

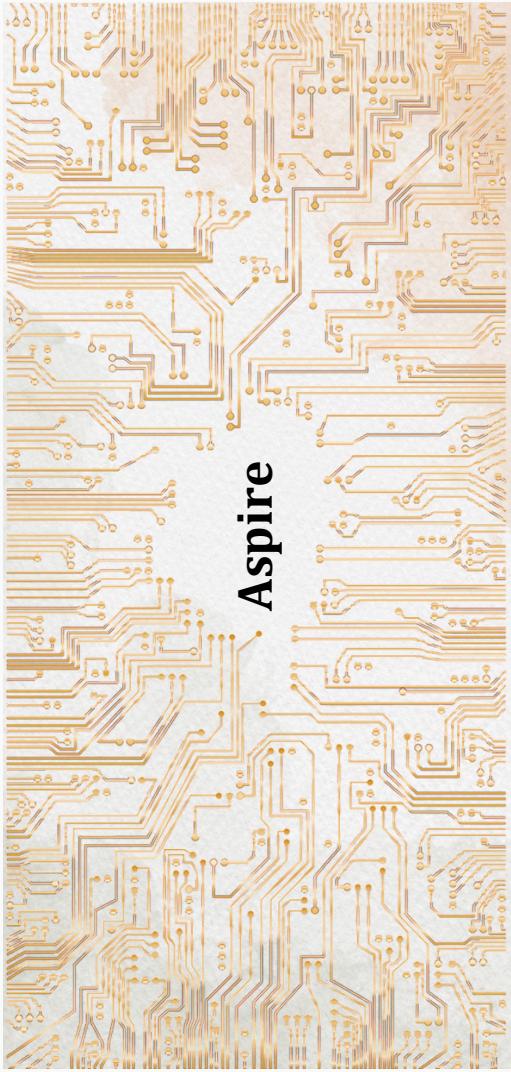


## A Critical Investigation of Spotify's Podcasters AI Translation Pilot

Case: No. 26-04-2024  
By Milica Hinic



Many thanks to Dr. Andrea Zeffiro and my fellow  
MA cohort in Course CMSTMM 720: Data  
Cultures for their care and support in the  
development of the Zine project.



3

"Artificial intelligence is not a substitute for human intelligence; it is a tool to amplify human creativity and ingenuity."

- Fei-Fei Li



About

I am an MA student in Communications and New Media, at McMaster University. My research interests include academic podcasting and knowledge mobilization. During my podcast research, I found accessibility in higher education teaching and learning to be a common topic.

Here, the agency of language and oral storytelling begins to take shape. Siobhán McHugh examines voice in today's contexts as being inclusive, diverse and equitable [1]. From this perspective, sharing international voices through podcasting breaks language barriers. Excellent!

**BUT WAIT!** There are some considerations to bear in mind when artificial intelligence (**AI**) is cloning your voice. For this reason, I am reporting on an AI-critical investigation of **Spotify's AI language translation Pilot**.

4



both the digital space and through an analog medium like this zine. For this reason, I have created Operation Aspire to generate **hope for balancing podcast power-knowledge structures**. Specifically, using voice(s) as data to **mobilize care** with sociotechnological futures.

### Mission

- Provide contextual data to Spotify users (podcasters and listeners)
- Identify the socio-technical design of artificial intelligence (AI) language translation. How it works!
- How data voice(s) are being synthesized and translated
- Future maintenance and care strategies

~ Let us Aspire.

## Operation Aspire

5

The **Zine**'s purpose is to provide podcasters and listeners with awareness about **voice(s) as data** and **AI language accessibility**. Operation Aspire reports on social impacts and why podcast infrastructural balance matters. The Zine is meant to reveal the hidden, misrepresented data and future possibilities. It is also a unique space to generate **multimodal conversations**. In other words, the zine demonstrates how knowledge sharing could happen in

both the digital space and through an analog medium like this zine. For this reason, I have created Operation Aspire to generate **hope for balancing podcast power-knowledge structures**. Specifically, using voice(s) as data to **mobilize care** with sociotechnological futures.



