

AIMC 2024 (09/09 - 11/09)

Endless Generative AI: A Practical Tutorial

Marco Amerotti Bob L. T. Sturm

Published on: Aug 24, 2024

URL: <https://aimc2024.pubpub.org/pub/p9tx21c3>

License: [Creative Commons Attribution 4.0 International License \(CC-BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

[KSPR-Workshop-slides.pdf](#)

4 MB

Endless Generative AI: A Practical Tutorial

Description

Generative AI is gaining more and more importance and power. From generative text models and image models to generative music and audio models, one can now synthesize artificial content that is difficult to distinguish from human-created content. Moreover, many such generative systems are openly available: the weights of very large models that require a great amount of power and data to train — such as Mistral or Stable Diffusion — are now easily downloadable and ready to use with an almost plug-and-play approach. How can we creatively exploit this new toolkit of generative models? What can be built with it, and how? What ethical predicaments surface in the generation and distribution of real-time endless generative content?

In this tutorial we target the case of endless artificial entertainment: we discuss creating an endless radio station entirely generated by AI via the orchestration of many different generative models publicly available online. We present our ideation and implementation process, alongside the challenges and issues we faced.

The workshop will last 2 hours, with a short 15-minute break. In the first 50 minutes, we present the radio and the techniques we use. In particular, we present an overview of our pipeline and what models are involved, focusing on generative models for text, TTS models, and generative models for music and audio, as well as the reasons behind these choices and what limitations they introduce.

After a short break, in the next 50 minutes, we analyze the generation pipeline and how each model is inferred and interacts with the others. In particular, we target:

- how to prompt each language model and how we address the incoherencies and issues with the generated material;
- how to create realistic speech from TTS models, and how to adapt the generated script to obtain specific effects (e.g. laughter, mumbling, etc)
- how to generate music mixes that are coherent in genre and style;

We finally present how all generated material is assembled and streamed online, and discuss ethical vulnerabilities and implications of such an exercise. Throughout the entire tutorial, we provide code examples from our codebase, as well as samples of the finite product and its intermediate stages.

At the end of the workshop, the participant

- will have gained insight and knowledge into the state of the art of generative AI;

- will be able to replicate our process with our code examples;
- will be able to assemble similar AI-powered pipelines to generate artificial entertainment;
- will have an overview of the problems and difficulties that can arise in the process and how they can be targeted.

Short Description

Generative AI has made huge advances in recent years, giving us a variety of publicly available models for text, speech, images, and audio. What can be built using this new toolkit of generative models? In this workshop, we present our approach to creating an endless radio station, completely generated by AI, alongside comprehensive code examples.

Organizers

Marco Amerotti is a research assistant at KTH, carrying out research in generative models for symbolic music and real-time interactive performance modeling. Bob L. T. Sturm is principal investigator of the [MUSAiC](#) project, funded by the European Research Council under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 864189). They are both responsible for the creation of the overall radio pipeline and maintenance of the codebase.

Preferred Length of Workshop

Two hours.

Technical and Space Requirements

The workshop will be conducted in person, with the option for online streaming. The workshop is thought for an audience of 20/30 *in situ* participants, with no limitations online.

A projector and a connected sound system are required.

Links to Supporting Media

The radio created by the approaches and techniques of this workshop is publicly available at the following link: <https://www.youtube.com/@KSPRStochasticPirateRadio>

Ethics Statement

The preparations of this workshop have been carried out in the context of the [MUSAiC](#) project, funded by the European Research Council under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 864189). The authors do not have any conflict of interest with any of the services and companies mentioned.

The ethics and sustainability issues connected with the continuous production and streaming of AI-generated content are explicitly addressed during the workshop. See the attendant paper at the conference, Sturm et al. “Stochastic Pirate Radio (KSPR): Generative AI applied to simulate commercial radio”, in Proc. AIMC 2024.