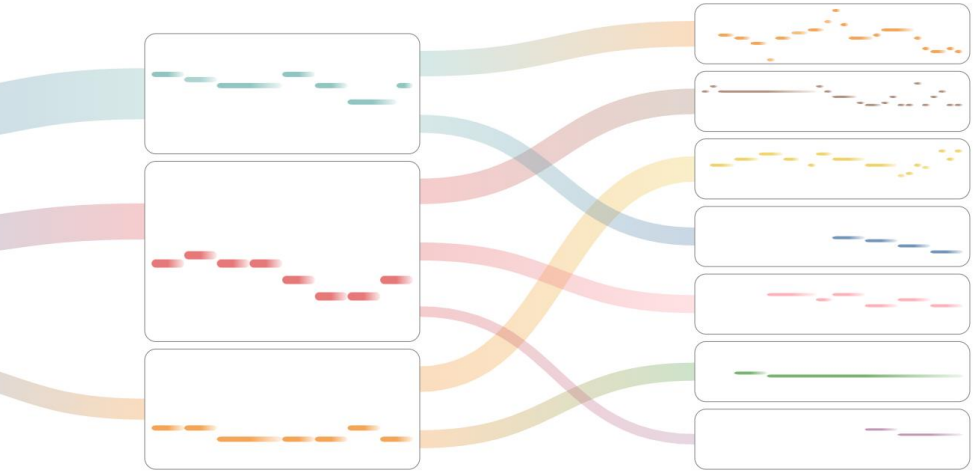
University of Stuttgart
Germany

Universität Stuttgart

AISA
Artificial Intelligence
Software Academy

Cyber Valley



Visualization for AI-Assisted Composing

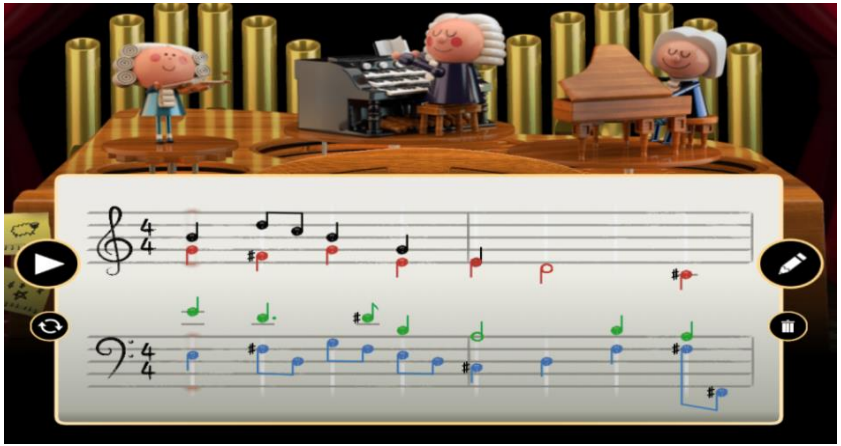
Simeon Rau[1], **Frank Heyen**[1], Stefan Wagner[2], Michael Sedlmair[1]

[1] VISUS, [2] ISTE, University of Stuttgart, Germany

ISMIR 2022, Bengaluru, India

Where are we currently?

- Many music generating AIs [1] exist, but hard to understand and control
- Users desire more control, authorship, and creative freedom [2]
- Current interactive tools lack (visual) exploration of choices
- Steering interfaces [3] can give more control



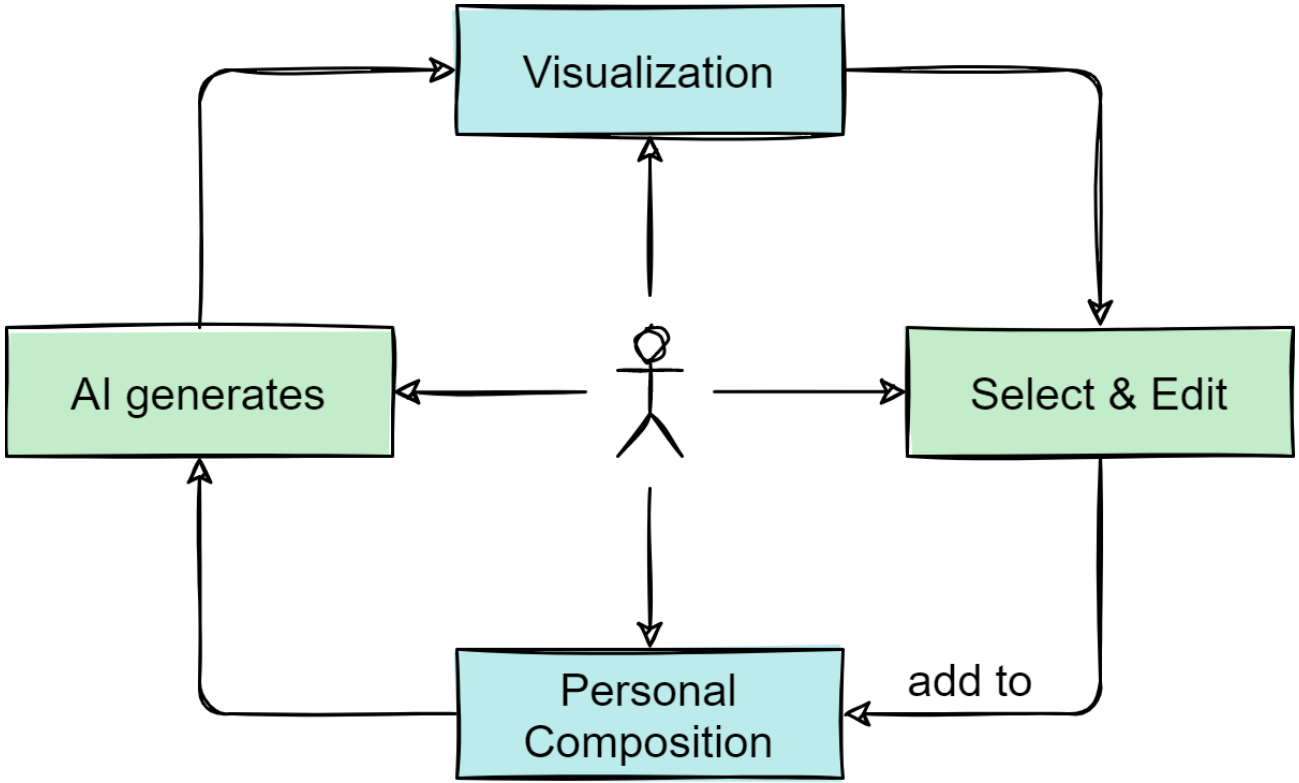
Bach Doodle [4]. Image source: Coconet: the ML model behind today's Bach Doodle, magenta.tensorflow.org/coconet

