Computer-Assisted Music Making Systems: Taxonomy, Review, and Live Coding

ISMIR 2023 Tutorial

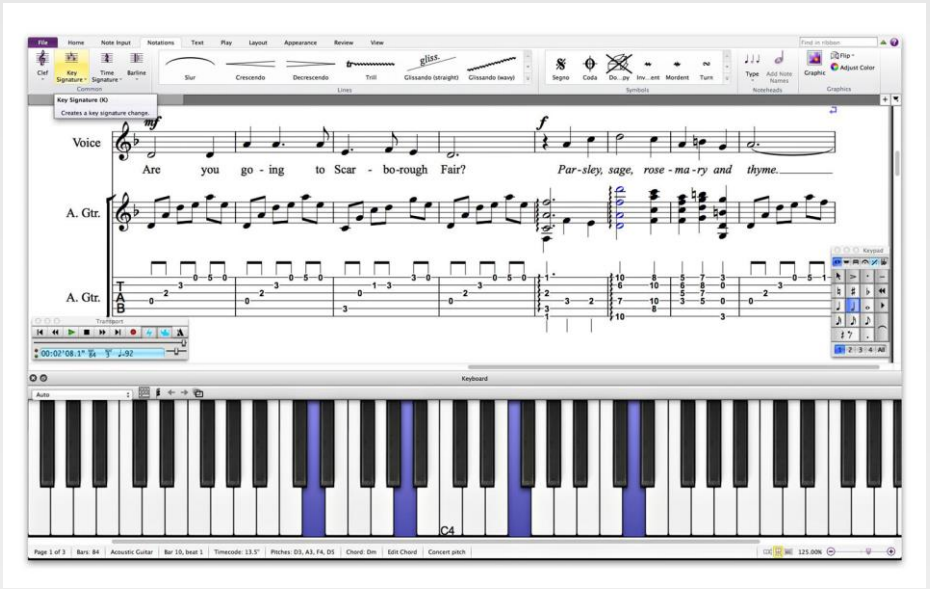
Christodoulos Benetatos †
cbenetat@ur.rochester.edu

Philippe Pasquier ‡
pasquier@sfu.ca

Zhiyao Duan †
zhiyao.duan@rochester.edu

† Audio Information Research Lab, University of Rochester, NY, USA
‡ Metacreation Lab, Simon Fraser University, BC, Canada

Music Making Tools



Music Making Tools

Credit: Alasabyss/Getty Images; Kelso Harper/Scientific American

https://www.scientificamerican.com/podcast/episode/artificial-intelligence-helped-make-the-coolest-song-youve-heard-this-week/ (March, 2023)

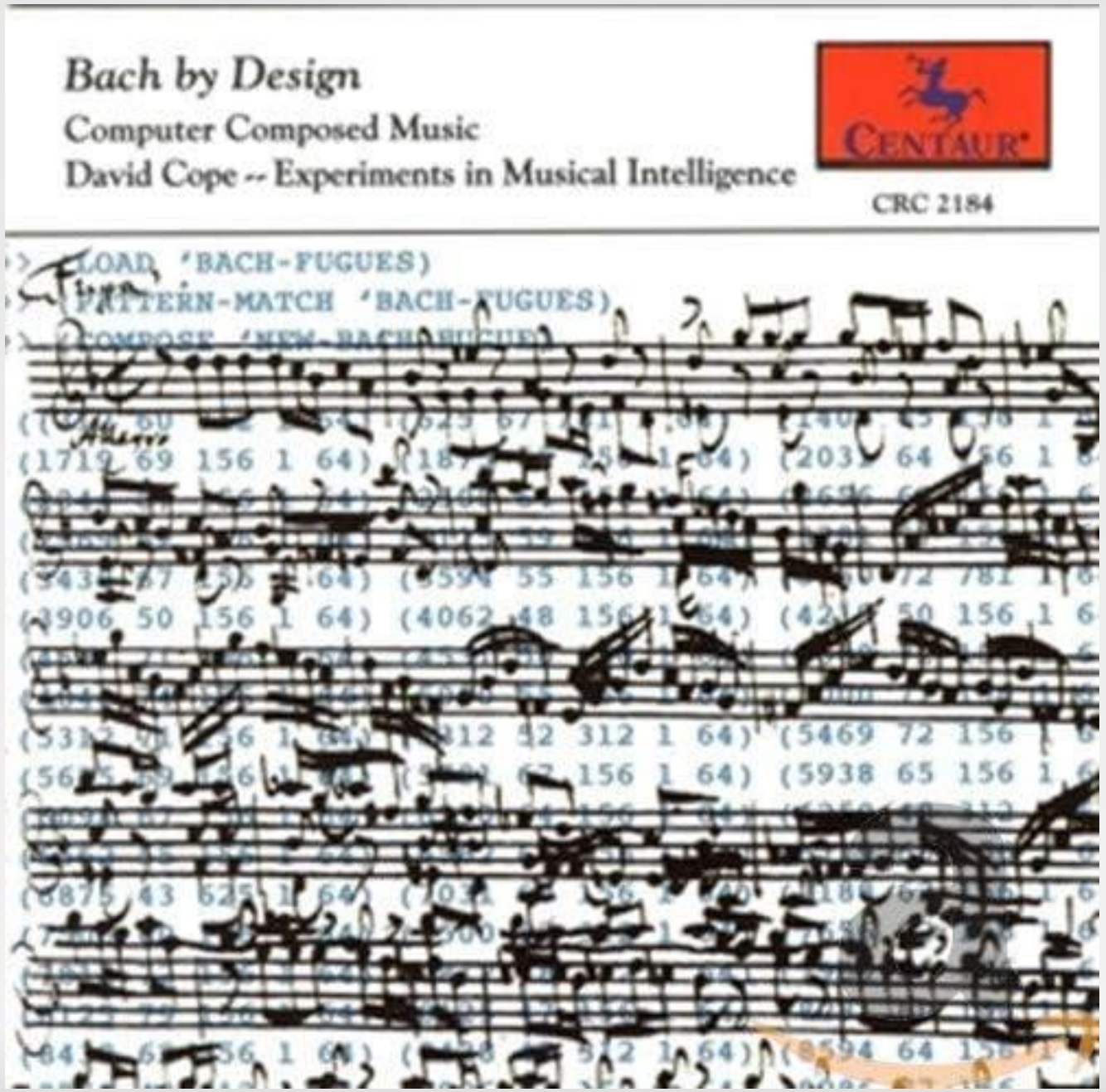
The world's first AI-composed music album is here, and it sounds amazing

CATCH TEAM | Updated on: 22 August 2017, 19:36 IST

Follo Follow @catchnews

Taryn Southern
Electronic Dance Music

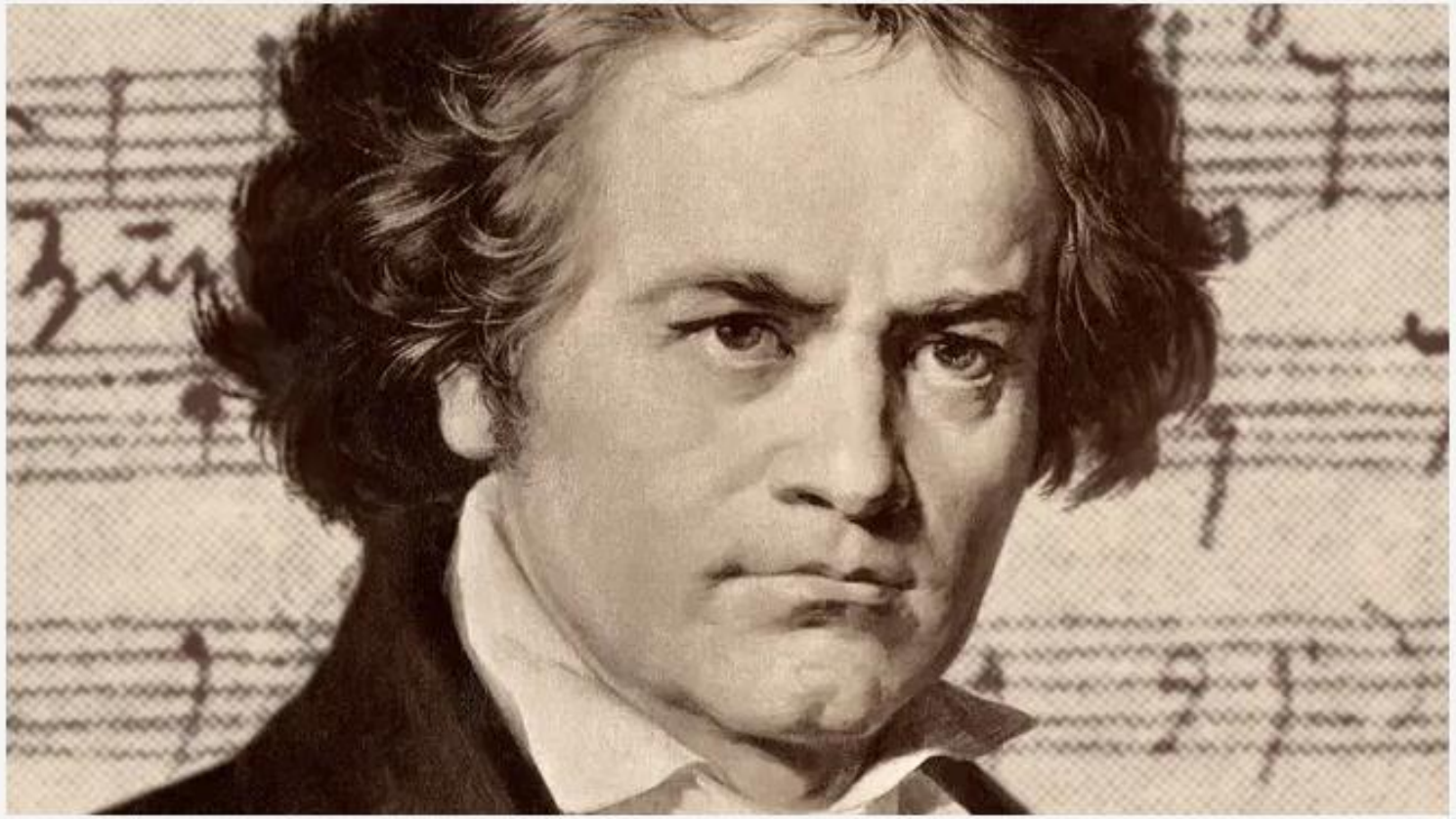
https://www.catchnews.com/entertainment-news/the-first-ai-composed-music-album-is-here-and-it-sounds-amazing-78436.html



David Cope's Experiments in Musical Intelligence (EMI)
"Bach by Design"
1993

Beethoven's unfinished Tenth Symphony completed by artificial intelligence

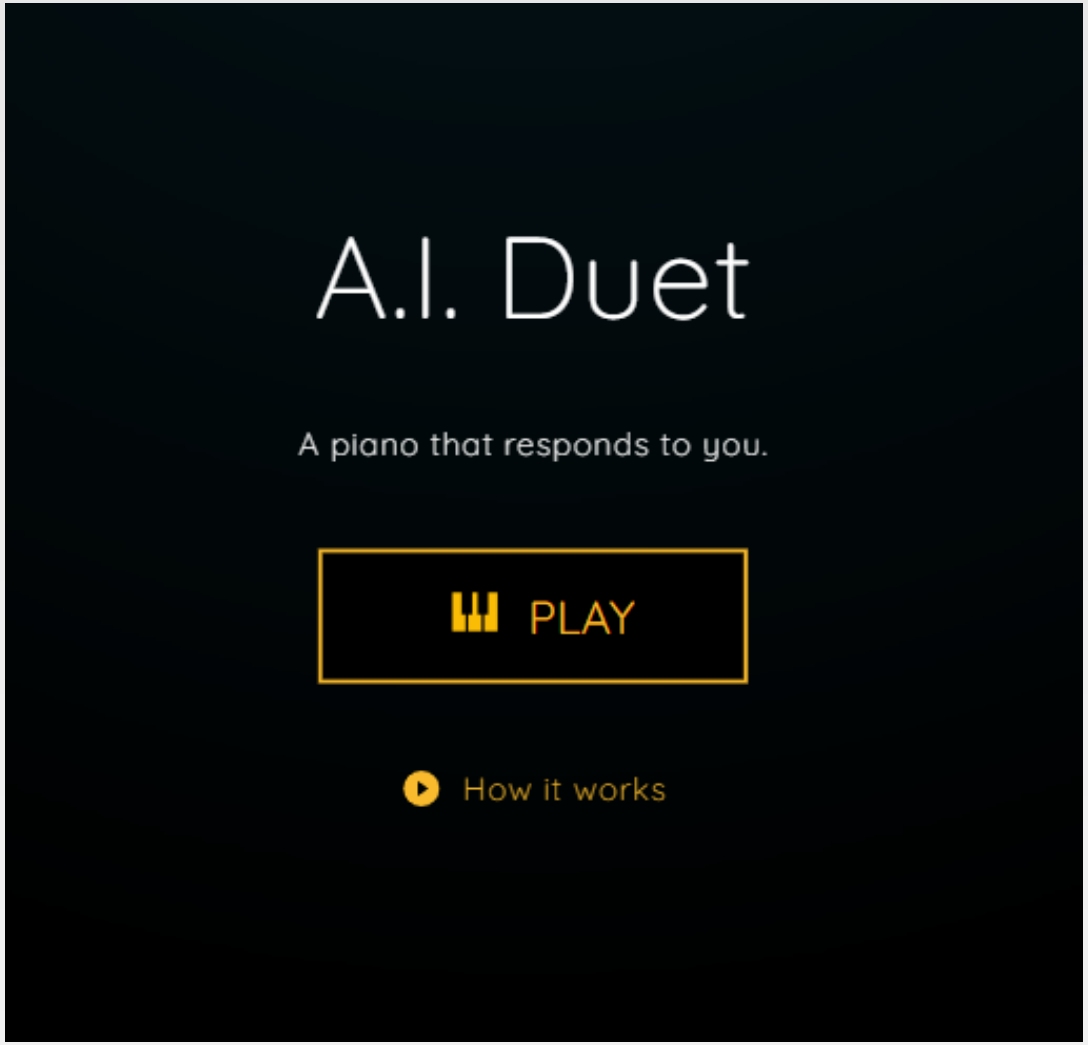
28 September 2021, 14:44 | Updated: 28 September 2021, 16:10



Beethoven's Tenth Symphony completed by AI. *Picture: Alamy*

https://www.classicfm.com/composers/beethoven/unfinished-tenth-symphony-completed-by-artificial-intelligence/

https://www.aisongcontest.com/

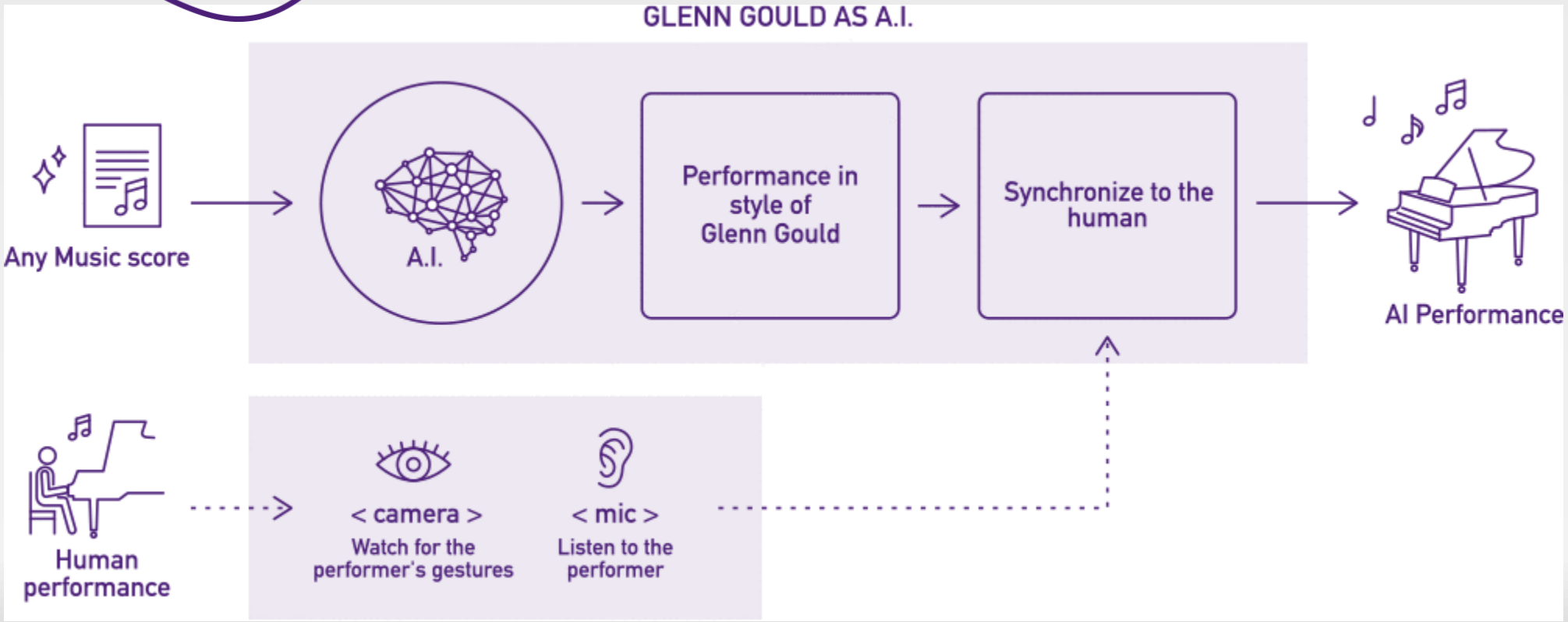


https://experiments.withgoogle.com/ai/ai-duet/view/

7 SEP. 2019

GLENN GOULD AS A.I.

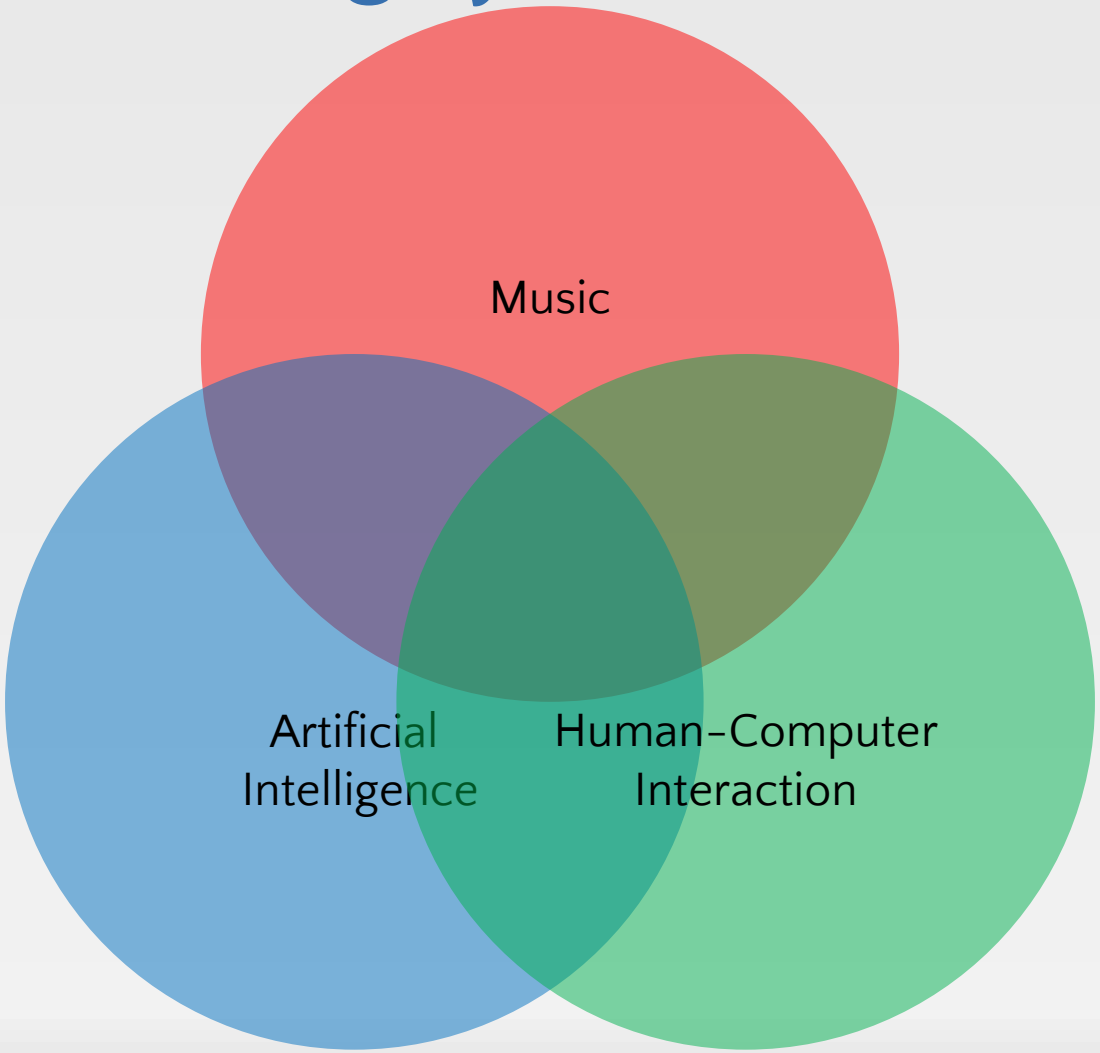
YAMAHA's Music with AI Project



https://www.yamaha.com/en/about/ai/dear_glenn/

https://www.shimonrobot.com/gallery Gil Weinberg's Shimon robot at Georgia Tech

Why is building computer-assisted music making systems interesting?



- Interdisciplinary
- It integrates research from multiple subareas of music information retrieval
  - Music analysis
    - Audio
    - Symbolic
    - Visual
  - Music generation
  - Music audio/visual/gesture synthesis
  - Interaction design and user Studies

Why are computer-assisted music making systems useful?

Music Creation

- Frees users from tedious work
- Lowers cost
- Improves accessibility

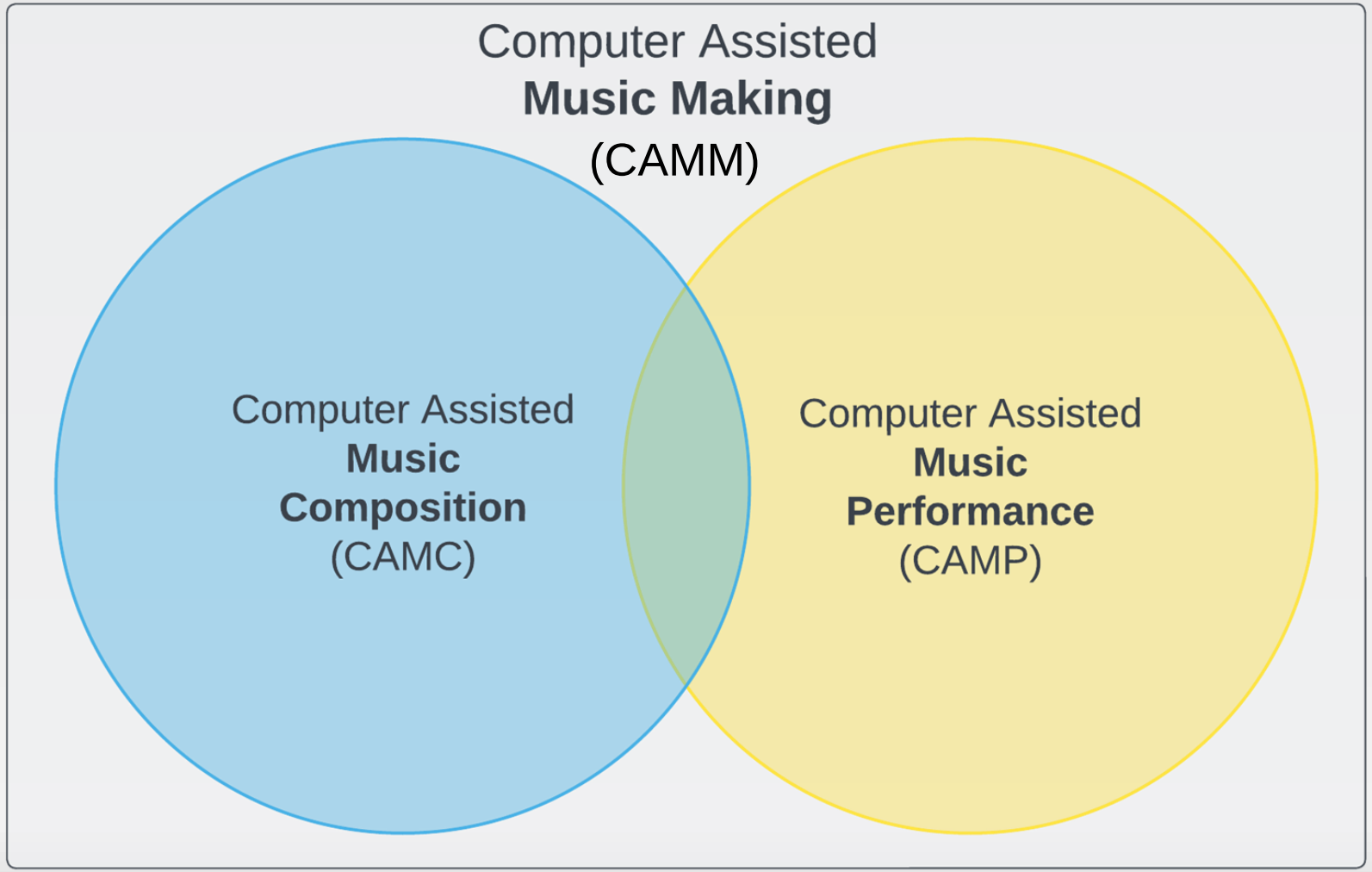
Music Education

- Composition
- Arrangement
- Improvisation
- Ensemble performance

Music Entertainment

- Novel content
- Various formats
- Anytime/anywhere

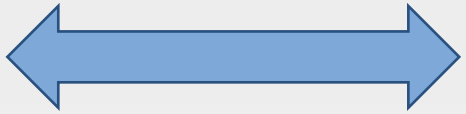
Two Categories



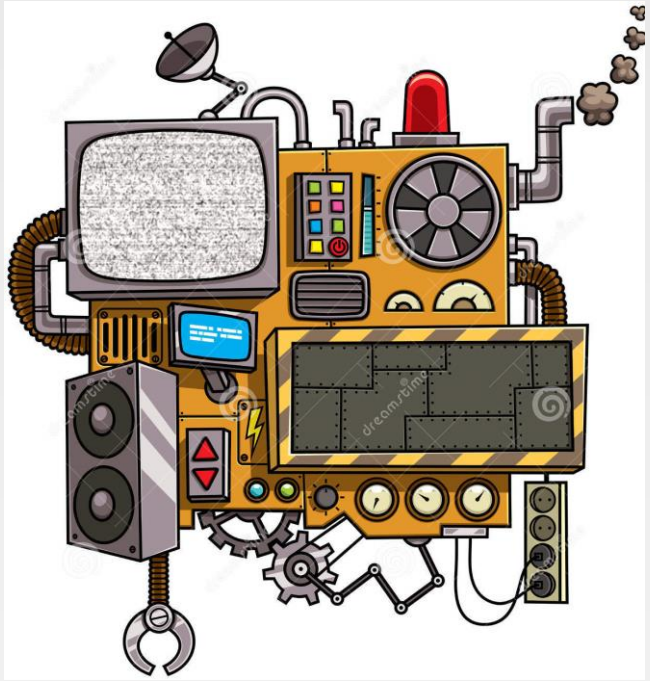
Computer-Assisted Music Making (CAMM) systems · · Interactive Music Systems

Human

Exchange of musical information



through some media (e.g., audio, visual, touch, gesture, brain signals)

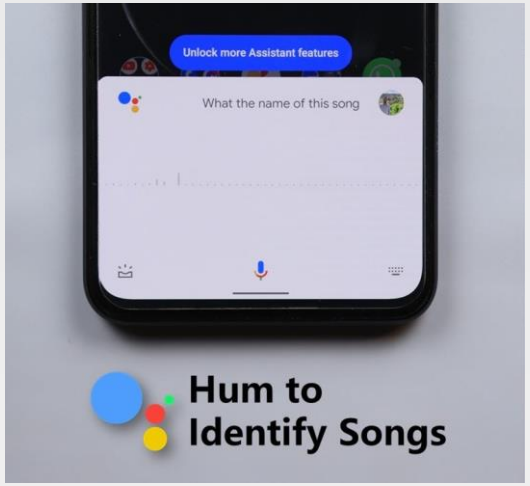


System (e.g., instrument, machine, agent, robot)

Are they interactive music systems?

Little interactivity

No intelligence



Not for fun

Interaction Needs to be the Primary Goal

"One does not 'use' an instrument to accomplish some ultimate goal: one plays it, and often that is the only goal."

---- McDermott, J., Gifford, T., Bouwer, A., & Wagy, M. (2013a). Should music interaction be easy? In S. Holland, K. Wilkie, P. Mulholland, & A. Seago (Eds.), Music and human computer interaction (pp. 29–48). London: Springer.

Goals of the Tutorial

- Provide high-level overview and taxonomy of research
- Provide more detailed review of representative systems
- Introduce Euterpe with live coding – a web framework for developing musical agents
- Discuss challenges and future research directions

Tutorial Outline

- (15 min – Zhiyao) Introduction
- (20 min – Christos) Taxonomy and examples of CAMM systems
- (45 min – Philippe) Review of music performance systems – musical agents
- (45 min – Philippe) Review of music composition systems
- (25 min – Christos) Live coding with Euterpe

- (30 min) Break

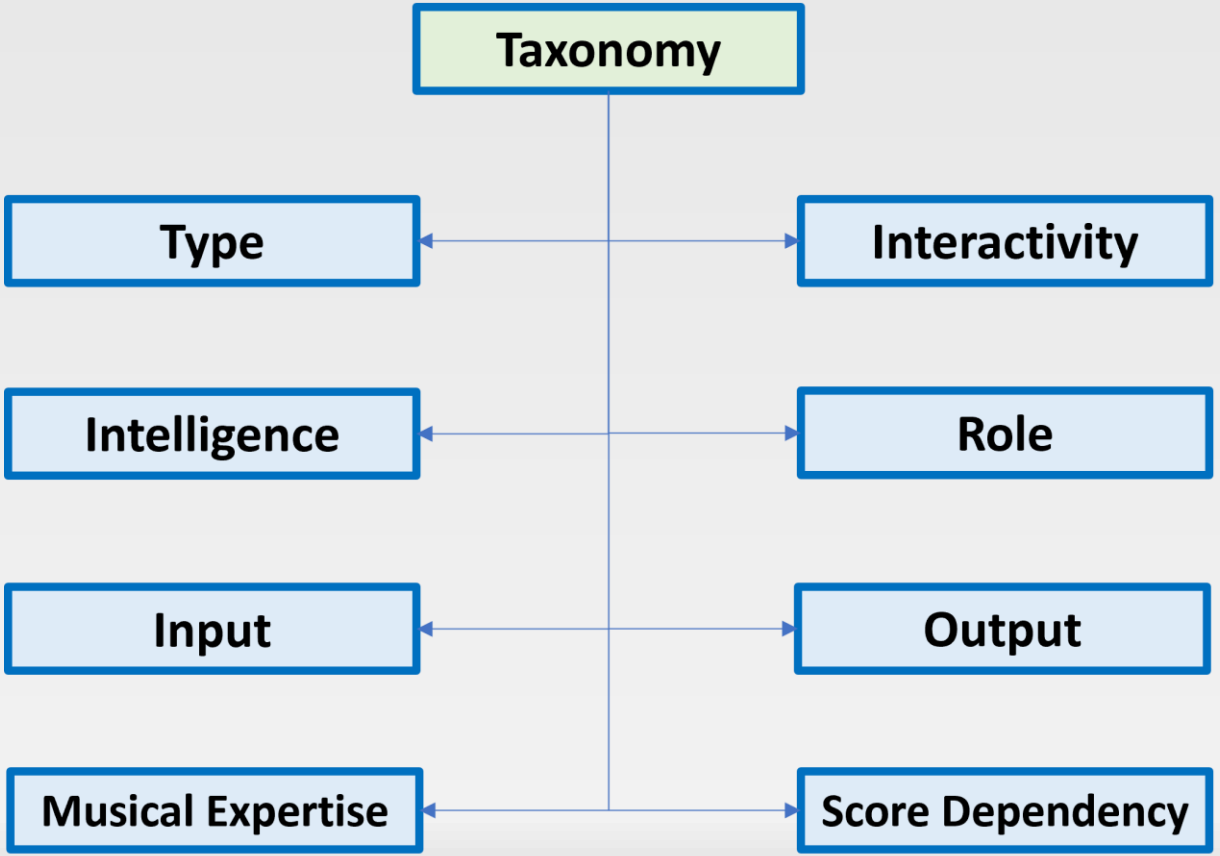
- (35 min – Christos) Continue live coding with Euterpe
- (15 min – Zhiyao) Challenges and future directions
- (10 min – All) Q&A

A Taxonomy

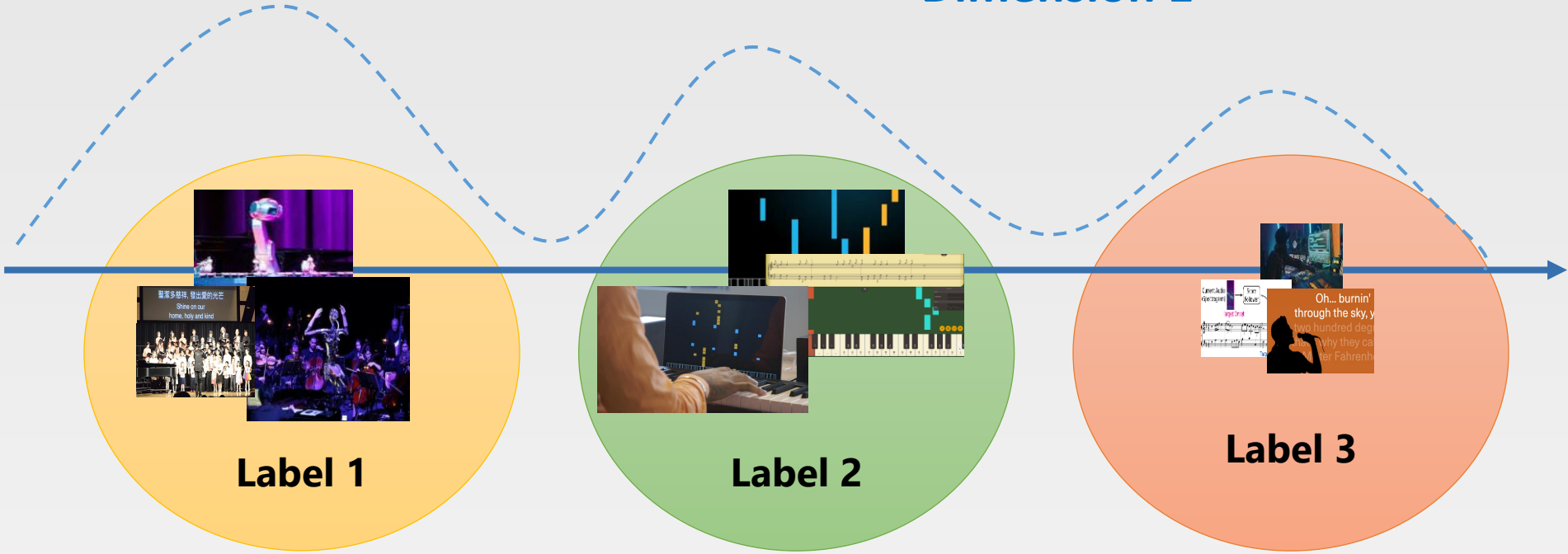
for Computer-Assisted Music-Making systems

Christodoulos Benetatos

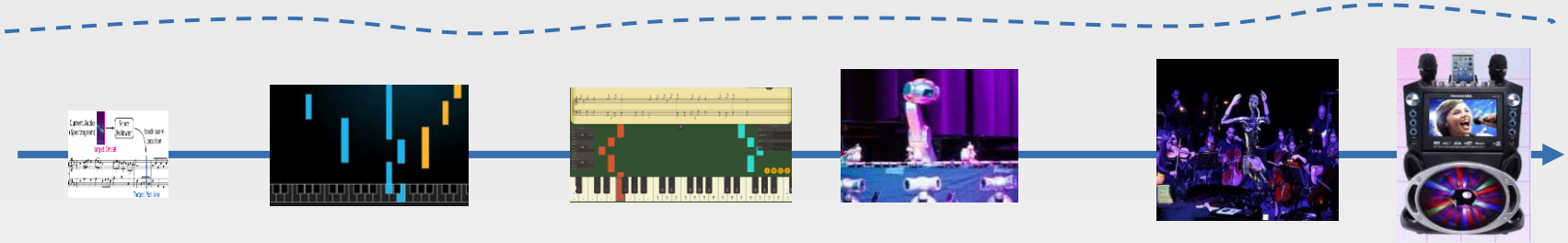
Dimensions


Taxonomy
Type
Interactivity
Intelligence
Role
Input
Output
Musical Expertise
Score Dependency

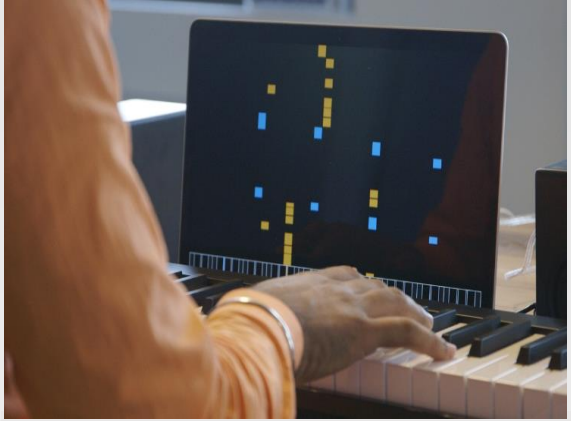
Dimension 1



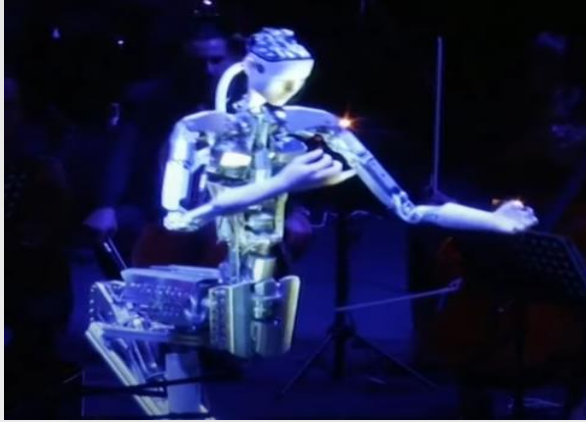
Dimension 2



Automatic Lyrics Display

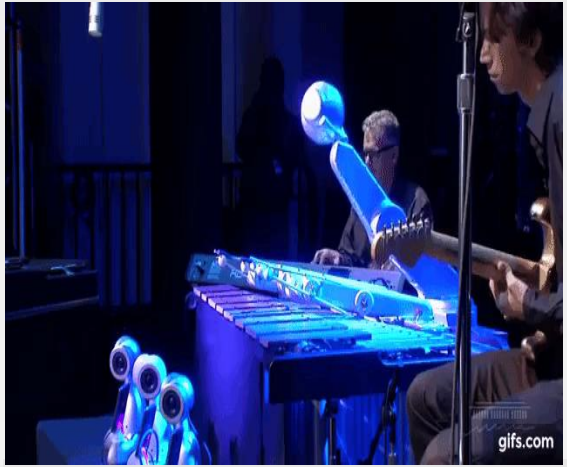


C & R Musical Partner

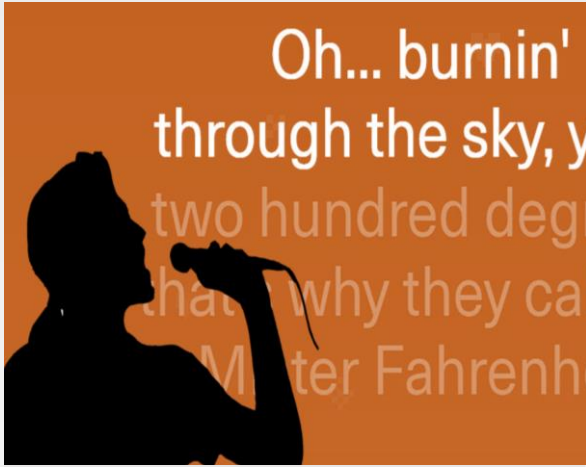


Robotic conductor

Intelligent Instrument

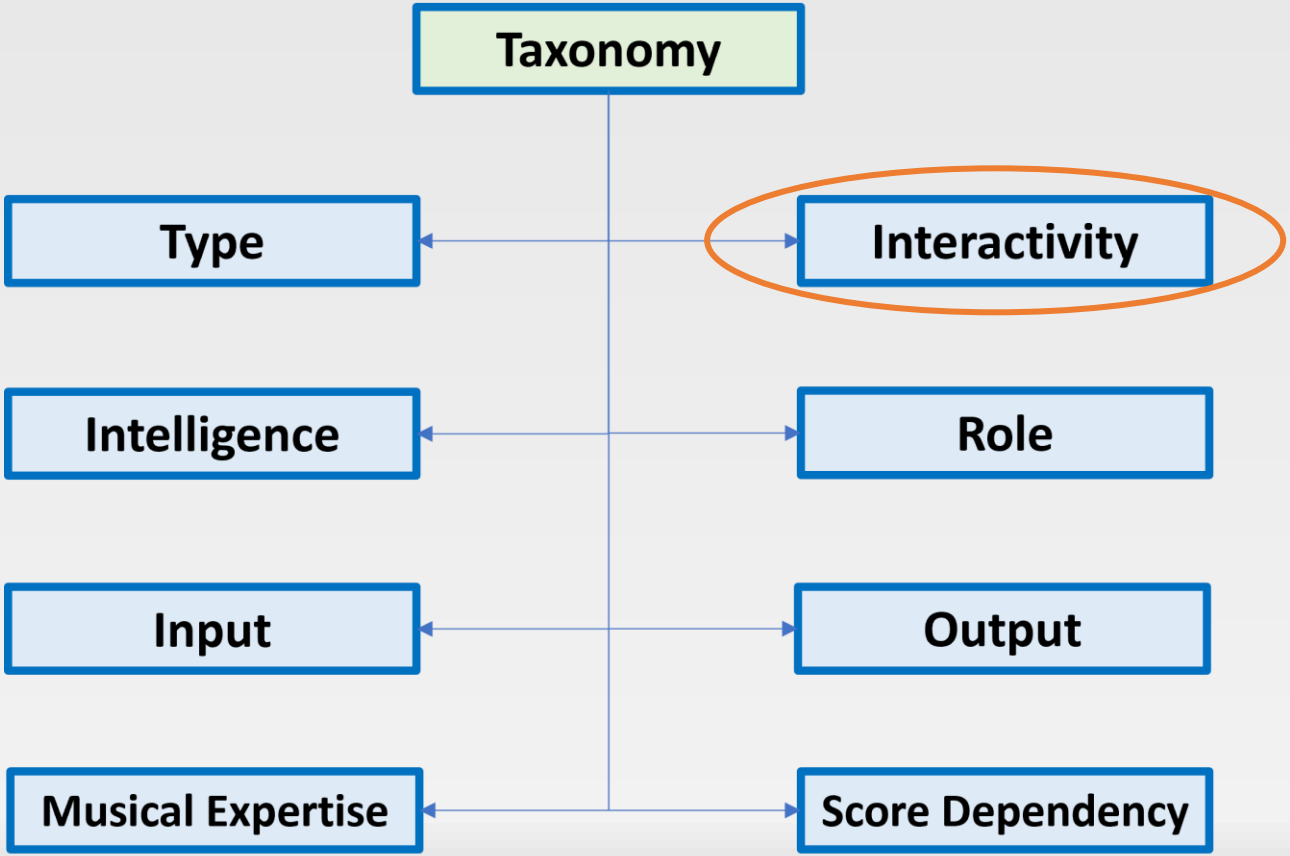


Robotic Ensemble Musician

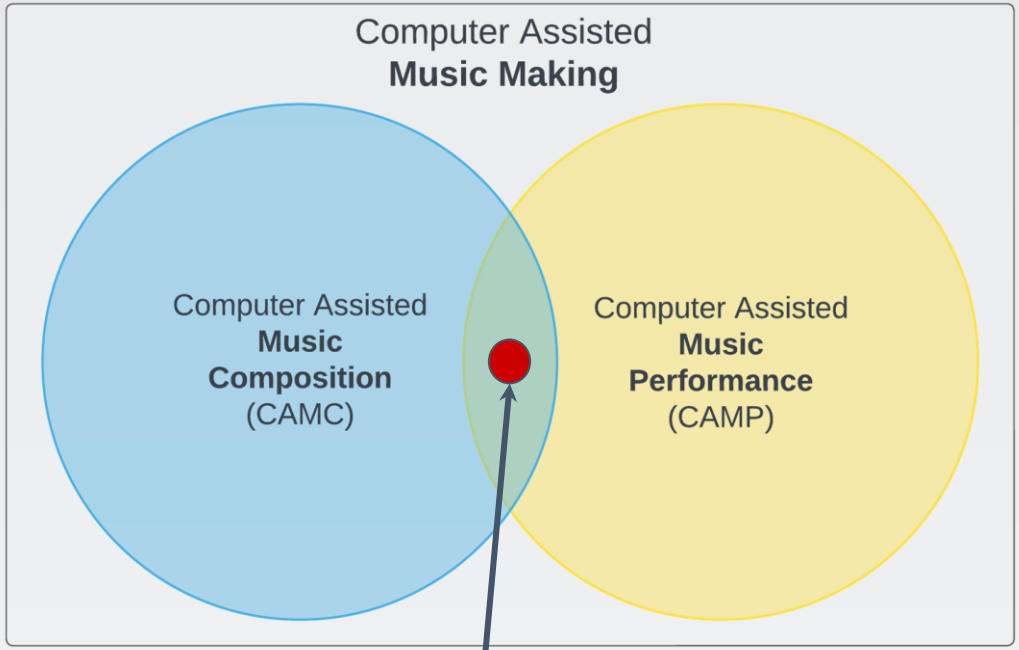


Karaoke

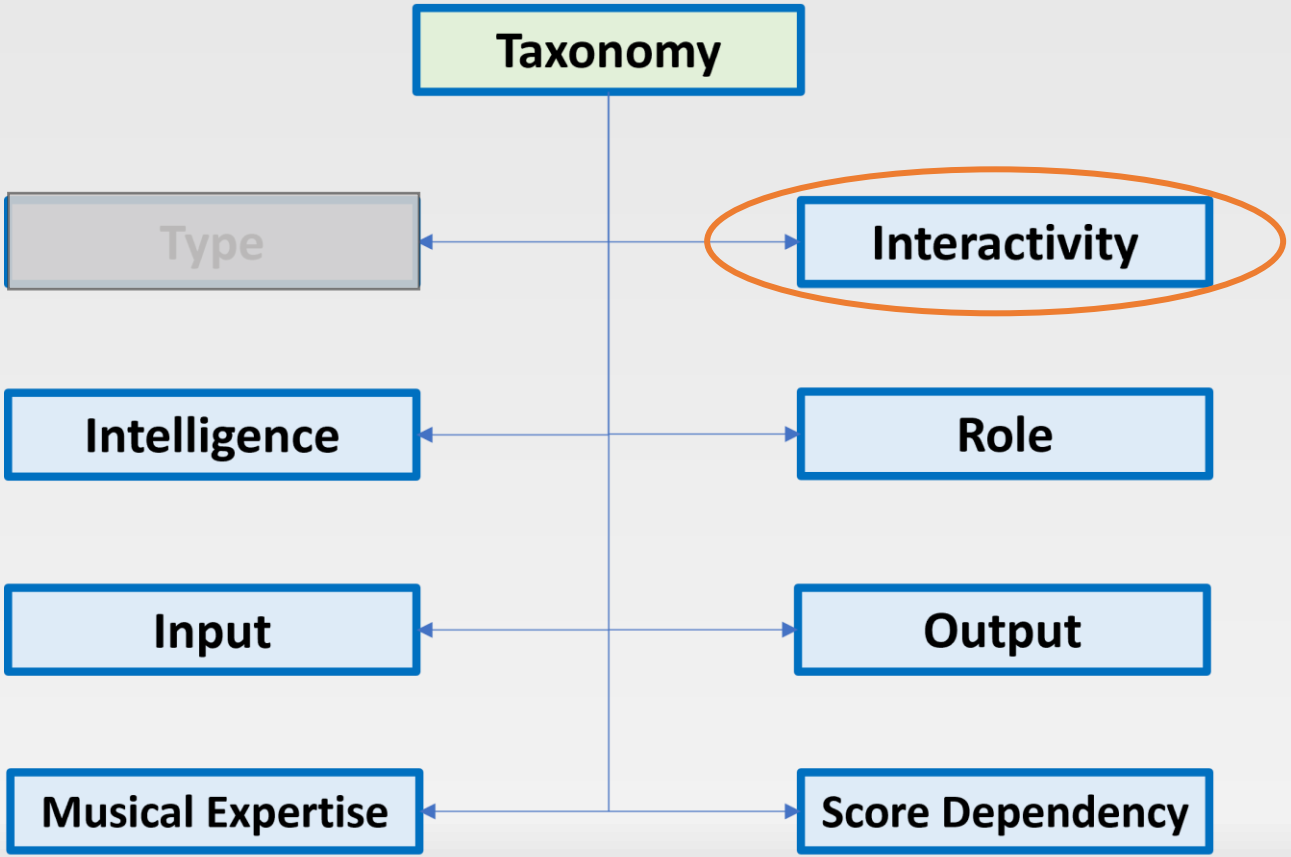
Dimensions


Taxonomy
Type
Interactivity
Intelligence
Role
Input
Output
Musical Expertise
Score Dependency

1. Type

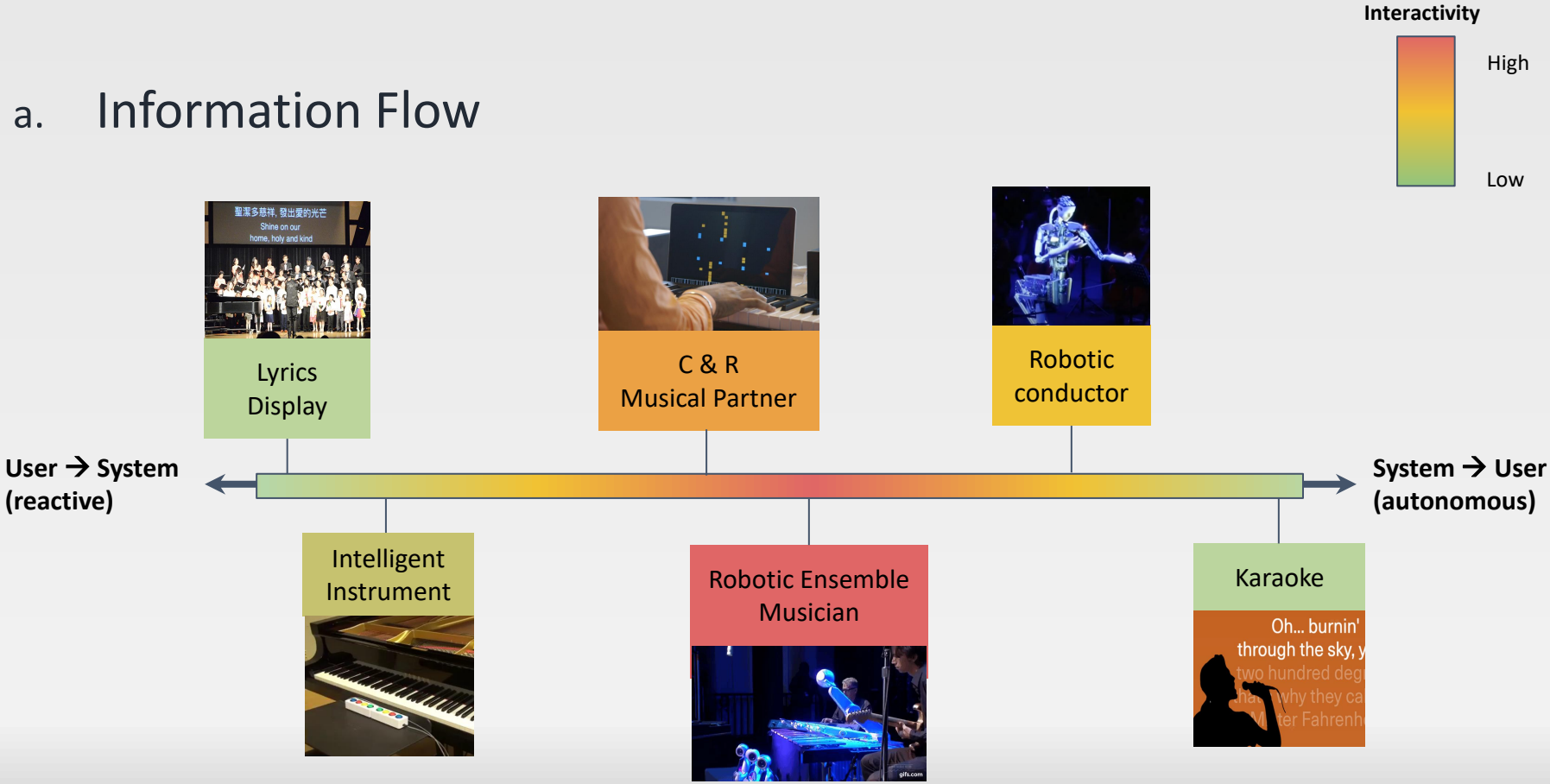

Computer Assisted
Music Making
Computer Assisted
Music
Composition
(CAMC)
Computer Assisted
Music
Performance
(CAMP)
Real-Time
Improvisation Agents

Dimensions

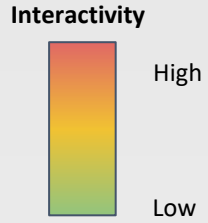

Taxonomy
Type
Interactivity
Intelligence
Role
Input
Output
Musical Expertise
Score Dependency

2. Interactivity

a. Information Flow

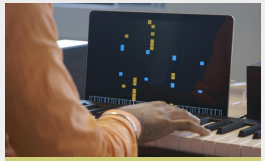


2. Interactivity



b. Responsiveness

- For every **input** event, there is **response** within specified **deadline**



C & R
Musical Partner

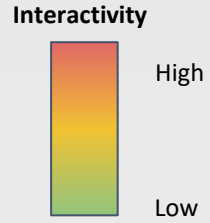
Robotic Ensemble
Musician

Lyrics
Display

soft ← → hard

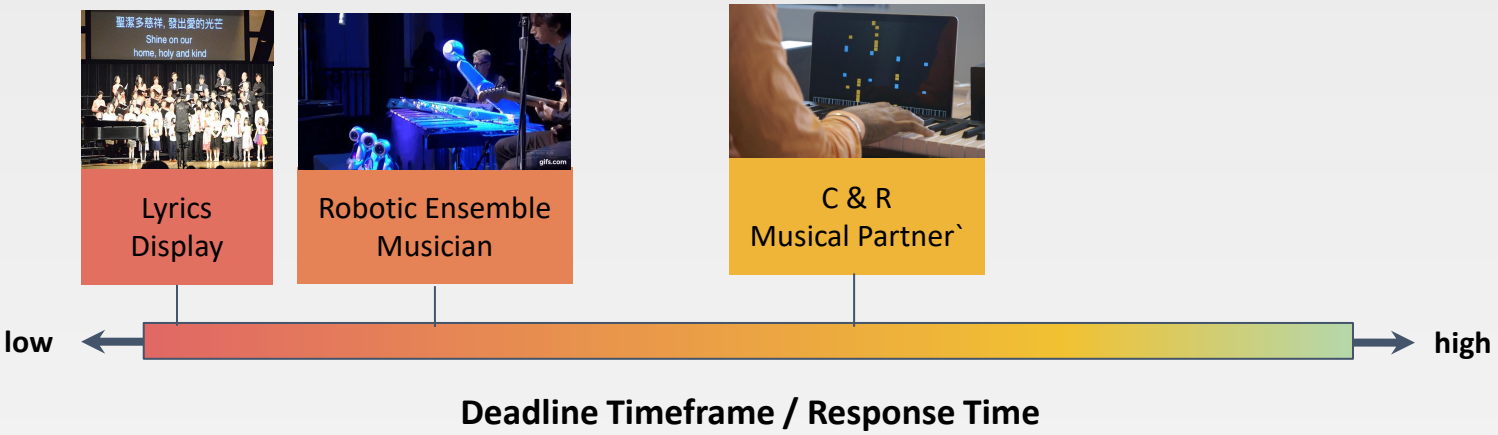
Deadline Type

2. Interactivity



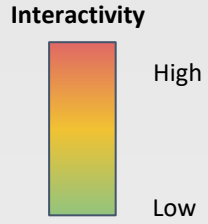
b. Responsiveness

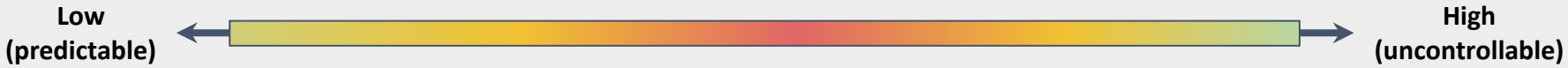
- For every **input** event, there is **response** within specified **deadline**



2. Interactivity

c. Stochasticity





2. Interactivity

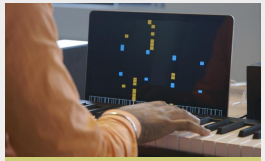
Interactivity

High

Low

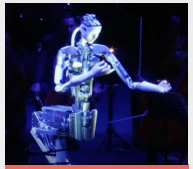
d. Synchronicity

- Percentage of concurrent task performance.



