A Multimodal Analysis of The AI Podcast "Bot Canon"

Deepthi Suresh[1], Jaisakthivel T[2]
[1]*Research Scholar, Department of Journalism and Communication, University of Madras, Chennai*
[2]*Assistant Professor, Department of Journalism and Communication, University of Madras, Chennai*
doi.org/10.64643/IJIRTV12I10-196102-459

***Abstract*—Podcasting has rapidly become a preferred medium for digital influencers to establish deeper, more personal connections with their audiences. Tim Paige, an audiobook narrator and voice artist once said "A podcast is a great way to develop relationships with hard-to-reach people". Additionally, AI technology is increasingly used in podcast production to enhance audio quality, streamline editing, and even generate content, making the medium more dynamic and easier for the creators. This study conducts a multimodal analysis of ten episodes from the AI Podcast Bot Canon, examining how narrative style, episode structure, sound effects, context, and choice of words work together to shape meaning. By looking beyond language alone, the research highlights how these elements contribute to storytelling about artificial intelligence. In an interview Stephen Hawking said "A development of full artificial intelligence could spell the end of the human race" (BBC, 2014). The findings will provide critical insights into whether AI storytelling in podcasts is sufficiently advanced to replicate human nuances or still evolving as a technology. Additionally, the study will evaluate the prospects of AI potentially replacing human roles in podcasting, highlighting both its current capabilities and limitations.**

***Index Terms*—Artificial Intelligence, Podcast, Multimodal Analysis, Storytelling, Content.**

I. INTRODUCTION

Podcasting has been a part of the entertainment filed since the early 2000s. Ben Hammersley in the year 2004, coined the term "podcast" in The Guardian newspaper article (Eya Chafroud, 2020). Throughout the years since, podcasting has grown rapidly both in transformation and popularity. In the initial stages, podcasting was seen and used for timepass or a hobby where the sole purpose was just entertainment. But today it has evolved into a form of mainstream media with multiple purposes of educating, informing, entertaining and spreading awareness. Millions of platforms like Apple Podcasts, Spotify, and Google Podcasts are available to listen as well as to create podcasts.

As of recent statistics, there are around 4.19 million podcasts available globally and around 4.6 million listeners throughout the world. This number is expected to reach 5.4 million by the end of 2024 (Anil Agarwal, 2024). India has emerged as the third-largest market for podcast listeners globally (Press Trust of India, 2024).

Artificial Intelligence is being used in numerous different fields today and is slowly taking control over the digital world. The introduction of AI into podcasting has paved way of making enhanced and faster creation of content, voice and distribution. It also naturally excites the listeners and draws them towards it due to curiosity. AI tools such as Whisper AI help in creating automatic transcription by converting spoken language into text within seconds which may normally take hours to prepare manually (Abid Ali Awan, 2023)

Natural Language Processing (NLP) helps machines to generate, understand and interpret human language (Jim Holdsworth, 2024). Machine Learning (ML) algorithms study and analyze the given data to make predictions or decisions (IBM). In podcasting, ML is used for recommendation systems, ad targeting, and audience analytics.

Podcasting using AI has made the creation of content much easier and helps reach a wider range of audience. AI is altering the podcasting field by offering tools that improve audio quality and personalize user experiences. The usage of AI has not attained its full potential yet but is expected to establish a wider spectrum in the field of podcasting in the coming days. Today podcasting can be created by any layman without any technical support with the use of AI. From creating content, using voice to preparing the transcripts, everything can be managed by one single person in less time. Normally, podcast creation can be

time consuming, but artificial intelligence has made it simpler and way easier to create a podcast.

Artificial intelligence has not only reshaped how podcasts are produced and consumed but also paved the platform for future advancements in the medium. The emergence of AI in podcasting will further enrich the experience for both creators and audiences which in turn help in the growth and development of podcasting.

The AI podcast Bot Canon represents a unique case where AI-generated narratives are central to its episodes. While much scholarly work on podcasting has emphasized linguistic features, thematic concerns, or audience engagement, less attention has been given to the multimodal dimensions of AI-driven podcasts. Multimodality, in this context, refers to the interplay of elements such as language, sound, structure, and context that work together to create meaning beyond spoken words.