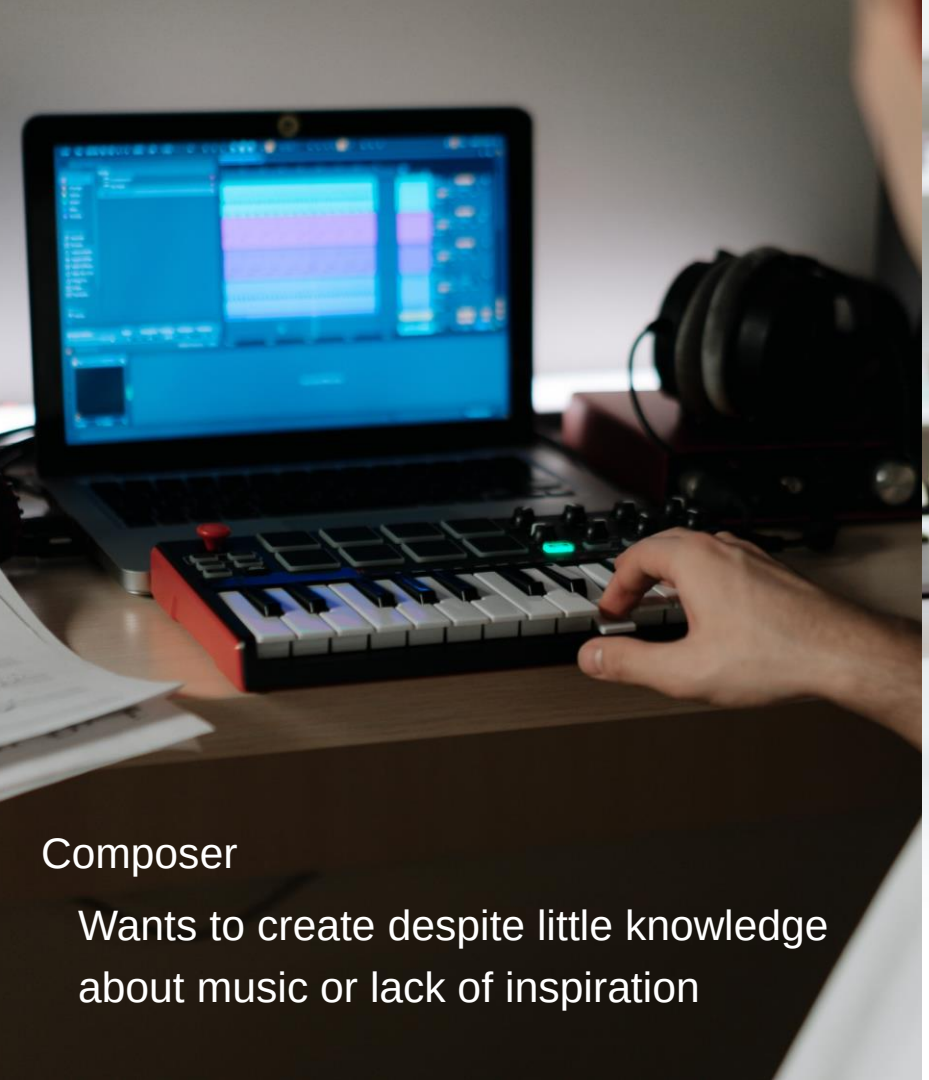
[1] Briot et al. *Deep learning techniques for music generation. (Springer)* 2020 **[2]** Huang et al. *AI song contest: Human-AI co-creation in songwriting.* (ISMIR) 2020 **[3]** Louie et al. *Expressive communication: A common framework for evaluating developments in generative models and steering interfaces.* (arXiv) 2021 **[4]** Huang et al. *The Bach Doodle: Approachable music composition with machine learning at scale.* (ISMIR) 2019

Our Goal

[MJx Music] Hip-Hop EP co-produced using Machine Learning, magenta.tensorflow.org/mj-hip-hop-ep

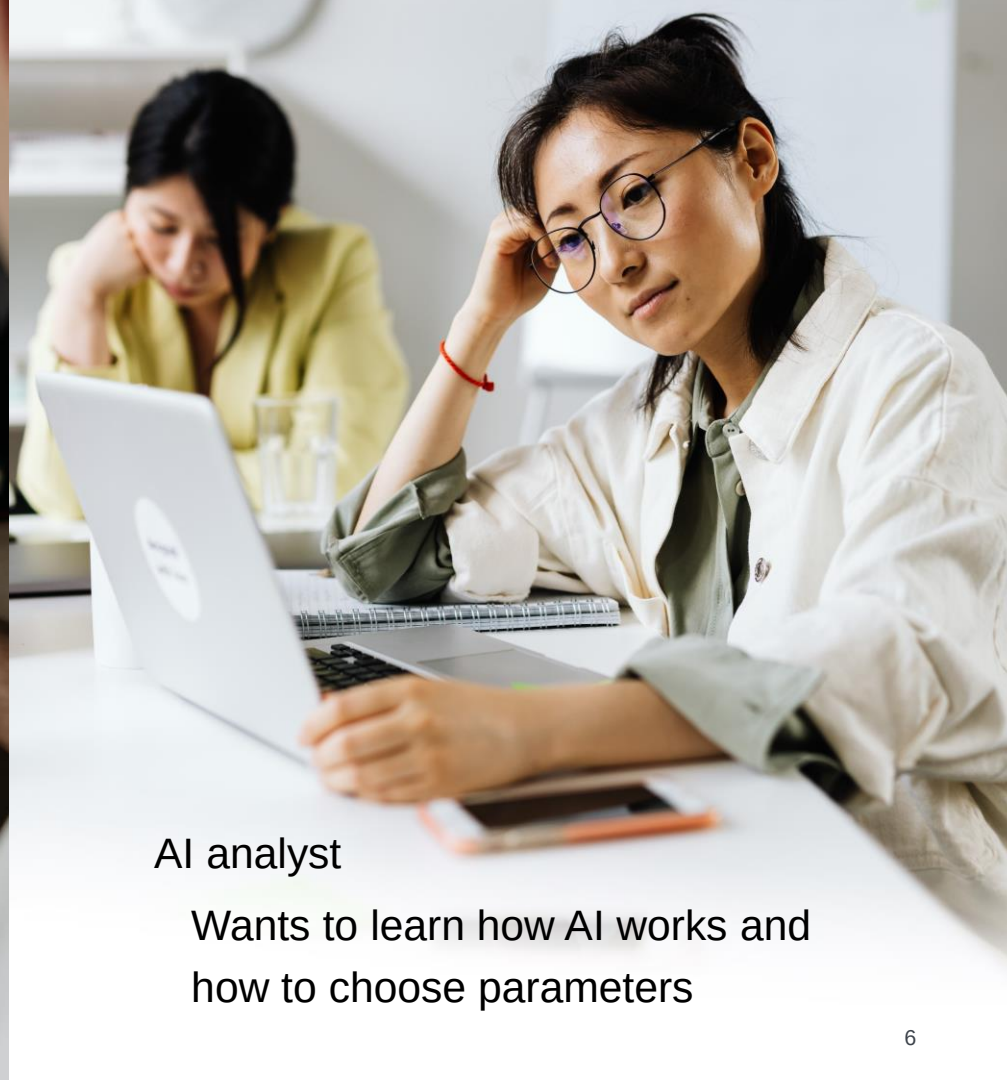
Our Idea


Visualization
AI generates
Select & Edit
Personal Composition
add to



Composer

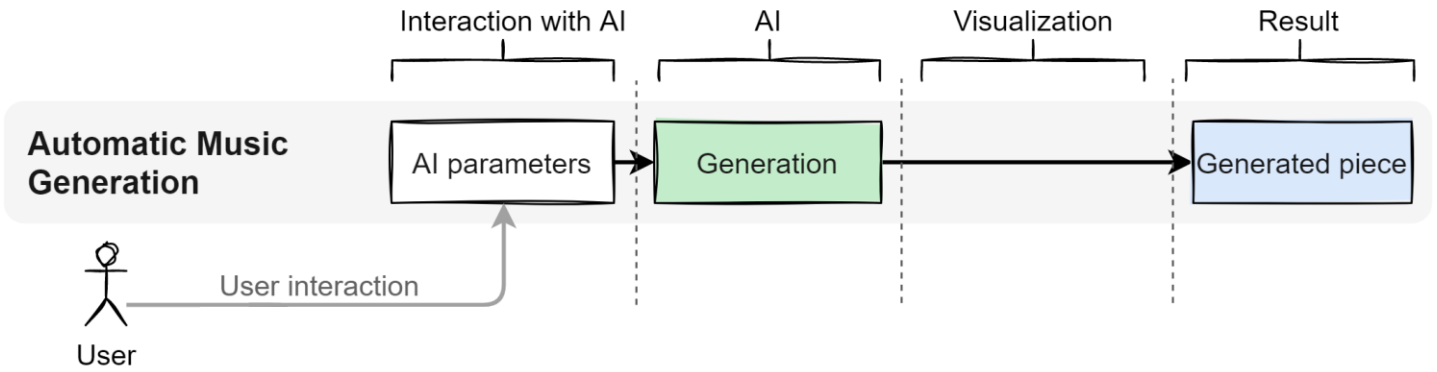
Wants to create despite little knowledge about music or lack of inspiration



