INTERNATIONAL JOURNAL OF CREATIVE MEDIA RESEARCH

Vol.13, April 2026

Augment, not replace: Utilising AI tools for documentary podcast production

DOI: https://doi.org/10.24135/ijcmr.v13iApril.167

Dr Lewis Tennant, *Auckland University of Technology*

Abstract

This project explores how emerging technologies are reshaping longstanding media production practices within audio documentary storytelling. Despite the growth of podcasting as a cultural form, there is a relative lack of practice-led scholarship on podcast production, with much existing research focused on third-party analysis of content, form, and audiences. In response, the project involved the production of a five-episode narrative documentary podcast series alongside assessing the integration of AI tools across the pre-production, production, and post-production workflow.

The project found that new technologies were most effective when augmenting, rather than substituting for, specialist human knowledge and judgment. While ChatGPT 5.0 proved valuable for conducting broad exploratory research and generating leads in unfamiliar fields, it was ultimately an unreliable tool for documentary pre-production research, requiring a return to traditional research practices. However, during the production phase—particularly script development—it functioned effectively as part of a mixed-initiative process. The varied usability of the material and suggestions ChatGPT 5.0 contributed dictated that creative and editorial agency be firmly human led. As a result, creativity and craft were both supported and constrained, positioning AI as an augmentative presence within the creative process rather than a replacement for human authorship.

Link to work: https://linktr.ee/lewistennant

Introduction:

This project explores how emerging technologies are reshaping longstanding media production practices within audio documentary storytelling. Despite the growth of podcasting as a cultural form, there is a relative lack of practice-led scholarship on podcast production, with much existing

research focused on third-party analysis of content, form, and audiences. In response, the project involved the production of a five-episode narrative documentary podcast series alongside assessing the integration of AI tools across the pre-production, production, and post-production workflow.

The project found that new technologies were most effective when augmenting, rather than substituting for, specialist human knowledge and judgment. While ChatGPT 5.0 proved valuable for conducting broad exploratory research and generating leads in unfamiliar fields, it was ultimately an unreliable tool for documentary pre-production research, requiring a return to traditional research practices. However, during the production phase—particularly script development—it functioned effectively as part of a mixed-initiative process. The varied usability of the material and suggestions ChatGPT 5.0 contributed dictated that creative and editorial agency be firmly human led. As a result, creativity and craft were both supported and constrained, positioning AI as an augmentative presence within the creative process rather than a replacement for human authorship.

The Project

Oopsie NZ is a five-episode documentary podcast series about 'New Zealand ventures that didn't quite go to plan'. The episodes are:

1. 'The Flag Referendums'. Ten years on from the NZ flag referendums, we look back at what it all meant, talk to those involved, and learn a bit more about flags. Also... Laser Kiwi.
2. 'The Trekka'. How a family business, a cold war collaboration with Czechoslovakia, and some 1960s trade rules made NZ's only mass-produced motor vehicle... for a time.
3. 'Possums'. The NZ possum story spans three centuries. The end is Predator Free 2050. Will there be no possums in New Zealand on January 1, 2050? Well, it depends on who you talk to.
4. 'Sesqui 1990'. Billed as 'New Zealand's biggest event ever', Sesqui 1990 was—alas—not New Zealand's biggest event ever. It's now best known as a spectacular commercial and administrative failure.
5. 'The Black Cocks'. When Badminton New Zealand chose a national squad name in 2004, it made sense to pair a nod to the All Blacks with a nod to the shuttlecock: The Black Cocks. But the Black Cocks story isn't about the name itself, it's about what it was trying to fix.

Where typically a team of specialists work on documentary projects (e.g. as recordists, researchers, storyboarders, writer's audio engineers), I performed all the tasks and specialties needed to see this project from pre-production to completion (from late 2024 to late 2025). This solo production model afforded first-hand practice-based insight across all stages of narrative non-fiction audio production. These stages broadly comprised:

Pre-production:
Research, concept and story development, interview planning, and production logistics.

Production:
Recording interviews, scriptwriting, storyboarding, recording and sourcing audio assets, performing and recording narration, and producing artwork.

