

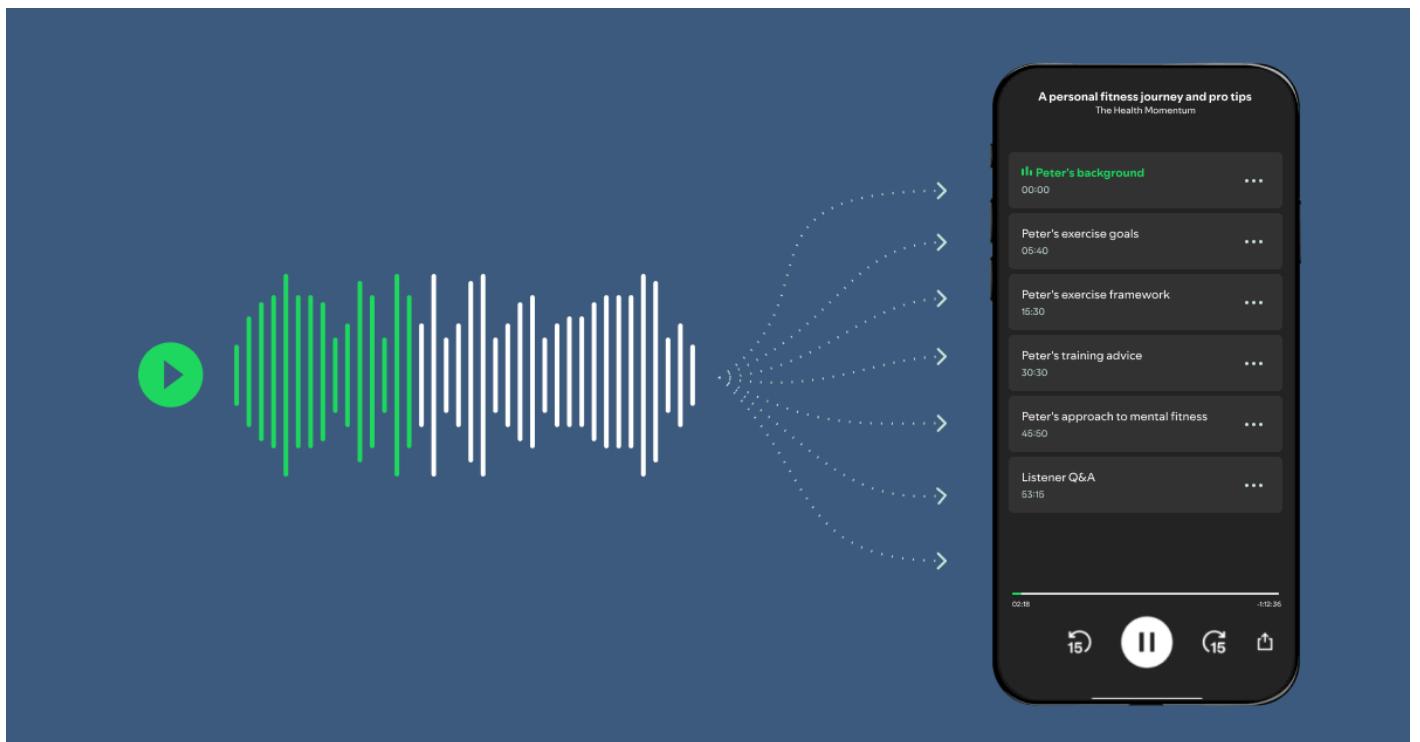


# PODTILE: Facilitating Podcast Episode Browsing with Auto-generated Chapters

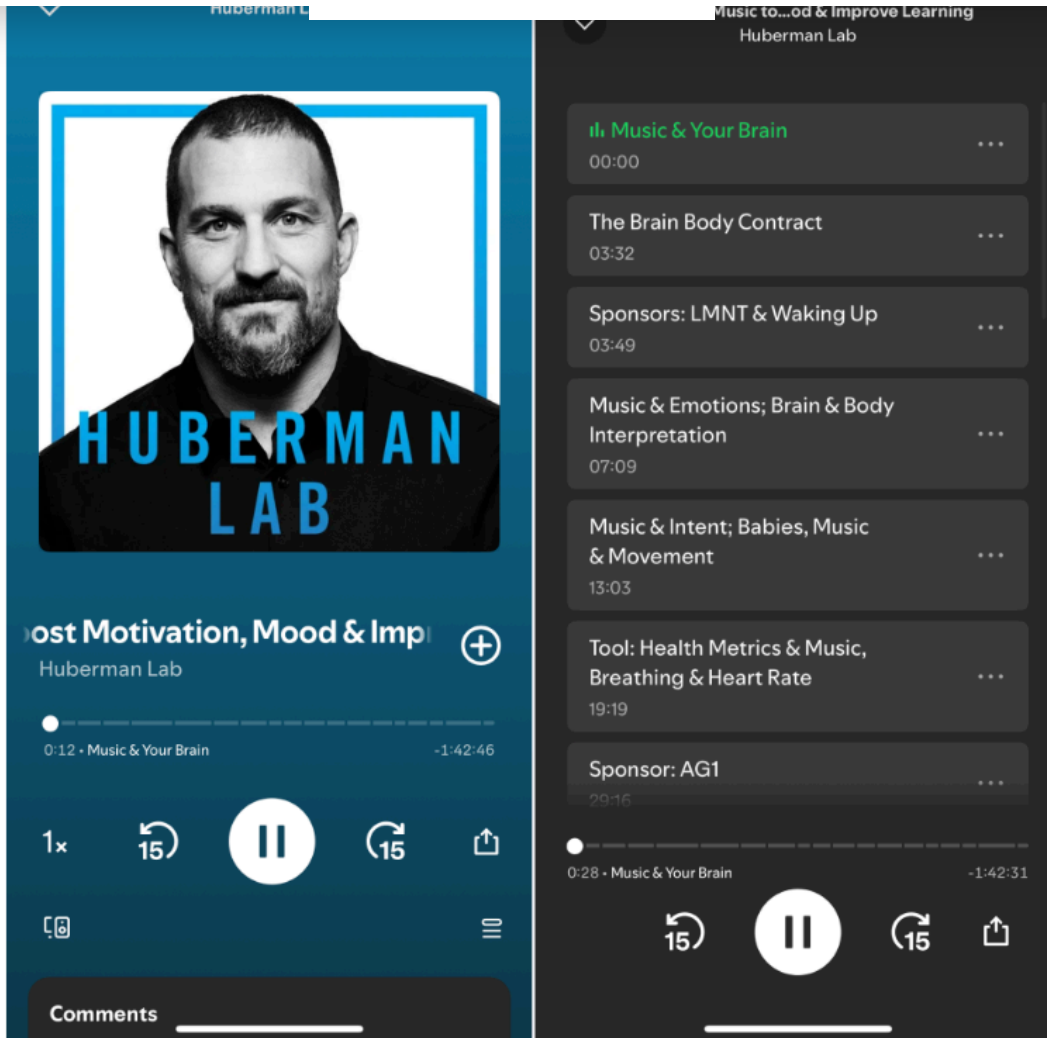
October 24, 2024



Published by Azin Ghazimatin and Ekaterina Garmash, Gustavo Penha, Kristen Sheets, Martin Achenbach, Oguz Semerci, Remi Galvez, Divya Narayanan, Ofeliya Kalaydzhyan, Ann Clifton, Paul N. Bennett, Claudia Hauff, Mounia Lalmas

