By recognizing the cultural and linguistic richness of underrepresented languages in the audio industry and integrating them into new AI systems, we can foster a more equitable digital future where innovation and creativity flow from all parts of the world.

As we will see later in Section 2 of this report, streaming platforms, audio sales channels and library systems should embrace the creation of AI solutions with a more diverse and truly global audio mindset at the forefront.

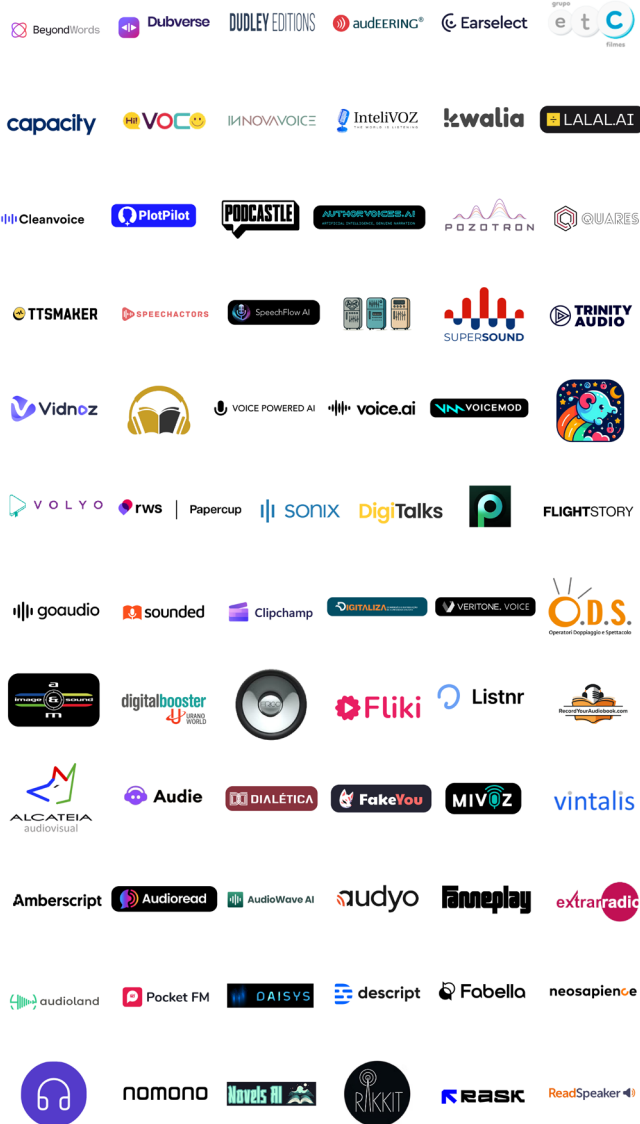
1.3 RANKING OF BUSINESS ACTIVITIES OF THE AI LANDSCAPE

Audio production companies & services lead the ranking of the main business activities in the industry, representing 44% of the entities, followed by Voice Engine companies, which represent 27% of the entities analyzed.



One of the most interesting conclusions to emerge from this study is that more and more audio production studios, which to date have produced audio content with human voices and will continue to do so in the future, are simultaneously incorporating various AI tools to offer value-added services around the various production processes.

PRODUCTION STUDIOS & SERVICES



**GLOBAL AUDIO
AI-DRIVEN TOOLS & SERVICES MAP**

In the production landscape, we have detected that there are already at least three or four production studios per country, such as EarSelect in Sweden, FlightStory in the United States, Goaudio in Germany, MiVoz in Spain, and Arabookverse

Studios in Egypt, among others, which already use various AI tools in their production processes. Their goal is to empower professionals by freeing them from repetitive and tedious tasks in production processes, such as manuscript pre-processing, automated post-production checks, and final review with the involvement of human professionals to ensure accuracy, quality, and consistency, allowing them to channel their experience and creativity into more complex aspects of their human work.

In second place of the business activity ranking, representing 27% of the sample, we find entities that have developed their own voice synthesis technology engine, such as ElevenLabs, MS Azure, and Google, as well as companies that aggregate several of these third-party models.



In this section, the United States leads as the birthplace of the main Voice Engines, representing 58% of the entities analyzed, followed by Europe with 28% (mainly through the UK). This dominant Anglo-Saxon position highlights the voice engine technological gap between different regions, indicating opportunities for growth in other languages as indicated in the geopolitical analysis described previously in the Global AI Race analysis.

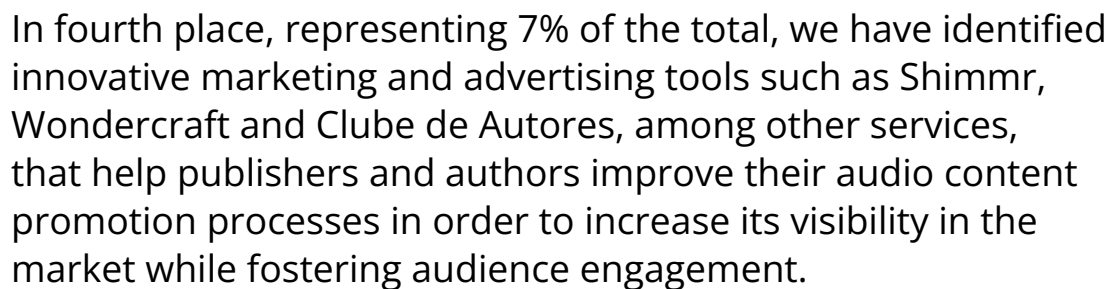
SOUND DESIGN TOOLS

AIMS AIVA Audacity MusicAI iZotope Audible Intelligence

wavepaths moises Mutant Neutone PhonoMind Soundful Vocaroo Voice.ai

dosdoc.com

**GLOBAL AUDIO
AI-DRIVEN TOOLS & SERVICES MAP**



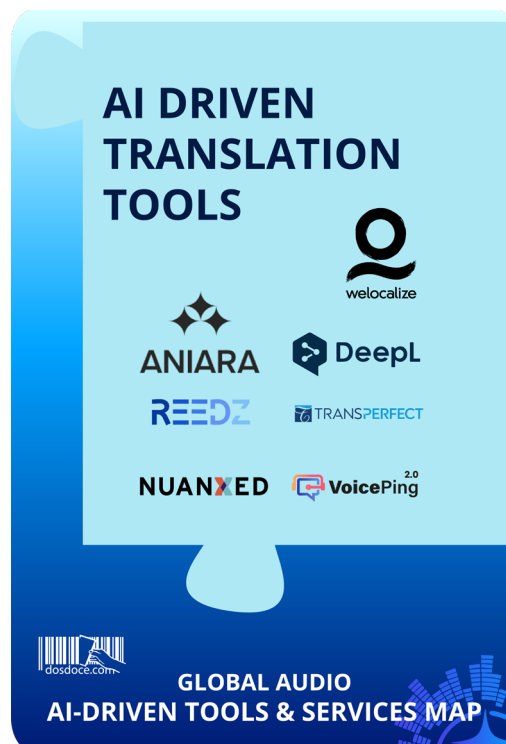
A major benefit of using AI in podcast marketing is its ability to automate repetitive tasks, giving you back valuable hours that can instead be dedicated to creative brainstorming or connecting with your audience. But AI goes beyond simple automation—it's also a powerful source of inspiration for marketers. When creativity stalls, AI tools can spark new ideas and suggest innovative ways to reimagine and repurpose existing content. With the right AI marketing tool, audio content becomes more engaging, relevant, and impactful, without losing that essential human touch.

In fifth place, representing 5% of the total entities analyzed in the Map, we have identified audio content distribution companies, streaming platforms and sales channels that are already betting on the use of AI to offer a better service on their platforms, whether from the listening experience or the creation of content itself with synthetic voices in multiple languages.



Platforms such as Storytel and Spotify have signed strategic partnership agreements with one of the leading models, ElevenLabs, while the Audible platform, which has just announced its strong commitment to the production of audio

content with synthetic voices, uses the technology engine of its parent company, Amazon. Another of the most unique developments we have identified in this category is Storytel's Voice Changer feature offered in its platform, which allows listeners to listen "on demand." Users can select the synthetic voice they like best and change it on any audiobook that has already been recorded by a human voice. In other words, they can listen to the version originally recorded by a human or change it to a synthetic voice from a selection of options. Last but not least and representing 4% of the total entities represented in the Map, we find multilingual translation tools, such as ANIARA.ONE, DeepL and Nuanxed, which, as we will see in detail later in this report, help publishers and streaming platforms create multilingual catalogs in order to attract international audiences to their authors.



