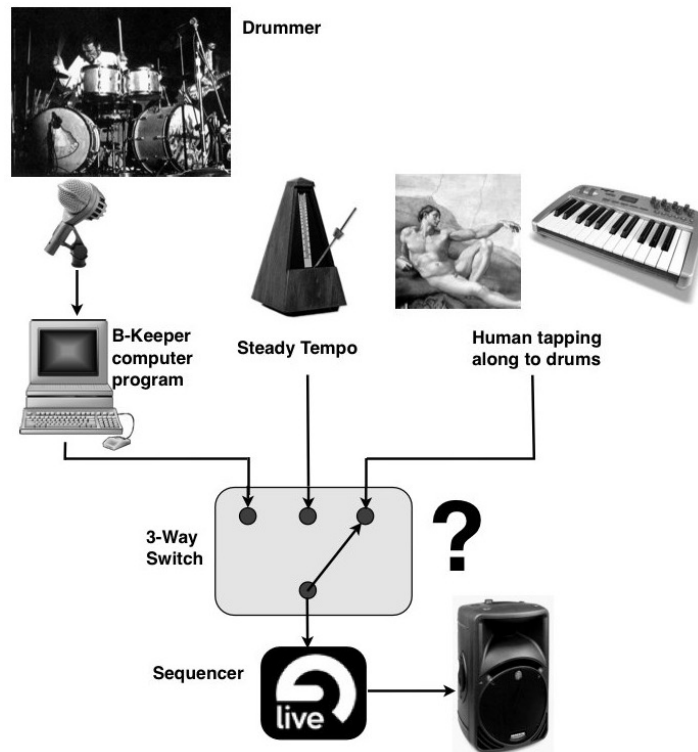# Explainable AI + Music

Prof. Nick Bryan-Kinns

Queen Mary University of London, UK

Media and Arts Technology Centre

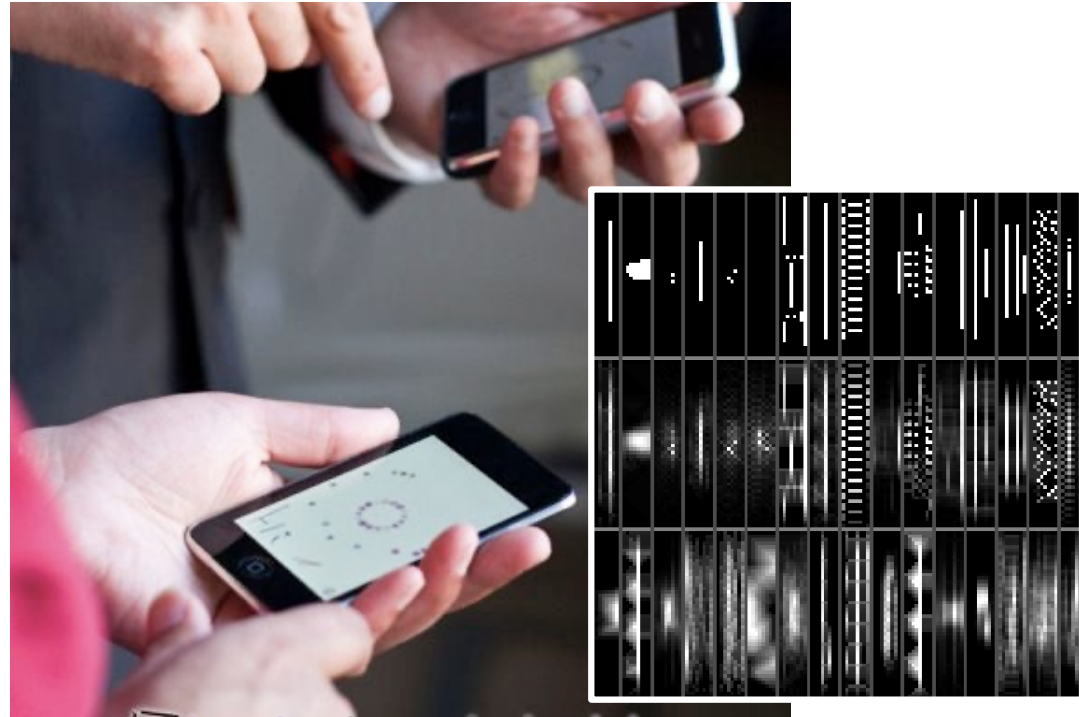AI and Music Centre for Doctoral Training

AI + HCI

Turing Test for AI drummers

Robertson, A., Plumbley, M. D., & Bryan-Kinns, N. (2008) A Turing Test for B-Keeper: Evaluating an Interactive Real-Time Beat-Tracker. In Proceedings of New Interfaces for Musical Expression (NIME'08), Genova, Italy.

AI + HCI

ML Analysis of Mutual Engagement

Bryan-Kinns, N. (2018). Case Study of Data Mining Mutual Engagement. In Proceedings of British HCI Conference 2018 (BCS-HCI 2018).

How People Interact with AI in Creative Settings

XAI + Music

Artificial Intelligence (AI)

Human Centred AI

eXplainable AI (XAI)

XAI+Music

XAI+Music Probotype

Open Research Questions

Artificial Intelligence

3 Waves of AI (Xu, 2019)

	First Wave (1950s–1970s)	Second Wave (1980s–1990s)	Third Wave (2006–)
Major advances in technologies	Early symbolism and connectionism school, production systems, knowledge inference, preliminary expert systems	Statistical model in speech recognition and machine translation, artificial neural network in pattern recognition, expert systems	Breakthroughs in applications of deep learning in speech recognition, pattern recognition, big data, high-performance computers
Human needs	Not satisfied	Not satisfied	Starting to provide useful and real problem-solving AI solutions
Focus	Technological solutions	Technological solutions	Integrated solutions: ethical design, technological enhancement, human factors design
Characteristics	Academia driven	Academia driven	Technological enhancement and application + a human-centered approach

Table 1. A comparison of the three waves of AI.

Human Centred AI (Shneiderman, 2020/1)

"goal is to amplify, rather than erode, human agency"

"computers should play a supportive role, amplifying people's ability to work in masterful or extraordinary ways"

"HCAI systems emerge when designers, software engineers, and managers adopt user-centered participatory design methods by engaging with diverse stakeholders"

eXplainable AI (XAI)

Gunning (2016) proposed that XAI aims to help people understand AI:

Why did the AI do that?