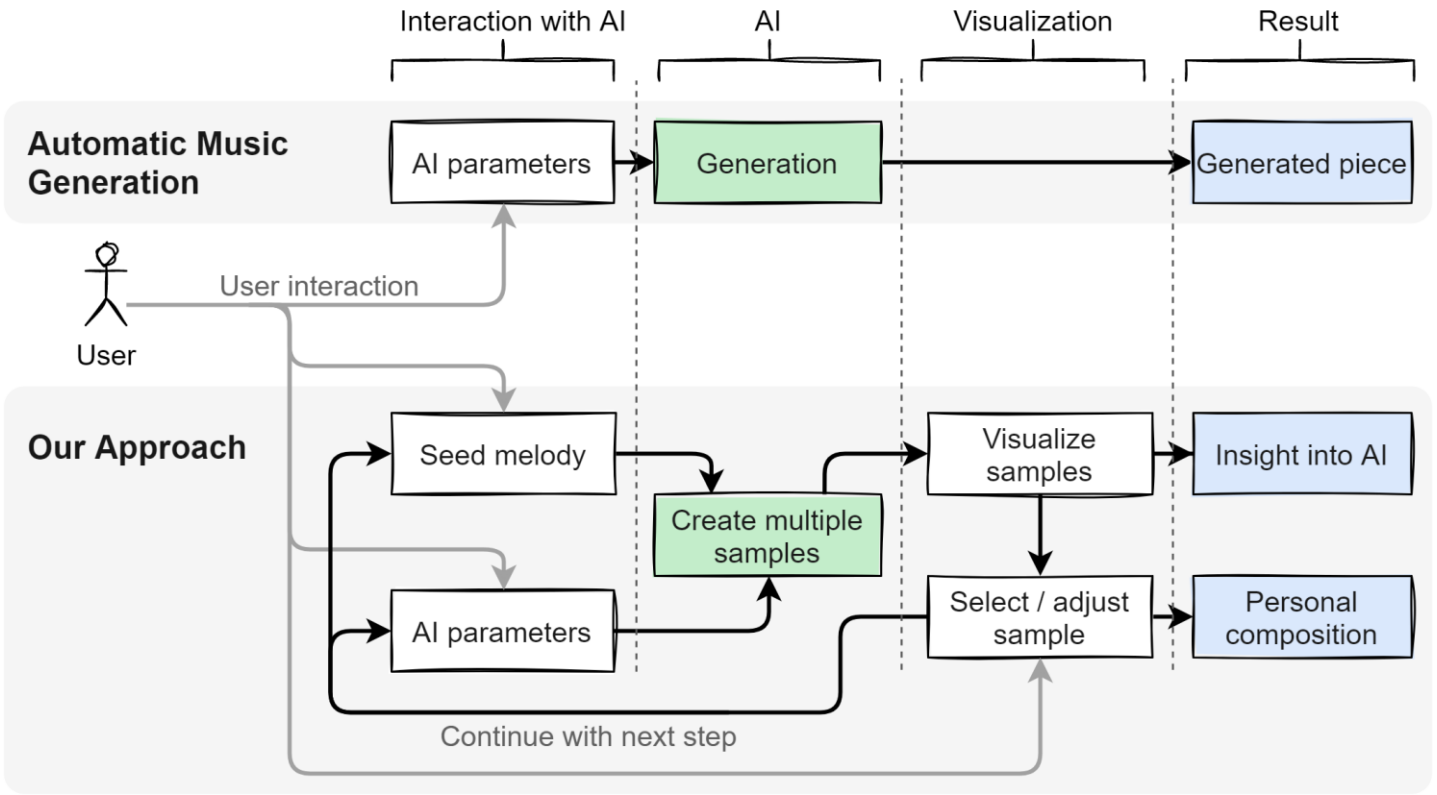
AI analyst

Wants to learn how AI works and how to choose parameters

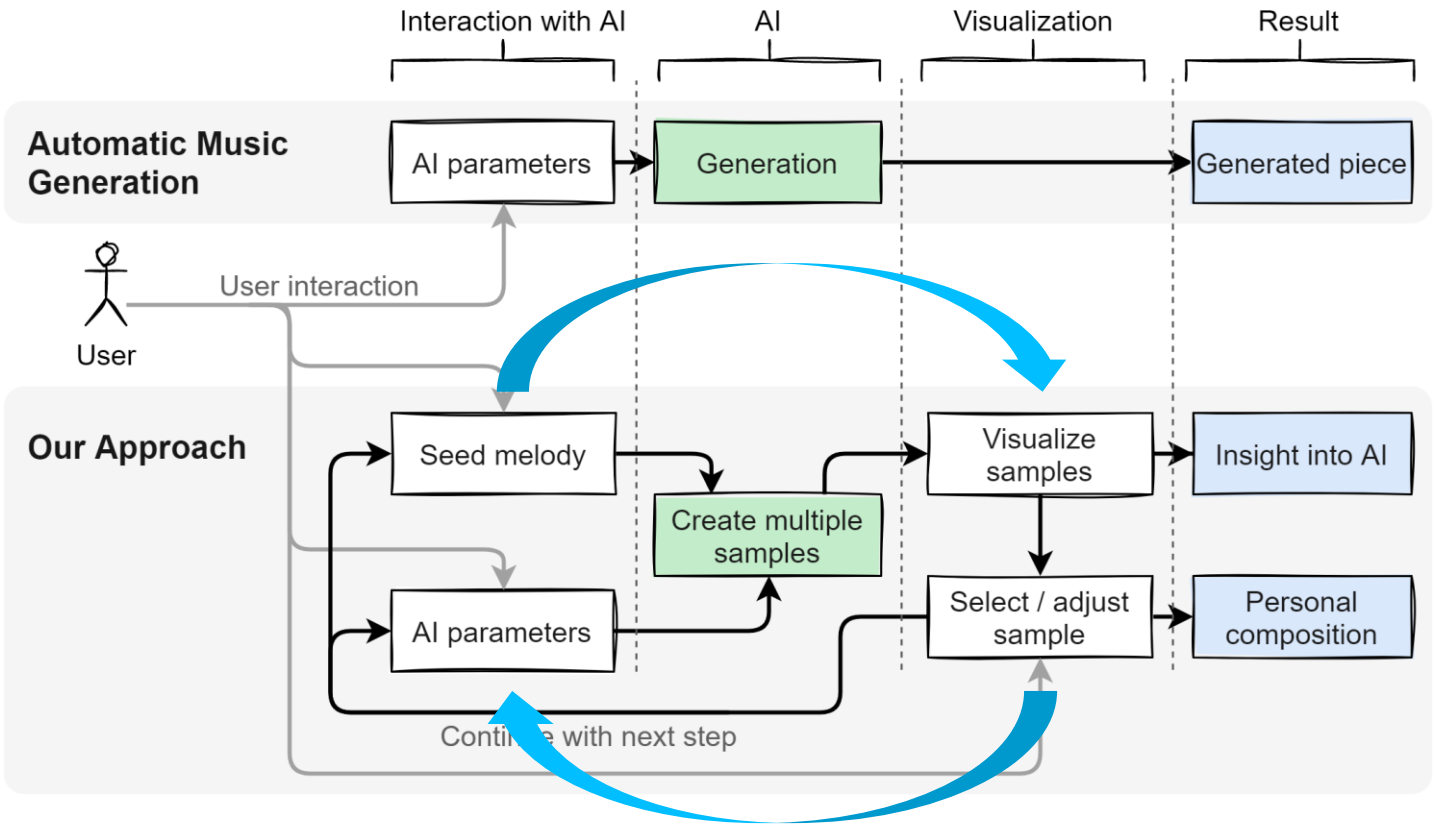
Workflow


Interaction with AI
AI
Visualization
Result
Automatic Music Generation
AI parameters
Generation
Generated piece
User interaction
User