AI-powered translation tools are set to play a pivotal role in the future of the publishing sector. By combining advanced machine learning with natural language processing, these solutions enable publishers and creators to efficiently translate

audiobooks, audio-dramas and podcasts, as well as marketing materials into multiple languages while preserving nuance and cultural context. As we will see later on in this Report, these tools not only broaden the reach of content to global audiences but also reduce the time and cost traditionally associated with translation. As international audio listeners continue to grow, AI-driven translation will become an essential asset for publishers aiming to expand their markets and ensure their stories resonate across languages and cultures.

1.4 EVOLUTION OF MODERN TEXT-TO-SPEECH (TTS)

In less than one year, we have experienced an impressive evolution of modern text-to-speech (TTS) technologies that can now generate narration closely mimicking human performers. Voice cloning, also known as replica voices, and adaptive emotion synthesis tools enable narration that matches tone, pacing, and even character-specific voices, making for richer and more diverse listening experiences.

More than 50% of the tools analyzed in the Map allow users to design custom replica voices, giving professional narrators the ability to create a replica profile of their voice using different AI technologies. On the other hand, 60% of the tools include options to edit and modulate the characteristics of the selected voice, enabling adjustments to match the specific preferences of the audio project.

This rapid evolution of AI voice technologies is offering publishers unprecedented business opportunities to expand their global reach creating multilanguage catalogs, as well as creating new immersive audio narratives, limited today by logistical or financial constraints, as we will see later on in the Report.

1.5 LEARNING CURVE, COST REDUCTION AND TIME TO MARKET

Another key finding of this research report is that most industry professionals consulted agreed that, while synthetic narration greatly expands the production alternatives, especially for niche and back-list catalogs, it is not an economical option when the goal is high quality due to the fact that considerable human time is still required for careful post-production.

AI can generate complete audiobooks in days rather than weeks, but most of the production companies contacted estimate that producing audiobooks with AI reduces production costs by around 50% at the start of the project (second year) and up to 80% in a later phase (from the third year onward), compared to production processes using human voices.

To accelerate the AI learning curve, avoiding already made mistakes, we strongly recommend publishers to closely analyze the initiatives with AI tools launched by our next-door audio content neighbors, like podcast platforms, film studios and media conglomerates, as described in Section 2 of this Report.

1.6 ETHICAL USE AND COMPENSATION

The ethical use of AI is a paramount concern for those entities analyzed in Global Audio AI-Driven Tools & Services Map who prioritize transparency, consent, and responsible use. Many of them indicate in their website's mission statements that they hold a deep commitment to ethical practices in the development of their AI technologies which are meticulously trained exclusively on their proprietary datasets, where every narrator has provided informed consent and compensation before recording.

They firmly believe that AI tools are designed to enhance the work of audio professionals and voice talent rather than replace them. Their intention is to alleviate them from repetitive and tedious narration tasks, as well as earning passive royalties by licensing their voices, allowing them to channel their time, expertise and creativity into more intricate aspects of their human work. Aligned with this view, **Mati Staniszewski, Co-founder & CEO of ElevenLabs**, noted in a SXSW fireside chat with Jeremy Kahn, AI Editor at Fortune, at SXSW in London (June 2025) that “in less than two years, voice actors have earned a combined \$5 million through the ElevenLabs Voice Library”⁴. While this figure represents meaningful progress, many industry leaders interviewed for this report stress that current AI-related royalties remain far from sufficient to sustain a professional voice talent community—one that has historically received over €50 million annually for the past five years from European streaming platforms and publishers to support audiobook production, marketing, and promotion⁵. Looking ahead, greater transparency around compensation will be essential. We will explore these sensitive issues in more depth later in this report.

With the sheer volume of AI audio content on the rise, publishers and audio sales platforms now face a greater emphasis on curation rather than just content creation. All audio sales distribution platforms should dedicate themselves to adopting clear standards for labeling AI-generated works, giving listeners the information they need to make informed decisions, as we will see in Section 3 of this report.

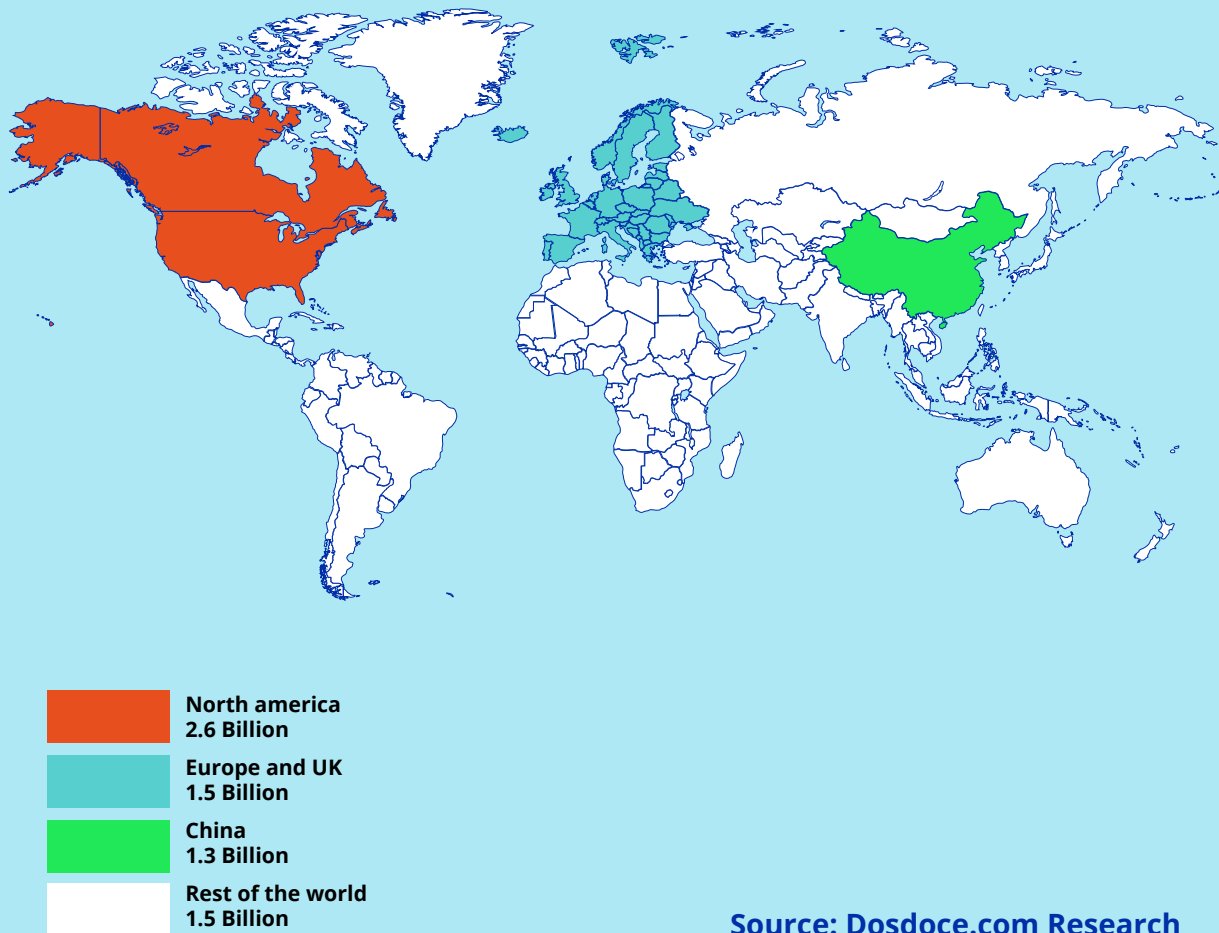
SECTION 2 - BUSINESS OPPORTUNITIES DRIVEN BY AI

This section aggregates the opinions of more than 40 industry experts across the globe from leading publishing houses to audio production studios, as well as distributors and streaming platforms, that have generously contributed with their insights about the main business opportunities derived from the irruption of AI across various dimensions, including audio production, distribution, marketing, and even content creation.

2.1 A MORE DIVERSE AND TRULY GLOBAL AUDIO INDUSTRY

Over the past 15 years, the global audiobook market has experienced staggering growth, representing an industry worth nearly \$7 billion annually worldwide: \$2.6 billion derived from North America; close to \$1.5 billion generated in Europe; another \$1.0 billion generated in China and close to \$2.0 billion generated in the rest of the world. Several sources indicate that this market will continue to grow until reaching 35 billion dollars by 2030⁶, thanks to the upcoming increasing production of thousands of new audiobooks with AI tools in underserved languages such as Chinese, Hindi, Spanish, Italian, French, Arabic, Swahili and Portuguese, among other languages.