C & R
Musical Partner`

Robotic Ensemble
Musician

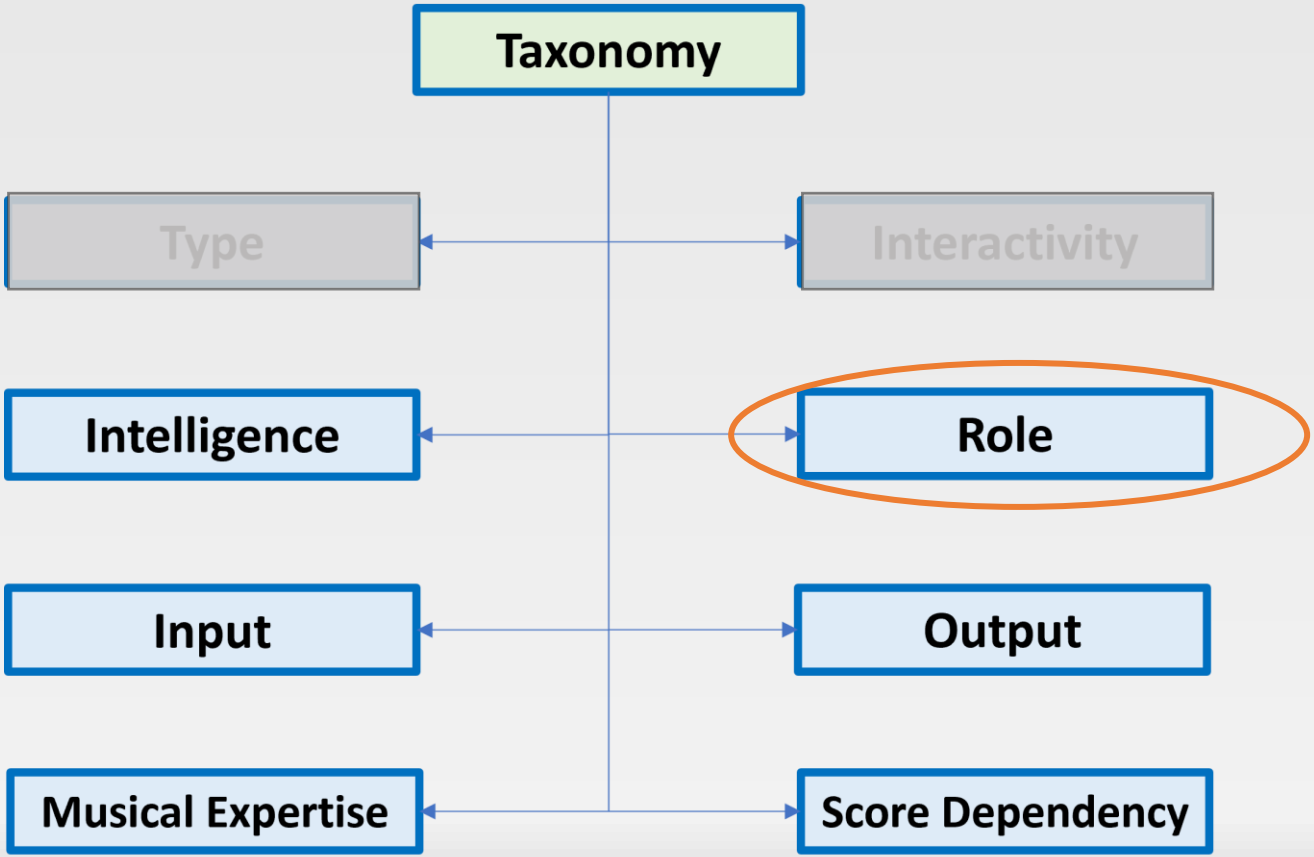


Robotic
conductor

low

high

Dimensions


Taxonomy
Type
Interactivity
Intelligence
Role
Input
Output
Musical Expertise
Score Dependency

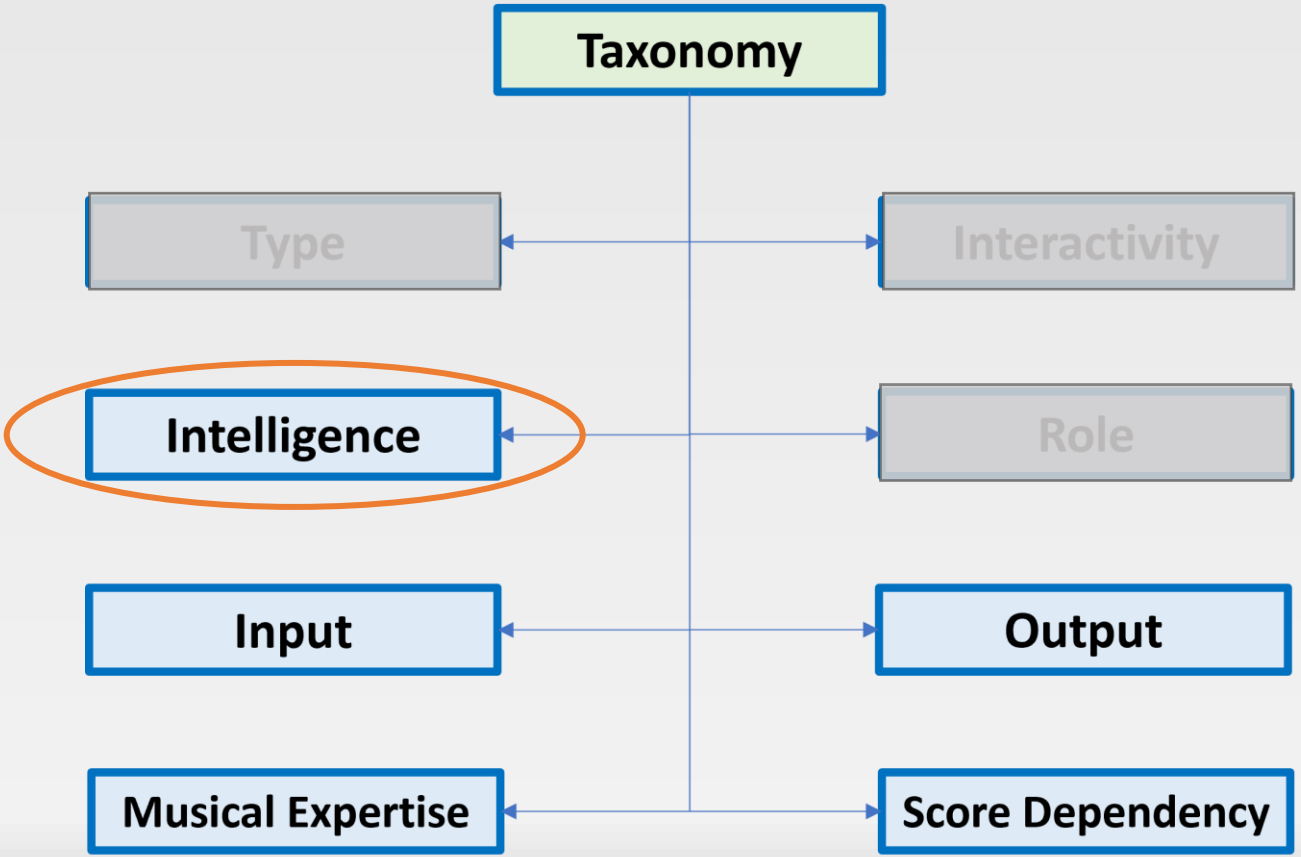
3. Role

a. Agents
   - Reactive
   - Partners
   - Leaders
   - Autonomous

b. Tools

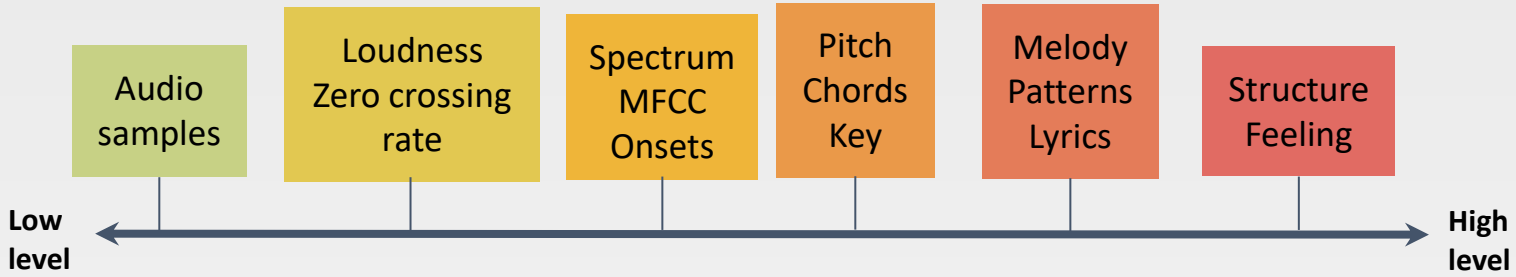
c. Instruments

Dimensions


Taxonomy
Type
Interactivity
Intelligence
Role
Input
Output
Musical Expertise
Score Dependency

3. Intelligence

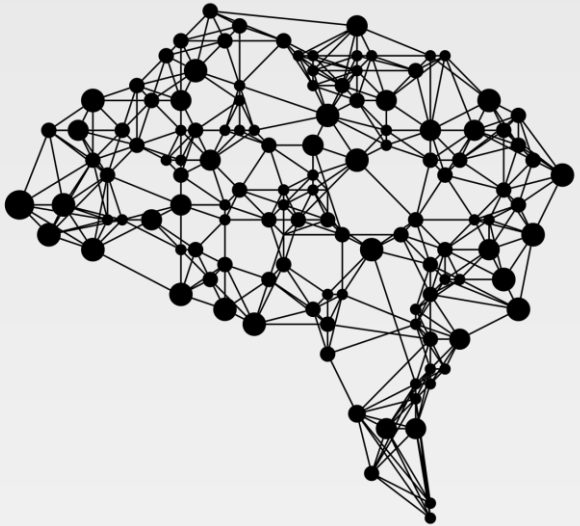
a. Perception



3. Intelligence

b. Cognition

- Genetic algorithms
- Neural networks
- Dynamic programming
- Etc.



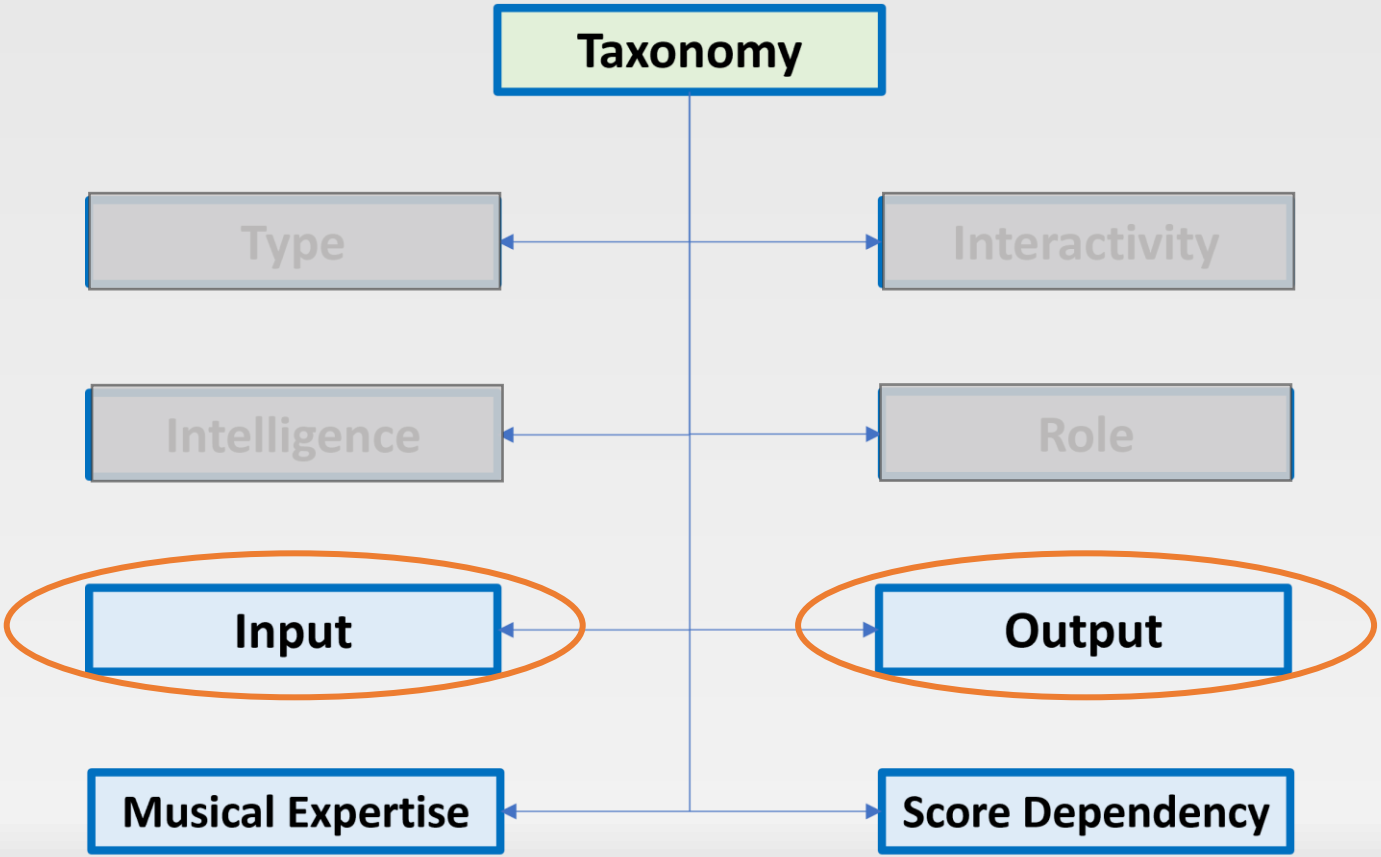
3. Intelligence

c. Expressivity

- Agents
  - Ability to make use of expressivity tools
- Instruments
  - Amount of provided expressivity tools

poco rubato

Dimensions


Taxonomy
Type
Interactivity
Intelligence
Role
Input
Output
Musical Expertise
Score Dependency

5/6. Input/Output

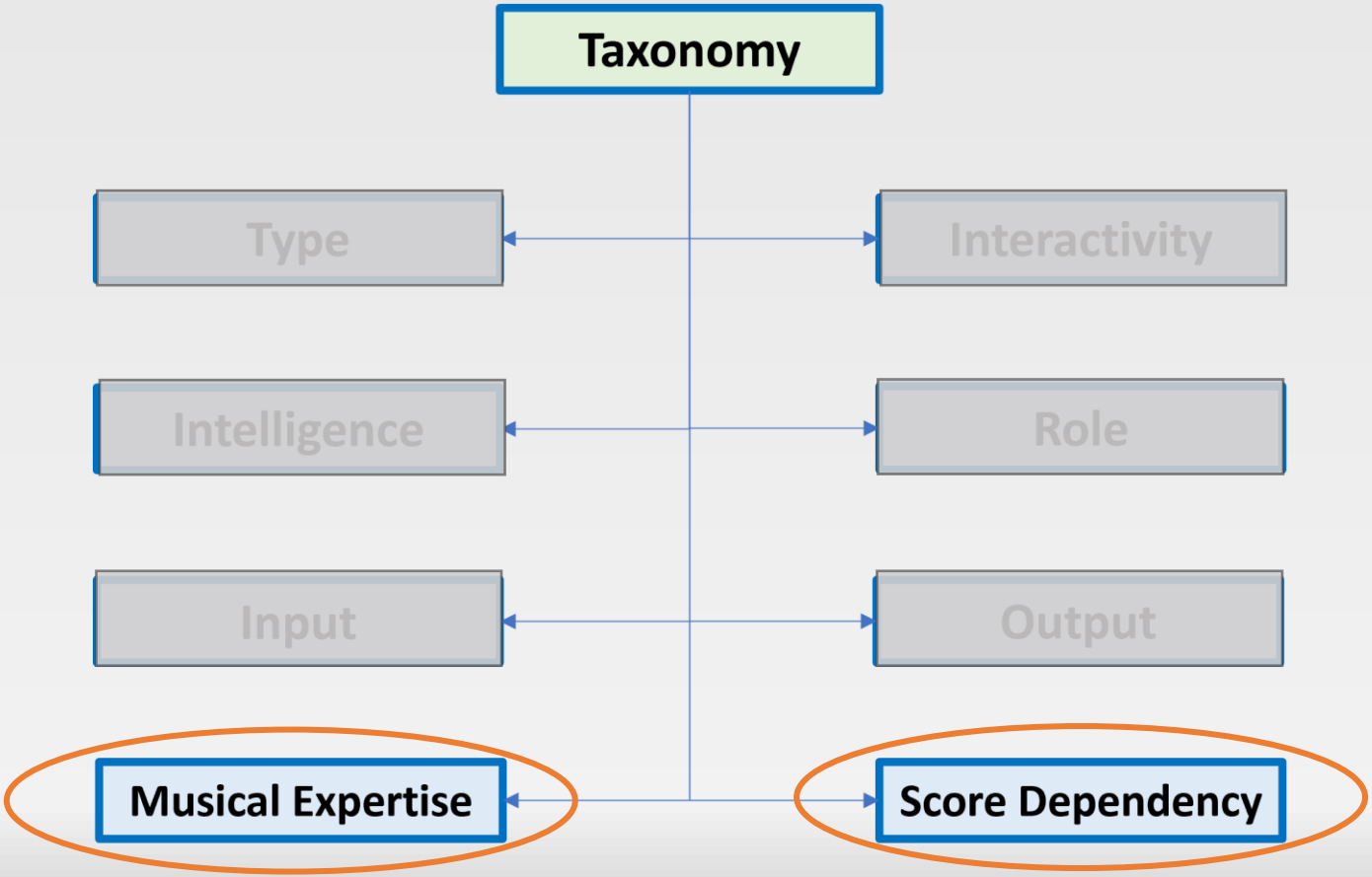
Input

- Audio
- Symbolic (MIDI, MusicXML)
- Text
- Visual (camera, depth sensors)
- Wearable Sensors

Output

- Audio
- Symbolic (MIDI, MusicXML)
- Text
- Visual Rendering (i.e dancing avatar)
- Motor Actions

Dimensions

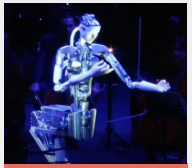

Taxonomy
Type
Interactivity
Intelligence
Role
Input
Output
Musical Expertise
Score Dependency

7. Musical Expertise

- Minimum required musical knowledge for successful operation
  - Low (non musicians)
  - Medium (amateur musicians)
  - High (professional musicians and producers)

8. Score Dependency

- Internally stored music score
  - match with music arriving at the input (score followers)
  - Improvisation over a given chord sequence
  - guide the systems' musical performance



Robotic conductor

Lyrics Display

Computer-Assisted (AI) Music-Making Systems
ISMIR 2023

Philippe Pasquier
Professor of Creative AI
School for Interactive Arts + Technology (SIAT)
Simon Fraser University, Vancouver, Canada.
With special thanks to Pr. Zhiyao Duan for slides marked * and Christodoulos Benetatos (marked +)

METACREATION
Lab for Creative Artificial Intelligence

SCHOOL OF INTERACTIVE ARTS + TECHNOLOGY

SFU SIMON FRASER UNIVERSITY
ENGAGING THE WORLD

Outline of the Tutorial

- Two broad types of Creative AI applications:
  - Embedded Generative Systems (Embedded Generation): live performance, interactive systems, games, …
  - Computer-Assisted Creativity (Interactive Generation): augmenting creative software with Creative AI.

Outline of the Tutorial

- Two broad types of computer-assisted music-making systems:
  - Musical Agents: online interactive generation
    - I.e., Computer-assisted music performance (CAMP in our intro)
  - Computer-Assisted Composition
    - i.e., Computer Assisted Music Composition (CAMC in our intro)

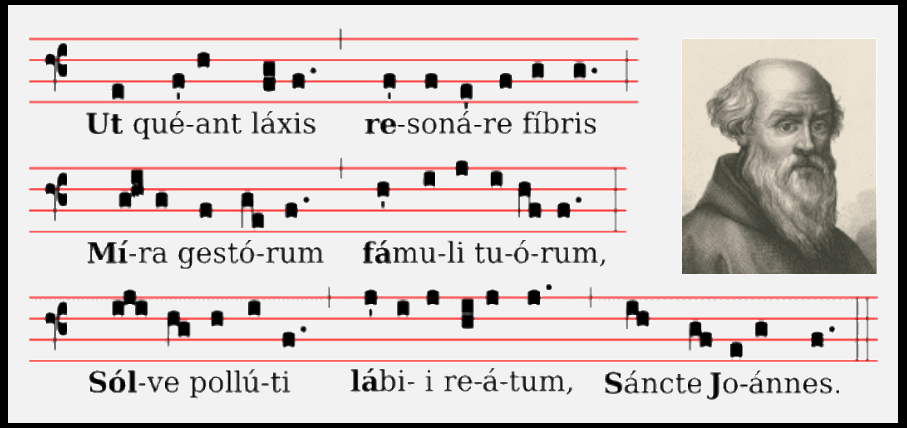
Outline of the Tutorial

- Historical precedents
- Musical Agents:
  - Cognitive agents
  - Reactive agents
  - Hybrid agents
  - Virtual Ecosystems (Artificial Life)
- Computer-Assisted Composition:
  - Audio domain
  - Symbolic domain

Historical Precedents

A brief History of music generation

- It does not start with computers.
- Guido d'Arezzo (one of the pioneer of musical notation) had the idea of an algorithmic composition associating a note to each vowels of a text as early as 1026
- Conceptual machines aside, it starts with early automaton



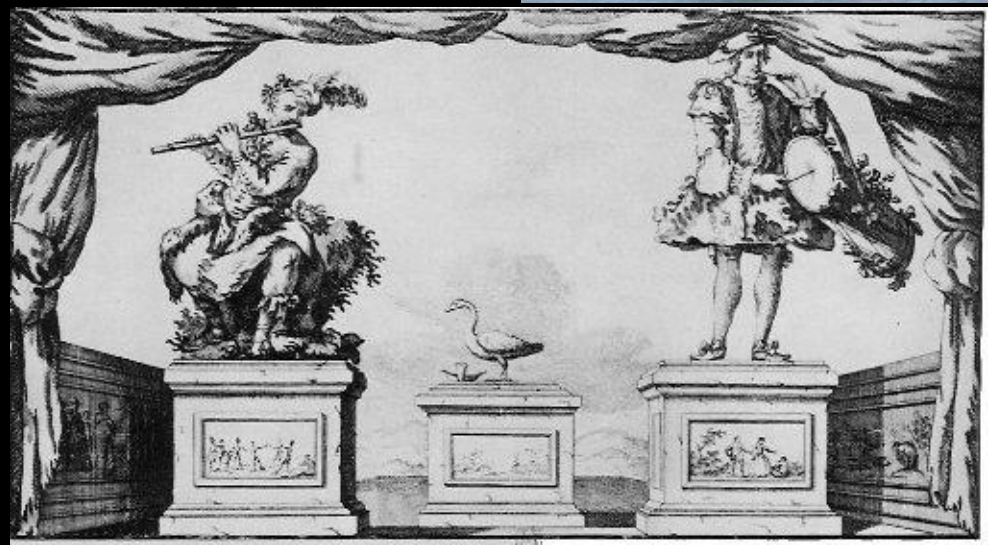
Early Automaton

- With the development of energy sources some processes start to be automatized, and more and more machines are being built.
- Very early, water was used and hydraulic energy started to be exploited.
- The hydraulic organ or **Hydrolis** was conceived 3rd century BC in ancient Greece. It does not need the human to blow air anymore.
- Fountains, which seem to defy the laws of gravity become a trend. The siphon that makes water travel upward is attracting curiosity (as it is magic to those that are not in the know)
- This is the emergence of automaton
- The polymath and mechanical genius, Al Jazari (12th century), is as known for his hydraulic automaton, as for his ingenuous engineering
- He produced a band of musical automaton.
- Al-Jazari created a boat with four automatic musicians that floated on a lake to entertain guests at royal drinking parties. It was programmable so that each automatic musician could could play different patterns.

Early Automaton

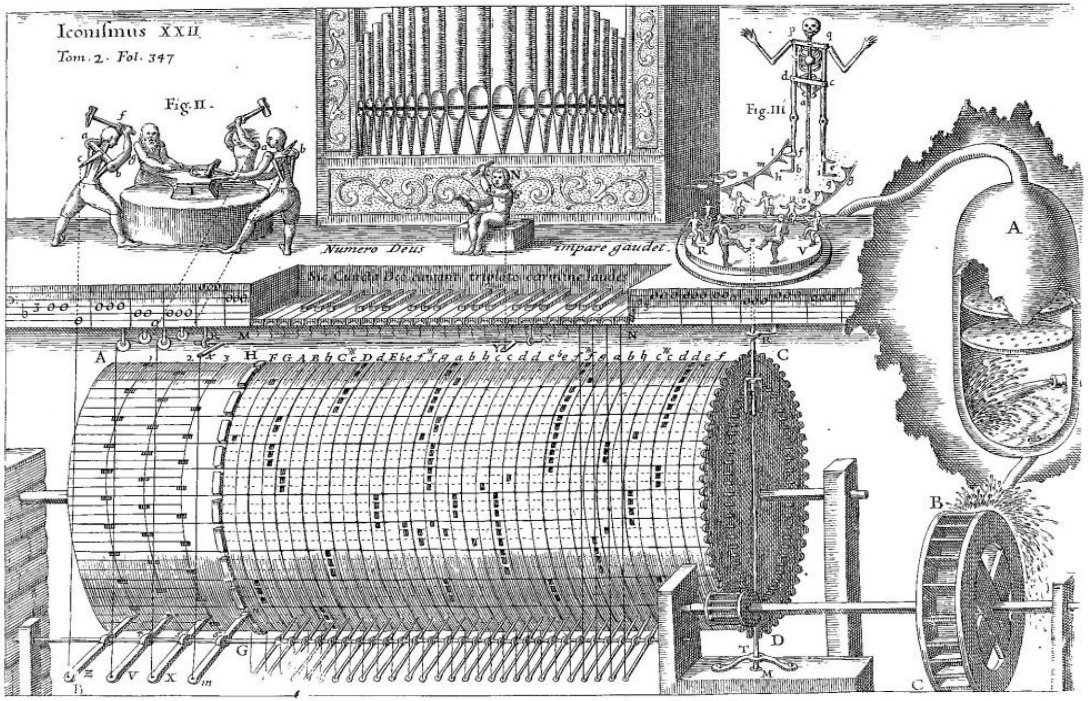
- Around the 14th century, and with the development of physics hydrolic energy is supplemented with mecanical energy and steam powered systems.
- Automatons become more common:

- A wide variety of automaton are produced ranging from pieces of furniture and instruments like the barrel organ, to androids and animal automaton like Vaucanson flute player, tambour player and duck.
- The duck, for example is made of over 400 moving parts, allowing the automaton to eat, digest and defecate.

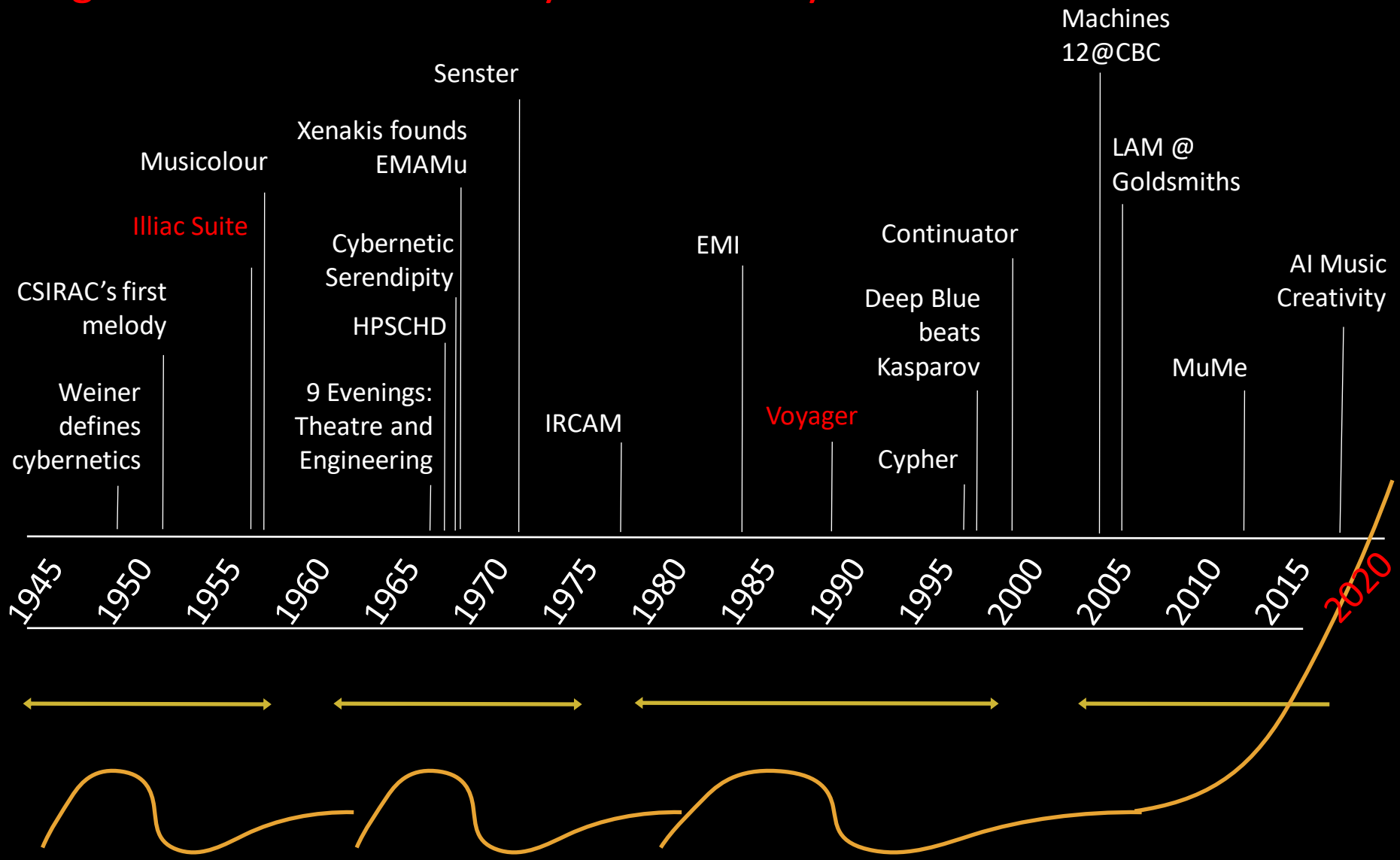


Early Automaton

- Besides the cam, the pin cylinder was invented.
- Although it was not thought of in terms of information and programming at the time, It did inspire the automatic loom, which in turn influenced the design of the first computers.
- Kircher, hydraulic organ with dancing skeleton from 1650.



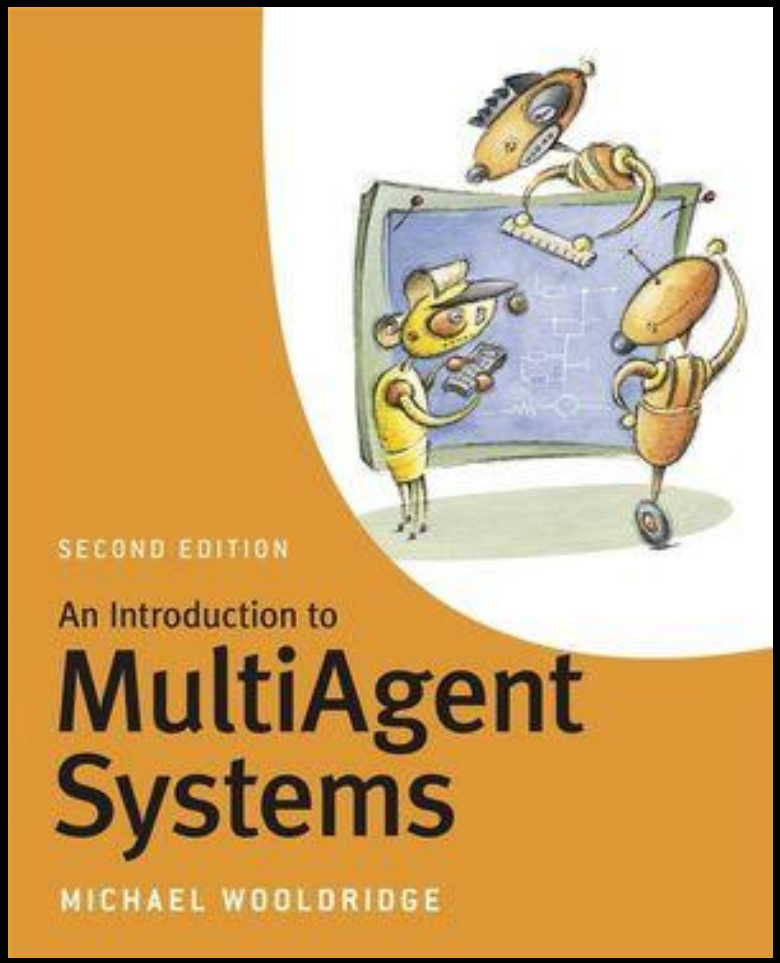
Digital Generative Music Systems History

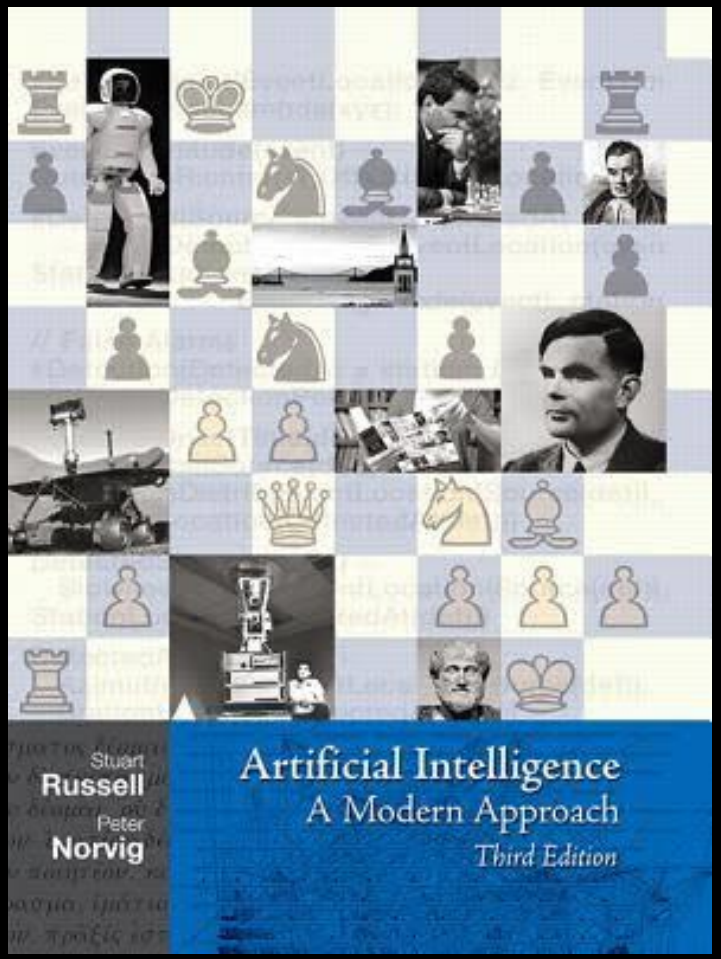


Musical agents

Musical agents are agents!

Agents and Multiagent Systems





Agents and Multiagents Systems

- An artificial agent is a computer system that is capable of autonomous action on behalf of its user or designer.

- A multiagent system is one that consists of a number of agents, which *interact* with their environment (including with one-another)

Agent architectures

- Three types of agent architectures:
  - Cognitive: maintain internal symbolic representations
    - Deliberative architectures: reasoning and planning
  - Reactive: no explicit representation of the environment and focus on behavioural rules
    - Reflex: no internal states (just mapping inputs to outputs)
    - Reactive: with internal states (but not cognitive)
  - Hybrid: mixing reactive and cognitive components to balance reactiveness and deliberativeness

Musical Agent

Routledge
Taylor & Francis Group

Journal of New Music Research

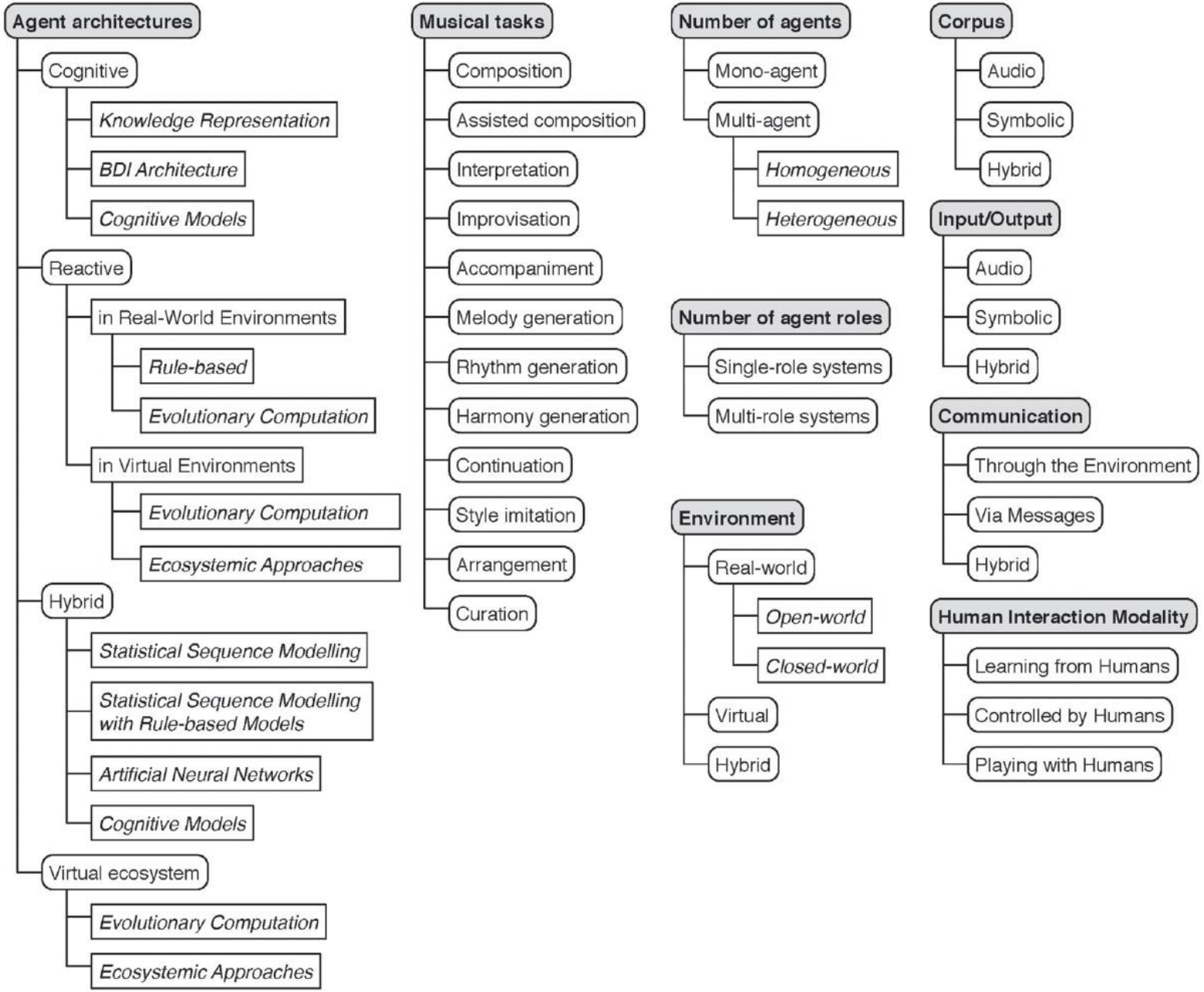
ISSN: 0929-8215 (Print) 1744-5027 (Online) Journal homepage: http://www.tandfonline.com/loi/nnmr20

Musical agents: A typology and state of the art towards Musical Metacreation

Kıvanç Tatar & Philippe Pasquier

To cite this article: Kıvanç Tatar & Philippe Pasquier (2018): Musical agents: A typology and state of the art towards Musical Metacreation, Journal of New Music Research, DOI: 10.1080/09298215.2018.1511736

To link to this article: https://doi.org/10.1080/09298215.2018.1511736


Agent architectures
Cognitive
Knowledge Representation
BDI Architecture
Cognitive Models
Reactive
in Real-World Environments
Rule-based
Evolutionary Computation
in Virtual Environments
Evolutionary Computation
Ecosystemic Approaches
Hybrid
Statistical Sequence Modelling
Statistical Sequence Modelling with Rule-based Models
Artificial Neural Networks
Cognitive Models
Virtual ecosystem
Evolutionary Computation
Ecosystemic Approaches
Musical tasks
Composition
Assisted composition
Interpretation
Improvisation
Accompaniment
Melody generation
Rhythm generation
Harmony generation
Continuation
Style imitation
Arrangement
Curation
Number of agents
Mono-agent
Multi-agent
Homogeneous
Heterogeneous
Number of agent roles
Single-role systems
Multi-role systems
Environment
Real-world
Open-world
Closed-world
Virtual
Hybrid
Corpus
Audio
Symbolic
Hybrid
Input/Output
Audio
Symbolic
Hybrid
Communication
Through the Environment
Via Messages
Hybrid
Human Interaction Modality
Learning from Humans
Controlled by Humans
Playing with Humans

#	System	Architecture	# of Agents	# of roles	Environment	Corpus	Input	Output	Communication	HIM	MuMe Task	Evaluation	Code	Public
	Cognitive Musical Agents													
①	VMMAS	knowledge representation	Multi-agent/ Heterogenous	Multi-role	Real-world	Symbolic	Symbolic	Symbolic	Env.	P	Comp., Accomp.	✓	Not shared	
②	Inmamusys	knowledge representation	Multi-agent/ Heterogenous	Multi-role	Real-world	Symbolic	Symbolic	Symbolic	Mess.	C	Comp., Accomp.	✓	Not shared	
③	Generating Affect	knowledge representation	Multi-agent/ Heterogenous	Multi-role	Real-world	Hybrid	None	Audio	Mess.	–	Comp.		Shared	✓
④	Coming Together	BDI	Multi-agent/ Homogenous	Single-role	Real-world	–	Symbolic	Hybrid	Hybrid	–	Comp.		Shared	✓
⑤	Indifference Engine	BDI	Multi-agent/ Heterogenous	Single-role	Hybrid	–	Audio	Audio	Hybrid	P	Improv., Comp.		Not shared	✓
⑥	MUSIC-MAS	BDI	Multi-agent/ Homogenous	Multi-role	Real-world	Symbolic	Symbolic	Symbolic	Hybrid	C	Assisted Comp., Style Im.	✓	Not shared	
⑦	HSMM	Cognitive	Mono-agent	Single-role	Real-world	Symbolic	Symbolic	Symbolic	Env.	P+C+L	Comp., Assisted Comp.„ Cont.		Not shared	✓
⑧	MusiCOG	Cognitive	Multi-agent/ Heterogenous		Real-world	Symbolic	Symbolic	Symbolic	Env.	P+L	Comp., Assisted Comp.		Shared	✓
⑨	MAMA	Cognitive	Multi-agent/ Homogenous	Single-role	Real-world	Hybrid	Hybrid	Hybrid	Hybrid	–	Accomp., Improv.	✓	Shared	
	Reactive Musical Agents	in Real-World Environments												
⑩	Cypher	Rule-based	Multi-agent/ Heterogenous	Multi-role	Real-world	Symbolic	Symbolic	Symbolic	Mess.	C	Comp.		Not shared	
⑪	Voyager	Rule-based	Multi-agent/ Heterogenous	Single-role	Real-world	–	Hybrid	Symbolic	Env.	P	Improv.		Shared	✓
⑫	Bob	Rule-based	Mono-agent	Single-role	Real-world	Symbolic	Symbolic	Symbolic	Env.	P+L	Improv.,Melody Gen.		Not shared	
⑬	ARHS	Rule-based	Multi-agent/ Homogenous	Single-role	Real-world	–	Audio	Audio	Env.	P	Improv.		Not shared	✓
⑭	LL:	Rule-based	Multi-agent/ Homogenous	Single-role	Real-world	–	Audio	Audio	Env.	P	Improv.	✓	Not shared	✓
⑮	Virtualband	Rule-based	Multi-agent/ Heterogenous	Multi-role	Real-world	Hybrid	Audio	Audio	Mess.	P+L	Style Im., Accomp.		Not shared	✓
⑯	Odessa	Rule-based	Mono-agent	Single-role	Real-world	–	Audio	Symbolic	Env.	P	Improv.	✓	Not shared	✓
⑰	Rhythms as···	Rule-based	Multi-agent/ Homogenous	Single-role	Real-world	–	Symbolic	Symbolic	Mess.	–	Rhythm Gen.		Not shared	
⑱	VirtuaLatin	Rule-based	Multi-agent/ Heterogenous	Multi-role	Real-world	Symbolic	Symbolic	Symbolic	Env.	P	Rhythm Gen.		Not shared	
⑲	DrumTrack	Rule-based	Mono-agent	Single-role	Real-world	–	Audio	Audio	Env.	P	Accomp., Rhythm Gen., Improv.		Not shared	✓
⑳	BBCut2	Rule-based	Mono-agent	Single-role	Real-world	Audio	Audio	Audio	Env.	P+L	Accomp., Rhythm Gen., Improv.		Shared	

			Heterogenous									
(24)	PIWeCS	Rule-based	Multi-agent/ Heterogenous	Multi-role	Real-world	Audio	Audio	Audio	Hybrid	C	Comp.	
(25)	CT: Freesound	Rule-based	Multi-agent/ Homogenous	Single-role	Real-world	Hybrid	Audio	Audio	Hybrid	–	Comp.	✓
(26)	Curatorial···	Rule-based	Multi-agent/ Heterogenous	Multi-role	Real-world	Symbolic	Symbolic	Symbolic	–	–	Curation, Comp.	
(27)	ParamBOT	Rule-based	Multi-agent/ Heterogenous	Multi-role	Real-world	Agents	–	Audio	Mess.	–	Curation, Comp.	
(28)	GenJam	Evolutionary Computation	Mono-agent	Single-role	Real-world	Symbolic	Symbolic	Symbolic	Hybrid	P+L	Improv.,Melody Gen.	
(29)	automated···	Evolutionary Computation	Mono-agent	Single-role	Real-world	–	Audio	Audio	Env.	P	Improv.	
(30)	Frank	Evolutionary Computation	Mono-agent	Single-role	Real-world	Hybrid	Audio	Audio	Env.	P+L	Improv.	
(31)	RGeme	Evolutionary Computation	Multi-agent/ Homogenous	Single-role	Real-world	Symbolic		Symbolic		–	Rhythm Gen.	
(32)	··· Tuning···	Evolutionary Computation in Virtual Environments	Multi-agent/ Homogenous	Single-role	Real-world	–	Symbolic	Symbolic	Hybrid	–	Assisted Comp.	✓
(33)	Frankensteinian···	Evolutionary Computation	Multi-agent/ Heterogenous	Multi-role	Real-world	Symbolic	–	Symbolic	Hybrid	–	Comp., Assisted Comp.	
(34)	Living Melodies	Evolutionary Computation	Multi-agent/ Homogenous	Single-role	Virtual ecosystem	Symbolic	–	Symbolic	Hybrid	–	Comp., Assisted Comp.	
(35)	Emergent···	Evolutionary Computation	Multi-agent/ Homogenous	Multi-role	Virtual ecosystem	Symbolic	Symbolic	Symbolic	Env.	–	Rhythm Gen.	
(36)	IMAP	Evolutionary Computation	Multi-agent/ Homogenous	Multi-role	Real-world	Symbolic	Symbolic	Symbolic	Env.	–	Interpretation	✓
(37)	RiverWave	Evolutionary Computation	Multi-agent/ Homogenous	Single-role	Virtual ecosystem	–	–	Output	Env.	–	Comp.	
(38)	Petri	Evolutionary Computation	Multi-agent/ Homogenous	Single-role	Virtual ecosystem	–	Computer Vision	Audio	Env.	C	Comp.	

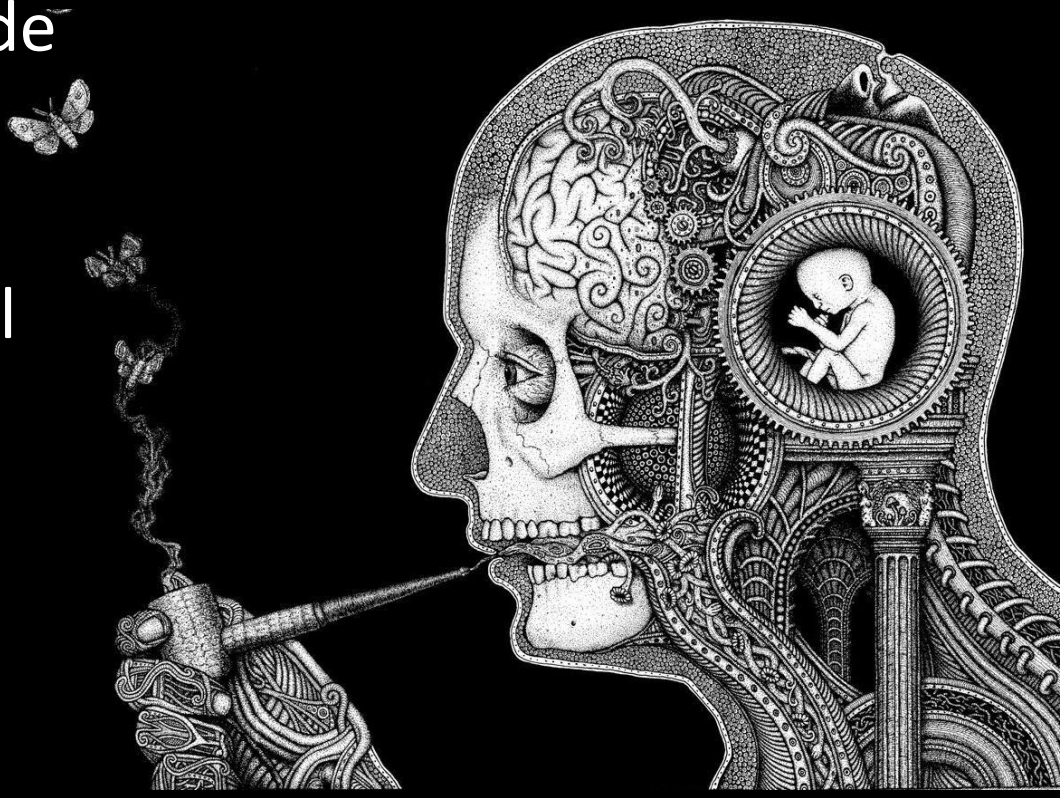
Table 1. Continued.

#	System	Architecture	# of Agents	# of roles	Environment	Corpus	Input	Output	Communication	HIM	MuMe Task	Evaluation	Code	Public
(42)	Nodal	Ecosystemic	Multi-agent/ Homogenous	Single-role	Virtual ecosystem	–	Symbolic	Symbolic	Mess.	C	Comp.		Shared	✓
(43)	OSCAR	Ecosystemic	Multi-agent/ Homogenous	Single-role	Virtual ecosystem	–	Symbolic	Symbolic	Env.	–	Comp.		Shared	✓
(44)	CT: Shoals	Ecosystemic	Multi-agent/ Homogenous	Single-role	Virtual ecosystem	Audio	Parameter	Parameter	Mess.	–	Comp.		Not shared	✓
(45)	earGram Actors	Ecosystemic	Multi-agent/ Homogenous	Single-role	Virtual ecosystem	Audio	–	Audio	Env.	C	Comp.		Not shared	✓
(46)	pMIMACS	Ecosystemic	Multi-agent/ Homogenous	Multi-role	Virtual ecosystem	–	Symbolic	Symbolic	Env.	–	Interpretation	✓	Not shared	
(47)	SDS	Ecosystemic	Multi-agent/ Homogenous	Single-role	Virtual ecosystem	–	Symbolic	Symbolic	Mess.	–	Melody Gen.		Not shared	
(48)	iMe	Ecosystemic	Multi-agent/ Homogenous	Multi-role	Virtual ecosystem	–	Symbolic	Symbolic	Env.	P	Comp., Assisted Comp.		Not shared	
	Hybrid Musical Agents													
(49)	POMDP	Statistical Sequence Modelling	Mono-agent	Single-role	Real-world	–	Symbolic	Symbolic	Env.	P+L	Improv., Style Im.		Not shared	
(50)	Continuator	Statistical Sequence Modelling	Mono-agent	Single-role	Real-world	–	Symbolic	Symbolic	Env.	P+L	Improv., Style Im., Accomp.		Not shared	✓
(51)	Beatback	Statistical Sequence Modelling	Multi-agent/ Homogeneous	Single-role	Real-world	–	Symbolic	Audio	Mess.	P+C	Rhythm Gen.	✓	Not shared	
(52)	Ringomatic	Statistical Sequence Modelling	Mono-agent	Single-role	Real-world	Hybrid	Symbolic	Audio	Env.	P	Rhythm Gen.	✓	Not shared	
(53)	Using FO···	Statistical Sequence Modelling	Mono-agent	Single-role		–	Symbolic	Symbolic	Env.	P+L	Improv., Style Im.		Not shared	✓
(54)	OMAX	Statistical Sequence Modelling	Mono-agent	Single-role	Real-world	Hybrid	Hybrid	Hybrid	Mess.	P+L	Improv., Style Im.		Not shared	✓
(55)	Anticipatory···	Statistical Sequence Modelling	Mono-agent	Single-role	Real-world	Symbolic	Symbolic	Symbolic	Env.	P+C+L	Improv., Style Im.		Not shared	
(56)	Improvagent	Statistical Sequence Modelling	Mono-agent	Single-role	Real-world	Symbolic	Symbolic	Symbolic	Env.	P	Improv.		Not shared	
(57)	Improtek	Statistical Sequence	Multi-agent/ Heterogenous	Multi-role	Real-world	Hybrid	Hybrid	Hybrid	Env.	P+C+L	Improv., Style Im.		Not shared	✓

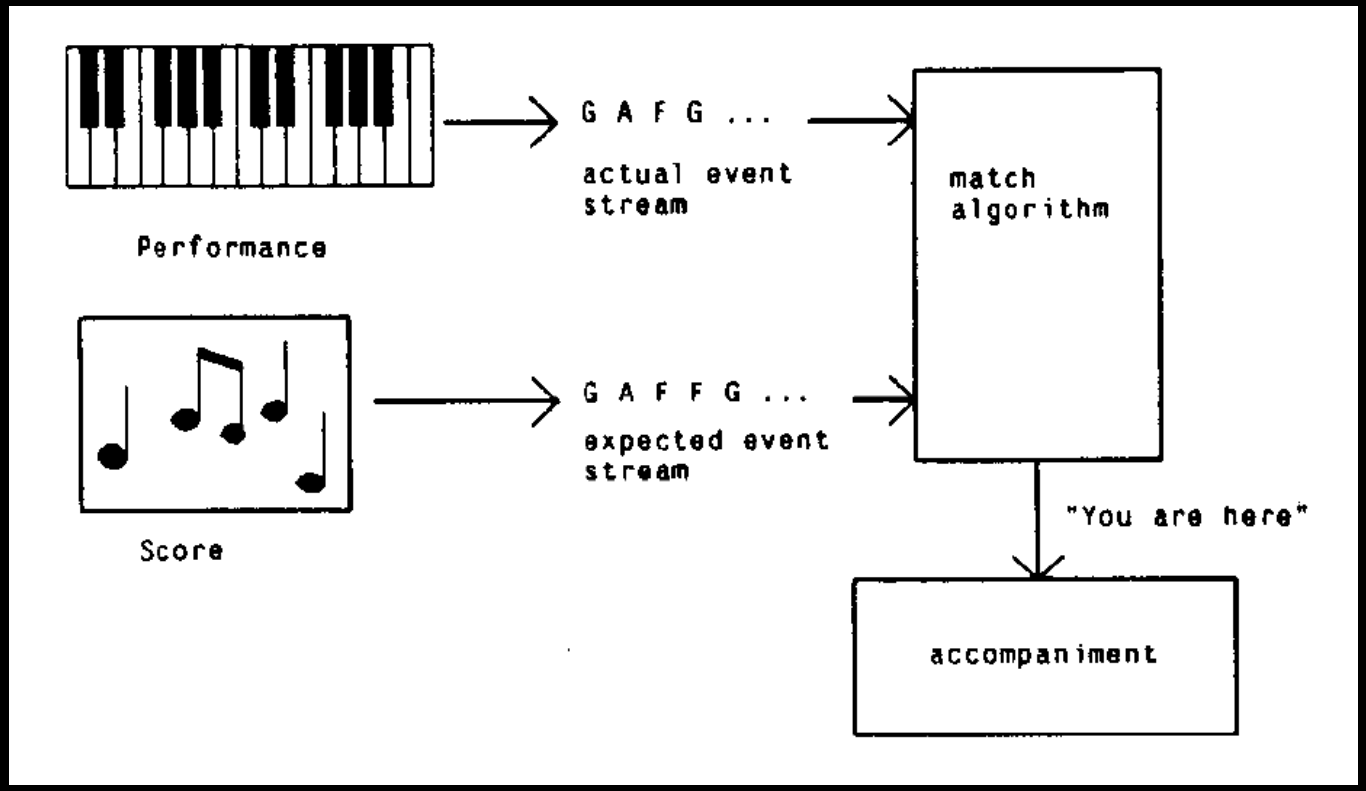
(58)	AO	Statistical Sequence Modelling	Mono-agent	Single-role	Real-world	Audio	Audio	Audio	Env.	P+C+L	Improv., Style Im.		Not shared	✓
(59)	PyOracle	Statistical Sequence Modelling	Mono-agent	Single-role	Real-world	Audio	Audio	Audio	Env.	P	Improv., Style Im.	✓	Shared	✓
(60)	VMO	Statistical Sequence Modelling	Mono-agent	Single-role	Real-world	Audio	Audio	Audio	Env.	P+C+L	Improv., Style Im		Not shared	✓
(61)	Filter	Statistical Sequence Modelling	Mono-agent	Single-role	Real-world	Hybrid	Audio	Audio	Env.	P	Improv.		Not shared	✓
(62)	SpeakeSystem	Statistical Sequence Modelling	Mono-agent	Single-role	Real-world	Symbolic	Audio	Audio	Env.	P+L	Improv.	✓	Shared	✓
(63)	ADTK	Statistical Sequence Modelling	Multi-agent/ Heterogenous	Multi-role	Real-world	–	Symbolic	Symbolic	Hybrid	C	Style Im., Improv.		Not shared	
(64)	CinBalada	Statistical Sequence Modelling	Multi-agent/ Homogenous	Multi-role	Real-world	–	–	Symbolic	Mess.	–	Rhythm Gen.	✓	Not shared	
(65)	Reactive Accompanist	Artificial Neural Networks	Mono-agent	Single-role	Real-world	–	Audio	Symbolic	Env.	P	Improv., Accomp.		Shared	
(66)	NN music	Artificial Neural Networks	Mono-agent	Single-role	Real-world	–	Audio	Audio	Env.	P+L	Improv.		Not shared	✓
(67)	… Live Algorithms	Artificial Neural Networks	Mono-agent	Single-role	Real-world	–	Audio	Audio	Env.	P	Improv.		Shared	✓
(68)	… Automated…	Artificial Neural Networks	Mono-agent	Single-role	Real-world	Symbolic	Symbolic	Symbolic	Env.	L	Improv.,Melody Gen.		Not shared	
(69)	ML.*	Artificial Neural Networks	Mono-agent	Single-role	Real-world	Hybrid	Audio	Audio	Env.	P+L	Improv.		Not shared	
(70)	Connectionist…	Artificial Neural Networks	Mono-agent	Single-role	Real-world	Symbolic	Symbolic	Symbolic	Env.	–	Rhythm Gen.		Not shared	
(71)	HARP	Cognitive	Mono-agent	Single-role	Real-world	Symbolic	Hybrid	Hybrid	Mess.	C	Assisted Comp., Improv.		Not shared	✓
(72)	Jambot	Cognitive	Mono-agent	Single-role	Real-world	–	Symbolic	Symbolic	Env.	P	Rhythm Gen., Improv., Accomp.		Not shared	
(73)	… Motivation…	Cognitive	Mono-agent	Single-role	Real-world	Symbolic	Symbolic	Symbolic	Env.	P	Improv.		Partially shared	✓
(74)	Mockingbird	Cognitive	Mono-agent	Single-role	Real-world	Hybrid	Audio	Audio	Env.	P+L	Accomp., Improv.		Not shared	
(75)	MAgentA	Cognitive	Mono-agent	Single-role	Real-world	Symbolic	–	Symbolic	Mess.	–	Comp.		Not shared	
(76)	FO with flow	Cognitive	Mono-agent	Single-role	Real-world	Symbolic	Symbolic	Symbolic	Env.	P+L	Improv.		Not shared	
(77)	MASC	Cognitive	Multi-agent/ Homogenous	Single-role	Virtual ecosystem	–	–	Symbolic	Mess.	C	Rhythm Gen.	✓	Not shared	✓
(78)	MASOM	Cognitive	Multi-agent/ Homogenous	Single-role	Real-world	Hybrid	Audio	Audio	Env.	P+L	Improv., Comp.		Partially	✓

Musical cognitive agents

- Musical cognitive agents:
  - Performing on their own
  - Performing alongside with humans
  - Helping humans to create new material

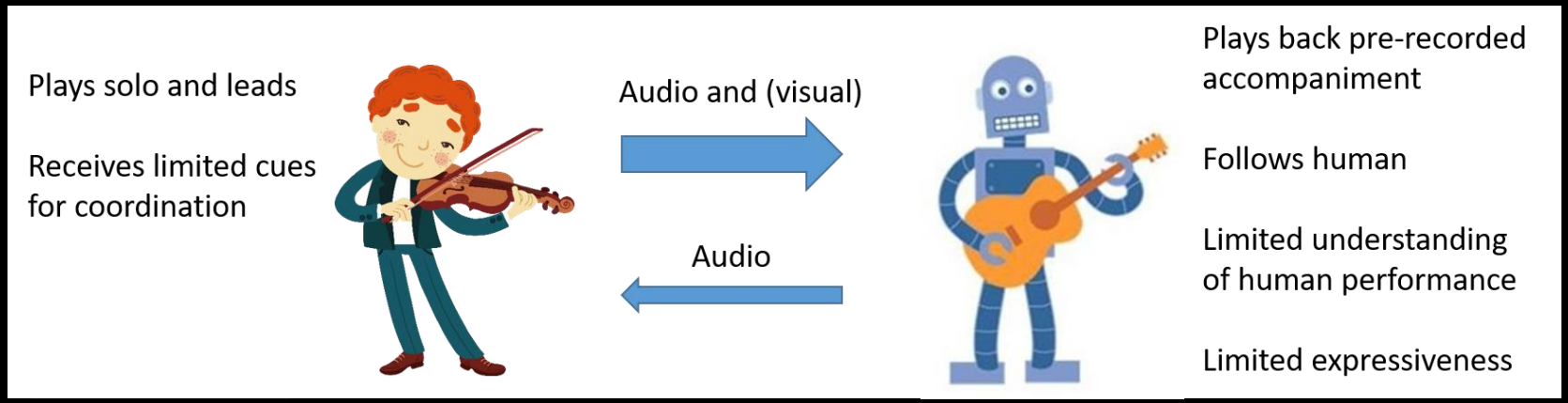


Early Acompaniment Agent



Online accompaniment interpretation system, Roger Danenberg, 1984.

Automatic Music Accompaniment Systems



Piano Tutor (Dannenberg et al., 1993)
Music Plus One (Raphael, 1999)
Antescofo (Cont, 2008)
Eurydice (Nakamura et al., 2015)
Humanoid Robot (Xia et al., 2016)

Musical Agents: Voyager (1986)

*

- Early example of “cognitive agent” working online, and interacting with live musician in the context of Jazz improvisation (free Jazz).
- The system was programmed in Forth in 1986
- Voyager Duo 4, George Lewis, 1986
  - Listens to MIDI (e.g., tempo, note spacing, melodic interval width, primary pitch material, octave range, microtonal transposition, and volume)
  - Improvises on many musical aspects (e.g., timbre, volume, microtonal transposition, tempo, tactus, note probability distributions, pitch interval range, and inter-onset time intervals)

*

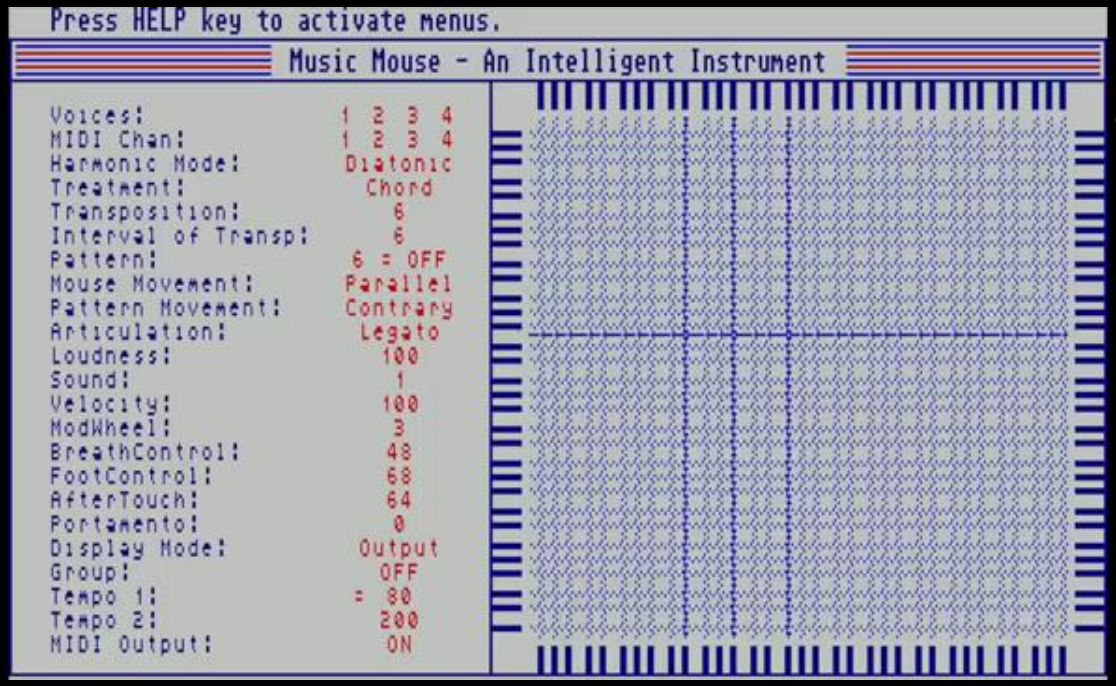
Voyager (George Lewis, 1999)

Musical Agents: Voyager

Lewis, George E. "Too Many Notes: Computers, Complexity and Culture in Voyager." *Leonardo Music Journal*, vol. 10, 2000, pp. 33-39.