This study examines five episodes of Bot Canon selected on the basis of their popularity. Each episode is analysed using five key elements that is narrative style, episode structure, sound effects, context and choice of words to understand how AI constructs meaning within the podcast form. By focusing on these dimensions, the research aims to highlight the extent to which AI storytelling in podcasts can approximate human-like creativity and where it continues to reveal technological limitations.

II. REVIEW OF LITERATURE

Katja Grace (2018) says that Researchers forecast that in the coming decade, AI will surpass human capabilities in a range of activities. For instance, they predict that AI will be able to translate languages as effectively as humans by 2024, write high-school essays at least as well as students by 2026, and handle truck driving tasks more proficiently than human drivers by 2027. AI is expected to in retail jobs by the year 2031 and it would also become extremely capable of writing a bestselling book by the year 2049. In surgery artificial intelligence could become more skilled than human surgeons. In the next 45 years it is expected by tech experts that there is a 50% chance for AI to outsmart human beings in all tasks and within the span of 120 years artificial intelligence might take over all human jobs.

Benjamin Biering (2021) says in a podcast streaming service that offers podcasts in different languages, it's really important to make sure that every user gets to hear podcasts they'll enjoy, no matter what language they speak. Usually, we figure out which podcasts to recommend to people by looking at things like titles and descriptions. But now, because computers are getting better at understanding what's said in different languages, we can use what they understand from the spoken words to suggest even better podcasts to our users. So, we're not just looking at the text information about the podcasts; we're also listening to them to give you recommendations that you'll like, even if the podcast is in a language you don't speak.

Jemily Rime (2022) says AI can do more than just transcribe audio for podcasts. It can also be used to make podcasts that are really unique and interesting while still keeping things balanced. This might mean making podcasts that change based on what listeners like, adding more ways for listeners to interact with the podcast, creating special sounds for each listener, or coming up with new ways for people to enjoy podcasts. Earlier, converting audio recordings into text was not commonly enjoyed by podcasters because of the time it consumes and the difficulty of the process. They had to listen to the audio and type it out simultaneously which makes is harder to transcribe. Sometimes they can tend to mishear the words and type it out incorrectly. Some would prefer hiring professionals but even that would be expensive and not every podcaster could afford it.

Today because of AI technologies such as Whisper AI that can convert audio recordings into text automatically and accurately within seconds, podcasters find it easier to create their work without much expense and in less time. This means researchers don't have to spend as much time and effort on typing. It's important to know that these AI programs are not meant to replace humans completely, but to make the process faster and more accurate.

Now, there are advanced computer programs like Whisper AI that can change audio into text automatically, and they do it very accurately. This means researchers don't have to spend as much time and effort on typing. It's important to know that these AI programs are not meant to replace humans completely, but to make the process faster and more accurate.

III. OBJECTIVE

- To investigate how AI is used in the creation of podcasts, with specific reference to Bot Canon.
- To understand the role of multimodal elements in shaping meaning in AI-driven storytelling.
- To evaluate whether AI-generated content can replicate human creativity and narrative depth.
- To identify the strengths and limitations of AI in podcast production.
- To examine the implications of AI storytelling for the future of podcasting as a communication medium

IV. THEORETICAL FRAMEWORK

This study is guided by multimodal theory, which emphasizes that meaning is constructed not through language alone but through the interplay of multiple semiotic modes. As Kress and van Leeuwen (2001) argue, communication in contemporary media relies on the integration of linguistic, visual, auditory, and contextual resources, each contributing uniquely to how messages are produced and understood. In the context of podcasting, these resources include spoken language, sound design, narrative organization, and contextual framing all of which shape the listener's experience.

Podcasts, as an audio-centric medium, rely heavily on sound and narrative performance, yet they are inherently multimodal because they combine language, sound effects, pacing, tone, and contextual framing to produce meaning. Multimodal theory provides a lens through which these interdependent elements can be analysed systematically rather than being considered in isolation.