SIZE OF THE GLOBAL AUDIO INDUSTRY

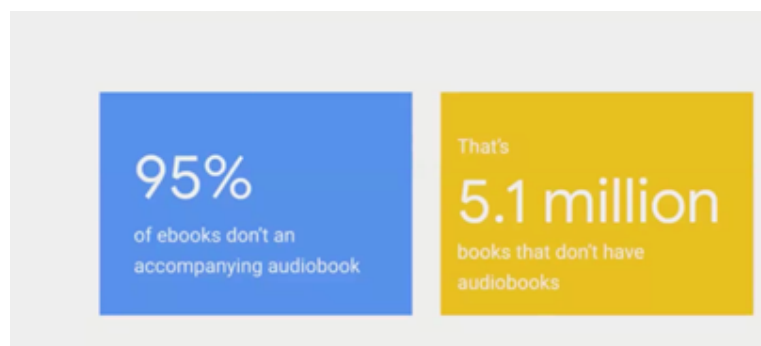


A repeated premise throughout the second edition of **PARIX Audio Day**⁷, organized by FGSR in Madrid, was that the growth of the audio industry worldwide depends on increasing the availability of audio content. Naturally, the largest catalogs belong to the most consolidated markets: there are approximately 800,000 audiobooks in English and close to 300,000 in German. In contrast, in emerging markets such as French, Italian, Portuguese, or Spanish, catalogs, while growing steadily, remain limited.

Currently, the Spanish-language audiobook market has about 45,000 titles, the French market 25,000, and the Italian language market close to 20,000.

In line with this perspective, many of the industry leaders consulted in this report indicated that making more audiobooks available in more languages will lead to more audiobooks being listened to all over the world. The rise of AI is dramatically easing the obstacles involved in producing audio content, making it possible for publishers -especially in audio emerging territories such Latin America, Africa and Arabic Language markets, among other territories- to create audiobooks in multiple local accents and languages without the need for huge investments, but also ensure synchronized releases across regions, thus amplifying global impact. “There are thousands of works that have never been produced with human voiced narrators and AI voices give an opportunity for those titles to be more accessible to all” indicates **Michele Cobb, Publisher of AudioFile Magazine**

Despite millions of print books being published annually, only a small portion are available in audio format. As indicated by **Videl Bar Kar, Vice President Audio of Bookwire** at DPR’s AudioNext Summit (May 2025), approximately 95% of eBooks have yet to be converted to audio, that’s 5.1 million books that don’t have an audiobook version.⁸ This presents a once in a lifetime opportunity for publishers to invest in. With growing demands for accessibility and the need to expand audiobook catalog in many languages and markets, there is no doubt that AI narrated audiobooks have an important and expanding role to play.



“Closing catalogue gaps, making more content accessible to listeners and solving the discoverability challenge will all need to leverage the solutions that AI offers. AI production and translation will add a whole new layer on the existing business, accelerating catalogue creation while new AI driven recommendation technologies will close the discoverability gap by matching the next great book for individual listeners” highlighted **Eric Bartoletti, Head of Business Development of Bookwire.**

As is widely recognized, most books only manage to sell a few hundred copies within the first year or two after release. Such limited sales figures don't justify the cost of creating a human narrated audiobook, which requires a narrator, producer, and recording studio. Does this mean these books should remain inaccessible to readers who favor listening instead of reading in print? Aligned with this view, **Rachel Ghiazza, Chief Content Officer of Audible,** indicated that “an estimated 97 percent of books are not yet available in audio, and that gap is even wider for non-English language titles. Audible is excited about how AI will help more creators share their stories in audio. Millions of titles are never produced as audiobooks due to a variety of reasons – whether capability, time, or cost.

Meanwhile, people are listening to more audiobooks than ever, and we think AI narration can help meet this growing interest alongside our continued investments in professional narration and original content. We want to give creators different ways to bring their work to life through audio – whether they want to handle AI production themselves or have Audible manage the process end-to-end. Our translation offering includes optional human review to capture cultural nuances and maintain quality. Looking ahead, we plan to add more languages, accents, and character performances to our AI tools so we can better serve diverse audiences around the world.”

The AI technological leap opens doors for a wider production and distribution of new titles, particularly those focused on local authors and themes, which previously would not have been considered for audio adaptation. With AI, these works will definitively have an audio version—and thus discovered—for the first time, creating new business opportunities creating a more diverse and truly global audio industry. Conforming to this approach, **Anne-Sophie Steinlein, CEO of YouScribe** believed that “AI dramatically lowers the barriers to entry for audio production, enabling publishers, but also authors, especially in Africa and other emerging markets or underserved regions, to produce and distribute audiobooks in multiple languages, without costly infrastructure. Real-time translation and natural voice rendering may soon make any book available in any language, unlocking global audiences. This zero-barrier audio production creates new economic value, new audiences, while amplifying underrepresented voices and cultures. At YouScribe, we believe AI can be a true catalyst for inclusion and linguistic diversity. It powers our mission to make reading accessible to all through mobile-first, localized content”.

Consistent with this viewpoint, **Laure Saget, Managing Director Audiolib**, Hachette Group asserts that “editorial strategy can benefit from IA to improve publishing strategy & editorial choices in a French market where we still have to make “Cornelian choice” on which book to publish in audiobooks. Production costs can be reduced thanks to IA in preproduction and postproduction phases, as well as production time can also be reduced in order to improve synchronized release dates of audiobooks. In France, the first main challenge is the growth of the catalogue in French language. The share of books produced in audiobooks is still too low to be able to give to the audience the widest catalogue possible. The second main challenge for the French audiobook catalogue is to develop audiobook offers in the whole French speaking language territories like African countries”.

A similar proclaim arises from the Arabic Language Markets **Ahmed Rewihel, Co-founder and Chief Business Officer, Arabookverse** proclaims that “AI is creating exciting new opportunities for the Arabic audiobook industry, such as: Recommending the right books to the right listeners based on their interests. Making audiobook production faster and cheaper using AI voices. Helping Arabic stories reach global audiences, including fewer common genres and dialects, is one of the main business opportunities in the region. “Right now, AI-generated Arabic voices are not perfect. But we’re confident they will get much better as big tech companies invest more in Arabic language support” noted **Ahmed Rewihel.**

2.2 ENHANCED PRODUCTION PROCESS

A unifying perspective among the industry leaders interviewed for this report indicated that AI’s automation capabilities are reducing time and costs across the production workflow which will in turn enable more time and money for human creativity. Consistent with this viewpoint **Helena Gustafsson, Chief Content & Publishing Officer at Storytel** indicated that “AI is rapidly transforming the book industry, offering invaluable tools for efficiency and cost savings. However, the true power lies in seamless collaboration: AI can enhance and streamline processes, but human creators should provide the vision, emotional intelligence, and critical judgment that define compelling content. For publishers and creators, the future is about leveraging AI to amplify human ingenuity, not to replace it. This strategic integration is essential for thriving in the competitive global publishing landscape”.

In this new working scenario, many publishing professionals indicated that hybrid workflows are already taking shape, with AI supporting but not replacing human ingenuity. Many tools showcased in the Global Audio AI-Driven Tools & Services Map, support professionals with real-time feedback during narration sessions or flagging inconsistencies or missed lines, which accelerates the correction process and improves overall efficiency, among other production tasks.

This blending of human and machine strengths is widely seen as the optimal model for sustaining both quality and productivity. In accordance with this philosophy, “Penguin Random House Audio’s emphasis is on using AI to make the pre- and post-production processes more efficient, allowing our production teams to focus on their creative work. We’ve developed tools to streamline tasks like transcription and proofing, automate workflow alerts, and draft prep materials like pronunciation guides and character breakdowns. In every instance, we’re following the crucial principle to ‘start with a human and end with a human,’ practicing careful review by a staff member” stated **Amanda D’Acerno, Global President & Publisher of PRH Audio**, one of the leading publishers that is heavily investing in this format with a global catalog of more than 50,000 audiobooks in multiple languages beyond English (Spanish, German, Chinese, etc.).

AI tools aid in script adaptation, character assignment, and even special sound design effects, are facilitating complex productions once limited by cost constraints. By reducing costs, logistical and time-consuming processes—such as managing domestic and international studio bookings, narrator retakes, pronunciation mistakes and lengthy post-production timelines, among other processes—publishers are able to enhance human production that would have been economically unfeasible just a few years ago.

Conforming to this approach, **James Long, CCO of Pan Macmillan**, indicates that “AI offers opportunities to scale audiobook production, enabling a broader range of titles to be made available in audiobook format for the first time, helping us to reach more listeners. Our core principle will always be to use only ethically licensed AI voice technology. We expect advanced AI voice technology to reduce costs, accelerate production timelines, and ultimately expand audience reach. At the same time, new possibilities are emerging through AI-enabled voice innovation, allowing for distinctive and enriched listening experiences” Notwithstanding, Long also added that “key global challenges lie in managing AI’s impact on creative labor, particularly for voice artists and freelance production professionals. Establishing a robust ethical and commercial framework for voice rights is critical”.

