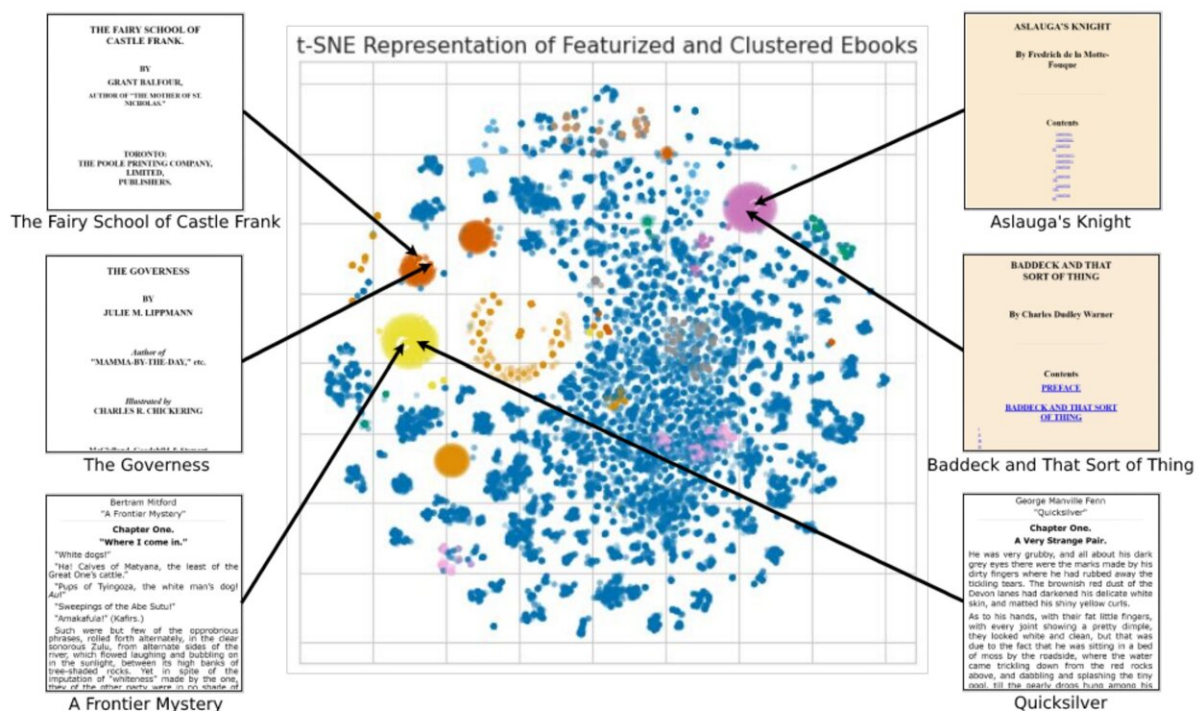


AI-generated speech brings a personal voice to books

September 20 2023, by Peter Grad



t-SNE Representation of Clustered Ebooks. Colored areas represent uniformly formatted clusters of books. Credit: *arXiv* (2023). DOI: 10.48550/arxiv.2309.03926

There's a bold new chapter in audiobooks.

Researchers from Microsoft, MIT and Project Gutenberg, which has

been hosting a digital archive of public domain literature since before the Internet, announced an initiative that brings natural-sounding AI-generated speech to books ranging from Randall Garrett's "After a Few Words" to "Zut and Other Parisians."

Automatic audiobook production is nothing new; it's been around for years. But the announcement of a new generation of audiobooks in the *arXiv* preprint "Large-Scale Automatic Audiobook Creation," details a new approach that generates a new dimension of realism with vocalizations powered by the latest generation of neural text-to-speech processes. It also shaves time and costs.

Current public-domain audio books suffer largely from robotic sounding narration. The new approach will generate narration with distinctive emotional nuance.

According to Microsoft's software engineer Brendan Walsh, "We use an automatic speaker and emotion-inference system to dynamically change the reading voice and tone based on context."

Narration is read in one voice while dialog by characters in the story is spoken in varying voices. The tone and style of speaking is determined by the neural inference system.

"This makes passages with multiple characters and emotional dialog more life-like and engaging," Walsh said.

Customers can adjust the sound of the voice, pitch, speed and intonation to their personal taste.

The researchers noted that they are preparing a live demonstration that will allow the public to generate an audiobook in their own voice. It will require only small samples of their voice that will be used to generate a

full volume.

The Wall Street Journal reported last April that DeepZen Ltd. has been using samples of the actor Edward Herrmann's voice for narrations of dozens of recent audiobooks. Interestingly, Herrmann died nearly a decade ago.

But with generative AI technology, samples of his voice were used to accurately construct a smooth dialog, complete with natural intonation, virtually indistinguishable from recordings of the late actor's actual voice.

Project Gutenberg has already posted about 5,000 books totaling 35,000 hours of speech online. Anyone can log on and listen, and the service is free.

They will soon offer the option to users to record their own books. Users will complete a voice profile by reading several sentences. Project Gutenberg will create an AI-generated voice that will be immediately available for users to listen to.

Users can recite a preface or dedication in their own voice, and then upload the complete text of their book. Customers will receive an email containing a link to their audiobook upon completion.

Soon, when mommy must work late and can't read a bedtime story to her 7-year-old son, he will need only call up his favorite audiobook and hear Mom's comforting [voice](#) bringing him tales of adventure.

Or aspiring actors can generate quick gifts for friends by sampling themselves for various roles in a Shakespearean play that brings characters alive with their own voices.

And, assuming legal cooperation with participating parties, who wouldn't jump at the opportunity to choose among the voices of Taylor Swift, Arnold Schwarzenegger or Morgan Freeman to narrate their own novel?

More information: Brendan Walsh et al, Large-Scale Automatic Audiobook Creation, *arXiv* (2023). [DOI: 10.48550/arxiv.2309.03926](https://doi.org/10.48550/arxiv.2309.03926)

Project page: [marhamilresearch4.blob.core.wi ... c/Website/index.html](https://marhamilresearch4.blob.core.windows.net/Website/index.html)

© 2023 Science X Network

Citation: AI-generated speech brings a personal voice to books (2023, September 20) retrieved 10 May 2024 from <https://techxplore.com/news/2023-09-ai-generated-speech-personal-voice.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--