Applying this theoretical perspective to the AI-generated podcast *Bot Canon* allows for an examination of how narrative style, episode structure, sound effects, context, and choice of words interact constructing meaning. It also highlights the extent to which AI-driven storytelling can replicate or differ from human communicative practices. Thus, multimodal theory serves as the foundation for analysing the complexity of AI's role in shaping storytelling within digital audio formats.

V. METHODOLOGY

This study adopts a qualitative multimodal analysis to examine how artificial intelligence constructs meaning within the podcast *Bot Canon*. A qualitative approach is appropriate because the focus is not on measuring audience responses or quantifying data, but on interpreting how narrative strategies, sound design, and linguistic choices interact to produce meaning in AI-driven storytelling.

Five episodes of Bot Canon were selected for analysis based on their popularity (measured through platform visibility, listener engagement, and ratings). The selection ensures that the episodes studied represent content that has resonated most with audiences, thereby providing insights into how AI-generated narratives appeal to listeners.

Each episode was analysed through five multimodal elements:

1. Narrative Style – how the storyline is structured and presented, including coherence, pacing, and engagement strategies.
2. Episode Structure – the overall organization, such as introductions, transitions, and conclusions.
3. Sound Effects – the use of background sounds, pauses, or music to enhance meaning.
4. Context – the thematic framing of AI within cultural, technological, or ethical discussions.
5. Choice of Words – lexical patterns, tone, and linguistic markers that reflect AI's approach to storytelling.

VI. DATA ANALYSIS

The podcast Bot Canon, developed by Hannah Keefer in 2021, serves as the primary data source for this study. The podcast is unique in that its content is generated using the AI Dungeon platform, while narration is delivered through the TTS Maker text-to-speech tool. Each episode begins with a prompt constructed from a randomly selected set of 100 words drawn from literary sources, which are then expanded into full narratives by the AI system. This process makes the podcast an ideal site for studying how artificial intelligence constructs and performs storytelling.

For this research, five popular episodes were selected: Jurassic Park, Peter Pan, Harry Potter and the

Sorcerer's Stone, Pride and Prejudice, and Charlie and the Chocolate Factory. These episodes were chosen on the basis of their visibility and rating, ensuring that the analysis focused on content that had the greatest impact on listeners.

VII. NARRATIVE STYLE

In the AI-generated podcast Bot Canon, the narrative style is marked by a formal and structured tone that differs significantly from the more natural, conversational approach often found in human-hosted podcasts. The narration is consistently delivered in the first-person point of view, which gives the impression of a direct recounting of events by the protagonist. While this perspective offers clarity and immediacy, it also limits the range of viewpoints and interpretations, creating a somewhat restricted narrative lens.

The pacing of the story is generally slow and deliberate. This appears to be a by-product of both the AI's language generation process and the delivery through the TTS (text-to-speech) system. While the slow tempo aids comprehension, it may reduce narrative tension and listener engagement, particularly in action-driven episodes such as Jurassic Park or Harry Potter and the Sorcerer's Stone.

A notable feature of the narration is its voice quality. The pronunciation is clear and calm, with no stammering, filler sounds, or verbal hesitations. However, the AI struggles to produce variation in tone. It does not achieve husky or expressive vocal textures and instead maintains a largely neutral delivery throughout the narration. As a result, the storytelling lacks dramatic flair or emotional resonance, particularly in dialogue-heavy moments where human narrators might rely on tone shifts, emphasis, or pauses to convey character and mood.

Dialogue is the dominant narrative device, with characters frequently engaging in exchanges that move the plot forward. However, the emotional dimension of the dialogue is underdeveloped. Characters speak in ways that advance the storyline mechanically, but without the nuanced expressions of tone, hesitation, or subtext that human authors often embed in conversation. As a result, listeners may find it difficult to connect with the characters on an emotional level.

The complexity of the AI-generated plots adds another layer of challenge. The narratives often contain intricate or unexpected turns, which, while showcasing the creative potential of AI, sometimes make the stories difficult for listeners to follow. The absence of clear signposting or emotional cues can lead to confusion, particularly for those unfamiliar with the source texts being reinterpreted, such as Pride and Prejudice or Charlie and the Chocolate Factory.