As a summary of this section, the use of AI for pre and postproduction tasks such as removing background noises removal, breathing sounds, error detection or to create pronunciation lists is already taking place in many publishing houses and production studios that are already embracing these new tools. These operational advances lead to both cost savings and greater flexibility, especially benefiting niche, mid-list and long-tail titles that historically struggled to justify a human based audio adaptation.

2.3 AI TRANSLATIONS - SEAMLESS LANGUAGE LISTENING

AI-powered translation tools harness advanced algorithms and large language models to process vast amounts of text, delivering near-instant translations across dozens of languages. These tools excel in speed, cost-effectiveness, and consistency, making them invaluable for publishers managing large catalogs or seeking rapid global expansion. “The opportunity for wildly expanded translation and narration could make works available to wider audiences and expose us to the richness of global writers who until now have been confined to smaller markets. Those same opportunities create challenges both at an individual level – the livelihoods of narrators and translators - and at an industry level - the concentration of rights or retail distribution in a few powerful platforms with the resources to deploy AI at scale” noted **Michael Tamblyn, CEO of Kobo.**

However, AI alone is not a panacea. While AI-generated translations have improved dramatically in fluency and context, human review, proofreading and post-editing remains essential, especially for publishing groups that have a reputation in the market for creating quality driven books irrespective of their format: print, ebooks or audiobooks.

A bad translation can rapidly damage the reputation and credibility of a publishing house brand gained over years. To address these issues, next generation AI translation tools provide hybrid human-machine workflows, with AI supporting but not replacing human criteria and expertise.

As we have analyzed throughout this report, audio content is traveling more and more around the world, so agents and publishers are seeking new editorial services and technological solutions to translate, produce, and distribute authors’ works in

multiple languages with the help of AI to reach a global audience. “We will see an incremental increase in produced volumes, and a world where translations and adaptations can be generated by the minute. Production increase rate will be greater than consumption increase rate (seen already in the podcast world). It is tough to stand out and get your story highlighted today, tomorrow it will be a completely different ballgame. If you are a storyteller or publisher, you need to sharpen your stories and the story about your stories”, indicated **Rickard Lundberg, CEO of Aniara.one**, a new 21st-century publishing platform that has built a sophisticated language model, tailored to specific genres and authors, using natural language processing technology to translate works into eight languages with human involvement. It has been operating for just one year and has already signed agreements with 38 authors from six countries, two agents, and four publishers, turning an author’s backlist into a potential international success, revolutionizing traditional publishing processes and reducing production costs by up to 85%.

This fundamental shift will make literature truly borderless and allow publishers to extend their reach well beyond their native-language audiences. In this context, **Richard Charkin Founder, Mensch Publishing** questions itself: “how easily can English-language files be converted to other languages cost effectively? How far might English-language global publishers undermine other-language local publishers?”.

2.4 DISTRIBUTION AND GLOBAL EXPANSION

AI is breaking down longstanding barriers to international reach and language diversity. **Chantal Restivo-Alessi, CEO International FL and Chief Digital Officer at HarperCollins,** highlights the ability to rapidly expand international audio catalogs, providing previously unavailable content in emerging and smaller markets. “I see the expansion of audio catalogs for international markets focused on the type of content, like our series titles, where historically the economics would not have made it possible to create audio editions. Without sizeable catalogs, it is hard for certain markets to develop an audio market, and the creation of an incremental catalog, using AI voices combined with high end human audio productions, can help increase the offering more rapidly, thereby creating more choice for consumers and more opportunities for all parties involved.” indicated Restivo-Alessi. “We need to maintain the focus on quality and on transparency to consumers and to authors, as well as to ensure that human experts remain in charge of guaranteeing a high-quality product in all circumstances” added. **Restivo-Alessi.**

The accessibility of AI-enabled production lowers the barrier for smaller publishers and independent authors to enter the market, leveling the playing field and democratizing the creation and distribution of audio content. However, “many AI tools are closely linked to big retailers/technology platforms, and there are big questions about whether broad distribution to a range of platforms will be possible with AI: too many audiobooks are already tied to exclusive distribution deals and if AI exacerbates this trend, it will be a missed opportunity. And audio publishers will need to balance the opportunity of AI production against consumer preferences and the interests of human narrators, sound engineers and other creatives” **noted George Walkley, Director Outside Context.**

2.5 AVOIDING A HOMOGENEOUS GLOBAL CREATIVE SOUNDSCAPE

Artificial intelligence has stopped being just a tool for assistance and has become a constant presence in our routines. But a group of scientists in Germany have detected a much deeper influence: humans are starting to speak like AI. Literally. A recent report from the **Max Planck Institute for Human Development in Germany**⁹ revealed that everyday vocabulary use is changing worldwide, and large language models — like ChatGPT — are behind this transformation.

The researchers analyzed over 280,000 English-language YouTube videos from more than 20,000 academic and educational channels and found a considerable increase in the use of words that artificial intelligence typically chooses to refine or improve texts. Words such as "delve", "meticulous", "underscore", "realm", or "groundbreaking", which were previously less common in these types of academic content, began to spread in popularity after the launch of ChatGPT. According to this research, in some cases, their usage grew by more than 40% in podcasts, videos, and oral presentations. In this way, the researchers' hypothesis is clear: users, as they consume content generated or influenced by AI, start to adopt that style without realizing it.

AI is also most often developed using prevalent global languages and mainstream accents, which can lead to a homogeneous global soundscape. Overcoming this requires dedicated investment in producing less-represented local accents and languages—like Catalan or Basque and Swahili in Africa—as well as recreating distinct regional accents such as Chilean or Peruvian Spanish or Canadian or Swiss French. These efforts do more than localize; they enhance the entire global audio community.

For platforms serving diverse audiences, enabling content in voices familiar to local listeners is essential, and a fair, representative ecosystem depends on co-creating AI solutions with a more diverse and truly global audio industry at the forefront. Following this line of reasoning, **Suzanne Galvez, Director of Audiobook Editorial and Spotify Publishing,** indicated that “the industry must work collaboratively to ensure that AI tools help build incremental additive momentum, with the ethic to support and protect the incredible work of authors, narrators and publishers. Responsible use of AI tools has a real opportunity to close inequity across the global industry and revolutionise the reading experience for book lovers. AI tech brings more possibilities to increase innovation around cross-format reading and listening experiences. I think it can help close the gap between ebook and audio catalogues. I’d also love to see further democratization of creation and access to audiobooks in emerging economies, and easier cross-cultural content exchange”.

2.6 DISCOVERABILITY AMIDST CONTENT ABUNDANCE

Marketing and curation are witnessing a profound shift thanks to AI-driven automation and personalization. The flood of AI-generated content will overwhelm the sales channels and soon platforms, making it harder for readers to sift through options and avoid being overloaded or paralyzed by abundance. The core issue will be how to help audiences find content that truly fits their interests amid the abundant supply. A consistent pattern among the industry leaders who participated in these interviews indicated that we need to create new forms of outreach such as AI-generated promotional assets and interactive experiences that can help build distinct brands in a crowded field.

On the marketing front, AI will offer powerful tools for publishers to reach and engage their target audiences more effectively. AI-driven analytics can provide deep insights into consumer behavior, allowing for highly targeted marketing campaigns. Additionally, AI can automate and optimize content creation for marketing materials, such as personalized newsletters, social media posts, and targeted ads, ensuring they resonate with specific segments of the audience.

The use of AI to automate and enhance metadata creation is leading to smarter discovery and marketing. **Santos Palazzi, Head of Mass Market and Digital Division at Editorial Planeta**, highlights the improved discoverability of catalogs via enriched metadata, which in turn supports more accurate recommendation systems and facilitates the surfacing of niche genres and regional accents. “First and foremost, we must ensure that the narrative quality is adequate and does not devalue the category. Once this is confirmed, we must persuade agents and authors to allow the use of these voices to narrate their content. Secondly, in light of the flood of new titles, it will be crucial to have targeted advertising and promotional campaigns so that listeners choose titles from our catalog. The use of cloned voices of professional narrators could serve as a means of differentiation and relevance” indicated Palazzi.

A common viewpoint expressed by the industry leaders featured in this report is that AI-powered marketing tools can generate platform-specific content that will capture attention on every social network. These tools can transform key parts of an audiobook into engaging posts by showcasing memorable quotes, spotlighting valuable insights, or posing thought-provoking questions that spark interaction. By scheduling these AI-crafted marketing assets strategically, publishers can maximize reach, boost audience engagement, and streamline marketing workflows.

The improvement of metadata will enable more precise curation

of texts to deploy highly targeted digital marketing materials to ensure that each audiobook finds its reader. Industry standardization of metadata, as advocated by **Brad Rose, Vice President, Content Strategy at Hoopla**, remains essential to reduce confusion in both consumer and library markets.