## Zine Playlist

7



8



## Case Summary

**Spotify** is a music, podcast, and audiobook streaming platform. The media organization is always looking for ways to improve its services for target audiences. With the recent advancements in AI technology (OpenAI's ChatGPT), Spotify developed an AI language translation app [2]. On SEPTEMBER 25, 2023, Spotify released a podcast pilot which would impact voices worldwide.

### What does this mean?

The pilot is like a test or trial where only a select few creator's episodes will be translated into Spanish, French and German. The translation involves artificial intelligence (AI) that will translate the podcasts using the podcaster's voice. [2]

Lex Freidman's podcast

Armchair Expert

The Diary of a CEO with Steven Barlett



Once audience and creator feedback is gathered and analyzed, the rollout will extend to other creator's podcasts.

**Voice Translations Hub:** You can view and listen to the AI episodes currently available on the Voice Translation Hub.



### Complaint Filed:

Hello, I have been a dedicated listener of Spotify. Why isn't my favourite podcast being translated into...

" Kiswahili " " العربية " " മുസ്ലീം " " српски " " 翻译中文 "

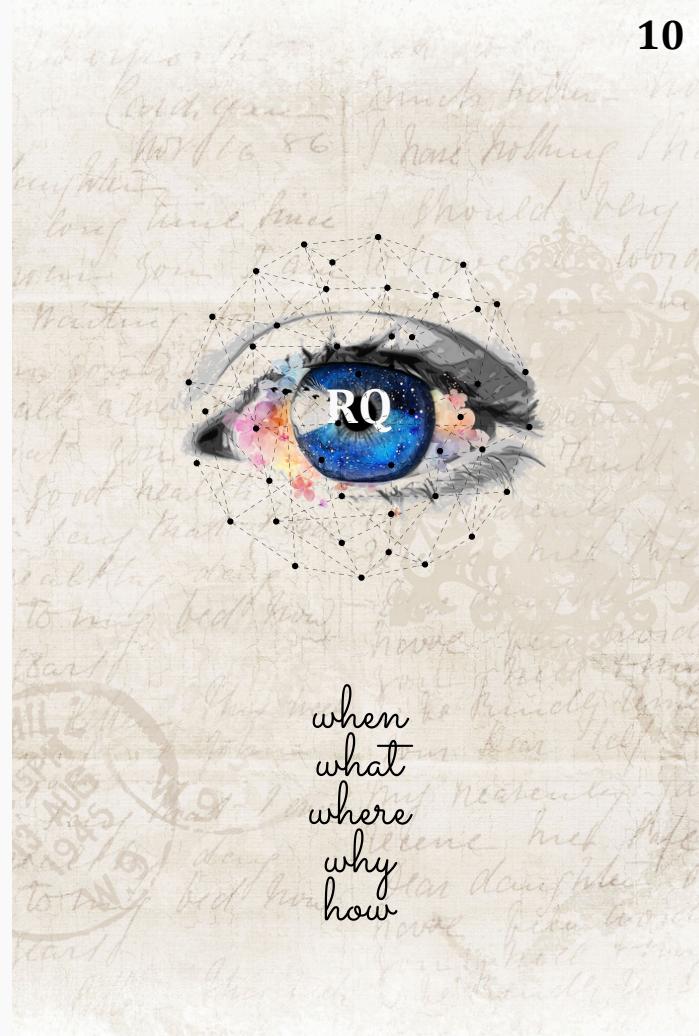
How does the AI voice replication work? Is it safe?

**Primary Investigator:** Milica Hinic  
**Type of Case:** AI language translation



9

## Spotify AI Translation Pilot



10

**11**

?

In what ways does Spotify's pilot project have sociotechnical influence?

How can we address some of the AI voice as data tensions?