Workflow


Interaction with AI
AI
Visualization
Result
Automatic Music Generation
AI parameters
Generation
Generated piece
User
User interaction
Our Approach
Seed melody
Create multiple samples
Visualize samples
Insight into AI
AI parameters
Select / adjust sample
Personal composition
Continue with next step

Workflow


Interaction with AI
AI
Visualization
Result
Automatic Music Generation
AI parameters
Generation
Generated piece
User
User interaction
Our Approach
Seed melody
Create multiple samples
Visualize samples
Insight into AI
AI parameters
Select / adjust sample
Personal composition
Continue with next step

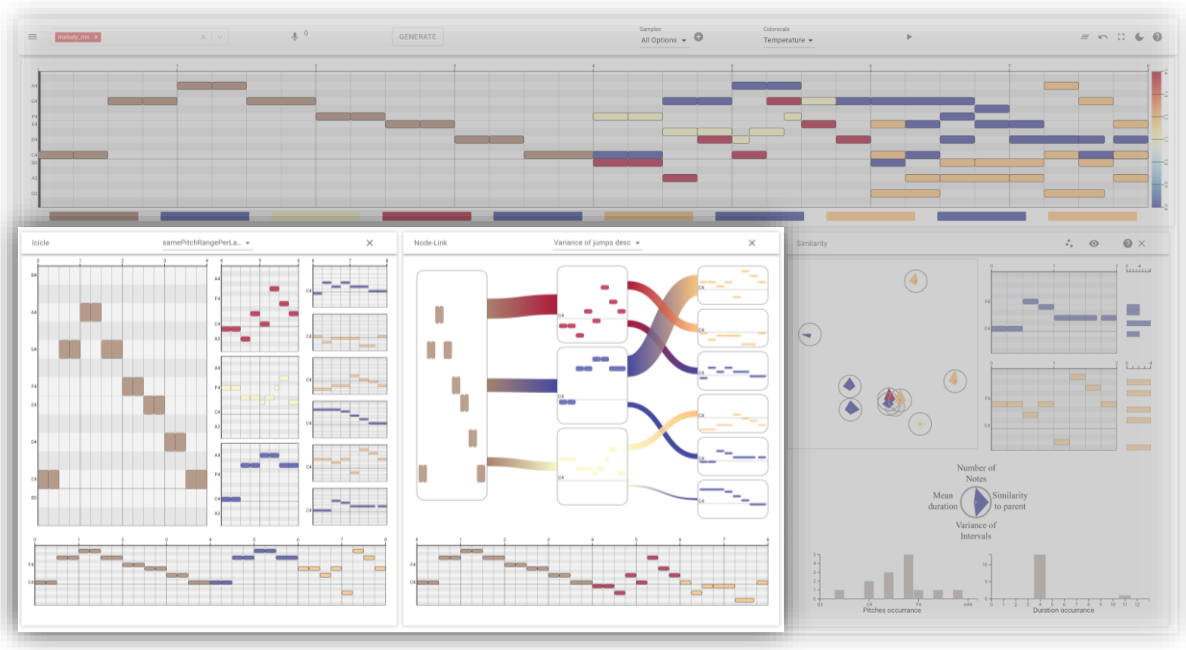
Design


GENERATE
Samples
All Options
Colorscale
Temperature
Icicle
samePitchRangePerLa...
Node-Link
Variance of jumps desc
Similarity
Number of Notes
Mean duration
Similarity to parent
Variance of Intervals
Pitches occurrance
Duration occurrance

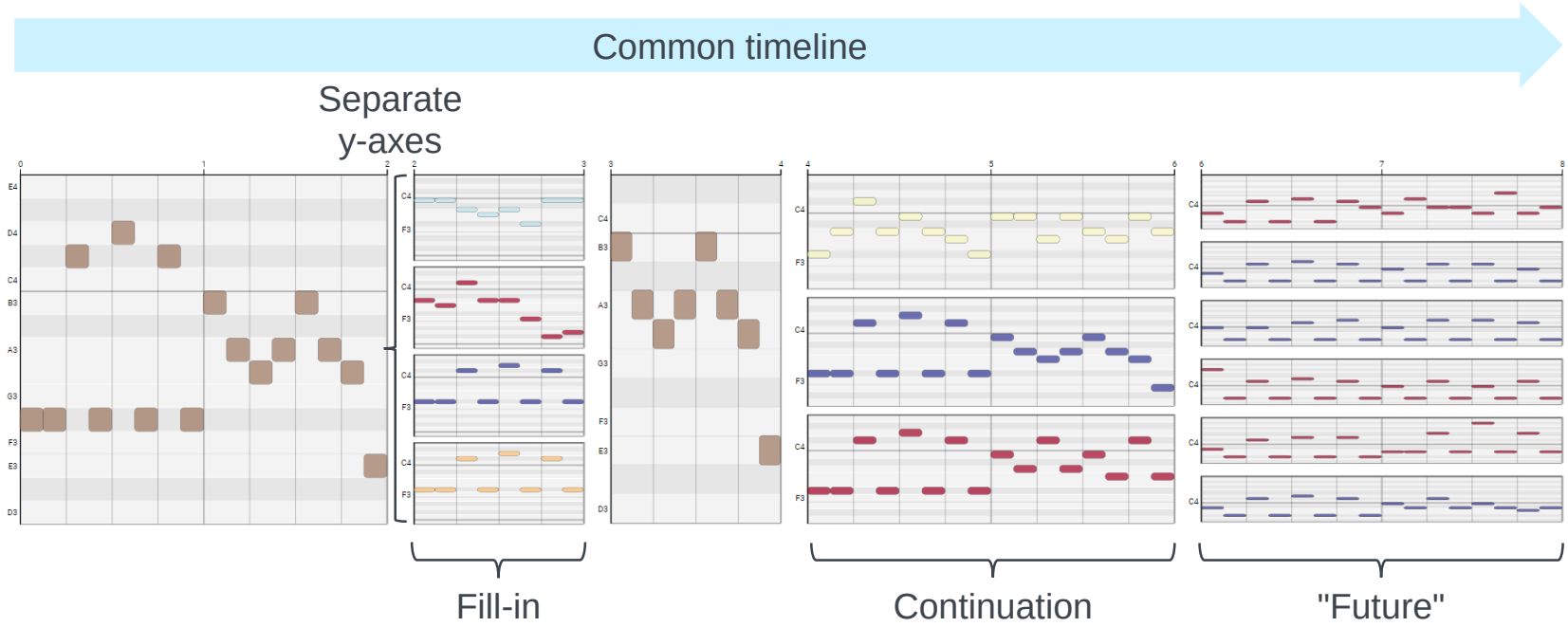
Design


GENERATE
Samples
All Options
Colorscale
Temperature
Icicle
Node-Link
Variance of jumps desc
Similarity
Number of Notes
Mean duration
Similarity to parent
Variance of Intervals
Pitches occurrance
Duration occurrance

