

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

ISMIR 2023 Tutorial

Christodoulos Benetatos †
cbenetat@ur.rochester.edu

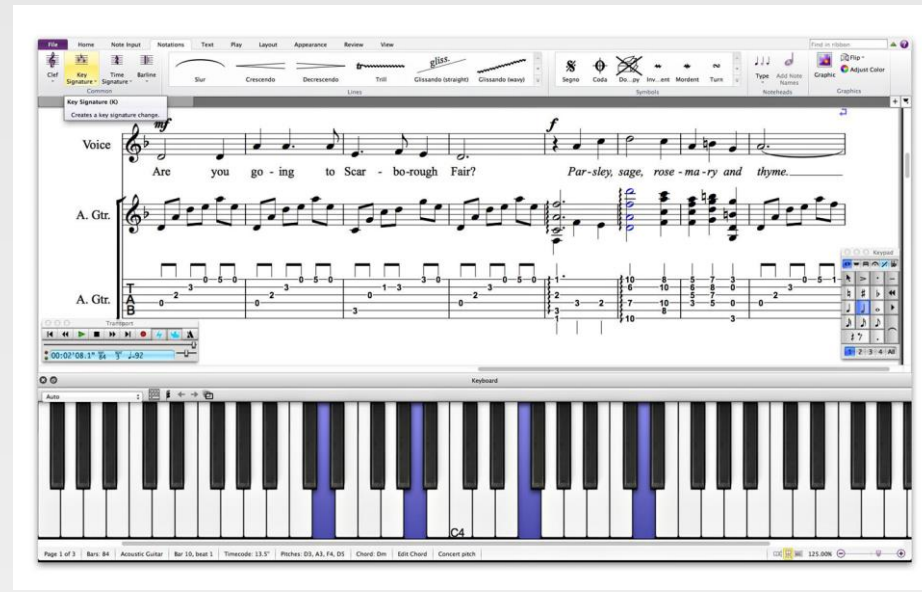
Philippe Pasquier ‡
pasquier@sfu.ca

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



Taryn Southern
Electronic Dance Music

Bach by Design

Computer Composed Music

David Cope -- Experiments in Musical Intelligence



CRC 2184

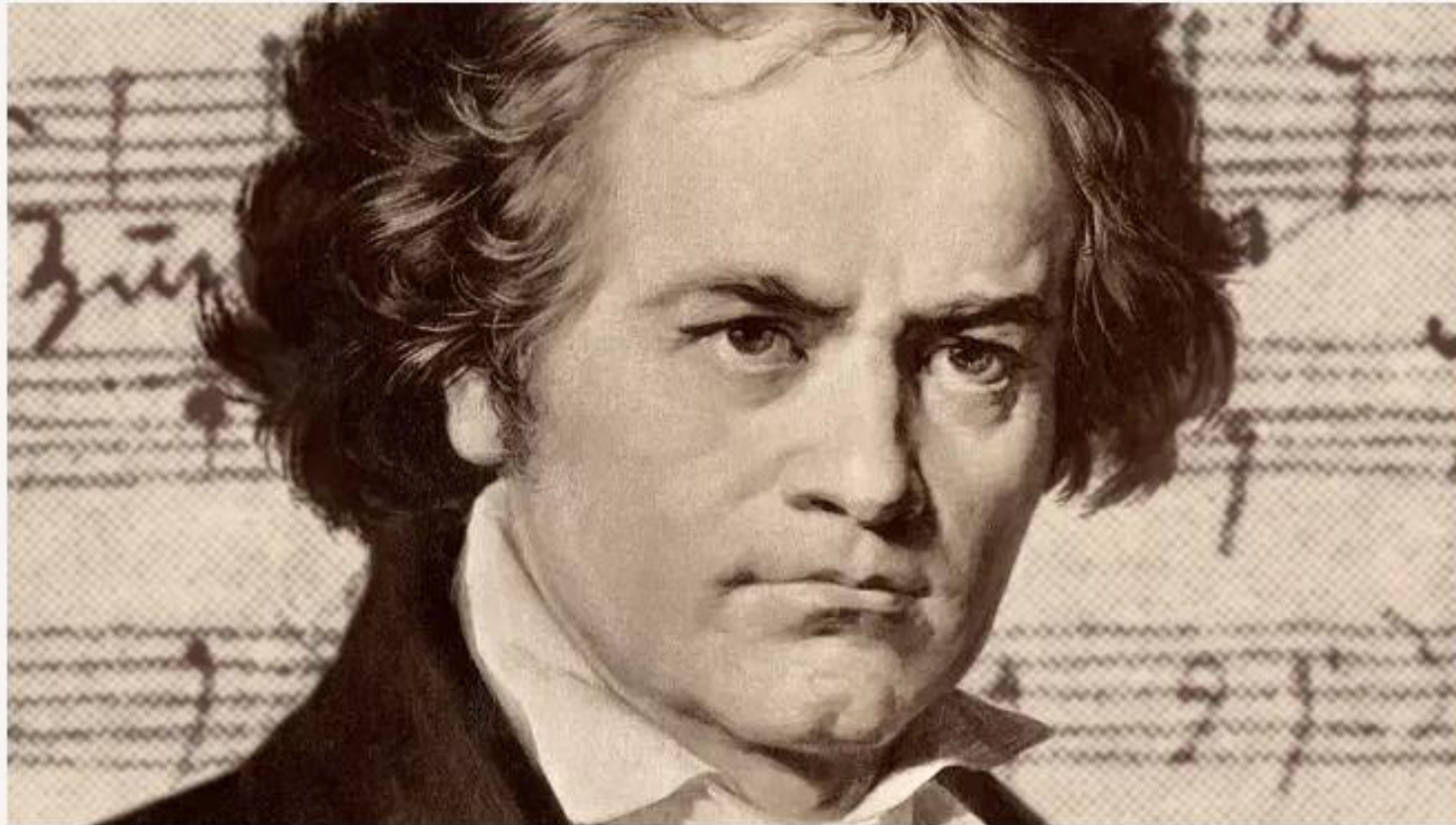
> LOAD 'BACH-FUGUES)
 > PATTERN-MATCH 'BACH-FUGUES)
 > COMPOSE 'NEW-BACH-FUGUES)

