

---

# Adaptive representation learning for the gestural control of deep audio generative models

---

**Co-supervised by**

Philippe Esling<sup>1</sup>

Frédéric Bevilacqua<sup>2</sup>

Geoffroy Peeters<sup>3</sup>

Sarah NABI

1st year PhD Student

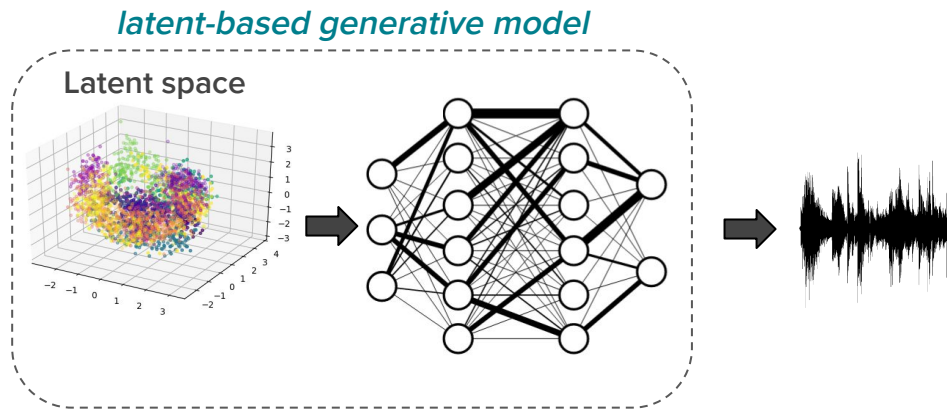
(01.01.2023 - 31.12.2025)

1. *Artificial Creative Intelligence and Data Science (ACIDS), Sound Analysis-Synthesis, IRCAM, UMR STMS - Sorbonne CNRS*
2. *Sound Music Movement Interaction (ISMM), IRCAM, UMR STMS - Sorbonne CNRS*
3. *Information Processing and Communications Laboratory (LTCl), Telecom Paris - Institut Polytechnique*

# Problem statement

---

Deep generative models for audio synthesis → new musical instruments

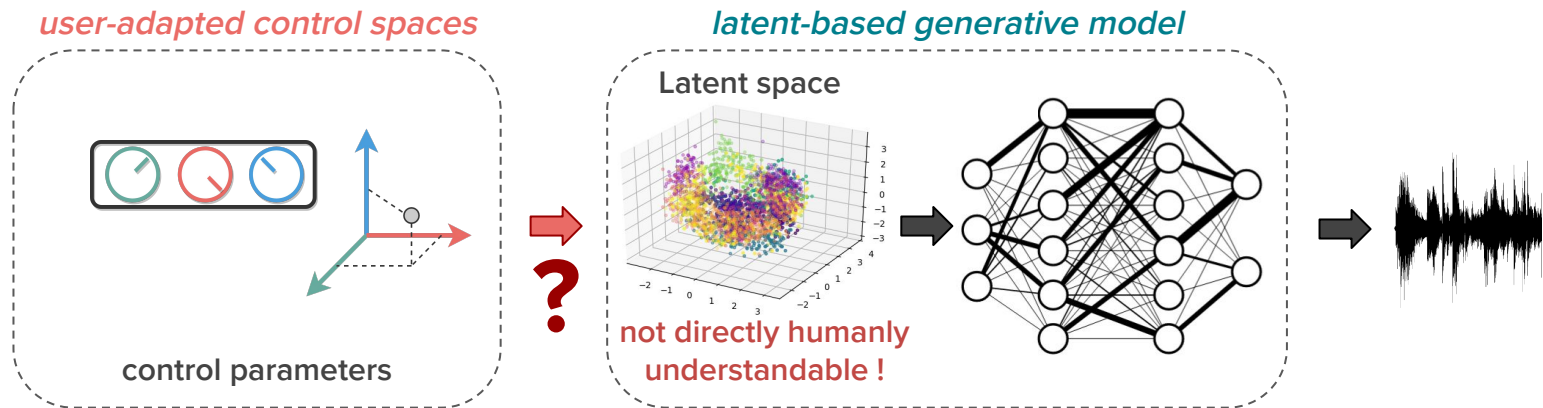


# Problem statement

Deep generative models for audio synthesis → new musical instruments

## Challenges

- lack of control and interaction
- existing methods rely on **lots of labeled examples**
- changing controls implies **retraining** the model