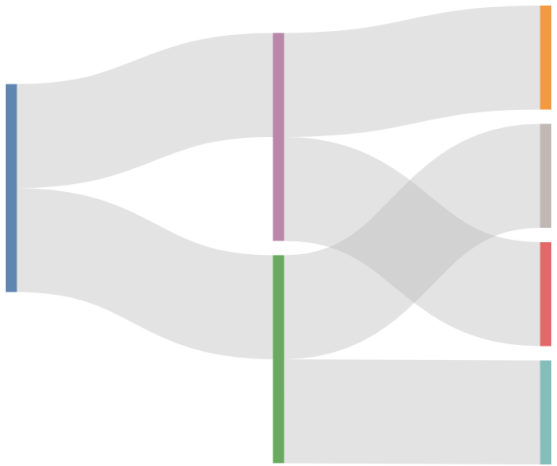
Design

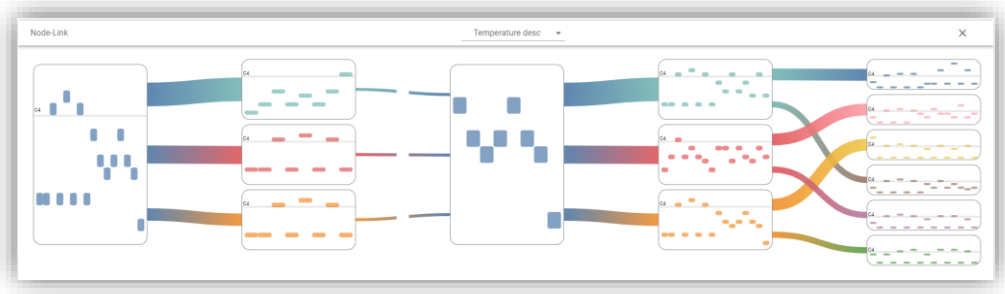

GENERATE
Samples
All Options
Colorscale
Temperature
Icicle
samePitchRangePerLa...
Node-Link
Variance of jumps desc
Similarity
Number of Notes
Mean duration
Similarity to parent
Variance of Intervals
Pitches occurrance
Duration occurrance

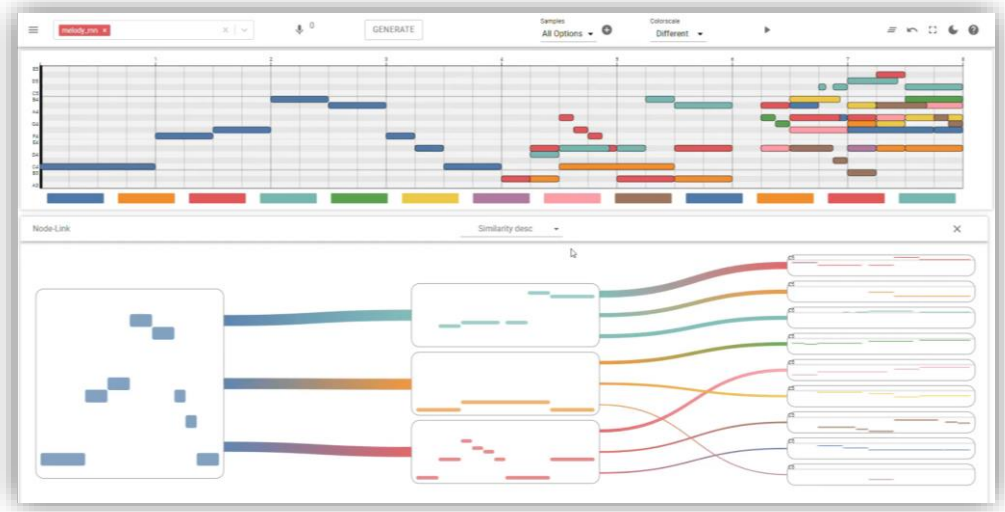
Icicle Plot


Common timeline
Separate
y-axes
Fill-in
Continuation
"Future"

Sorting



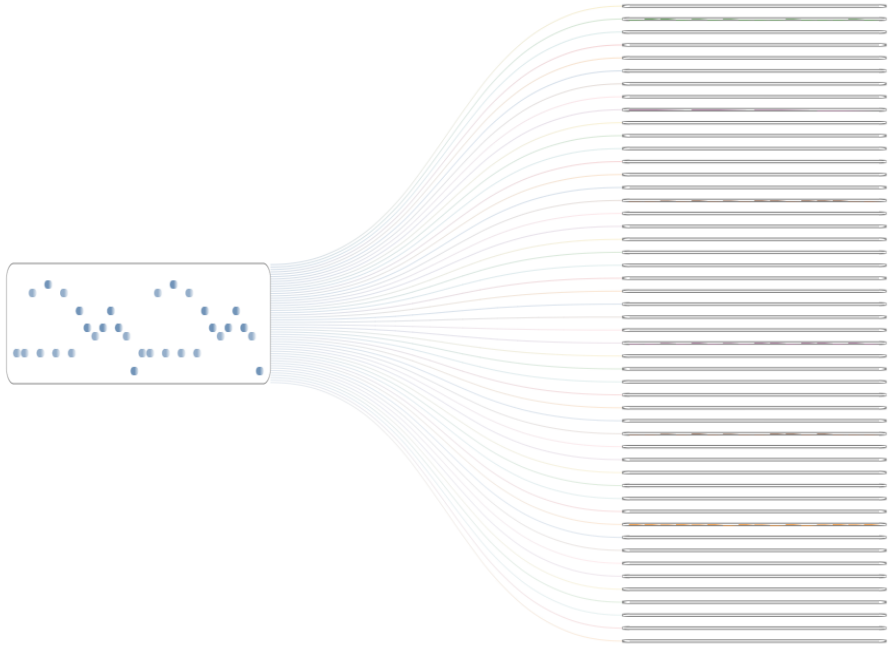

Node-Link
Temperature desc


GENERATE
Samples
All Options
Colorscale
Different
Node-Link
Similarity desc

Fill-In


melody_mm
GENERATE
Samples
All Options
Colorscale
Different
Node-Link
Similarity desc
Provenance
Icicle
Node-Link
Similarity
Correlation

What if we want to see more alternatives?

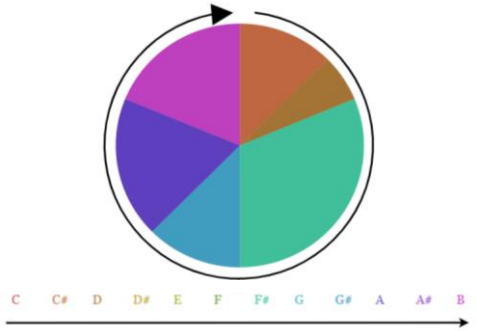


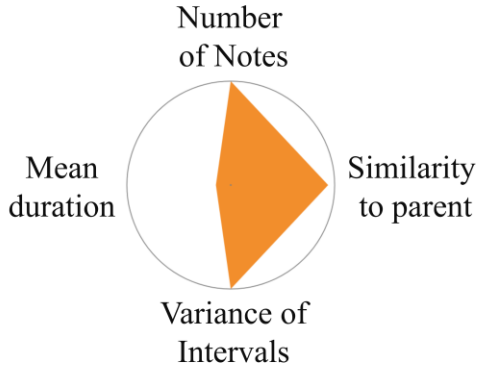
Does not scale!

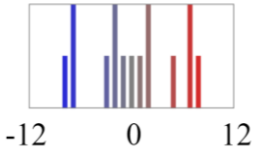
Design

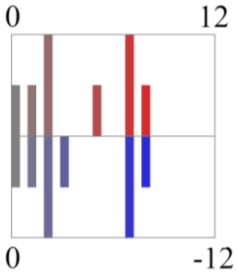

GENERATE
Samples
All Options
Colorscale
Temperature
Icicle
samePitchRangePerLa...
Node-Link
Variance of jumps desc
Similarity
Number of Notes
Mean duration
Similarity to parent
Variance of Intervals
Pitches occurrance
Duration occurrance