Allegro

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

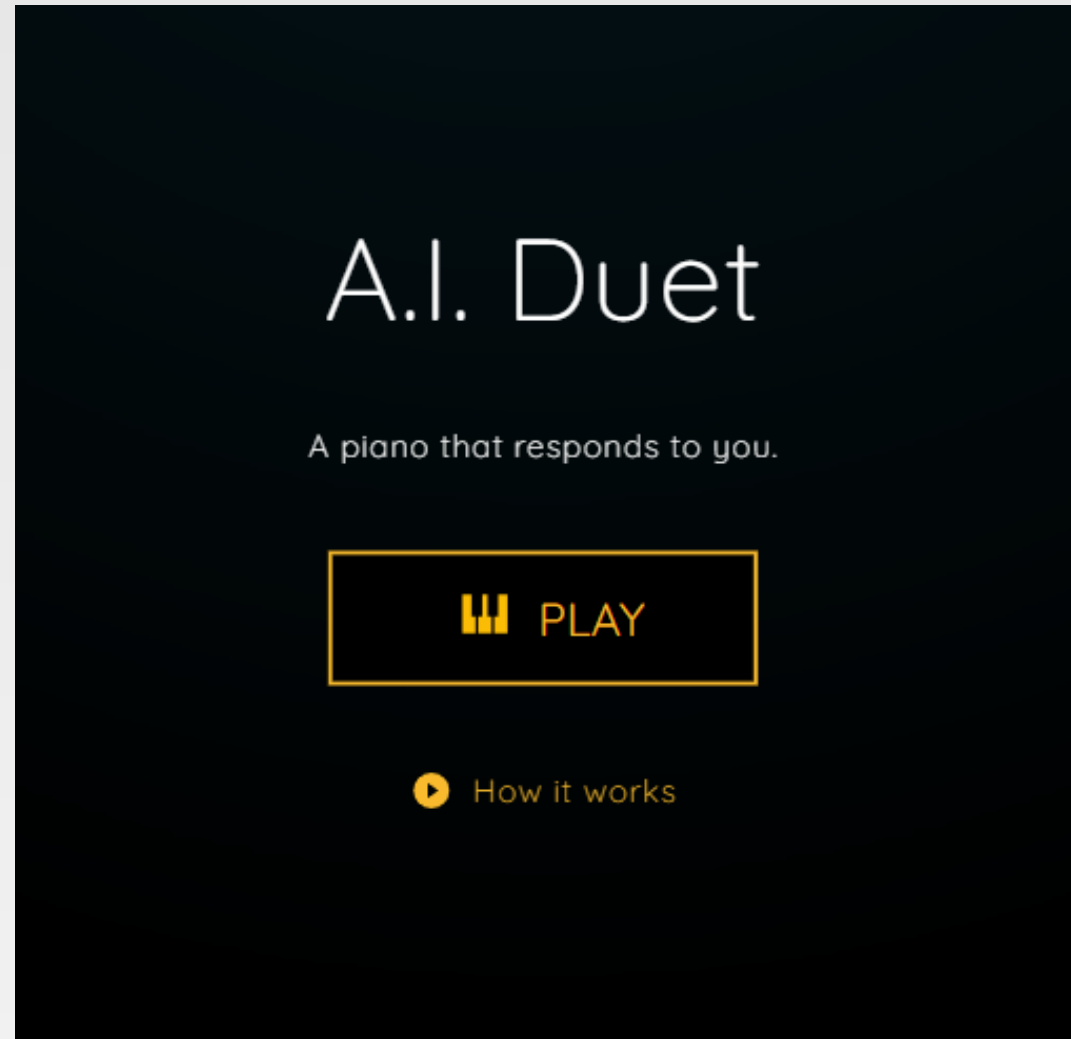
<https://www.classicfm.com/composers/beethoven/unfinished-tenth-symphony-completed-by-artificial-intelligence/>

Hello World! Let the AI Song Contest 2023 begin.