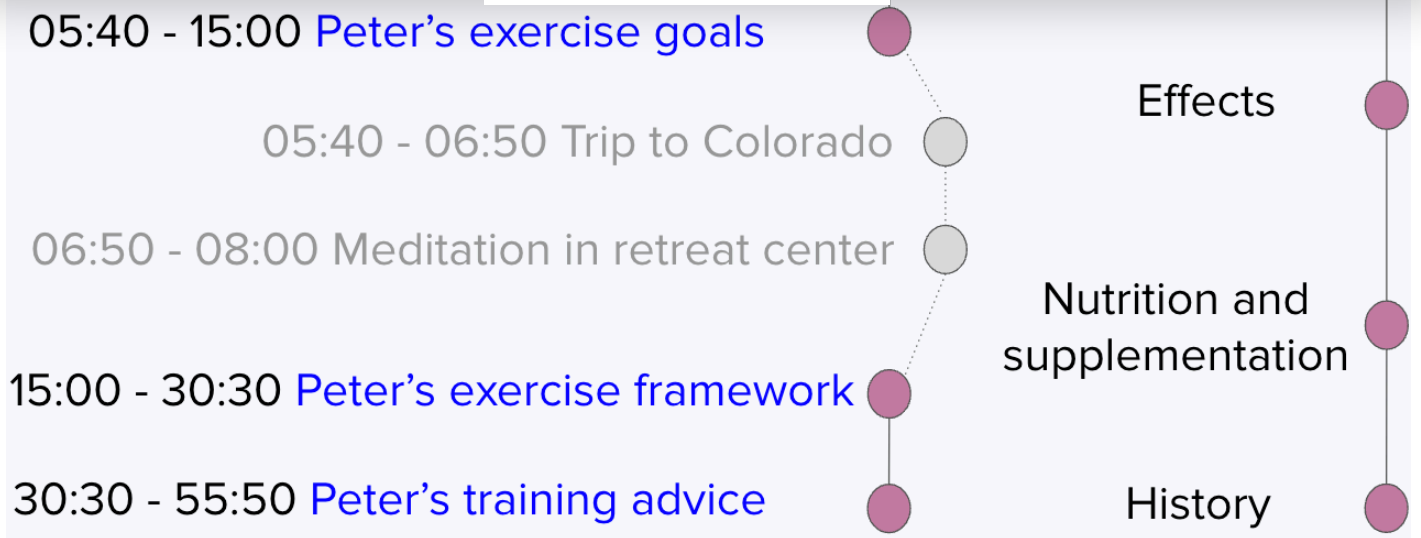


Listeners often find it challenging to navigate long podcast episodes due to their long duration. This makes it difficult for them to understand the overall structure and locate specific sections of interest. A useful tool to address this issue is podcast chapterization, where the content is divided into segments labeled with titles and timestamps. Although podcast creators can provide chapters with their episodes, this is rarely done.



*Examples of podcast chapters.*

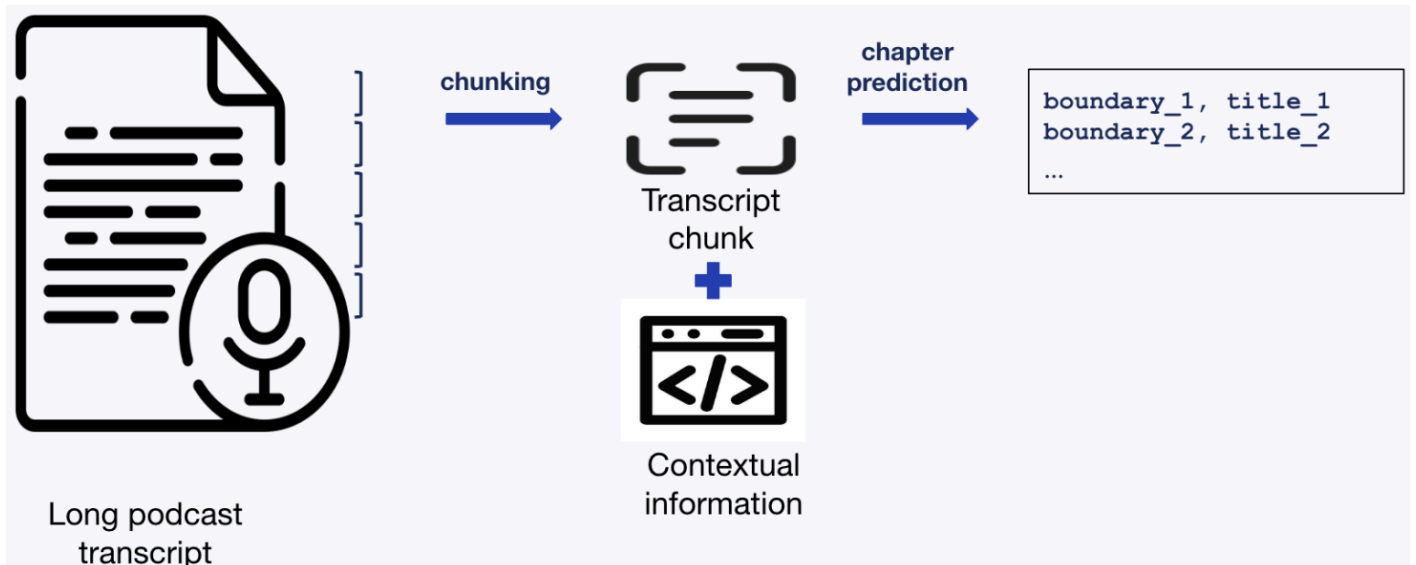
To extend the benefits of chapterization to more podcasts in our catalog, we have developed a machine learning-based chapterization model. This model is trained in a supervised way using creator-provided chapterizations of their podcast episodes. The podcast domain presents unique research challenges compared to previous work on chapterization and semantic segmentation. Podcasts are often conversational and lack a specific structure, with speakers sometimes diverging from the main topic for short periods. Additionally, episode transcripts are typically long, requiring efficient processing methods. In this post, we describe our solution for automating podcast chapterization that addresses these challenges.



*Comparison of typical podcast structure (left) and typical Wikipedia structure (right).*

## PODTILE

We employ large language models and develop PODTILE, an LLM-based model that simultaneously generates chapter boundaries and titles for the input transcript.



