

**When filling in this template please keep the font, size, formatting and line spacing already selected in this file**

## **An Oracle of Sounds**

### **Sampling potential sonic futures using text-sound multimodal models**

#### Abstract

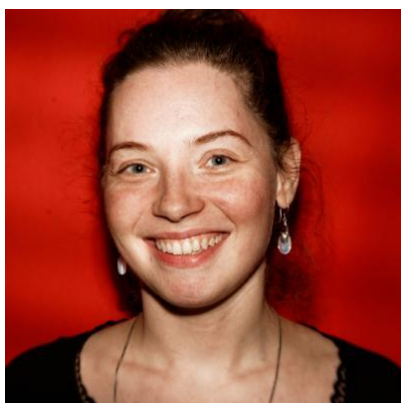
During the past few years, text-sound multimodal machine learning models have become increasingly widespread. These models can generate a new sound given a text prompt, or an audio file. Given an audio prompt, the model aims to, either suggest a plausible continuation of the input sound, or provide a sound that is similar to the input sound. These models are based on very deep architectures that require vast amounts of data and significant computational power for their training. The state of the art models of this type are commercial models. Open source text-sound multimodal models are research prototypes usually trained on generic music and audio in the public domain (for example Meta's MusicGen or AudioLDM2). These research models typically generate around 10 seconds of audio per prompt, Within this context I have been researching the following questions.

In which ways do the nonlinearities inherent in the algorithmic generative process (also called inference or prediction) of text-sound multimodal models influence time in the compositional process? And, as these algorithms sample from plausible futures, what can the artistic process of imagining these potential sonic futures be? Moreover, what are the challenges and methods of aesthetic negotiation that emerge as the artist attempts to assert aesthetic agency over these black boxes? And lastly, where does the issue of interpretability come to play in the iterative process of model prompting?

I will describe the ways I have been exploring these questions in my research, prompting both with text and with my own sound inputs. I will also present a set of criteria and methods for artist-model aesthetic negotiation that have emerged through my practice and I will share methods of co-creation and of reclaiming agency within my artistic process.

- For masters and doctoral students: I wish for this session to be an individual presentation and not a student-supervisor dialogue.

Research Presentations



**Zoi Efstathiou**  
**Norwegian Academy of Music (Oslo, Norway)**  
**zoie@nmh.no**

Zoi (Zoe) Efstathiou, pianist and electroacoustic music composer, originally from Greece, is known for her work on extended techniques and on expanding the timbral possibilities of the piano.

She holds a Master in Improvisation from the Academy of Music and Drama at the University of Gothenburg and an Advanced Postgraduate Diploma (Soloist) in Composition from the Rhythmic Music Conservatory in Copenhagen.

She also holds a Bachelor degree in Mathematics from Lund University, awarded with distinction, that includes advanced courses in machine learning.

Her music has been performed by her or others at Nordic Music Days Festival, A L'Arme Festival, Onassis Cultural Center, Levande Musik, Only Connect, Copenhagen Jazz Festival, Monopiano Festival, Copenhagen Piano Festival, Samtida Musik, Alice, The Nordic Embassies Berlin, Njord Festival, among others.

Currently, she is a Research Fellow in Artistic Research at the Norwegian Academy of Music.