Interactive Trio - George Lewis (2011)

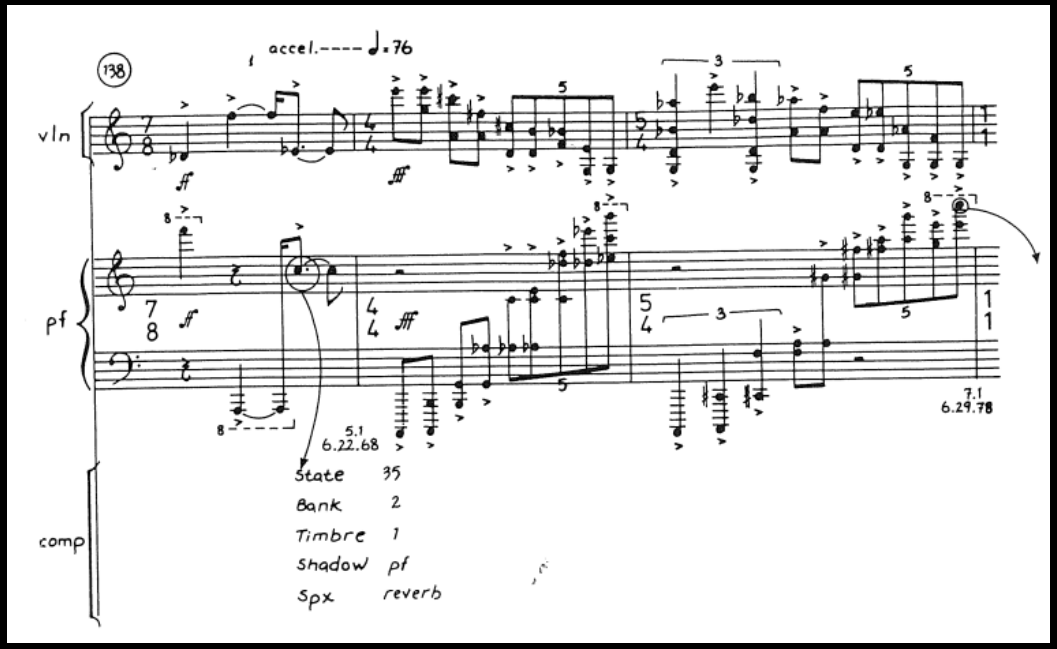
Music Mouse (http://musicmouse.com/) (Laurie Spiegel, 1986)



- Rule-based music harmonization and improvisation
  - User moves mouse in 2D space, controlling 2 voices
  - System generates the other 2 voices
  - User uses keyboard commands to control orchestration, harmonic mode, tempo, etc.

https://www.youtube.com/watch?v=D-mmEvGOopk

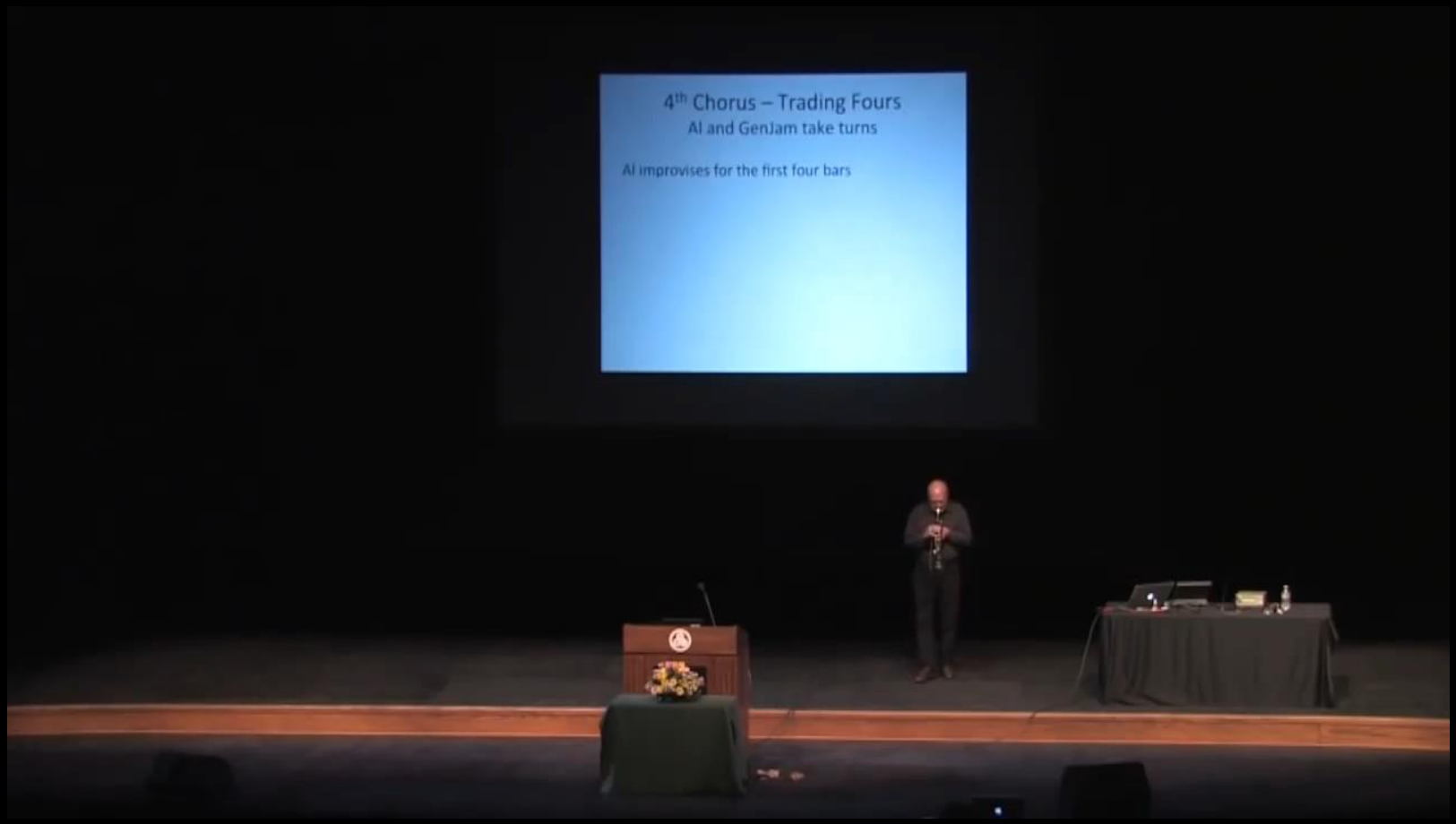
Cypher (Robert Rowe, 1992)



- Multi-agent system responding to human MIDI input in real time
  - Listener analyzes MIDI input (e.g., vertical density, attack speed, loudness, register, duration and harmony, beats, tonal pivots, etc.)
  - Player produces musical output in a virtually deterministic way

(Rowe, Interactive Music Systems, 1993)

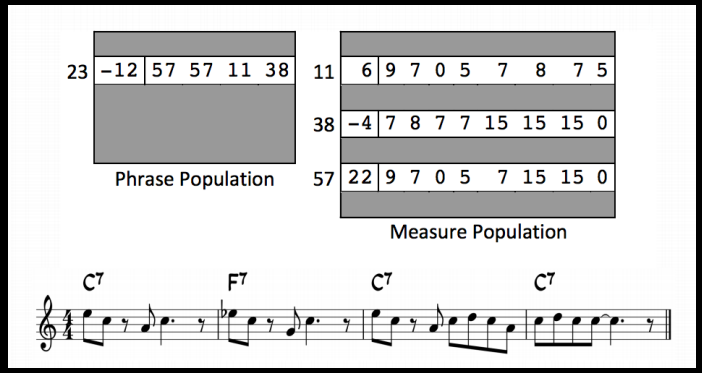
GenJam (Al Biles, 1993)



https://www.youtube.com/watch?v=rFBhwQUZGxg

GenJam

- In order to improvise Jazz solos, GenJam is co-evolving t wo populations of melodic ideas:
  - A measure population of 64 individuals: chromoso mes are made of 8 genes that each map to an 8th no tes. Each gene in a measure is encoded by four bits, with value 0 for rest, 15 being a hold, and 1-14 bein g the notes events that are mapped to an actual MI DI note through a set of scales that corresponds to t he chord being played during that measure.
  - A phrase population with 48 individuals: A phrase is made of 4 measures each encoded by 6 bits.
- Musically meaningful operators:
  - The measure mutations operate at the note level and include transposition, rotation, sorting, inversion, retrograde, …
  - The phrase mutations operate at the measure-pointer level and include reverse, rotation, sequencing, …



GenJam, Al Biles, 1993.

METACREATION
Lab for Creative Artificial Intelligence

SFU

Continuator
(François Pachet, 2002)

- Continuing music in the same style
  - Modeling user MIDI input sequences with a variable-order Markov model and builds pre-fix trees
  - Random traversals of trees to generate continuations

Omax-Ofon
(Assayag, Bloch, & Chemillier, 2006)

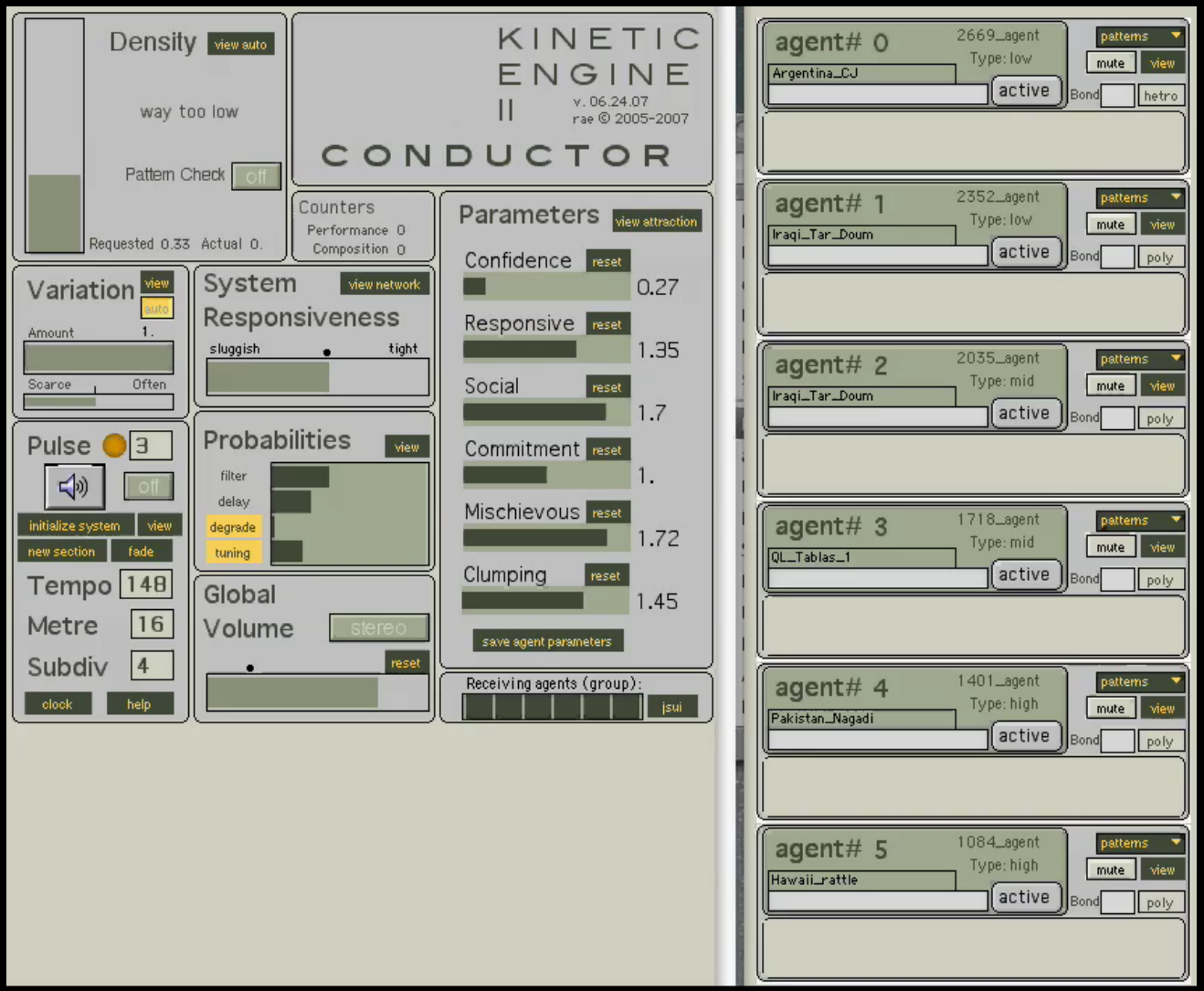
https://www.youtube.com/watch?v=2jFpGQbrcag

- Improvising based on what users just played
  - Modeling note sequences with factor oracle (a finite state automaton for efficient string matching)
  - Sampling sub-sequences to play back
  - Supports MIDI/audio input and multi-player/system settings

Shimon (Hoffman & Weinberg, 2006)

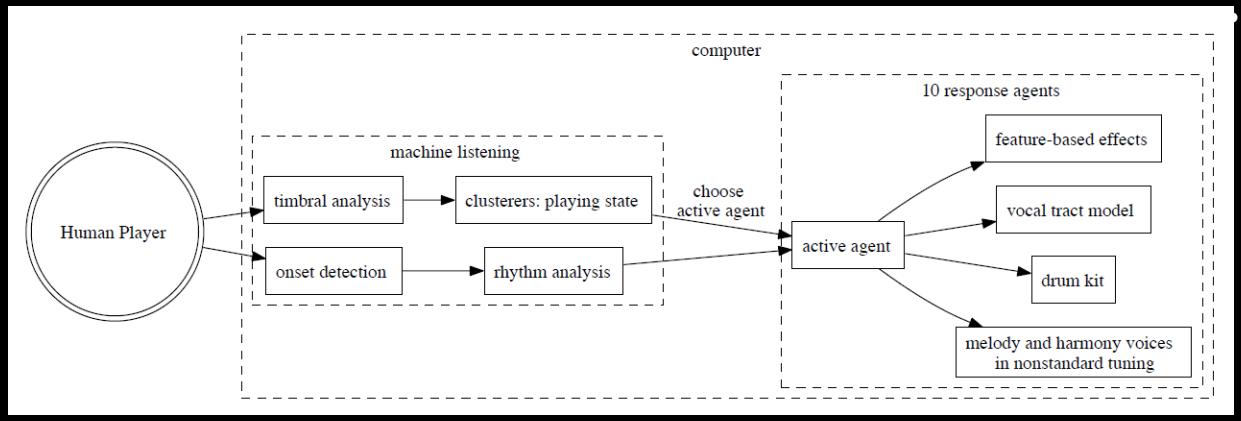
- A robotic marimba agent for interactive improvisation
  - Physical embodiment greatly helps the audience to enjoy the performance
  - Beat tracking and chord matching to adapt to human's tempo variation
  - Improvisation includes the choreographic aspect of the movement

Kinetic Engine (Multi-Agent)



Kinetic Engine (2005-2007), Arne Eigenfeldt

LL (ListeningLearning), Nick Collins, 2009

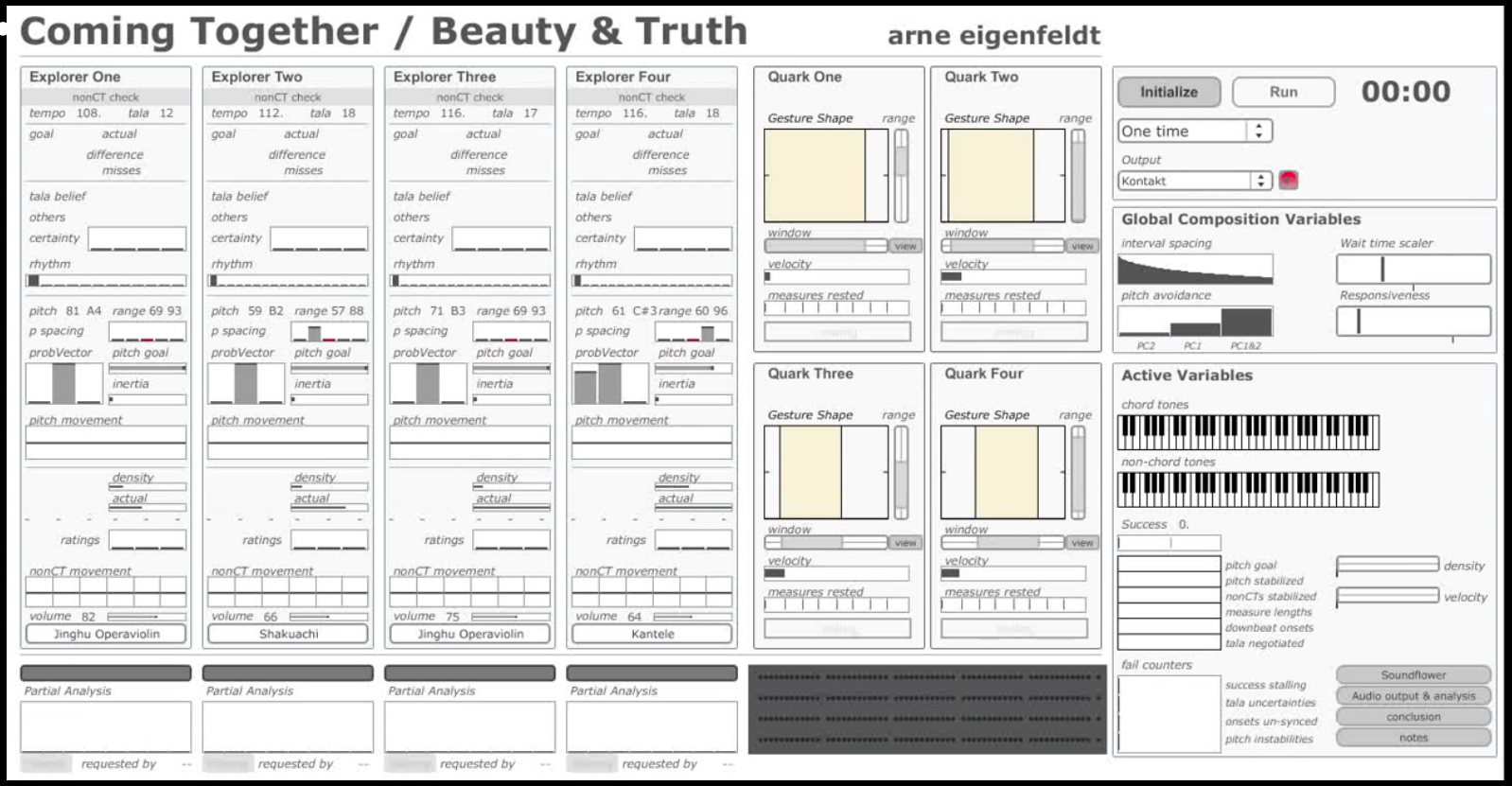


Rule-based system for free improvisation with humans

- Rhythm tracking: onset, inter-onset interval
- Silence detection: perceived loudness
- Timbral state clustering: using low-level acoustic features
- Generation: choose among 10 agents to follow the human's timbral state

Musical MultiAgents Systems

- Coming Together, Arne Eigenfeldt, 2010
  - Using the BDI architecture
  - Play 20s from 4:25 to 4:45



Reflexive Looper (Pachet et al., 2013)

https://www.youtube.com/watch?v=oquvn8GybRs

- A system allowing users to play with past virtual copies of themselves
  - Takes simultaneous MIDI and audio input: MIDI for analysis and audio for resynthesis
  - Uses an SVM classifier trained on MIDI data to classify the mode of user playing: bass, chords, and melody
  - Resynthesizes the other modes using past input audio

MUSEBOTS framework (2016-ongoing)

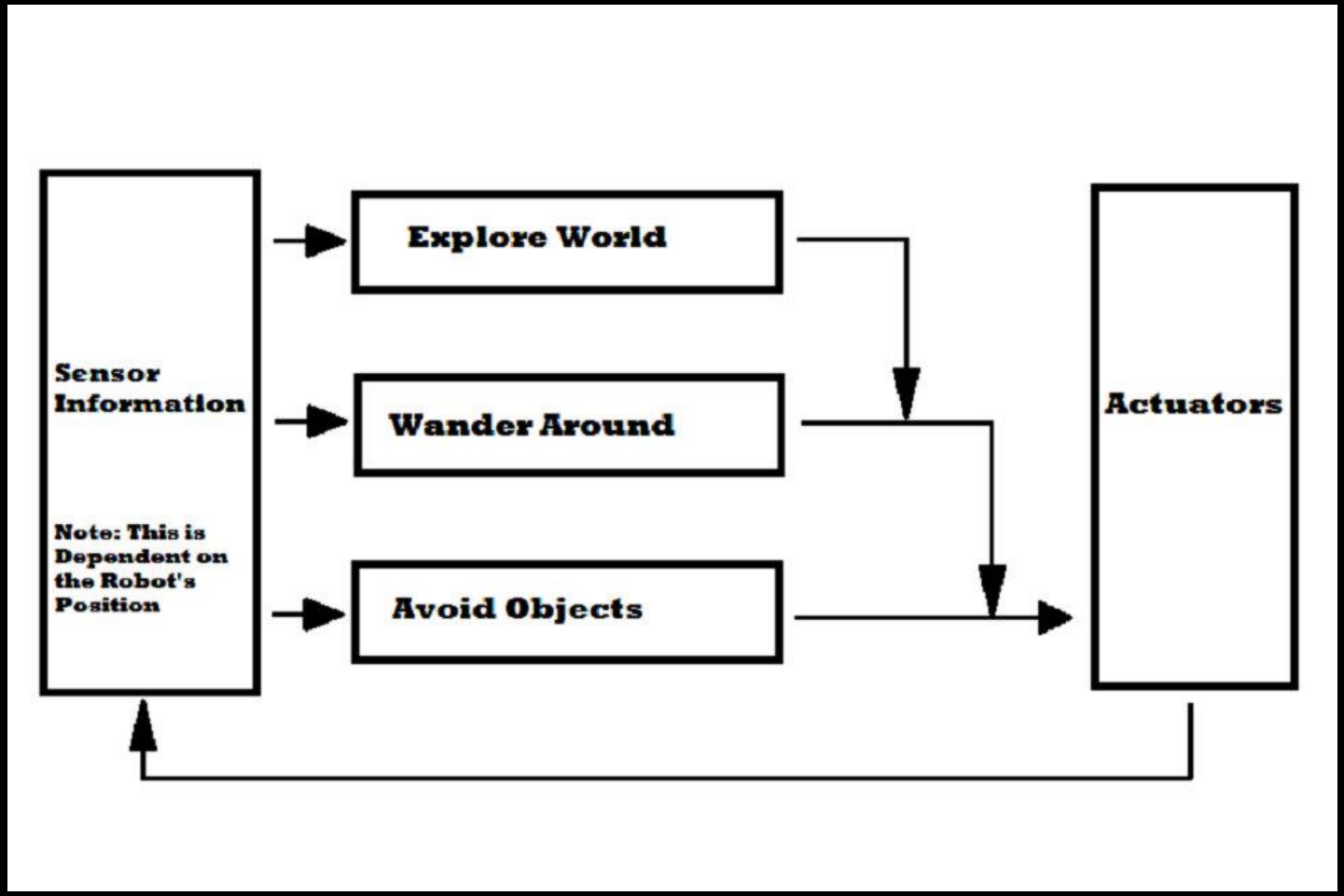
- A communication and interaction protocol for inter-agent operability.
- A collection of MAX (python, PD, …) musical agents:
  - Drummer bot
  - Bass bot
  - Melody bot
  - Conductor bot
  - …
- Developed by a variety of artists/researchers
- Performing together or alongside with humans

Brown, Andrew, Horrigan, Matthew, Eigenfeldt, Arne, Gifford, Toby, Field, Daniel, McCormack,Jon , Interacting with Musebots, New Interfaces for Musical Expression (NIME), 2018.

Agent architectures

- Three types of agent architectures:
  - Cognitive: maintain internal symbolic representations
    - Deliberative architectures: reasoning and planning
  - Reactive: no explicit representation of the environment and focus on behavioural rules
    - Reflex: no internal states (just mapping inputs to outputs), e.g., Braintenberg vehicle.
    - Reactive: with internal states (but not cognitive)
  - Hybrid: mixing reactive and cognitive components to balance reactiveness and deliberativeness

Subsumption Architecture

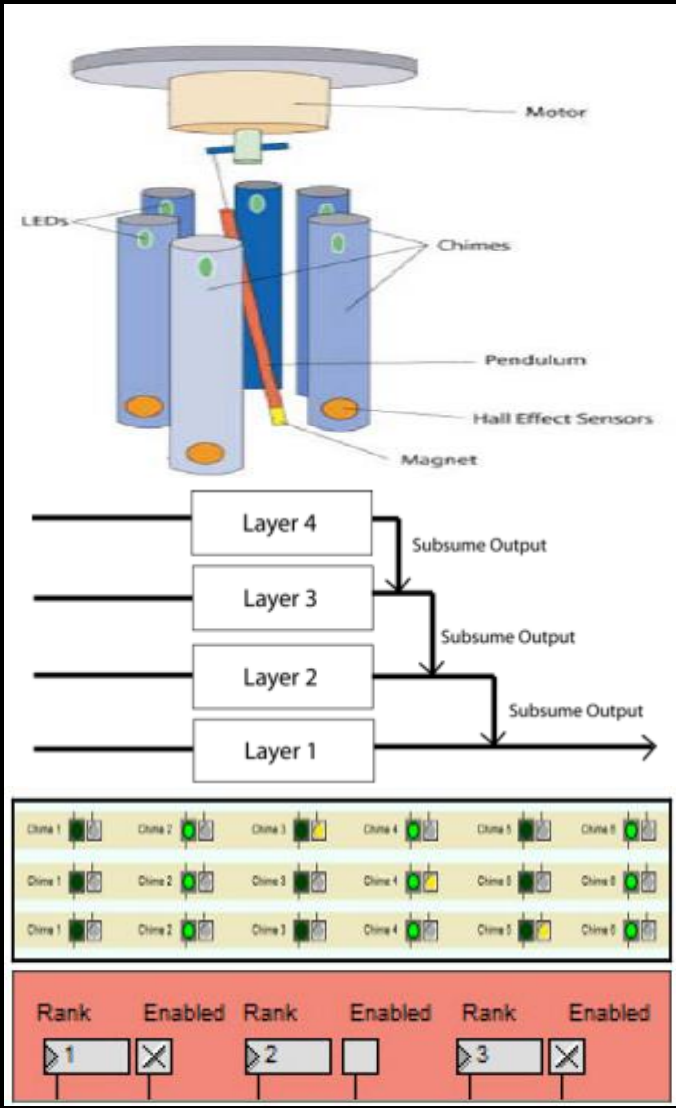


Musical Metacreation

With Aaron Levisohn
ACM ACE 2008

BeatBender: multi-agent rhythm generation

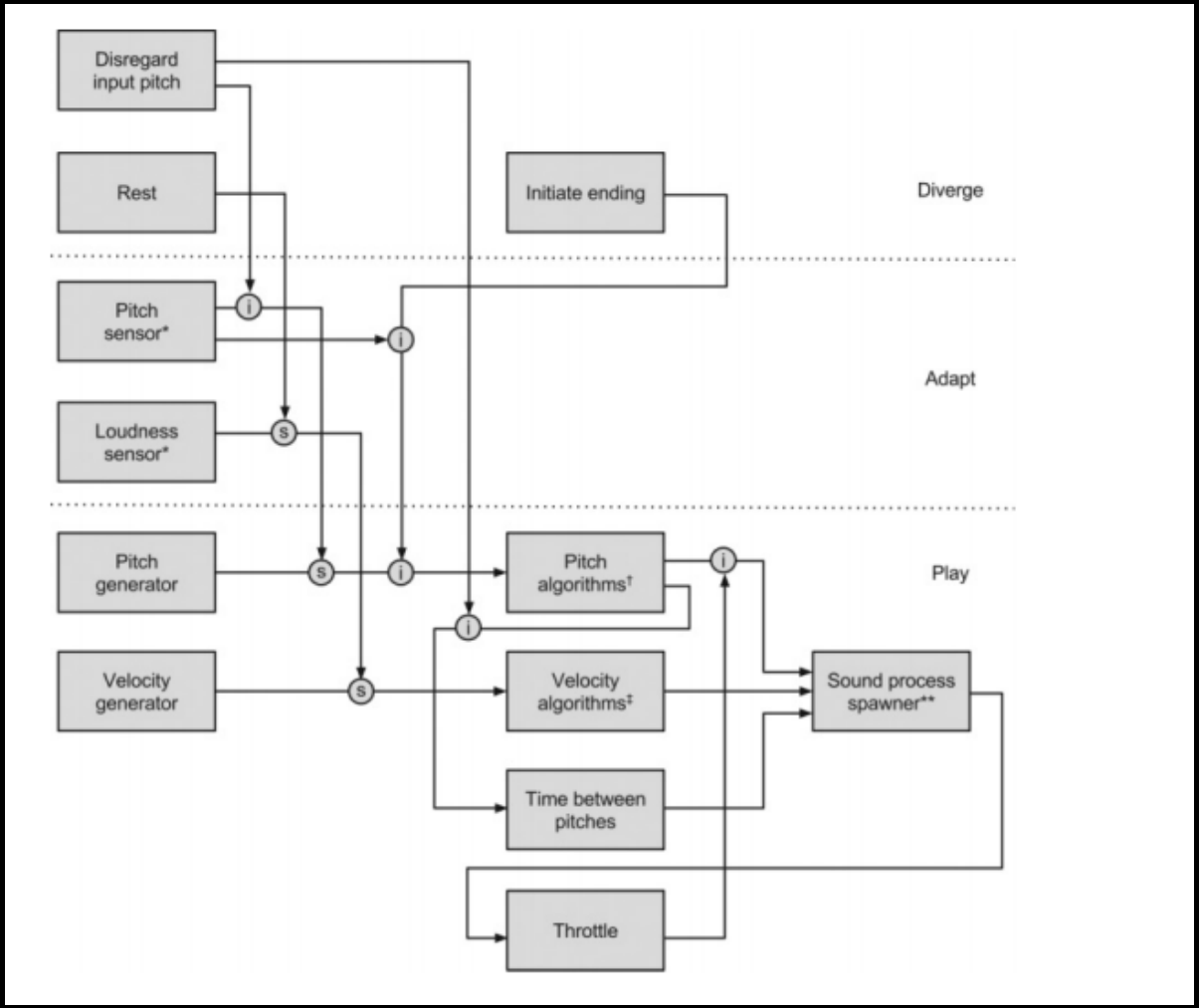
- Challenge: non-corpus based generation of rhythmic patterns
- Our approach:
  - Using reactive agents to create rhythmic patterns
  - Using subsumption agent architecture
- Experiments on a sample of 10+10 rhythms show that:
  - Humans prefer BeatBender rhythms over human composed ones
  - They find them more natural (less artificial)



Musical MAS – BeatBender

- The system models a drum circle with agents based on the subsumption architecture with four types of behavioral rules:
  - Neighborhood rules react to the status of the neighbors agents
  - Directed rules react to the status of specific agents
  - Collective rules react to the global activity of all the active agents
  - Temporal rules that use the history of the agent state
- Experiments show that complex rhythmic structures can be generated this way.

Musical agent – Odessa



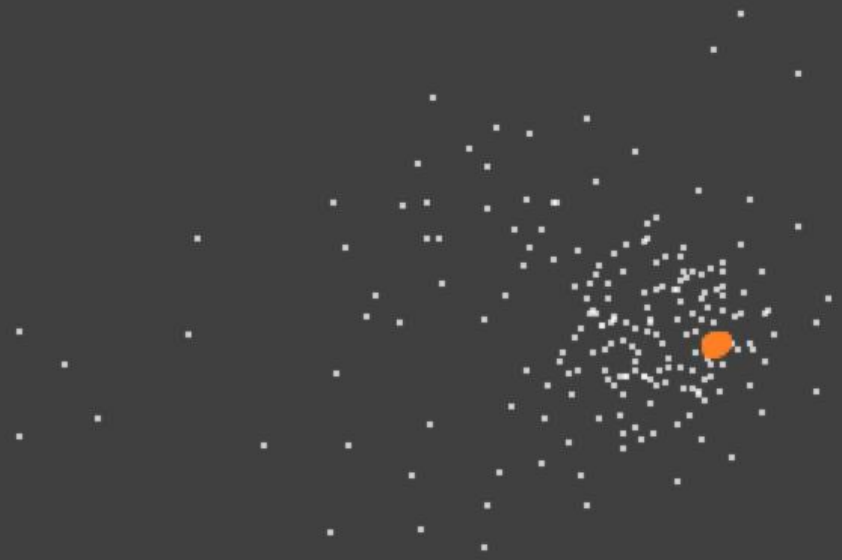
The Odessa musical agent, Adam Linson, Chris Dobbyn, George Lewis, and Robin Laney, 2012.

Musical agent – Odessa

- Here is an excerpt of the system in an improvisation with Adam Linson playing the double bass.

Boids and Swarms

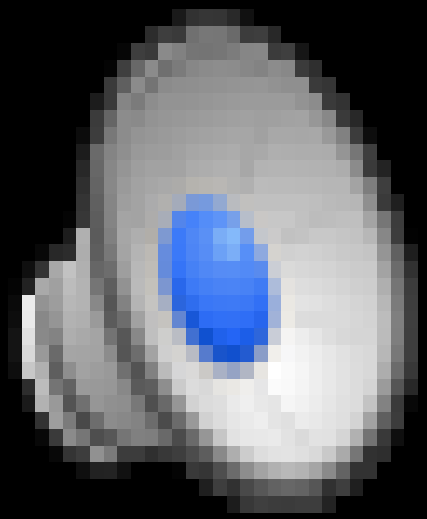
- A basic boid agent is implementing three simple behavioral rules:
  1. Avoidance: move away of a flock that is too close.
  2. Imitate: fly in the average direction/speed of the flock by averaging the velocity and direction of the other boids in the neighborhood.
  3. Center: Minimize exposure to the flock exterior by drifting towards the perceived center of the flock.



Reactive agents – Swarm Music

Excerpt of Autumn Leave, Time Blackwell, Swarm Music CD, 2002.

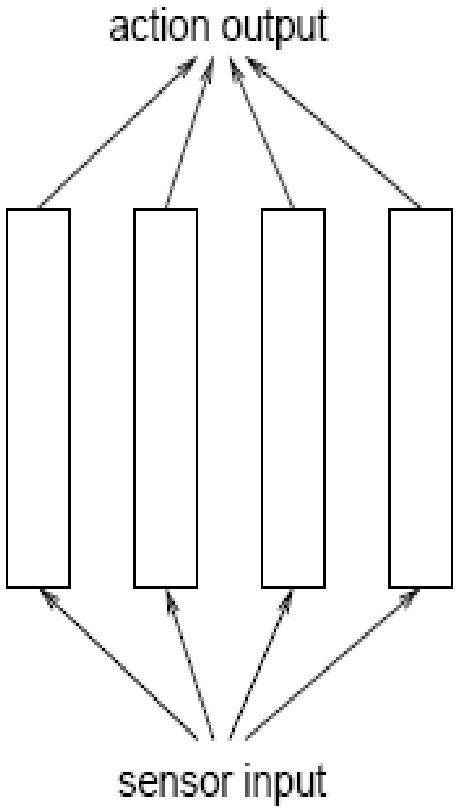
Musial agents – Porto Actors with Eargram



Porto actors with Eargram, Peter Beyls, Gilberto Bernardes, and Marcelo Caetano, 2015.

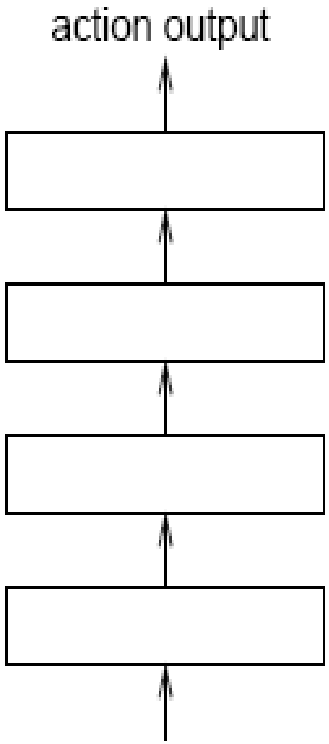
Layered Hybrid Architectures

Horizontal Layering

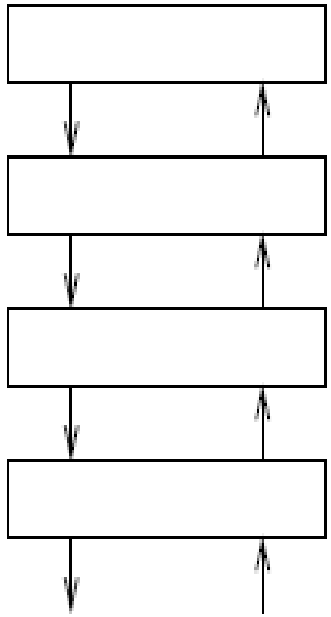


Vertical Layering

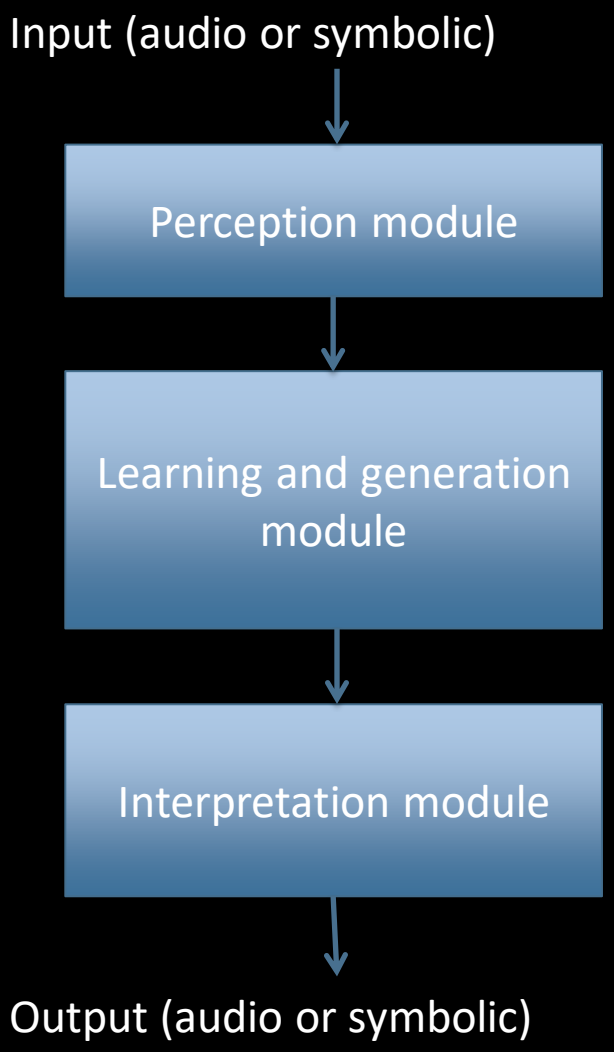
one-pass control



two-pass control



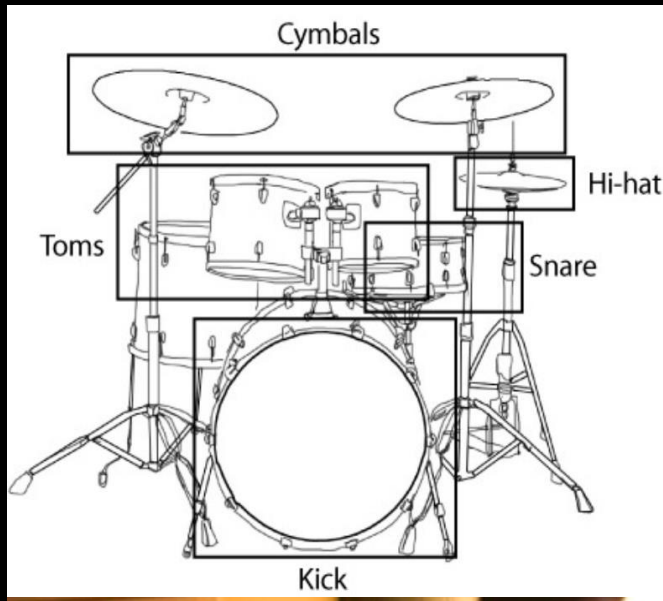
Generic Musical Agent Architecture



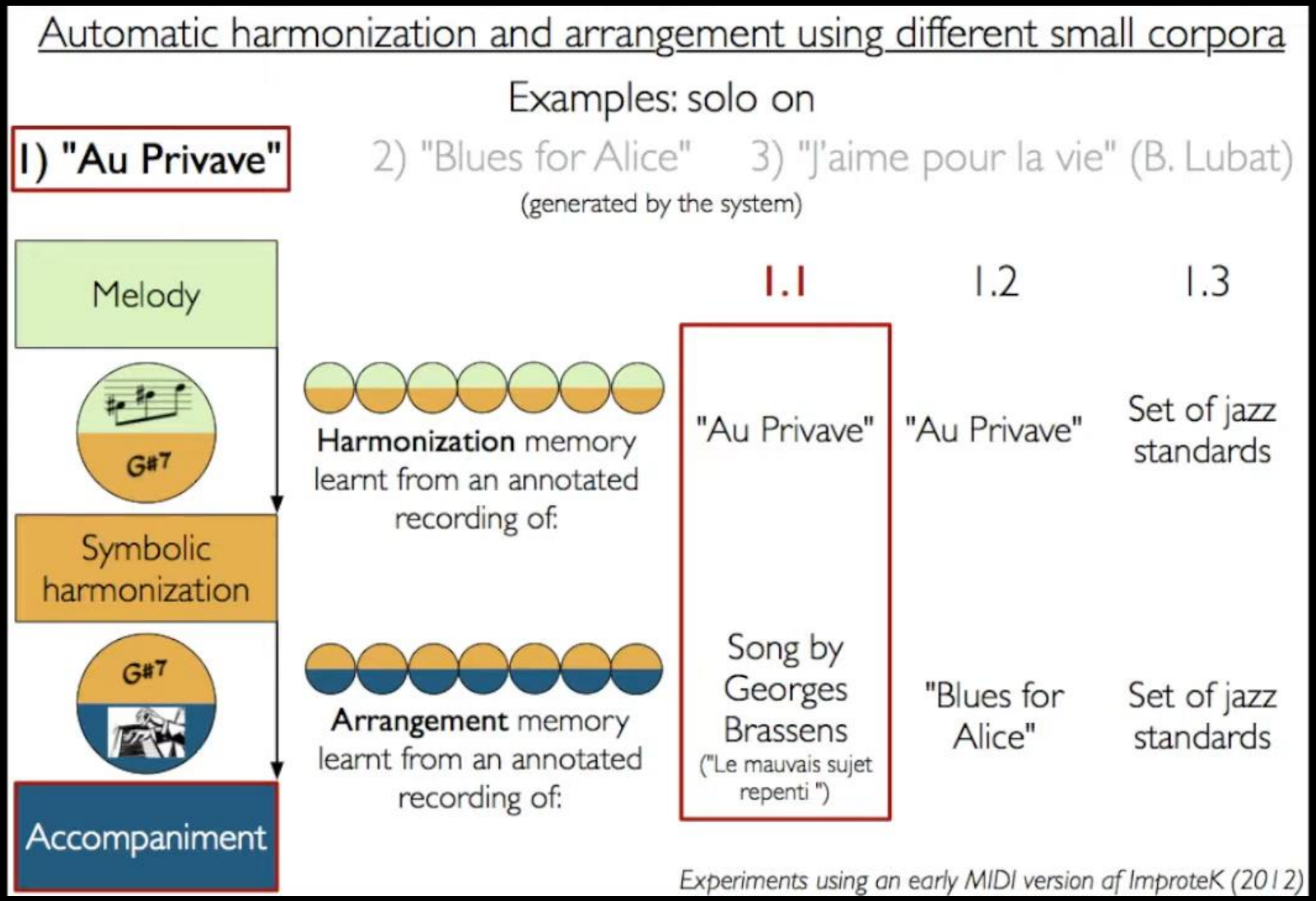
With Andrew Hawrishkewich
NIME 2010

BeatBack: Interactive Percussion System

- Multiagent interactive drum kit.
- Our solution:
  - Variable Order Markov models (VOMM)
  - Drum zoning
  - Call-response and accompaniment
- Empirical Evaluation:
  - Intrinsic Motivation Inventory
  - Quantitative analysis (MIDI+Matlab)
  - Humans prefer BeatBack to their usual drum kit both in learning and exploratory tasks



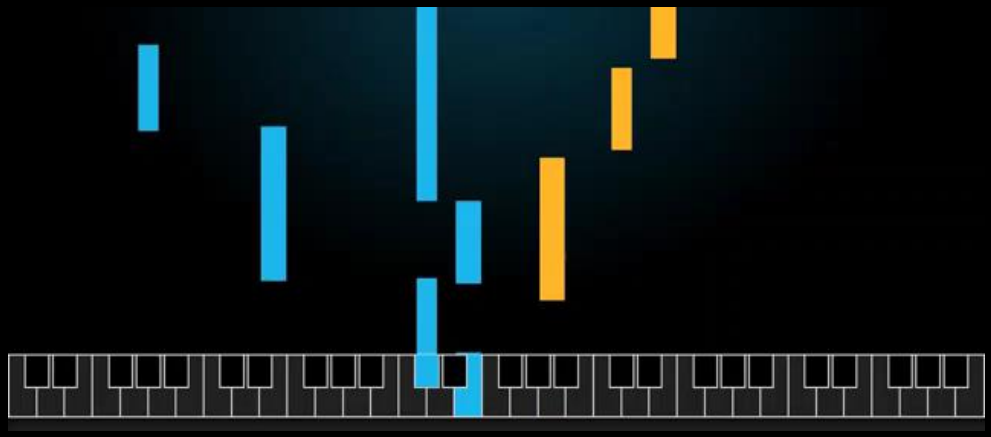
ImproteK, Jerome Nikka et al. 2012-ongoing



Musical agent – OMAX (2013)

Excerpt of a recording a variant of the system using the Variable Order Audio Oracle algorithm by Cheng-I Wang and Shlomo Dubnov, MUME, 2013.

A.I. Duet, Google Magenta, 2016



- A neural network model that responds to tunes played by the user on a MIDI keyboard using a similar style

Dyci2 agents, Ircam, 2017-ongoing

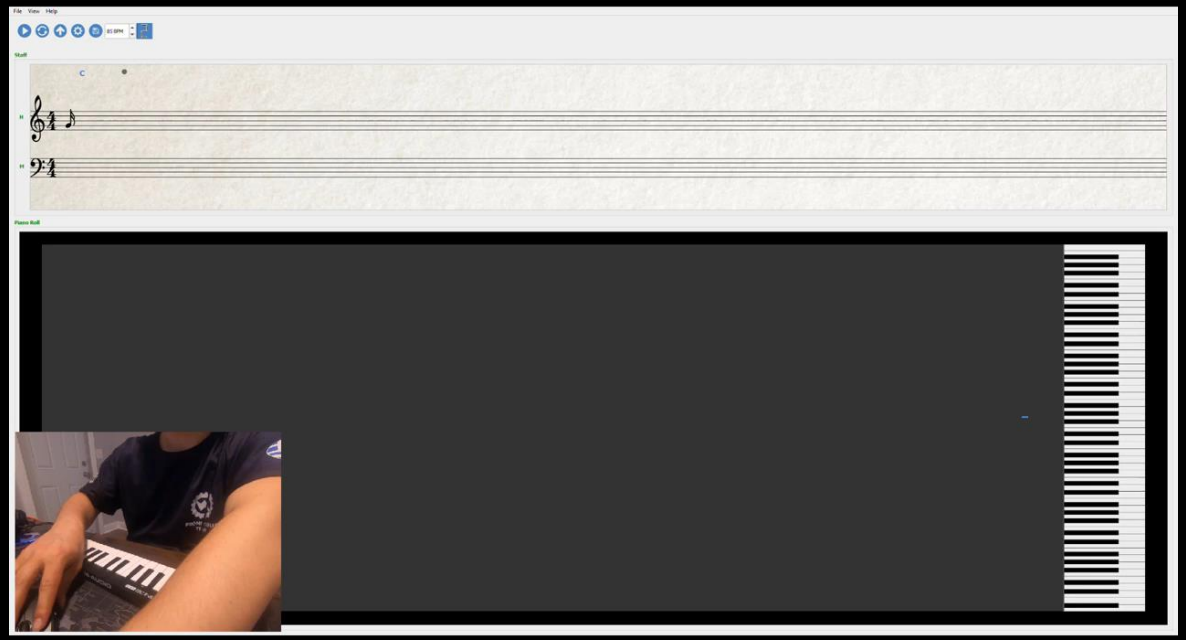
- Nika, J., Déguernel, K., Chemla, A., Vincent, E., & Assayag, G. (2017, October). Dyci2 agents: merging the" free"," reactive", and" scenario-based

Piano Genie, (Donahue, Simon, & Dieleman, 2019)

https://www.youtube.com/watch?v=YRb0XAnUpIk

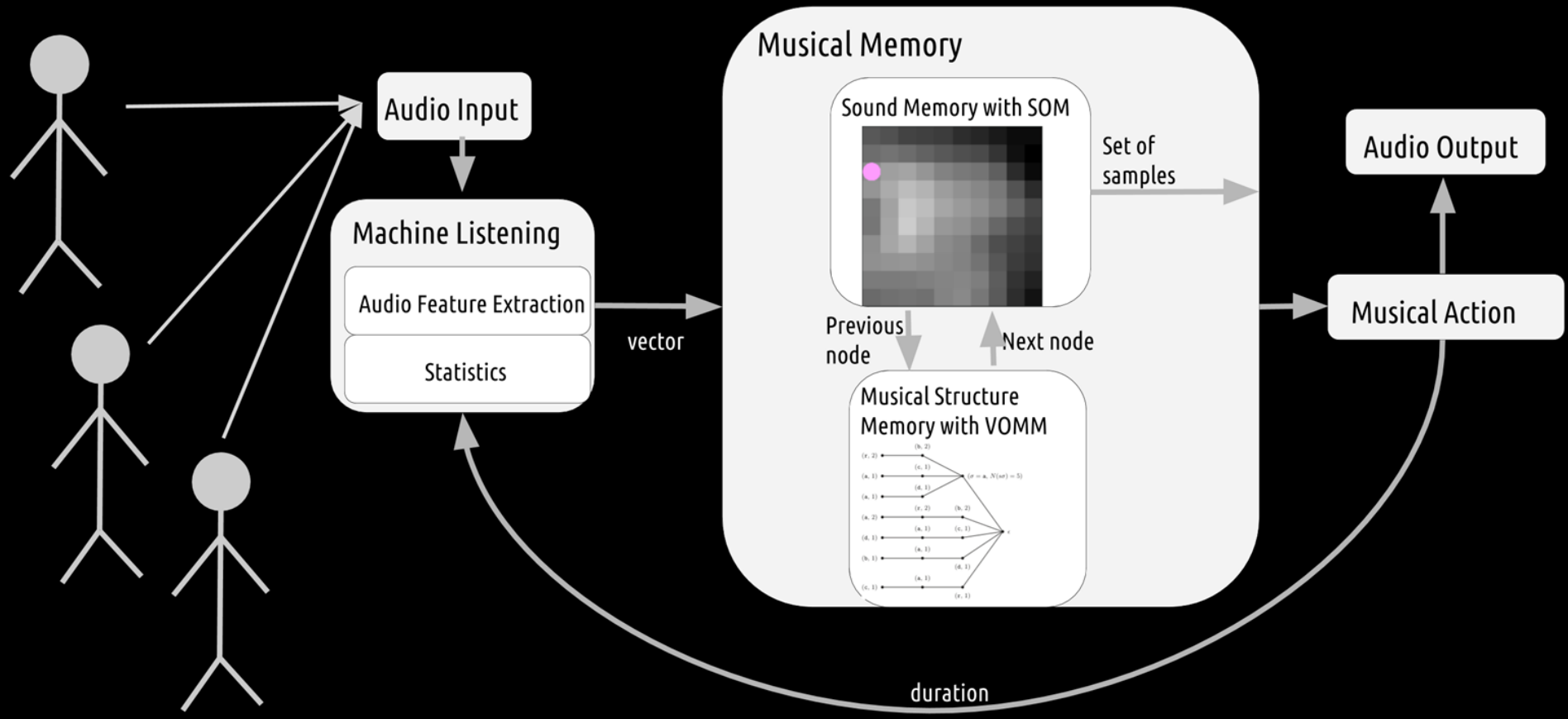
- Allowing users to improvise piano music on an 8-button controller
  - Uses an autoencoder to map note sequences in the 88-d space (corresponding to the 88 piano keys) to sequences in the 8-d space
  - Trained on 1400 piano performances by skilled pianists

BachDuet (https://bachduet.com/), Benetatos & Duan, 2019. *



- A neural network based system to allow human-AI duet improvisation in the style of Western counterpoint
  - Trained on outer voices of 370+ Bach chorales
  - Relatively equal role between human and AI – 6:4
  - Only supports MIDI input and fixed tempo

MASOM - Live performance



Tatar, K. & Pasquier, P. (2017). MASOM: A Musical Agent Architecture based on Self-Organizing Maps, Affective Computing, and Variable Markov Models. In Proceedings of the 5th International Workshop on Musical Metacreation (MuMe 2017)

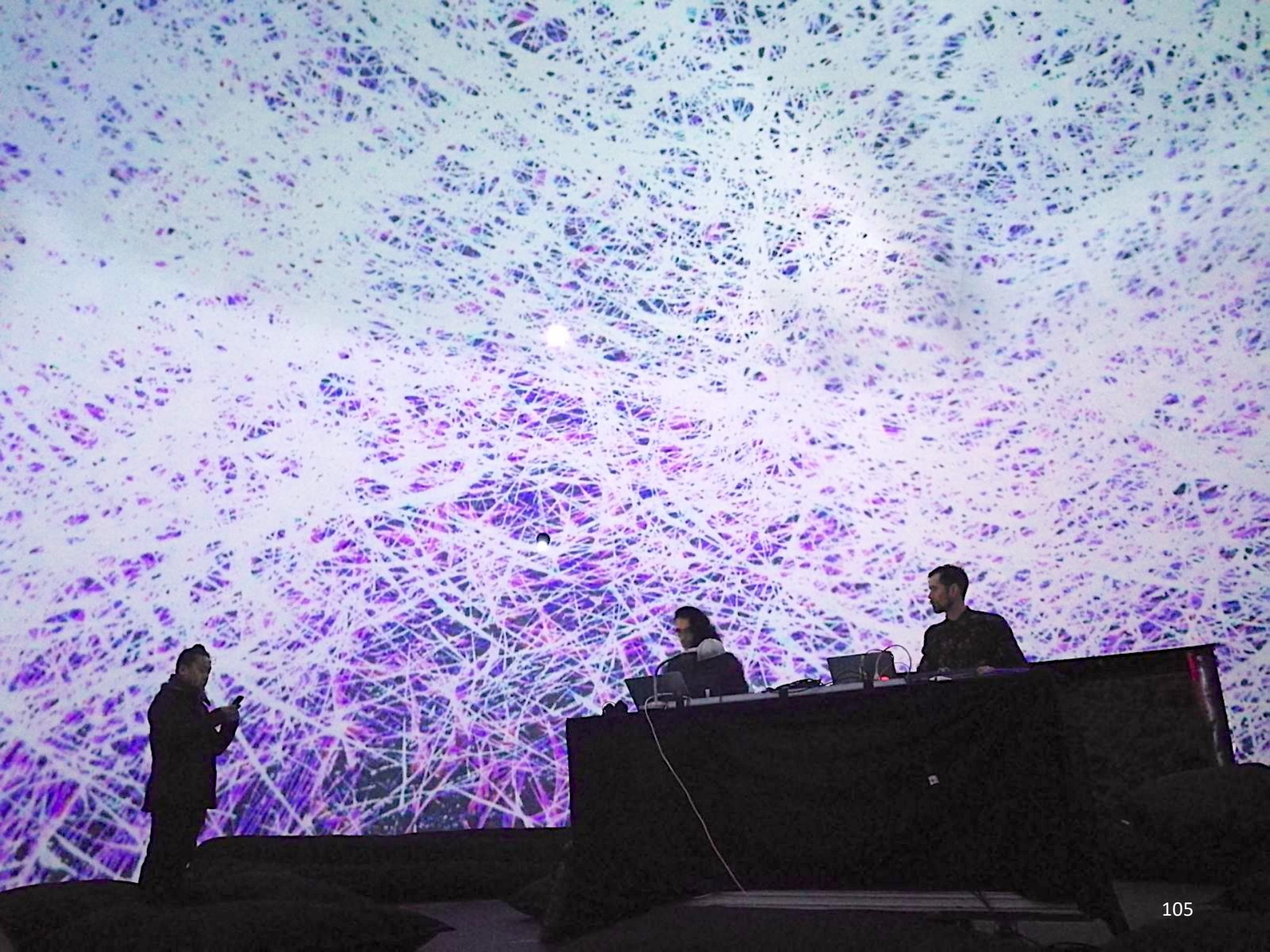
METACREATION
Lab for Creative Artificial Intelligence

SFU


/MASOM-factor-v1_04/REVIVE-1
IN
OUT
drop a folder here!
MASOM
START
Gain: -96 dB
Fade out
500 ms
Self-Organizing Map
STATS
Valence
Arousal
Mean
Std
Min
Max
NumData
0.8
Congruence
141
Perceived node
231
Played node
0.73
Sample Length
Randomize-feature-vis-OFF
MFCC_3-m...
SOM dimensions:
a x a
16
/MASOM-factor-v1_04/input
soundfile
KIVANC-16.wav
0 : 0 : 0
Transpose
0 st
Open
Play
Loop
Pan
0.00
Master Gain
0.00 dB
Audio On
DSP
/MASOM-factor-v1_04/output
Limiter
DC Filter On
Clear
Preamp
Threshold
Lookahead
Release
0.0 dB
-1.0 dB
64 samples
300 ms
Record
aiff
int24
Audio On
14.88
DSP
1
2
Pan
0.00
Master Gain
0.
dB

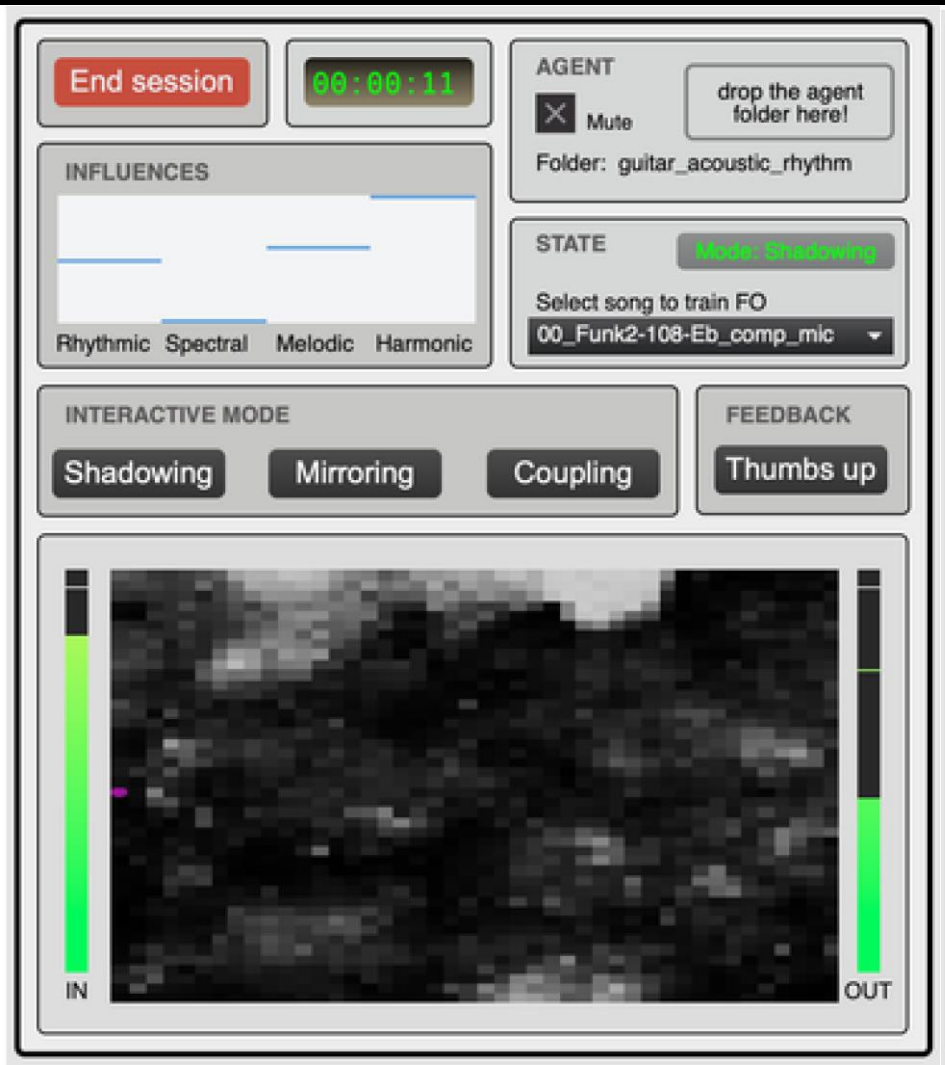
iOTA

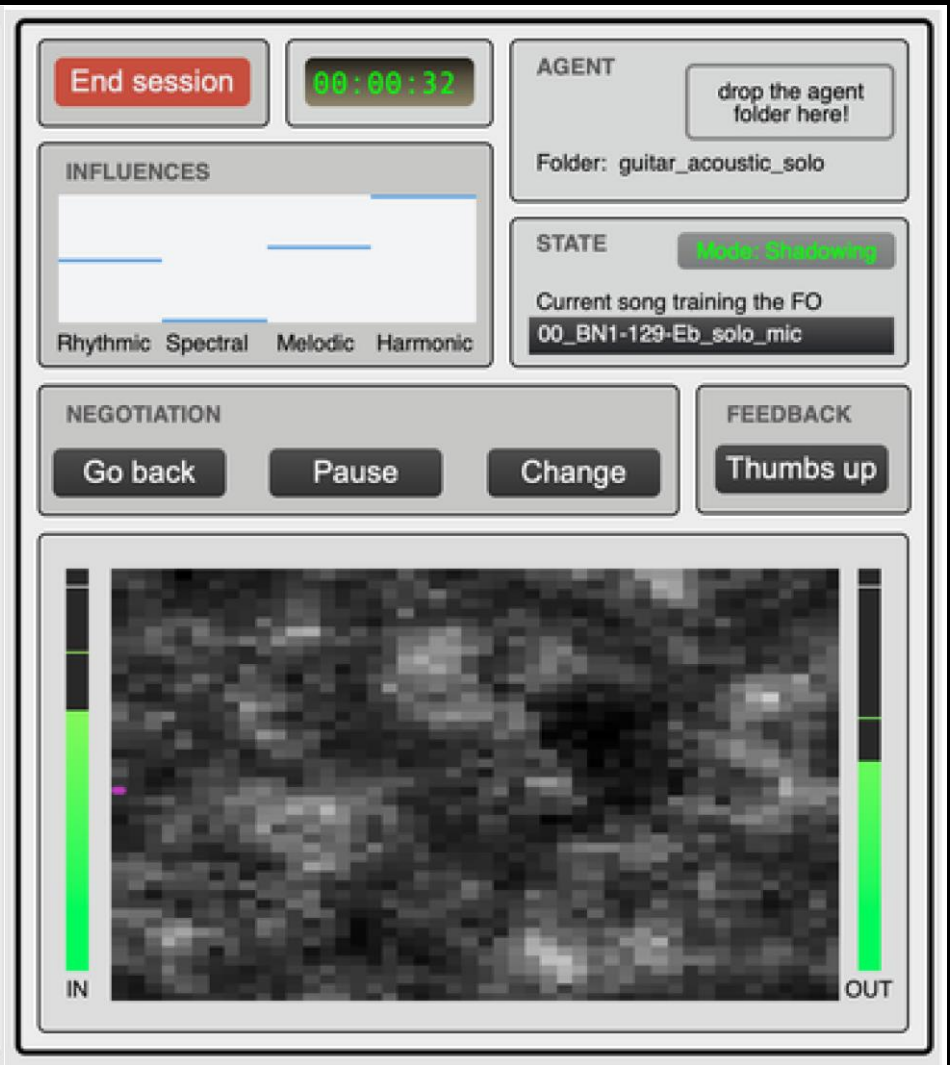
(collaboration with OUCHHH and Audiofil)



SpireMuse: musical agent

With Notto J.W. Thelle, NIME 2021.
Best paper Award.