Why not something else?

When does the AI succeed?

When does the AI fail?

When can people trust an AI?

AI + Music

Long history

Ada Lovelace (1843), Illiac Suite (1957)

Challenging requirements

Real-time, robust, reliable & co-creative, engaging, emotive

Music is huge market

Music composition, recording and performance, IP protection, computer games, KTV apps, etc. …

Major companies care about it

Apple, Google, Sony, Alibaba, Tencent, ByteDance, …

XAI and Music

AI-Human Co-creation

Survey: XAI Music (89 Papers)

Role of the AI: Lubart (2005) "How can computers be partners in the creative process: Classification and commentary on the Special Issue"

Interaction with the AI: Cornock & Edmonds (1973) "The Creative Process Where the Artist Is Amplified or Superseded by the Computer"

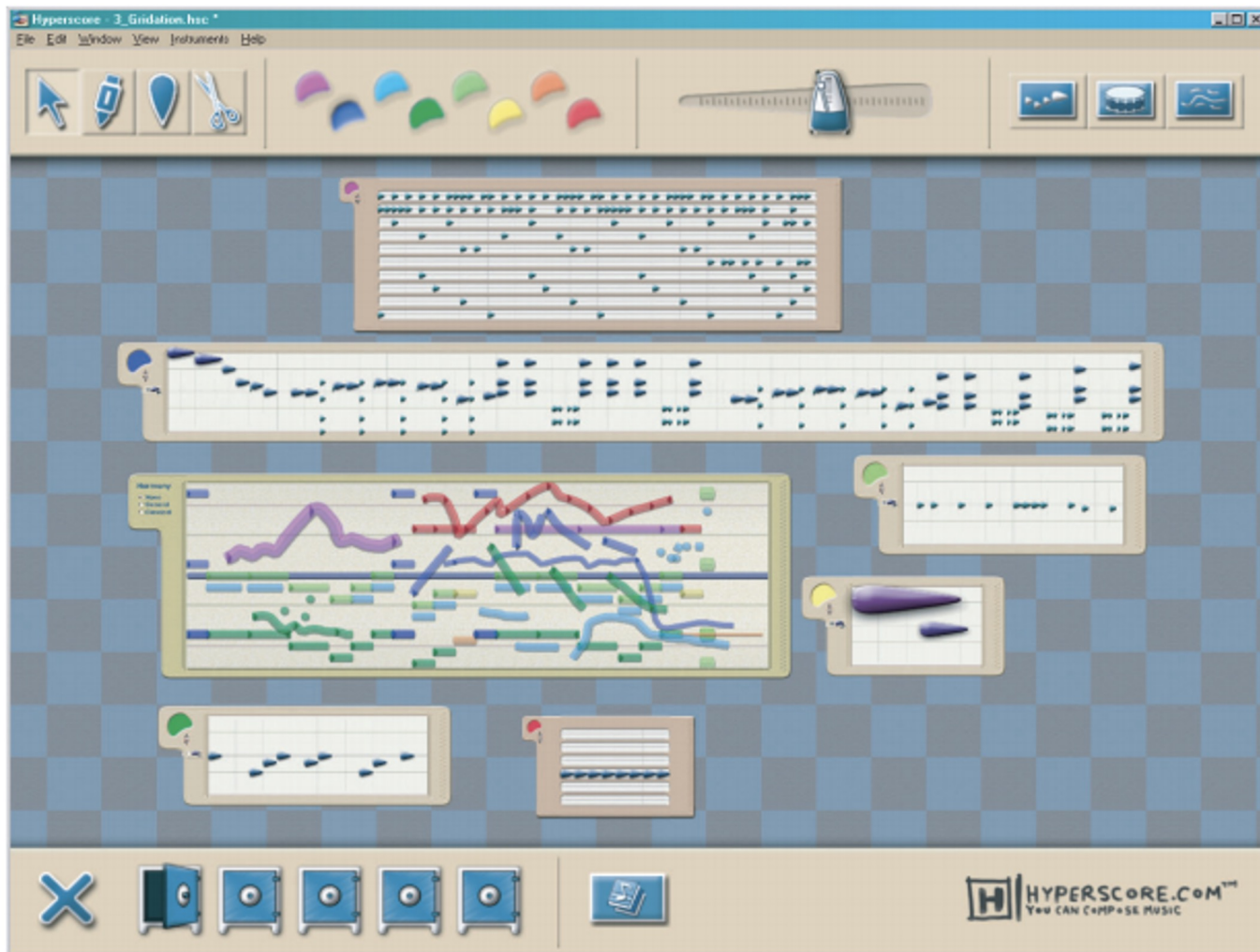*Common ground with the AI*: Clark & Brennan (1991) "Grounding in Communication"
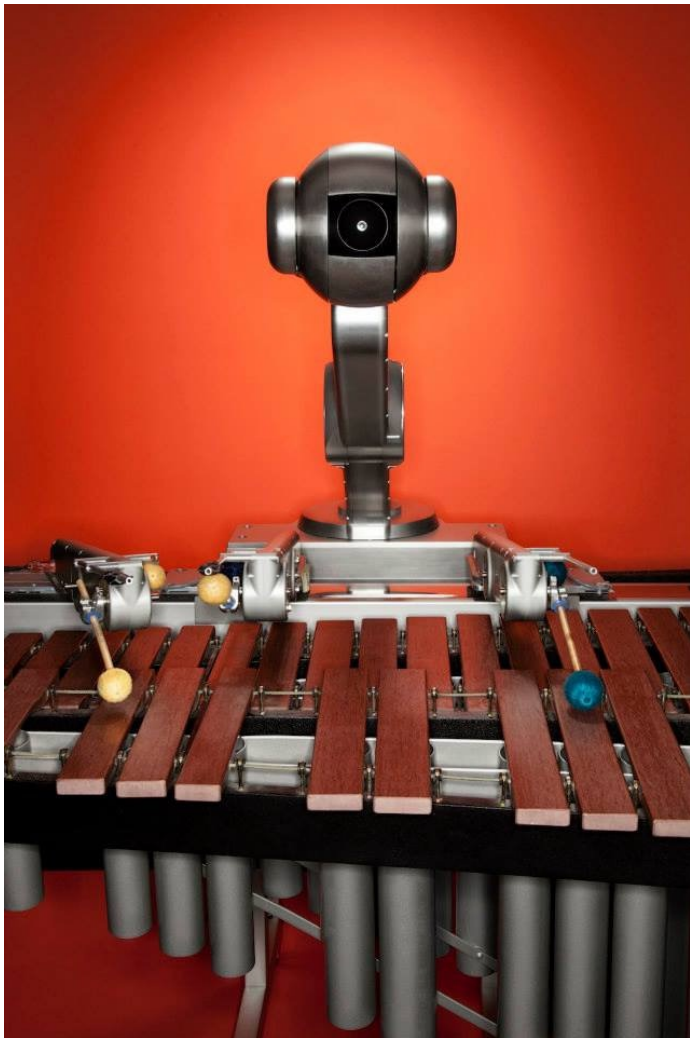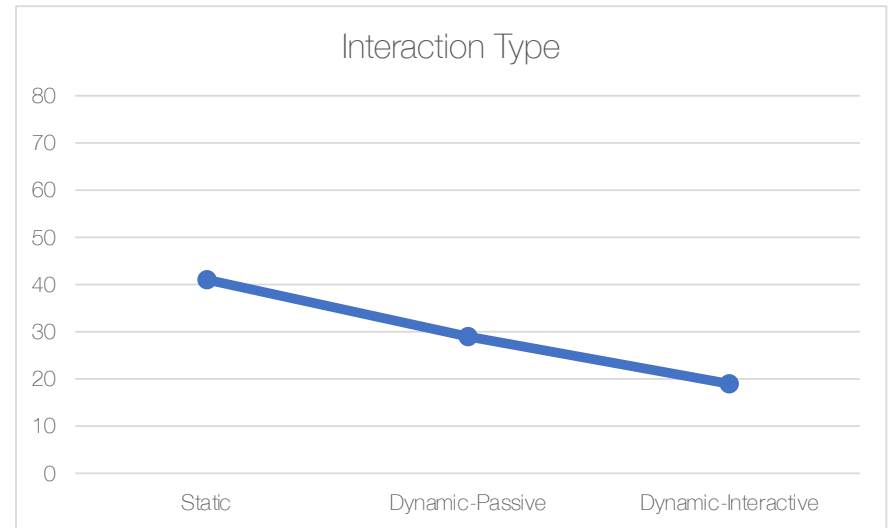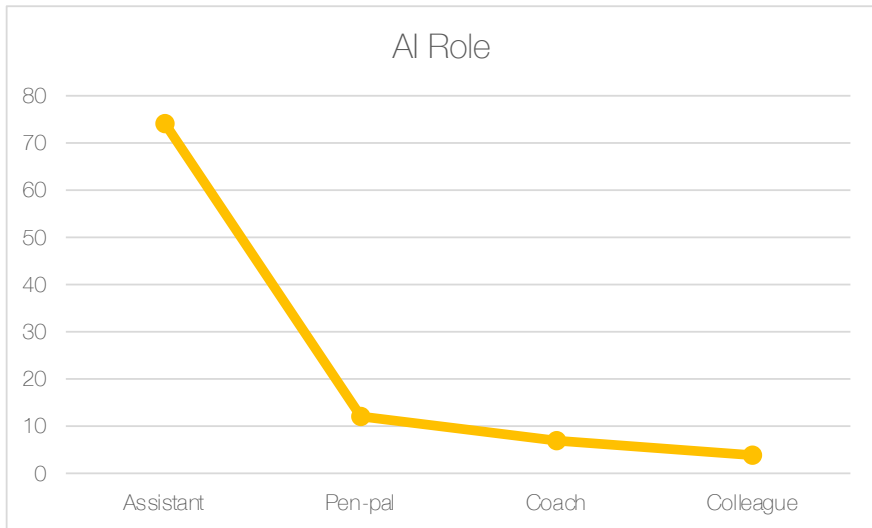
Shimon
Robotic Marimba Player
Georgia Tech