November 4 in A Coruña, Galicia (Spain)

The AI Song Contest is an international competition showcasing the creative potential of human–AI co-creativity in the songwriting process. Teams consisting of musicians, researchers, music producers, data scientists, developers – and anyone else interested in the combination of music and artificial intelligence (AI) – collaborate to create a song with AI as a creative partner.

<https://www.aisongcontest.com/>

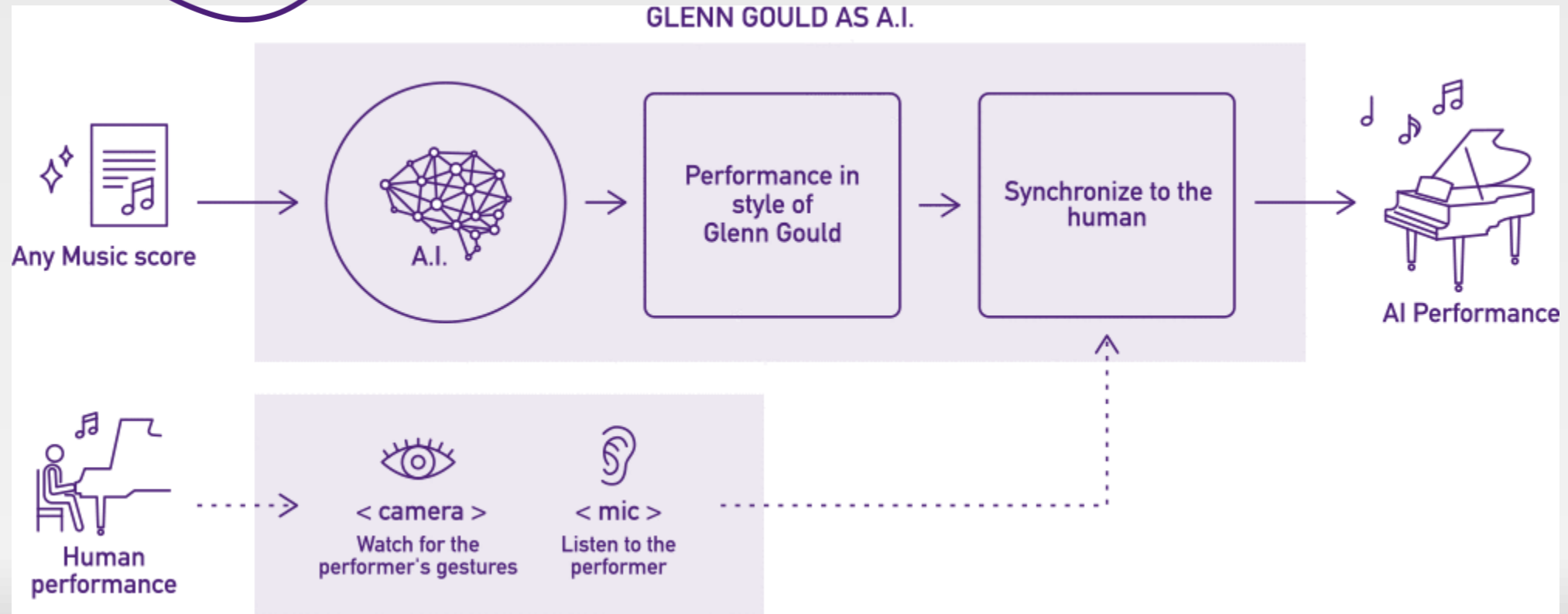


<https://experiments.withgoogle.com/ai/ai-duet/view/>

7 SEP. 2019

GLENN GOULD AS A.I.