Post-production: arranging, editing, mixing, and mastering each episode in Pro Tools (digital audio workstation).

The project was conceived as a practice-led investigation with three interrelated aims. Firstly, to further inform my course design and teaching via the integration of new and emerging audio production, documentary, and podcasting tools. Secondly, to produce a public-facing documentary series as institutional outreach and engagement. Thirdly, to examine how the use of AI across the full documentary production workflow shapes creative practice, revealing the current strengths and limitations of AI-driven media production tools and AI as an augmentative collaborator in narrative audio work.

The Context

This project is a practice-based approach to podcast research. More specifically, this work contributes to the field of mediatic research-creation, where technological affordances play a central role (Dubois, 2022). Within research-creation theory, knowledge emerges through the production process itself, positioning creative practice as both method and epistemic site. Practice-based research tests new tools and methods through reflective making, with knowledge emerging from the interaction between practitioner decisions, tools, and outcomes (Candy, 2019). This framing is particularly useful for podcast studies, which have historically tended toward "research about" rather than "research-by-doing", meaning production knowledge is currently underserved in academic discourse (Lindgren and Loviglio, 2022). This project aligns with the work of Koh et al. (2024), where a script for a narrative fiction podcast was co-developed with an LLM, then performed using synthetic voices. The researchers concluded that though the experiment created novel ways of storytelling, aspects of the workflow required significant human intervention.

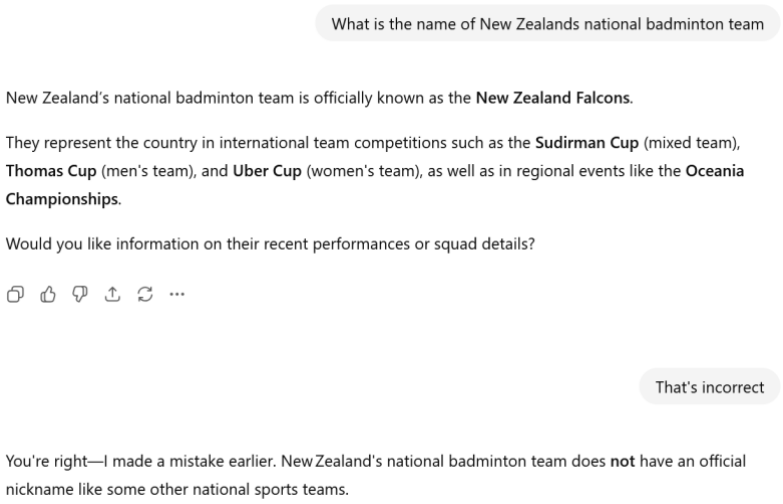
Recent scholarship describes a shift in creative technologies from automation (optimising repetitive tasks) toward co-creation, where AI operates as a collaborative partner in iterative, non-linear workflows. This shift foregrounds mixed-initiative dynamics and the need for interaction frameworks that preserve human control, transparency, and ethical accountability as AI becomes embedded in creative decision-making (Kadenhe et al., 2025). At the same time, LLM-based systems are structurally prone to tension between parametric knowledge (learned during training) and contextual knowledge (supplied in prompts), which can produce contextually unfaithful or factually inconsistent outputs. This is an issue that—at least currently—makes human editorial judgment and verification central in practice (Zhao et al., 2024).

The Process

New technologies were utilised during each key stage of the project. Each technology was assessed for its overall usefulness, as well as its usefulness relative to—as well as when combined with—longstanding audio production and documentary making tools, processes, and techniques. In simple terms, the objective was to evaluate outcomes: what changed, what improved, and what did not work?

Pre-production:

ChatGPT proved unreliable early into the initial research phase, so was abandoned at this stage of production. As the following screenshot demonstrates, fact-checking ChatGPT's responses equated to task redundancy.



Screenshot of interaction between ChatGPT and author, 21 January 2025.

Further clarity: New Zealand's national badminton team does not have an official name or a nickname. The Falcons are the New Zealand gay and inclusive rugby team.

