227. Essays, podcasts and NotebookLM

I LISTEN TO quite a few podcasts. And given more time, knowledge, and certain skills that I almost certainly don't have, I have considered turning my own essays on Gaia science into audio/podcast form. The goal would be to make them more accessible, and a little less formal. And I have experimented on a couple of occasions: reading them aloud, or doing so after re-writing them in a different style. My attempts were rather disappointing!

NOTHING PREPARED me for the remarkable results of converting the pdf documents to audio format using AI. Specifically, I have converted the essays into a discussion-type dialogue between two avatars.

I will explain below what I have done, and the tools that I used, but let me first urge anyone curious to listen to one, perhaps taking a look at the corresponding pdf source first. So far, I have put essays 1–51 on a dedicated YouTube channel @GaiaEssays. Here are a few representative examples, linked directly:

- Essay 005: An Input Catalogue, or . . . (16:22)
- Essay 013: The distance to the Pleiades (12:04)
- Essay 024: Occultations of Europa and Triton (14:02)
- Essay 033: Nearby stars (19:42)
- Essay 035: Stellar flybys (16:03)
- Essay 037: Ultra-wide binaries (19:07)
- Essay 047: Iterative solution: implementation (15:28)
- Essay 050: The DIVA project (13:38)

THE SPEECH is rather fast, and the accent is American, so the resulting audio may not be so easy for non-native English speakers to follow. For the prototype software that I used, Google NotebookLM, it is not possible to change the delivery speed, the voices, or the accents. One of the reasons for using YouTube as host medium is that the listener can request (Google-generated) subtitles, and a slower playback speed (e.g. at 75%).[1]

LET ME SAY a little about the principles of AI underlying the creation of these audio files before being more specific on NotebookLM, and then commenting on the resulting accuracy and fidelity in my trials.

First, some background. The generative AI chatbot ChatGPT burst onto the scene at the end of 2022. Developed by OpenAI, it is a chatbot based on 'large language models' (LLM), such as GPT-4o. Considered to represent a significant breakthrough in the field of artificial intelligence, ChatGPT is a *pre-trained AI model designed to engage in natural language conversations, utilising sophisticated techniques from Natural Language Processing, Supervised Learning, and Reinforcement Learning to comprehend and generate text comparable to human-generated text'* (Roumeliotis & Tselikas, 2023).

In turn, Large Language Models are a type of machine-learning designed for natural language processing tasks such as language generation. These models acquire predictive power regarding grammar, syntax, semantics, and ontologies (definitions and relations) inherent in human language.

An accessible introduction to LLMs, and ChatGPT (GPT refers to 'generative pre-trained transformer'), is given in this YouTube video by Stephen Wolfram

Basically, LLMs model how text is generated. They extend text in a statistically meaningful way, essentially one word (or 'token') at a time. They are based on neural nets, and with their many parameters trained with self-supervised learning on a vast amount of text; Wolfram refers to 175 billion *training* parameters for ChatGPT.

In this 2023 video, he also refers to ChatGPT as making use of the results of `CommonCrawl`, which *'maintains a free, open repository of web crawl data that can be used by anyone'*. This www repository comprises the billions of web pages and millions of books and other documents that have been indexed.[2]

---

[1] Anyone wanting to take these concepts further could change the voices (even using one's own voice!) using other AI software which first transcribes the audio file into a (voice-labelled) text stream (e.g. `Descript`), then re-records the text using a library of available avatar voices (e.g. `ElevenLabs`).

[2] The 'indexed web' refers to the portion of the internet that search engines like Google have discovered, crawled, and stored in their databases, making it searchable. The number of such pages is updated daily at www.worldwidewebsize.com, and was estimated at around 50 billion at the start of 2025.

THE IMPACT of ChatGPT led to the development of several other LLMs, including Gemini developed by Google DeepMind (released in May 2023, and as Gemini 2.5 Pro Experimental in March 2025), and the Chinese DeepSeek (in early 2025). In contrast to previous LLMs, Gemini is 'multimodal', viz. not trained on text alone, but making use of multiple date types simultaneously, including text, images, audio, video, and computer code.