Robotic Flute Player
Tsinghua
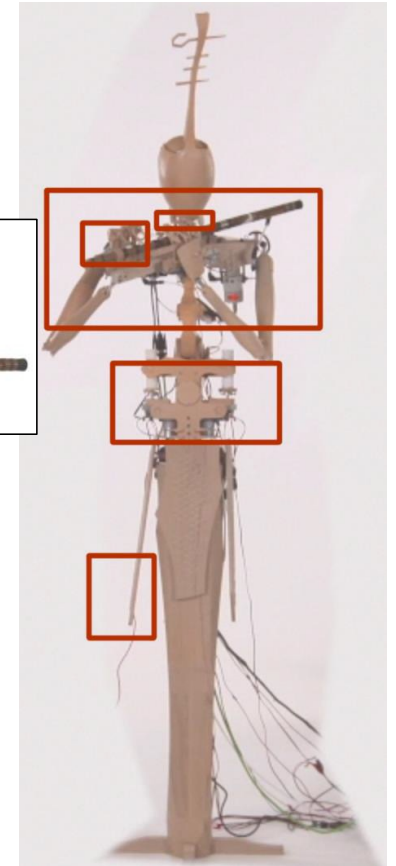
Hyperscore
MIT Media Lab

N. Bryan-Kinns

Research Challenges

Lack of XAI for music (and the arts in general)

    What makes a good XAI for music?

    How can XAI contribute to making AI more co-creative?