Overall, the narrative style reflects both the strengths and shortcomings of AI-driven storytelling: it is coherent and formally structured but lacks the spontaneity, emotional depth, and intuitive pacing that characterize human storytelling.

VIII. EPISODE STRUCTURE

The episode structure of Bot Canon follows a relatively consistent framework across the selected episodes, reflecting the programmed design of the podcast. Each episode typically begins with a short introduction, where the AI-generated narrative sets the scene and introduces key characters or environments. Unlike human-hosted podcasts, there is little emphasis on a warm or engaging opening; instead, the episodes immediately move into the story world, creating a functional but somewhat abrupt start.

The body of the episodes unfolds in a linear fashion, but the organization of events often reveals discontinuities. For instance, in Jurassic Park, the plot transitions from setting descriptions to action sequences with minimal narrative bridges, giving the impression of loosely connected story fragments rather than a seamless progression. Similarly, in Pride and Prejudice, character interactions sometimes appear out of sequence, with sudden shifts in dialogue that challenge the coherence of the story flow. These structural gaps highlight the limitations of AI in maintaining traditional narrative arcs.

Episodes generally progress toward a climax, but the build-up is less deliberate compared to human-authored storytelling. The transitions between rising action, conflict, and resolution tend to be abrupt, with limited foreshadowing or narrative scaffolding. For example, in Harry Potter and the Sorcerer's Stone, key plot events appear suddenly, without the gradual development that would normally sustain suspense or emotional investment.

The conclusion of episodes is similarly abrupt, often ending when the AI has exhausted the narrative thread rather than reaching a purposeful resolution. In Charlie and the Chocolate Factory, the story halts mid-action,

leaving listeners with a sense of incompleteness. This structural limitation demonstrates that while AI can generate plotlines, it struggles with narrative closure feature essential to listener satisfaction in storytelling. Overall, the episode structure in Bot Canon reflects a pattern of functional organization rather than artistic design. The episodes are organized with a clear beginning, middle, and end, but the transitions and resolutions lack polish. This reveals how AI, while capable of sequencing story elements, often fails to replicate the nuanced structuring techniques that human storytellers employ to maintain coherence, suspense, and narrative satisfaction.

IX. SOUND EFFECTS

One of the most striking features of Bot Canon is its minimal reliance on sound effects. Unlike many contemporary podcasts that incorporate background music, ambient sounds, or transitional audio cues to enrich the listening experience, Bot Canon adopts a bare-bones auditory design. The episodes are almost entirely voice-driven, relying on the AI-generated narration without significant layers of sonic embellishment.

This absence of sound design produces both strengths and limitations. On the one hand, the lack of auditory distractions allows listeners to focus solely on the story being told. The formality and clarity of the TTS-generated voice become the primary mode of communication, emphasizing the words and structure of the AI-crafted narrative. On the other hand, this minimalism also strips the stories of emotional texture and atmosphere. For instance, in an action-driven episode like Jurassic Park, the absence of roaring dinosaurs, environmental sounds, or suspenseful background music diminishes the tension and drama that listeners would typically expect from such a narrative.

Transitions between narrative segments are also handled without audio cues. In human-produced podcasts, sound effects often serve as structural markers, signalling shifts in time, space, or perspective. In Bot Canon, the lack of these auditory signals can make shifts in the storyline feel sudden or disjointed. For example, in Pride and Prejudice, scene changes between conversations appear abruptly, without the subtle auditory cues that could otherwise prepare the listener for a transition.

Additionally, the voice quality of the TTS system itself becomes a kind of sound effect. Its mechanical rhythm, flat intonation, and lack of emotional inflection underscore the artificiality of the storytelling. In Charlie and the Chocolate Factory, for example, the whimsical tone of the original text is muted by the monotonous delivery, reducing the playfulness that characterizes the source material. Rather than supplementing the narrative with immersive sounds, the TTS voice both delivers the content and, unintentionally, highlights the limitations of AI-generated performance.

In sum, the soundscape of Bot Canon is marked by austerity and minimalism. The podcast forgoes the immersive qualities of rich sound design, instead foregrounding the words themselves. While this highlights the textual dimension of AI storytelling, it also reduces the emotional resonance and imaginative engagement of the listening experience, exposing a key limitation in AI-driven podcast production.