Not enough evaluations/comparisons

Davis, Tommy, Kasey Pocius, Vincent Cusson, Marcelo M. Wanderley, and Philippe Pasquier. “eTu{d,b}e: Case Studies in Playing with Musical Agents.” In *Proceedings of the 2023 conference on New Interfaces for Musical Expression*, 2023.

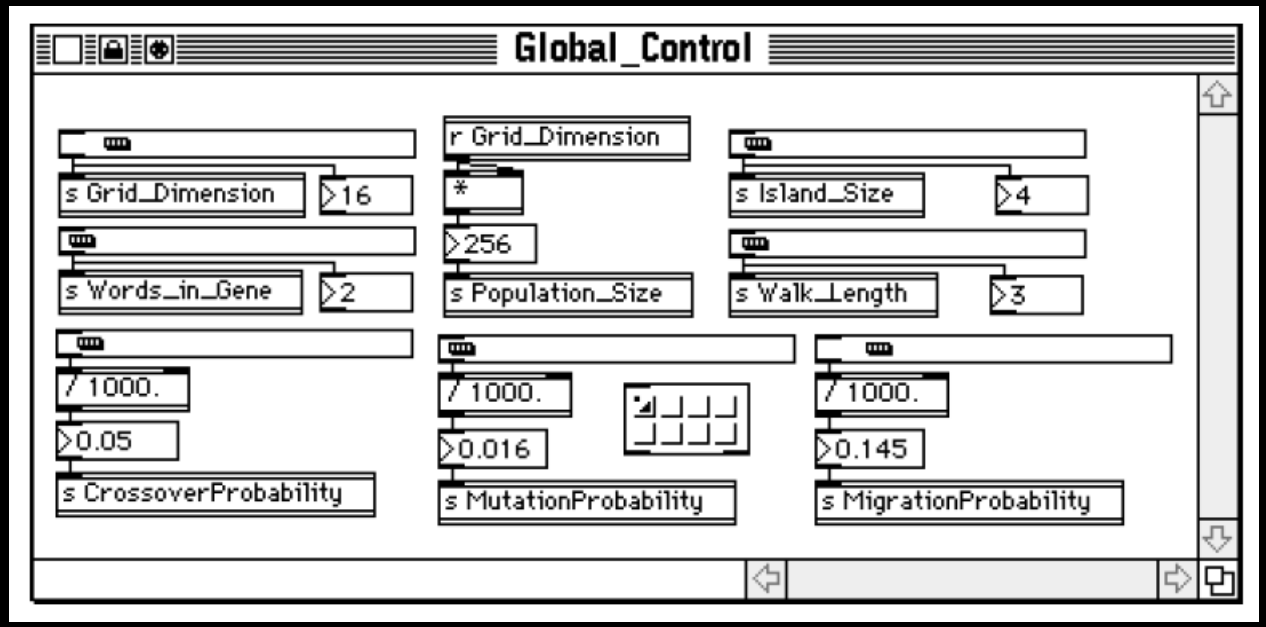
Outline of the Tutorial

- Historical precedents
- Musical Agents:
  - Cognitive agents
  - Reactive agents
  - Hybrid agents
  - Virtual Ecosystems (Artificial Life)
- Computer-Assisted Composition:
  - Audio domain
  - Symbolic domain

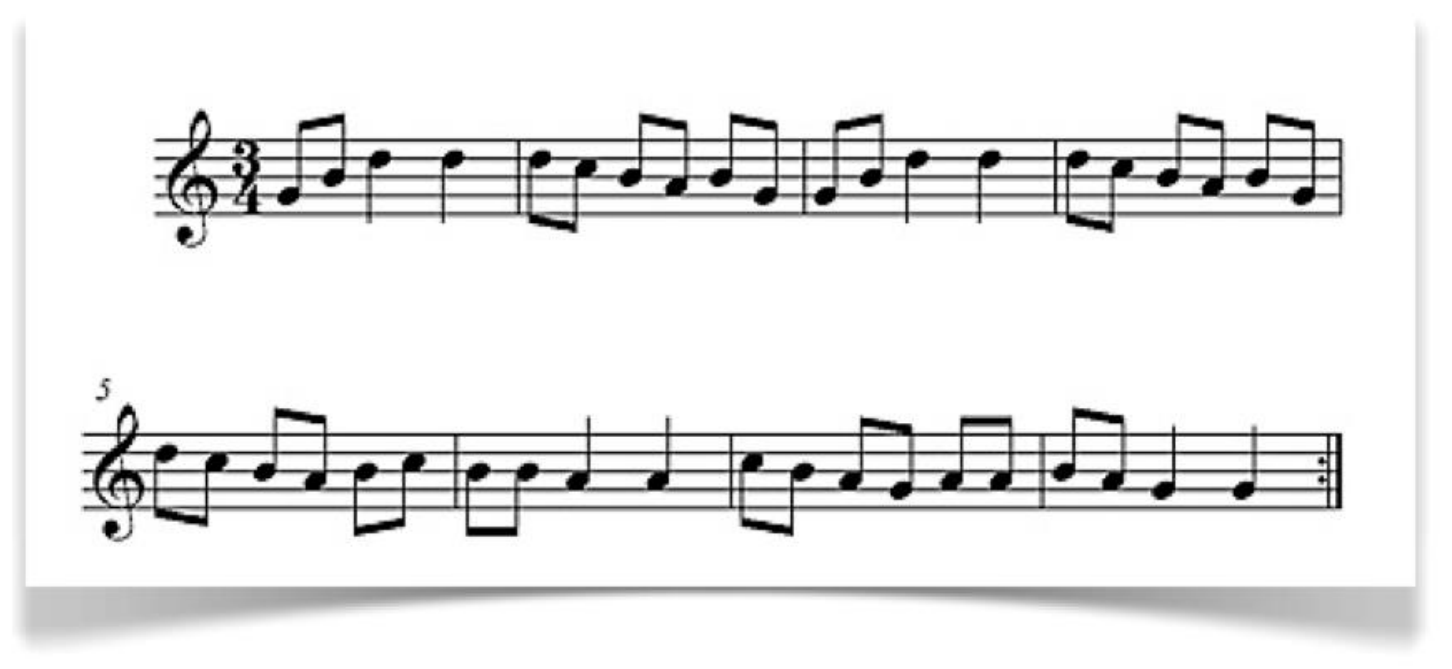
(Evolutionary) ecosystem

- Mixing agents, ecosystems, and evolutionary computing

Purely sonic ecosystems

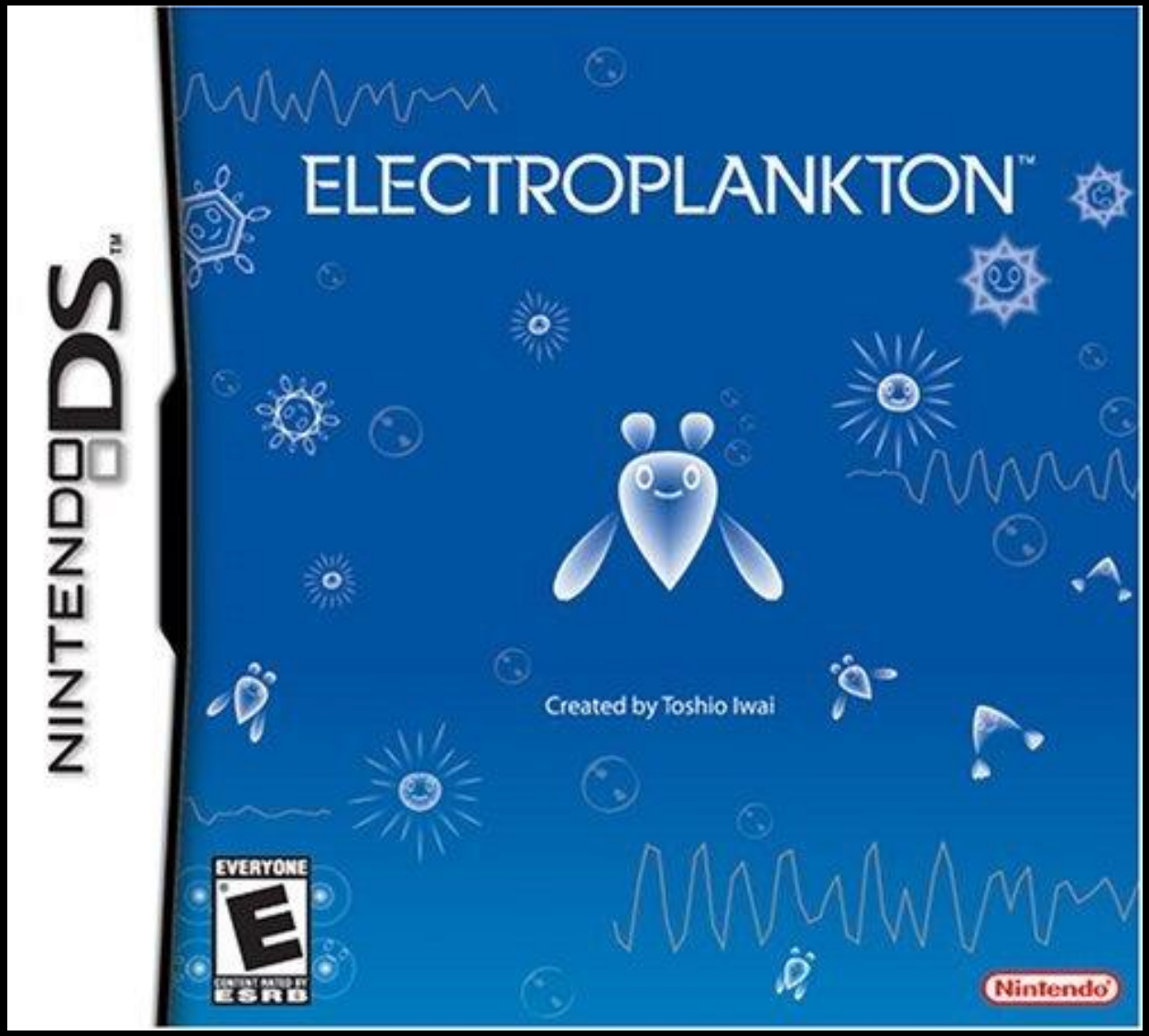


Sonomorphs: An application of genetic algorithms to the growth and development of musical organisms, Gary L. Nelson, 1993.



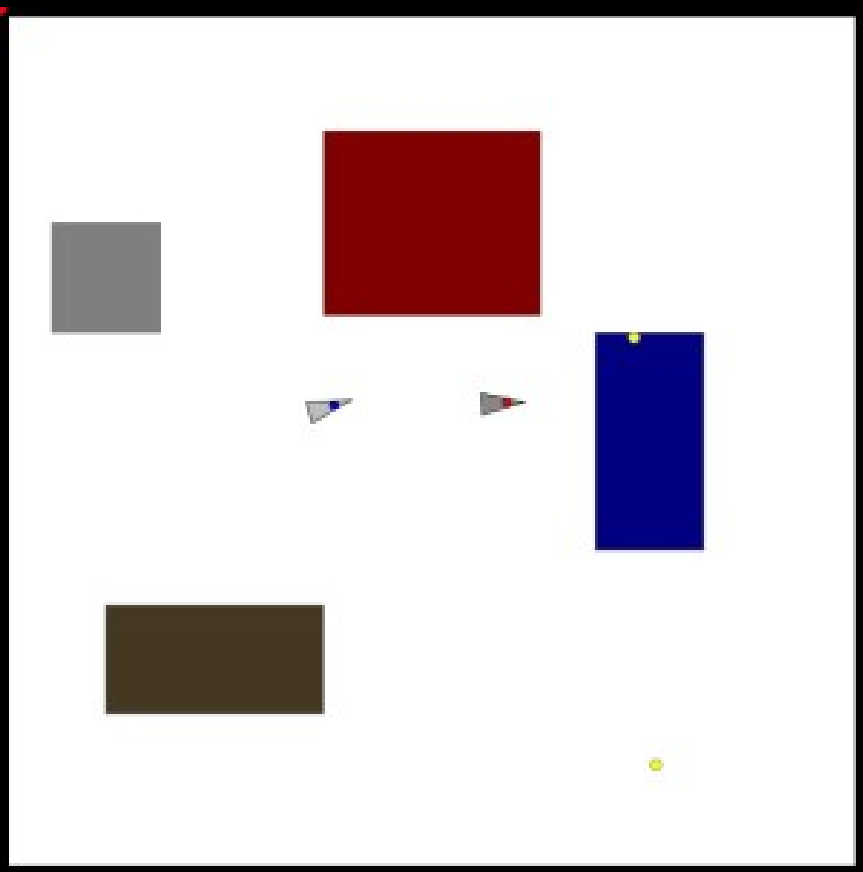
Living melodies, Palle Dahsteldt and Mats G. Nordhal, 2001.

Listening Sky, Rodney Berry, Wasinee Rungsarityotin, Alan Dorin, 2001.



ElektroPlancton, Indieszero, Nintendo DS, 2005.

Amar, Arne Eigenfeldt, 2009.



Genesynth, Anees Vartakavi, 2013.

Conclusion on Agents and MAS

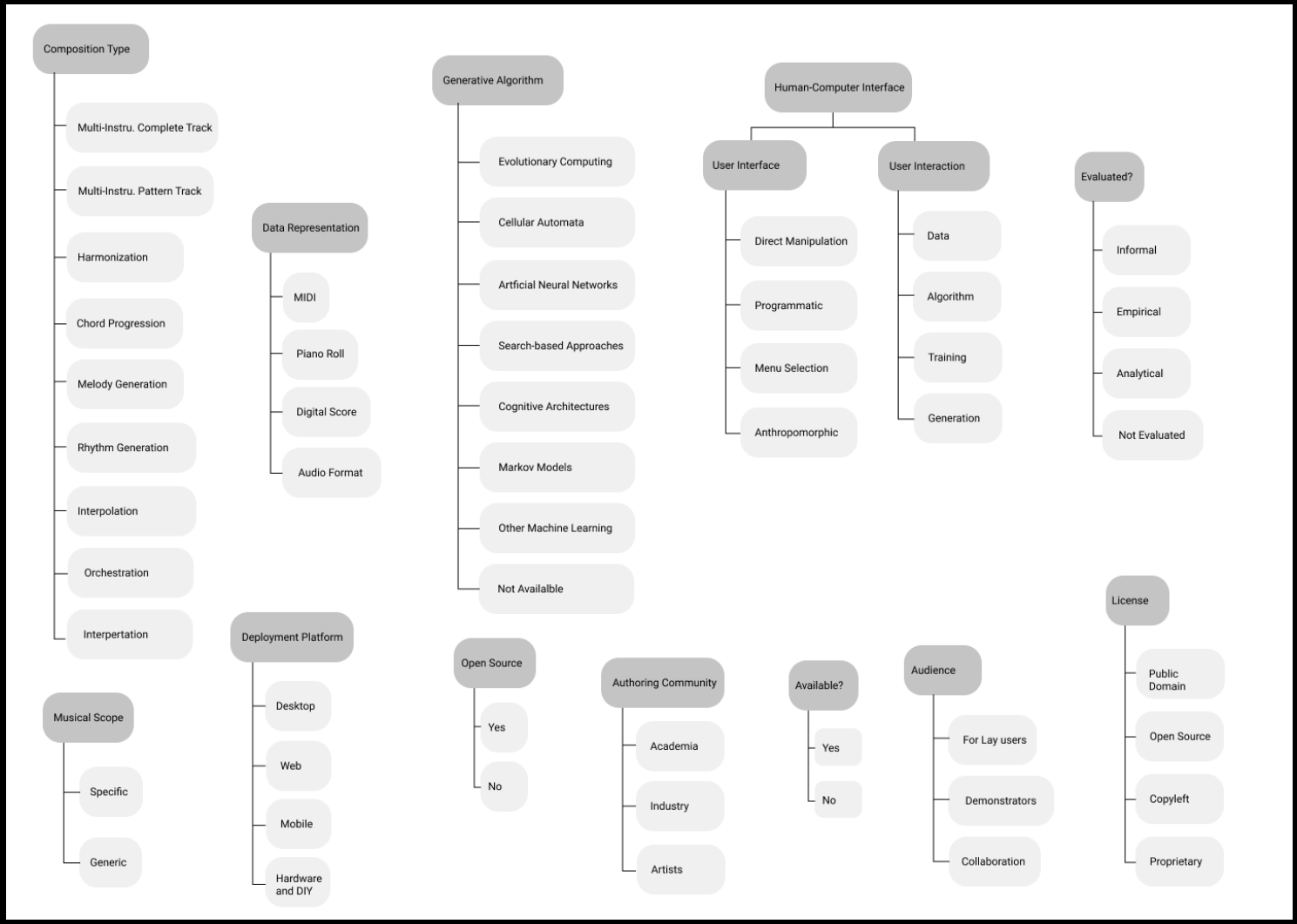
- Cognitive approaches propose a top-down solution to:
  - Agent design: the agent architecture, and its decision process.
  - Society design: The organization of the MAS uses roles, conventions and protocols. Group goals, are broken down into individual goals, themselves broken down in sub-goals, reified as intentions, achieved through planning sequences of actions.
- Reactive AI proposes a bottom-up emergent solution to:
  - Agent design: the agent behavior emerges from the interaction between its behavioral rules
  - Society design: the MAS behavior emerges from interaction the agents with their environment.
- Hybrid architecture mary both approaches

Pros and Cons of musical agents

- Pros:
  - Online
  - Agents offer and anthropocentric conceptualization
  - Modelling Flexibility
- Cons:
  - More a modeling paradigm than an algorithm
  - Complexity of real-time (machine listening, anticipation models,...)
  - Complexity of decentralized systems in the MAS case
  - Hard to Evaluate: see E{tudbe}, NIME 2023.

Computer Assisted Composition

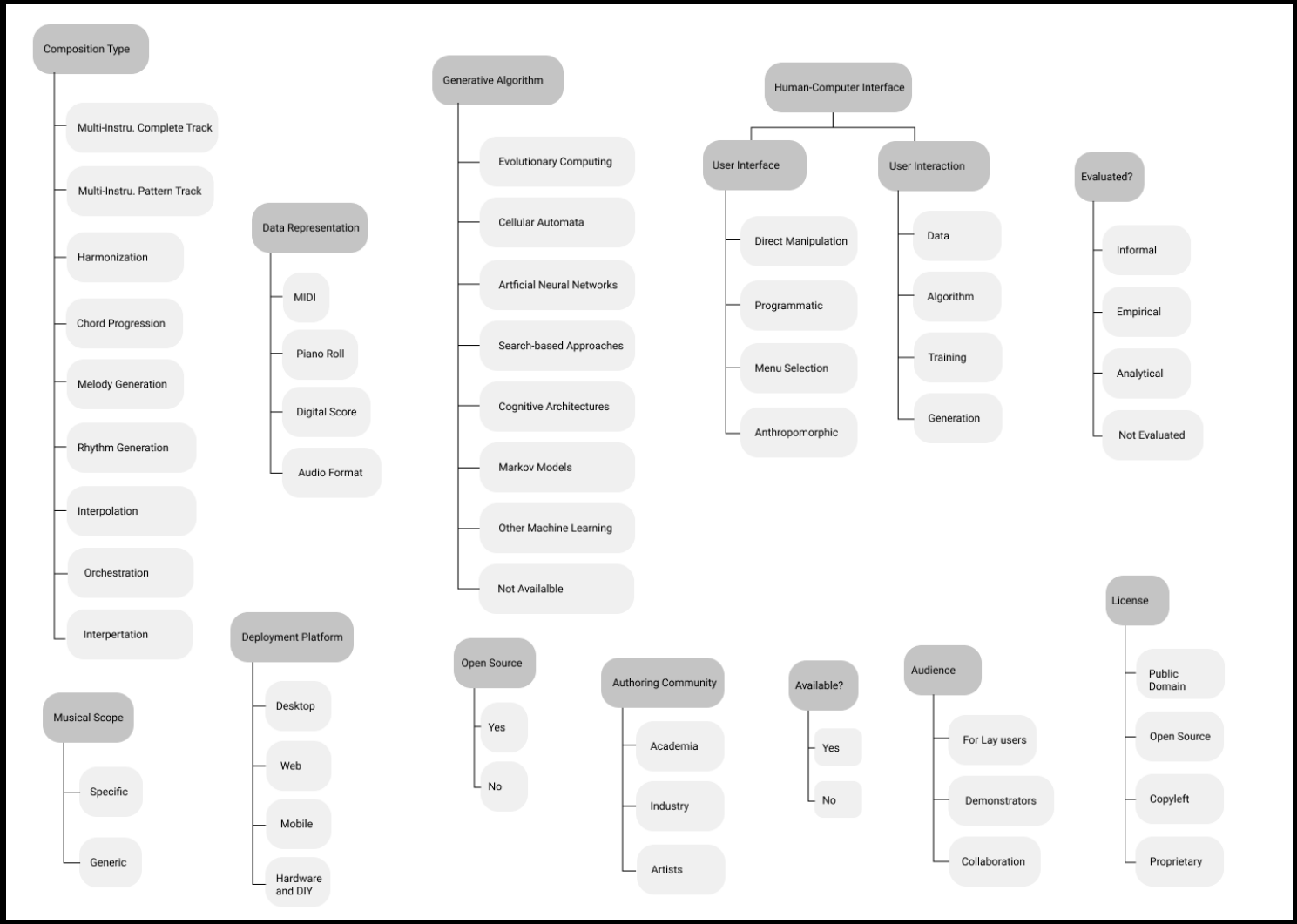
Taxonomy of CAC Systems



Taxonomy of CAC systems (75+)

CAC Systems	Year	Composition Task	Data Representation	Algorithm	Audience	Deployment Platform	Evaluation Method
Magenta Studio	2019	Melody, Chords, Rhythm	MIDI	Deep Learning	Amateurs	Desktop / Ableton	Formal
MuseNet	2019	Multi-Track	MIDI	Deep Learning	Amateurs	Web Demo	Formal
MMM4Live	2021	Multi-Track	MIDI	Transformer	Pro	Ableton	*Pending*
Apollo	2018	Multi-Track	MIDI	Machine Learning	Amateurs + Pro	Web	*Pending*
FolkRNN	2018	Multi-Track	MIDI	RNN	Amateurs + Pro	Web	Informal
Flow Machines	2013	Melody	MIDI	Markov Models	Pro	iPad / Mac VST	Formal
AIVA	2016	Multi-Track	Audio	Proprietary	Amaterus	Web	-
Spliqs	2016	Multi-Track	MIDI	Proprietary	Lay users	iPad	-
Jukedeck	2012	-	-	Proprietary	-	Web	-
Amper Music	2014	Multi-Track	Audio	Proprietary	Content creators	Web	-
Melody Sauce	2019	Melody	MIDI	Proprietary	Amateur, Pro	VST	-

Taxonomy of CAC Systems



CAC Systems examples

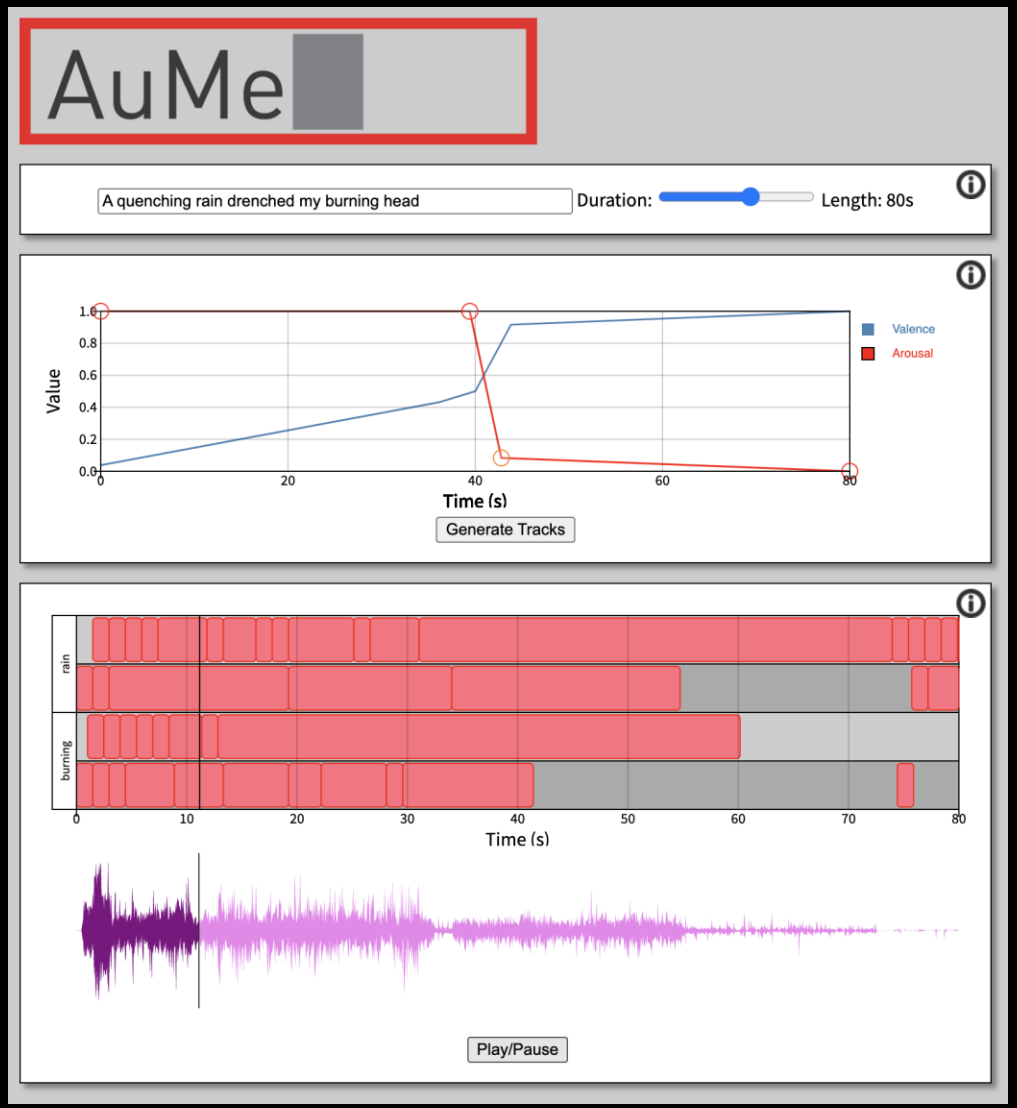
- **Audio output:**
  - AUME (2012)
  - Amper Score (2014) – video scoring
  - Mubert Pro (2016) – generative music stream
  - AIVA (2016) *
  - Boomy (2018)
  - Ecrett (2018)
  - Soundraw (2020)
  - MusicML (2023)
  - SongR (2023)
  - Riffusion (2023)
  - Halo (2018-2023)

- **MIDI output:**
  - EMI (1995-)
  - Imporvisor (2005-2020, Pr. Bob Keller,
  - Harmonic Progression Generator (2010)
  - MusiCog, Manuscore (2012)
  - Style Machine Lite (2012-2016)
  - FlowMachines (2013-2021)
  - FolkRNN (2018)
  - MuseNet (2019)
  - DrumVAE (2019)
  - Magenta Studio (2019)
  - Orchidea (2020, orchestration)
  - Draw and Listen (2022)
  - MMM4Live (2023)
  - Calliope (2023)
  - MMM-C (2023)

Audio Metaphor

With Miles Thorogood, Jianyu Fan et al.,
ICCC, JNMR, SMC, 2012-2022.

Try it at: https://audiometaphor.ca/



Prompt + duration

+

Valence and arousal affect curves

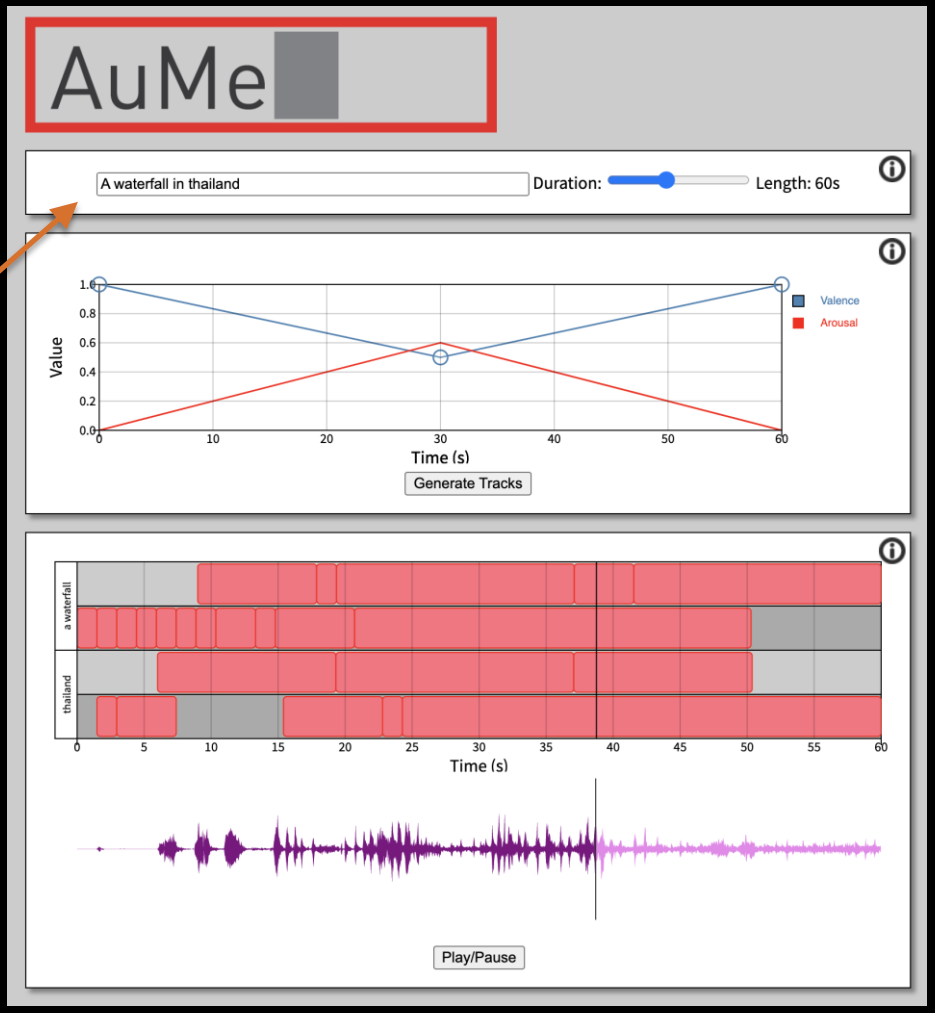
Multitrack sequence, mix (AAC)

Audio rendering (44k, stereo)

Audio Metaphor

With Miles Thorogood, Jianyu Fan et al.,
ICCC, JNMR, SMC, 2012-2022.

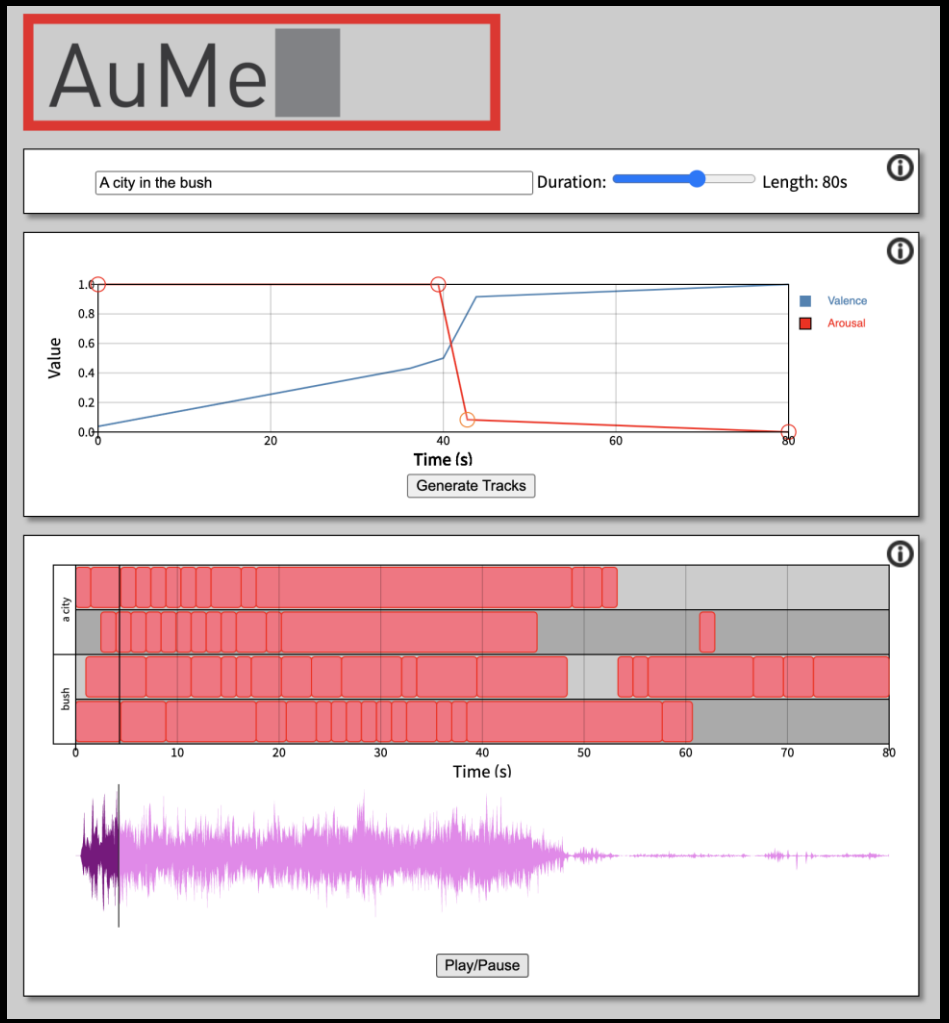
A waterfall in Thailand



Try it at: https://audiometaphor.ca/

Audio Metaphor

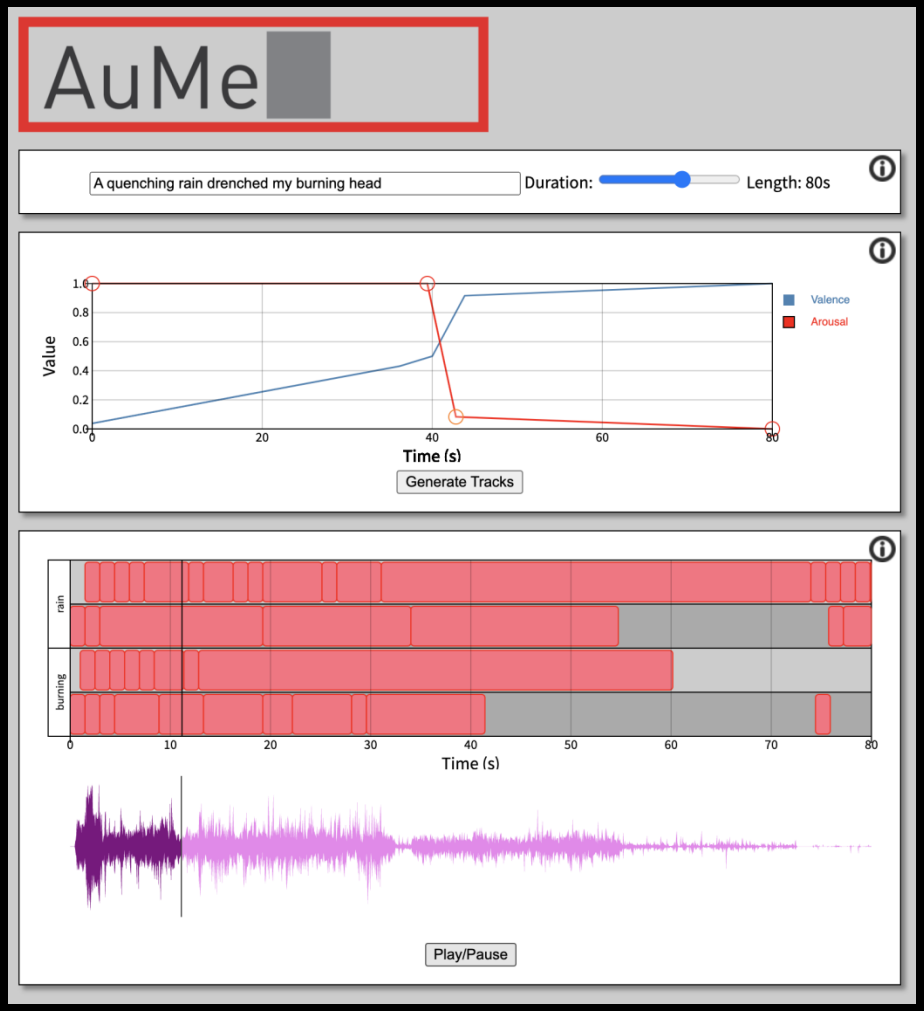
A city in the bush



Try it at: https://audiometaphor.ca/

Audio Metaphor

A quenching rain drenched my burning head

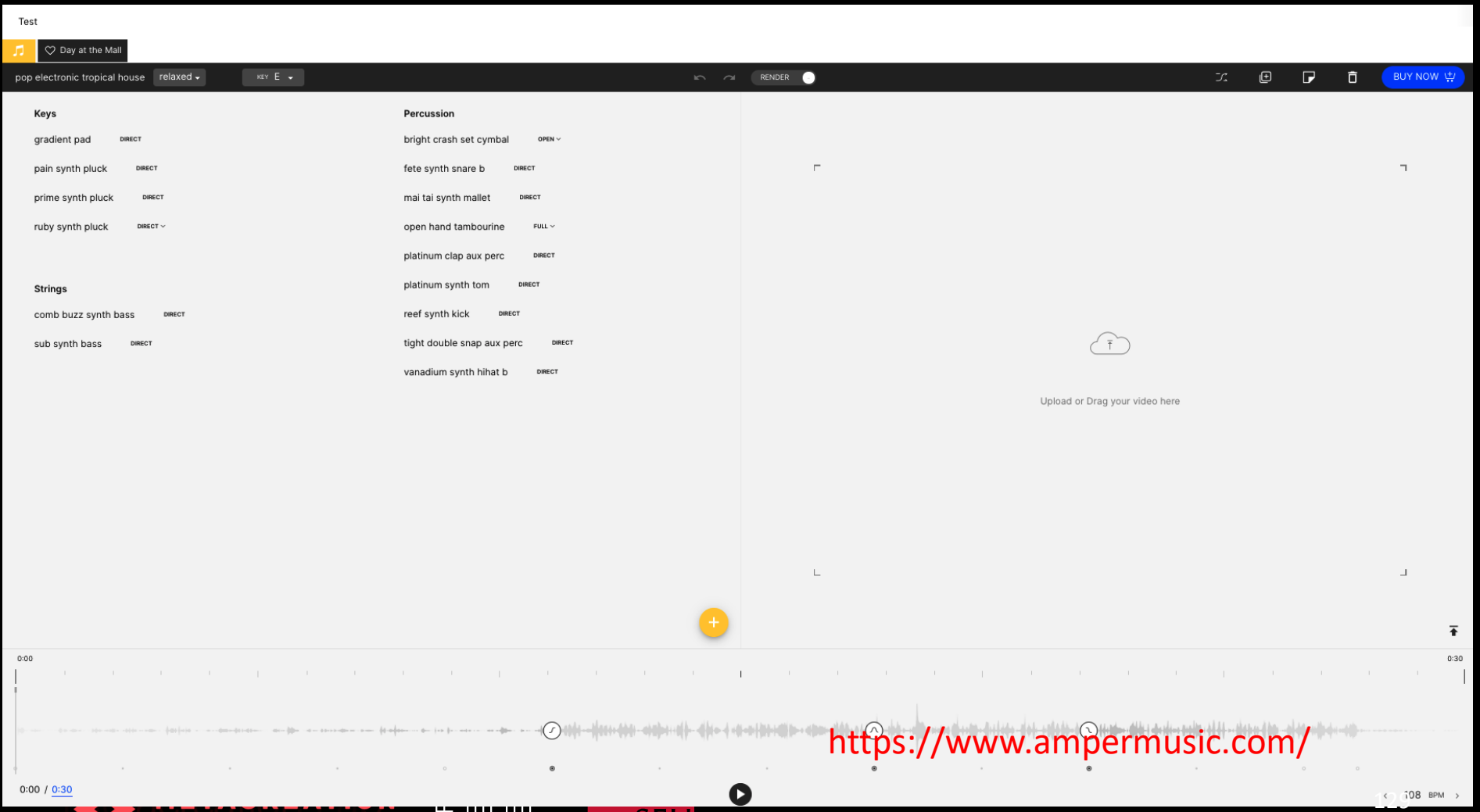


Try it at: https://audiometaphor.ca/

Amper Score

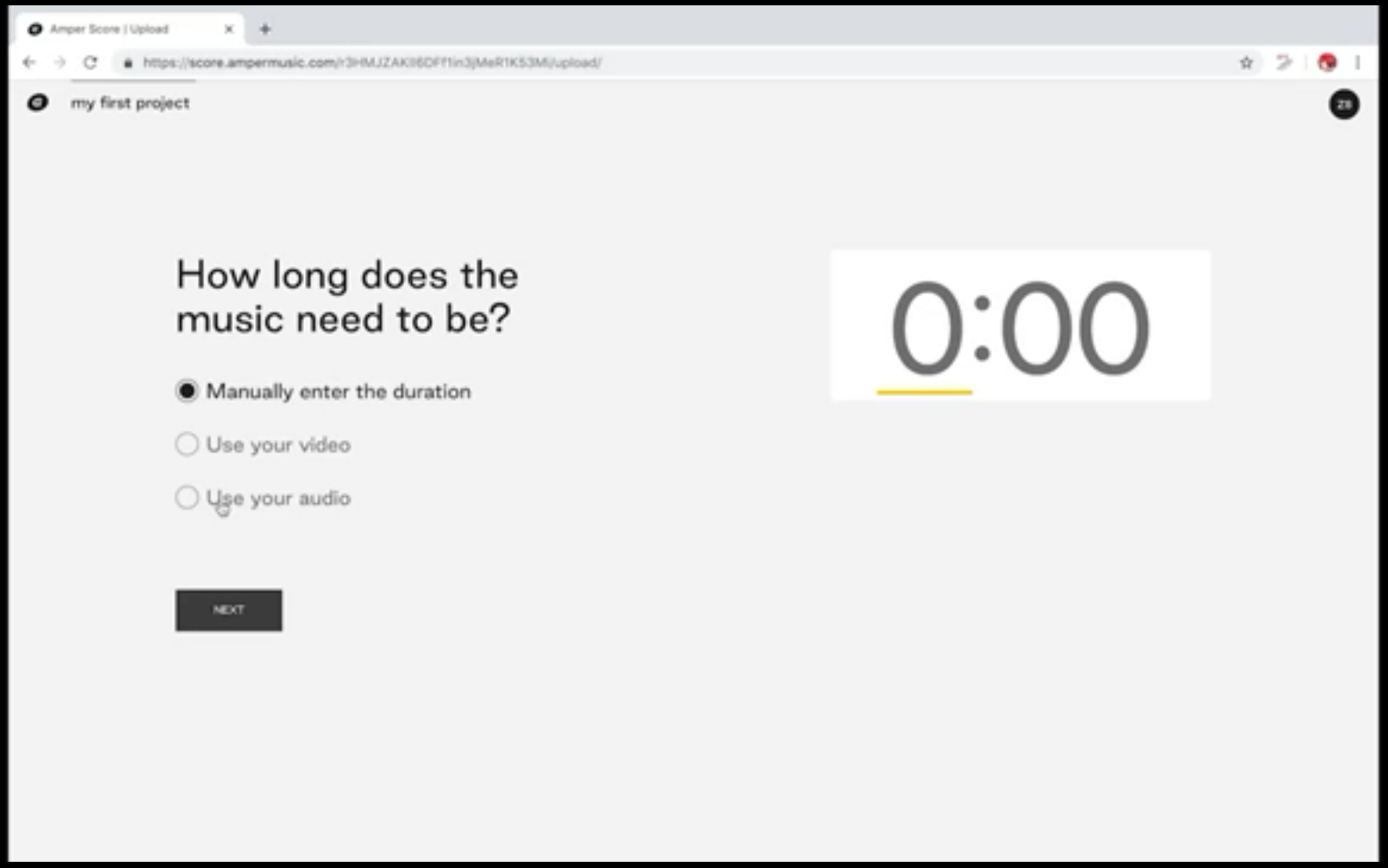
- Year: 2014
- Composition Task: Multi-Track composition for video (soundtrack generation)
- Data Representation: MIDI (unsure, but sounds audio/sample based)
- Algorithm: Proprietary
- Audience: Amateurs, Content creators
- Deployment: Web Application

Amper Score



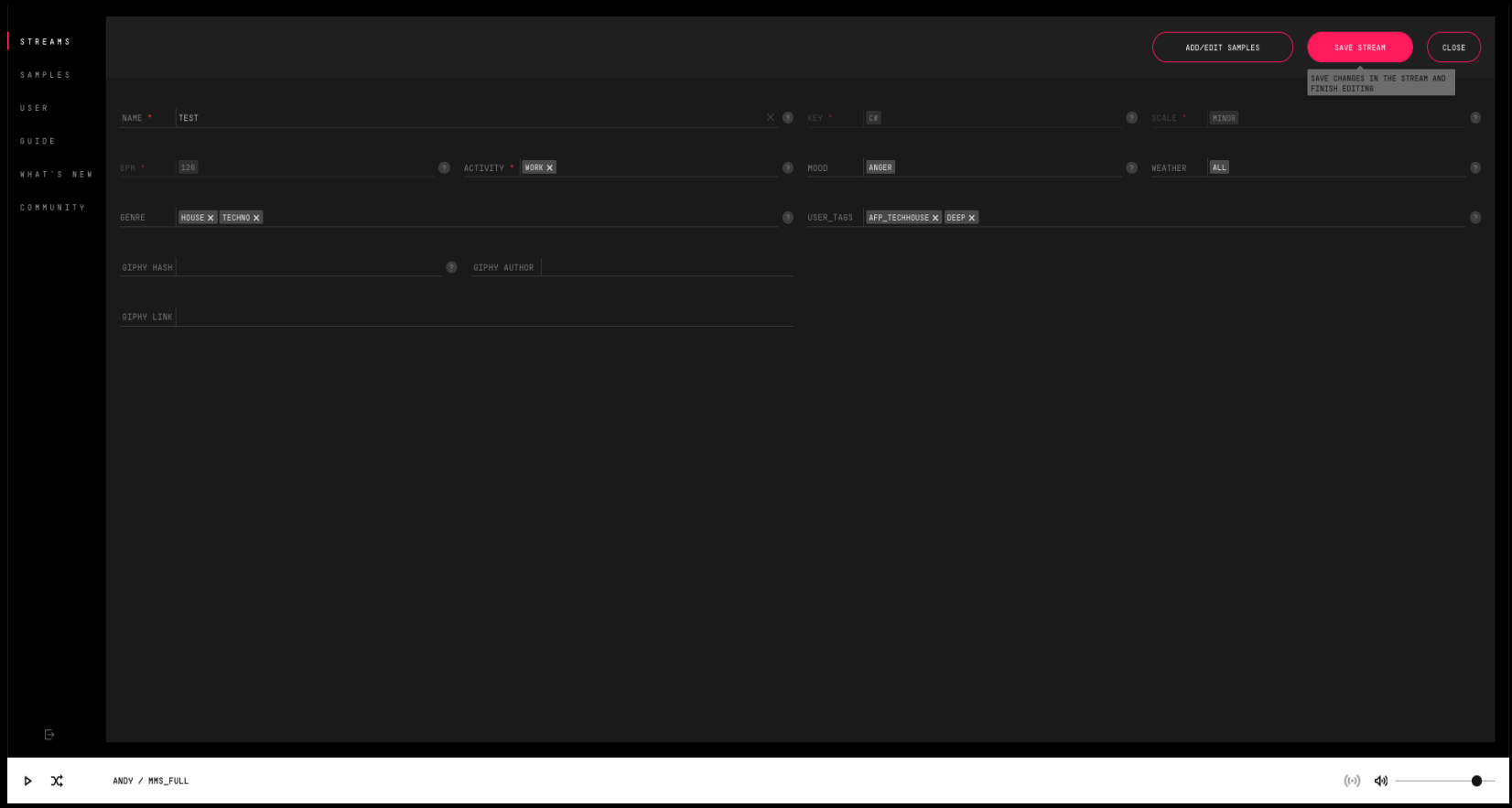
https://www.ampermusic.com/

Amper Score



Introducing Amper Score

Mubert



https://beta.mubert.com/

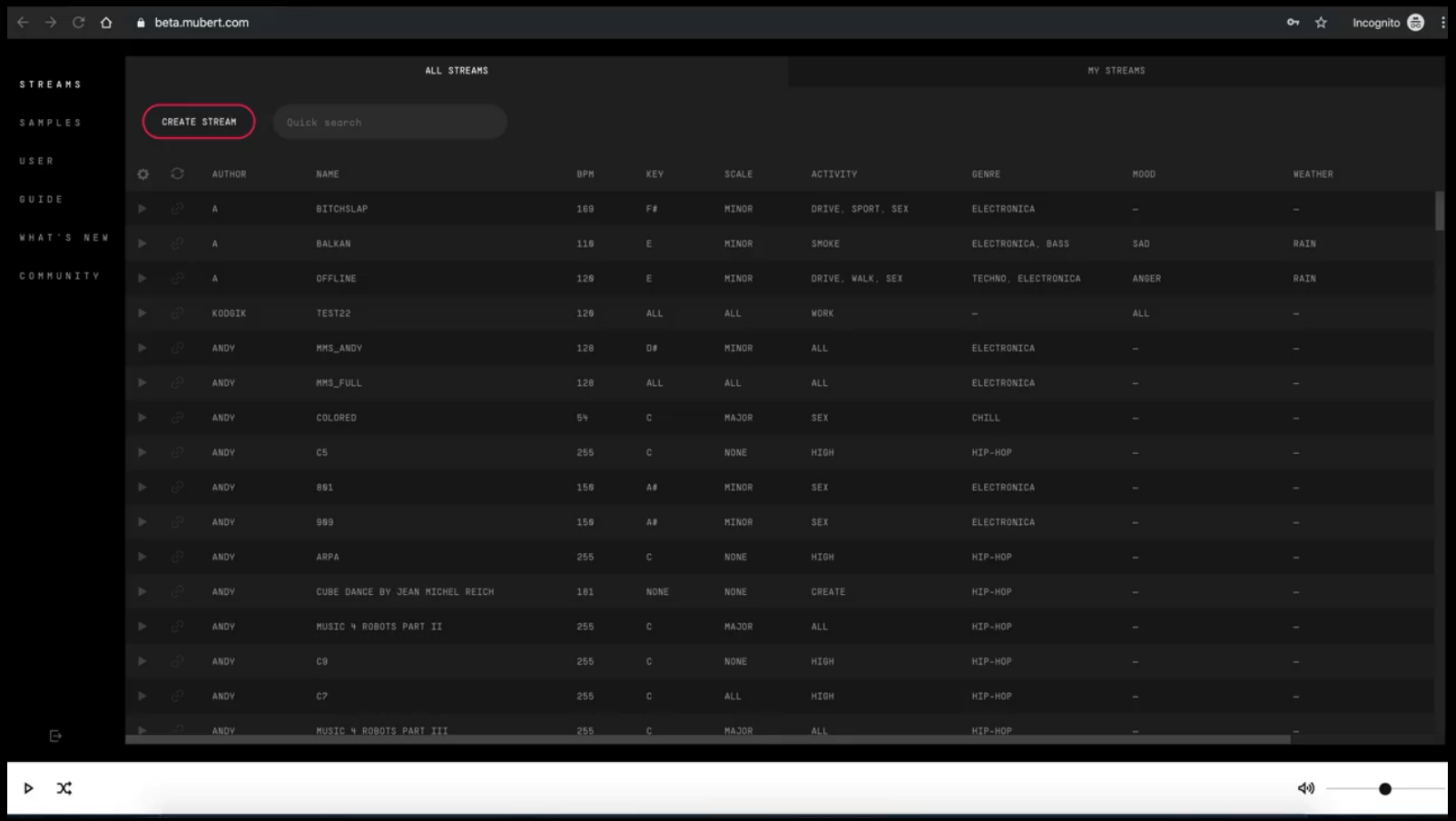
METACREATION
Lab for Creative Artificial Intelligence

SFU

Mubert Pro

- Year: 2016
- Composition Task: Multi-Complete Track (the output is an audio “stream”, a specific generative engine)
- Data Representation: Audio samples (of specific lengths)
- Algorithm: Proprietary
- Audience: Advanced and Enthusiasts
- Deployment: Web Application

Mubert



[How to Create a Generative Music Stream — Mubert Pro]

Mubert

- Andy – High life 
- Dancingteeth – Speed of Thought 
- Ceadfra – Sweetye 
- Ilberezkin – House of mub 

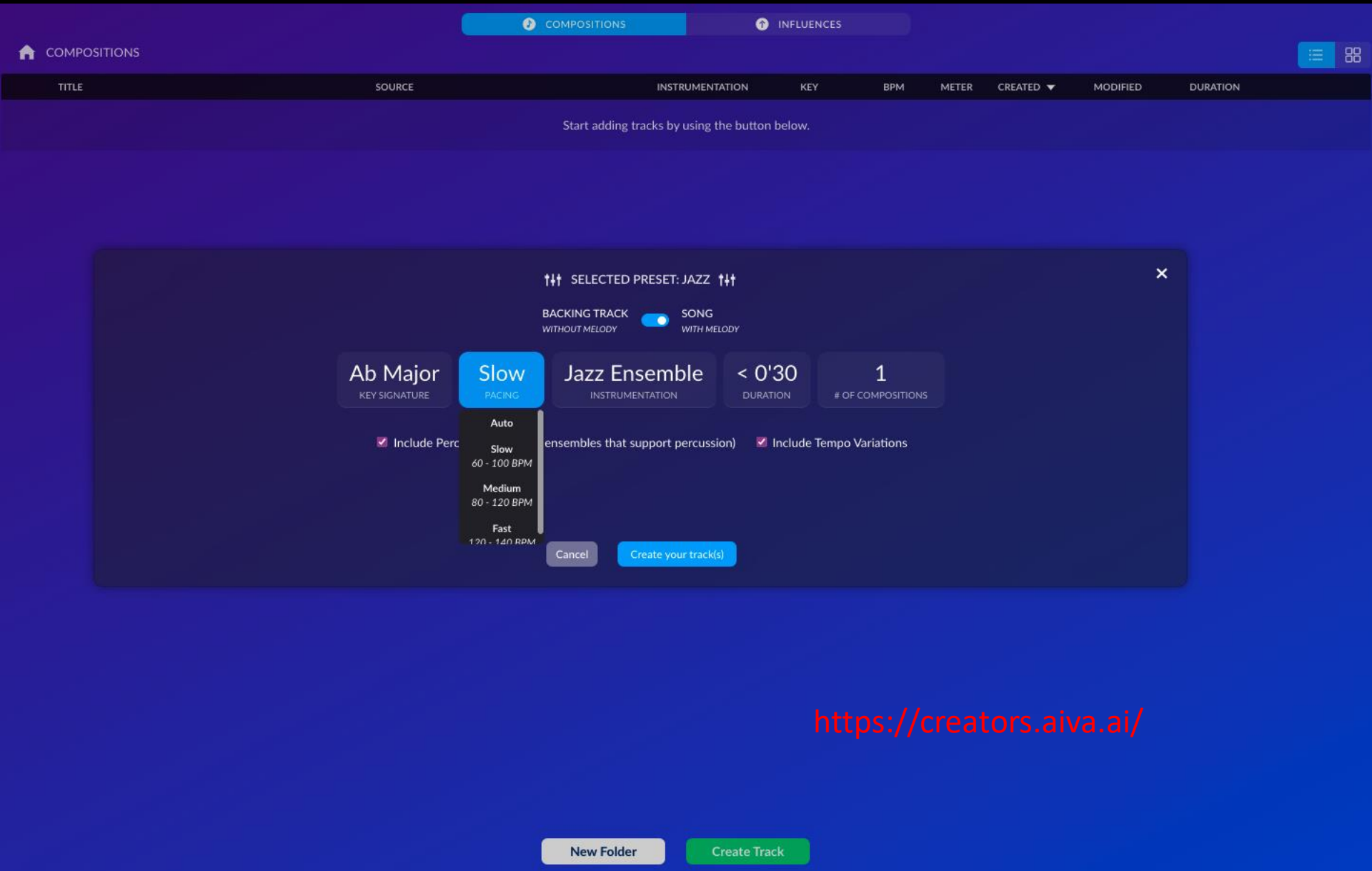
METACREATION
Lab for Creative Artificial Intelligence

SFU

AIVA

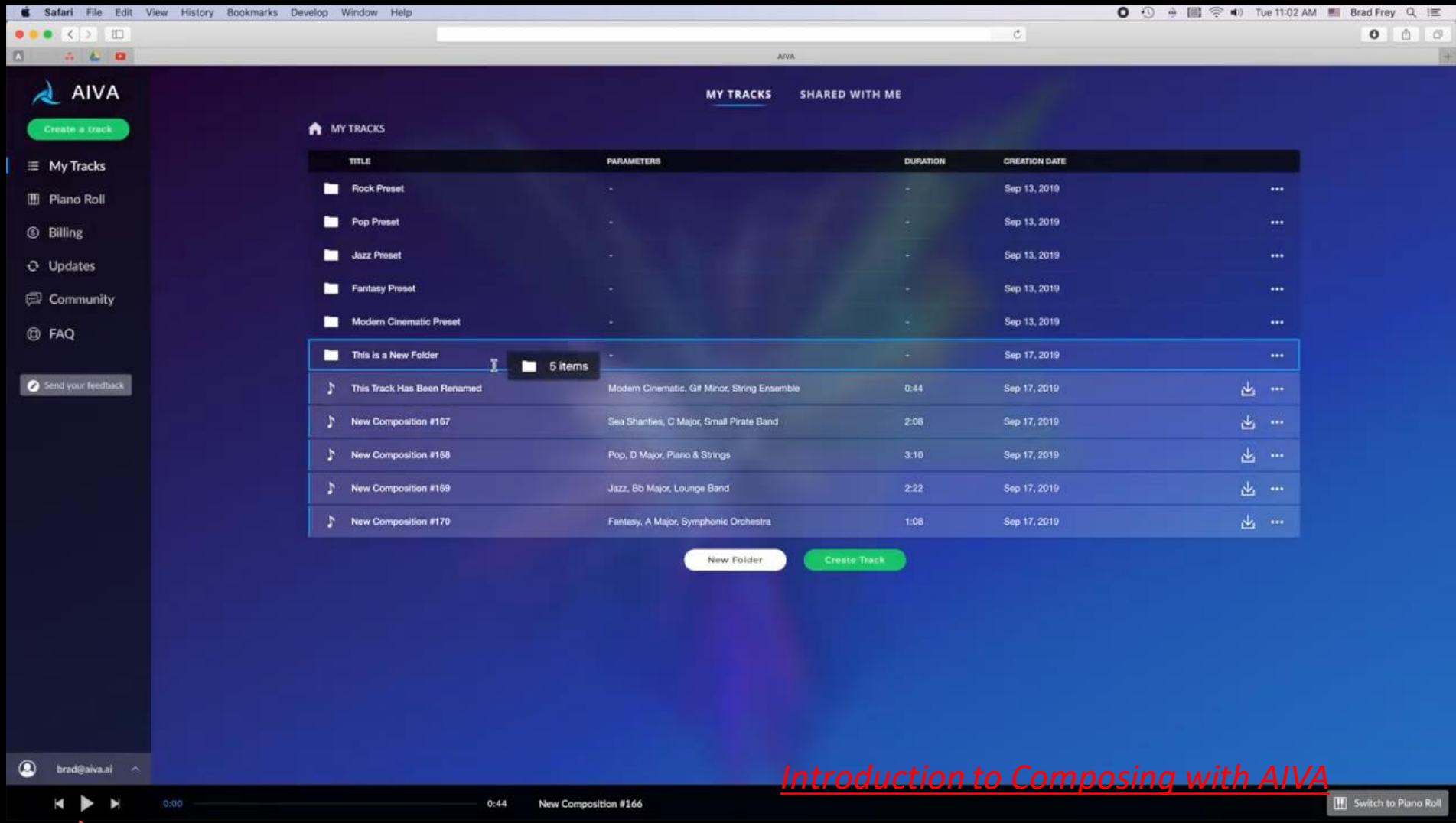
- Year: 2016
- Composition Task: Multi-Complete Track
- Data Representation: Audio/MIDI
- Algorithm: Proprietary
- Audience: Amateurs
- Deployment: Web Application