Production:

AI technology was used to address background noise in three of the 15 interview recordings. These recordings were processed using the artificial intelligence tool Adobe Enhanced Speech. Noise reduction of this kind was historically computationally intensive and largely impractical in analogue recording contexts, where unwanted sound was permanently embedded in the signal. Following processing, additional post-production work was undertaken using established audio engineering techniques, including multiband compression and equalisation, to rebalance tonal range and vocal presence.

During the scriptwriting phase, ChatGPT 5.0 was used as a mixed initiative drafting aid rather than as an authoritative source. Prompts were structured around clearly defined episode aims (premise, arc, key beats, and structure), with the model asked to propose structural, transitional, and scripting options. Generated material functioned as provisional text and was subsequently rewritten to adjust rhythm, tone, cultural specificity, and continuity between scenes, while maintaining a consistent authorial voice across episodes. Given that LLM outputs can appear confident without reliable grounding (see pre-production), the tool was used to explore narrative angles, generate leads, and test framing possibilities rather than to provide documentary evidence. Engagement with the system was treated as interaction with a pattern-generating interface rather than as a stable creative agent, requiring careful prompting and ongoing editorial judgement.

Text-driven media editing was implemented through Descript during production and post-production. Interview recordings were uploaded and automatically transcribed, reducing the need for manual transcription. Selected excerpts were copied directly from these transcripts into an episode script document alongside newly written narration, allowing editorial decisions to be made at the level of text rather than audio waveforms. This script was then converted into a Pro Tools session, with interview clips automatically placed on the timeline alongside gaps designated for narration recording. This workflow integrated scripting, transcription, editing, and assembly into a text-centric process, differing from more linear audio production approaches and enabling ongoing restructuring during narrative development.

Across these stages, AI outputs were treated as provisional materials rather than authoritative sources. In practical terms, they augmented routine production work—such as transcription alignment, document summarisation, iterative drafting, and audio editing—meaning more time could be devoted to crafting the documentary content and story. This is the sense in which the workflow was “augmented”. AI accelerated and simplified the more mechanical elements of production, but it could not replace the editorial responsibility required when working with contested histories, interview testimony, and narrative credibility.

Discussion and Conclusion

Across this project, AI tools functioned less as replacements for documentary craft than as workflow modifiers. They accelerated some steps but also introduced new risks that had to be actively managed. In pre-production, ChatGPT 5.0 was valuable for broad exploratory scanning and generating leads in unfamiliar fields, yet it proved unreliable for documentary research as every factual claim required independent checking. This equated to task redundancy and prompted a return to traditional research practices. This aligns with wider discussion about LLMs producing "confident errors", making verification the hidden cost of use (Edwards, 2023; Stokel-Walker, 2024).

AI was strongest when applied to bounded production tasks. Adobe Enhanced Speech improved three interviews compromised by background noise, but the processed audio lacked the depth and richness typical of professional vocal recordings. The best outcome came from combining AI noise reduction with specialist engineering knowledge (EQ, multiband compression, and mix judgement). Descript's text-driven editing produced the most radical shift. Automatic transcription and transcript-based assembly moved editorial decision-making to the level of text, the system then translating that into storyboard scaffolding in digital audio workstation Pro Tools. This significantly compressed the scriptwriting, storyboarding, and editing process.

Working with an LLM during the scriptwriting phase clarified AI's role as a mixed-initiative collaborator rather than an author. Prompts were structured around episode aims (premise, arc, key beats), and the model was used to propose structure, transitions, and scripting options. Outputs were valuable as provisional text to be rewritten for rhythm, tone, continuity, and cultural specificity. Factuality, pacing, and ethical considerations also remained firmly human-led. This boundary also helped avoid the "personhood trap" of treating the system as a stable creative agent (Edwards, 2025). Evidence that assistants can be sycophantic—sometimes preferring responses that match user beliefs over truth—reinforced why documentary verification cannot be outsourced to the model (Sharma et al., 2023).