Research Question(s)

**12**

**EVIDENCE AGAINST**

A research group at Roskilde University, Denmark provides a detailed account of how Spotify's AI translation works.

They argue that the model is **data-driven** and is based on **surveillance capitalism**. Here, the sociotechnical **patented design** seeks to **extract human behaviours** [3].

The software then further analyzes accents, replicates, translates, and transcribes language into sound/text. Does this mean Spotify will have control and power over these "voiceprints"? [3].

**Accessibility** is the company's goal, however, misrepresentation and misuse are a public concern. Who are the powerful players? Why are only certain voices being replicated? And who does this impact the most?

Top Beneficial Owners	Ordinary Shares Number	Percent
Daniel Ek	30,056,376	15.3%
Martin Lorentzon	21,476,345	10.9%
Balaji Gifford & Co	23,657,054	12.0%
Tencent	16,631,969	8.4%

## JUSTIA Patents

Identification of taste attributes from an audio signal

Patent number: 10891948

Abstract: A system, method and computer product are provided for processing audio signals. An audio signal of a voice and background noise is received by speech recognition is performed to identify the voice. There is a retrieval of media metadata corresponding to the speech content, and environmental metadata corresponding to the background noise. There is a determination of preferences for media content corresponding to the content metadata and the environmental metadata, and an output is provided corresponding to the preferences.

Type: Grant

Filed: February 21, 2018

Date of Patent: January 12, 2021

Assignee: SPOTIFY AB

Inventor: Stéphane Hulaud



### DEFENCE

Spotify defends its socio-technical design through the use of OpenAI technology and its approved patent.

**How it works:** The tool automatically matches the AI-generated voice to the speaker's tone and style. This means a podcast host can record an episode in English, and then the program can translate it into other languages in the podcaster's voice [4]. -> A genius!

Artificial intelligence is not new to our organization. Our AI is what makes our streaming platform a success. We have been using AI tools for **content personalization** [5].

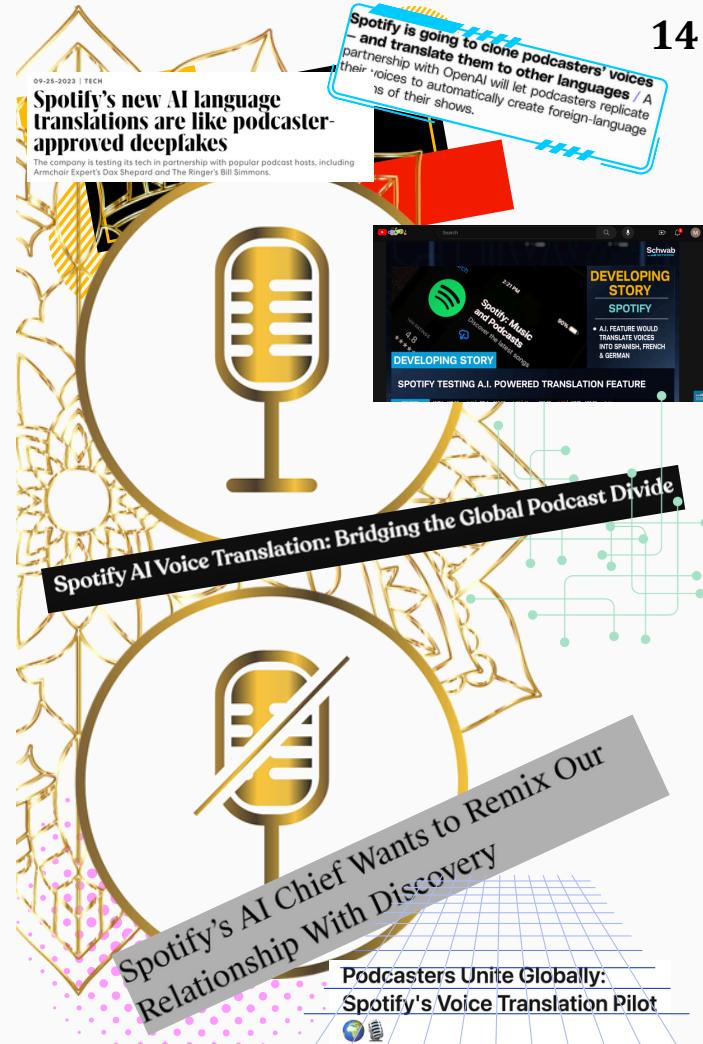
The AI translation feature provides listeners content in their preferred language without losing the original voice's authenticity. Who wouldn't want that? "It's like having a universal remote for Spotify podcasts"[6].