AIVA



https://creators.aiva.ai/

AIVA



Introduction to Composing with AIVA

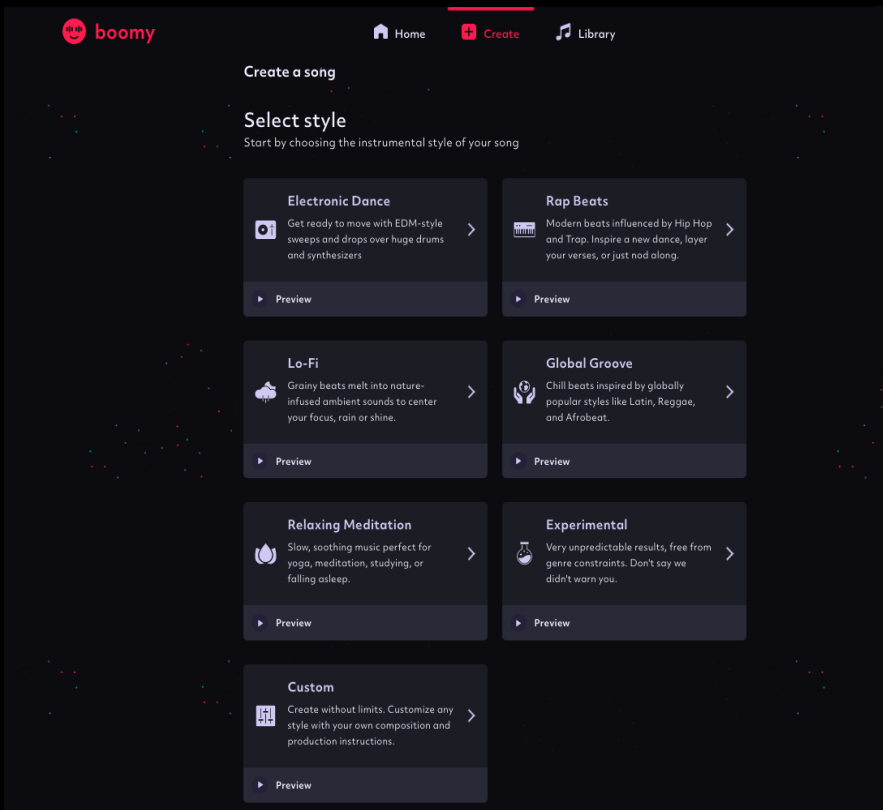
METACREATION
Lab for Creative Artificial Intelligence

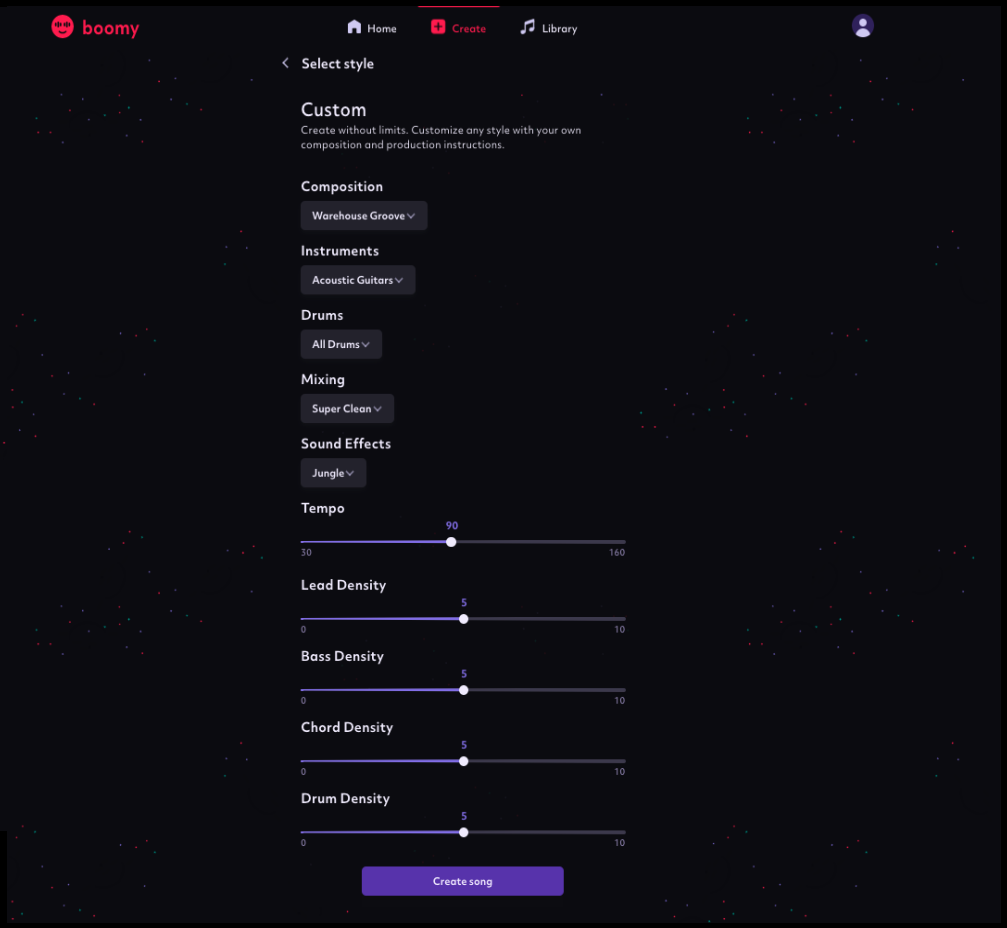
SFU

Boomy

- Year: 2018
- Composition Task: Multi-Track
- Data Representation: Audio
- Algorithm: Proprietary
- Audience: Amateurs and Content Creators
- Interaction: GUI with simple menu selections.
- Deployment: Web Application

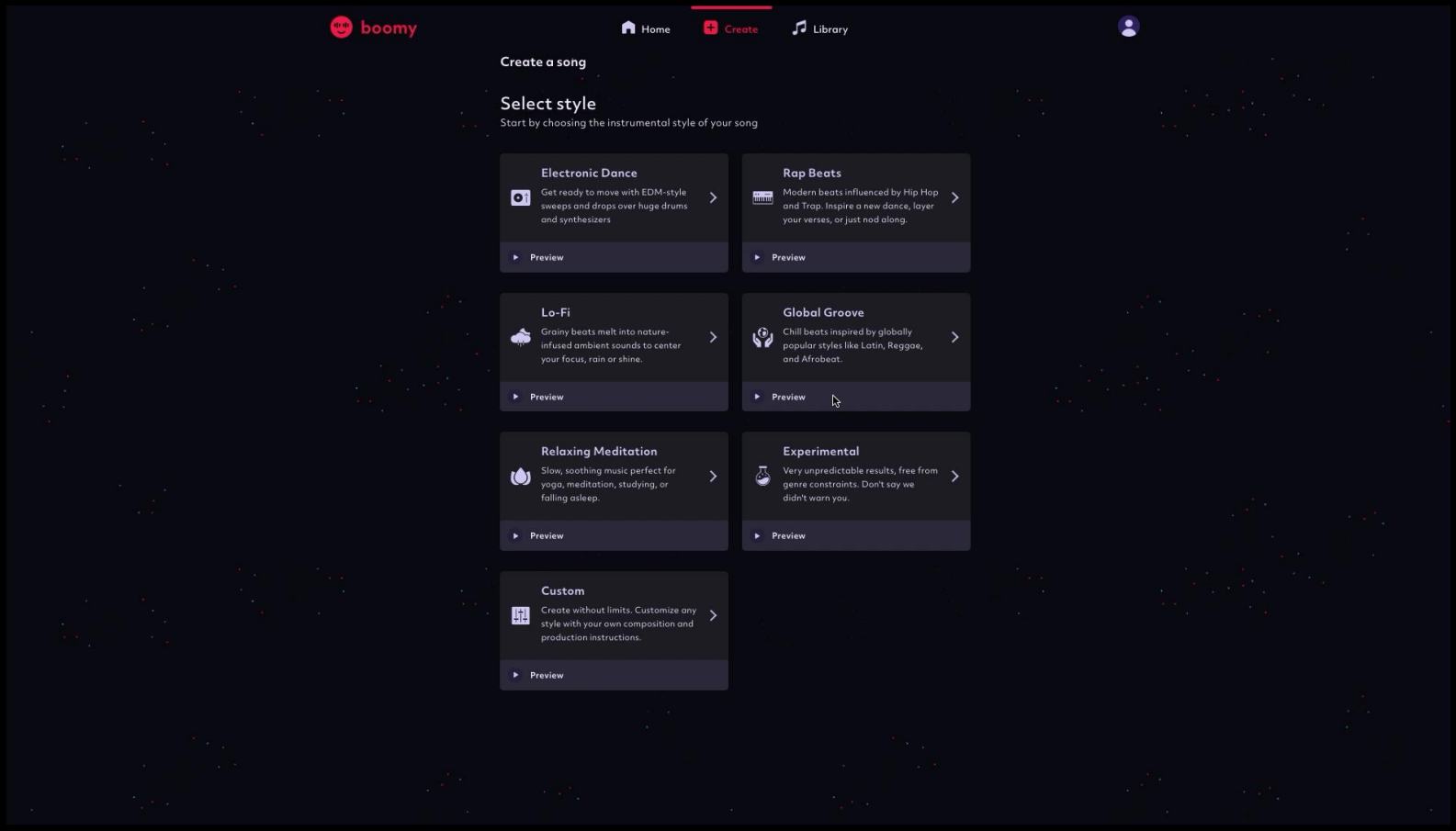
Boomy





https://boomy.com/style

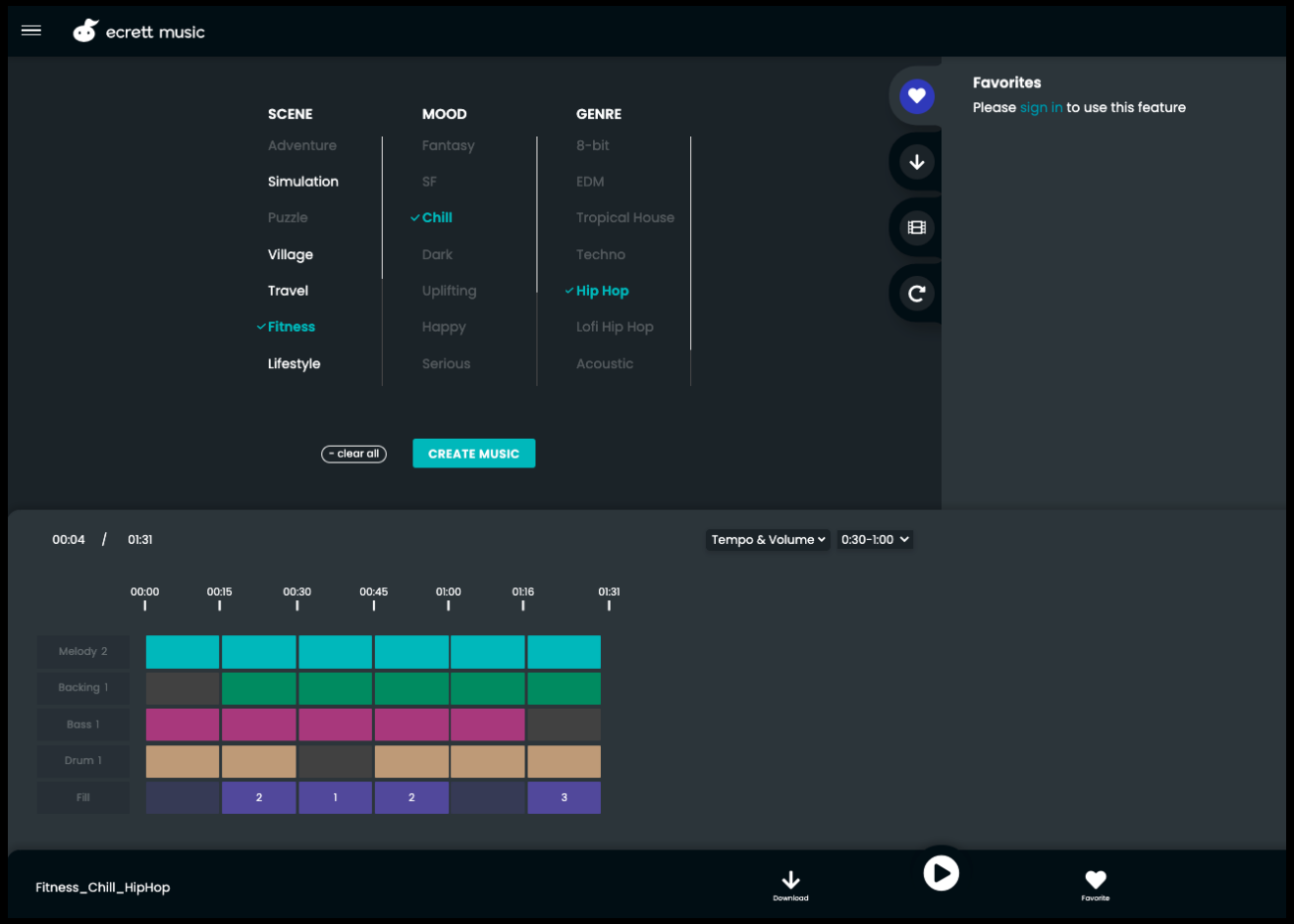
Boomy


boomy
Home
Create
Library
Create a song
Select style
Start by choosing the instrumental style of your song
Electronic Dance
Get ready to move with EDM-style sweeps and drops over huge drums and synthesizers
Preview
Rap Beats
Modern beats influenced by Hip Hop and Trap. Inspire a new dance, layer your verses, or just nod along.
Preview
Lo-Fi
Grainy beats melt into nature-infused ambient sounds to center your focus, rain or shine.
Preview
Global Groove
Chill beats inspired by globally popular styles like Latin, Reggae, and Afrobeat.
Preview
Relaxing Meditation
Slow, soothing music perfect for yoga, meditation, studying, or falling asleep.
Preview
Experimental
Very unpredictable results, free from genre constraints. Don't say we didn't warn you.
Preview
Custom
Create without limits. Customize any style with your own composition and production instructions.
Preview

Ecrett

- Year: 2018
- Composition Task: Multi-Track
- Data Representation: Audio (Bars)
- Algorithm: Proprietary
- Audience: Content Creators
- Deployment: Web Application

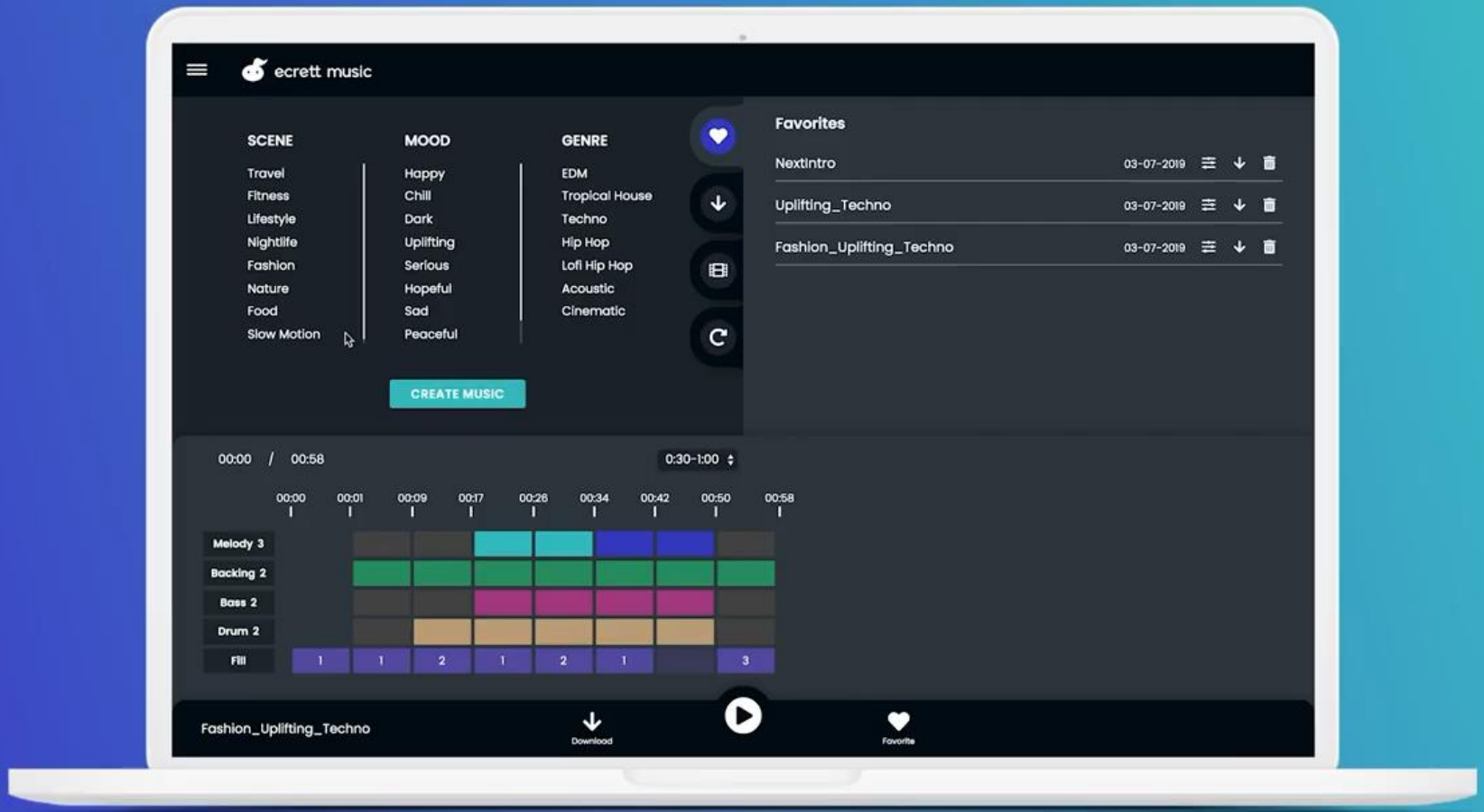
Ecrett



https://ecrettmusic.com/

METACREATION
Lab for Creative Artificial Intelligence

SFU



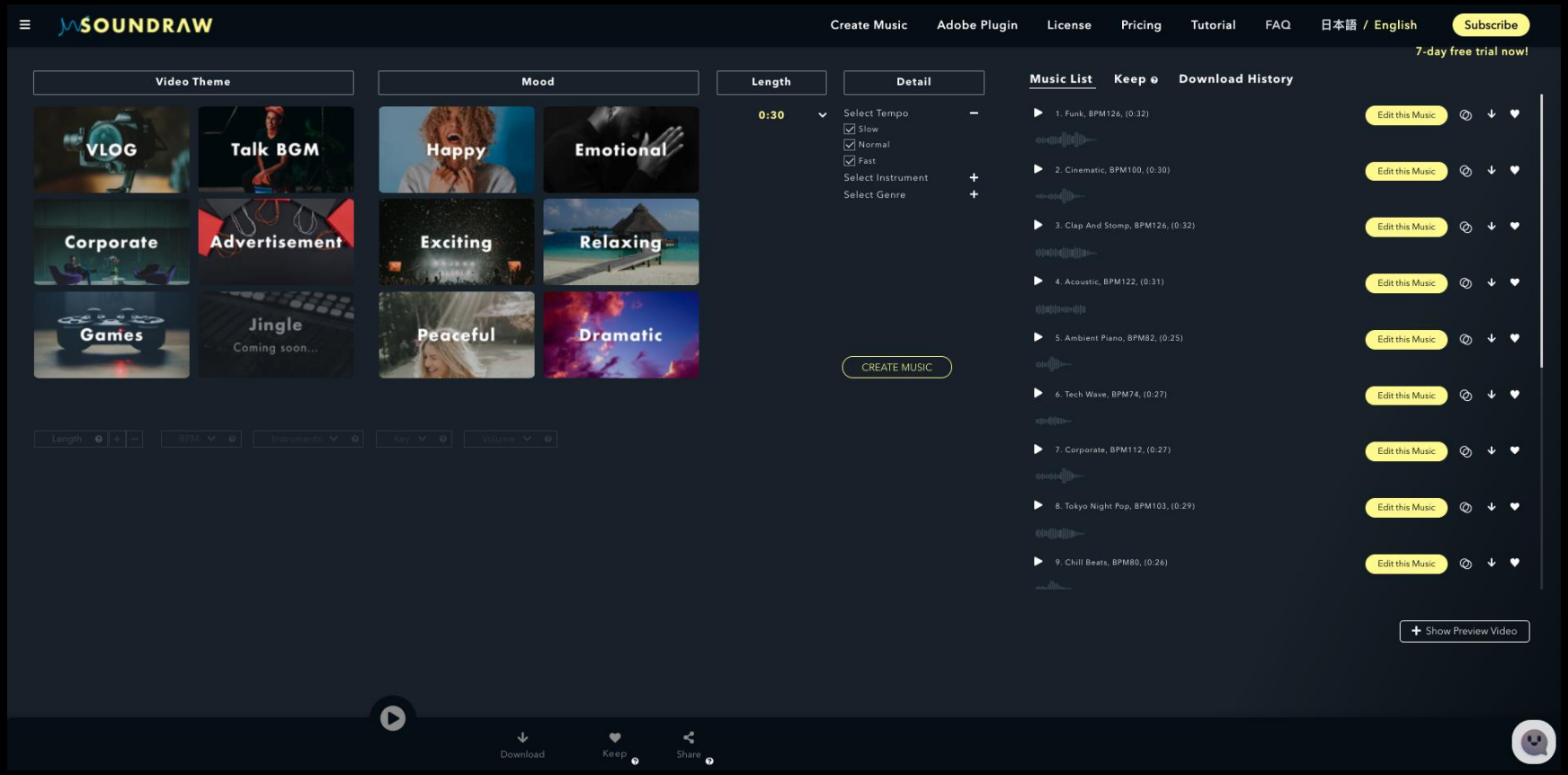
To create music, select at least one from SCENE / MOOD / GENRE.

[How to Create Royalty Free Music for YouTube Video | ecrett music]

METACREATION
Lab for Creative Artificial Intelligence

SFU

Soundraw

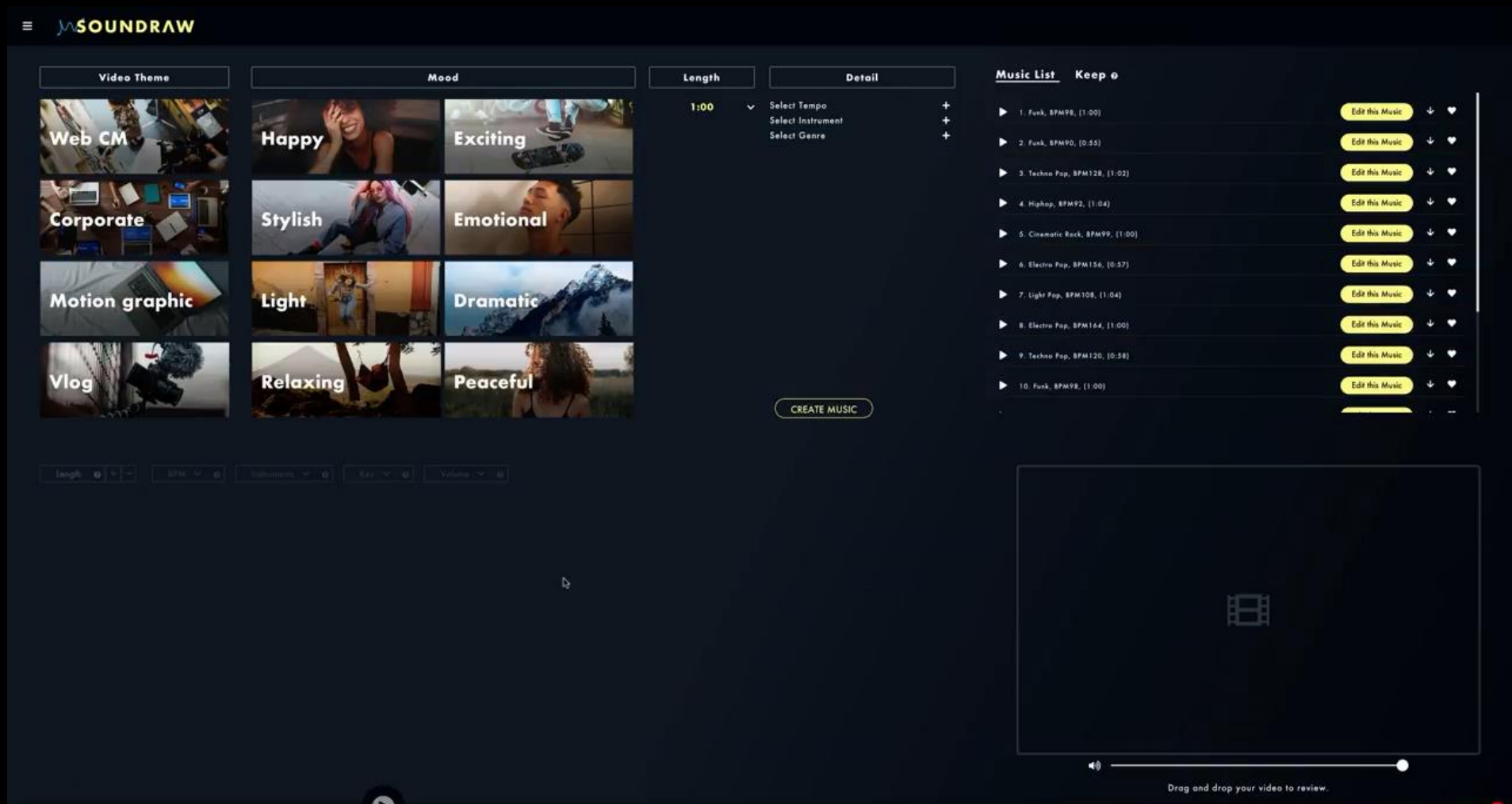


https://soundraw.io/play

Soundraw

- Year: 2020
- Composition Task: Multi-Track
- Data Representation: Audio
- Algorithm: Proprietary
- Audience: Content Creators
- Deployment: Web Application

Soundraw



METACREATION
Lab for Creative Artificial Intelligence

SFU

[Easily Create Unique Royalty-Free Music Using AI [SOUNDRAW.io Review]]

MusicML (2023)

MusicLM: Generating Music From Text

| paper | dataset |

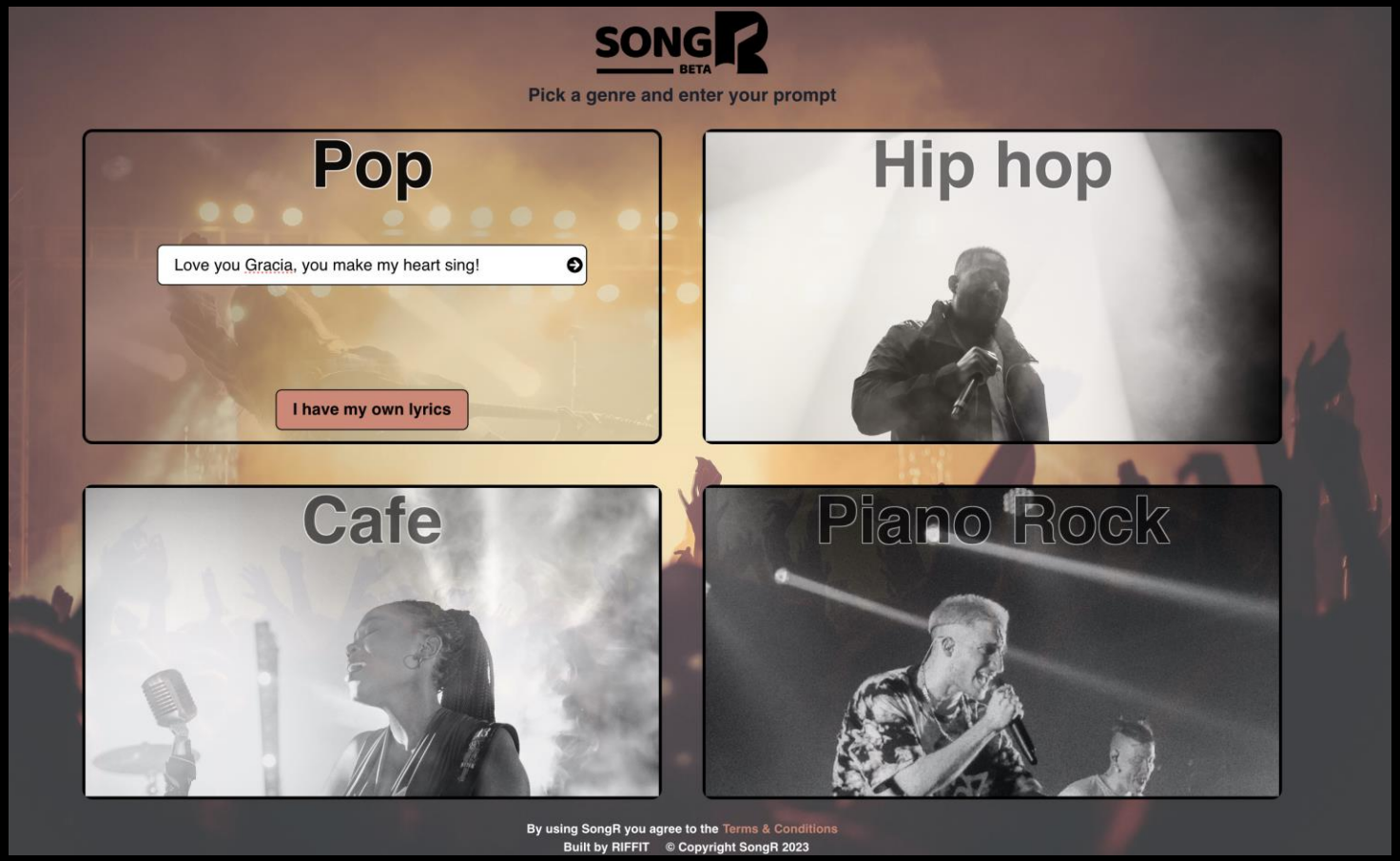
Andrea Agostinelli, Timo I. Denk, Zalán Borsos, Jesse Engel, Mauro Verzetti, Antoine Caillon, Qingqing Huang, Aren Jansen, Adam Roberts, Marco Tagliasacchi, Matt Sharifi, Neil Zeghidour, Christian Frank
Google Research

Abstract We introduce MusicLM, a model generating high-fidelity music from text descriptions such as *"a calming violin melody backed by a distorted guitar riff"*. MusicLM casts the process of conditional music generation as a hierarchical sequence-to-sequence modeling task, and it generates music at 24 kHz that remains consistent over several minutes. Our experiments show that MusicLM outperforms previous systems both in audio quality and adherence to the text description. Moreover, we demonstrate that MusicLM can be conditioned on both text and a melody in that it can transform whistled and hummed melodies according to the style described in a text caption. To support future research, we publicly release MusicCaps, a dataset composed of 5.5k music-text pairs, with rich text descriptions provided by human experts.

Audio Generation From Rich Captions

Caption	Generated audio
The main soundtrack of an arcade game. It is fast-paced and upbeat, with a catchy electric guitar riff. The music is repetitive and easy to remember, but with unexpected sounds, like cymbal crashes or drum rolls.	▶ 0:00 / 0:00

SongR (2013)



METACREATION
Lab for Creative Artificial Intelligence

SFU

SONGR BETA

Pick a genre and enter your prompt

Pop

Hip hop

Love you Gracia

Edit Your Pop Lyrics!

Prompt: "Love you Gracia, you make my heart sing!"

Love you Gracia, you make my heart sing!
You brighten up my day,
With your smile and your laughter,
You make everything feel okay.
I love spending time with you,
Whenever we're together,
I feel so happy and free,
I know that this is true love.

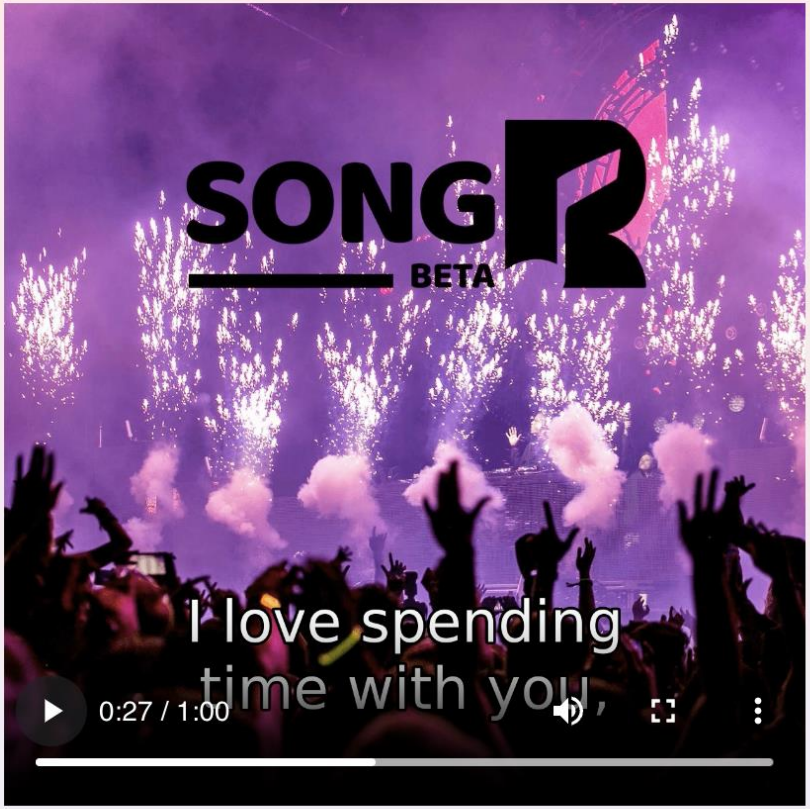
279 characters remaining

Natalie Kevin

Clear Try Again Render

By clicking "Render" you agree to abide by our Terms & Conditions

SONG BETA

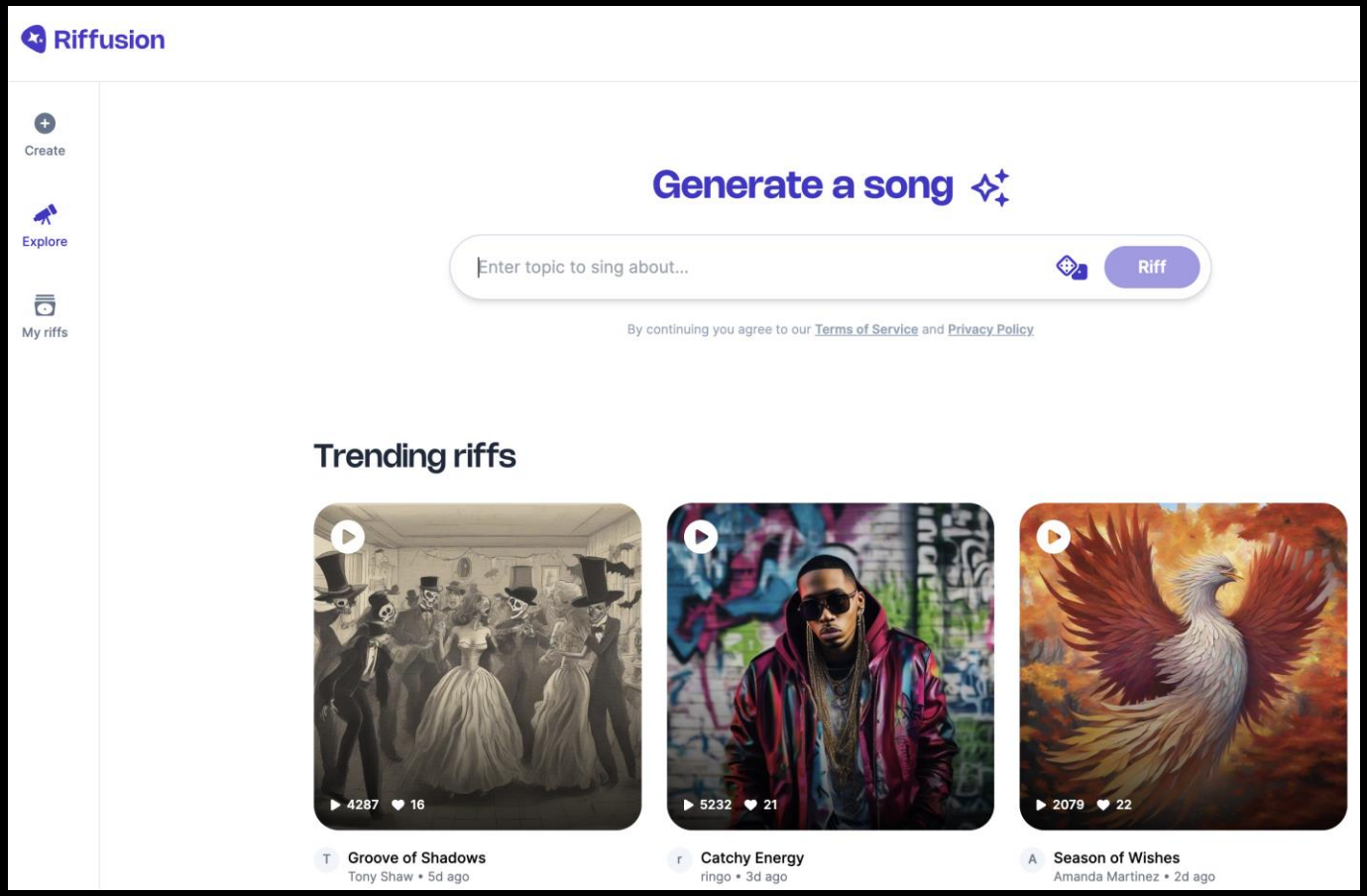


Restart

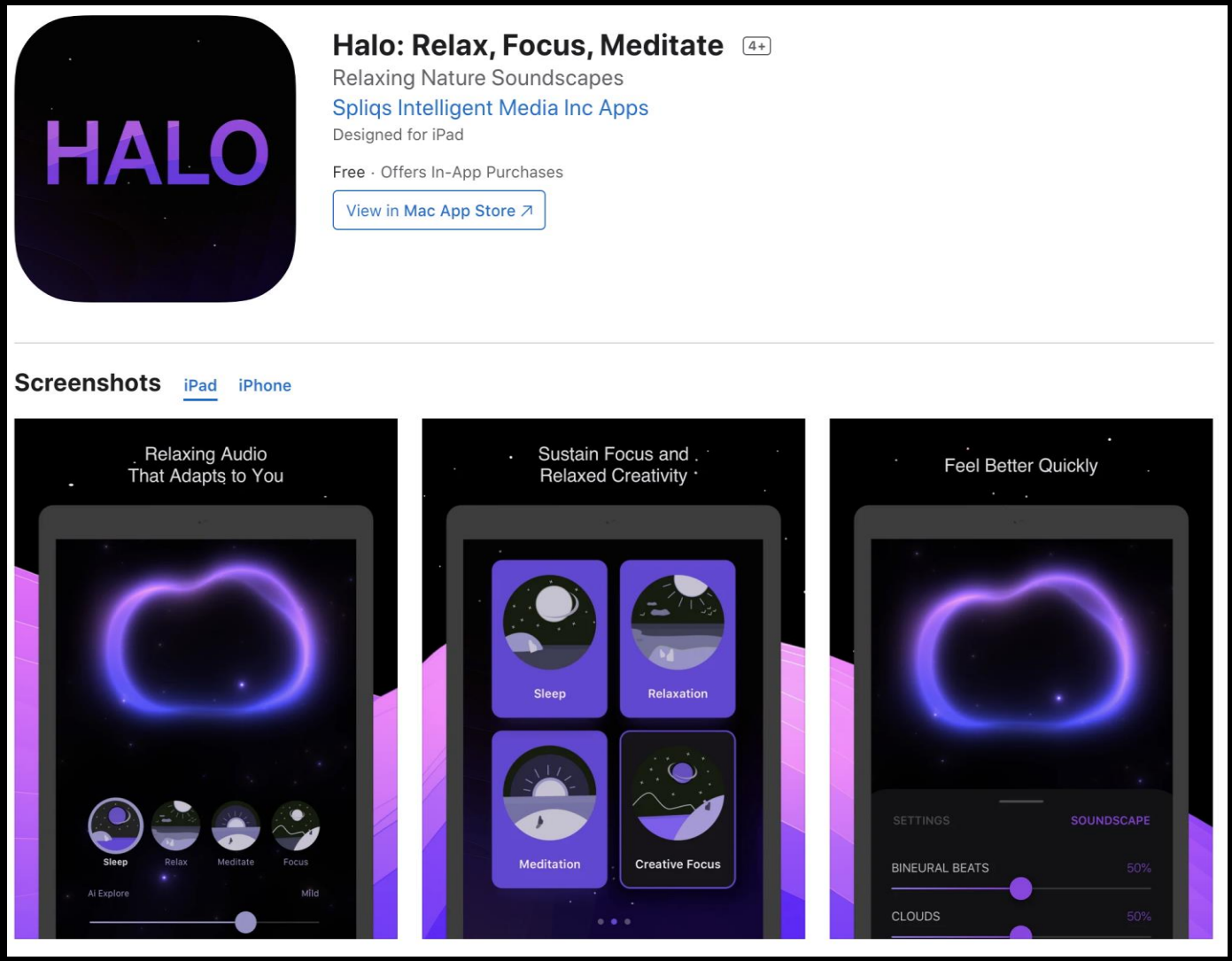
Download

Share Your Creation!

Riffusion (2023)



Functional music generation.



By Lab alumni Nicolas Gonzales and Dr. James Maxwell (2018-2023)

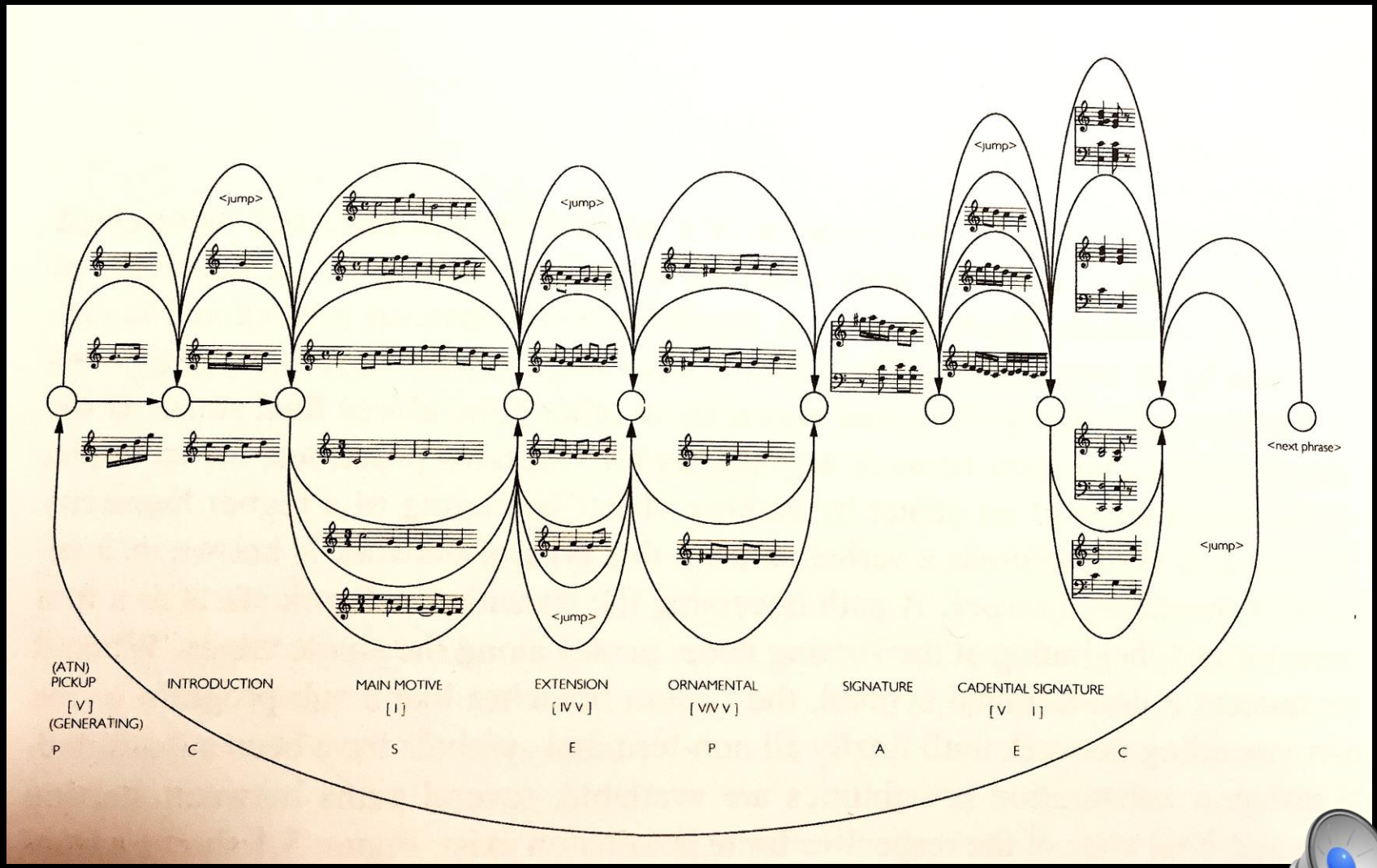
Computer-assisted composition
MIDI/Symbolic generation

CAC Systems examples

- **Audio output:**
  - AUME (2012)
  - Amper Score (2014) – video scoring
  - Mubert Pro (2016) – generative music stream
  - AIVA (2016) *
  - Boomy (2018)
  - Ecrett (2018)
  - Soundraw (2020)
  - MusicML (2023)
  - SongR (2023)
  - Riffusion (2023)
  - Halo (2018-2023)

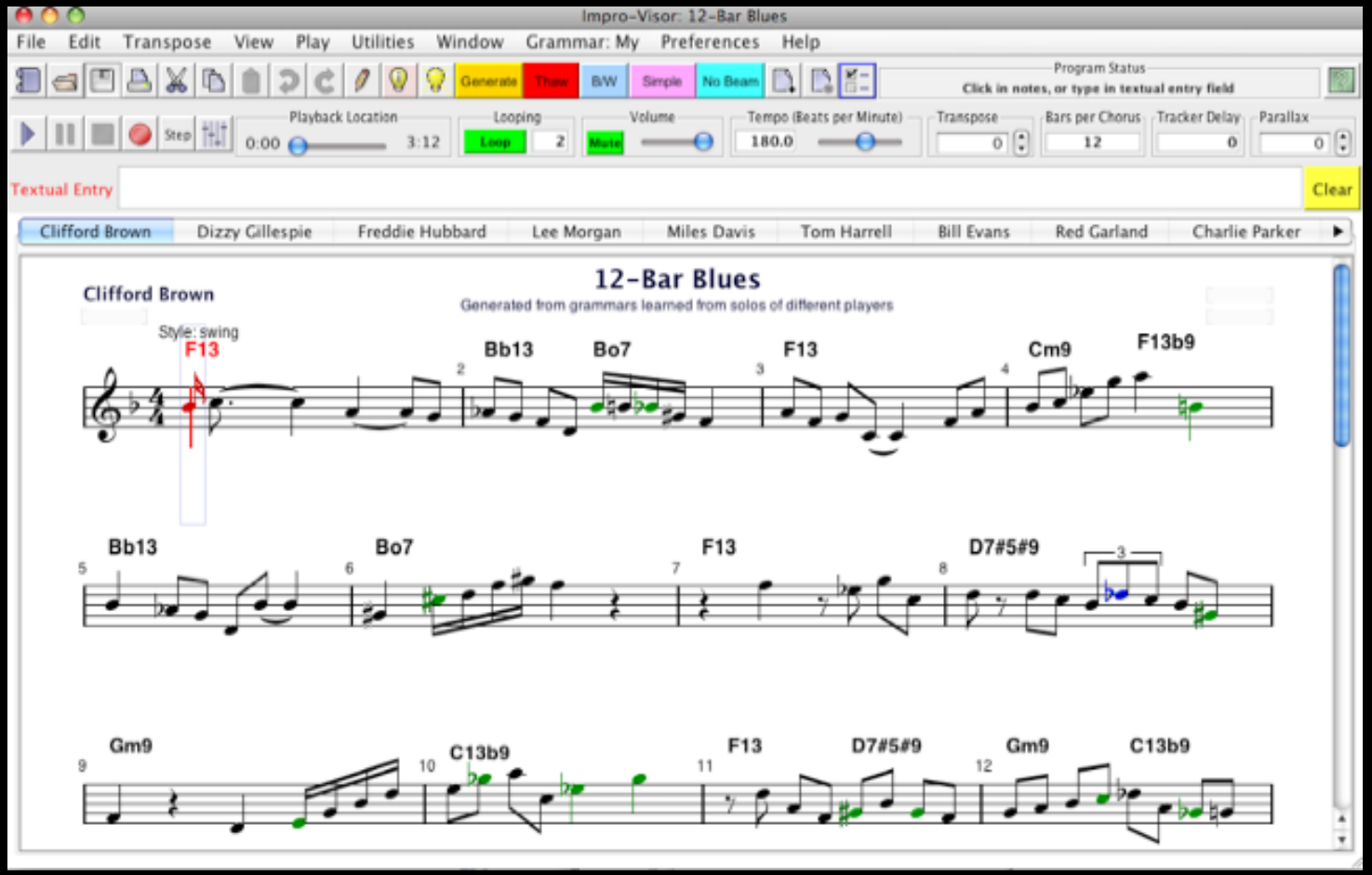
- **MIDI output:**
  - EMI (1995-)
  - Imporvisor (2005-2020, Pr. Bob Keller,
  - Harmonic Progression Generator (2010)
  - MusiCog, Manuscore (2012)
  - Style Machine Lite (2012-2016)
  - FlowMachines (2013-2021)
  - FolkRNN (2018)
  - MuseNet (2019)
  - DrumVAE (2019)
  - Magenta Studio (2019)
  - Orchidea (2020, orchestration)
  - Draw and Listen (2022)
  - MMM4Live (2023)
  - Calliope (2023)
  - MMM-C (2023)

Augmented Transition Networks (ATN)



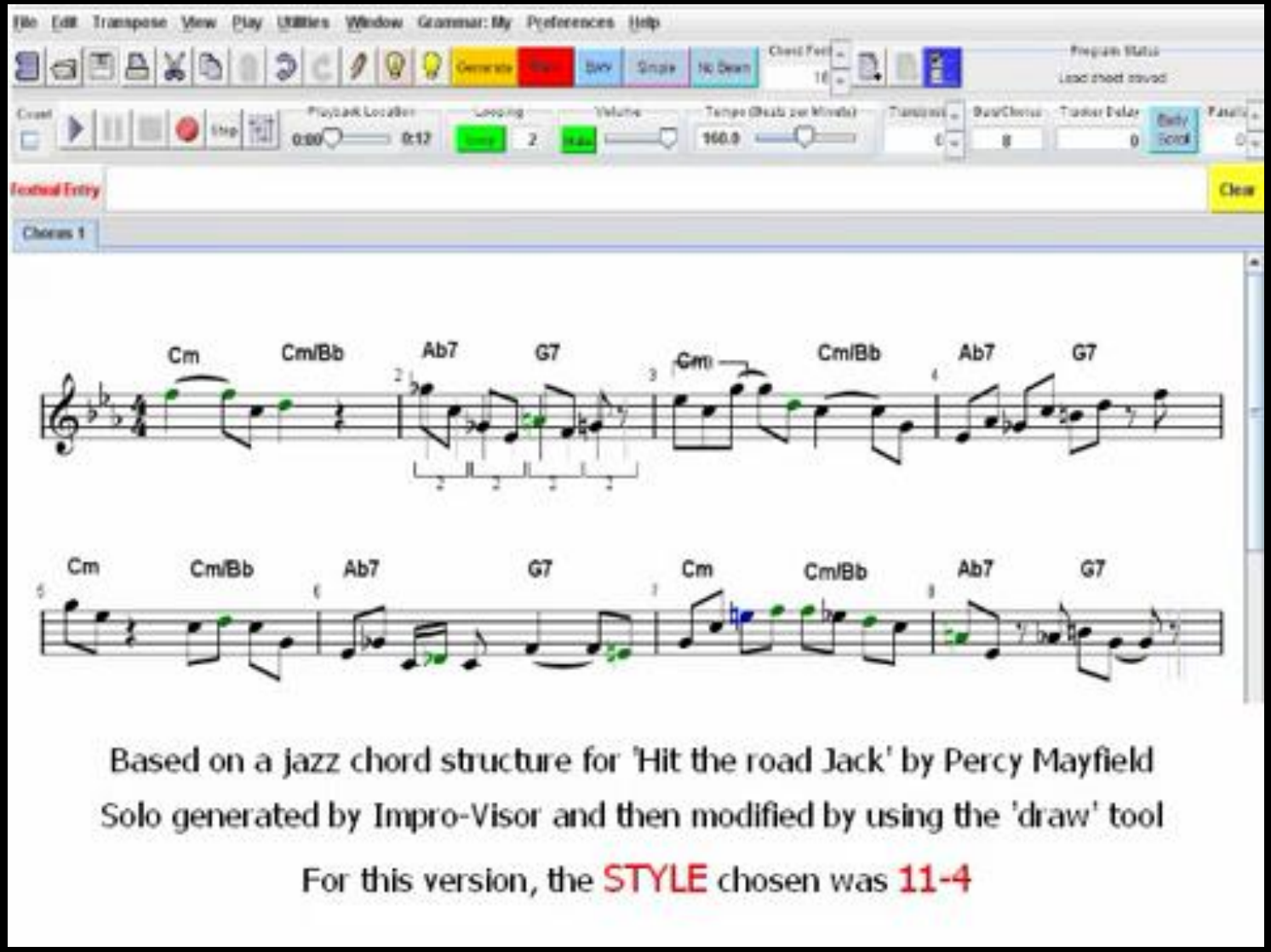
David Cope EMI (Experiments in Musical Intelligence) is a system that does style imitation using a recombinant approach based on ATN (1996-).

Impro-Visor



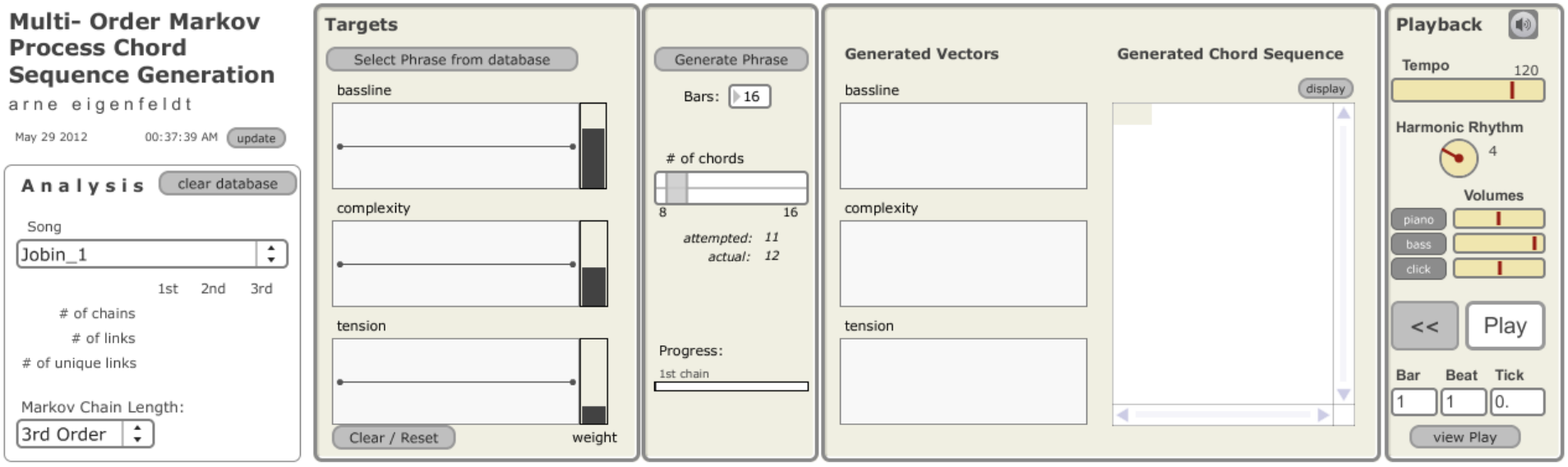
Bob Keller et al. (2005-2020)

Impro-Visor


File Edit Transpose View Play Utilities Window Grammar: My Preferences Help
Festival Entry
Clear
Chorus 1
Cm Cm/Bb Ab7 G7 Cm Cm/Bb Ab7 G7
Cm Cm/Bb Ab7 G7 Cm Cm/Bb Ab7 G7
Based on a jazz chord structure for 'Hit the road Jack' by Percy Mayfield
Solo generated by Impro-Visor and then modified by using the 'draw' tool
For this version, the STYLE chosen was 11-4

With Arne Eigenfeldt
ICCC-2010

Harmonic Progression Generator

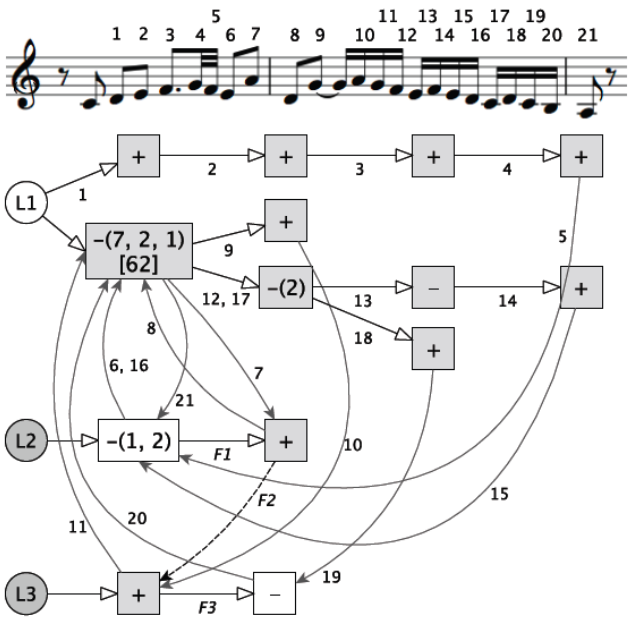


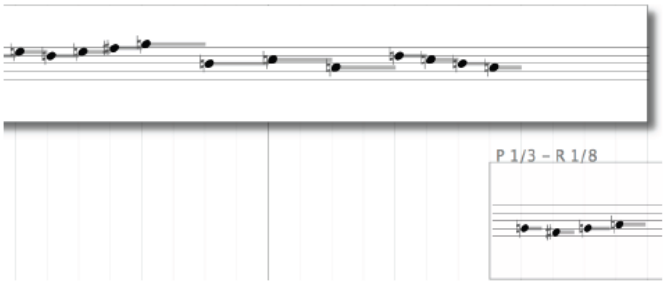
- It is doing style imitation, at human-competitive levels.
- The system is available online (open source and free!).
- It has been used by composers and the company Teenage Engineering (Sweden) when producing the Absolut Blank iPhone app.

