

Weaving·memory·matter

Steerability and embodiment of
latent audio models through
interactive machine learning

Gabriel Vigliensoni
gabriel.vigliensoni@concordia.ca



Ethical and Responsible AI Music Making Workshop
CCI/UAL, London
17 July, 2024

Problems and questions

Steerability and embodiment of generative audio models

- How to navigate unlabelled, high-dimensional latent spaces?
- How to perform real-time interaction with latent models?
- How to provide long-term coherence?
 - * Long-term coherence is important in linear media and music
 - * This is still missing in SOTA generative music systems

Neural audio synthesis

Architectures

- Jukebox by OpenAI (2020)
- DDSP by Google Magenta (2020)
- RAVE by Caillon and Esling (2021)
 - nn~ (Max/Pd)
 - *Tasks:* timbre transfer and latent manipulation

Neural audio models

Datasets

Different music types:

Acoustic-harmonic (~32h)

Acoustic-percussive (~19h)

Electro-harmonic (~22h)

Electro-percussive (~15h)

Aventures Sonores (~5h)

vigliensi (~8h)

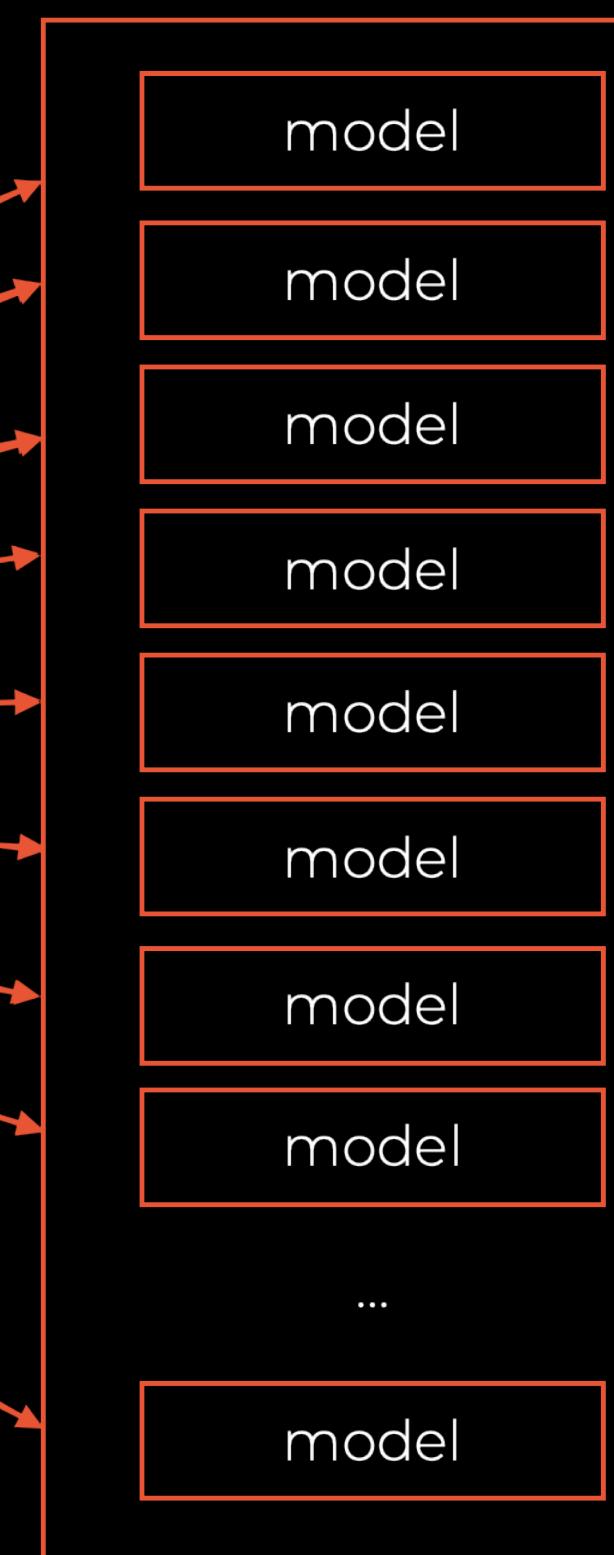
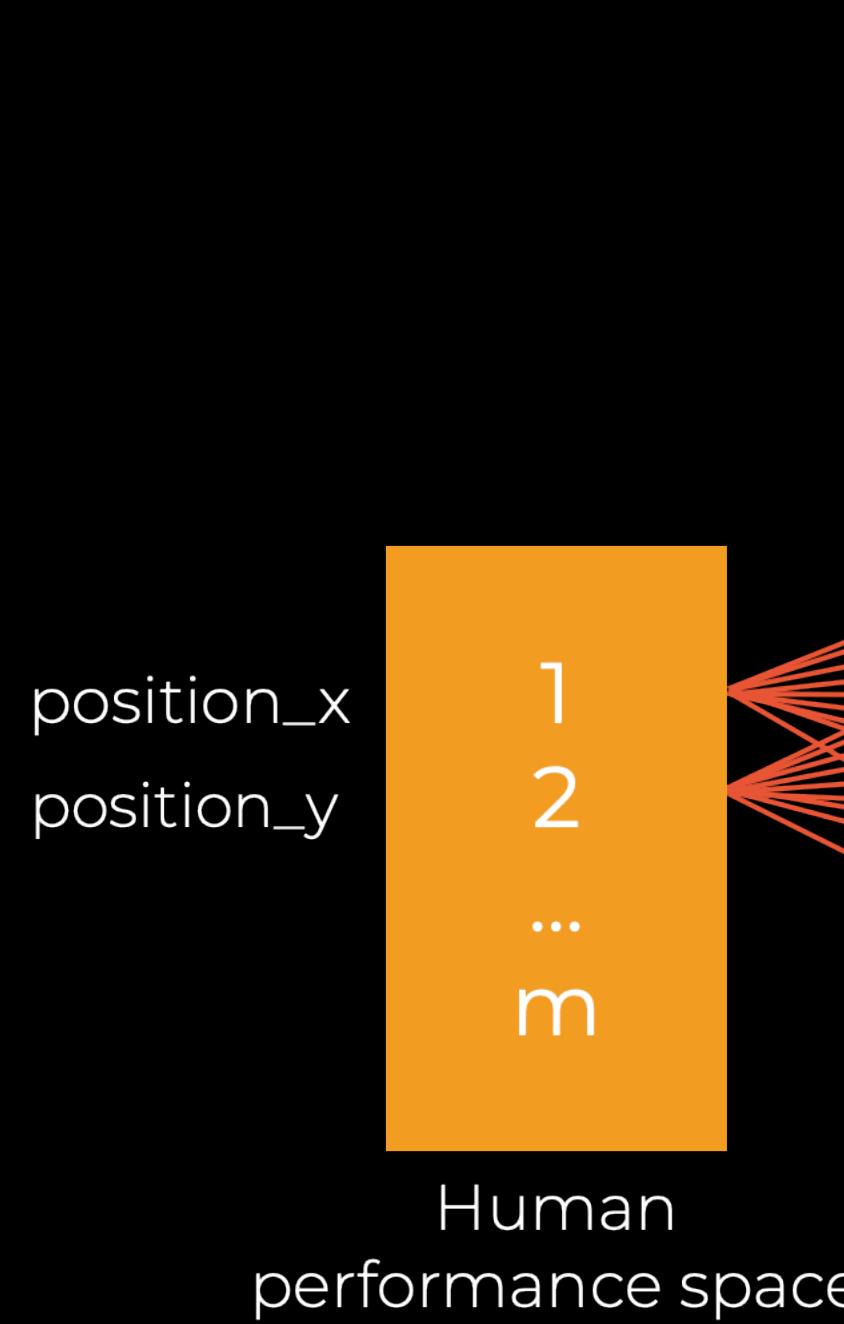
Waterlab (~4h)

Archivo Sonoro Museo de la Memoria y los Derechos Humanos. Santiago, CL (~62h)