Glyphs

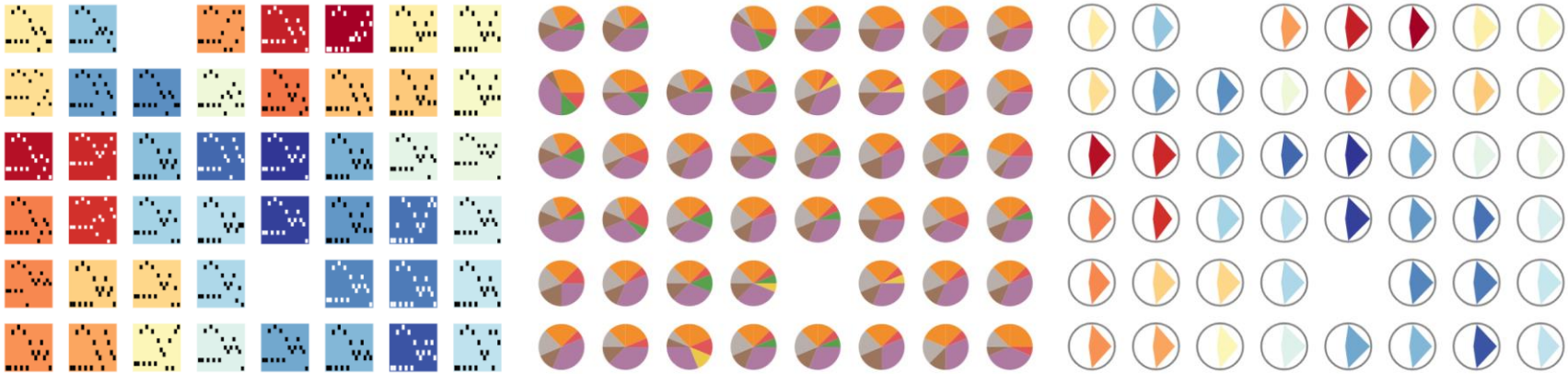

C C# D D# E F F# G G# A A# B


Number of Notes
Mean duration
Similarity to parent
Variance of Intervals


-12
0
12


0
12
0
-12

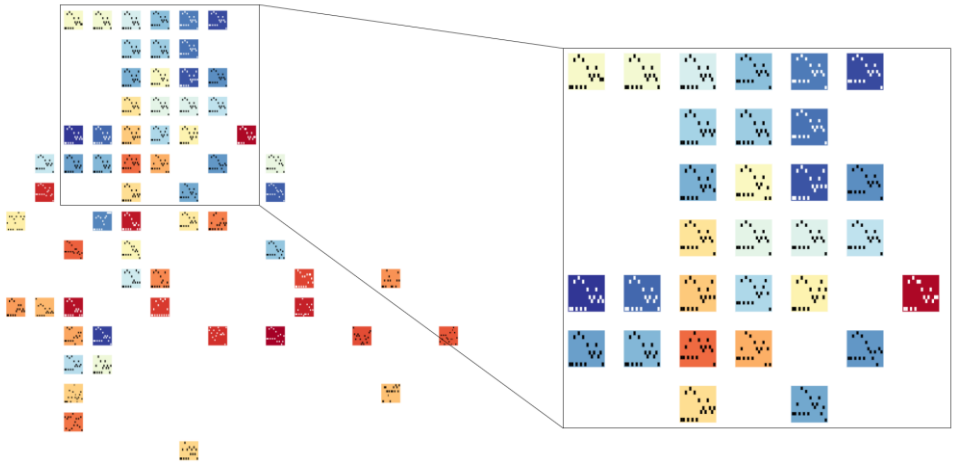
Glyphs

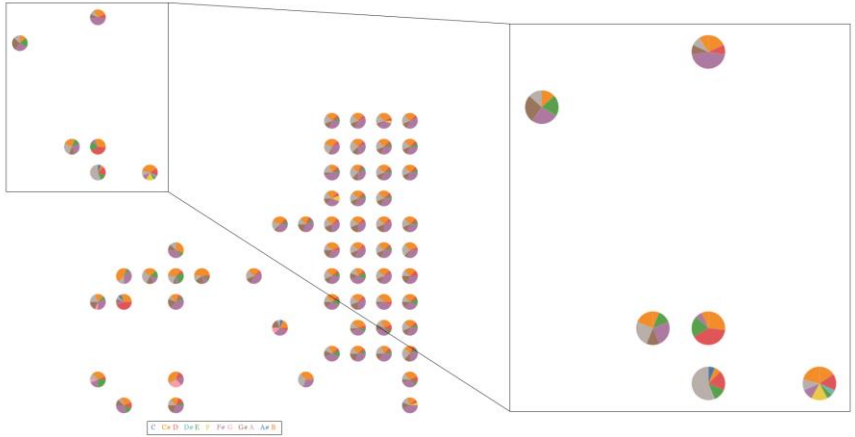


low temperature

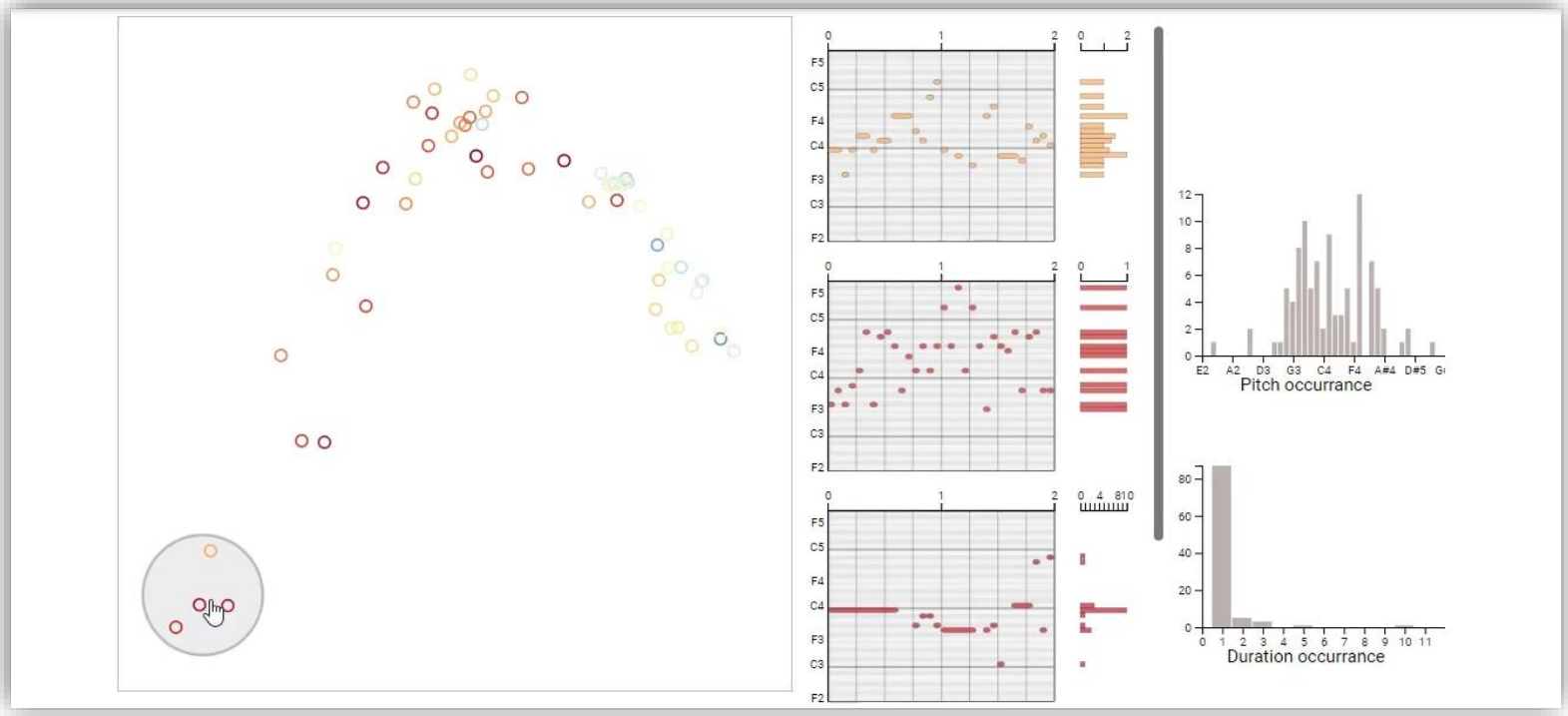
high temperature

Glyphs

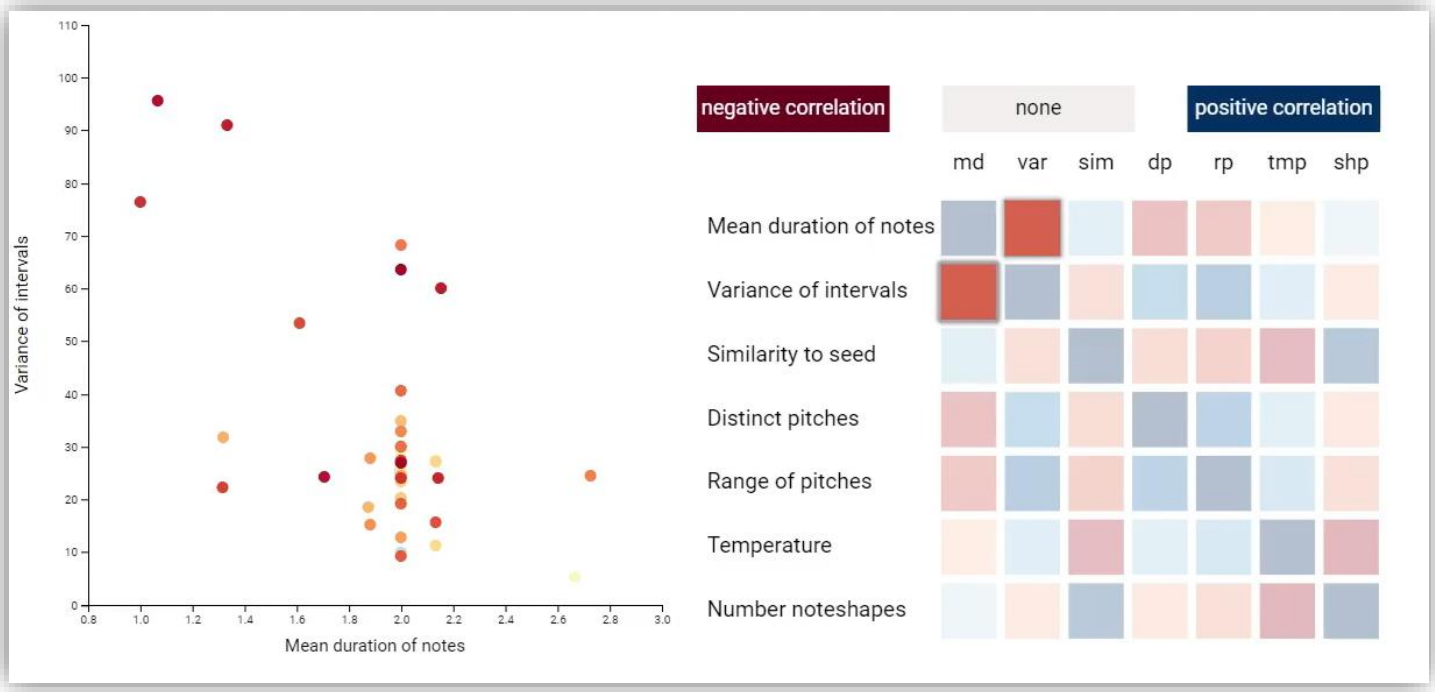




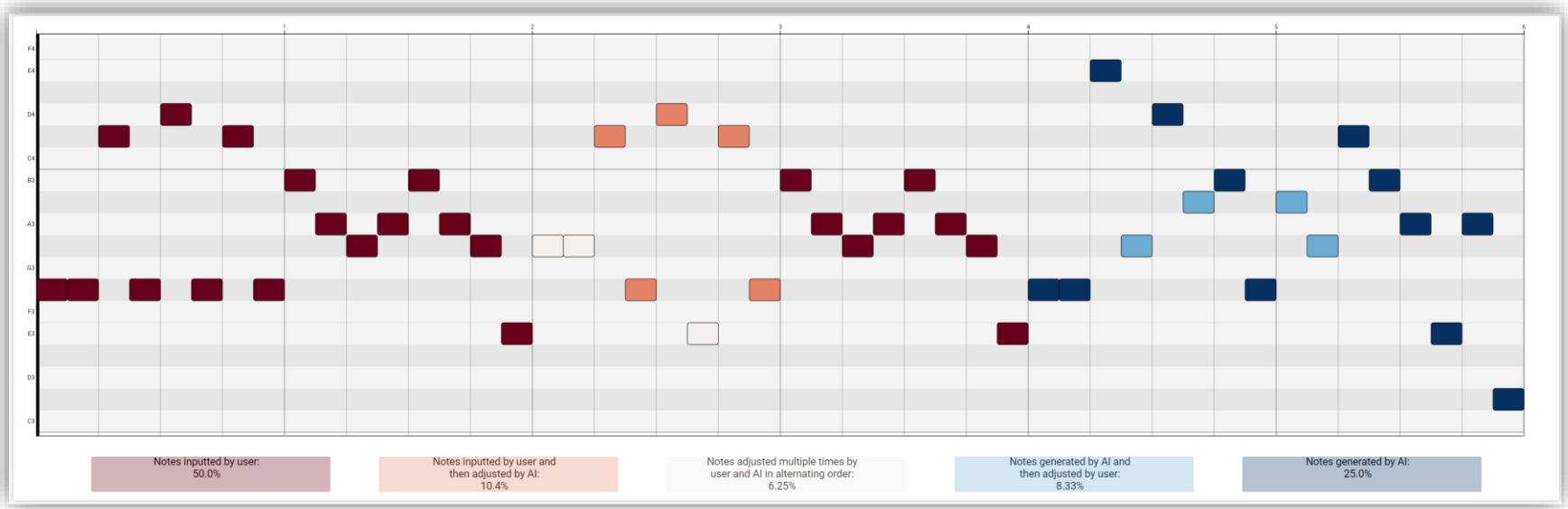
Scatterplot



Correlation



Authorship


Notes inputted by user:
50.0%
Notes inputted by user and then adjusted by AI:
10.4%
Notes adjusted multiple times by user and AI in alternating order:
6.25%
Notes generated by AI and then adjusted by user:
8.33%
Notes generated by AI:
25.0%

Evaluation

Pair analytics study with five composers as domain experts