X. CONTEXT

The context of Bot Canon is central to understanding both its creative process and its reception. The podcast operates at the intersection of artificial intelligence, literature, and digital media, reimagining well-known cultural and literary works through an AI-generated lens. Each episode draws from a literary prompt of 100 randomly selected words, which anchors the story within a broader cultural text. By choosing iconic narratives such as Jurassic Park, Peter Pan, Harry Potter and the Sorcerer's Stone, Pride and Prejudice, and Charlie and the Chocolate Factory, the podcast situates itself within the cultural memory of its audience, creating a sense of familiarity even as the AI reshapes the original plots.

However, this act of recontextualization highlights both the creative potential and limitations of AI storytelling. On one hand, listeners are invited to compare their prior knowledge of these canonical works with the AI's reinterpretation, sparking curiosity about how machines "imagine" familiar worlds. On the other hand, the AI's lack of cultural sensitivity or narrative intention means that significant thematic or emotional layers from the original texts are often flattened. For example, the social critique embedded in Pride and Prejudice or the themes of childhood innocence in Peter Pan are not meaningfully

addressed, as the AI primarily recombines words and phrases without deeper awareness of context.

The podcast also operates within a broader technological and cultural discourse about AI. By using platforms like AI Dungeon for content generation and TTS Maker for narration, Bot Canon embodies the possibilities of machine-led creativity. This positions the podcast not just as entertainment, but as a commentary on the evolving role of AI in creative industries. When considered alongside public concerns such as Stephen Hawking's 2014 warning about the risks of AI the podcast reflects a cultural moment where fascination with artificial intelligence coexists with uncertainty about its implications.

In Pride and Prejudice, cultural sensitivity would mean recognizing the social norms, class dynamics, and gender roles of 19th-century England, and weaving them into the narrative in a way that reflects their importance. AI, however, often misses these nuances it might reproduce dialogue or plot fragments without capturing the deeper commentary on society that Austen intended. In Peter Pan, cultural sensitivity could involve handling themes of childhood, imagination, and colonial undertones (e.g., portrayals of the "Indians" in the original story) with care. An AI retelling may overlook these complexities and reproduce them mechanically or even inappropriately. In Charlie and the Chocolate Factory, cultural sensitivity would involve keeping the whimsical tone but also recognizing the moral lessons and social satire embedded in Dahl's work. AI, by contrast, might only reproduce the surface-level plot.

Overall, the context of Bot Canon reveals how AI storytelling is situated at the boundary between cultural heritage and technological innovation. By reworking well-known literary texts, the podcast demonstrates AI's capacity to produce novel narratives while also exposing its struggles with depth, coherence, and cultural sensitivity.

XI. CHOICE OF WORDS

In Bot Canon, the choice of words reflects the AI's reliance on structured, grammatically correct language generation. The podcast consistently uses formal phrasing, even when reimagining well-known stories that are originally playful, whimsical, or emotionally charged. This formality manifests in the way sentences are built they are always complete, syntactically correct, and neatly constructed, showing none of the irregularities or broken sentence patterns that often appear in natural human speech.

The vocabulary employed often includes straightforward, functional expressions such as "going to," "because," "all right," and "see you later." While these phrases contribute to a sense of clarity and accessibility, they also create a somewhat predictable and mechanical rhythm in the dialogue. The AI avoids colloquialisms, idioms, or region-specific variations that might otherwise add cultural or emotional depth to the narrative. A striking feature of the AI's language is the uniform delivery of words, with no emphasis or stress placed on particular terms or phrases. This lack of vocal stress means that dialogues do not highlight emotion, urgency, or irony, which are often conveyed through intonation in human speech. As a result, even moments of high tension such as encounters with dinosaurs in Jurassic Park or magical conflicts in Harry Potter and the Sorcerer's Stone are delivered in the same steady tone as calmer scenes in Pride and Prejudice or Charlie and the Chocolate Factory. This deliberate neutrality contributes to both strengths and weaknesses in the storytelling. On the one hand, it ensures clarity and avoids misunderstandings.