Difficult to deploy XAI directly in music

    XAI is functional

    XAI is new
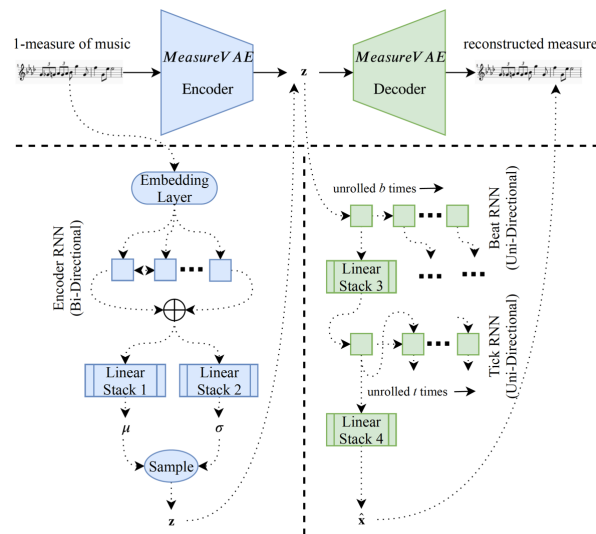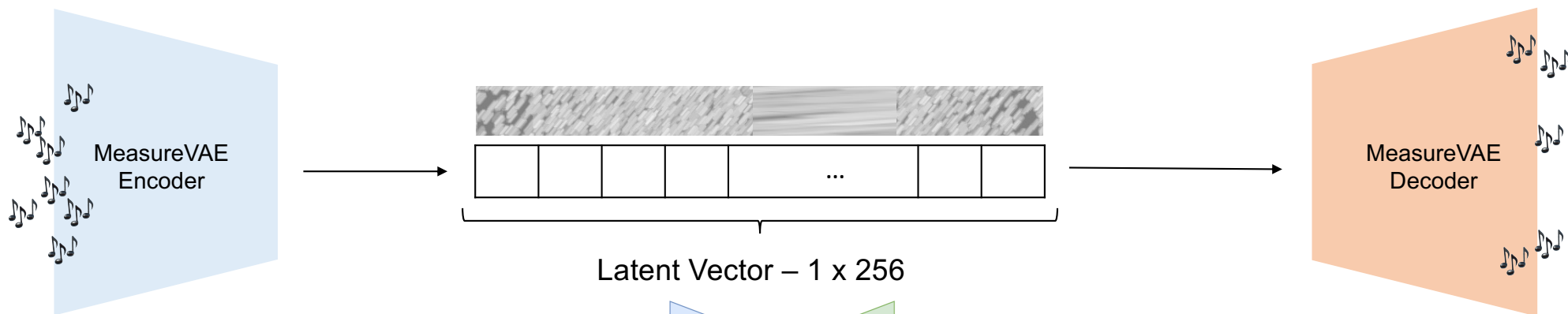
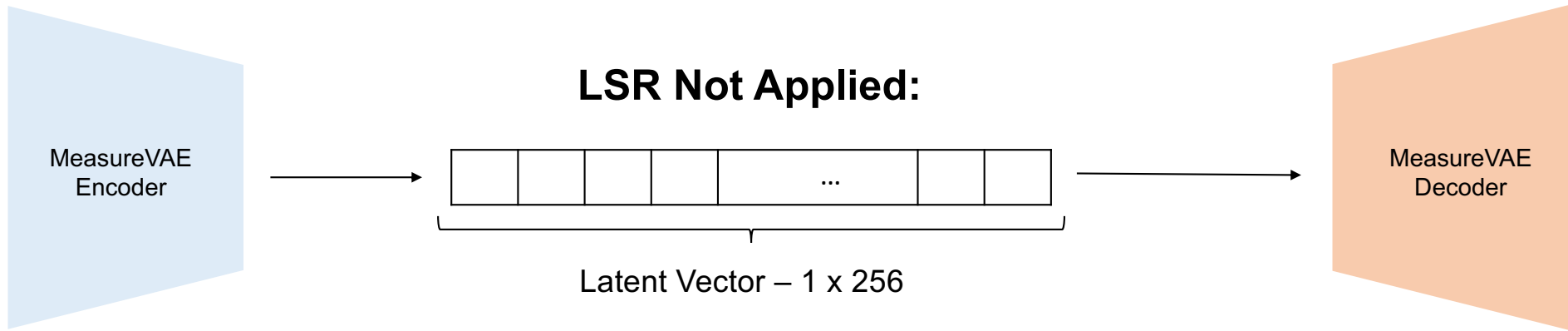   Do we really want to understand everything?

XAI & Music Prototype

Deep generative models

Latent spaces

MeasureVAE Encoder

Latent Vector – 1 x 256

MeasureVAE Decoder

1-measure of music → *MeasureVAE* Encoder → z → *MeasureVAE* Decoder → reconstructed measure

Encoder RNN (Bi-Directional)

Embedding Layer

Linear Stack 1 — μ

Linear Stack 2 — σ

Sample

z

Beat RNN (Uni-Directional)

unrolled *b* times →

Linear Stack 3

Tick RNN (Uni-Directional)

unrolled *t* times →

Linear Stack 4

x̂

A. Pati, A. Lerch, and G. Hadjeres, "Learning to traverse latent spaces for musical score inpainting,"arXiv preprint arXiv:1907.01164, 2019.

LSR Not Applied:

MeasureVAE Encoder

				...		

Latent Vector – 1 x 256

MeasureVAE Decoder

LSR Applied:

MeasureVAE Encoder

Rhythmic Complexity   Note Range   Note Density   Average Pitch Interval

					...		

Latent Vector – 1 x 256

MeasureVAE Decoder

Demo https://xai-with-lsr-ui.vercel.app/
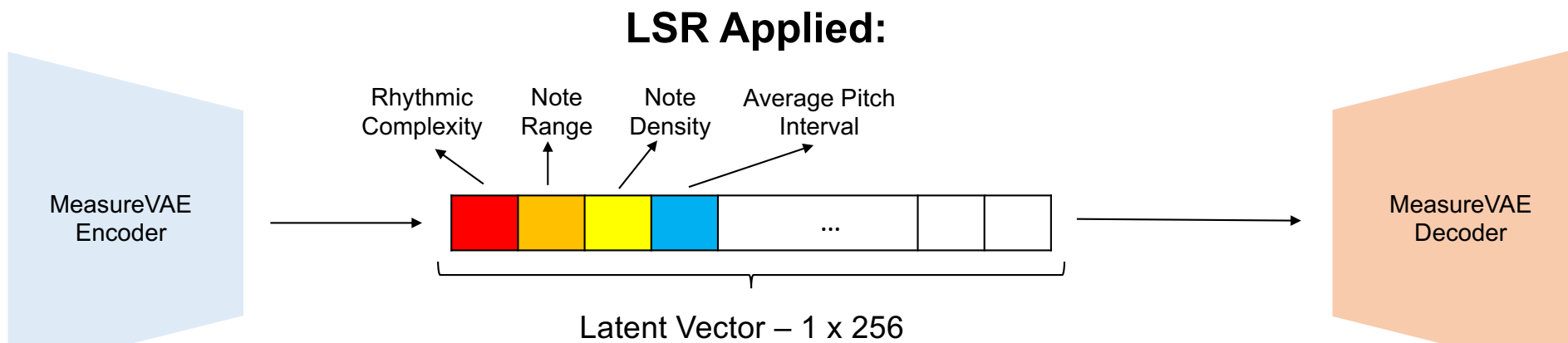
Explainable AI via Latent Space Regularisation