Our approach to steerability and embodiment

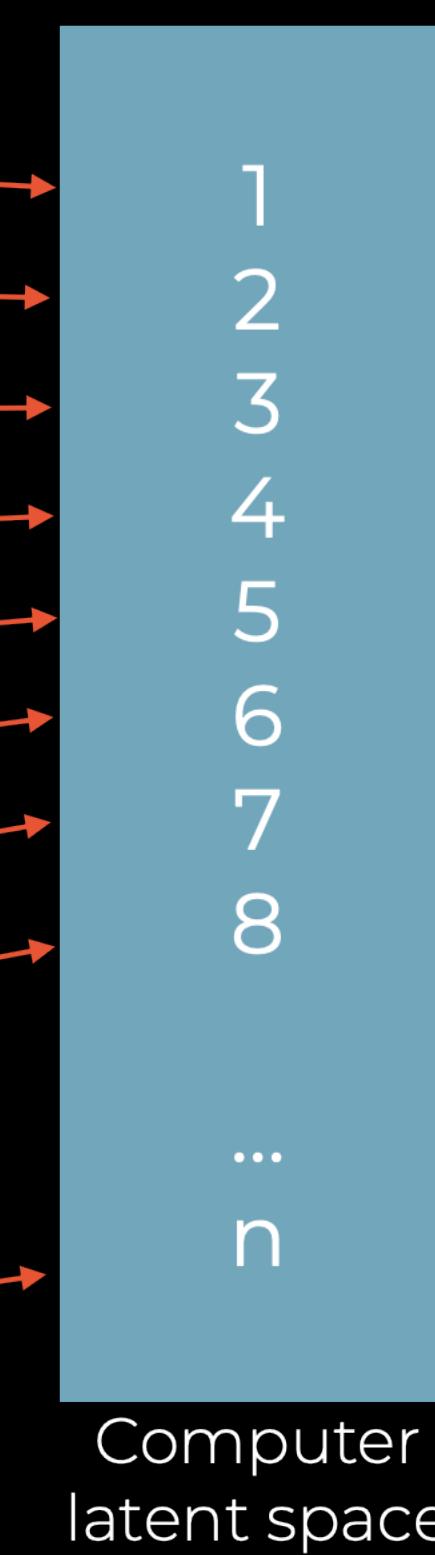
Interactive machine learning as a mapping tool

INPUTS



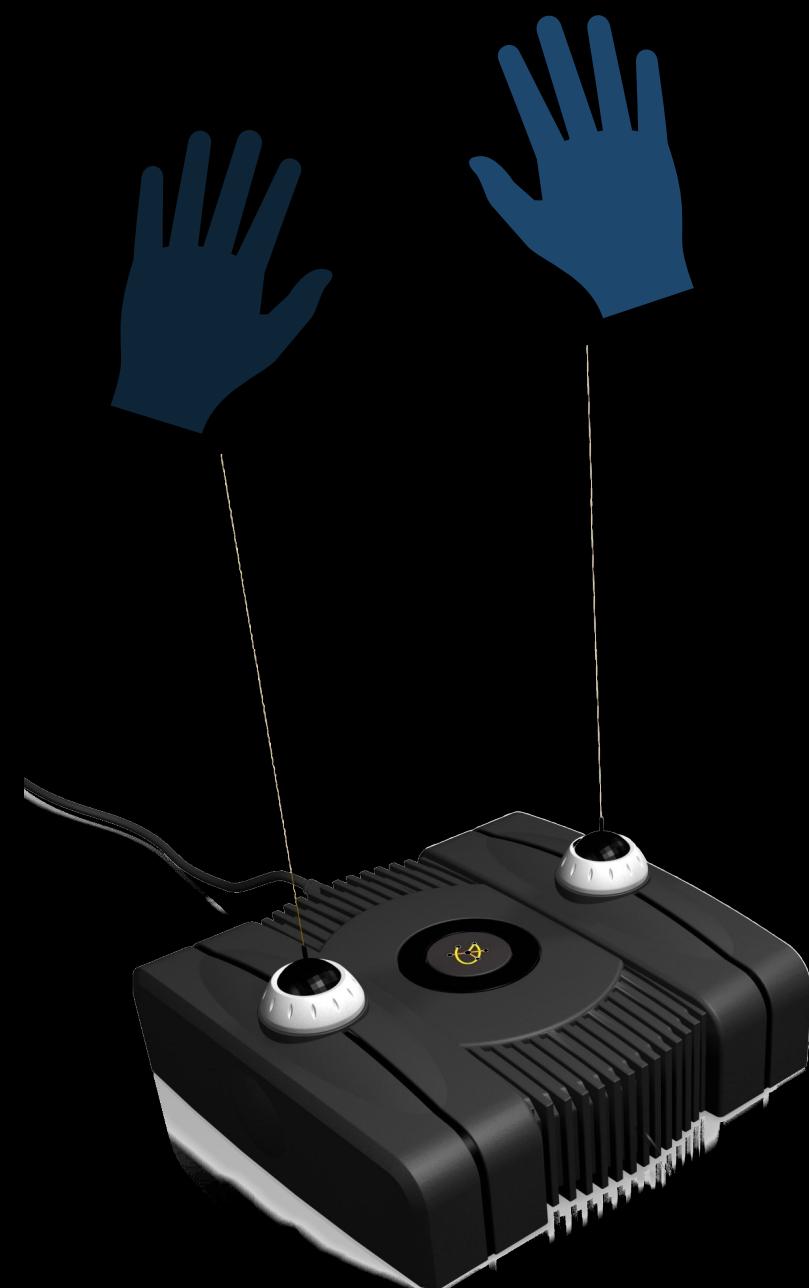
Interactive machine
learning application