YAMAHA's Music with AI Project

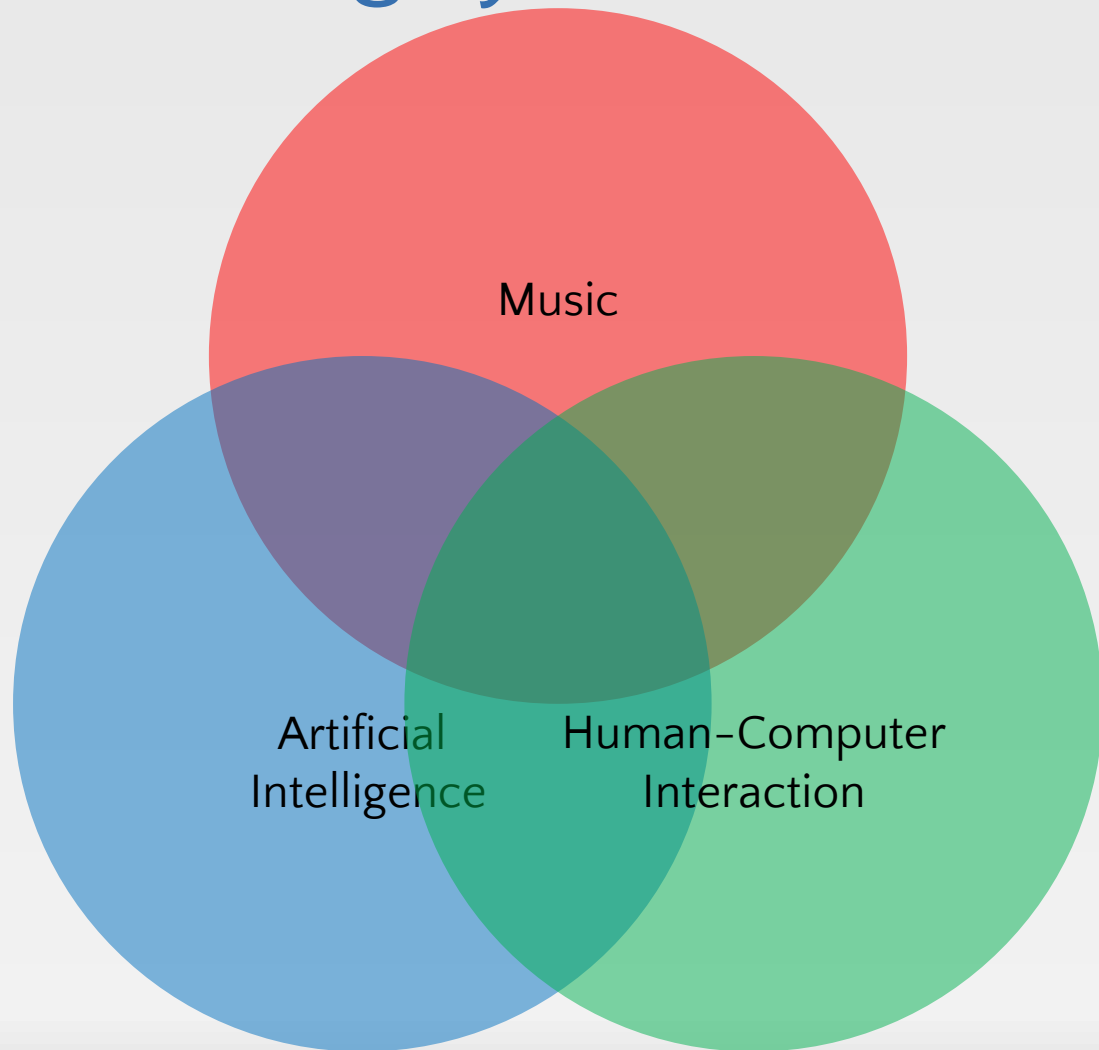


https://www.yamaha.com/en/about/ai/dear_glenn/



<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

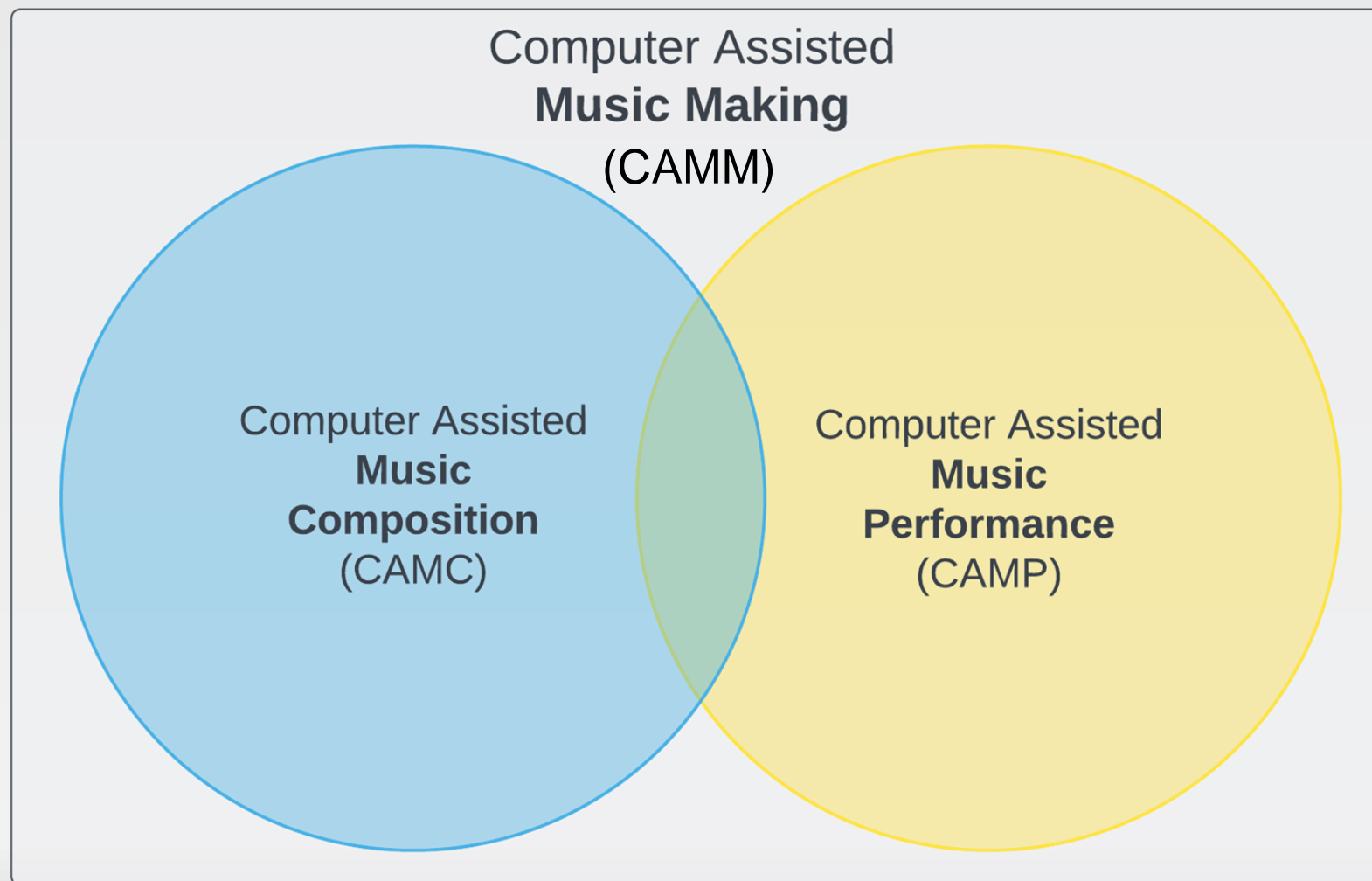
Music Education

- Composition
- Arrangement
- Improvisation
- Ensemble performance

Music Entertainment

- Novel content
- Various formats
- Anytime/anywhere

Two Categories