"...very interesting as an analysis tool, even independent of the AI"

"As human caught in own preferences [...] AI produces parts we would have not even tried [...] If we think about something we would throw it away, but if we see and hear it, it could be interesting"

"...not used to look at the data"

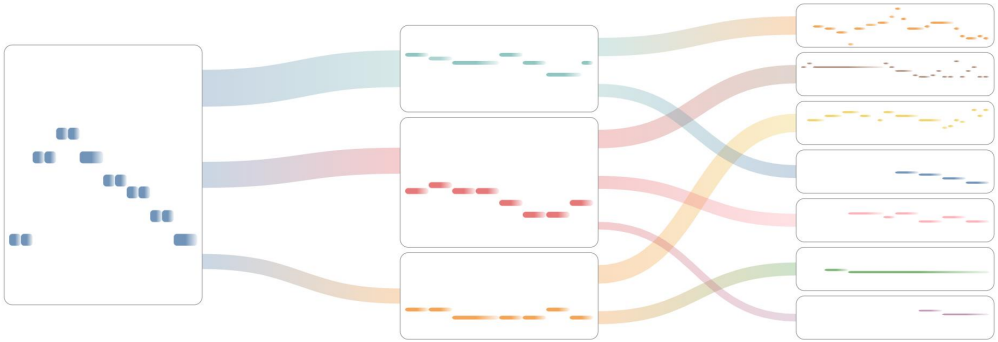
"...see more quickly which samples are interesting and which are not"

"I can see the same musical structure..."

- Skeptical but open towards using AI
- Saw potential in our AI-*assisted* approach
- Found visualizations helpful but unfamiliar

Conclusion & Future Work

- Composing is hard & complete automatization is not practical
- We propose interactive co-composition with AI using visualization
- Promising preliminary results with many future possibilities



More details and a live demo available!