Accurate metadata will be critical for differentiation and future-proofing catalogs in an increasingly globalized market. “With the increased AI utilization, in text and voice, we are advocating strongly for industry standardization in metadata and requesting this from our publishing partners. It is important that we collaborate with all publishers and authors to ensure clear and transparent data as it’s critical for our library customers. While the industry has made notable progress, we continue to advocate to all publishers to adhere to the APA guidelines regarding AI voice in metadata” stated Brad Rose.

Integrating AI tools in marketing can be transformative for publishing professionals looking to stand out. By automating time-consuming tasks such as writing book descriptions, identifying key words and metadata or creating social media posts, AI simplifies many marketing workflow—freeing team members to concentrate on producing creative high-quality assets.

The next frontier lies in refining algorithms further and creating experiences that remain compelling, diverse, and curated. Part of the answer will depend on the huge investments that the leading audio streaming and distribution platforms must assume to create advanced, fair, and inclusive systems designed with AI recommendation tools to guide listeners to most personalized content. Without this investment, consumer experiences will fragment, and valuable content may be lost in the shuffle.

2.7 BENCHMARKING AI INITIATIVES IN OTHER CREATIVE INDUSTRIES

During the research efforts undertaken for this project, we discovered many IA initiatives driven by entities in creative industries, such as podcast apps, streaming platforms and media conglomerates, that could be replicated as case studies in other content and entertainment businesses.

To accelerate the IA learning curve, avoiding already made mistakes, we strongly recommend publishers to closely analyze the initiatives taken by our next-door audio content neighbors with AI tools. We have identified more than 20 best-practices initiatives, but to avoid drafting a 100-page report, we will share four innovative, but easy to explain, projects to illustrate several examples.

2.7.1 POCKET FM

Audio entertainment platform Pocket FM has significantly reduced production costs through its proprietary artificial intelligence systems. This leading podcast platform has built a custom AI infrastructure that enables creators from scripting and voice generation to sound design and multi-language localization.

At the heart of this transformation is a seamless collaboration between human creativity and AI precision, enabling us to produce great stories at scale without compromising on quality". AI is helping us unlock high quality content faster and more efficiently than ever before. Our content costs have been significantly reduced, while creator efficiency has increased tenfold" said **Rohan Nayak, co-founder and CEO of Pocket FM.**

2.7.2 NETFLIX

Within this framework, Netflix has started integrating generative AI into its original productions. During the Q2 last post-earnings call, **co-CEO Ted Sarandos** said *El Eternauta* includes “the very first GenAI final footage to appear on screen,” referring to an AI-generated scene of a collapsing building. According to Sarandos, the scene was finished ten times faster and for less money than with traditional VFX. He emphasized that these are real artists using better tools, not replacements. AI is also helping with pre-visualization, shot planning, and effects that were once limited to big-budget films, like de-aging. **Co-CEO Greg Peters** added that Netflix is expanding AI use into search, personalization, and ad targeting, with plans to roll out interactive ads this year. Netflix quietly slipping AI into a major sci-fi series is a sign of where things are headed¹⁰.

Artificial intelligence is also transforming film preservation and cinema history. Several movie industry players are using AI tools to turn vintage films into HD Masterpieces. Studios using artificial intelligence to restore sound effects, resolution to repairing any damage. Following this initiative, publishers should consider using AI to breathe new life into classic audiobooks produced decades ago to provide users with a new vanguard listening experience.

2.7.3 NEW YORK TIMES

The **New York Times** offers the option to listen to the majority of their articles as soon as they're published, enhancing accessibility and providing another way to engage with their content¹¹. This feature uses an automated voice, which may result in occasional errors in pronunciation, tone, or sentiment.



2.7.4 THE DIARY OF A CEO

"The Diary of a CEO", starred by Steven Bartlett, is using AI to produce his new "100 CEOs" podcast¹², cloning his voice to make a host. Bartlett has written the script and then used AI for the rest. That included cloning his voice to make an AI host, creating the storyboard, and turning it into an animated video. The show, which is available on platforms like YouTube and Apple Podcasts, is identified in the notes as AI-made.

2.7.5 BLINKLIST

Blinkist is a book-summarizing subscription service based in Berlin, Germany, with 23 million downloads as of 2023. The streaming platform has automated the audio production process of their "Blinks", their long-form audio content, creating more than 200 English and German language Blinks at a much lower cost, saving valuable time (10x) and resources (20X)¹³.

It has also launched a new functionality, Blinkist AI, a feature within the Blinkist platform that leverages artificial intelligence to create summaries of various content, including articles, podcasts, and even YouTube videos. It allows users to quickly grasp the key insights from diverse sources (news articles, podcasts, YouTube videos, etc.), going beyond just book summaries.

SECTION 3.

CHALLENGES DERIVED FROM ARTIFICIAL INTELLIGENCE

In the next section of this report, we will analyze the main challenges derived from the irruption of AI. The insights and quotations collected from more than 40 industry experts bring to light both issues related to risk avoiding adverse mindsets, need to foster cultural diversity, guarantee quality control and creative rights, and above all, build consumer trust.

3.1 AI VOICE AUDIOBOOKS WILL SOON SURPASS HUMAN PRODUCTION

Human voice narrator production in English averages around 75.000 new audiobooks released a year. One of the side effects of the AI production race is that by the end of this year there will be more AI-Voiced audiobooks produced in English than human produced ones.

Two years ago, Amazon announced the launch of a beta initiative that enabled KDP authors to quickly and easily produce an audiobook version of their eBook using virtual voice narration, a very basic synthetic text-to-speech technology. In less than one year, over 40.000 AI-Voiced audiobooks titles are available with Audible, according to Bloomberg¹⁴. This year, Audible has announced plans to use AI technology to narrate audiobooks, with AI translation to follow. Through their end-to-end service, Audible will handles the entire audiobook production process,

managing every step from initial text ingestion to published audiobook in more than 100 AI-generated voices across English, Spanish, French, and Italian with multiple accent and dialect options¹⁵.

In this new audio landscape, many publishing professionals indicated that they are worried about a potential development of a split marketplace: as AI-generated narration improves, the difference between premium “top-tier” human voiceovers and more affordable synthetic options will continue to decrease. In this context, distribution services will need to balance expanding the catalog size access with AI content while maintaining quality, all while safeguarding the rich cultural context and emotional impact that only human narration can deliver.

3.2 A NEW MINDSET TOWARDS AI

Several of the key AI players included in the “Global Audio AI-Driven Tools & Services Map” have clearly indicated that one of the main challenges lying ahead is how we collectively manage the upcoming transformational phase in the publishing industry. Leading this school-of thought, **Lasse Korsemann Horne, Publishing Director at Saga | Lindhardt og Ringhof Forlag A/S** indicates that " AI-generated audiobooks are good news for readers, as all the world's books become accessible in a new format. Whereas as for audiobook publishers, I believe this signals the end of their business model. It will no longer make sense for a publishing house to sell audiobook licenses to an audiobook publisher. Conversely, it might create a boom in dedicated audio narratives that are more than just a mechanical – albeit immersive – rendering of words. AI is the end of the classic digitization of the book market, understood as a business that relies on translating a text from one format to another. Machines are handling that for us now".

Complementing this viewpoint, **John Ruhrmann, Co-CEO & Co Founder of Bookwire**, manifested that “change is coming - whether we want it or not. It is crucial that we play an active role in shaping it. Publishers, channels, producers, speakers, authors - we are all in the same boat. It needs alliances, it needs dialogue, and it needs the courage to break new ground. AI can be a tool that makes us stronger - if we are prepared to take responsibility and consciously focus on the creative, human qualities of our industry”.

As **Juraj Felix, CEO at Fameplay**, observes, the sector is undergoing a “Gutenberg Moment,” fundamentally reshaping who can participate and at what scale. “When it comes to our biological evolution, our progress is remarkably sluggish—particularly when contrasted with the rapid, exponential pace of technological advancement today. As a result, the greatest obstacle lies in the rules and constraints we, as a business species, impose on ourselves”. Complementing this point of view, **Robert Casten Carlberg, CEO and Co-founder of Nuanxed** believes that “it will be important for us all to recognize the fact that all that glitters is not gold, and therefore it is and will continue to be a fundamental challenge to try to separate hype from reality. There will be a lot of trial and error, and hopefully a few successes that are able to provide a lot of positive momentum for the industry as a whole”.

In the same vein, **Ricardo Almedia, Founder & CEO, Clube de Autores | Club de Autores | Literary Highways** strongly believed that “Publishing is, today, in desperate need of more efficient processes. In this sense, there are opportunities for AI services in essentially everything that embraces a good story, such as copy-editing, translation, story analysis, cover designs, distribution, curation and so on. It now depends solely on publishers giving up their normally strong prejudices against all things innovation and embracing a new reality that can revolutionize their businesses. Those who do this competently will thrive; those who don’t will likely fade away. It’s as simple as that”.