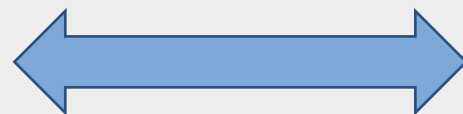


Computer-Assisted Music Making (Camm) systems · · Interactive Music Systems

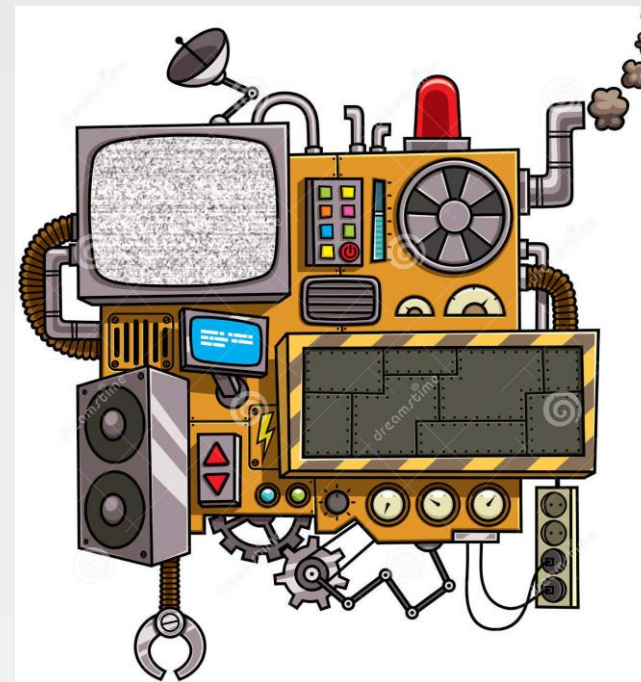


Human

Exchange of
musical
information



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



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

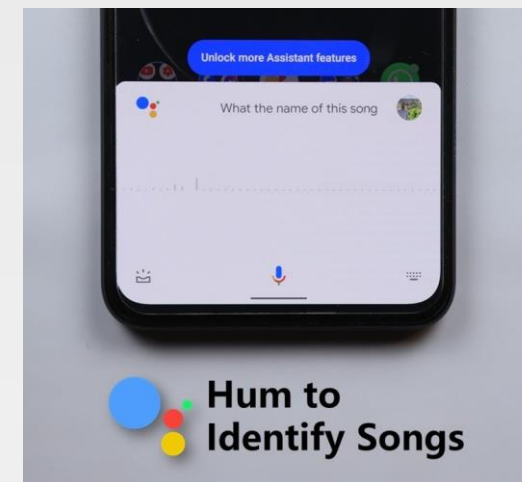
Are they interactive music systems?