THIS BRINGS ME to Google's NotebookLM (where the LM refers to 'language model'). It was launched in 2023, with its 'audio overview' feature of interest here added in September 2024 (using Gemini 2.0). Google describes it as *a virtual research assistant that can summarise facts, explain complex ideas, and brainstorm new connections–all based on the sources you select*.

To create the audio version, I go to Google NotebookLM, upload an essay in its original pdf form, then in the right-hand panel, under 'Audio Interview/Deep Dive Conversation', click on 'Generate'. The two-host format conversation is generated within a couple of minutes.

Reaction to these products is generally along the lines of 'amazing', 'unnerving', and 'scary'. Even Stephen Wolfram, in the video I referred to above, described the results of these Large Language Models as being *remarkable that it works as well as it does*.

The fact that the output is generated from my self-contained essays is particularly relevant when aiming to create a factually reliable audio version. Indeed, I have found the results to be essentially 100% accurate, and remarkably well constructed. They include 'back-and-forth' banter, and sufficient informalities of the spoken language (interjections, hesitations, etc.) to be almost indistinguishable from human speech. They do not contain what AI/LLM jargon refers to as 'hallucinations'.

CURRENTLY, NotebookLM provides just two voices: one male, one female (named Jaimie and Jamie by Google). So until a wider choice becomes available, there will probably be a lot of similar-sounding audios appearing on YouTube, Spotify, and beyond.[3]

To be super-critical, some spoken mannerisms (including a liberal sprinkling of surprised reactions) can be a little repetitive. Also, the voices alternate between one being the expert and the other the interrogator. And the generator seems to have been program-loaded with a fair degree of 'wow'- and 'amazing'-type sentiment, which nevertheless works quite well in this context.

---

[3]For anyone wanting to experiment further (and in addition to the voice options that I mentioned in footnote 1), there are even the prospects of using the resulting audio file to create a *video* podcast, with self-chosen avatars. Here, generative AI web apps like `Hedra` allow for uploading audio and imagery, then mixing them to create a *video* that animates the image with the audio.

I AM HUGELY IMPRESSED with the audio files generated by NotebookLM from the essays that I've processed so far. They are comprehensive, accurate, didactic, and engaging, and they seem to fill in bits of background information to give the whole thing meaning. They even pick out key information from a table or figure.

If you are one of the Gaia 'insiders', and already know something of the history of astrometry (essay 3), the objects that are so important as Galactic tracers (essay 6), about on-board detection (essay 7), or the Pleiades distance controversy (essay 13), do listen – and prepare to be astonished! They far exceed anything that I could contemplate doing myself.

Multiple source documents can be loaded and processed. In the case where I have written something on one topic, then provided an update a year or two later, NotebookLM can create a useful merging of the two, taking into account some repeated preamble. I have tried this on my two essays on OB associations, #18 and #223, but I will probably not use this sort of merging.

THERE ARE OTHER FEATURES of NotebookLM, perhaps slightly less relevant in this context, but which seem quite powerful as general tools when assessing other complex source material. The first is the immediate creation of a short summary of the source. Here is the summary created from essay #5 (on source selection), retaining NotebookLM's use of capitalisation:

The provided text contrasts the approaches of the Hipparcos and Gaia space missions for observing stars. **Hipparcos** *required a meticulously prepared input catalogue of around 100 000 pre-selected stars due to technological limitations that allowed only one star to be observed at a time.* **Constructing this catalogue** *in the 1980s was a complex, international effort involving significant coordination and debate over which stars were most important to observe.* **Gaia**, *on the other hand, observes every object brighter than a specific magnitude as it scans the sky,* **eliminating the need for a pre-defined observing program** *and enabling the detection of transient events and a much larger number of stars.* **This shift** *in methodology was driven by technological advances and the lessons learned from the challenges faced during the Hipparcos mission.*

Other options in NotebookLM include creating FAQ sheets, briefing notes, and timelines from the material provided (all these are accessible from below the Audio Overview panel). For example, the 'timeline' option generates a chronological listing of the cited references, with their main results, and a list of the authors referenced.

In the central panel, below the chat summary, is another button labelled 'Mind Map'. This results in an interactive visual representation of the source content, with nodes that can be further expanded where relevant.

I PLAN to convert the rest of my essays, more than 200, to this audio discussion format, with the goal of publicising the achievements of Gaia more widely. I would appreciate wider feedback, via the YouTube channel.