Latent Space Regularisation

Rhythmic Complexity, Note Range, Note Density, Average Pitch Interval

MeasureVAE Encoder → Latent Vector – 1 x 256 → MeasureVAE Decoder

Input MIDI

2D Display Mode: Train Data Contr | RC Surface Map | NR Surface Map

Click on the Pad to Activate: Off

Note Range Level Latent Dimension = 1 (5.4 to -6.1)

Rhythmic Complexity Level Latent Dimension = 0 (-5.4 to 6.6)

2D Display Mode: Train Data Contr | ND Surface Map | API Surface Map

Click on the Pad to Activate: Off

Average Pitch Interval Level Latent Dimension = 3 (10.0 to -4.0)

Note Density Level Latent Dimension = 2 (-2.8 to 3.2)

Output MIDI (Input Variations)

0:00 / 0:02

Latent Vector Values:

Rhy Complx: 6.6   Note Range: 4.3   Note Density: 3.2   Avg Pitch Interval: 10.0

Augmentation of Pati, Ashis and Lerch, Alexander. (2019). Latent Space Regularization for Explicit Control of Musical Attributes.

Berker Banar, 2021.

Open Research Questions

Musical features

AI models

Training sets

Nature of explanation

Co-design of AI

Artistic practice

UX design

Open Research Questions

Musical Features

Nature of Explanation

UX Design Challenge

Musical Features

4 typical low-level features:

Rhythmic Complexity   Note Range

Note Density   Average Interval Jump

There are many many more…

Higher-level features:

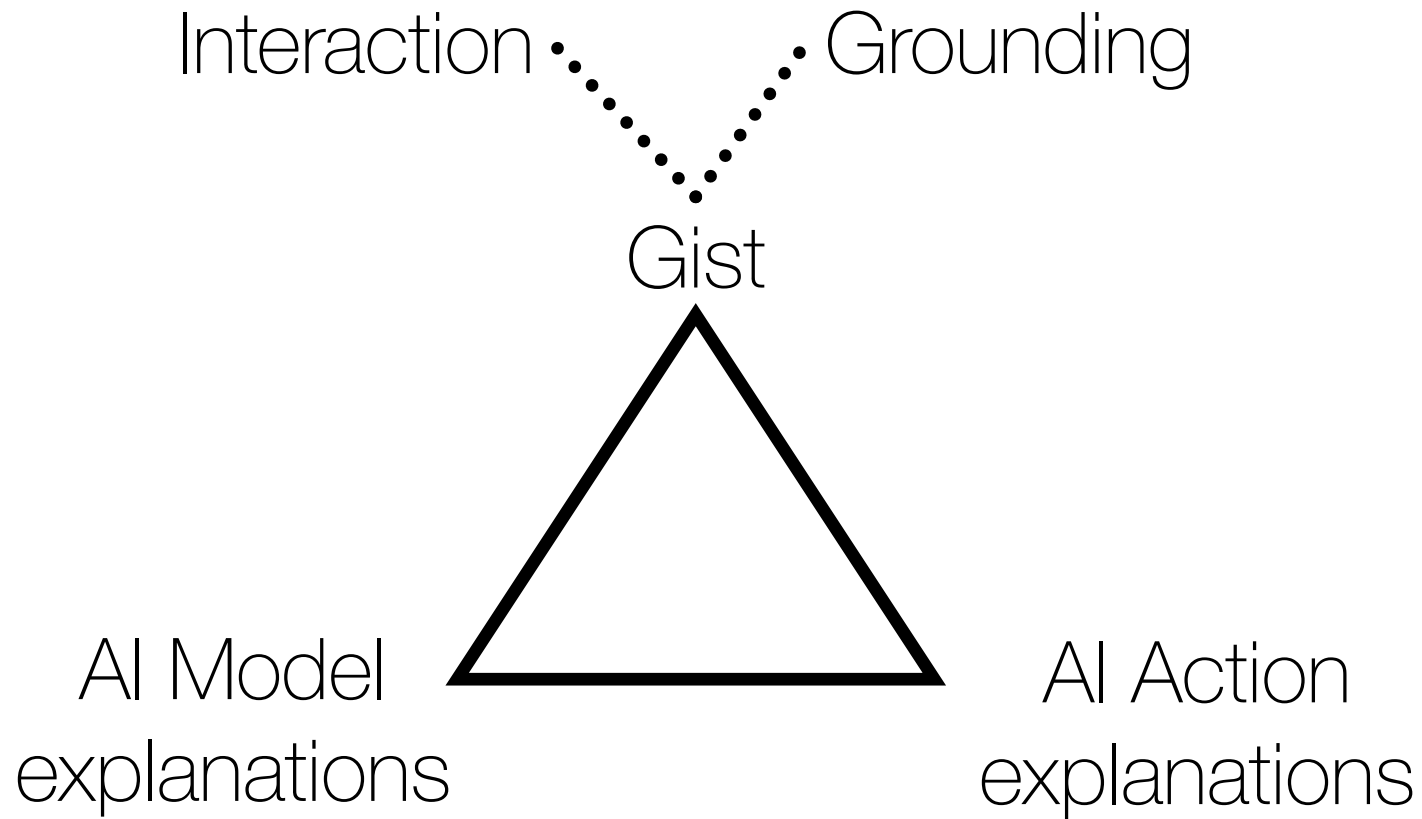Style   Texture   Emotional valence

Many many many more…

Nature of Explanation

How much explanation?

What aspects of the AI are most important to explain?