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., & Wagyu, 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

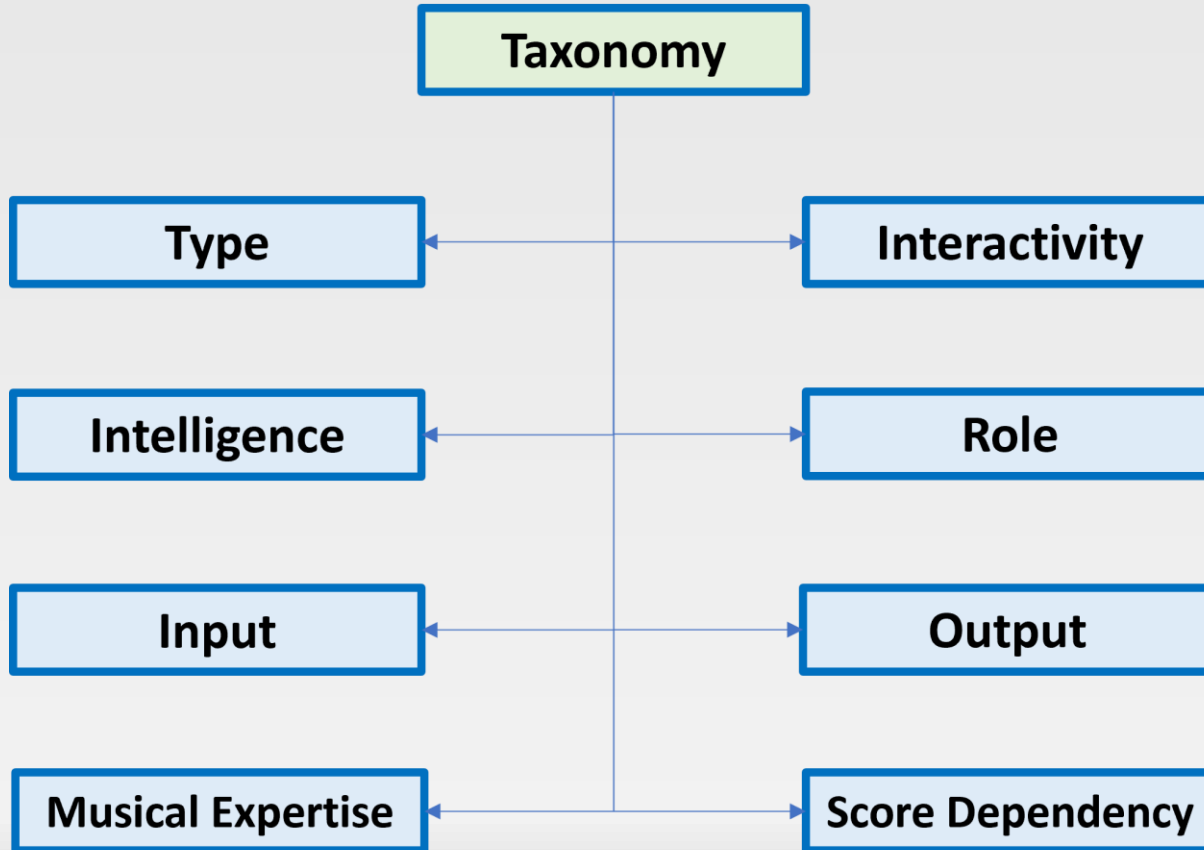
- (35 min - Christos) Continue live coding with Euterpe
- (15 min – Zhiyao) Challenges and future directions
- (10 min - All) Q&A

A Taxonomy

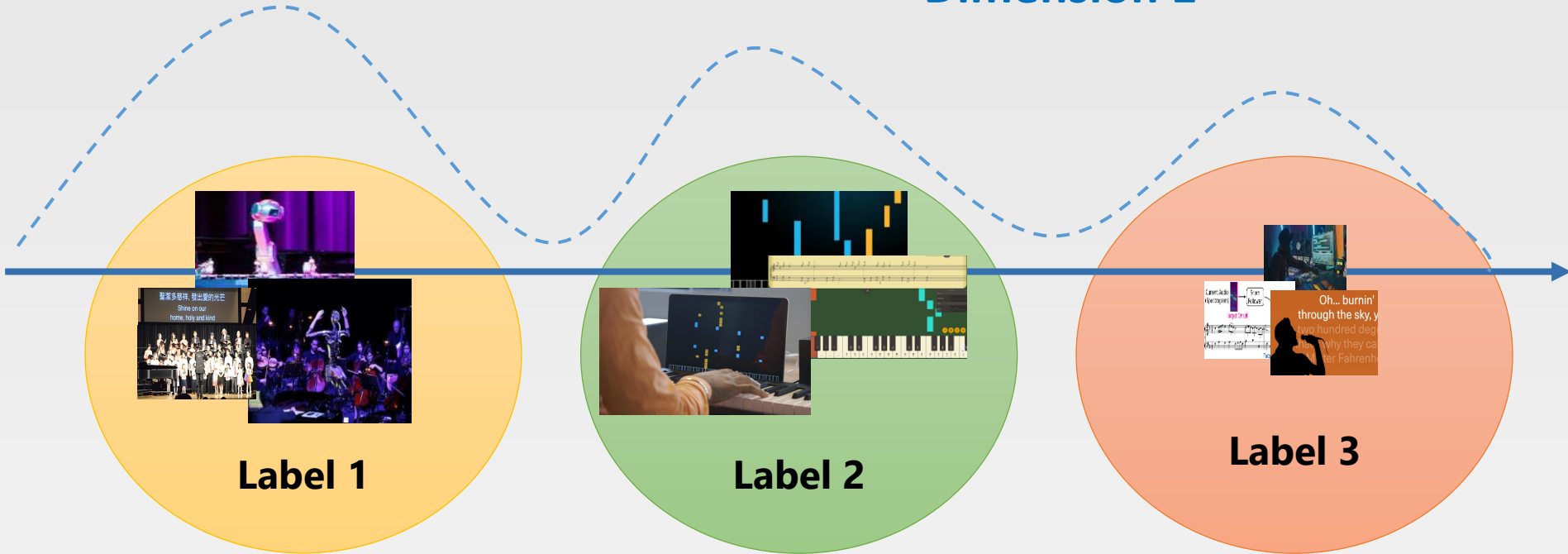
for Computer-Assisted Music-Making systems