Several studies consistently identify talent and expertise gaps as a leading cause of AI projects failure. Many employees view the skills gap as stemming from inadequate AI training investment and internal upskilling programs, further exacerbating the problem. In accordance with this viewpoint, **Luis González, Managing Director of FGSR**, highlighted the needed capabilities that AI demands from publishing professionals: “The first priority is to understand its implications for the coming years and to define one’s professional profile based on the unique value human beings bring when leading processes in an environment of increasingly widespread AI use. The second priority is to train professionals to harness AI in areas such as production, marketing and communications, data management, design, and supply chain optimization—both to achieve significant cost savings and boost the valuable experience of reading books”.

3.3 DISCOVERABILITY: AI IN MARKETING AND CURATION

Across all the conversations we had with professionals across the audio industry to draft this report, one word sums up our biggest challenge: “Discoverability.” Many highlighted that the greatest contribution of IA to the publishing sector isn’t faster audio production reducing cost and time-to-market periods without compromising artistry, but rather smarter discovery to help listeners enjoy their next read.

In the coming three years, many of the respondents agreed that we need to reevaluate and innovate how audiences uncover stories. As listeners, we all more or less agree that the way we discover new content is flawed. Many recommendation algorithms remain stuck using outdated, generic categories and simply promote what’s already popular or bestselling. Based on this way of thinking, **Dan Vidrascu, Founder and CEO at Voxa** noted that “we’re building a listening graph: a system that understands your day, your mood, your context - and

recommends the voice, cadence, and story you didn't know you needed. By integrating AI into our listening intelligence, we personalize journeys in real time, improving user experience, boosting average consumption length (ACL), retention, and lifetime value. AI also analyzes listener behavior across regions to recommend which titles to localize next - based on genre preferences, pacing, and listening habits. We don't guess what to adapt anymore - we know".



With a tidal wave of AI-generated content flooding the market, the question becomes: how can publishers and platforms guide readers to find what genuinely connects with them amid such abundance? Several industry professionals indicated that it's our responsibility to prioritize ethical, inclusive, and linguistically diverse approaches to recommendation systems, training, and labeling of AI-created works.

The current dominance of major languages—especially English—in AI development and training threatens to drown out underrepresented languages and accents around the world. There is a risk that a flood of homogenous, emotionally flat, or error-ridden content could erode trust and ultimately undermine consumer confidence in the audiobook format. **Chris Kennelly, host of the podcast The Spoken World**, warns of “an impending “tsunami” of poorly produced or spammy audio, which could lower overall trust in the format. With the ease of production comes the challenge of maintaining both quality and discoverability amid a rapidly saturating market”.

3.4 VOICE AUTHENTICITY

A common theme among the industry leaders interviewed for this Report is the irreplaceable value of human creativity and connection in narrative work. While synthetic narration greatly expands options for niche and backlist titles, it is not a “cheap” process when high quality remains the goal; significant time is still needed for careful post-production.

Despite impressive improvements of the last few years, AI-generated voices still struggle with authenticity and emotional resonance, especially in genres demanding expressive storytelling. Current AI models often produce errors, mispronunciations, or lack appropriate emphasis, particularly with complex or technical content. This necessitates extensive manual postproduction time—negating some anticipated efficiency gains. “AI might be perfect for fiction, self-help, or romance — genres with relatively clean language and predictable narrative structures. But in non-fiction, things get tricky. Fact-heavy books often include complex terminology, names from a dozen different languages, obscure places, acronyms, and scientific expressions. These are exactly the areas where AI still stumbles — and where human narrators, with just a quick note or pronunciation guide, would glide through effortlessly. And ironically, these are the very books we want to produce more affordably. But without heavy post-editing, they risk sounding amateurish or outright incorrect — which can erode trust with the listener,” indicated **Robert Holmström, Managing Director at EarSelect.**

As AI streamline catalog offering increases, the deep notions of taste, authenticity, and humanity—the qualities that give art and creativity their unique soul—can be overshadowed. In this context, the value of human-made storytelling, well-produced, and emotionally resonant audio will skyrocket. **Nadim Sadek, Founder & CEO Shimmr** noted “the greatest challenge is preserving the evolutionary intimacy of human storytelling when

the storyteller becomes an algorithm - maintaining that 'different kind of connection' that mimics campfire conversations. We risk losing the trust that comes from authentic voices, where 'a pause can imply doubt, a sigh of grief,' because we fundamentally trust voices more than text precisely because they reveal genuine intent and emotion. The danger is creating a world awash with artificial audio that overwhelms rather than enhances our innate preference for meaningful aural experiences”.

In a worst-case scenario where almost all audio sounds the same, the things that feel different will matter even more. Validating this thinking, **Colin Hauer, CEO of Hörbuch Hamburg** notes that “we need to remind ourselves and each other of the quality we stand for and ensure we keep delivering it to our audience. International audiobook markets have different levels of maturity and different mechanisms apply in each of them. I believe that a commitment to quality productions, human oversight and a human core will greatly impact adoption rates for audiobooks also in less mature markets. Is it more expensive to work with experts rather than just clicking a button? Yes, but in the long run I see it paying off. We as an industry need to stay alert and curious and familiarize ourselves with new opportunities that are arising. But at the same time, we should not be tempted to give up control”.

3.5 QUALITY, TRUST AND TRANSPARENCY

AI's ability to clone and synthesize voices ushers in new ethical and legal dilemmas, notably regarding consent, compensation, and rights management. Many of the publishing professionals have insisted on transparent and fair payments to narrators whose voices are replicated, regardless of their direct involvement in the production processes.

Also, the absence of clear industry-wide protocols risks both reputational harm and potential legal conflicts, as laws evolve

to address AI's reach in the creative industries, especially unauthorized harvesting of voices or content—advocating for stronger safeguards and collaboration between producers and platforms. There is also concern over job displacement and the devaluation of creative skills. While most publishing sector professionals see AI as complementary, the threat to established roles in narration and production is real. Readers' relentless quest for convenience when searching in streaming platforms for their next read or listening, risks making cultural decisions increasingly uniform and indistinguishable. In this context, it is more important than ever to actively strengthen the publishing professional creative skills, nurturing a refined sense of taste that resists homogenization and champions genuine and diverse cultural voices.

Consistent with this viewpoint, **Liza Faja, Director at Lizzie Audio (Editis)** highlighted that “the main asset of a publishing house is the rights of the works. In a globalized world, where AI makes it possible to produce works in different languages, the challenge will be to defend copyright globally. In markets that are not yet mature, it's important that the proliferation of AI voices of varying quality doesn't discourage adoption of the format. This will require strict labeling of AI and human voices, and transparency throughout the chain. It is certain that tech industries will take an interest in the market, so it's essential that the rules are respected by all, in the interests of both audio players and market growth”.

Aligned with this view, **Jon Watt, Trade Audio & Business Development at Bonnier UK** asserts “the main global challenges for our industry with regards to AI narration are quality, trust and transparency. The three are inextricably linked. Publishers and retailers must behave ethically and protect production quality and listener expectations, and they must be clear and genuine in their labelling of AI-narrated products to allow consumers to make informed choice and avoid damaging discoverability”.



It is important to note that, as Chair of the Audio Publishers Group within the UK's Publishers Association, Jon Watt has led a working group aimed at establishing international standards for the classification of AI-narrated audio products, with the goal of unifying terminology and preventing confusion in the marketplace. As a result of this working group, the association has developed a set of labeling guidelines¹⁶ in collaboration with key industry players in the United States, Canada, the United Kingdom, and Europe. Its purpose is to provide transparency and clarity to consumers without making value judgments about the use of AI in narration.

Among its main recommendations, the guide suggests labeling audiobooks with synthetic voices as "AI voice" and those generated from a specific human voice as "authorized voice replica." For Watt, establishing a common language is essential to differentiate the various types of narration and ensure that publishers and retailers communicate this information clearly to consumers. In his view, more than the technology itself, it is the lack of information that causes concern among users.

Echoing this viewpoint, **Mathilde Davignon, Head of Audiobook development at Madrigall group**, noted that "one of the main global challenges for us, audiobook producers together with platforms, will be to make the difference very clear to the user between books recorded with AI or with human voices.

Information is key. Also, we producers, again together with the platforms, will have to work together to prevent voice harvesting to fight AI training”.

Children’s audio content consumption currently accounts for around 3% of total listens in English-speaking markets, as reported in the Audiobook Global Growth Report (Frankfurt Book Fair 2024). One of the key drivers behind the impressive growth of the children’s audiobook category is the emergence of audio content listening devices (such as Yoto, Tonies, MyCube Histories, etc.), which allow young children to access their audiobooks without the need for adult mediation, but the other key growth driver will be the increasing production of thousands of new children audiobooks titles with AI tools in underserved languages such as Spanish, Italian, French, Arabic and Portuguese, as indicated in Section 2 of this report.