OUTPUTS

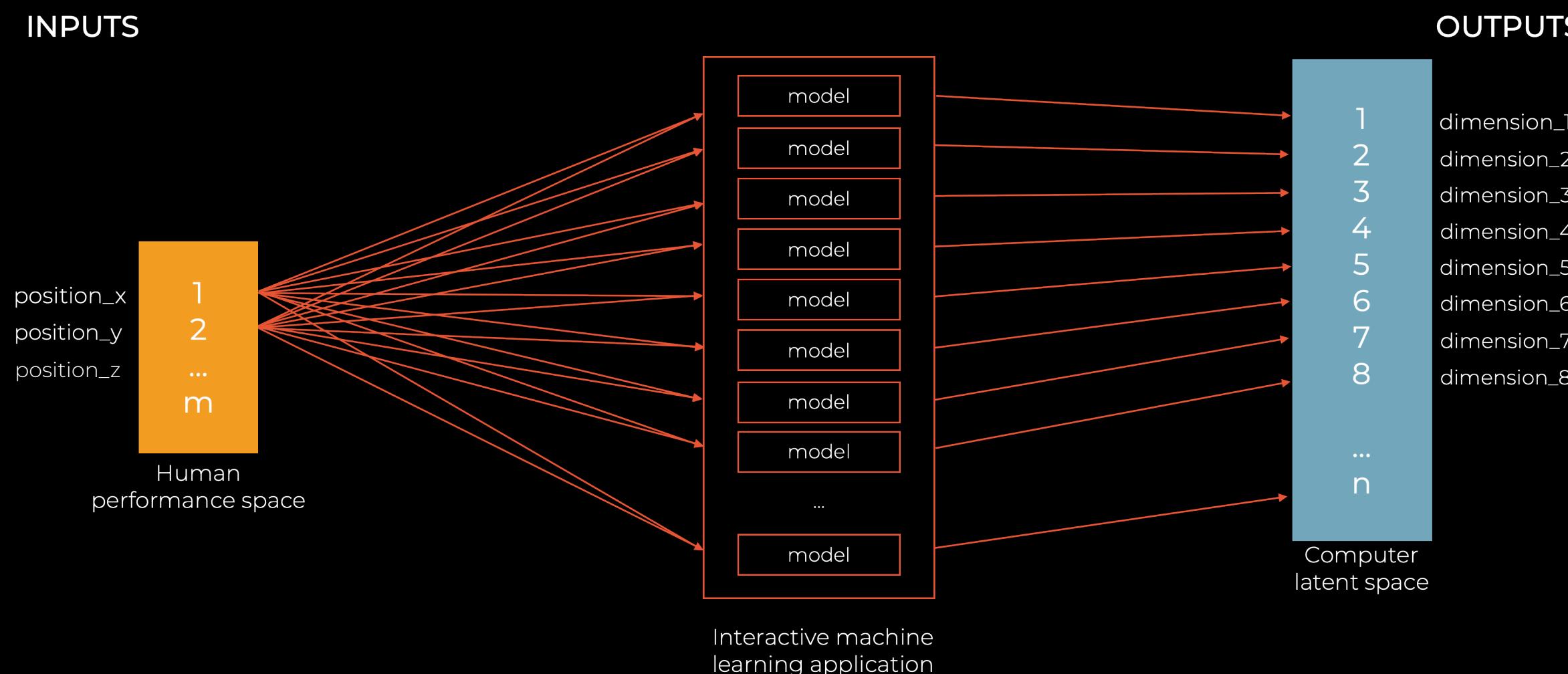


Our approach to steerability and embodiment

Interactive machine learning as a mapping tool



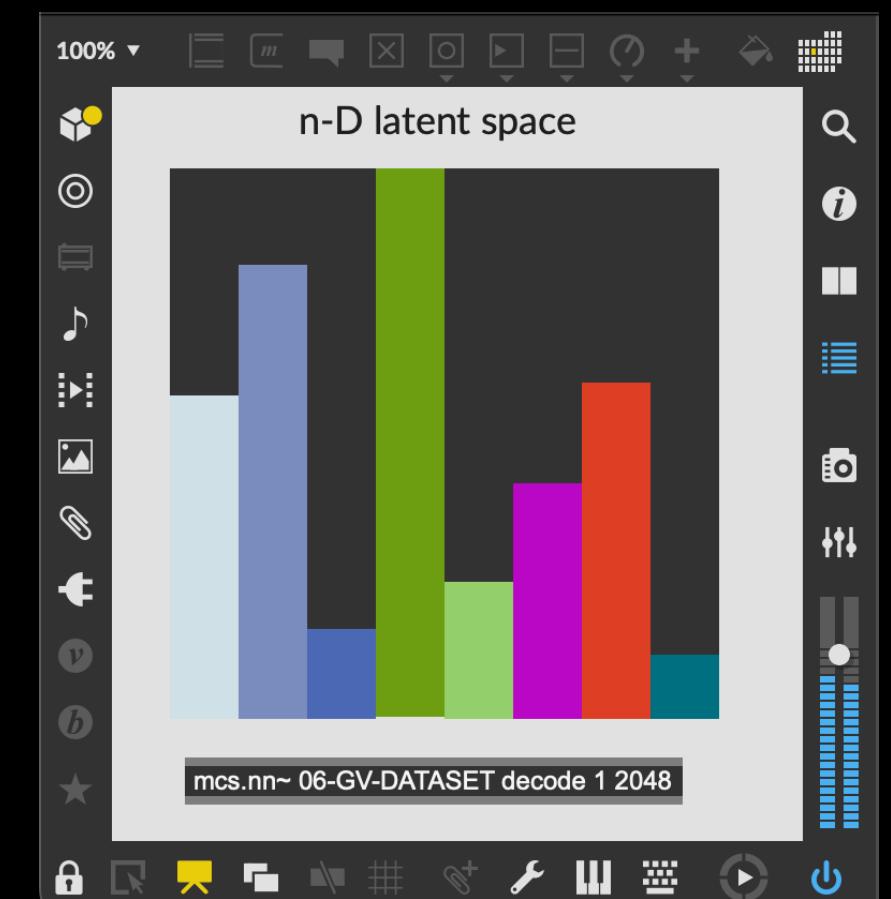
Gestural controller



Wekinator or FluCoMa as mapping tool applications



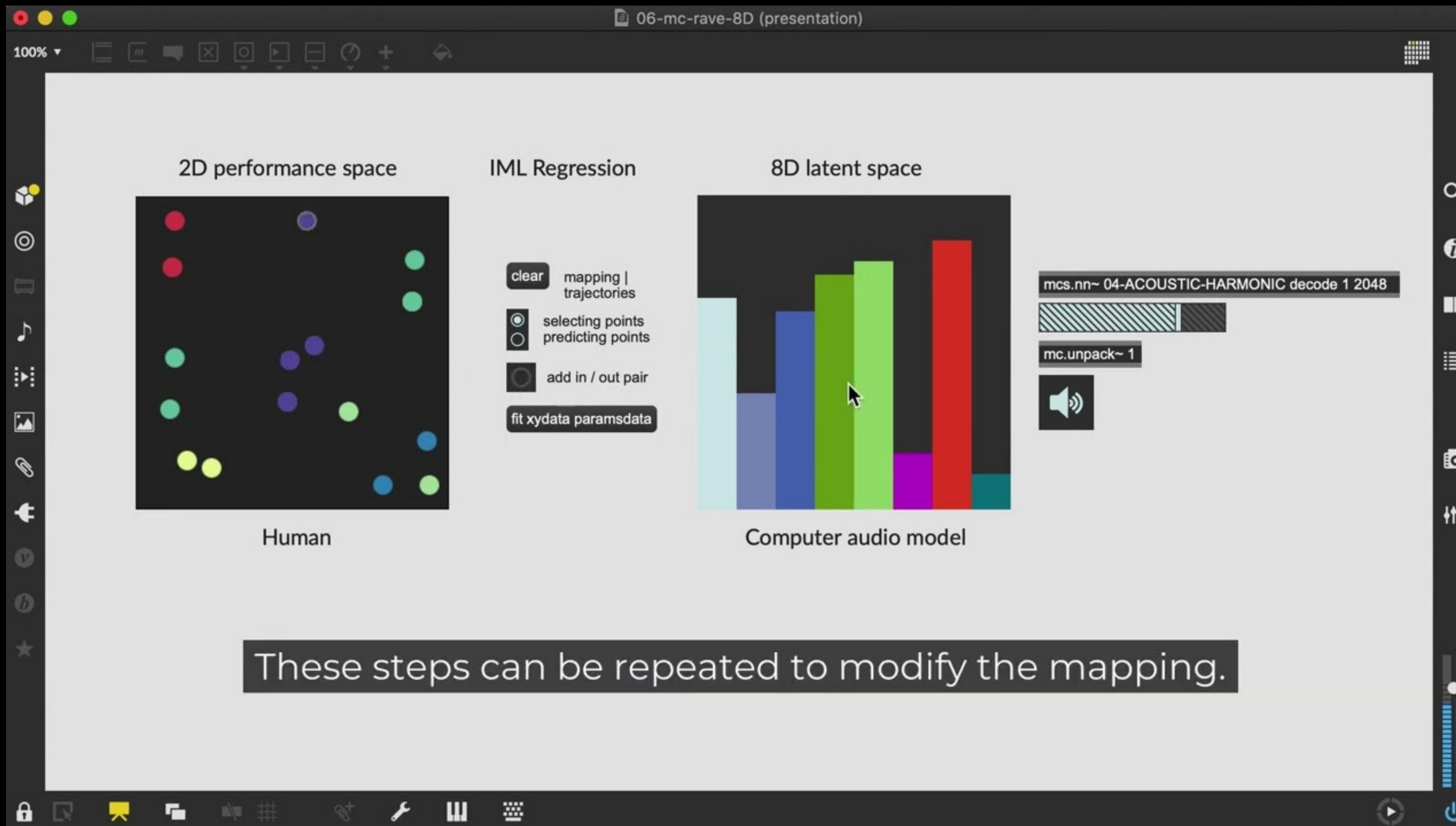
RAVE
Caillon and Esling (2021)



RAVE models for NAS

Mapping physical to digital space

2DoF to 8DoF



Embodying a generative model

2DoF to 8DoF



18 Festival Internacional de Música y Nuevas Tecnologías. CMMAS, México. September 2022.

Embodying a generative model

6DoF to 16DoF



With sound artist dedosmuertos in La Villette, Paris. July 2022.

Embodying a generative model

1434DoF to 8DoF



Viglienzi, Gabriel, and Rebecca Fiebrink. 2023. Re•col•lec•tions: Sharing sonic memories through interactive machine learning and neural audio synthesis models In *Creative AI track of the 37th Conference on Neural Information Processing Systems (NeurIPS 2023)*.

Generative audio systems

Many unknowns

Training

How much **training data**?

How much **consistency and variability** in the data?

When to **stop training**?

Performance

How to explore a **high-dimensional latent space**?

How to explore an **unlabelled** latent space?

How to explore a latent space that **changes for each model**?

Steering generative audio systems

Insights

- * Exploration generates knowledge
- * Work with a well-known dataset (e.g., your custom dataset!)
- * IML approach to interact with high-dimensional latent space is useful
- * IML enables performers to introduce long-term temporal coherence lacking in generative audio systems

Thanks!

Gabriel Vigliensoni
gabriel.vigliensoni@concordia.ca