Predictable explanations and surprising explanations

Gist - Co-creation with AI relies on us getting the gist of what an AI is doing

Interaction ⋯ ⋯ Grounding

Gist

AI Model
explanations

AI Action
explanations

UX Design Challenge

eXplainable AI

4+ dimensions

Interactive

Real-time generation

Learnable/ Intuitive

Surprising/ Engaging

Rhythmic Complexity

Note Range

Note Density

Average Interval Jump
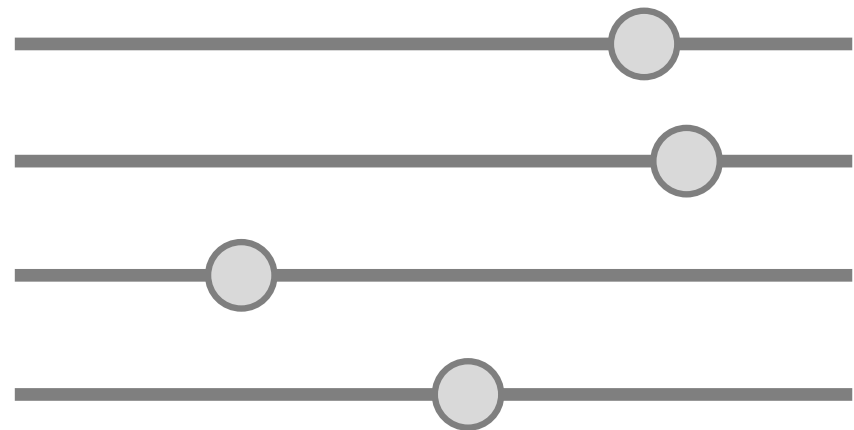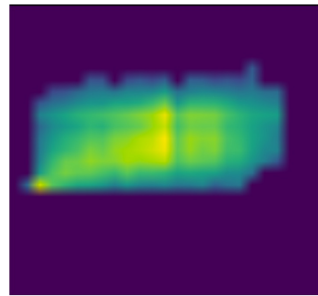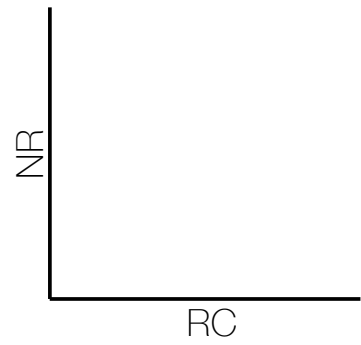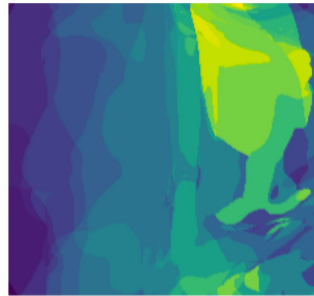
(a) TDRCNR   (b) RC   (c) NR

Rhythmic Complexity
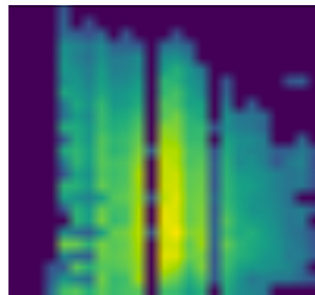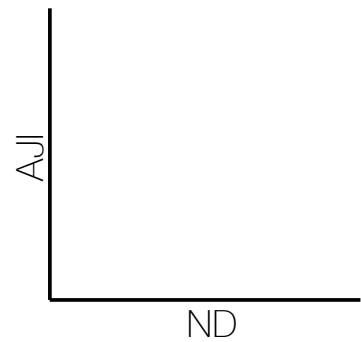
Note Range
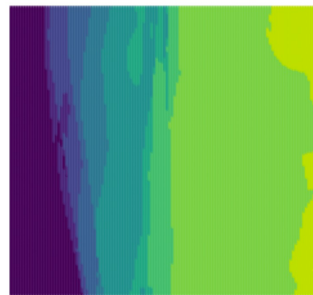
(d) TDNDAP   (e) ND   (f) AIJ
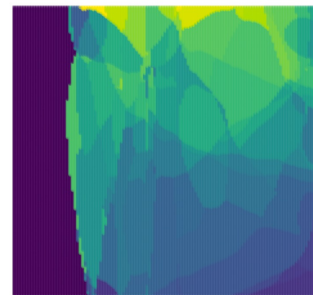
Note Density