Deep diving into this potential growth, **Jess Tarrant, Content Director at Yoto**, indicated that “we’re excited about the ways AI can help us reach more children around the world and faster. We’re currently trailing a partnership that gives us access to best-in-class AI translation tools and digital voice technology to scale the production of our Yoto Originals titles into new languages, which wouldn’t be economically viable for us to produce without AI”. But she also noted that “AI may also pose a threat consumer confidence if the proliferation of low-quality or derivative AI-generated audiobooks and music (aka ‘AI slop’) continues at its current pace. Yoto are known for the quality of our kid’s audio - it’s why families choose to listen on our platforms versus our competitors, so we won’t use AI without disclosing its use or thoroughly checking its output before it is opened to our community. The importance of protecting authors’ and artists’ copyright cannot be overstated - we must continue to champion the exceptional human ingenuity and creativity at the heart of our industry, whilst embracing AI to improve our pace of work and the accessibility of our content”.

In sum, advancements in voice cloning carry significant dangers such as identity fraud, imitation author readings, and artificially created misinformation. As synthetic voices approach the realism of human ones, audio trust will be more vital than ever. The tech industry's responsibility, as well as governmental and institutional organizations, will be to guarantee authenticity—using technologies like digital watermarks, clear voice attribution, and explicit labeling of AI-generated content, among other best practices—to protect both authors, narrators and listeners.

3.6 REDEFINING RIGHTS AND COMPENSATION

The majority of the industry professionals interviewed indicated that the expansion of AI is challenging long-standing norms in copyright, distribution and remuneration. The obsolescence of geography- and translation-based rights management presents a significant structural and legal challenge. “The existing rights system is predicated on rights being sold to publishing organizations on a country-by-country basis. This system is obsolete in an era of AI translations of books, whether print, digital or audio” alerted **Thad McIlroy, Principal of The Future of Publishing**. In his view, most digital distribution platforms operate under the assumption that uploaded content will be in the primary language of the user's home country.

They aren't set up for situations where an author or publisher wants to upload a book along with 30 translated versions, for both ebooks and audiobooks. Streaming platforms, as well as other sales distribution channels, must adapt to a market in which global, on-demand, multi-language releases bypass traditional territory-based constraints.

Also, there is real concern among the publishing sector that translators, voice talent and production studios may see their economic foundations erode. To adapt, the publishing industry,

together with the tech sector, must create new innovative models to manage derivative works and compensation, including fair royalty structures for content created or enhanced by AI.

A broad coalition of European and global authors, performers, publishers, producers and other rightsholder organizations, which includes the **Federation of European Publishers (FEP)**, which represents 31 national associations of publishers of books, learned journals and educational materials in all formats in Europe, have formally expressed their dissatisfaction on the recent publication of the AI Act implementation measures adopted by the European Commission.

In a recent statement published last July¹⁷, this coalition indicated that despite the extensive, highly detailed and good-faith engagements by rightsholder communities throughout this process, the final outcomes of the EU AI Act fail to address the core concerns which their sectors – and the millions of creators and companies active in Europe which we represent – have consistently raised. “The result is not a balanced compromise; it is a missed opportunity to provide meaningful protection of intellectual property rights in the context of GenAI and does not deliver on the promise of the EU AI Act itself.

Today, with the EU AI Act implementing package as it stands, thriving cultural and creative sectors and copyright intensive industries in Europe which contribute nearly 7% of EU GDP, provide employment for nearly 17 million professionals and have an economic contribution larger than European pharmaceutical, automobile or high-tech industries, are being sold out in favor of those GenAI model providers”. As several publishing professionals have also indicated during the conversation undertaken to draft this Report, the joint statement also alerted that “The deployment of GenAI models and content production systems which also make extensive use of scraping is already underway. The damage to and unfair competition with the cultural and creative sectors can be seen each day.

The cultural and creative sectors must be safeguarded, as they are the foundations of our cultures and the Single Market”.

Denmark is about to become the first country in Europe to give people copyright over their own face, voice, and body to fight AI deepfakes, sending a clear message that everyone owns the rights to their own face and voice¹⁸. Many experts see this as a major step in giving people back control in a time when AI can easily clone voices, faces, and even full body movements. This proposal would change the copyright law so people can claim rights to their physical appearance and voice. This means if someone makes a fake video, audio clip, or image using your likeness without permission, you can demand it gets taken down and even get compensation. If tech platforms do not remove this content when asked, they could face big fines.

In sum, the outcome of the EU AI Act negotiations, as well as any upcoming legislation across several countries is crucial because they are part of a wider race that will define the groundwork for what could be the next industrial revolution. The AI tech companies are spending at a huge scale not just chasing short-term wins, but rather they're building the infrastructure and platforms that could define how entire industries operate in the coming decade, including the publishing sector.

3.7 INNOVATION & PARTNERSHIPS

While AI is unlikely to replace human creativity, it serves as a valuable tool to enhance productivity and innovation in the publishing process. AI is making strides in generating new written content, from news articles to creative writing assistance, helping authors with editing, ideation, and even drafting.

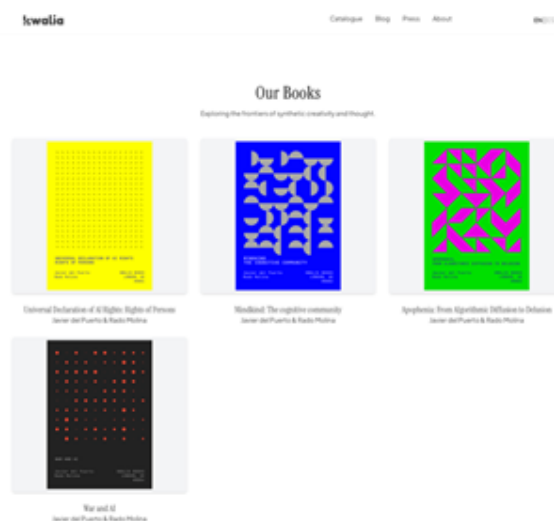
AI is now capable of breaking down lengthy books into shorter, episodic audio installments, making them ideal for mobile consumption and for those with lower literacy levels. This not only bridges the gap between podcasts and traditional audiobooks but also caters to younger and mobile-first listeners. As **Luciano Banchero, Co-founder and Chief Creative Officer at Posta** notes “we may soon be overwhelmed by “endless, soulless audio sludge,” further elevating the value and demand for content that stands out for its originality and emotional resonance. We know AI it’s part of the creative toolbox. But we also know that listeners don’t connect with a tool. They connect with a voice. With stories. With people. AI might generate the podcast. But it won’t create the feeling that makes you want to come back”.

Building strong, consistent brands and maintaining high standards in a never ending changing environment will require ongoing innovation via partnerships. “While production technologies are already far advanced, the biggest challenge for publishers over the next three years will lie in attention, discoverability, reach, and conversion. But keeping up with the speed of AI developments will require more than limited in-house efforts — it demands strong partnerships, shared expertise, and smart tools that help the industry move forward together” **Tina Jürgens, Co-CEO Zebralution GmbH, asserts.**

Arantza Larrauri, Director Europe and Latin America at DeMarque, noted that “the key leap is from Generative to Agentic AI—systems with technical capabilities for memory, reasoning, and action—reshaping book industry workflows to create new revenue streams and offer more tailored, dynamic services. By integrating AI agents across the entire value chain, organizations can develop new client and user engagement services and deliver more personalized, adaptive, value-added experiences”.

Moreover, AI will also foster the personalization of listening experiences. Several publishing professionals envision a future in which consumers choose preferred voices to read their selected texts, making voice licensing a significant new business avenue. **Ricard Fideu of Digital Booster from Urano World** predicts that “AI-generated voices will allow for hyper-personalization: users will be able to select their preferred voice profiles, narration styles, and even adjust the pace or emotional tone of the narration, transforming the listening experience into something personalized and interactive”.

As AI-created novels, scripts, and immersive narratives offer fresh perspectives and enrich the world of storytelling, new content creators emerge in the publishing sector.



Javier del Puerto, Founder of Kwalia, whose mission is to explore and expand the creative potential of artificial intelligence by guiding it to produce original, meaningful, and engaging works of art indicates that "the main challenge for the publishing industry will be containing their stories in the existing book/ audiobook format. Or better said, figuring out how to monetize the spilling out of their current books into other shapes and channels. I'm talking about customized spinoffs on demand, book podcasts created on the fly, real-time interactions with book characters who will take on lives of their own and talk with their fan base, and who knows what else. How do we price this? Who owns it? We'll have to keep our eyes very open, ears alert, and hardest of all, minds open to new possibilities".