XII. FINDINGS

The multimodal analysis of five popular episodes of Bot Canon (Jurassic Park, Peter Pan, Harry Potter and the Sorcerer's Stone, Pride and Prejudice, and Charlie and the Chocolate Factory) reveals the strengths and limitations of AI-driven podcast storytelling across narrative style, episode structure, sound effects, context, and choice of words.

The narrative style of the podcast is formal, clear, and deliberate. The AI narrator maintains a calm, steady, and neutral tone throughout, free from stammering, hesitations, or filler sounds. While this ensures clarity and avoids distractions, it strips the narrative of the dramatic intonations, emotional highs, and subtle tonal shifts that human narrators naturally employ. As a result, suspenseful moments in Jurassic Park or whimsical passages in Peter Pan are delivered with the same measured pace as the social commentary in Pride and Prejudice. The heavy reliance on dialogue, coupled with a lack of expressive delivery, makes the stories comprehensible yet emotionally flat and, at times, monotonous.

In terms of episode structure, Bot Canon demonstrates consistency and predictability. Each episode follows a linear progression with a clear introduction, main narrative, and conclusion. This structure highlights the AI's ability to maintain coherence and flow, ensuring that listeners can easily follow the storyline. However, the rigid format also reduces narrative dynamism. Stories that originally thrive on surprise, suspense, or imaginative leaps—such as Harry Potter or Peter Pan—lose their vibrancy when confined to a mechanical structure. The AI's systematic ordering of events ensures neatness but reveals its inability to replicate the flexible and creative pacing often seen in human-driven storytelling.

The analysis also reveals the near absence of sound effects in the podcast. Unlike traditional podcasts, which use background music, ambient noise, or pauses to heighten immersion, Bot Canon features uninterrupted narration. There are no breaths, dramatic pauses, or filler sounds, which keeps the delivery clean and distraction-free. However, this minimalism also diminishes the immersive potential of the podcast. Intense scenes such as dinosaur chases in Jurassic Park or magical battles in Harry Potter lack auditory texture, making them feel as neutral and subdued as domestic scenes in Pride and Prejudice. The lack of sound design underscores the technical efficiency of the AI but also its detachment from emotional depth.

The context of the podcast situates it within the broader technological and cultural discourse on artificial intelligence. By retelling culturally significant works of literature, Bot Canon showcases AI's capacity to reproduce language accurately and sustain coherent narratives. Yet the analysis shows that cultural sensitivity remains a major limitation. Subtleties such as Jane Austen's social satire in Pride and Prejudice, the colonial undertones in Peter Pan, or the moral allegories in Charlie and the Chocolate Factory are flattened into literal retellings. This reflects the AI's inability to engage with cultural, historical, or symbolic dimensions of storytelling. While it succeeds in replicating form, it fails to capture meaning at deeper cultural levels. Finally, the choice of words is formal, precise, and consistent. The AI favours complete, grammatically correct sentences, using functional phrases such as "going to," "because," "all right," and "see you later." This ensures clarity and grammatical accuracy, but also introduces a mechanical rhythm. Unlike human storytellers, the AI places no emphasis on particular words or phrases, resulting in uniform delivery across both dramatic and mundane dialogues. The absence of idiomatic expressions, colloquialisms, and figurative language further contributes to a flattened emotional register. Thus, while the language is clean and error-free, it lacks stylistic richness and cultural texture.

Overall, the findings suggest that Bot Canon demonstrates AI's technical competence in podcast production, offering clear pronunciation, consistent structuring, and error-free language generation. However, it falls short of capturing the emotional, cultural, and stylistic nuances that make human storytelling engaging and immersive. The analysis highlights both the promise and the limitations of AI-generated storytelling, showing that while the technology is effective in form, it remains incomplete in meaning.