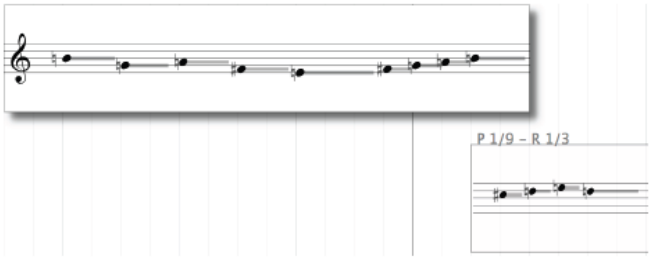
Musical Metacreation

With James Maxwell and Arne Eigenfeldt
Sound and Music Computing 2011
Int. Computer Music Conference, 2012

- Closure-based Cueing Model (CbCM)
- Challenge: learning and generating music (symbolic)
- Solution:
  - Another attempt at a hierarchical, learning, model of musical cognition
  - Based on notions from the musical perception and cognition literature
- Validation: it actually works!
  - Applied in the ManuScore computer-aided composition software
  - Used for actual compositions (instrumental contemporary music): presented in concerts.
  - Empirical evaluation with 42 participants: could not segregate

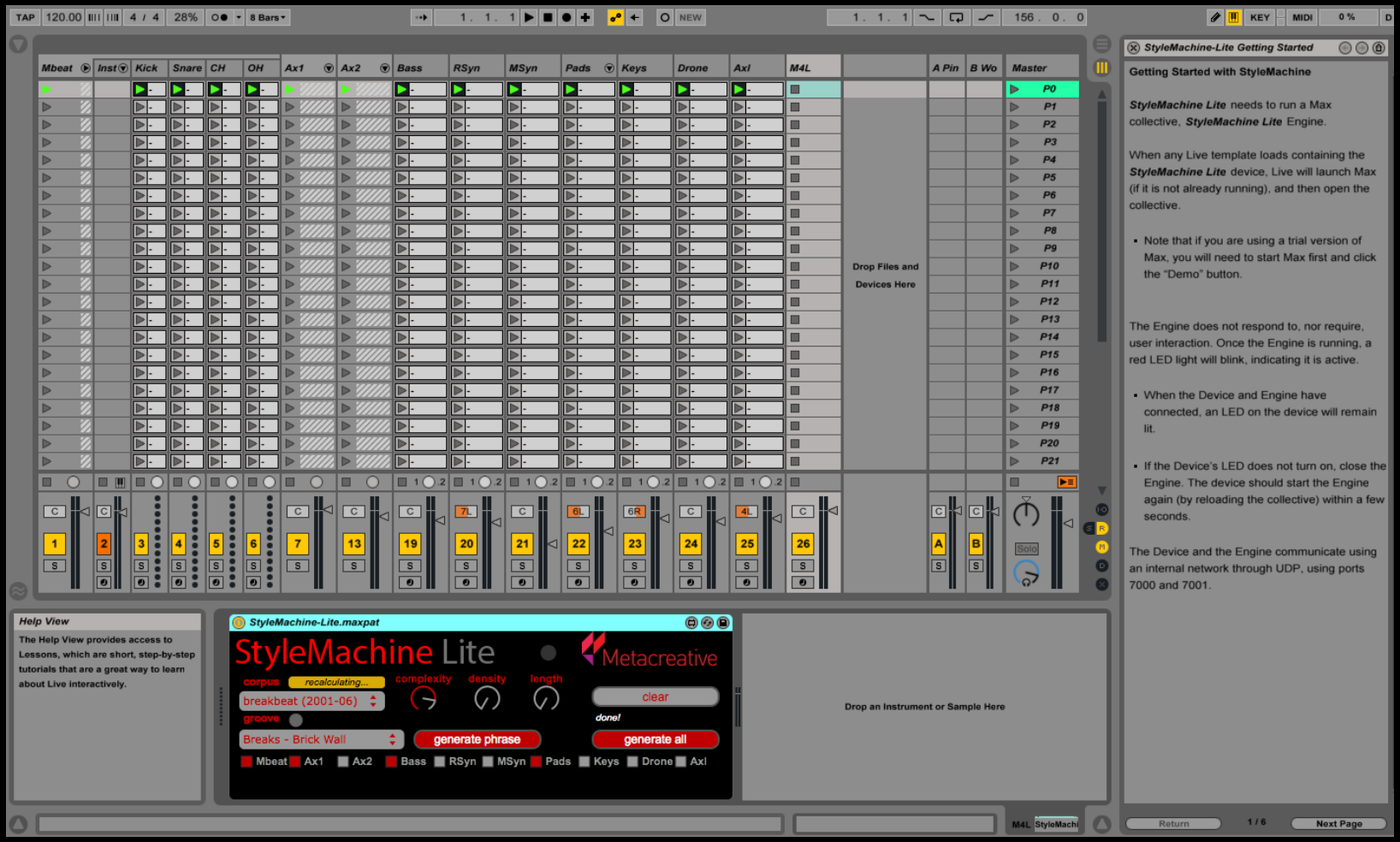






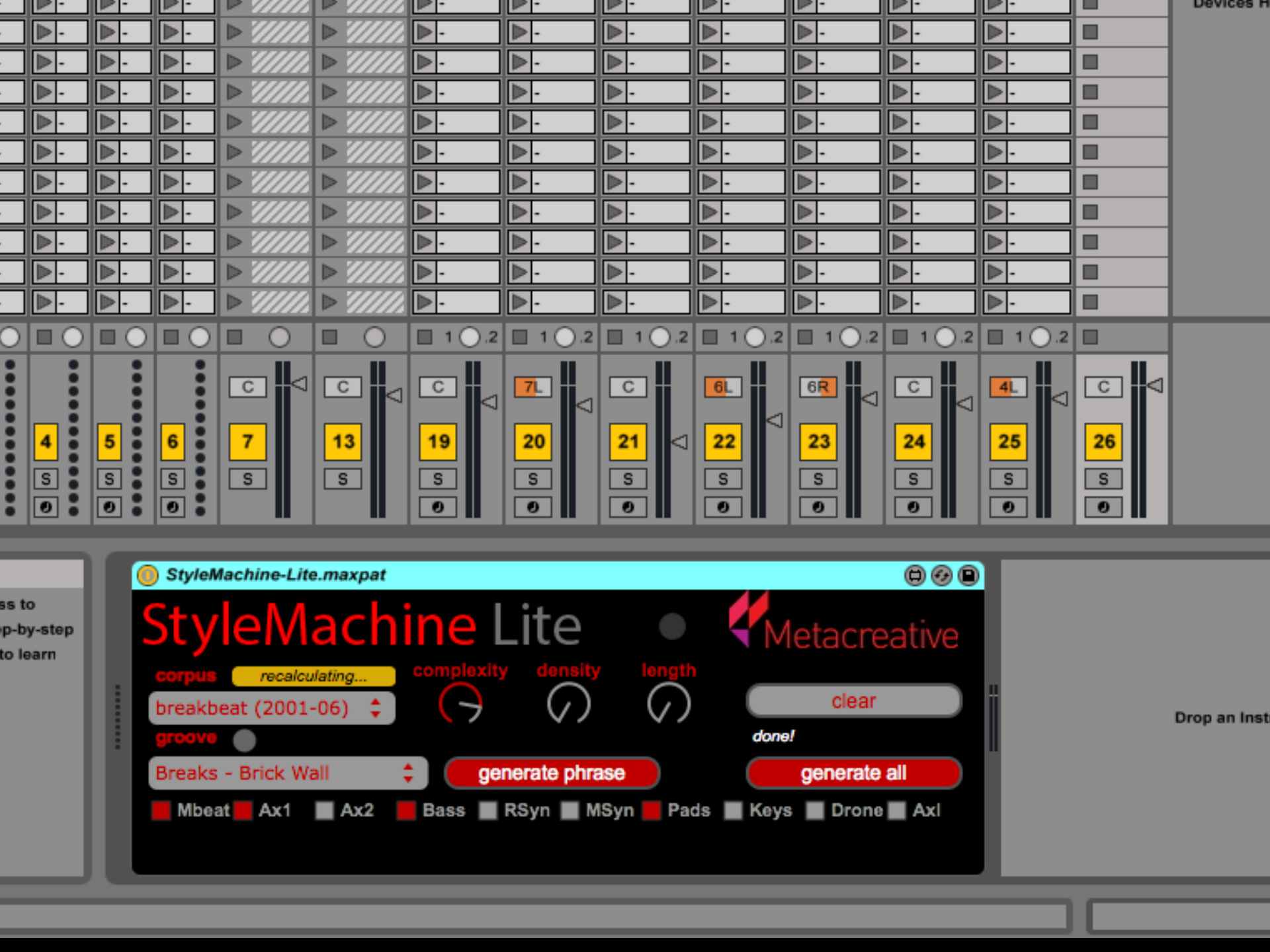
MusiCog

Experiri. MusiCog used by James B. Maxwell for computer-assisted composition in the Manuscore environment. Yaletown string quartet, 2011.



The StyleMachine Lite, by Metacreative Technologies, does corpus-based style imitation of Electronic Dance Music (EDM) since 2014.

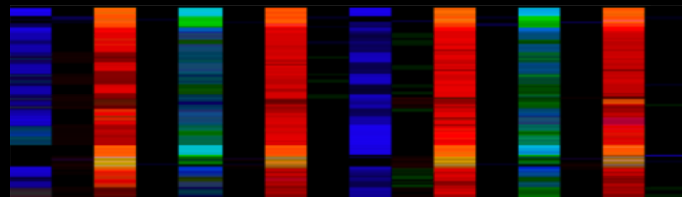
METACREATION
Lab for Creative Artificial Intelligence
SFU

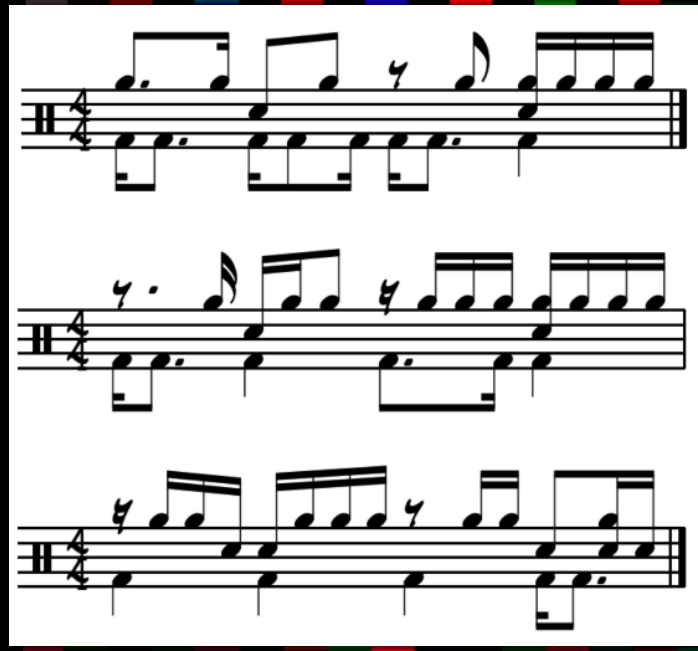

StyleMachine-Lite.maxpat
StyleMachine Lite
Metacreative
corpus
recalculating...
breakbeat (2001-06)
groove
Breaks - Brick Wall
complexity
density
length
clear
done!
generate phrase
generate all
Mbeat
Ax1
Ax2
Bass
RSyn
MSyn
Pads
Keys
Drone
Axl
Drop an Inst

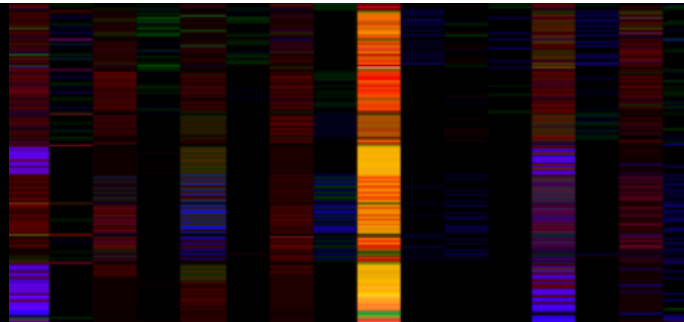
Style Machine

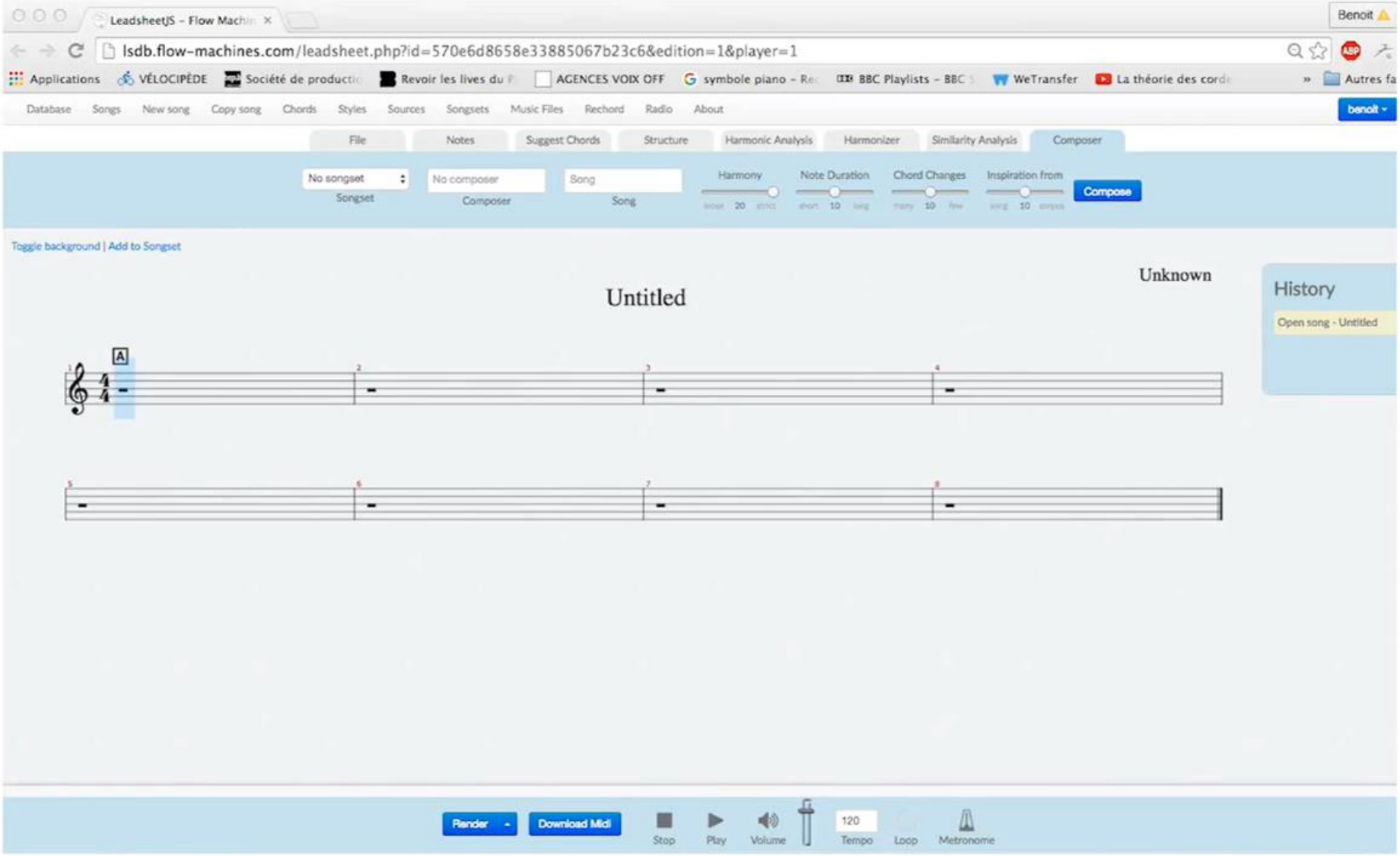
With Arne Eigenfeldt, Christopher Anderson
Sound and Music Computing 2011
Computation Creativity, 2013
GECCO, 2013

- Generative EDM
- How?
  - Manual analysis of corpus by experts (composers, producers)
  - Our machine learning algorithm Genetic algorithm / VOMM
- Validation: ongoing!
  - Confuses classifiers: pieces gets classified properly!
  - Confuses listeners
  - Public shows: Algorave ISEA, …
  - Album on ChordPunch (UK)





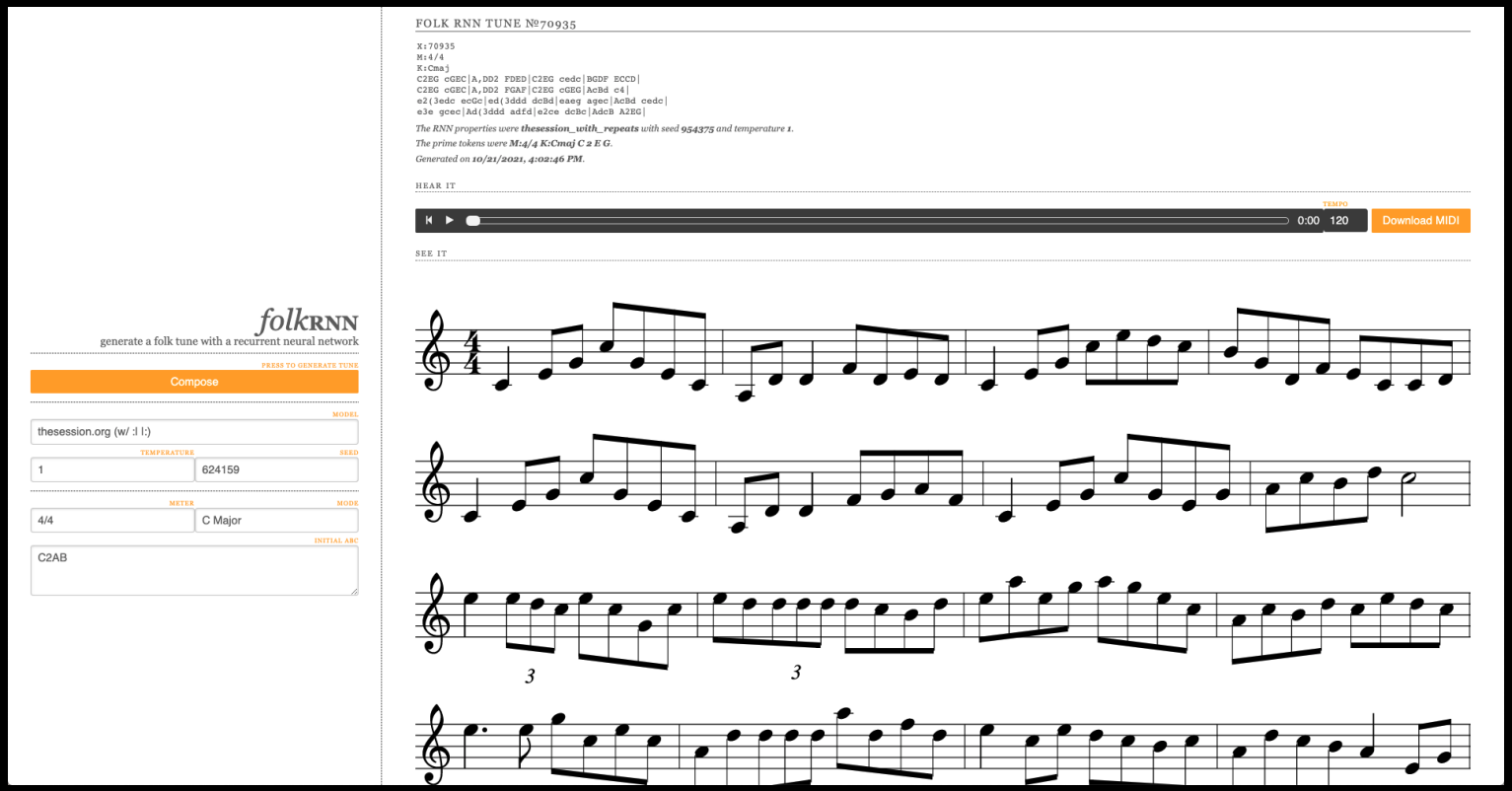




Flow Machines (Sony CSL, Francois Pachet), 2016.

SKYGGE - Magic Man [Official Music Video]
composed with Flow Machine in 2017/2018 (music composition + lyrics).

FolkRNN



https://folkrnn.org/

FolkRNN

- Year: 2018
- Composition Task: Melody Generation
- Data Representation: ABC Notation
- Algorithm: Recurrent Neural Networks
- Audience: Amateurs and Professionals
- Deployment: Web Application

Sturm BL, Santos JF, Ben-Tal O, Korshunova I. Music transcription modelling and composition using deep learning. arXiv preprint arXiv:1604.08723. 2016 Apr 29.

FolkRNN

- Bob Sturm + folk-rnn v2 (beamsearch n=2) - Week 5: Mickey Fitternaly's (2020)

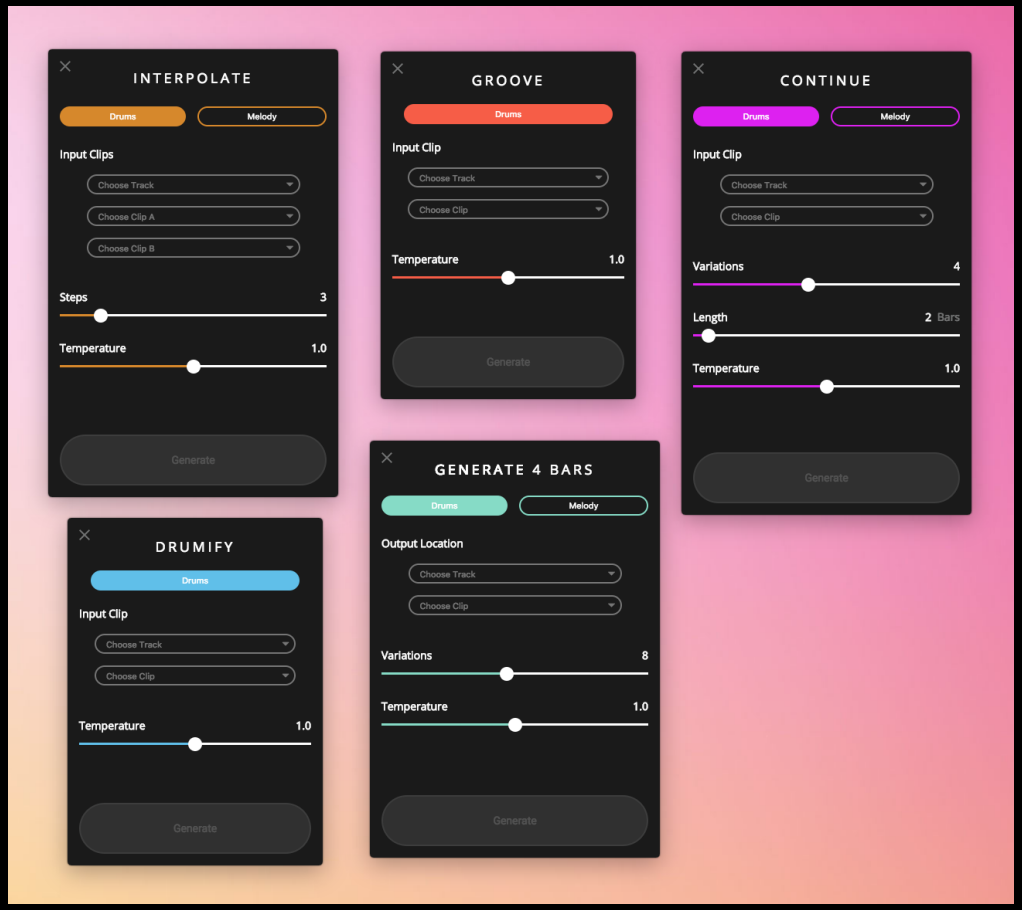
- Bastard Tunes - 2nd movement (2017)

- Bob L. Sturm + folk-rnn - March to the Mainframe (performed by Esemble x.y) (2017)

FolkRNN

folkrnn.org [2018]

Magenta Studio



https://magenta.tensorflow.org/studio/

METACREATION
Lab for Creative Artificial Intelligence

SFU

Magenta Studio

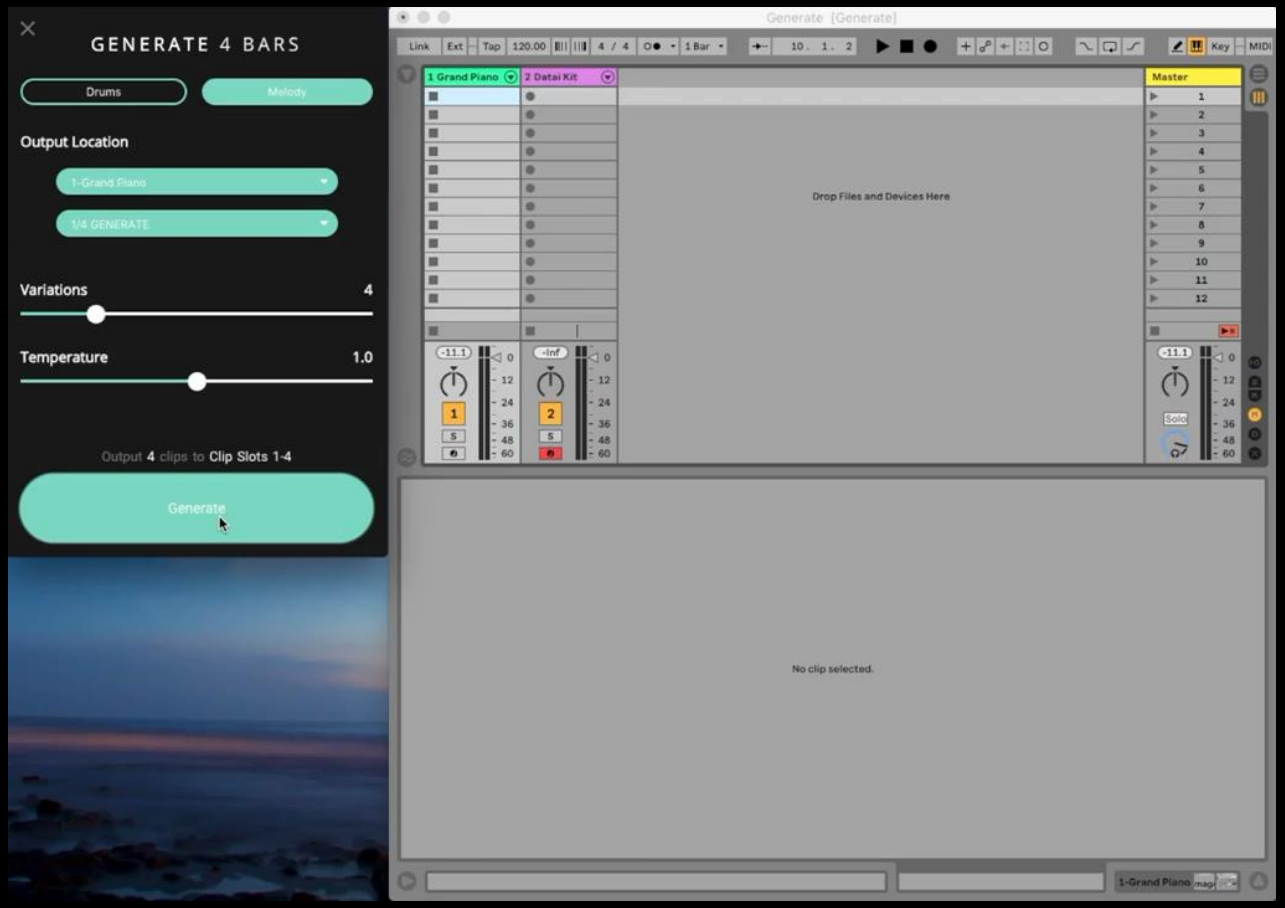
- Year: 2019
- Composition Task: Multi-Track
- Data Representation: MIDI
- Algorithm: MusicVAE, MusicRNN, GrooVAE
- Audience: Ableton users
- Deployment: Standalone desktop and Ableton plugins

Roberts A, Engel J, Mann Y, Gillick J, Kayacik C, Nørly S, Dinculescu M, Radebaugh C, Hawthorne C, Eck D. Magenta studio: Augmenting creativity with deep learning in Ableton live.

Magenta Studio

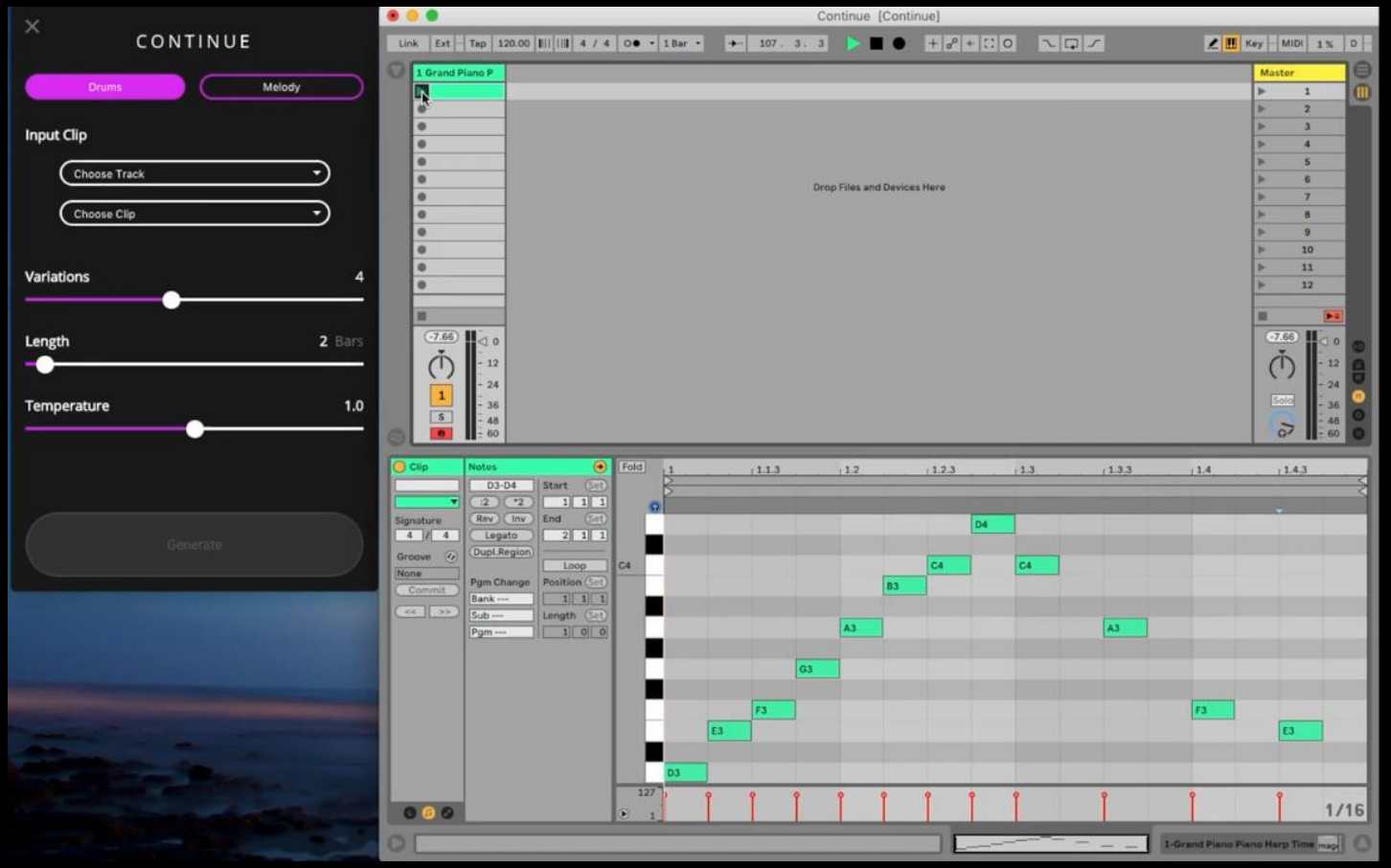
- YACHT – SCATTERHEAD (Chain Tripping // Album) (2019)

Magenta Studio



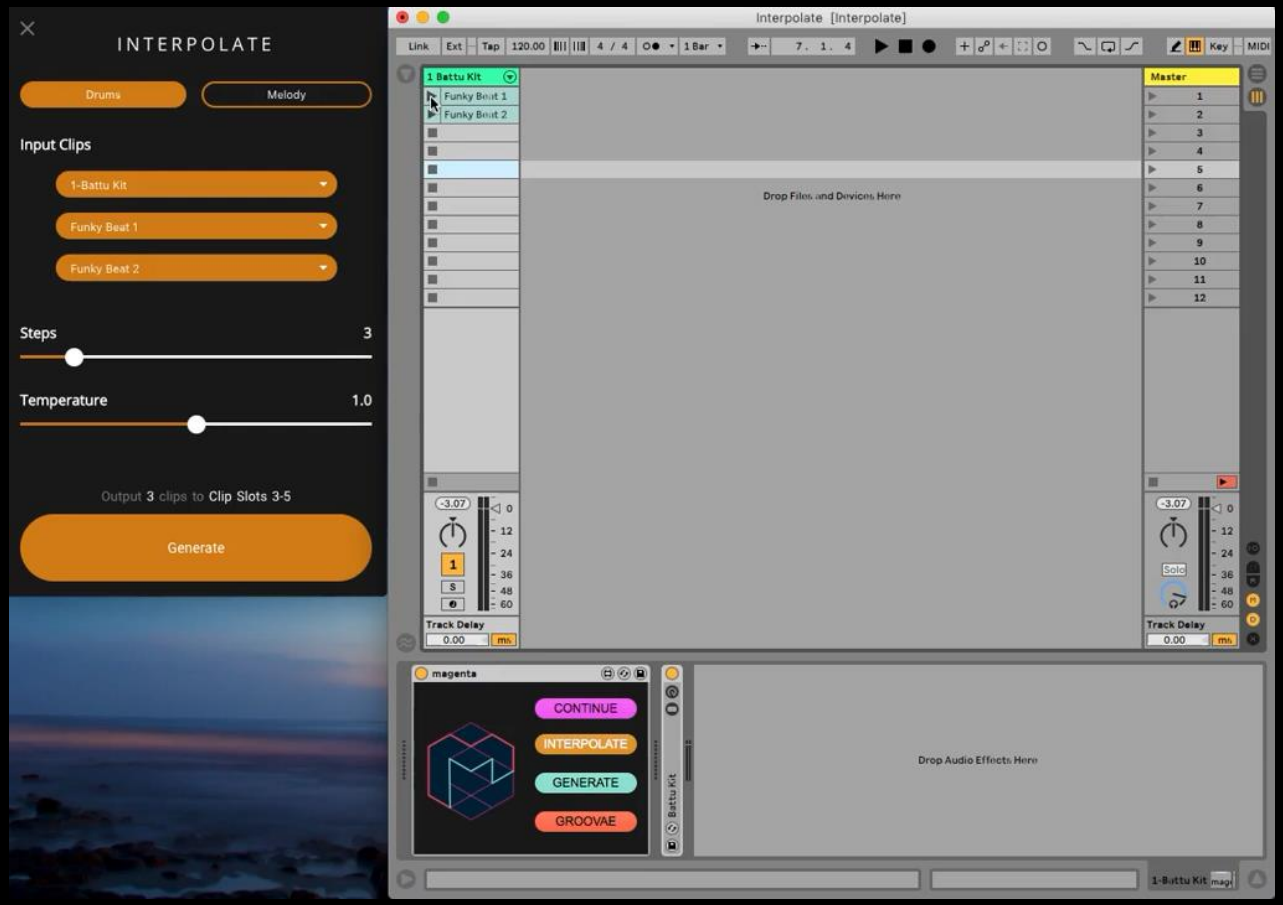
Magenta Studio: Generate Plugin

Magenta Studio



Magenta Studio: Continue Plugin

Magenta Studio



[Magenta Studio: Interpolate Plugin]

METACREATION
Lab for Creative Artificial Intelligence

SFU

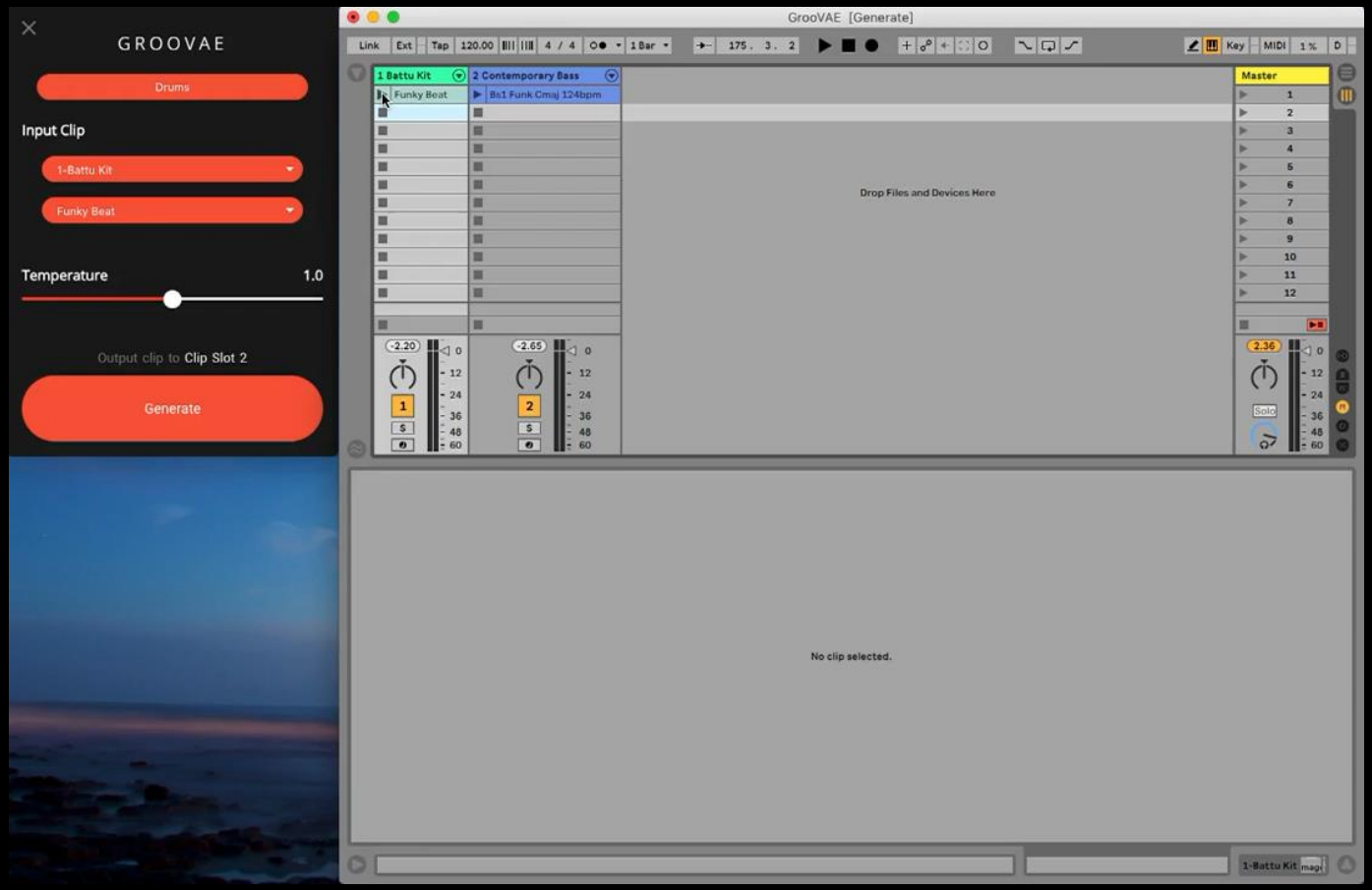
Magenta Studio

[Magenta Studio: Drumify Plugin (Bass2Drum)]

METACREATION
Lab for Creative Artificial Intelligence

SFU

Magenta Studio



Magenta Studio: Groove Plugin

MuseNet

https://openai.com/blog/musenet/

METACREATION
Lab for Creative Artificial Intelligence

SFU

MuseNet

- Year: 2019
- Composition Task: Multi-Track
- Data Representation: MIDI
- Algorithm: Sparse Transformer
- Audience: Lay users
- Deployment: Web demo

Child R, Gray S, Radford A, Sutskever I. Generating long sequences with sparse transformers. arXiv preprint arXiv:1904.10509. 2019 Apr 23.

MuseNet

- Prompt_ First 5 notes of Chopin Op. 10, No. 9.mp3 
- Prompt_ Jazz Piano-Bass-Drums.mp3 
- Prompt_ Bon Jovi and the first 6 notes of Chopin Op. 27, No. 2 

MuseNet

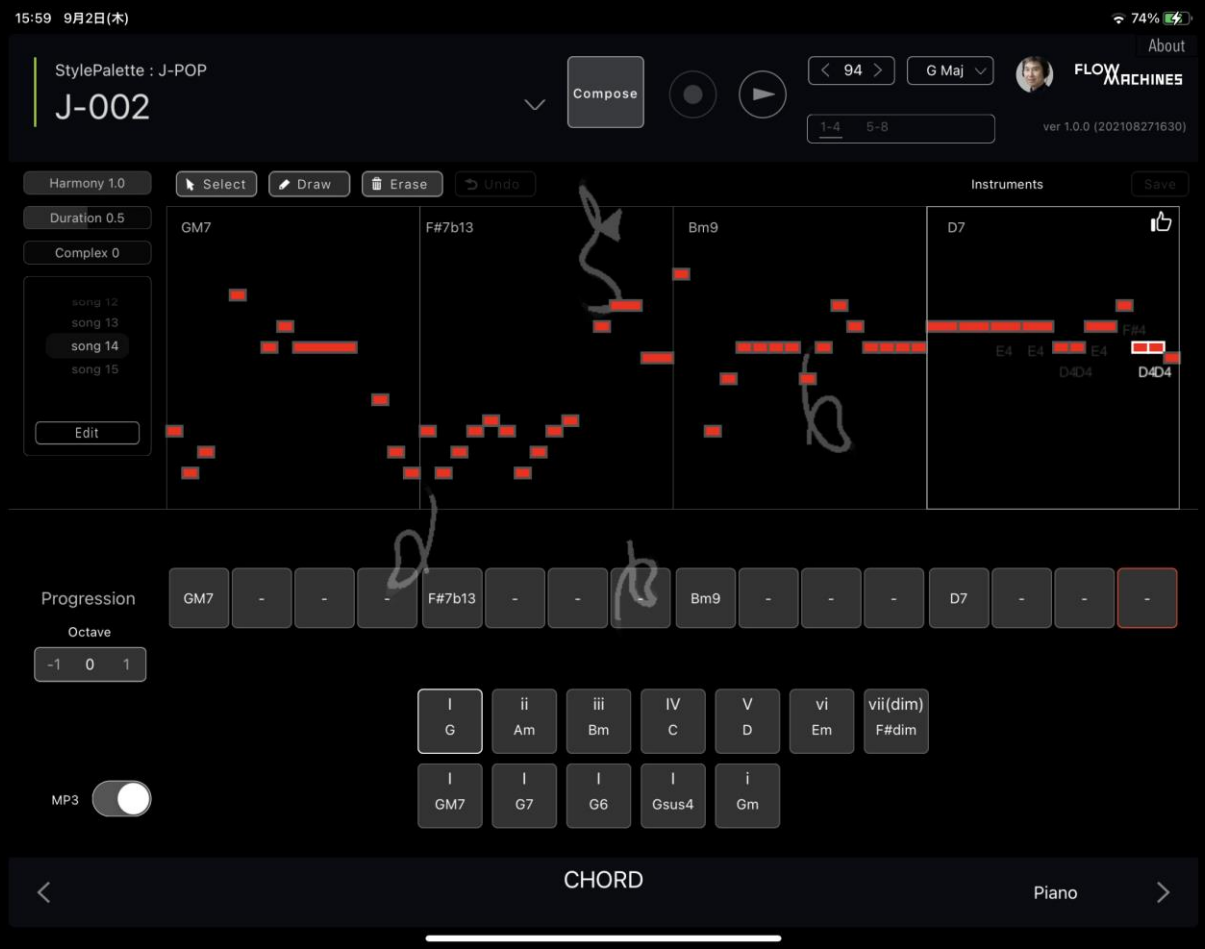
Some of MuseNet's limitations include:

- The instruments you ask for are strong suggestions, not requirements. MuseNet generates each note by calculating the probabilities across all possible notes and instruments. The model shifts to make your instrument choices more likely, but there's always a chance it will choose something else.
- MuseNet has a more difficult time with odd pairings of styles and instruments (such as Chopin with bass and drums). Generations will be more natural if you pick instruments closest to the composer or band's usual style.

Composer and Instrumentation Tokens

We created composer and instrumentation tokens to give more control over the

Flow Machines



https://www.flow-machines.com/

METACREATION
Lab for Creative Artificial Intelligence

SFU

Flow Machines

- Year: 2013 /2021
- Composition Task: Multi-Track
- Data Representation: MIDI
- Algorithm: Proprietary
- Audience: Lay users
- Deployment: iOS App

Flow Machines

- Benoît Carré - Daddy's Car (2016)
- SKYGGE feat Kiesza, Stromae and The Bionix - Hello Shadow (Hello World Album, 2018)

Flow Machines (2021)

Flow Machines Mobile Walkthrough | Flow Machines (English)

METACREATION
Lab for Creative Artificial Intelligence

SFU

DrumVAE

https://vibertthio.com/drum-vae-client/

METACREATION
Lab for Creative Artificial Intelligence

SFU

DrumVAE

- Year: 2019
- Composition Task: Multi-Track
- Data Representation: MIDI
- Algorithm: DrumVAE
- Audience: Lay users
- Deployment: Web Application

Thio V, Liu HM, Yeh YC, Yang YH. A minimal template for interactive web-based demonstrations of musical machine learning. arXiv preprint arXiv:1902.03722. 2019 Feb 11.

DrumVAE

DrumVAE: Latent Inspector (2018) by Vibert Thio

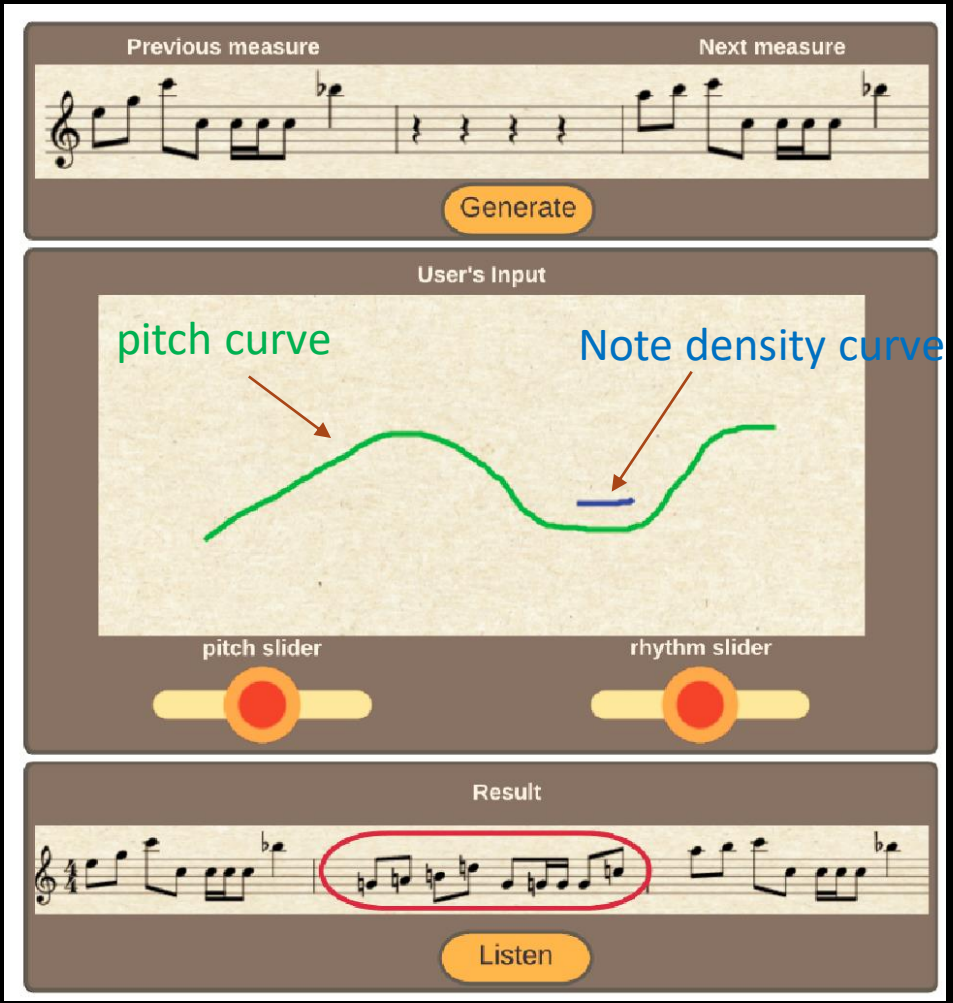
METACREATION
Lab for Creative Artificial Intelligence

SFU

Draw and Listen: A Sketch Based System for Music Inpainting

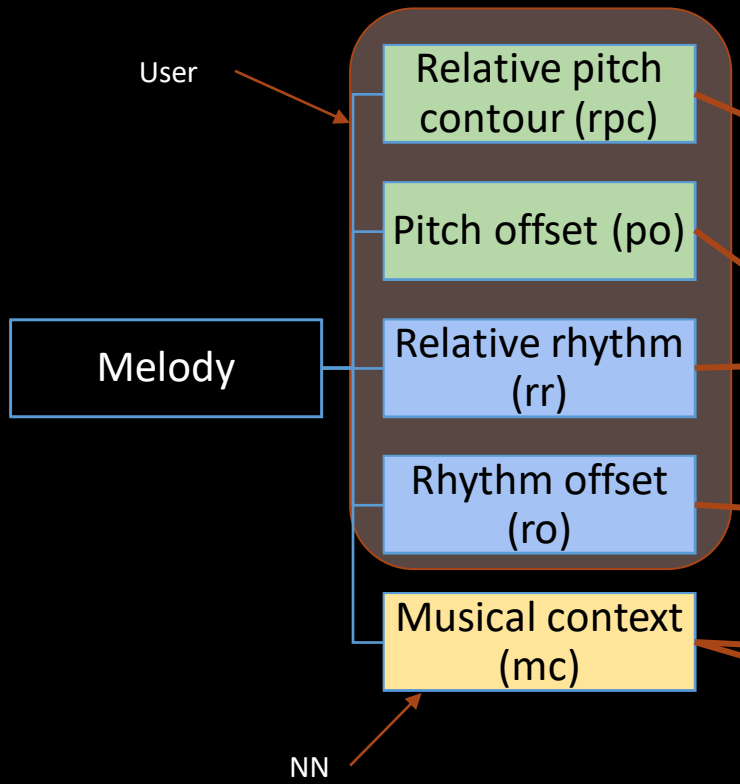
Christodoulos Benetatos, Zhiyao Duan, TISMIR 2022.

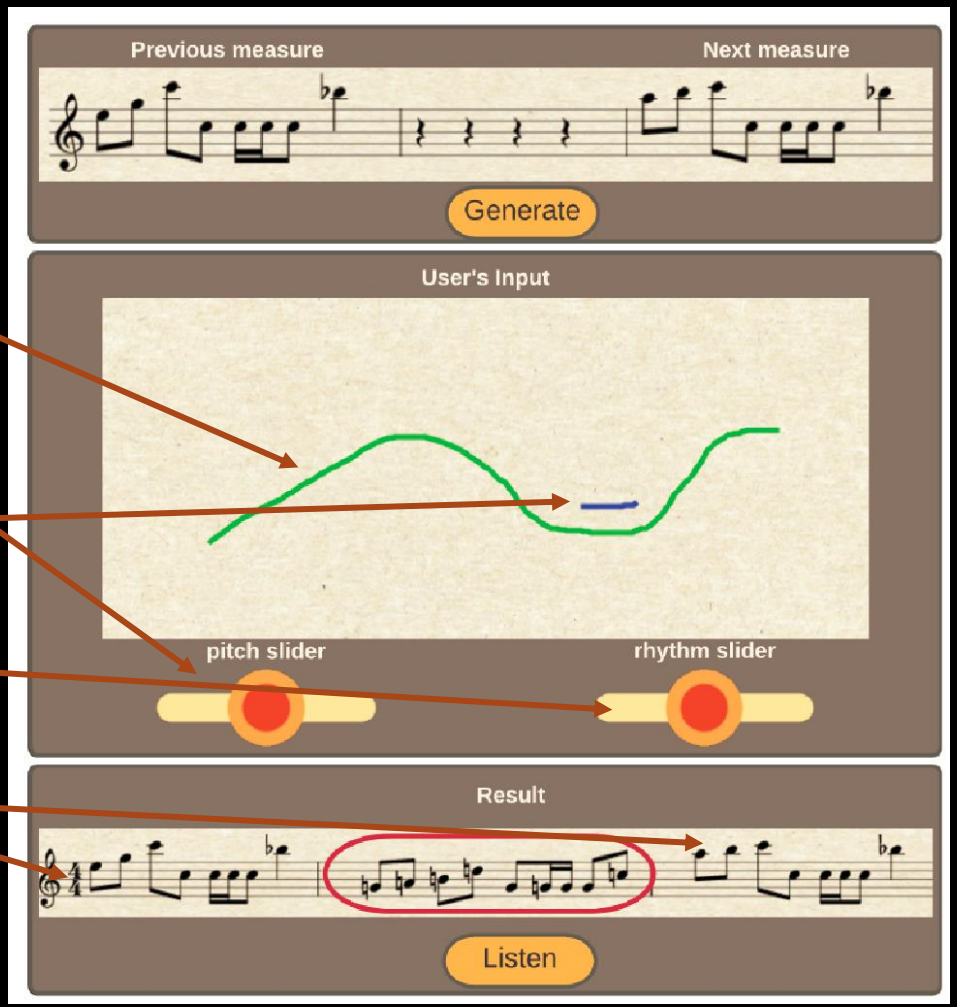
- Draw curves to guide inpainting
  - Intuitive
    Non-musicians understand notions
    of low-high pitch and sparse-dense rhythm
  - No music theory knowledge required



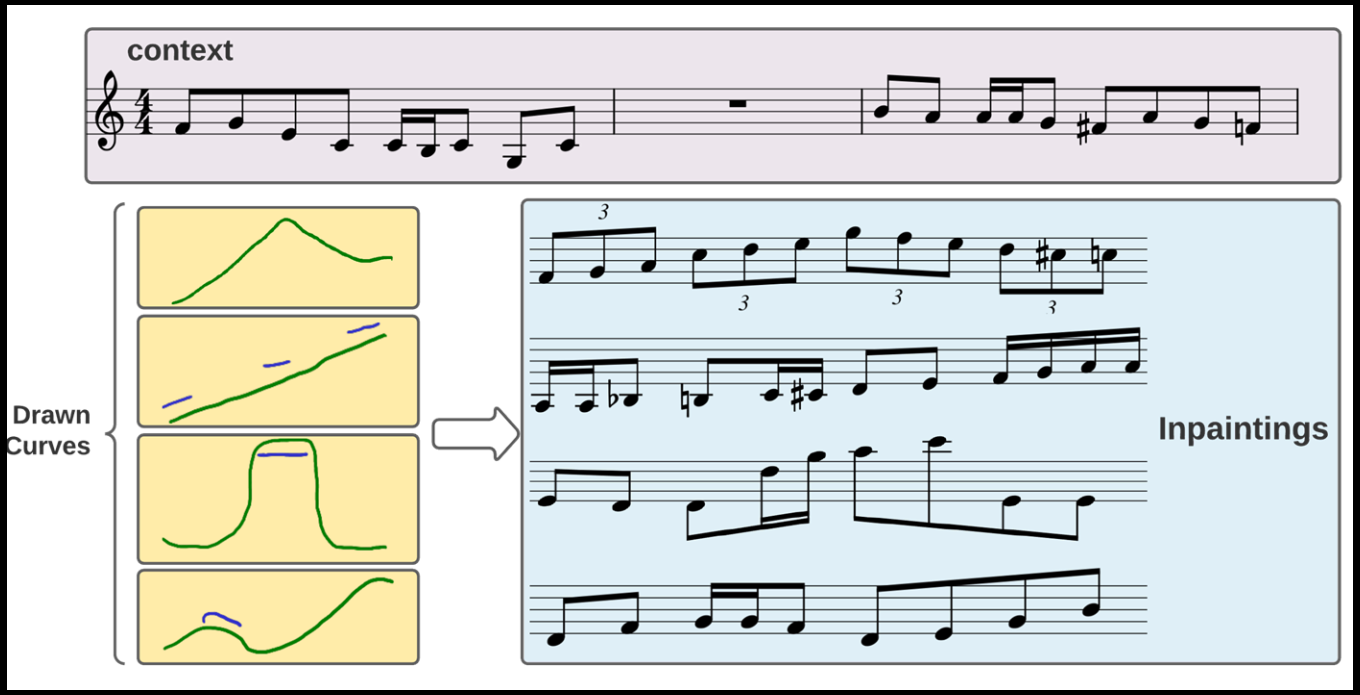
Draw and Listen: A Sketch Based System for Music Inpainting

Proposed Melody Decomposition

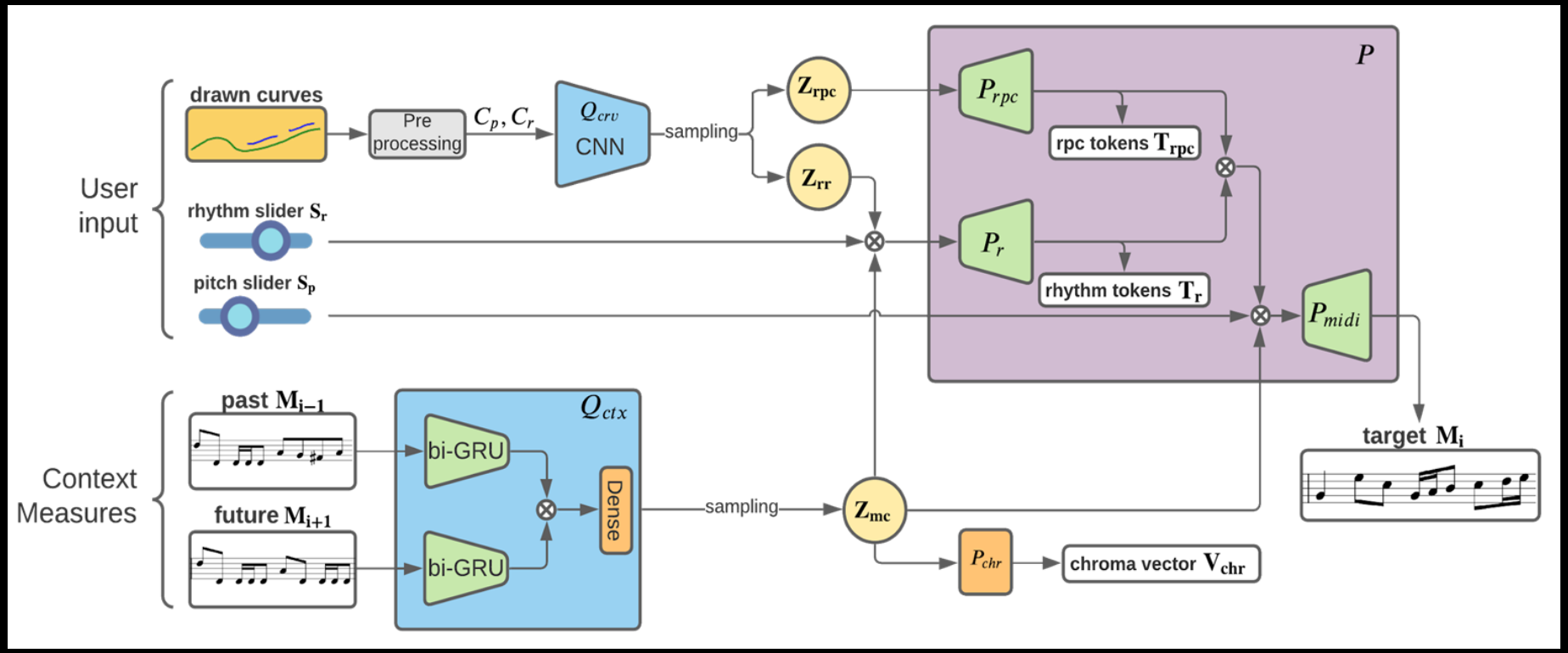




Examples
4 different user inputs for the same context measures



Model Architecture
Based on Variational Auto-Encoders (VAE)



With Jeff Ens, Arxiv, 2020, ISMIR 2021

Multitrack Music Machine: MMM

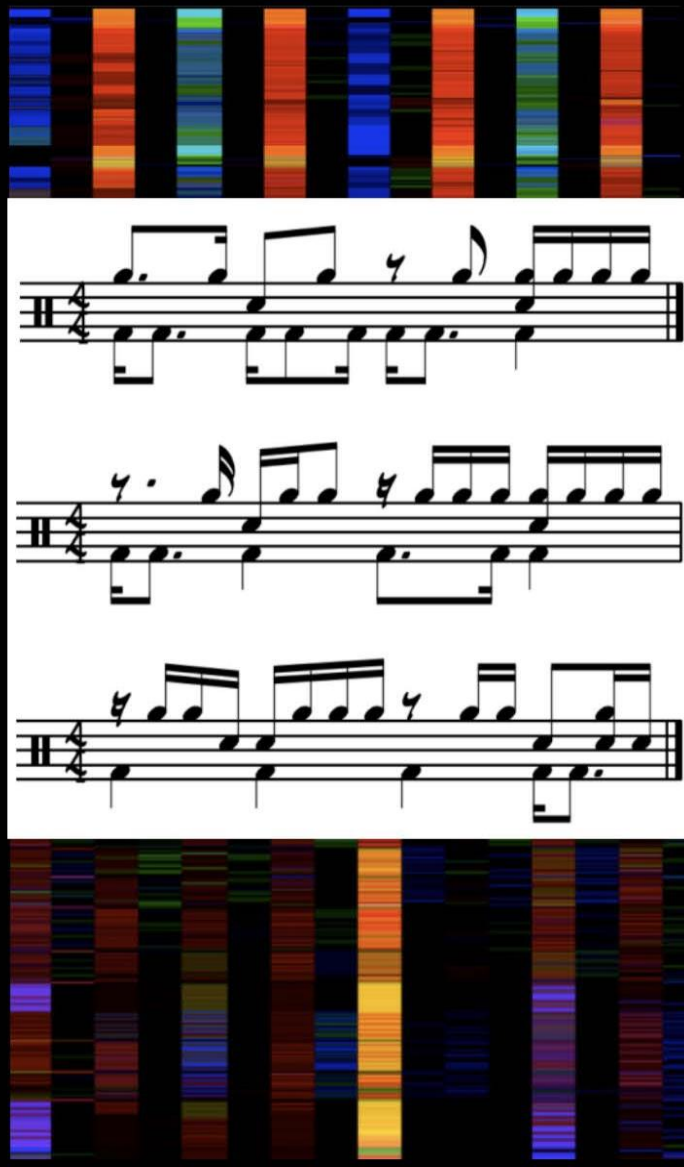
- Symbolic music generation model based on the GPT-2 Transformer architecture: 16 bars of attention.
- More versatile than its competition

System	Output Specifications					Generation Tasks	
	Number of Tracks	Number of Instruments	Fixed Schema	Drums	Track-Level Polyphony	Infilling	Attribute Control
MMM	⋆	128	-	x	x	x	Section 4
MuseNet Payne [2019]	10	10	-	x	x	-	x
MuseGAN Hong *et al.* [2019]	4	4	x	x	x	-	-
LahkNES Donahue *et al.* [2019]	4	4	x	x	-	-	-
CoCoNet Huang *et al.* [2017]	4	4	x	-	-	x	-
MusicVAE Roberts *et al.* [2018]	3	3	x	x	-	-	-
MusIAC Guo *et al.* [2022]	3	3	x	-	x	x	x
SketchNet Chen *et al.* [2020]	1	-	x	-	-	x	x
Pati *et al.* [2019]	1	-	x	-	-	x	-
Mittal *et al.* [2021]	1	-	x	-	-	x	-
Chang *et al.* [2021]	1	-	x	-	x	x	-
Chi *et al.* [2020]	1	-	x	-	x	x	-
Tan and Herremans [2020]	1	-	x	-	x	-	x
Wang and Xia [2021]	1	-	x	-	x	-	x
Wang *et al.* [2020]	1	-	x	-	x	-	x

With Jeff Ens, ISMIR 2021.