# Problem statement

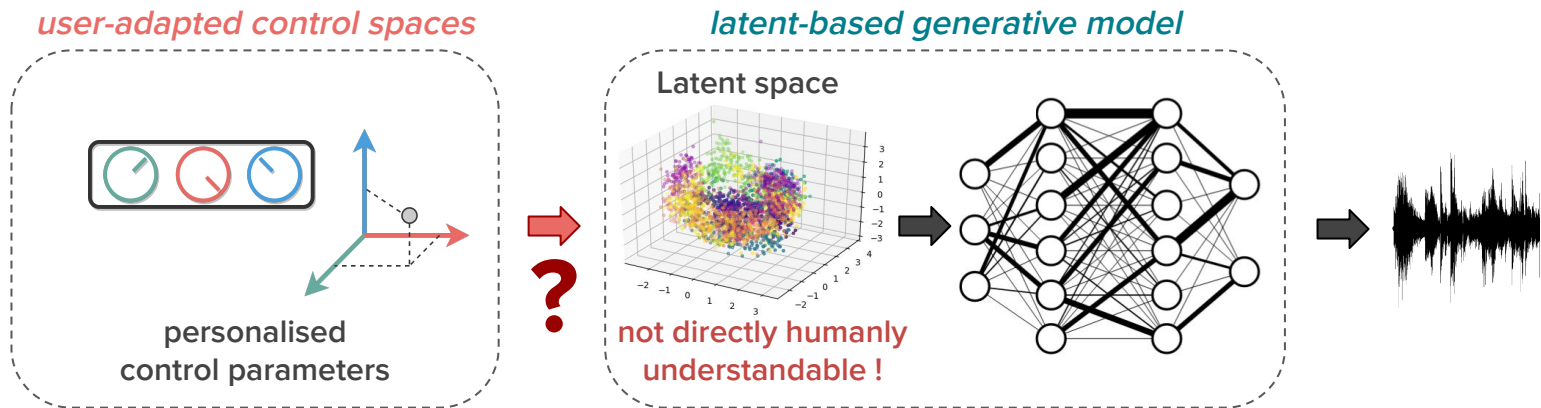
## Deep generative models for audio synthesis → new musical instruments

### Challenges

- **lack of control and interaction**
- existing methods rely on **lots of labeled examples**
- changing controls implies **retraining** the model

### Musical creation and applications

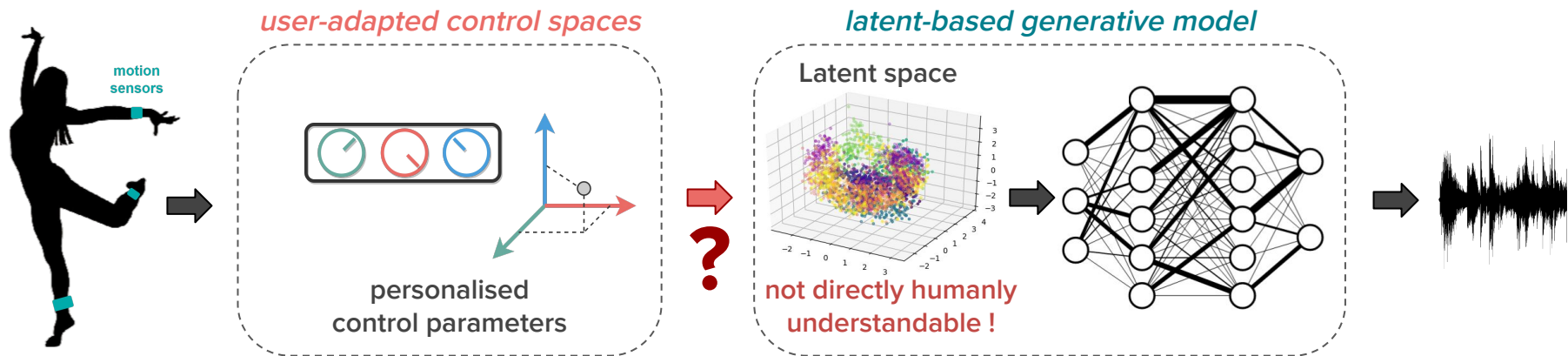
- rely on **limited datasets** that closely match a given artistic intent
- **iterative and exploratory** process
- need for different **levels of intuitive** and **personalised** controls
- **motion** and **sound** are **intrinsically bounded**



# Objective

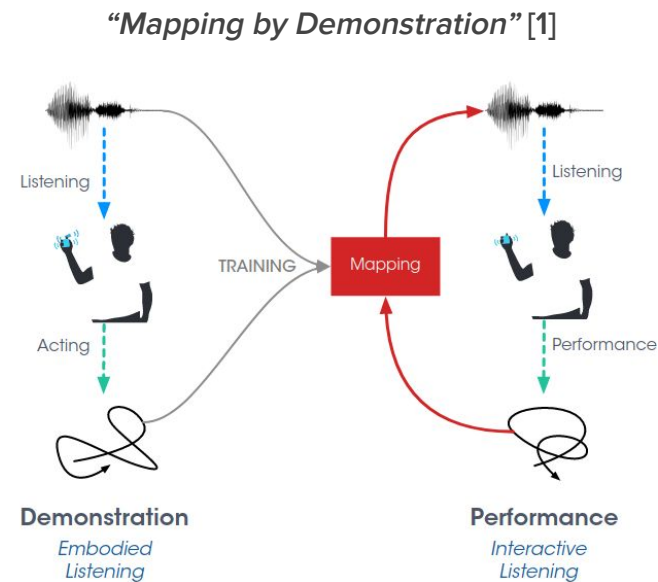
Deep generative models for audio synthesis → new musical instruments

Thesis goal: design **user-adapted control methods** based on **few examples** with a **multimodal motion-to-sound generation approach** to promote co-creative human-machine interaction



# Research directions

- **Self-Supervised Learning**
  - bypass the need of annotations
  - multimodal representations
- **Few-Shot Learning**
  - customize controls with few paired examples
- **Implicit Mapping Strategy [1]**
  - iteratively explore and adapt the control space
  - select dynamically few motion-sound paired examples



*Thank you !*