Average Interval Jump

http://www.jackarmitage.com/svelte-pianoroll-deploy/
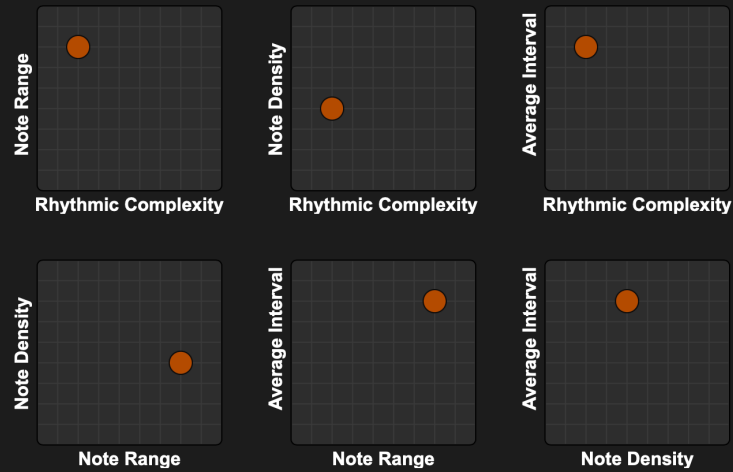
Google Magenta inpainting
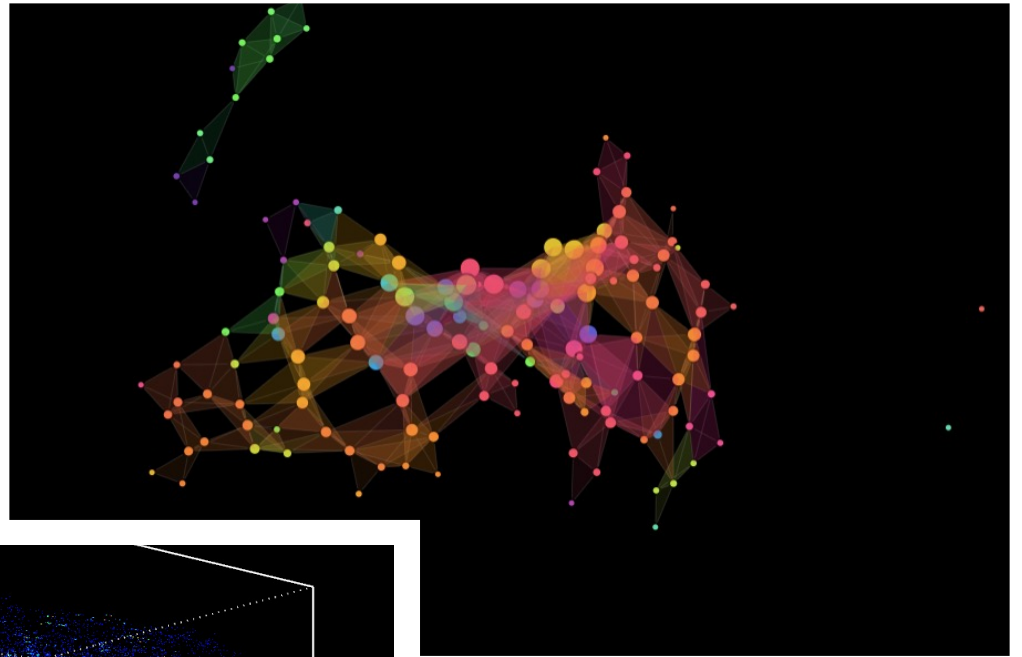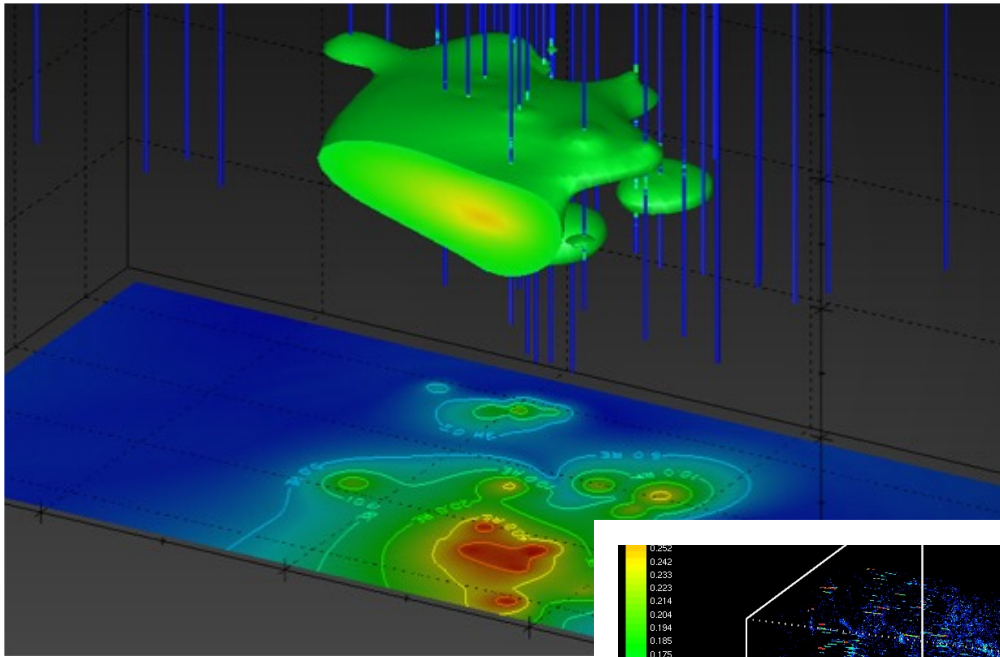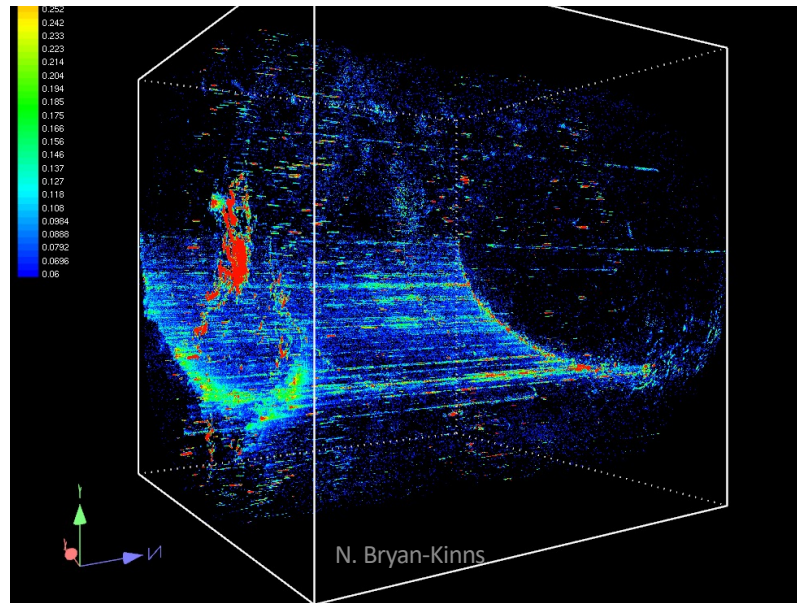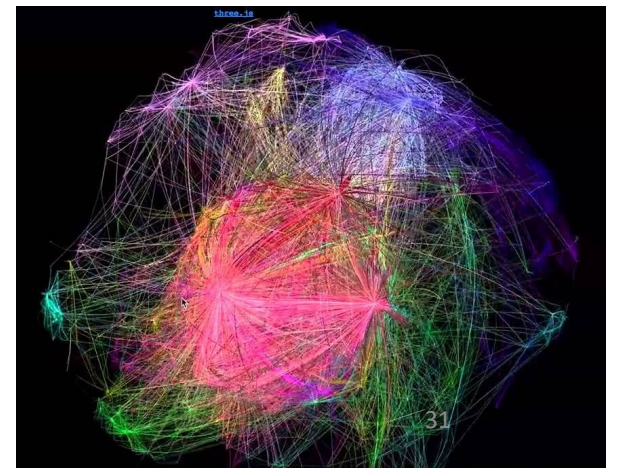
N. Bryan-Kinns