VISUALIZATION FOR AI-ASSISTED COMPOSING

Simeon Rau[1*] Frank Heyen[1*] Stefan Wagner[2] Michael Sedlmair[1]

[1] VISUS, University of Stuttgart, Germany
[2] ISTE, University of Stuttgart, Germany
* Equal contribution

simeon.rau@visus.uni-stuttgart.de, frank.heyen@visus.uni-stuttgart.de

ABSTRACT

We propose a visual approach for interactive, AI-assisted composition that serves as a compromise between fully automatic and fully manual composition. Instead of generating a whole piece, the AI takes on the role of an assistant that generates short melodies for the composer to choose from and adapt. In an iterative process, the composer queries the AI for continuations or alternative fill-ins, chooses a suggestion, and adds it to the piece. As listening to many suggestions would take time, we explore different ways to visualize them, to allow the composer to focus on the most interesting-*looking* melodies. We also present the results of a qualitative evaluation with five composers.

1. INTRODUCTION

Composing music is a challenging task, especially for beginners. Complex music theory, different rules and patterns in a wide range of genres, missing experience, or even stagnating inspiration make it difficult to express intended ideas and emotions. These aspects can result in frustration and leave the composer unsatisfied or even unable to finish a piece. Artificial intelligence (AI) might potentially mitigate such situations by generating music that provides inspiration or even fits the composer's needs directly.

Current AI-driven approaches for music generation allow creating parts or whole pieces in audio [1] or symbolic form [2]. Although automatic composition can be useful, for example as background music in videos or games [3], it often lacks personality and structure [3, 4]. A fine-grained steering of the generation process, beyond parameters such as overall tempo and feel, is often not possible. Therefore, a piece that is fully generated by AI might not satisfy expectations [5]. Although AI-generated music can show potential creativity [6], the current state of AI music will most likely not replace human composers as a whole [5].

To address above problems, we propose a user-centered approach for AI-assisted composing. Here, users have more control, acting as a leading composer who always has the last word, while the AI serves as assistant that supports composers instead of replacing them [4, 5, 7]. When assisted with suggestions for continuations, fill-ins, and replacements, users can iteratively elaborate on these and make progress towards a personal creation. Composing step by step often yields better results than selecting from multiple completely generated ones and gives a greater feeling of satisfaction and authorship [8].

To support the communication between AI and human, we propose using interactive visualization. As it allows to quickly spot and filter the most interesting suggestions, visualization facilitates choice. This means users do not have to spend time sifting through heaps of less interesting suggestions and can spend their time more efficiently, for example by only listening to a few melodies per group of similar ones. Furthermore, visualization can improve understanding of how the underlying AI works [9], and allow for a more effective usage and better steering of the AI.

We make two primary contributions: (1) We designed four interactive visualizations to assist the user in choosing and investigating AI-generated suggestions. The first two represent a graph structure of suggestions [10–12], while the third displays larger numbers of melody samples, by encoding their similarities and characteristics. Our fourth visualization shows the correlation between melody samples. (2) We evaluated our approach through a qualitative study with five composers, who generally liked our approach and were curious in analyzing the AI and using visualizations to find interesting samples. Results show that our representations were unfamiliar to most participants, but made interaction with the AI more accessible.

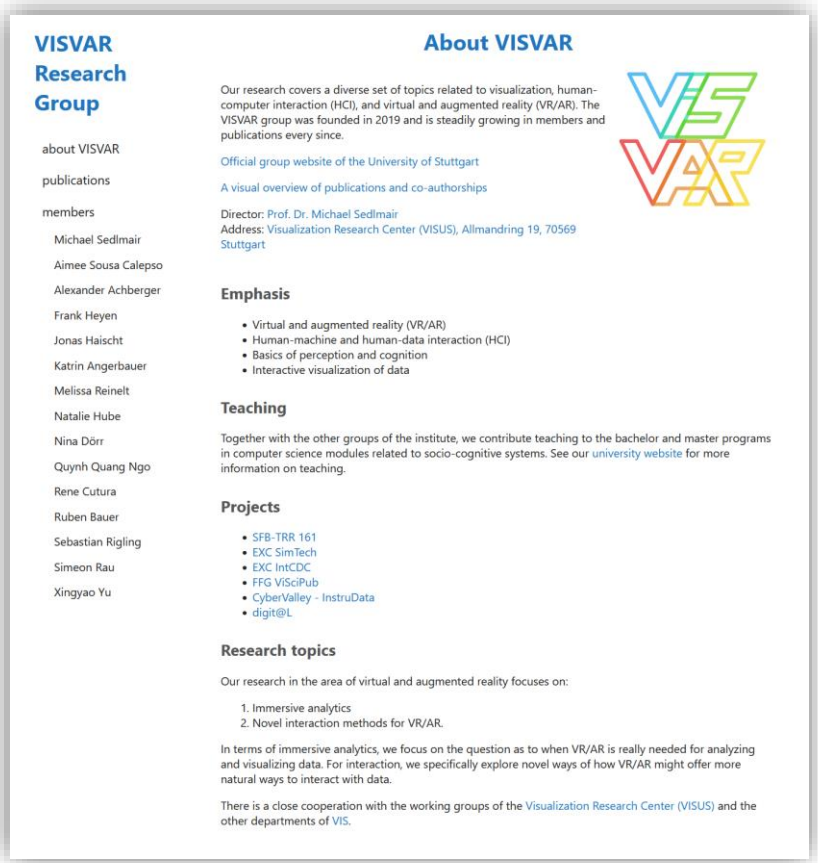
Our supplemental material [1] contains a live version of our prototype, the full source code, and additional details.

2. RELATED WORK

Briot et al. [13] surveyed deep learning for music generation and compare objectives, data representations, architectures, and the output that consists of audio [1] or symbolic data [14]. They cover recurrent neural networks [2, 15], variational autoencoders [16–18], and generative adversarial networks [19, 20]. Other work [14, 21] shows combinations of above techniques for different tasks. Recently, transformers [22–24] showed promising results and there also exist agent-based or heuristic algorithms [25].

[1] github.com/visvar/vis-ai-comp

More research on our group's website.



VISVAR Research Group

about VISVAR

publications

members

Michael Sedlmair
Aimee Sousa Calepso
Alexander Achberger
Frank Heyen
Jonas Haischt
Katrin Angerbauer
Melissa Reinelt
Natalie Hube
Nina Dörr
Quynh Quang Ngo
Rene Cutura
Ruben Bauer
Sebastian Rigling
Simeon Rau
Xingyao Yu

About VISVAR

Our research covers a diverse set of topics related to visualization, human-computer interaction (HCI), and virtual and augmented reality (VR/AR). The VISVAR group was founded in 2019 and is steadily growing in members and publications every since.

Official group website of the University of Stuttgart

A visual overview of publications and co-authorships

Director: Prof. Dr. Michael Sedlmair
Address: Visualization Research Center (VISUS), Allmandring 19, 70569 Stuttgart

Emphasis

- Virtual and augmented reality (VR/AR)
- Human-machine and human-data interaction (HCI)
- Basics of perception and cognition
- Interactive visualization of data

Teaching

Together with the other groups of the institute, we contribute teaching to the bachelor and master programs in computer science modules related to socio-cognitive systems. See our university website for more information on teaching.

Projects

- SFB-TRR 161
- EXC SimTech
- EXC IntCDC
- FFG ViSciPub
- CyberValley - InstruData
- digit@L

Research topics

Our research in the area of virtual and augmented reality focuses on:

1. Immersive analytics
2. Novel interaction methods for VR/AR.

In terms of immersive analytics, we focus on the question as to when VR/AR is really needed for analyzing and visualizing data. For interaction, we specifically explore novel ways of how VR/AR might offer more natural ways to interact with data.

There is a close cooperation with the working groups of the Visualization Research Center (VISUS) and the other departments of VIS.

Simeon Rau – Simeon.Rau@visus.uni-stuttgart.de

Frank Heyen – Frank.Heyen@visus.uni-stuttgart.de