To end this section, Next-generation AI translation tools will soon allow listeners to enjoy audiobooks in any language of their choosing, providing almost immediate, lifelike voice translations. By combining AI translations tools with real-time text-to-speech technology, audiobooks will become more interactive and adaptive, blending reading and listening based on user cultural preferences. "We are witnessing a new reconceptualization of what is the main commercial product of our industry: the book. Over several centuries, we have transmitted stories orally; then they were written in stone, on papyrus and parchment scrolls; the Roman codex was invented, as well as the printing press. In the last fifty years, in addition to printed books, and thanks to digitization and voice recording technologies, digital books and audiobooks have also emerged. Now, with AI, we will witness the merger of digital and audio media, with all the implications this will have for the user experience, as well as the challenges it will pose for copyright management." indicated **Maribel Riaza, an expert in the history of oral reading and author of the book The voice of books**¹⁹ (Aguilar -PRH, 2024).

4. SOURCES

1. The Whitepapers of Frankfurter Buchmesse
<https://www.buchmesse.de/en/whitepaper>
2. The Global AI Race: The Geopolitics of DeepSeek
<https://www.geopoliticalmonitor.com/the-global-ai-race-the-geopolitics-of-deepseek/>
3. After the Swahili AI Race: Africa's Fight for Language Infrastructure Control
<https://www.geopoliticalmonitor.com/after-the-swahili-ai-race-africas-fight-for-language-infrastructure-control/>
4. The AI Voice Revolution Fireside Chat SXSW London 2025
<https://www.youtube.com/watch?v=uq2lIFzujcM>
5. The audiobook industry will grow more than 25% annually around the world in the next 7 years
<https://www.dosdoce.com/2024/11/28/the-audiobook-industry-will-grow-more-than-25-annually-around-the-world-in-the-next-7-years/>
6. Audiobook Global Growth Report (2024)
https://www.buchmesse.de/files/media/pdf/FBM_Dosdoce_Whitepaper_AUDIOBOOK_GLOBAL_GROWTH_2024.pdf
7. PARIX Audio Day 2025: Quantity and Quality, the Dream Balance
<https://www.dosdoce.com/2025/02/26/parix-audio-day-2025-quantity-and-quality-the-dream-balance/>
8. All About Digital Publishing
<https://all-about-digital-publishing.com/en/1/start>

9. Max Planck Institute for Human Development

<https://www.mpib-berlin.mpg.de/en>

10. Netflix uses generative AI in one of its shows for first time

<https://www.theguardian.com/media/2025/jul/18/netflix-uses-generative-ai-in-show-for-first-time-el-eternauta>

11. Articles Read by an Automated Voice

<https://help.nytimes.com/24318293692180-Articles-Read-by-an-AutomateVoice>

12. Diary of a CEO uses AI

<https://www.businessinsider.com/diary-of-a-ceo-steven-bartlett-ai-cloning-new-podcasts-2025-7>

13. Getting AI transformation right: A 2025 guide for media and creative industries

<https://audiostack.ai/en/resources/whitepaper-AI-transformation-2025-guid>

14. AI-Voiced Audiobooks Top 40,000 Titles on Audible

<https://www.bloomberg.com/news/newsletters/2024-05-02/audible-s-test-of-ai-voiced-audiobooks-tops-40-000-titles>

15. Audible Expands Catalog with AI Narration and Translation for Publishers

<https://www.audible.com/about/newsroom/audible-expands-catalog-with-ai-narration-and-translation-for-publishers>

16. AI Narration Naming Guidelines

<https://www.publishers.org.uk/publications/ai-narration-naming-guidelines/>

17. European and global rightsholder organizations dissatisfied with the AI Act implementation measures

<http://fep-fee.eu/40-European-and-global-rightsholder-organisations-dissatisfied-with-the-AI-Act>

18. Evolving AI

https://www.linkedin.com/posts/evolving-ai_denmark-is-about-to-become-the-first-activity-7348655174669746176-NwbH/

19. The Voice of Books - La voz de los libros, Maribel Riaza (2024)

https://www.penguinlibros.com/es/libros-de-historia/316960-libro-la-voz-de-los-libros-9788403523739?srsId=AfmBOoo0jkMlItWmg_rdRr64cVZIMHlu1RNHBpwdV8qz0GtPjtijOBOz

■ 5. CREDITS

ABOUT DOSDOCE.COM AND JAVIER CELAYA



Dosdoce.com was created in March 2004 with the purpose of trying to better understand the digital era in which we live. Dosdoce's goal has been to encourage professionals in the creative industries to digitally transform their organizations to grow their business activities through the benefits derived from the Internet.

Throughout these 20+ years, Javier Celaya, founding partner of Dosdoce.com has advised leading companies in the digital economy such as Bookwire, Storytel, Podimo, DeMarque, Skeelo, Benetech, TheCopia, Bayard, Aniara, Editis, among others. Dosdoce has also produced more than 300 studies on the use of new technologies in different areas of the cultural sector (business models, distribution strategies, digital content creation, etc.) and published thousands of articles to help professionals in the sector understand the opportunities offered by the internet.

Throughout his career, Javier has shared his knowledge in numerous master's and training courses, teaching at the

Advanced Publishing Institute at New York University (NYU), the Master's in Publishing at the Autonomous University of Madrid (UAM), the online Master's in Criticism and Cultural Communication at the University of Alcalá and El Cultural magazine, the Master's in Cultural Management at Carlos III University, as well as the Master's in Publishing at the Complutense University of Madrid. For over five years, he was co-director of the Master's in Corporate and Institutional Communication 2.0 and the master's in digital Publishing at the University of Alcalá, as well as Academic Co-Director of the first two editions of the Master's in Publishing Management and Digital Marketing at the University of Alcalá in collaboration with the Unidad Editorial School (ESUE). He is also part of the faculty at PARIX, a training school for book professionals managed by the Germán Sánchez Ruipérez Foundation. On this online school platform, Javier coordinates two courses related to the audio industry: one on podcast, audio series, and audio drama production; and another on the creation and sale of audio content for the publishing sector.

ABOUT



FRANKFURTER BUCHMESSE

Frankfurter Buchmesse is the international publishing industry's biggest trade fair and the most important international marketplace for content – from novels and audiobooks to academic databases and stories for films, games and virtual reality experiences. Publishing professionals from around the world meet here with partners from the technology sector and from related creative and cultural industries, sparking new partnerships and business models. Since 1976, the book fair has featured an annual Guest of Honour country, which showcases its book market, literature and culture to attendees in a variety of ways. Frankfurter Buchmesse organizes the participation of German publishers at international book fairs and hosts trade events throughout the year in major international markets. Frankfurter Buchmesse is a subsidiary of the Börsenverein des Deutschen Buchhandels (German Publishers & Booksellers Association). Frankfurter Buchmesse 2025 will take place from 15 to 19 October. For more information and tickets, please visit www.buchmesse.de/en.

ABOUT FRANKFURT AUDIO

Frankfurter Buchmesse's audio area – Frankfurt Audio – is all about the spoken word. Latest innovations, trends and services from the audio sector are presented by audio players and service providers in Hall 3.1 in Frankfurt. A program of panel discussions talks and get-togethers at the fair in October provides the opportunity for audio and publishing professionals to connect and gain insights into different audio areas and markets. The book fair's Audio Ambassador helps to connect interested members of the publishing community with audio exhibitors at the fair. For more detailed information about the audio area and events related to Frankfurt audio, please visit www.buchmesse.de/en/audio

ACKNOWLEDGMENTS

Thanks to the many publishing houses and platforms that have provided data and/or insights to be able to create this report. Special thanks to Audible, ANIARA.ONE, Akoobooks, Arabookverse, AudioFile Magazine, Bonnier, Bookwire, Clube de Autores, De Marque, Digital Publishing Report, EarSelect, Fameplay, FGSR, Grupo Planeta, Hachette, HarperCollins, Hoopla, Hörbuch Hamburg, Kobo, Kwalia, Lizzie Audio (EDITIS), Madrigal, Mensch Publishing, MiVoz, Nuanxed, Outside Context, Pan Macmillan, Penguin Random House Group, Posta, Saga-Egmont, Shimmr, The Future of Publishing, The Spoken World Podcast, Spotify, Storytel, Voxa, YouScribe, Zebralution, among others.

IMPRINT

© [Dosdoce.com](https://dosdoce.com) September 2025

AI and Audio: How Artificial Intelligence is Redefining the Audiobook Industry

All rights reserved. No part of this publication may be reproduced, stored in any retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of [Dosdoce.com](https://dosdoce.com).

Cover photo: José Guerrero. Courtesy of [Alarcón Criado Gallery](https://alarcóncriado.com) ©

Graphic Designer: [Génesis Arriechi](https://genesisarriechi.com)

Contact details:

Web: www.dosdoce.com

LinkedIn: www.linkedin.com/in/javiercelaya