XIII. CONCLUSION

This study set out to examine how artificial intelligence engages with storytelling in the podcast Bot Canon through a multimodal analysis of five selected episodes: Jurassic Park, Peter Pan, Harry Potter and the Sorcerer's Stone, Pride and Prejudice, and Charlie and the Chocolate Factory. By analysing narrative style, episode structure, sound effects, context, and choice of words, the study highlights both the technical strengths and critical limitations of AI-driven podcasting. The findings show that AI demonstrates a high degree of technical proficiency. The narration is clear, precise, and free of common human imperfections such as filler sounds, hesitations, or mispronunciations. Episode structures are orderly and predictable, ensuring that narratives unfold in a coherent manner. Similarly, the choice of words reflects grammatical accuracy and consistency, with complete and syntactically correct sentences throughout. From a purely mechanical perspective, the AI succeeds in presenting stories that are intelligible, polished, and error-free, thereby showcasing its ability to function as a reliable content generator.

However, the analysis also uncovers significant shortcomings when the podcast is evaluated against the expectations of human storytelling. The narrative style, while clear, lacks expressive variation, emotional depth, and dramatic pacing. The absence of sound effects, background music, and tonal

modulation further flattens the storytelling experience, creating a delivery that is functional but not immersive. More critically, the AI struggles with cultural sensitivity and interpretive nuance. Literary works such as Pride and Prejudice or Charlie and the Chocolate Factory, which rely heavily on satire, symbolism, and cultural context, are reduced to surface-level retellings. The AI reproduces plot and dialogue but cannot engage with the deeper historical, cultural, or moral dimensions embedded in the original texts.

Taken together, these findings indicate that while AI is effective at replicating form, it remains inadequate in capturing meaning. The technological discourse around Bot Canon highlights the potential for AI to supplement or support podcast production, particularly in areas such as clarity, efficiency, and structure. Yet the cultural discourse surrounding storytelling emphasizes qualities—such as emotional resonance, creativity, and cultural awareness—that AI is still unable to replicate convincingly.

Therefore, this study concludes that AI-generated storytelling in podcasts is best understood as an emerging and experimental form, one that demonstrates the possibilities of automation but also exposes its limitations when compared to human creativity. Rather than replacing human podcasters, AI currently serves as a tool that can streamline production processes while leaving the most nuanced aspects of storytelling to human interpretation. Future research could extend this analysis by examining audience reception, comparing AI-generated podcasts across genres, or exploring hybrid approaches where AI and human inputs are combined. In sum, Bot Canon illustrates both the promise and the constraints of AI in podcasting. It represents a new frontier where technology intersects with narrative art, raising important questions about the future of storytelling, creativity, and cultural meaning in the digital age.

REFERENCE

[1] Balling, L. W., Townend, O., & Helmink, D. Widex Moment Sound Quality in Real Life–Not Just for Experts.

[2] Feng, T., & Narayanan, S. (2023). Foundation Model Assisted Automatic Speech Emotion Recognition: Transcribing, Annotating, and Augmenting. arXiv preprint arXiv:2309.08108.

[3] Grace, K., Salvatier, J., Dafoe, A., Zhang, B., & Evans, O. (2018). When will AI exceed human performance? Evidence from AI experts. Journal of Artificial Intelligence Research, 62, 729-754.

[4] Keller, M. A. (2024). Voices of Tomorrow: An Interdisciplinary Podcast Exploring Perspectives on Artificial Intelligence.

[5] Khanna, D., Bhushan, R., Goel, K., & Juneja, S. (2023). Summarizeai-summarization of the podcasts. Available at SSRN 4628657.

[6] Lau, T. T. (2019). When AI Becomes a Part of Our Daily Lives. environment, 6, 20.

[7] McHugh, S. (2022). The power of podcasting: Telling stories through sound. Columbia University Press.

[8] Pojoni ML, Dumani L, Schenkel R. (2023). Argument-Mining from Podcasts Using ChatGPT.

[9] Rime, J., Pike, C., & Collins, T. (2022). What is a podcast? Considering innovations in podcasting through the six-tensions framework. Convergence, 28(5), 1260-1282.

[10] Wang, S., Ning, Z., Truong, A., Dontcheva, M., Li, D., & Chilton, L. B. (2024, July). PodReels: Human-AI Co-Creation of Video Podcast Teasers. In Proceedings of the 2024 ACM Designing Interactive Systems Conference (pp. 958-974).