Spotify's **backend labour (AI + humans)** continues to innovate and enhance the listening experiences of our audiences.



13

## How AI Translation Works



14

15

### EVIDENCE AGAINST

Audra Simpson's concepts of **ethnographic refusal** discuss historical accounts of language translations being exchanged and categorized by militarized territory, ethnically defined, and primarily understood through the lens of the Colonized [7]. We question: Does AI translate these perceptions? Spotify's initial release was intended for Spanish, French and German. Do they continue to be the **voice of the dominant culture or Empire?** Or if we examine this concern from a capitalistic **framing** [3] Operation Aspire asks where are our voice data? Who has the power?

...  
"harvesting voice data from their users"  
( Bek et al.2021 )

### DEFENCE

The **user-generated data** prompts which language is most dominant based on listeners' demographics. We are only giving listeners what they want [4]. The AI-language translation tool identifies these needs and does its job, it translates.

Similarly, other companies are personalizing podcasts. A multilingual podcast production company **Studio Ochenta** based in Paris helps edit, consult and produce podcasts and/or audiobooks for creators [8]. Like us, they enhance the global listening experience and tailor content to specific **target audiences**. In this case, the model works in a similar way where Spanish and French are the dominant languages of translation. (Note: We cannot speak to AI use in their translations).

VOICE (s) as DATA  
Power-Knowledge

16



**WHAT NOW**

Operation Aspire has considered all the evidence of the **VISIBLE** and **invisible**

**What we do know**

AI language translation IMPACTS

**Universal assumptions**

In closing, the Spotify AI translation pilot is still under investigation as the data is further being collected, analyzed and sent out for processing through human and non-human means. In the meantime...

**17**

**HOW DO WE ENSURE SPOTIFY'S AI TRANSLATION is accessible, fair and safe?**

**Since the pilot was launched, there have been no research updates on the Spotify for Podcasters web page [9]. Also, a lack of Spotify AI translation reporting was found within mainstream media. Meanwhile, social media content creators are experimenting with language translation apps.**

**visible**

**Closing Statements**

## A collective practice

**18**

Like a community, this framework is built on sharing and working together. Here the **information maintainers** support bodies of information [10]. It is honest, it gives back, builds trust, and it is a collective that is isn't afraid of **not knowing**. The group includes a variety of folks - admins, teachers, engineers, technologists, storytellers, etc. They're goal is to continue maintenance from the backend of things and to the front facing public end.

## What do they do?

Instead of fixing things when they are broken, they continuously evaluate information systems and work flows. Guided by these principles, Operation Aspire shares some maintenance and care possibility for future Spotify AI translation practices.

"Care is a human trouble, but this does not make of care a human-only matter".