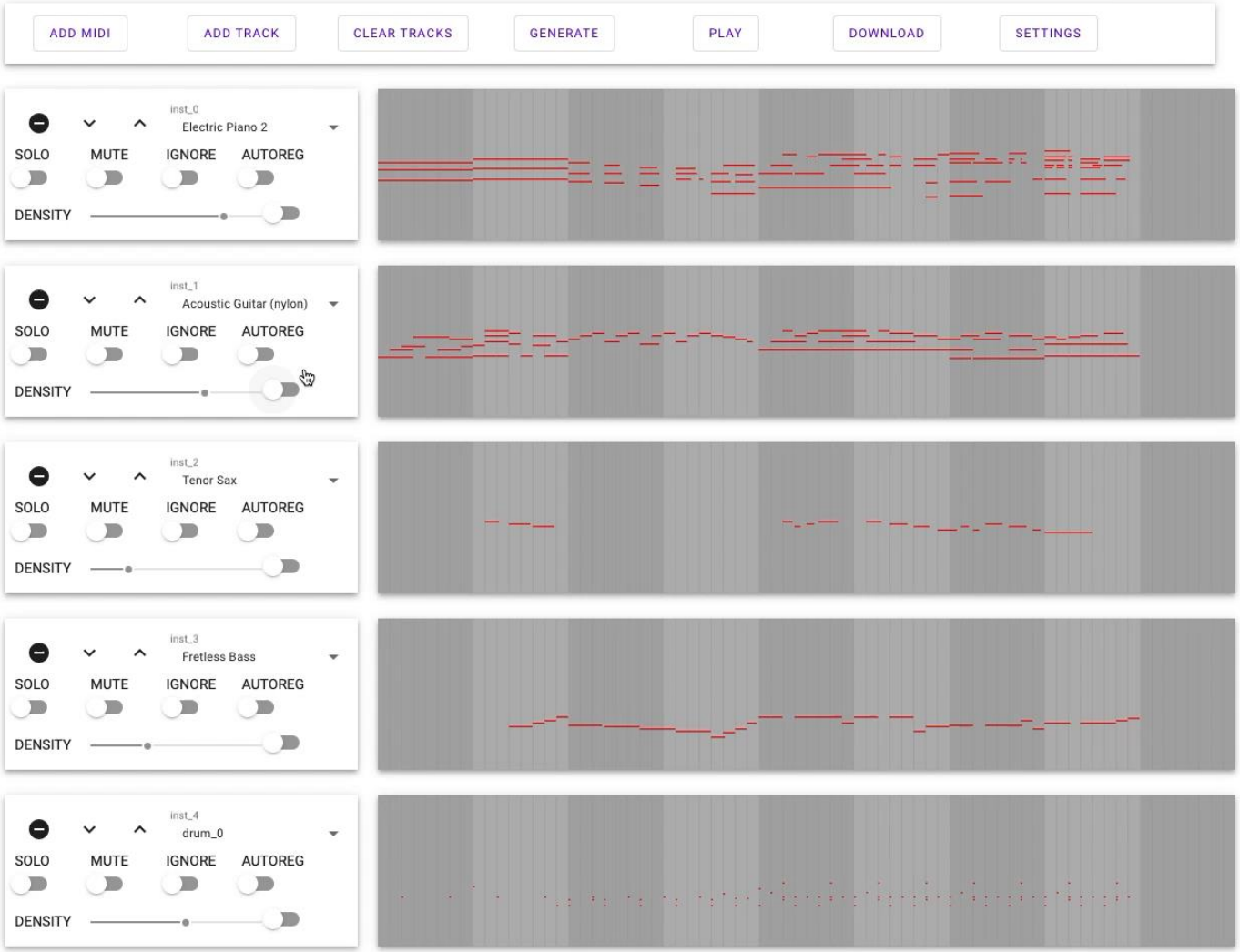
MetaMIDI Dataset

- **MIDI in/out: symbolic music representation**
- **MetaMIDI Dataset:**
  - 445,631 MIDI files
  - 221,504 MIDIs with Artist/Title metadata
  - 237,236 MIDIs matched to 10M Audio
- 3-10x Lahk Midi Dataset (existing)
  - Available on Zenodo (1k+ downloads)

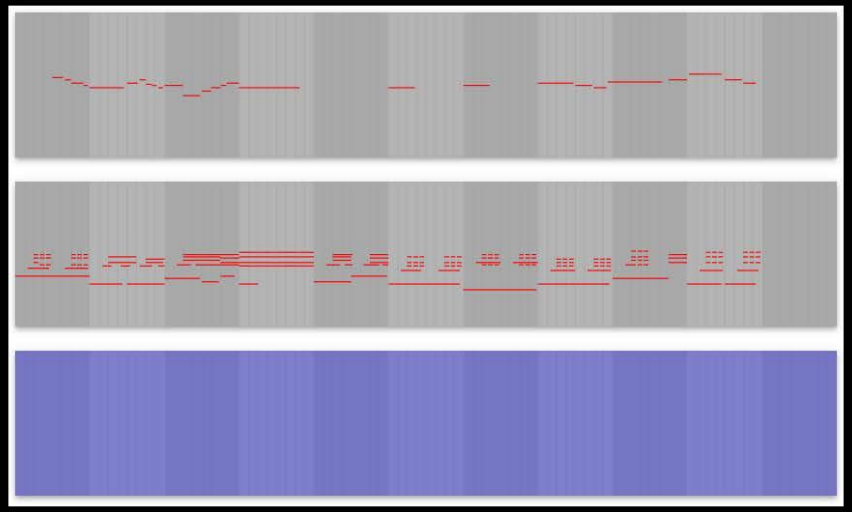


MMM: Multitrack Music Machine

with Jeff Ens, Arxiv 2020, ISMIR 2021

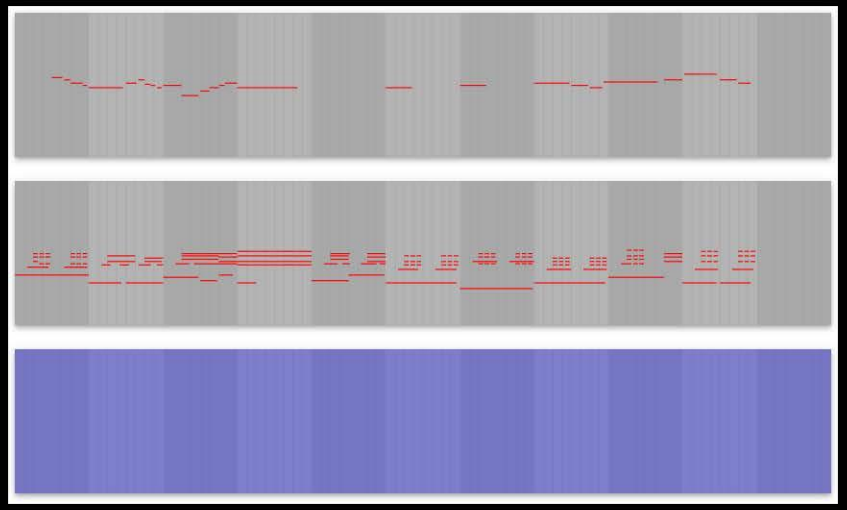


Flexible Generation (track in-filling)



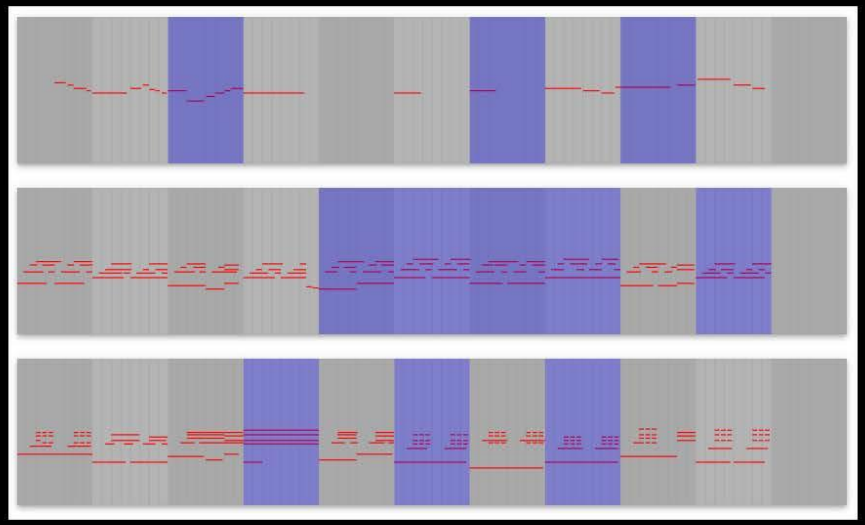
Generate a new track (shown in blue) conditioned on a set of tracks.

Flexible Generation (track in-filling)



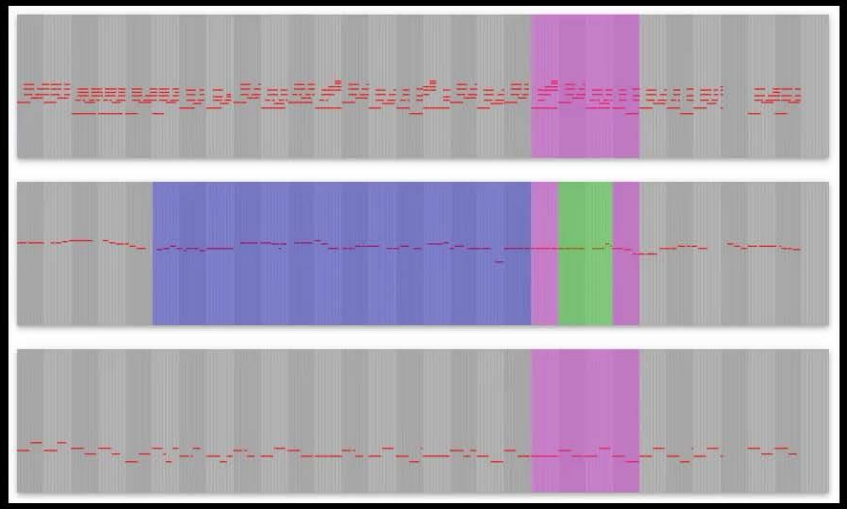
RE-Generate a track (shown in blue) conditioned on a set of tracks.

Generation (Bar in-filling)



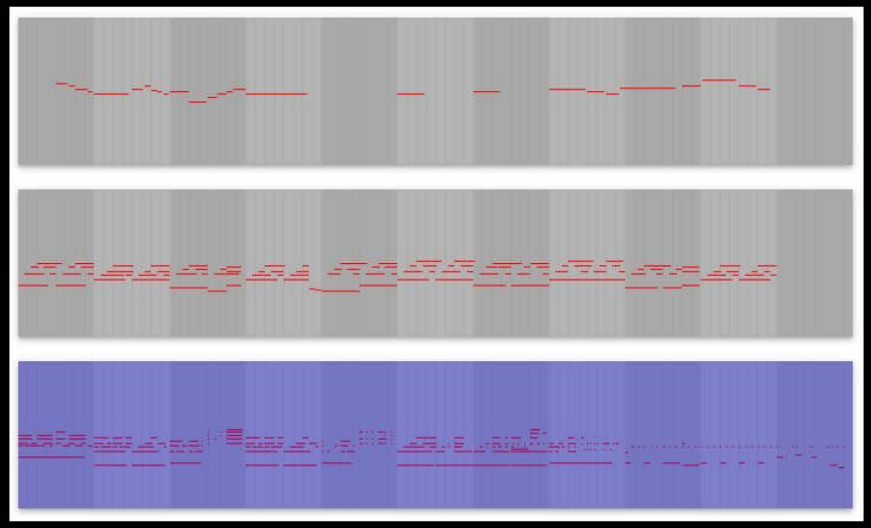
Re-generate some bars (shown in blue) conditioned on the remaining bars.

Flexible Generation



Generation of longer sections

Flexible Generation (controls)



You can control the note density of a generated track (shown in blue).

MMM Generation Control

With Jeff Ens, ISMIR 2020.

ISMIR 2020.

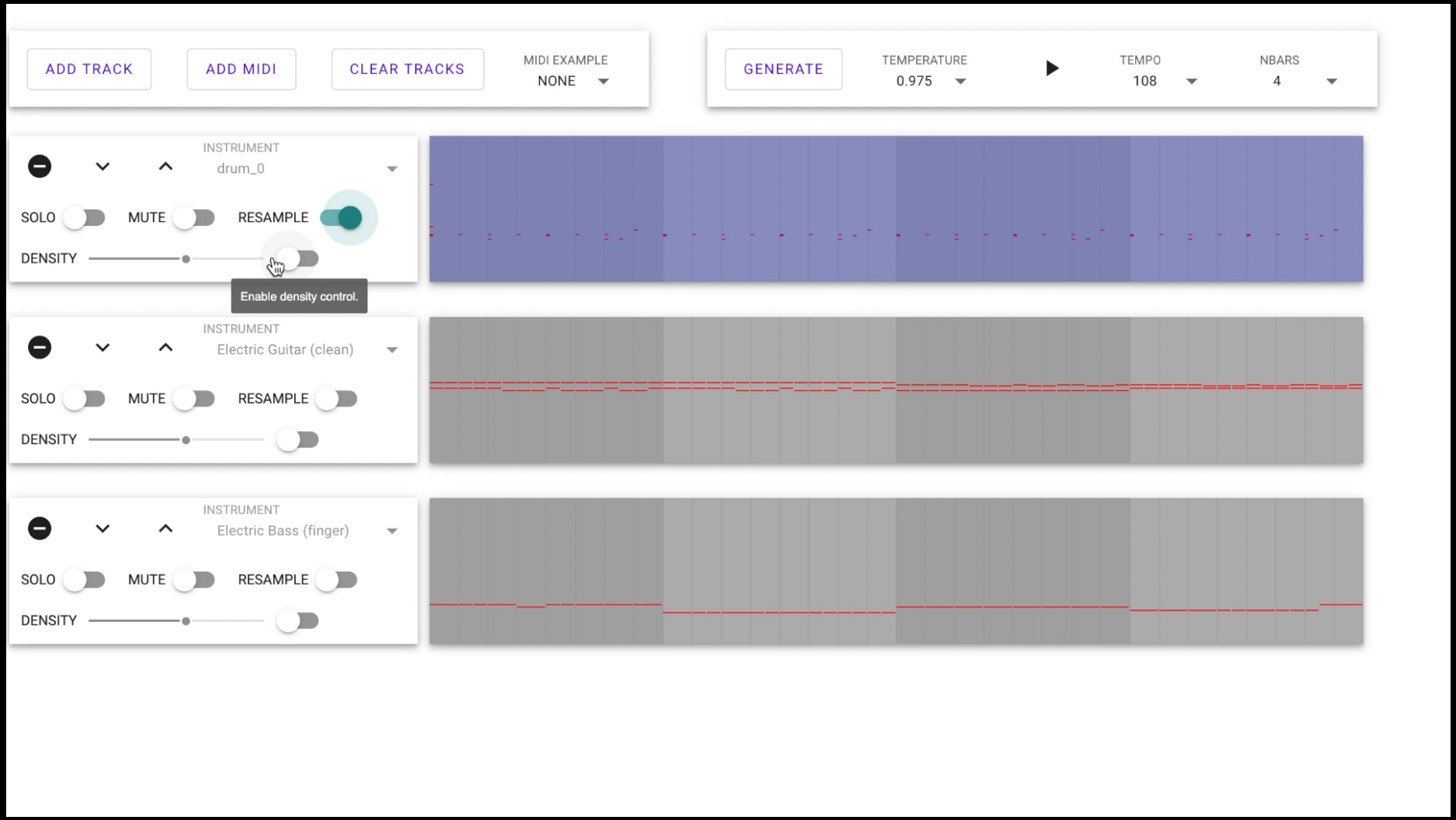
- Increased number of generation controls:
  - Time signature
  - Key, mode, forbidden notes, …
  - Note density (track relative)
  - Note duration (value or range)
  - Amount of silence
  - Min / Max polyphony (or monophony)
  - With/without interpretation (velocity)
  - Musical style control (reggae, pop, disco, Bach, Beatles, Metallica, …)
  - Spotify-like attributes: Danceability, …
  - Affective control: Valence / Arousal / Tension
  - Interpretation (velocity, micro-timing and groove)

Interpretation (velocity, micro-timing and groove)

MMM: Multitrack Music Machine

with Jeff Ens, ISMIR 2020

Demo available: https://jeffreyjohnens.github.io/MMM/

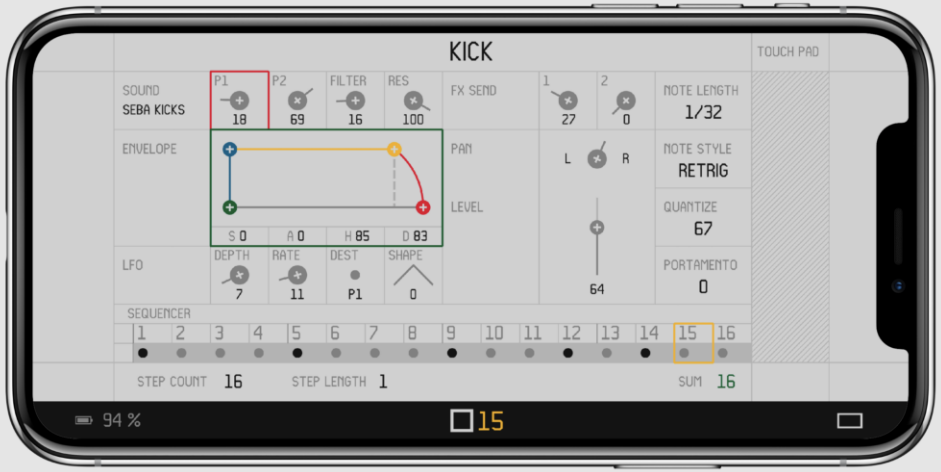


Applications

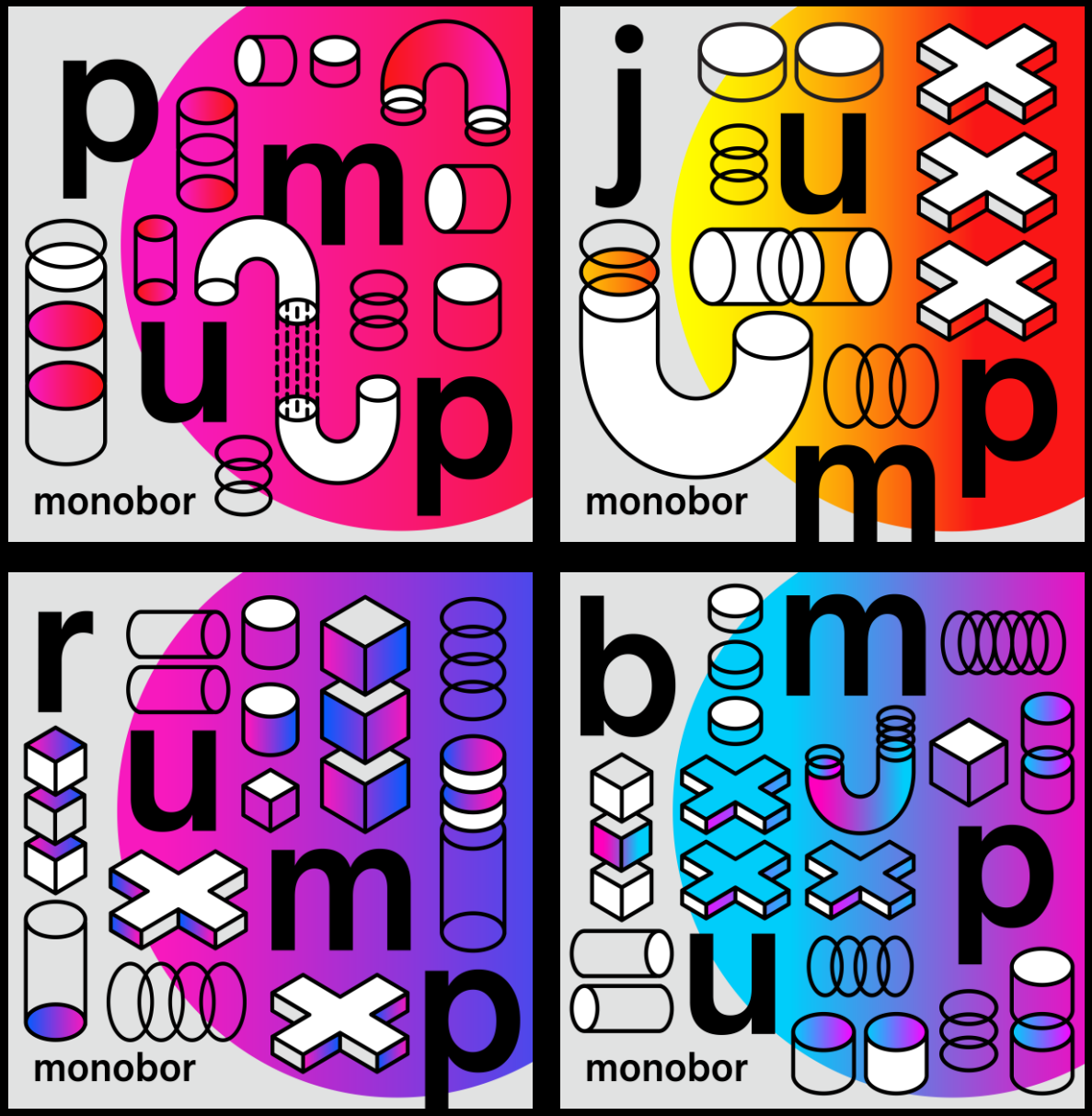
Integrating MMM in existing interfaces

Integrating to Synthesizers

teenage engineering

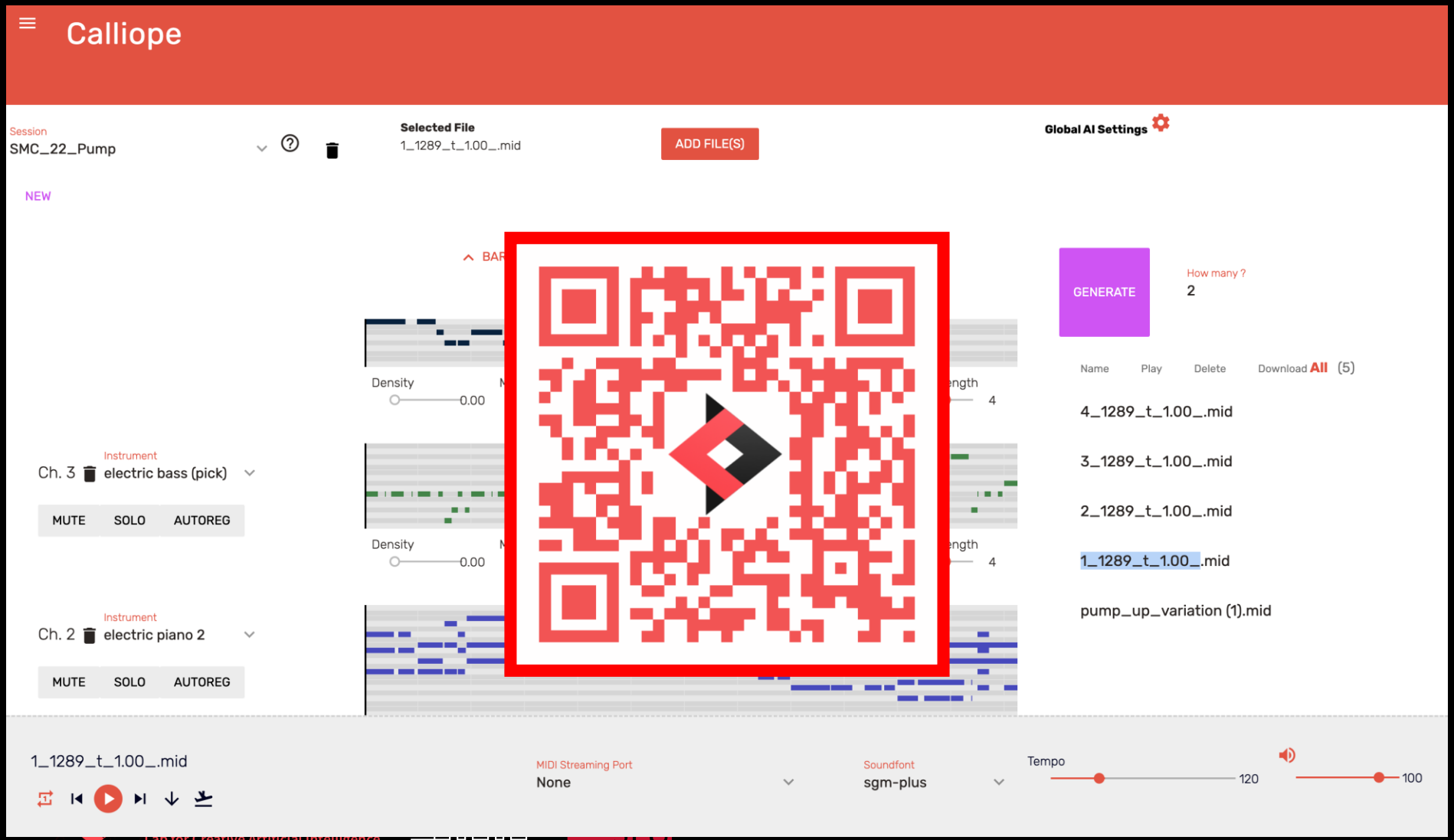


Making and Releasing Music



Calliope: Online System

With Renaud Bougueng Tchemeube, Jeff Ens, SMC *2022*, *ICCC 2022*, *C&C2022*.



Calliope

- Year: 2021
- Composition Task: Multi-Track
- Data Representation: MIDI
- Algorithm: MMM
- Audience: Amateurs and Professionals
- Deployment: Web Application

Calliope: Online System

With Renaud Bougueng Tchemeube, Jeff Ens, SMC *2022*, *ICCC 2022*, *C&C2022*.

METACREATION
Lab for Creative Artificial Intelligence

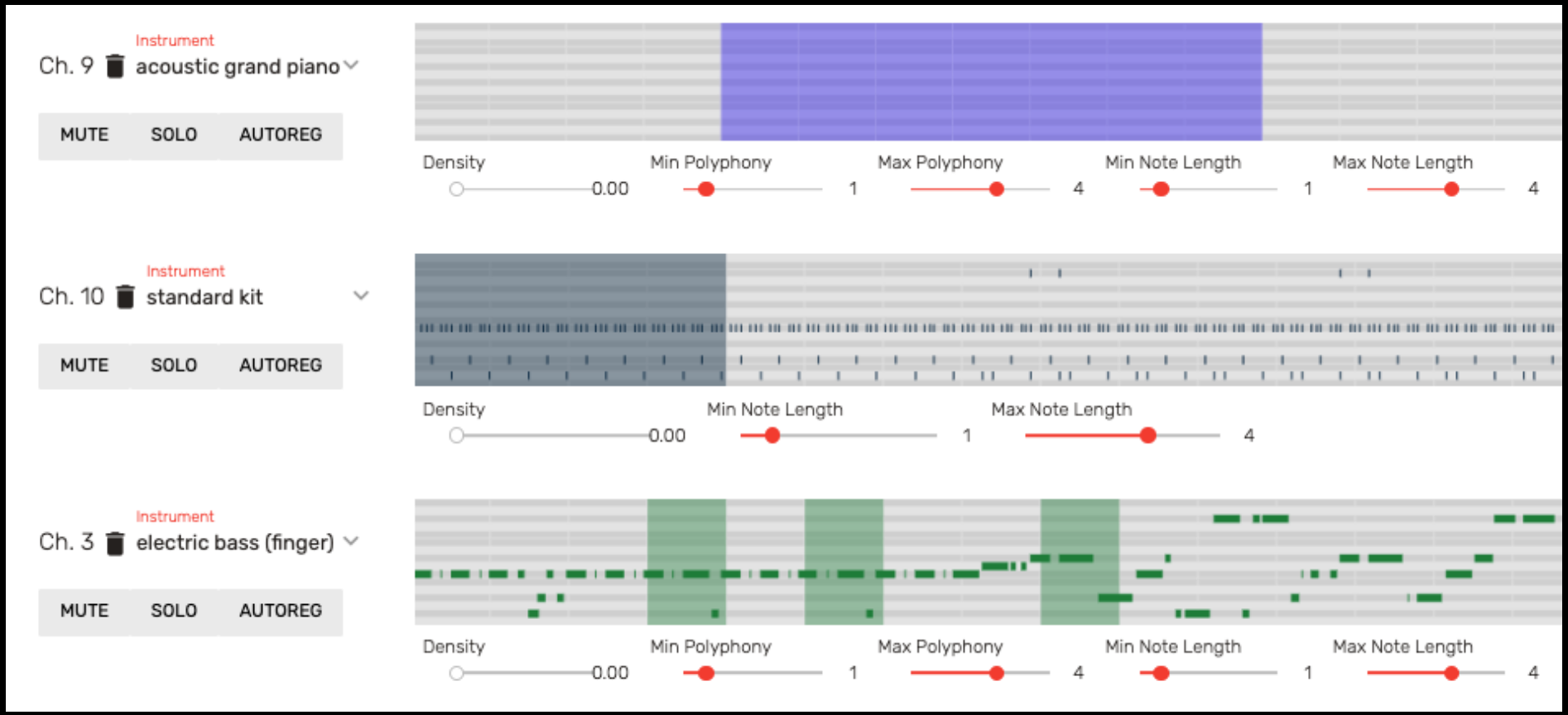
SFU

Apollo

- Apollo medley sessions
- Apollo medley session – variation 1
- Apollo medley session – variation 2

- Apollo EP – Hard Disk Fever (Original)
- Apollo EP – Hard Disk Fever (Melody variation)
- Apollo EP – Hard Disk Fever (New Brass section)

Bar Selection & Local Settings



METACREATION
Lab for Creative Artificial Intelligence

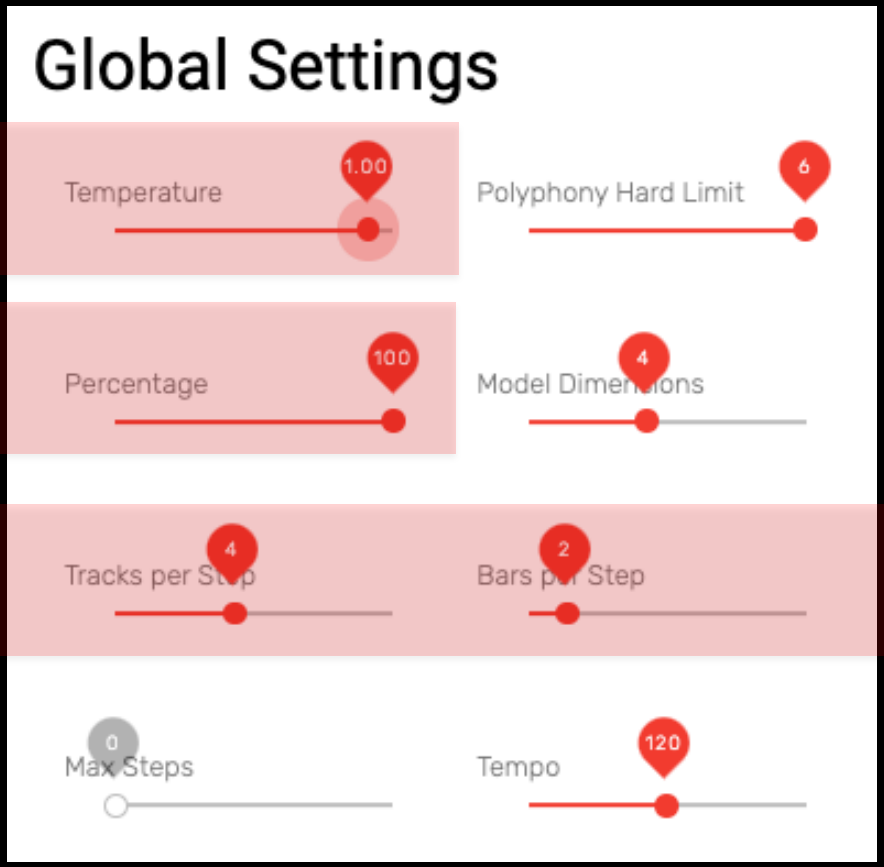
SFU

Global Settings

Craziness / entropy

Variation engine

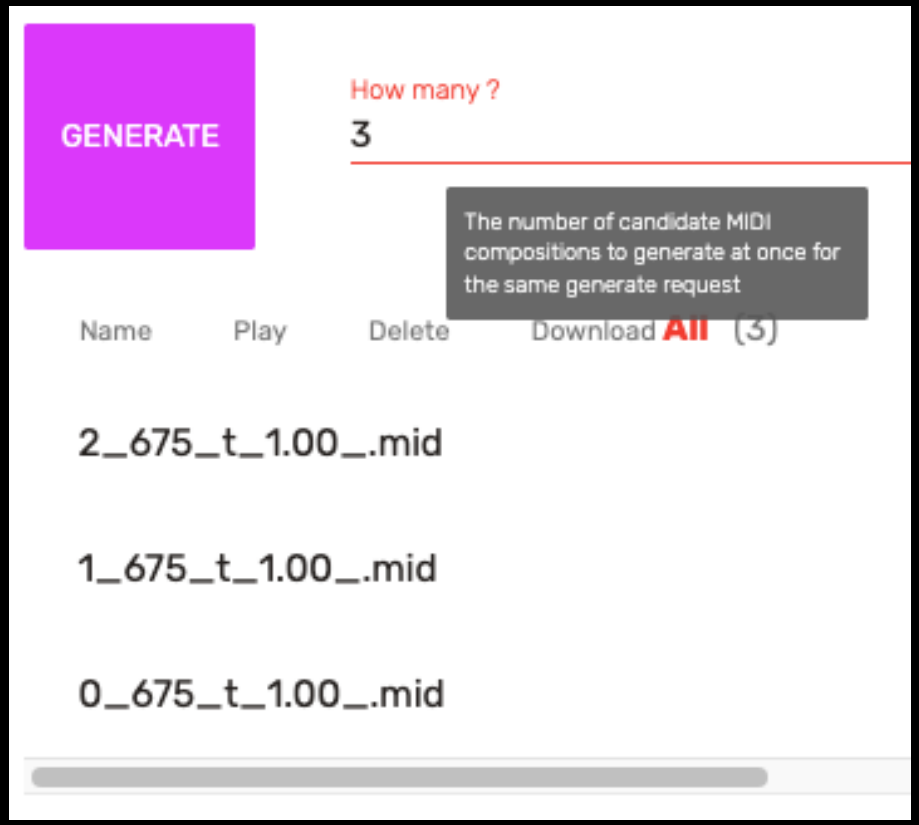
Sampling control



METACREATION
Lab for Creative Artificial Intelligence

SFU

Batch Generation and Ranking



- Batch Generate your in-filling to toggle through options
- Generate 10s, 100s, 1000s of variations of entire compositions
- Manage variations by ranking them according to similarity to a given track.

Jeff Ens and Philippe Pasquier. "Quantifying Musical Style: Ranking Symbolic Music based on Similarity to a Style". ISMIR, 2019, pp. 870–877.

A Cross-Domain Analytic Evaluation Methodology for Style Imitation, Jeff Ens, Philippe Pasquier, *In Proceedings of the 9th International Conference on Computational Creativity, ICCC 2018.* **Best Paper Award**

METACREATION
Lab for Creative Artificial Intelligence

SFU

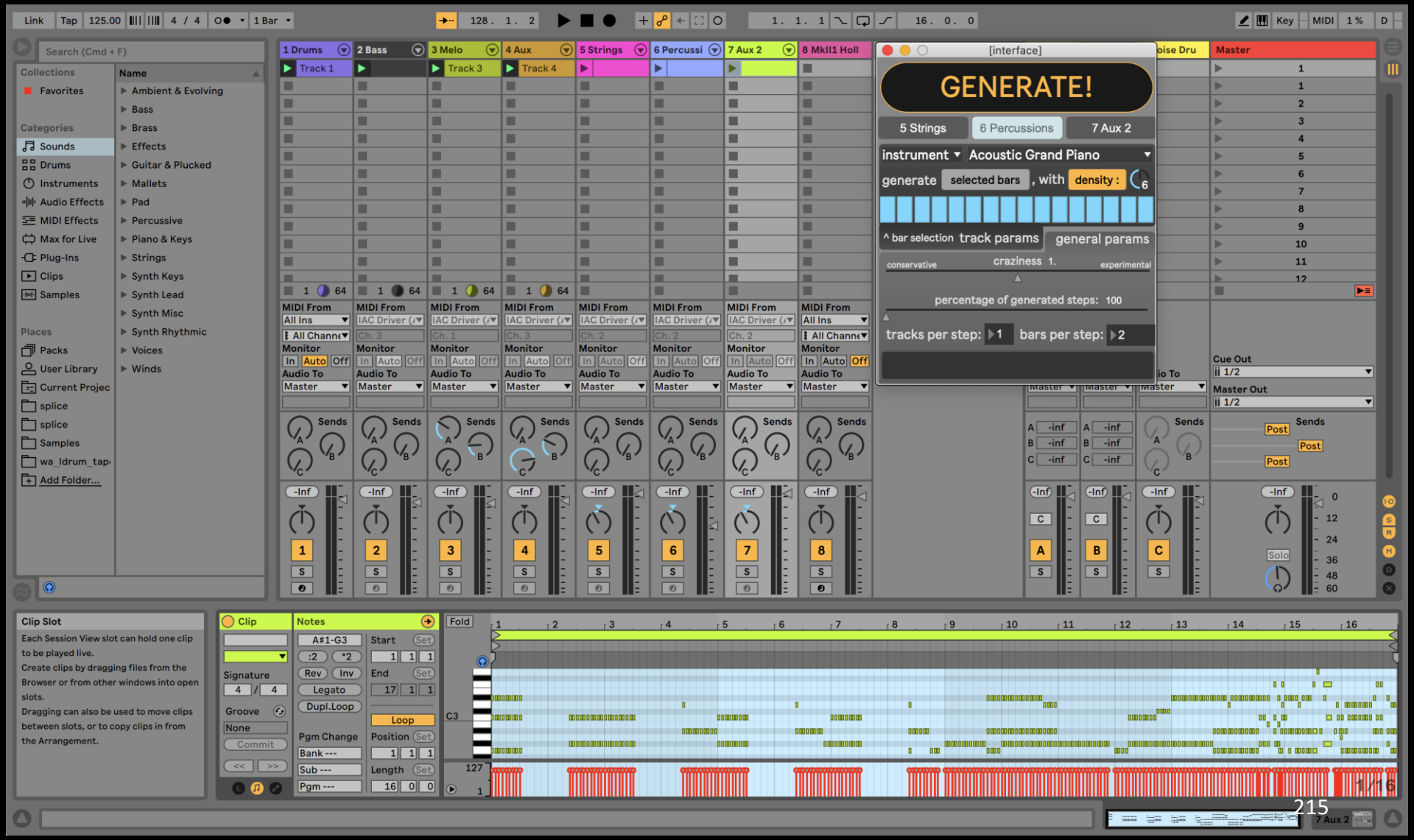
MMM4Live

- Year: 2021
- Composition Task: Multi-Track
- Data Representation: MIDI
- Algorithm: MMM
- Audience: Amateur, Professionals
- Deployment: Max For Live plugin

Ens J, Pasquier P. Mmm: Exploring conditional multi-track music generation with the transformer. arXiv preprint arXiv:2008.06048. 2020 Aug 13.

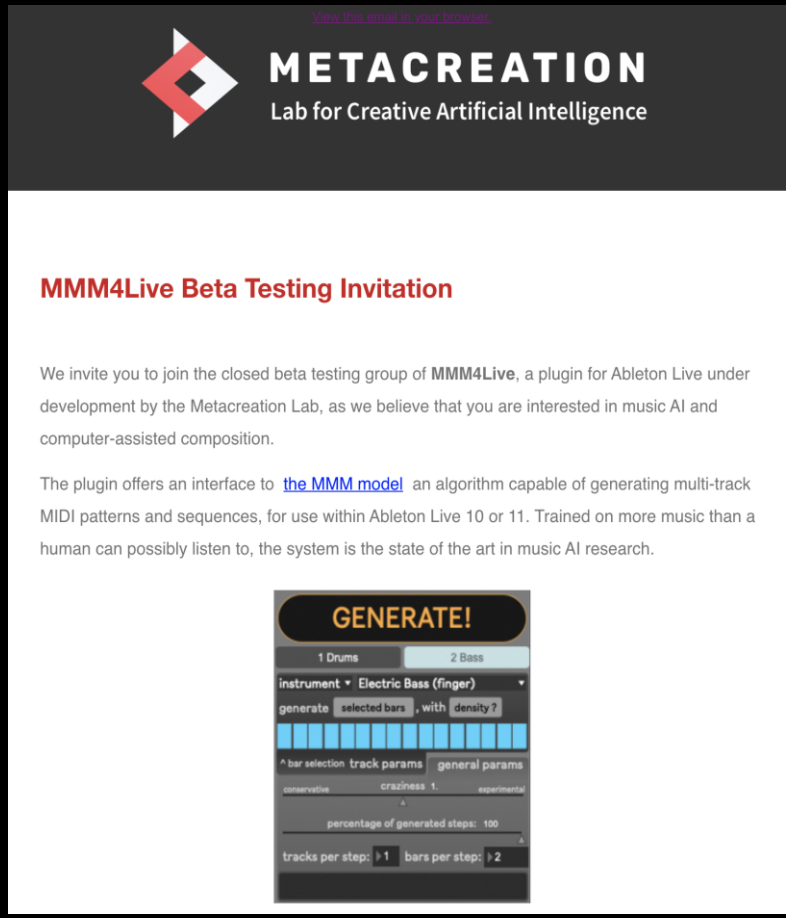
Alpha testing MMM4Live

https://metacreation.net/mmm4live/



Onboarding beta-testing of MMM4Live

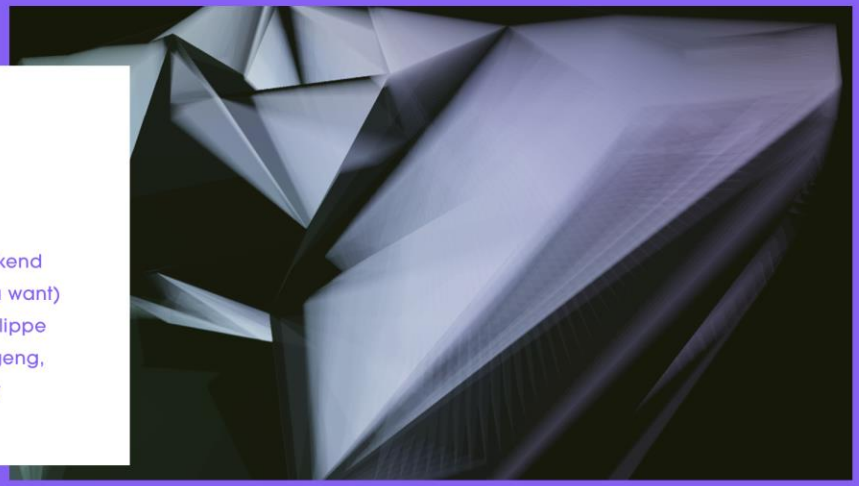
- Works with Ableton Live 10 and 11, MacOS (M).
- To get on our alpha/beta-tester list, please email me: ppa12@sfu.ca
- To be released publicly in Fall 2023 (?).



SONG CONTEST 2021

About Participants Jury AI & Creation day FAQs Blog Press Contact

Metacreation Lab



AI Song Contest 2021 / participants

TEAM / Metacreation Lab
SONG / A song about the weekend (and you can do whatever you want)
TEAM MEMBERS / Cale Plut, Philippe Pasquier, Jeff Ens, Renaud Bougeng, Tara Jadidi, and Dimiter Zlatkov

LISTEN AND EVALUATE

A song about the weekend (and you can do whatever you want)

https://www.aisongcontest.com/

Evaluating Human-AI Interaction with MMM-Cubase: A Creative AI System for Music Composition

Author Name
Affiliation
email@example.com

Abstract

With the rise of artificial intelligence (AI), there has been increasing interest in human-AI co-creation in a variety of artistic domains including music as AI-driven systems are frequently able to generate human-competitive artifacts. Now, the implications of such systems for the musical practice are being investigated. This paper reports on the thorough evaluation of the user adoption of the Multi-Track Music Machine (MMM) as a minimal co-creative AI tool for music composers. To do this, we integrate MMM into Cubase, a popular Digital Audio Workstation (DAW), by producing a "1-parameter" plugin interface named MMM-Cubase, which enables human-AI co-composition. We conduct a 3-part mixed method study measuring usability, user experience and technology acceptance of the system across two groups of expert-level composers: hobbyists and professionals. Findings indicate no significant difference between the two groups while informing on the potential of incorporating such capable co-creative tools, particularly variations with improved controllable interfaces, into the music

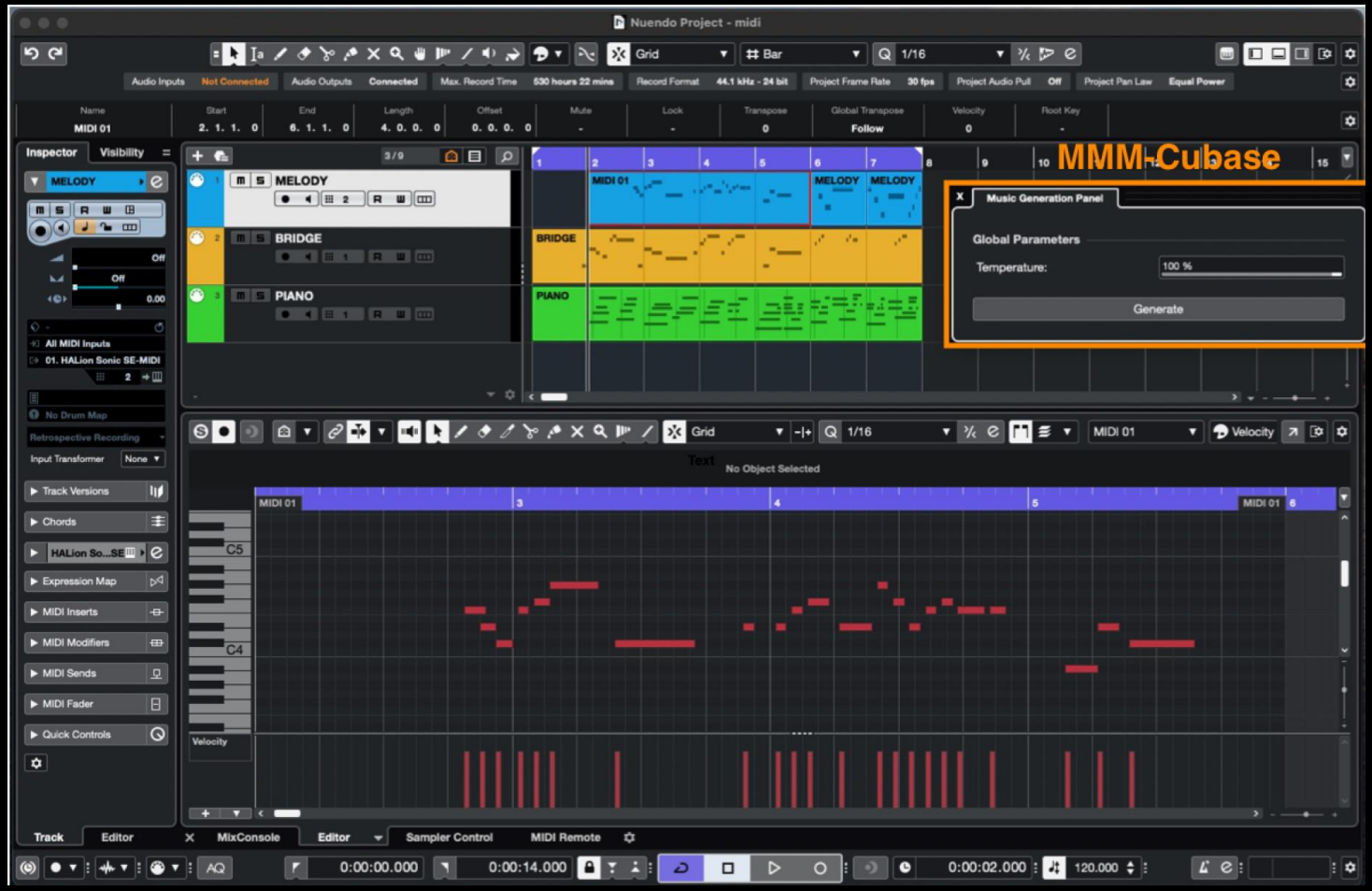
Figure 1: MMM-Cubase's Interface in Cubase

challenges, Ens et al. develop the Multi-Track Music Machine (MMM) [Ens and Pasquier, 2020a], a machine learning (ML) music system capable of generating multi-track symbolic music in a controlled manner. MMM is a powerful and highly controllable generative model with the ability to fully

Usability and Acceptability Evaluation (CuBase)

- **Competence (quality)**: Is the system perceived as reliable and competent at its task?
- **Efficiency**: Does the system allow saving time or effort?
- **Agency (control, expressivity)**: Does the user feel authorship over the output of the co-creative process?
- **Phenomenology (authorship, trust)**: Besides the surface-level experience, what are the felt and affective, subjective, impacts of using such systems?

Usability and Acceptability Evaluation (CuBase)



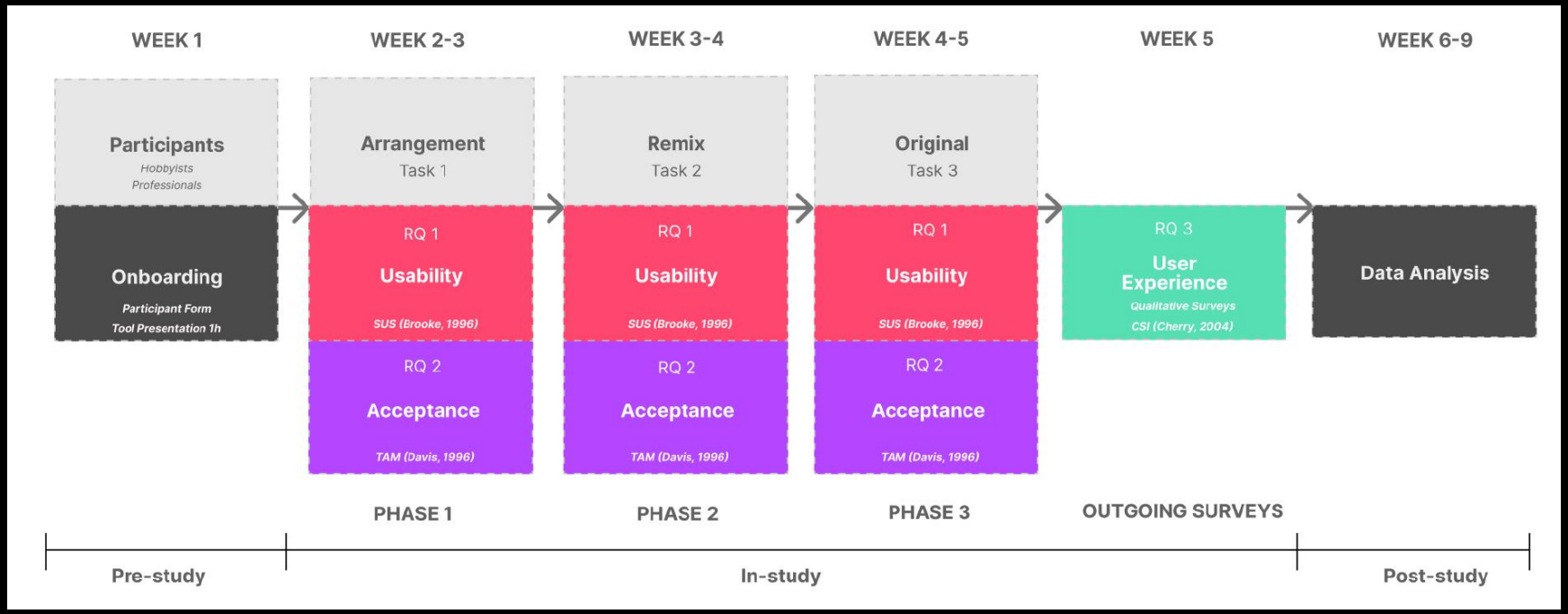
METACREATION
Lab for Creative Artificial Intelligence

SFU

steinberg

Study Protocol

With Renaud Bougueng Tchemeube, Jean-Baptiste Roland, Maryam Safi, Cale Plut, Jeff Ens, Submitted, IJCAI.

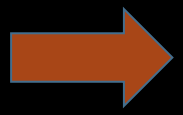


- Task 1: arrangement (adding 3 tracks to a 16 bars 4 tracks motif)
- Task 2: variation (of a 16 bars composition motif)
- Task 3: original composition (based on seeds of their choice)

METACREATION Lab for Creative Artificial Intelligence

Study Conclusions

- MMM-Cubase is a usable, computer-assisted co-creative interface for multitrack music composition.
- It provides for a decent user experience with good creativity support but lacks expressivity and control.
- The tool has decent acceptance and self-predicted future use.



Composers saw the system as a source of inspiration, with heavy editing of the output. They felt they maintained authorship over the final result.

Lessons and Challenges

- Like everywhere, ANN are making a foray.
- Audio generation is catching up with symbolic generation!
- Controllable factors need to be further explored (affective computing, …)
- Moving beyond style imitation (novelty/quality search,…)
- More needs to be done: cognitive modelling, agent learning, machine listening, ...
- Interaction design and user experience research needed.

Conclusion

- Musical Agent are progressing, but still lots of room for improvements.
  - We still need better models (transformers)
  - Embodiment and multi-modality
- Computer-assisted composition systems are ready to be deployed, but many more challenges to be addressed:
  - Interface design
  - Control and expressivity

www.Kadenze.com

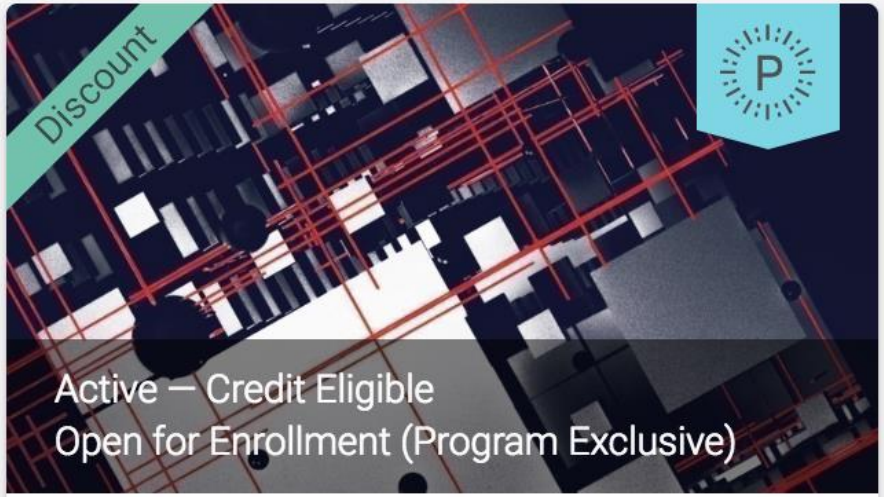
kadenze

Generative Art and Computational Creativity

Simon Fraser University

Philippe Pasquier with special guest Arne Eigenfeldt

VIEW



Advanced Generative Art and Computational Creativity

Simon Fraser University

Philippe Pasquier

VIEW

- Find us at: www.metacreation.net
- Watch the AI Music Creativity Conference AIMC (online + FREE) https://aimusiccreativity.org/
- 4th AI Music Creativity Conference (AIMC), Sussex University, Brighton, UK: https://aimc2023.pubpub.org/,
- 5th AI Music Creativity Conference (AIMC), Oxford, UK.
- Read the Special Issue on Music AI of the Journal for the Simulation of Music Creativity.