Navigation of 4+ Dimensions
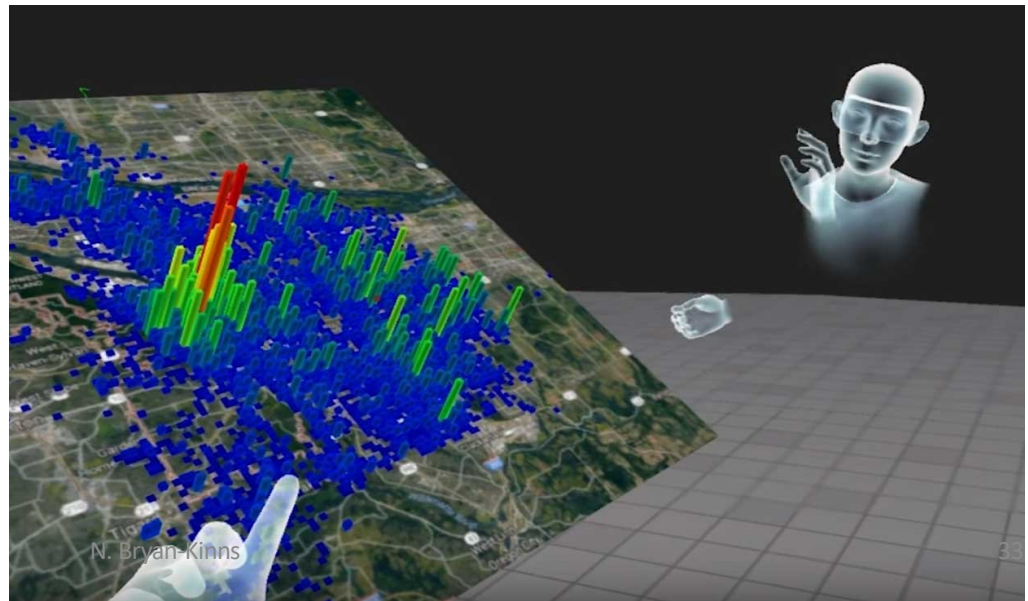
3D input?

Gestural interaction?

Tangible UIs?

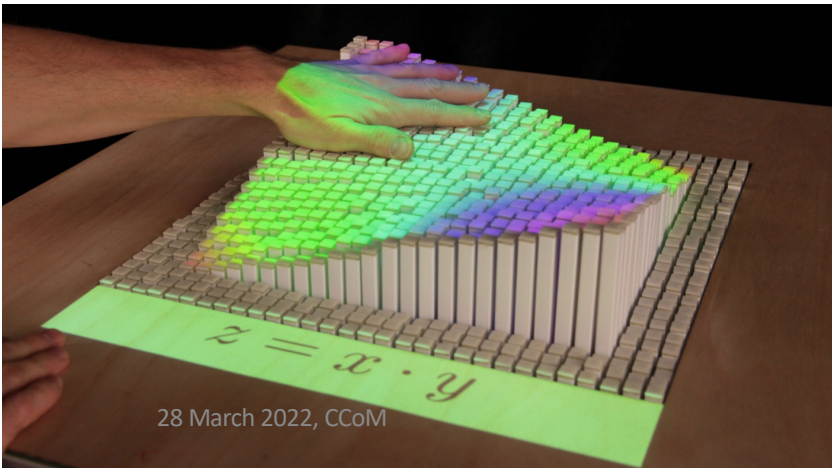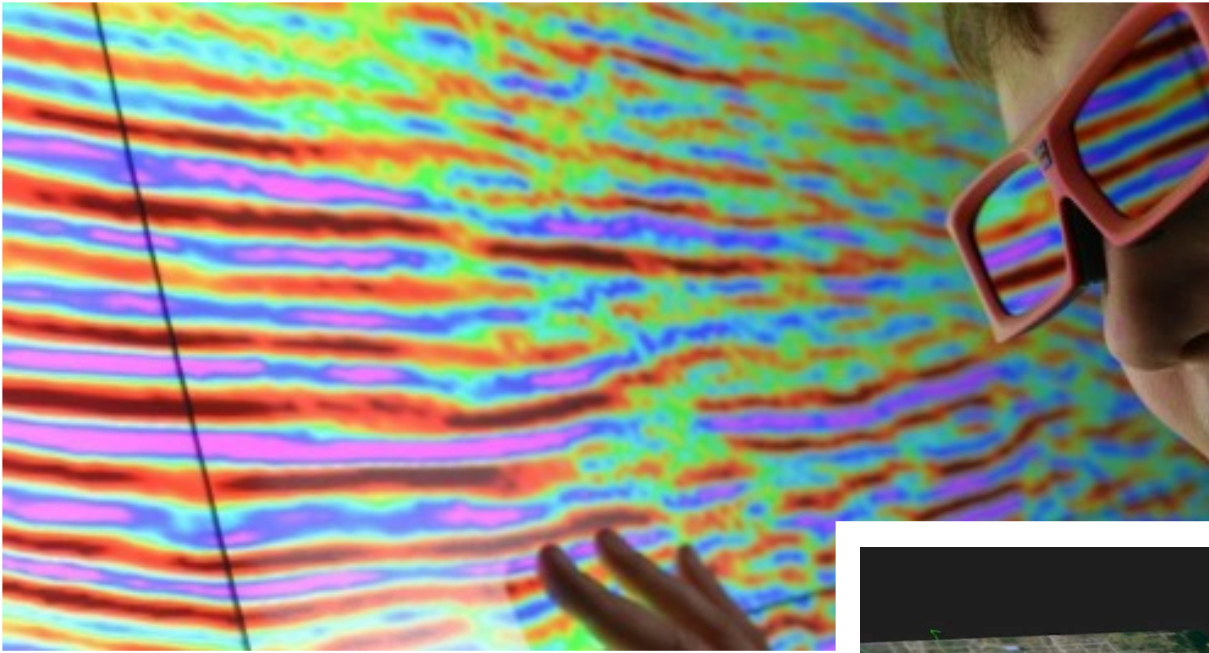Embodied control?

BUT - Balance between navigation and music making
  Music making is primary task
  XAI navigation is secondary task

$z = x \cdot y$

XAI + Music

Artificial Intelligence (AI)

Human Centred AI

eXplainable AI (XAI)

XAI+Music

XAI+Music Probotype

Open Research Questions

谢谢

Explainable AI + Music

Prof. Nick Bryan-Kinns

Queen Mary University of London, UK