de la Bellacasa

**time** & patience for self, and others (including AI)

**Openness** to the many ways of knowing

**Care** with AI translation data engagements  
**Hope** for fair socio-technological futures

## AI Maintenance and Care

## References

- [1] McHugh, S. (2022). *The power of podcasting: Telling stories through sound*. Columbia University Press.
- [2] Spotify. (2023, September 25). Spotify's AI voice translation pilot means your favorite podcasters might be heard in your native language. For the Record. <https://newsroom.spotify.com/2023-09-25/ai-voice-translation-pilot-lex-fridman-dax-shepard-steven-bartlett/>
- [3] Bek, A., Gammelgaard, A., Mills, A., Thomsen, M. (2021). Voice Data, Power Structures and Technological Enframing A Critical Examination of Spotify as Irrational Regress. Roskilde University.[https://rucforsk.ruc.dk/ws/files/77901714/A\\_critical\\_examination\\_of\\_Spotify\\_\\_\\_final\\_project.pdf](https://rucforsk.ruc.dk/ws/files/77901714/A_critical_examination_of_Spotify___final_project.pdf)
- [4] Kaput, M. (2024, January 26). How Spotify uses AI (and what you can learn from it). Marketing AI Institute. <https://www.marketingainstitute.com/blog/spotify-artificial-intelligence#:~:text=Spotify's%20AI%20models%20recommend%20music,want%20to%20listen%20to%20next.>
- [5] Chan, L., Hogaboam, L., & Cao, R. (n.d.). AI in Media and Entertainment. In *Applied Artificial Intelligence in Business* (pp. 305–324). Springer International Publishing. [https://doi.org/10.1007/978-3-031-05740-3\\_20](https://doi.org/10.1007/978-3-031-05740-3_20)
- [6] FreeYourMusic. (n.d.). Spotify's AI Voice Translation Pilot. <https://freeyourmusic.com/blog/spotify-ai-voice-translation-pilot>
- [7] Simpson, A. (2007). On ethnographic refusal: Indigeneity, voice, and colonial citizenship. *Junctures* 9:67–78.
- [8] Ochenta Studio. (2023). About Paris Podcast Company. Studio Ochenta. <https://www.ochentastudio.com/en/about>
- [9] Spotify. (2023). Spotify for podcasters - research. Spotify for Podcasters -. <https://podcasters.spotify.com/resources/research>
- [10] The Information Maintainers, Olson, D., Meyerson, J., Parsons, M. A., Castro, J., Lassere, M., Wright, D. J., Arnold, H., Galvan, A. S., Hswe, P., Nowwiskie, B., Russell, A., Vinsel, L., & Acker, A. (2019). Information Maintenance as a Practice of Care. <https://doi.org/10.5281/zenodo.3251131>
- [11] de la Bellacasa, M. (2017). *Matters of Care*. University of Minnesota Press.
- [12] Edwards, B. (2023, September 27). Spotify tests using AI to automatically clone and translate podcast voices. *Ars Technica*. <https://arstechnica.com/information-technology/2023/09/spotify-tests-using-ai-to-automatically-clone-and-translate-podcast-voices/>
- [13] Foucault, M. (1978). Right of death and power over life. In *The History of Sexuality* (Vol 1). Pantheon. (pp. 135–159).
- [14] Kitchin, R. (2014). *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. London: Sage.
- [15] Ricaurte, P. (2019). Data epistemologies, the coloniality of power, and resistance. *Television & New Media*, 20(4), 350–365. <https://doi-org.libaccess.lib.mcmaster.ca/10.1177/1527476419831640>
- [16] Roush, T. (2023, September 26). Spotify will translate podcasts into other languages using AI. *Forbes*. <https://www.forbes.com/sites/tylerroush/2023/09/25/spotify-will-translate-podcasts-into-other-languages-using-ai/?sh=7b0784122ee>
- [17] Weber, S. (2008). "Visual images in research," *Handbook of the arts in qualitative research: Perspectives, methodologies, examples, and issues*. J.G. Knowles (Ed.). (pp. 42–55). Thousand Oaks, CA: SAGE Publications.
- [18] The Public. (n.d.). An introduction to: Zines. <https://thepublicstudio.ca/files/DIY-No2-Zines.pdf>
- [19] Livio, M., Pasek, Rayner, S. (2023). A Mostly Screen-Free, Zine-Full, Remote- Participation Conference on Experimental Methods for Research and Research Exchange. Experimental Methods and Media Lab The Low-Carbon Research Methods Group.

19