Social Sciences and Humanities Research Council of Canada

Conseil de recherches en sciences humaines du Canada

Canada

NSERC CRSNG

Canada Council for the Arts

Conseil des Arts du Canada

Live Coding with Euterpe

Christodoulos Benetatos

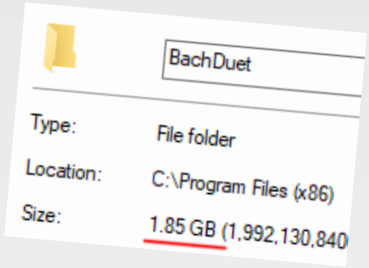
Euterpe: A Web Framework for Interactive Music Systems

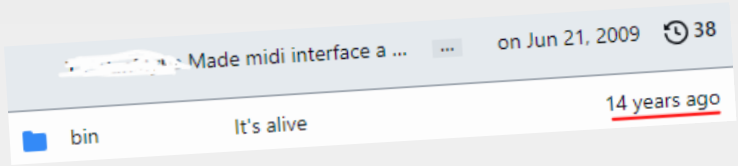
Yongyi Zhang*, Christodoulos Benetatos*, Zhiyao Duan

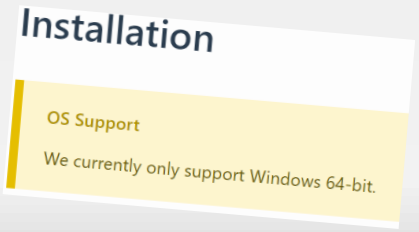
JAES 2023

Problem Statement

- Research **stops** at
  open sourcing the core algorithms
- Prototype systems that are **not easily accessible**
  - Large executable files
  - Unmaintained codebases
  - Platform dependent implementations
  - Complicated installation processes

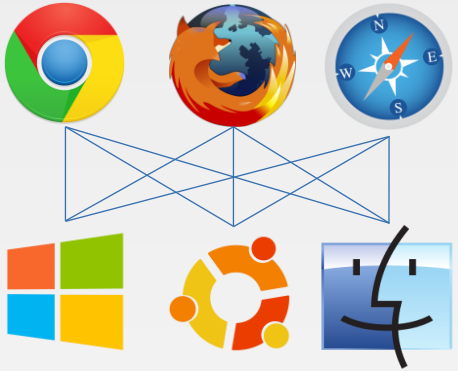






A Solution …

- Promote the development of **web** musical systems
  - Pros
    - Utilize the web's natural cross-platform compatibility
    - End-users are familiar with the browser environment
    - No installation required



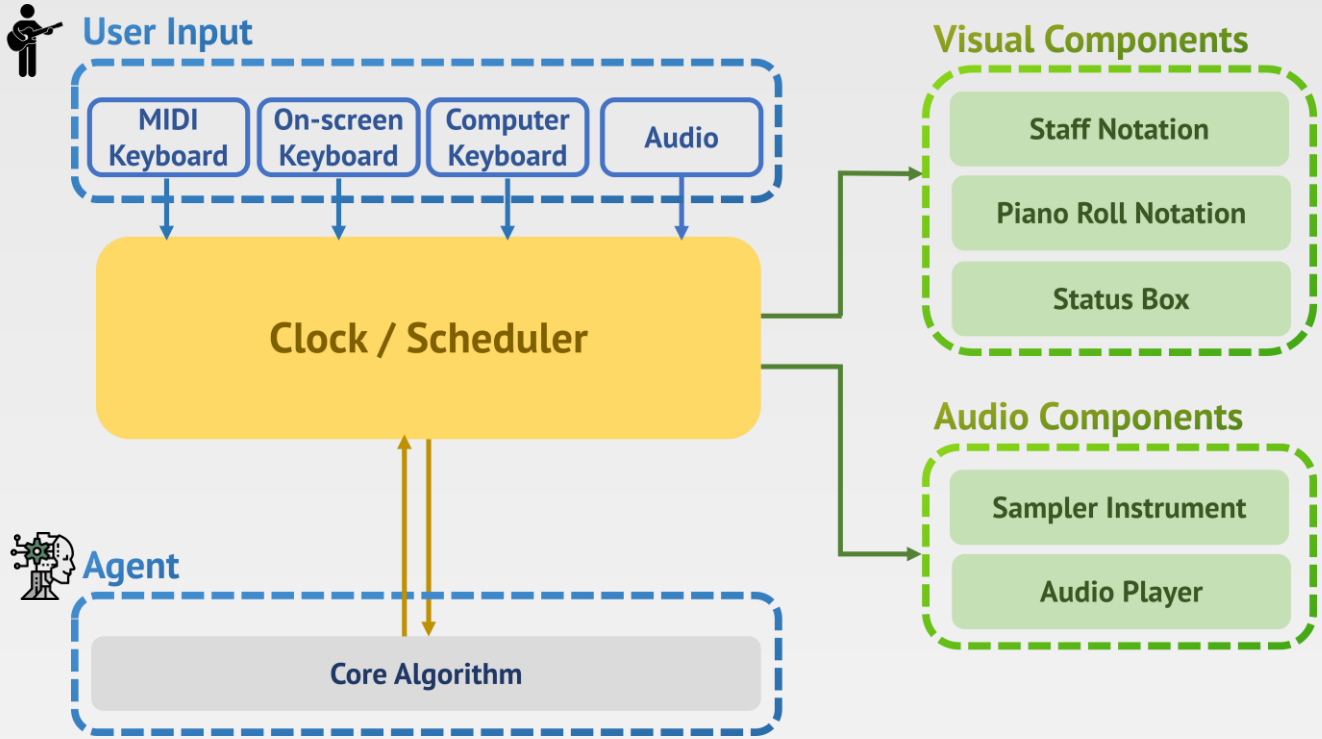
A Solution …

- Promote the development of **web** musical systems
  - Cons
    - Knowledge of web programming is required (JavaScript, CSS, HTML)

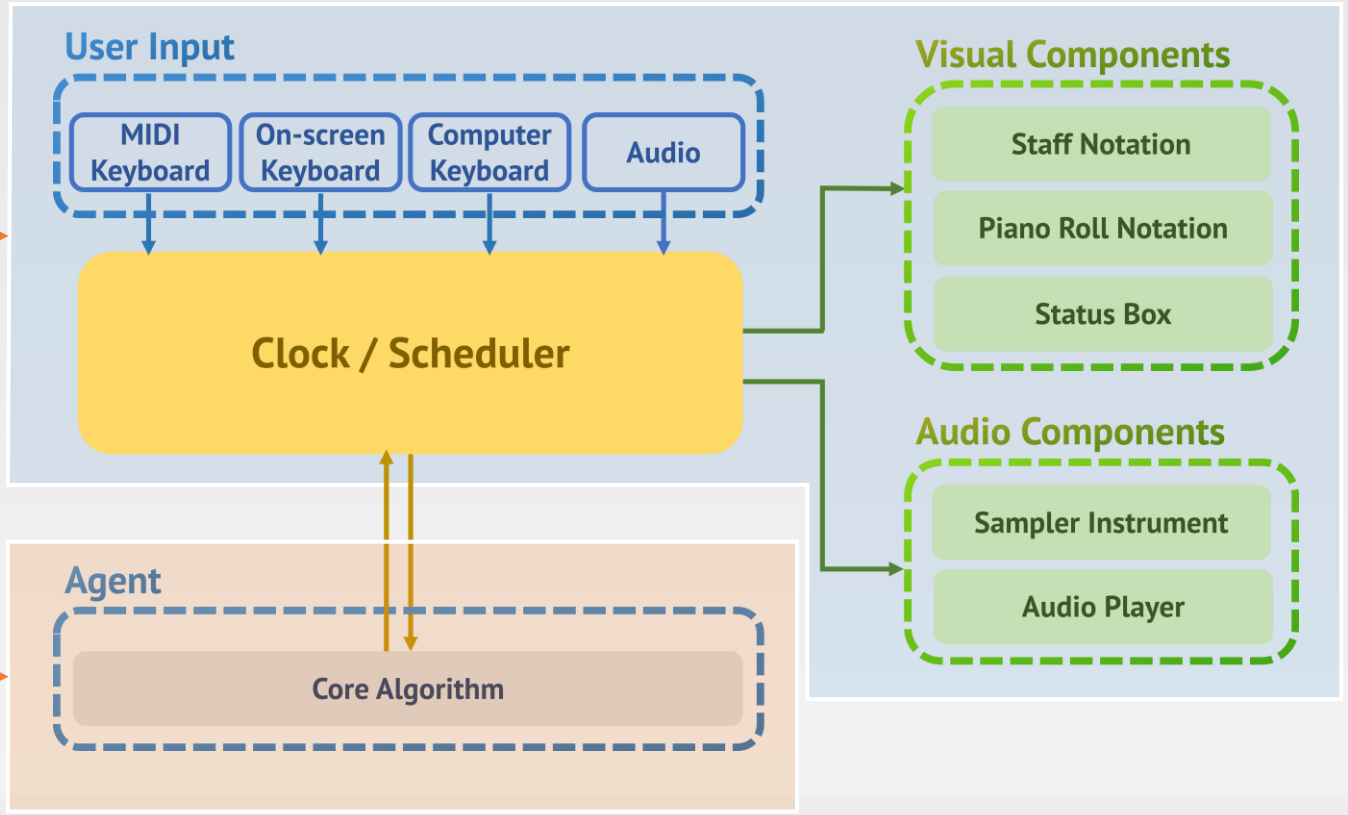
Euterpe's Goal

- **Alleviate challenges** associated with **web programming**
  - Offer ready-made submodules for common system components
  - Developers focus solely on their system's unique features

Generic IMS Architecture



Generic IMS Architecture

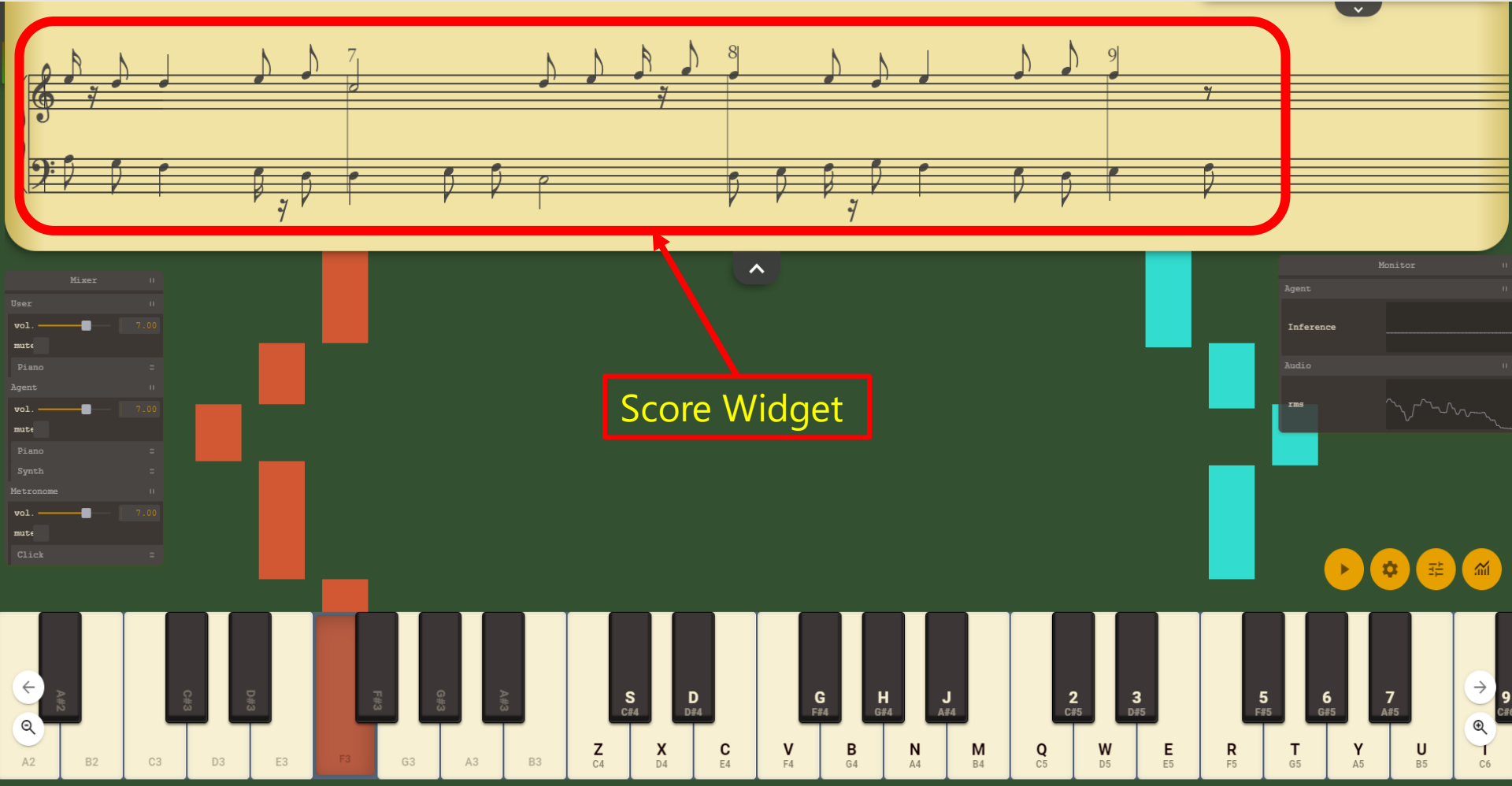


Design

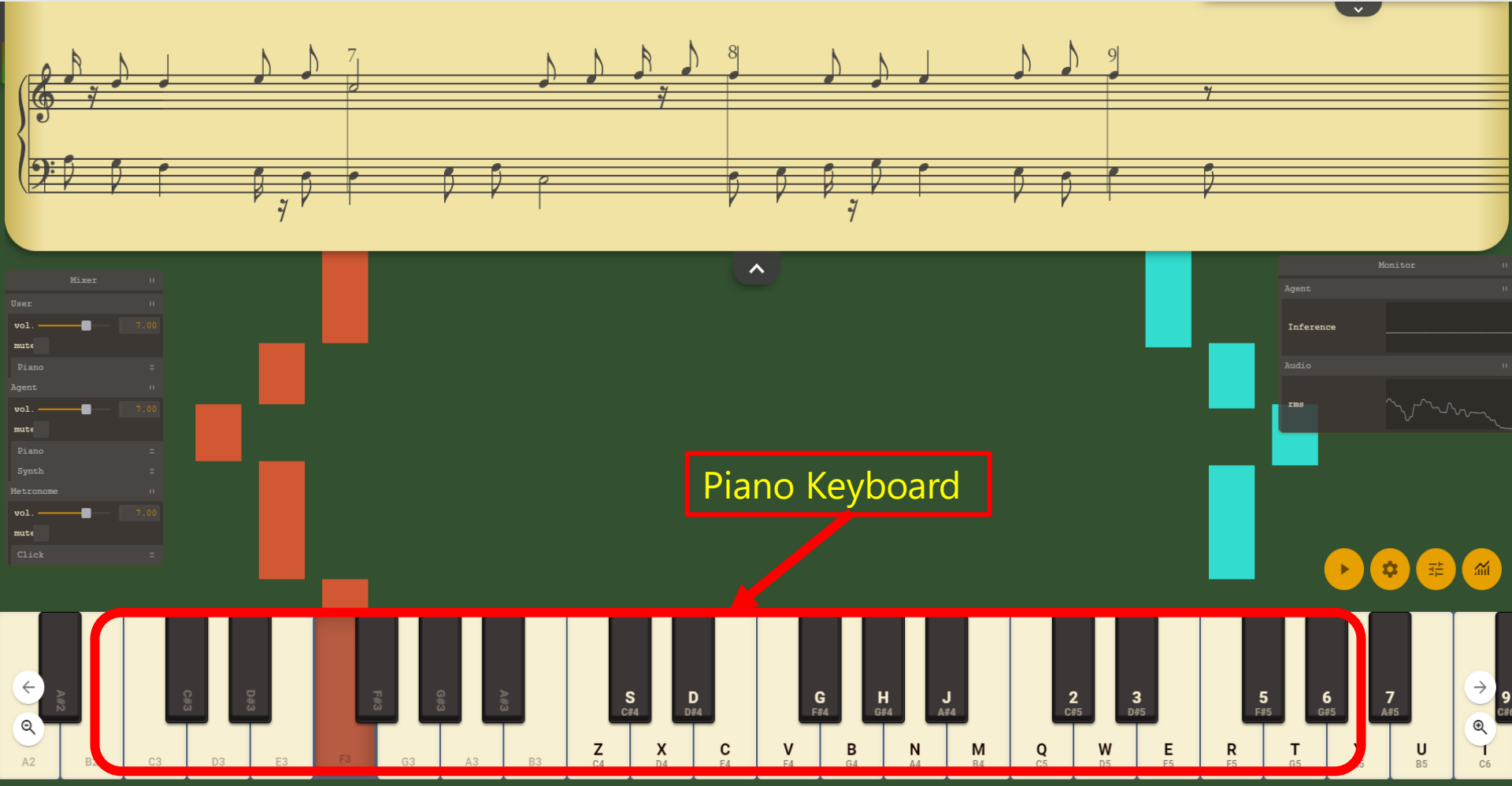
- Modular
  - **Separate** the **Agent code** from the peripheral components

- Configuration files
  - Allow app setup and **customization without** writing **JavaScript**

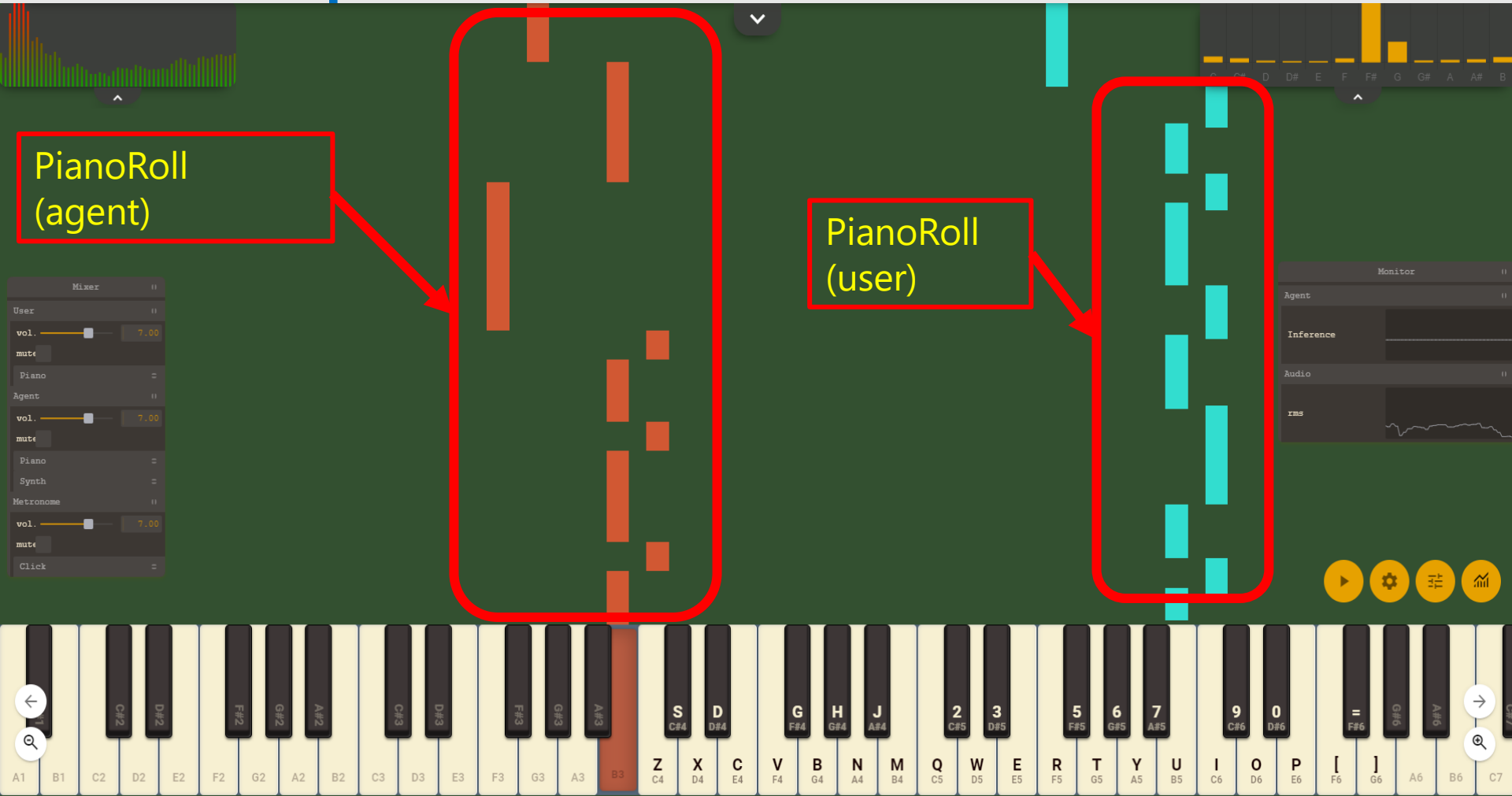
Visual Components



Visual Components



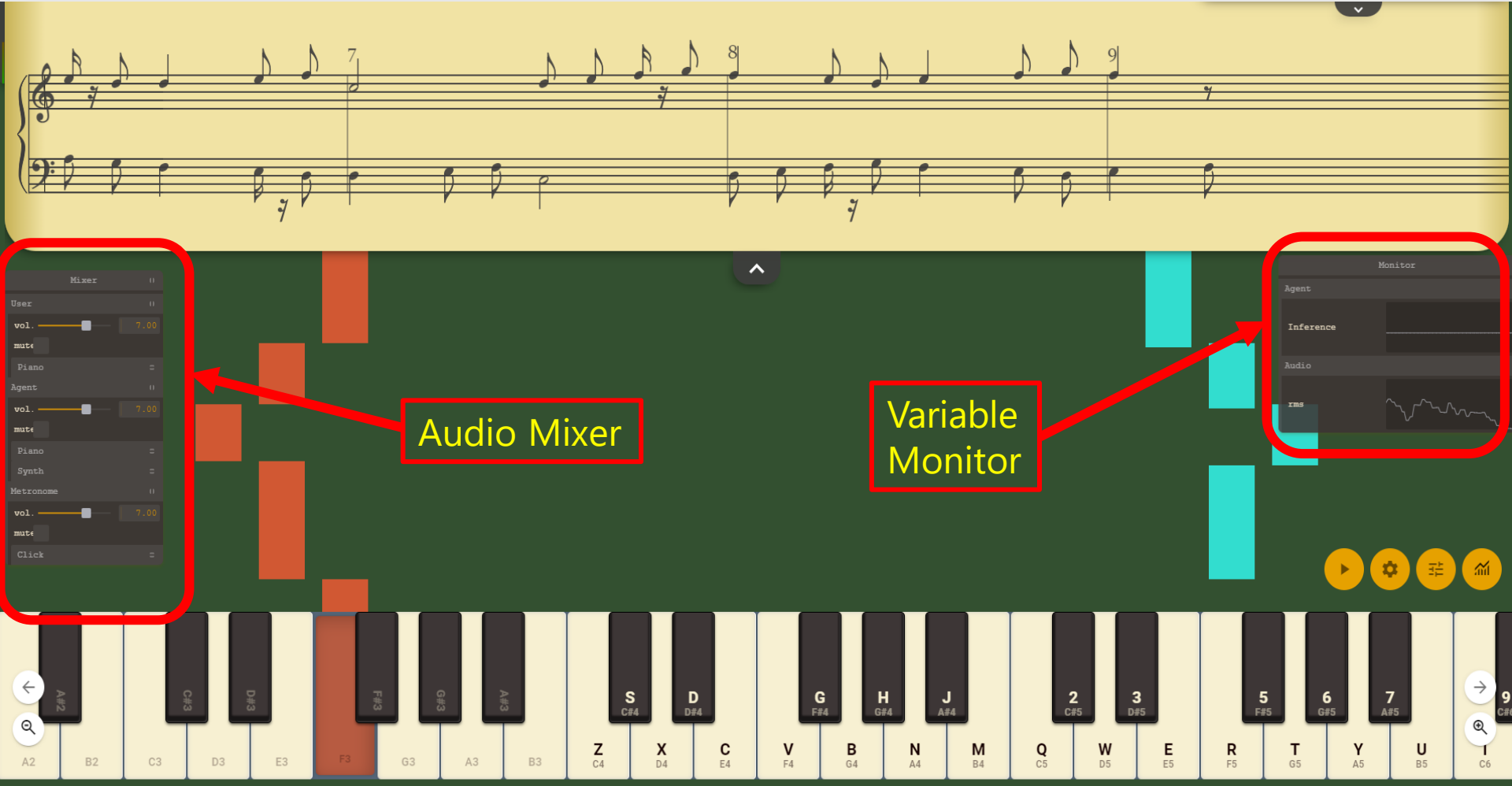
Visual Components



Configuration

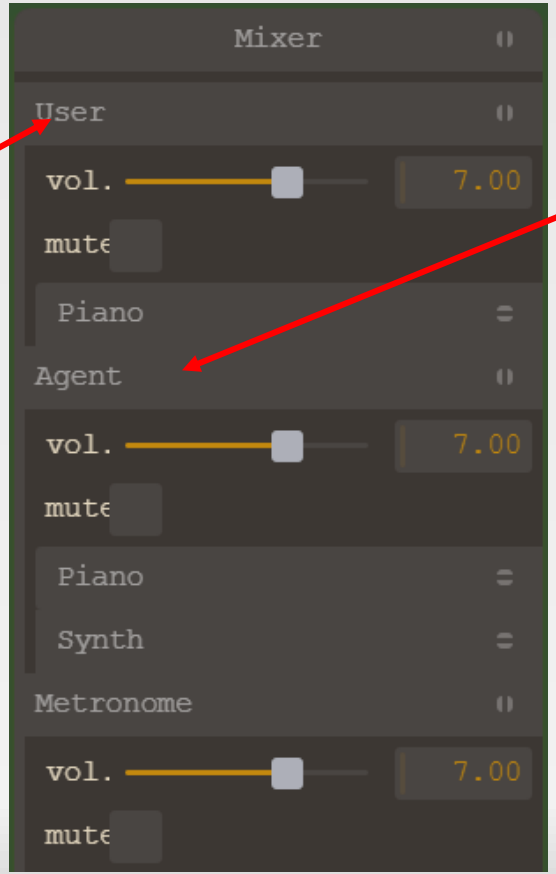
```
gui:
  score:
    status: true
  pianoRoll:
    status: true
    human: true
    agent: true
  keyboard:
    status: true
    octaveStart: 2
    octaveEnd: 6
```

Visual Components



Configuration

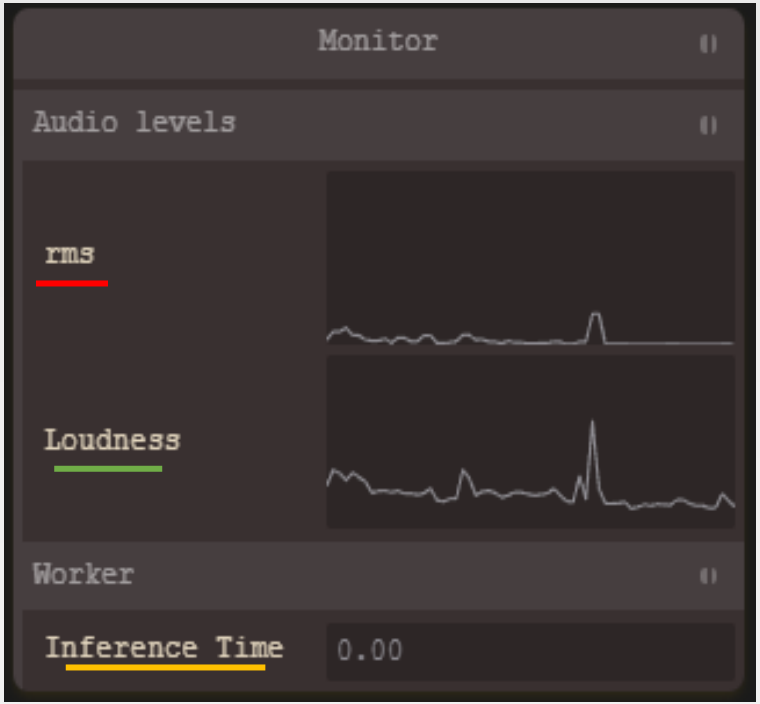
```yaml
players:
  human:
    label: 'User'
    mute: false
    volume: 5
    instruments:
      - id: "piano"
        label: "Piano"
        mute: false
        volume: 5
        default: true
```



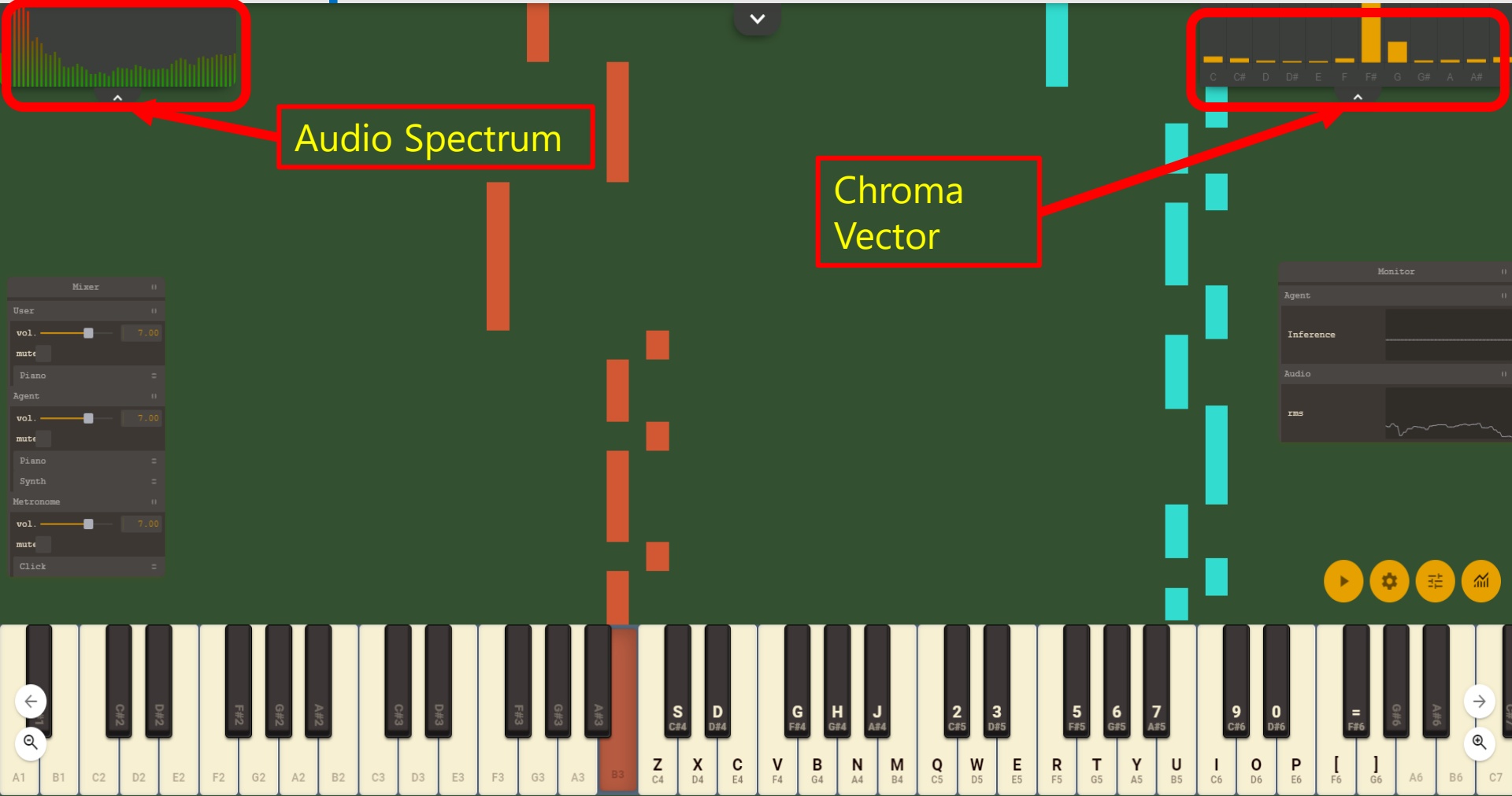

```yaml
players:
  agent:
    label: 'Agent'
    mute: false
    volume: 5
    instruments:
      - id: "piano"
        label: "Piano"
        mute: false
        volume: 7
        default: true
      - id: "synth"
        label: "Synth"
        mute: false
        volume: 8
```

Configuration

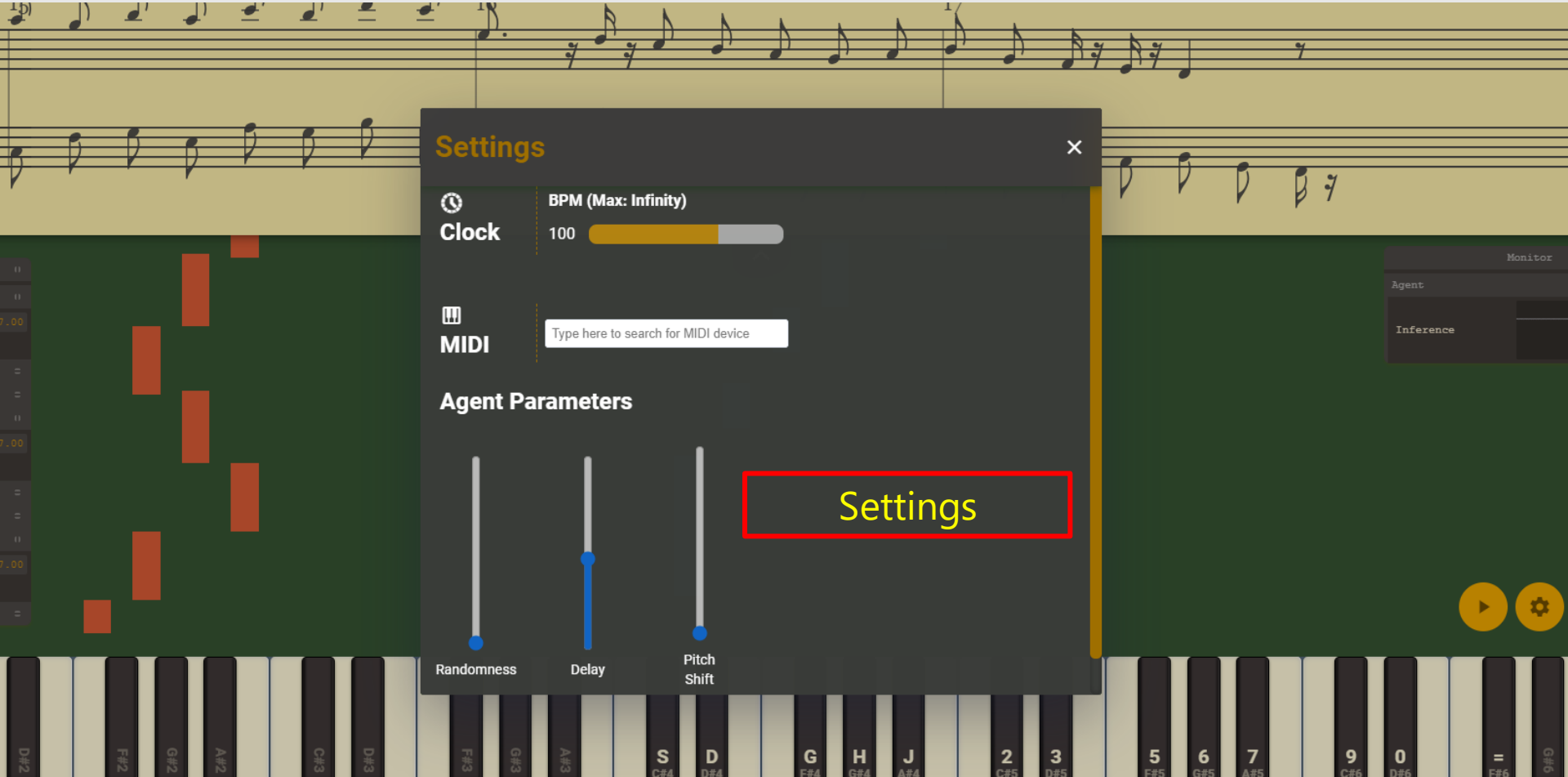
```yaml
monitor:
  title: "Monitor"
  structure:
    - label: "Audio levels"
      parameters:
        - id: 0 # id's must be unique
          label: "rms" # choose any name
          interval: 50 # in ms
          graph: true
          min: 0
          max: 0.2
        - id: 1
          label: "Loudness"
          interval: 50 # in ms
          graph: true
          min: 0
          max: 100
    - label: "Worker"
      parameters:
        - id: 2
          label: "Inference Time"
          interval: 100 # in ms
          graph: false
          min: 0
          max: 30
```



Visual Components



Visual Components

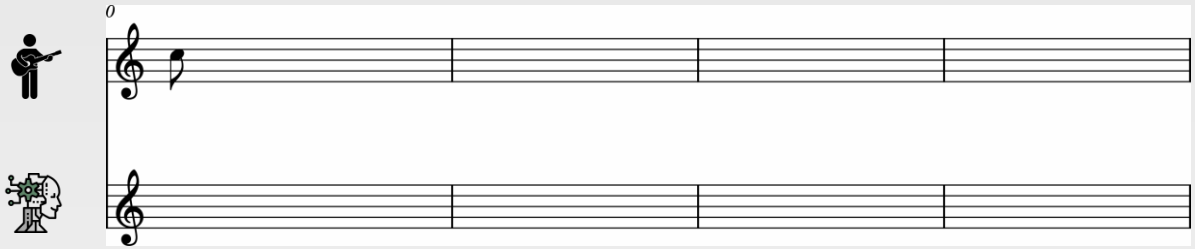


Configuration

```yaml
settingsModal:
  sliders:
    - id: 1
      label: "Randomness"
      value: 0
      min: 0
      max: 10        You, 3
    - id : 2
      label: "Delay"
      value: 8
      min : 1
      max : 16
    - id : 3
      label: "Pitch Shift"
      value: 0
      min: 0
      max: 24
```

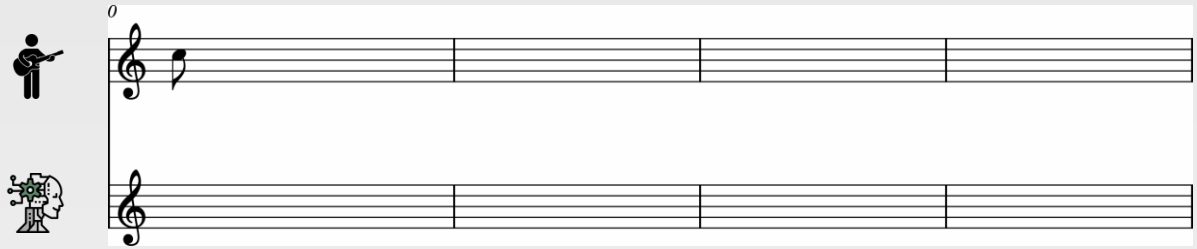
Paradigms of Music Interaction

- Call & Response

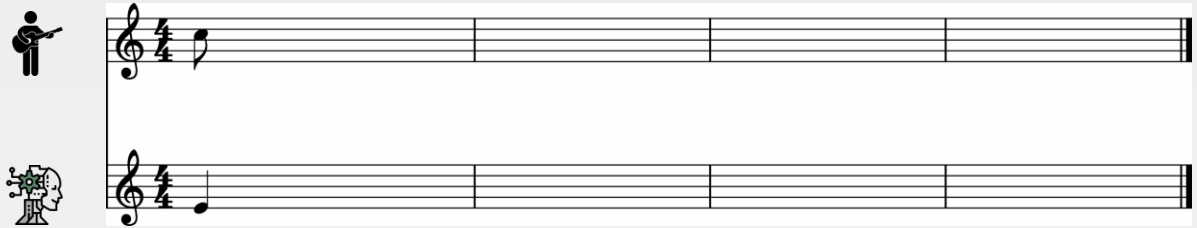


Paradigms of Music Interaction

- Call & Response

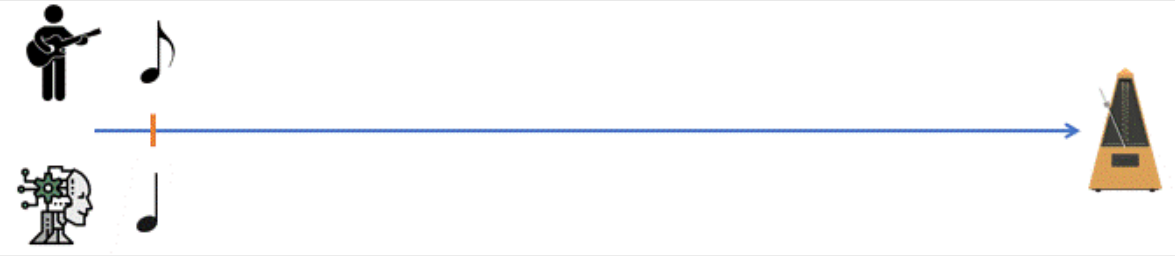


- Simultaneous



Paradigms of Music Interaction

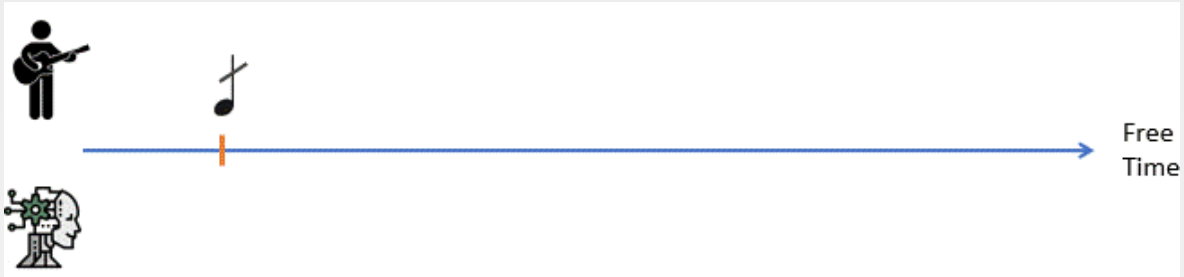
- Grid-based



Paradigms of Music Interaction

- Grid-based

- Event-based



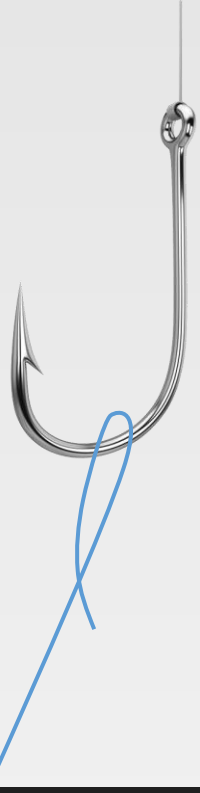
Configuration

```yaml
title: "Euterpe"
interactionMode:
  noteMode: true
  audioMode: false

noteModeSettings:
  eventBased:
    status: false
  gridBased:
    status: false
```

```yaml
audioModeSettings:
  windowSize: 1024
  hopSize: 512

clockSettings:
  # ---- OPTION 1 --- #
  # 16th-note grid on 4/4
  ticksPerBeat: 4
  timeSignature:
    numerator: 4
    denominator: 4
  defaultBPM: 100
  # ---- OPTION 2 --- #
  clockPeriod: null
```

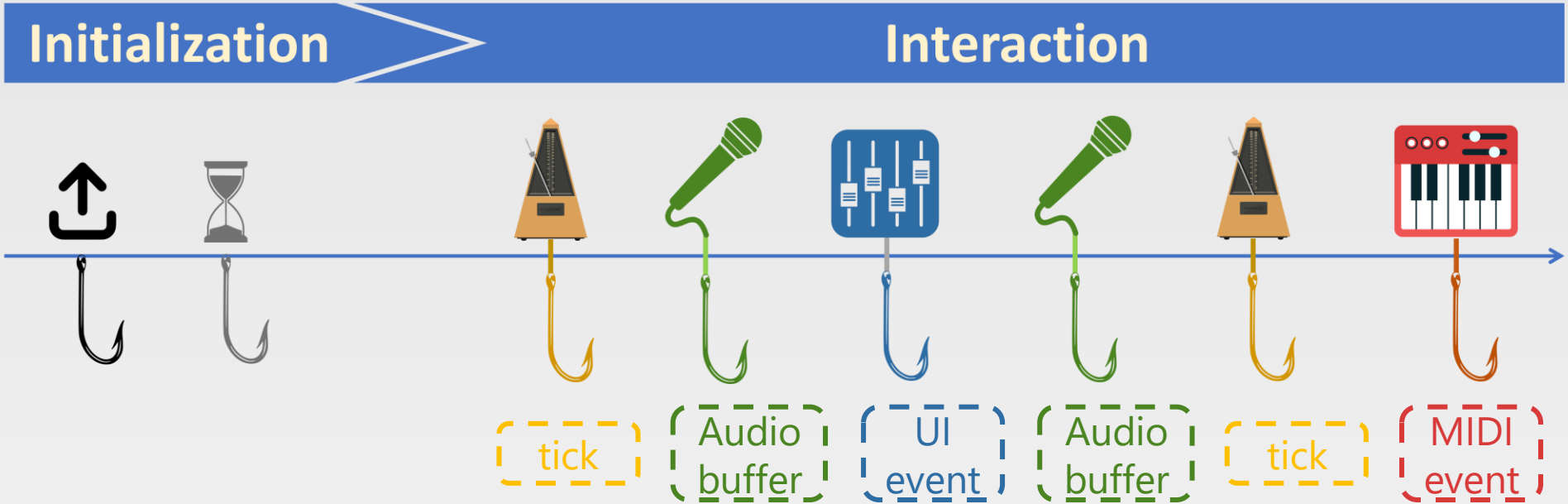
Agent

- Provides 6 hook functions
  - **Empty** functions to be filled.
    - Invoked **automatically** at specific events or stages within the interaction
    - Can be activated/deactivated from the configuration file



```
function hook(event){
    // your code
}
```

Euterpe Lifecycle



Agent - Hooks

- loadExternalFiles()
  - Load external resources useful for the Agent

- loadAlgorithm()
  - Core algorithm initialization
  - Checkpoint fetching
  - NN model loading
  - warmup NN

Agent - Hooks

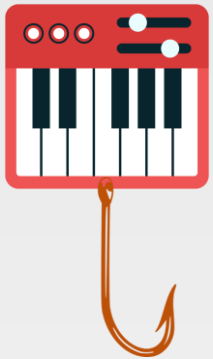
- updateParameter(id, value)
  - Invoked when the user interacts with the GUI (buttons, sliders etc.)
  - The Agent's hyper-parameters are updated

Agent - Hooks

- processClockEvent(tick)
  - Invoked periodically based on the Clock's "tick"
  - Used on a time-grid based interaction

Agent - Hooks

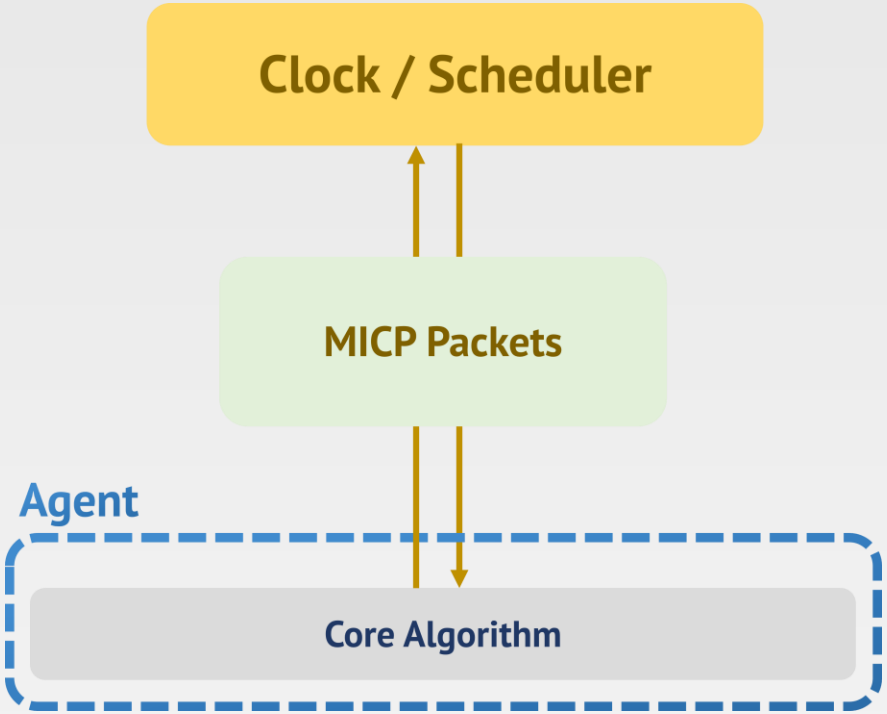
- `processNoteEvent(event)`
  - Invoked when a MIDI note is received
  - Used in “event-based” mode



Agent - Hooks

- `processAudioBuffer(buffer)`
  - Invoked when a new audio buffer is available
  - Every hopSize samples

Music Interaction Communication Protocol - MICP



MICP – NoteEvent

- **player** : Agent or User
- **instrument** : Which sampler instrument to use for playback
- **device** : The user's input device (i.e MIDI keyboard)
- **type** : Note_On, Note_Off or Note_Hold
- **name/midi/chroma** : Info about the note (i.e C4, 60, 0)
- **channel/velocity** : Midi specific info
- **createdAt** (tick, seconds) : When was this note created/generated (timestamp)
- **playAfter** (tick, seconds) : Play the note with a delay
- **duration** : The duration of the note (optional)

Coding Session

Online Guide :

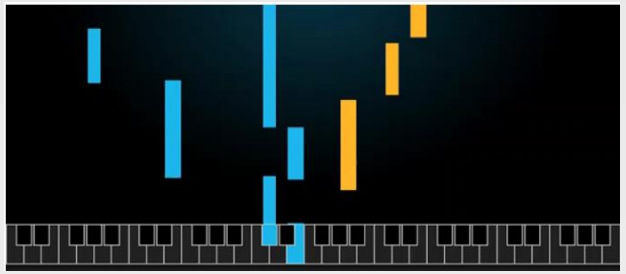
https://xribene.github.io/

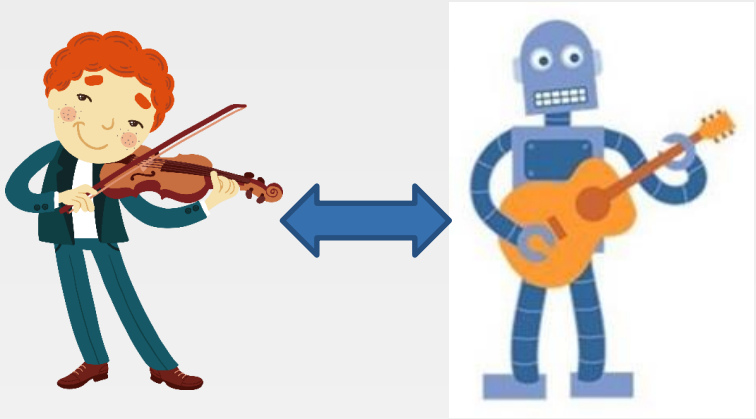
Challenges and Research Directions

Zhiyao Duan

Moving from Symbolic to Audio

- Audio signals offer much more musical expressiveness, but
  - Most existing musical agents work in the symbolic domain
  - Agents work in the audio domain often only analyze low-level features (e.g., Voyager, LL) or monophonic audio (e.g., GenJam, Omax-Ofon)
- Need more robust music analysis algorithms
  - Real-time beat tracking and rhythm analysis
    - Beat tracking for percussion-less music input, e.g., singing voice [Heydari et al., SingNet, 2023]
  - Fine-grained polyphonic pitch tracking to analyze pitch fluctuations e.g., vibrato
  - Robust score following to performance mistakes, improvisation, and structural changes
- Need expressive audio synthesis and coordinating it with human performance on timing, dynamics, and timbre

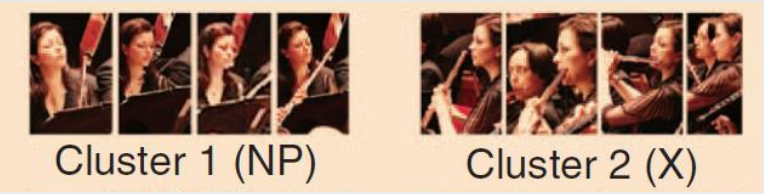




Incorporating the Visual Modality

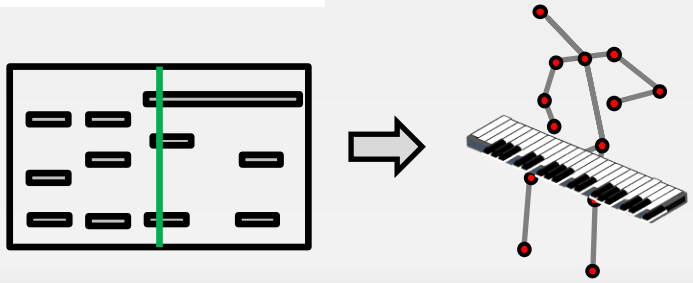
- Music performance is audiovisual in nature
  - Visual performance is important in musical expression
  - Musicians use visual cues to coordinate on timing, dynamics and intention
- Need algorithms to analyze various aspects of visual performance
  - Instrument recognition, body movement, facial expression, fingering motion
  - Audiovisual association and joint analysis
- Need real-time expressive visual rendering and coordinating it with human performance
- 2019 ISMIR Tutorial on Audiovisual Processing Processing

- Shoutout to Music Session #6 AI Pianist Performance by Juhan Nam's team at KAIST



[Bazzica et al., 2016]

[Li et al., 2019]



[Li et al., 2018]

Personalization

- Existing CAMM systems are “standardized”
- Can we make them adapt to user behaviors, habits, and preferences?
  - Learning from rehearsal: Chris Raphael’s Music Plus One system [Raphael, 2002]
- Some ideas
  - Provide more options in preference settings
  - Learn from interaction history
  - Provide feedback to users after interaction
  - Make suggestions on improvisation

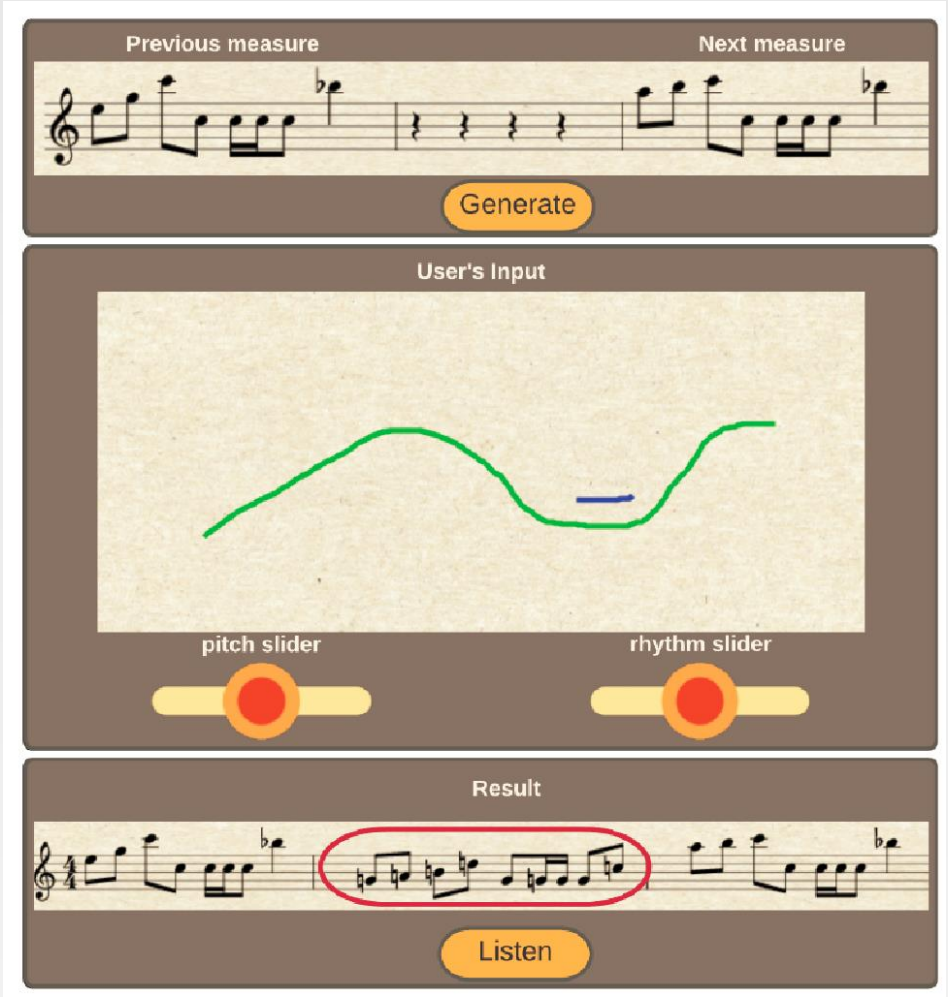
Chris Raphael’s Music Plus One system demo

Improving Accessibility

- Music theory is hard
- Music instruments are not easy to learn

- Many existing CAMM systems seem to be even harder to interact with
  - Require music background + software literacy

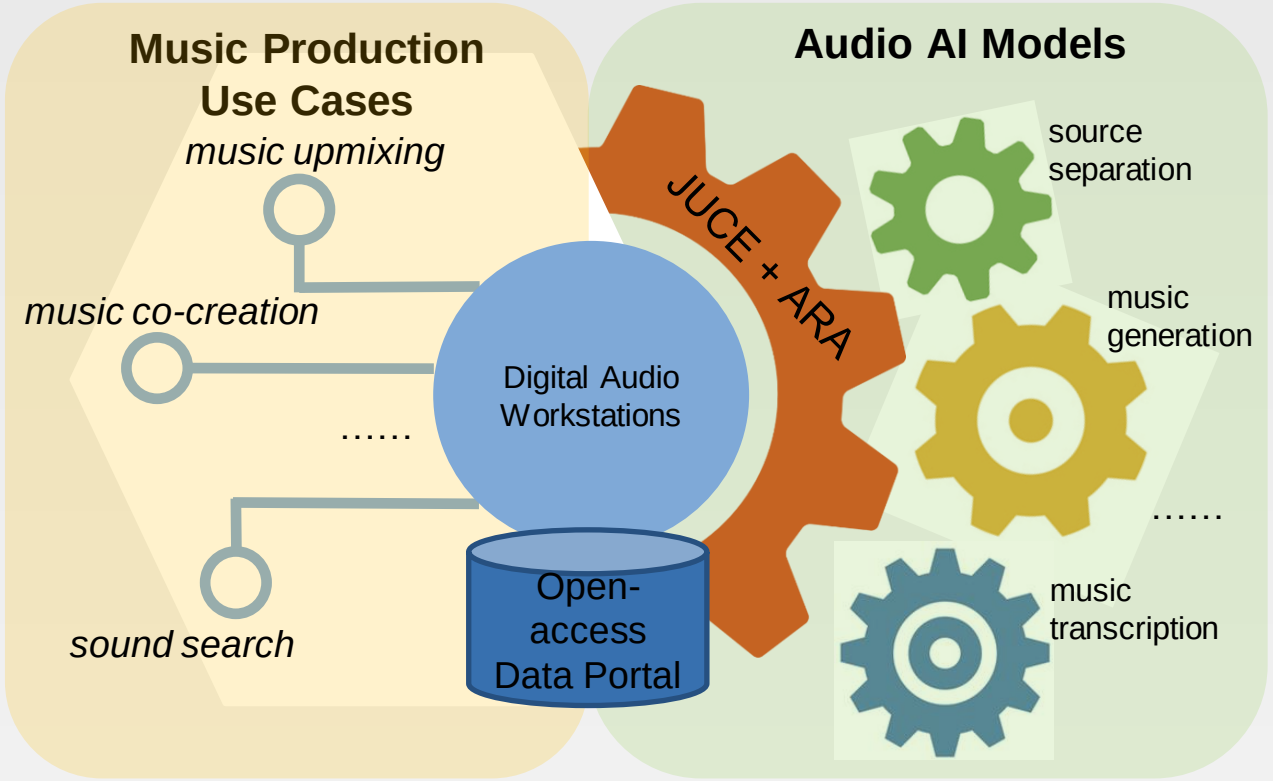
- Lower the barrier to entry?
  - Use music AI to augment users' capabilities



Draw&Listen [Benetatos & Duan, 2022]

Bridging Music AI and Music Production

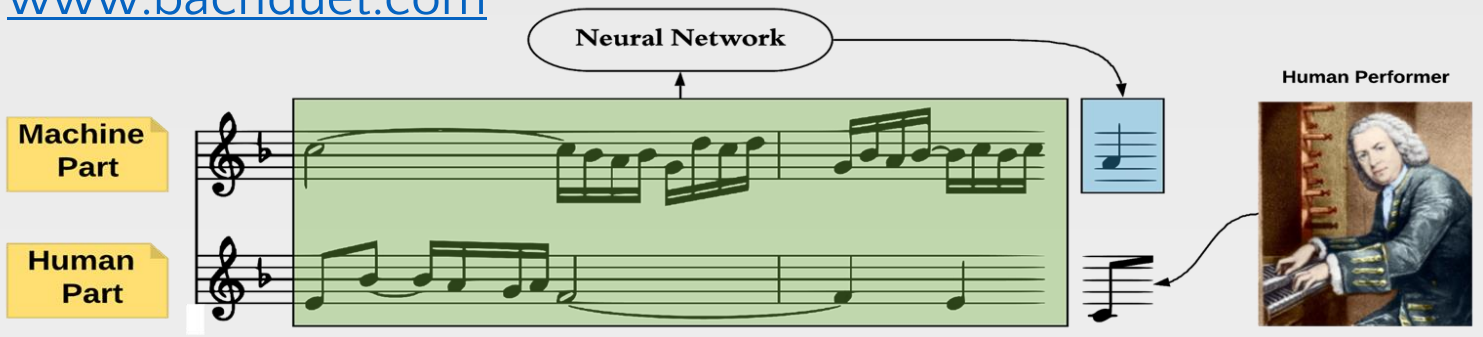
- Music AI models are released on a weekly basis, but most of them only stay in the labs
- Tools for music making (e.g., notation software, DAWs) have limited AI functionalities
- Idea: Build (open-source) tools to bridge the gap
  - Euterpe [Zang et al., 2023]
  - Hosted, Asynchronous, Remote Processing (HARP) for audio AI plug-ins [Garcia et al., 2023]
  - Commercial software: Neutone



TEAMuP NSF project (2022-2026)

Developing CAMM for Education

- Music education
  - Practicing counterpoint improvisation with BachDuet [Benetatos et al., 2020] www.bachduet.com



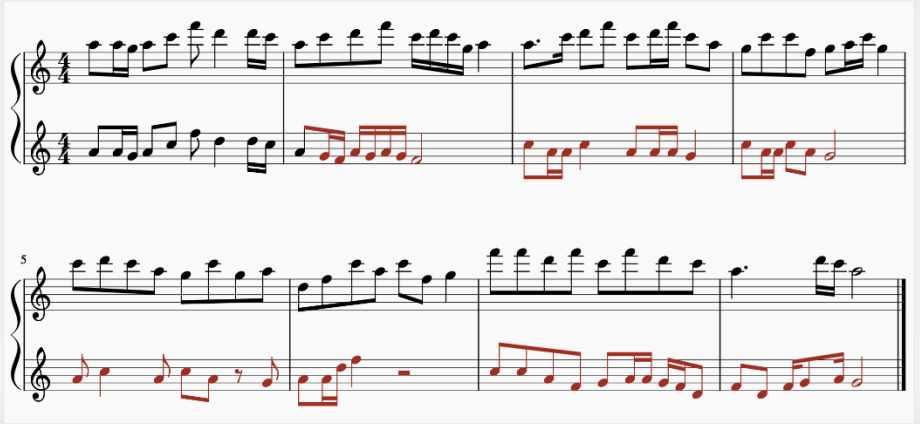
- Computing education
  - EarSketch by Jason Freeman at Georgia Tech: "EarSketch helps students learn to code in Python or JavaScript through manipulating loops, composing beats, remixing sounds, and applying effects to a multi-track digital audio workstation"
  - TunePad by Michael Horn at Northwestern: "TunePad is a free online platform for creating music with the Python programming language."

EarSketch

TUNEPAD

Diversifying Music Styles

- Existing CAMM systems focus on Western music styles
- One challenge for diversifying styles is the lack of training data
- For example: Counterpoint composition in Chinese folk music style is an important direction in Chinese music composition
  - One idea for automating this task is to use inverse reinforcement learning to fuse Western counterpoint with Chinese folk styles



Demo for “When counterpoint meets Chinese folk melodies“ [Jiang et al., 2020]

Incorporating Large Language Models

- LLMs (e.g., ChatGPT) represent the most significant AI advances in recent years
- Music generation with LLMs
  - Google's MusicLM: Generating high-fidelity music from text input, [Agostinelli et al., 2023]
  - Meta's MusicGen: Controllable text-to-music generator, [Copet, et al., 2023]
  - Tutorial #3: Transformer-based Symbolic Music Generation: Fundamentals to Advanced Concepts, Stylistic Considerations, Conditioning Mechanisms and Large Language Models, by Berker Banar, Pedro Sarmento, and Sara Adkins
- These models allow people to use natural language to guide the music generation process
- Look forward to more interaction mechanisms and user control flexibilities


Thank You!

References

- Agostinelli A, Denk TI, Borsos Z, Engel J, Verzetti M, Caillon A, Huang Q, Jansen A, Roberts A, Tagliasacchi M, Sharifi M. Musiclm: Generating music from text. arXiv preprint arXiv:2301.11325. 2023 Jan 26.
- Christodoulos Benetatos and Zhiyao Duan, Draw and listen! A sketch-based system for music inpainting, Transactions of the International Society for Music Information Retrieval, vol. 29, pp. 2288-2292, 2022.
- Christodoulos Benetatos, Joseph VanderStel, and Zhiyao Duan, BachDuet: A deep learning system for human-machine counterpoint improvisation, in Proc. International Conference on New Interfaces for Musical Expression (NIME), 2020, pp. 635-640.
- Copet J, Kreuk F, Gat I, Remez T, Kant D, Synnaeve G, Adi Y, Défossez A. Simple and Controllable Music Generation. arXiv preprint arXiv:2306.05284. 2023 Jun 8.
- Freeman J., Magerko B., Edwards D., Mcklin T., Lee T., Moore R. (2019). "EarSketch: engaging broad populations in computing through music," in Communications of the ACM.
- Hugo Flores Garcia, Christodoulos Benetatos, Patrick O'Reilly, Aldo Aguilar, Zhiyao Duan and Bryan Pardo, “HARP: Bringing deep learning to the DAW with hosted, asynchronous, remote processing, NeurIPS 2023 Workshop on Machine Learning for Creativity and Design.
- Bochen Li, Akira Maezawa, and Zhiyao Duan, Skeleton plays piano: online generation of pianist body movements from MIDI performance, in Proc. International Society for Music Information Retrieval Conference (ISMIR), 2018, pp. 218-224.

References

- Bochen Li, Karthik Denish, Chenliang Xu, Gaurav Sharma, and Zhiyao Duan, Online audio-visual source association for chamber music performances, Transactions of the International Society for Music Information Retrieval, vol. 2, no. 1, pp. 29–42, 2019.
- McDermott, J., Gifford, T., Bouwer, A., & Wagy, M. (2013a). Should music interaction be easy? In S. Holland, K. Wilkie, P. Mulholland, & A. Seago (Eds.), Music and human computer interaction (pp. 29–48). London: Springer.
- Mojtaba Heydari, Ju-Chiang Wang, and Zhiyao Duan, SingNet: a real-time singing voice beat and downbeat tracking system, in Proc. IEEE Internationsl Conference on Acoustics, Speech and Signal Processing (ICASSP), 2023.
- Nan Jiang, Sheng Jin‡, Zhiyao Duan, and Changshui Zhang, When counterpoint meets Chinese folk melodies, accepted by The Thirty-fourth Conference on Neural Information Processing Systems (NeurIPS), 2020.
- Christopher Raphael, C. 2002. A bayesian network for real-time musical accompaniment. In Dietterich, T.; Becker, S.; and Ghahramani, Z., eds., Advances in Neural Information Processing Systems, NIPS 14. MIT Press.
- Yongyi Zang*, Christodoulos Benetatos*, and Zhiyao Duan, Euterpe: A web framework for music interaction and creation, Journal of Audio Engineering Society, 2023. (* equal contribution)