*Podcast chapterization with PODTILE.*

We use LongT5 with a 16k input token limit as our base LLM. This LLM is useful for capturing long-distance dependencies in podcast episodes. For transcripts longer than the 16K input limit, we split text into smaller chunks and process them



Predicting chapters based on local chunks can result in a loss of global context essential for accurate chapter prediction. To address this limitation, we enrich each chunk with **global contextual** cues to help preserve overall coherence. Specifically, we leverage **static context**, including metadata like episode titles and descriptions, and **dynamic context**, which maintains a record of previously generated chapter titles. This dynamic context acts as a working memory to inform the generation of future chapters. We include both types of information as text into the input to the PODTILE model.

Input:	
<b>The title of the document is:</b> A Personal Fitness Journey & Pro Tips	} <b>Static context</b>
<b>The description of the document is:</b> Today, we dive into an inspiring journey of physical transformation with Peter! ...	
<b>The previous chapter titles are:</b> Peter's background   Peter's exercise goals	} <b>Dynamic context</b>
The chunk text is: ... <b>S135:</b> Very ambitious goals! <b>S136:</b> That's right. <b>S137:</b> So, Can you walk us through your approach ...	} <b>Chunk text</b>
Output:	
S137 := Peter's exercise framework	

*PODTILE's input and output format. The static context contains the episode's title and description, and the dynamic context consists of the earlier chapter titles.*

## Evaluation and Findings

We evaluated PODTILE on an internal podcast dataset using title and boundary accuracy metrics commonly used in text segmentation tasks. Specifically, PODTILE demonstrated an 11% improvement in title accuracy compared to the strongest baseline. Moreover, we found that for very long podcasts, which required chunking



underscores the effectiveness of our modeling in capturing global (static and dynamic) context.

Needs chunking?	$\Delta \text{ROUGE}_{F1} \%$	$\Delta \text{SBERT}_{F1} \%$	% in Test data
NO	6.61	2.92	20
YES	12.08	5.51	80

*Comparison of the amount of improvement in long transcripts vs. shorter ones that do not need chunking.*

We also conducted a qualitative comparison of chapter titles generated by PODTILE with those from the baseline. We found that, by leveraging static and dynamic context, PODTILE's titles are more informative.

PODTILE	Ground-truth	Baseline
Sigma MC-11 Adapter	Gear of the Week: Sigma's MC-11	Sigma Mount Converter
Planet of Lana Reviews	Planet of Lana Reviews	Lana Review
Sandeep's Family Background	About Sandeep's Background	Birth

*Comparison of chapter titles generated by PODTILE against those from the baseline. PODTILE's chapter titles are more informative.*

In April 2024, we began a limited roll-out of our chapterization model.

We anticipated that by broadening the availability of chapters, high-quality auto-generated chapters will lead to an increase in engagement. In fact, we observed an 88.12% increase in chapter-initiated plays in the first month of the roll-out.

We also conducted an experiment to assess the impact of indexing chapter titles on search effectiveness. Using the TREC podcast dataset, designed for short segment



descriptions enriched with chapter titles. The results showed a 24% increase in R@50 when chapter titles were included in the episode description, demonstrating that chapter titles effectively summarize transcripts, which then enhances retrieval effectiveness.

## Conclusions

We introduced PODTILE, a solution for supervised podcast chapterization that effectively models the global context of the episode. PODTILE addresses the challenges of extreme length, long-distance dependencies, and low structuredness. PODTILE outperforms state-of-the-art baselines in offline evaluation, with particularly notable improvement for extremely long podcasts. The deployed solution has substantially increased our catalog coverage, and analysis of user interaction data highlighted its value for less popularity shows. Additionally, we evaluated PODTILE's usefulness in other downstream tasks: in offline evaluation, we showed that adding chapters to episode descriptions increases episode search quality.

For more information, please refer to our paper:

[PODTILE: Facilitating Podcast Episode Browsing with Auto-generated Chapters](#)

Azin Ghazimatin and Ekaterina Garmash, Gustavo Penha, Kristen Sheets, Martin Achenbach, Oguz Semerci, Remi Galvez, Divya Narayanan, Ofeliya Kalaydzhyan, Ann Clifton, Paul N. Bennett, Claudia Hauff, Mounia Lalmas

CIKM 2024

---

SHARE



---

CATEGORIES



## Related articles

**Personalizing Audiobooks and Podcasts with graph-based models**



## Accelerating Creator Audience Building through Centralized Exploration

### Sign up for research updates

By clicking sign up you'll receive occasional emails from Spotify. You always have the choice to adjust your interest settings or unsubscribe.

Your Email

Sign Up







[Newsroom](#) [Spotify Jobs](#) [Spotify.com](#)

[Spotify R&D Engineering](#) [Spotify R&D Design](#)

[Legal](#) [Privacy](#) [Cookies](#) [About Ads](#)

© 2024 Spotify AB