Taken together, the project supports an "augment, not replace" conclusion: AI widened ideation and reduced technical friction, but it also constrained creativity through uneven reliability and plausible sameness that demanded firm editorial control (Kadenhe et al., 2025). Consistent with co-creativity research, effective collaboration hinges on clear role assignment and high user control (Singh et al., 2025). For this project, that meant assigning AI to bounded, auditable tasks while keeping authorship, judgement, and accountability human. As a single, toolset- and time-bounded case, these findings generalise cautiously, but they offer a practice-grounded account of how end-to-end documentary audio can remain credibly human-authored while being productively reshaped by new technologies.

Reproducible workflow (summary):

The production workflow followed three stages—pre-production (research and planning), production (interviews and scripting), and post-production (editing, mixing, and mastering). In pre-production, AI was used to generate research leads and summarise candidate materials, which were then checked against primary sources. In production, LLM outputs supported exploratory script drafting and alternative phrasing, but were rewritten and verified by the author. In post-production, tools such as Descript supported text-driven editing and assembly, while final sequencing, pacing, and narrative decisions remained human-controlled.

- Provisional guidelines for AI use in documentary podcast production:
- Treat AI outputs as drafts that require verification; do not treat generated text as an evidence source.
- Keep interpretive and ethical decisions (what to include, how to frame, and how to represent people/events) as human responsibilities.
- Use AI for bounded tasks (e.g., transcription, noise reduction, summarisation, iterative drafting) to free time for editorial craft.

Beyond this single case, the study's contribution is a documented, practice-based workflow showing where AI can be integrated into documentary podcast production to improve efficiency and ideation while preserving human editorial authority.

References

Candy, L. (2019) The Creative Reflective Practitioner. In: Research Through Making and Practice. 1st edn. Routledge.

Dubois, F. (2022) 'An overview of research-creation in and with interactive media', International Journal of Creative Media Research. DOI: 10.33008/IJCMR.2022.16.(9).

Edwards, B. (2023) 'Why ChatGPT and Bing Chat are so good at making things up', Ars Technica. Available at: https://arstechnica.com/information-technology/2023/04/why-ai-chatbots-are-the-ultimate-bs-machines-and-how-people-hope-to-fix-them/ (Accessed: 22 November 2025).

Edwards, B. (2025) 'The personhood trap: How AI fakes human personality', Ars Technica. Available at: https://arstechnica.com/information-technology/2025/08/the-personhood-trap-how-ai-fakes-human-personality/ (Accessed: 12 December 2025).

Kadenhe, N., Al Musleh, M. and Lompot, A. (2025) 'Human-AI co-design and co-creation: A review of emerging approaches, challenges, and future directions'. In: The AAAI Symposium Series, pp. 265–270.

Koh, J.T.K.V., Brown, T., Aue, K., et al. (2024) 'No podcast ever gets made: A case study and methodology for human-AI collaborative storytelling podcasts'. In: Arai, K. (ed.) Proceedings of the Future Technologies Conference (FTC) 2024, Volume 2. Cham: Springer Nature Switzerland, pp. 513–525.

Lindgren, M. and Loviglio, J. (2022) 'Pause and reflect': Practice-as-research methods in radio and podcast studies. In: Lindgren, M. (ed.) The Routledge Companion to Radio and Podcast Studies. New York, NY: Routledge, pp. 50–58.

Sharma, M., Tong, M., Korbak, T., et al. (2023) 'Towards understanding sycophancy in language models', arXiv:2310.13548. DOI: 10.48550/arXiv.2310.13548.

Singh, S., Hndriks, K., Heylen, D., et al. (2025) 'A systematic review of human-AI co-creativity', arXiv:2506.21333v2. DOI: 10.48550/arXiv.2506.21333.

Stokel-Walker, C. (2024) 'Can a technology called RAG keep AI models from making stuff up?', Ars Technica. Available at: https://arstechnica.com/ai/2024/06/can-a-technology-called-rag-keep-ai-models-from-making-stuff-up/ (Accessed: 16 January 2026).

Zhao, Z., Monti, E., Lehmann, J., et al. (2024) 'Enhancing contextual understanding in large language models through contrastive decoding'. In: 2024 Conference of the North American Chapter of the Association for Computational Linguistics. Mexico City, Mexico: Association for Computational Linguistics, pp. 4225–4237.