Christodoulos Benetatos

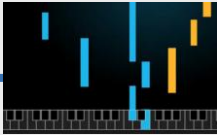
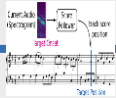
Dimensions



Dimension 1



Dimension 2

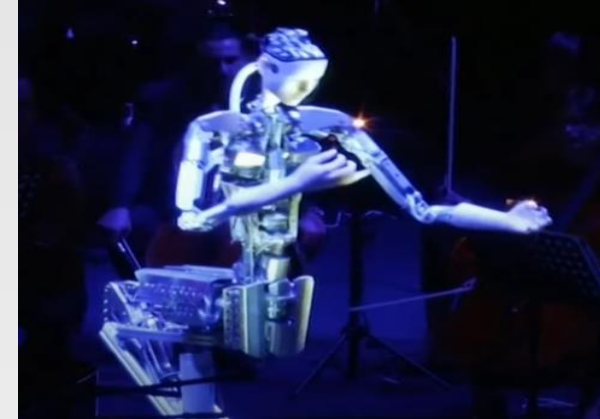




Automatic Lyrics Display



C & R Musical Partner



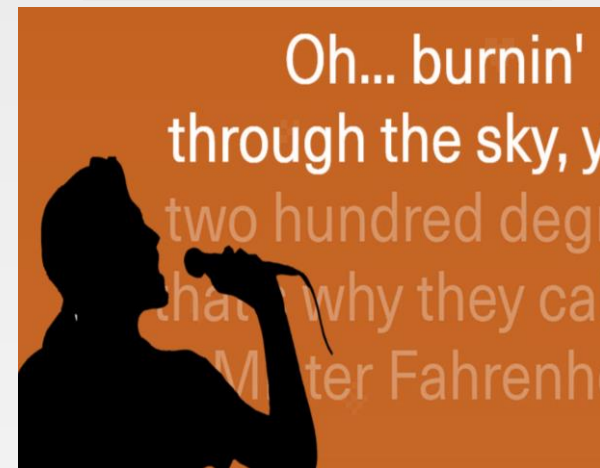
Robotic conductor



Intelligent Instrument

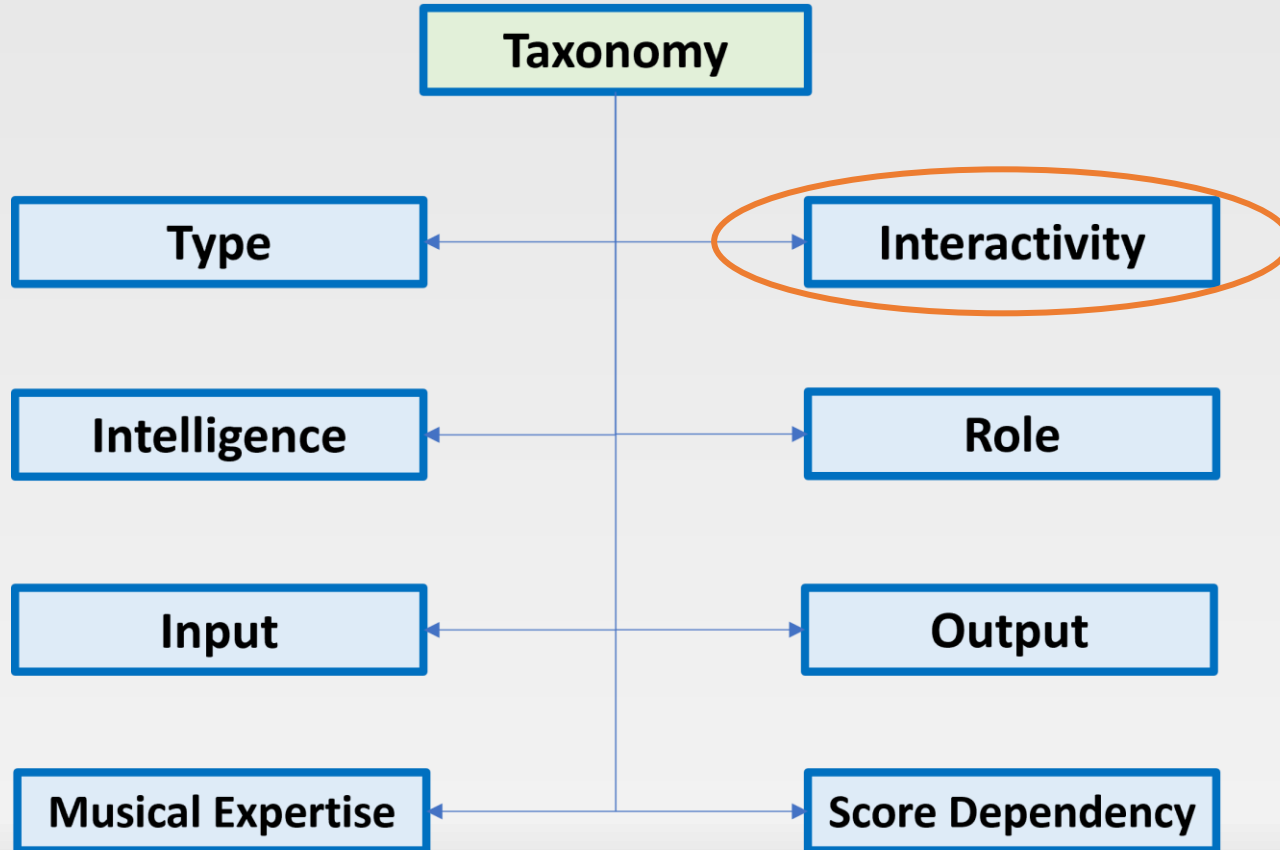


Robotic Ensemble Musician

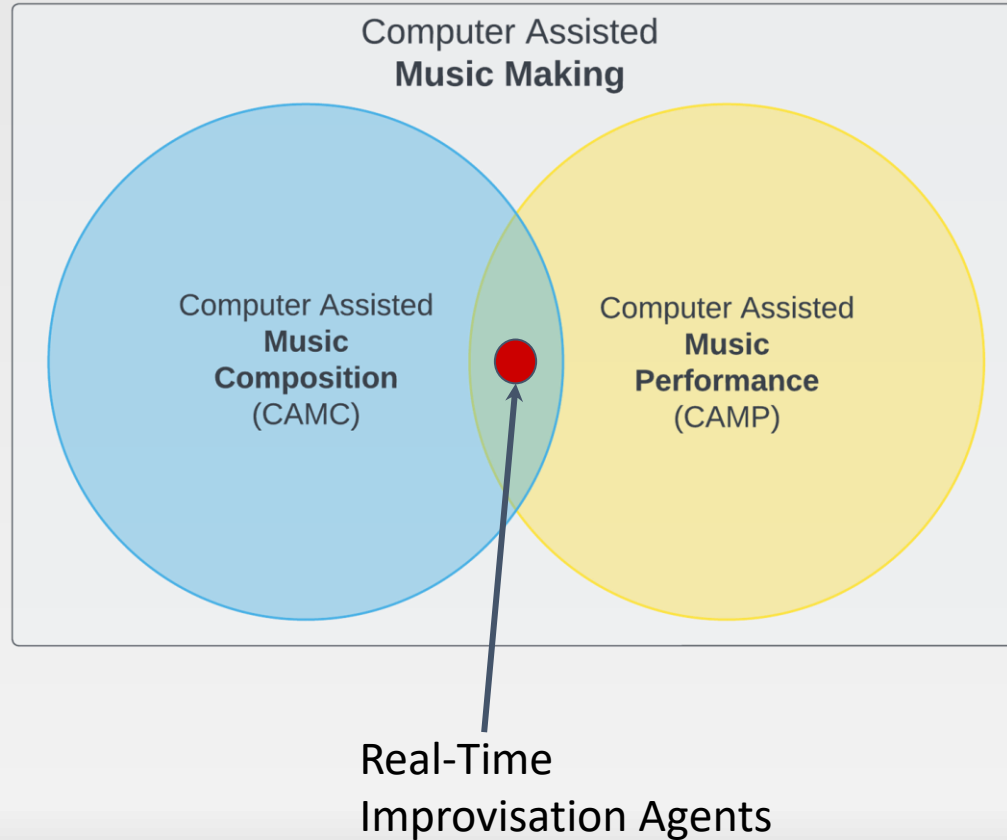


Karaoke

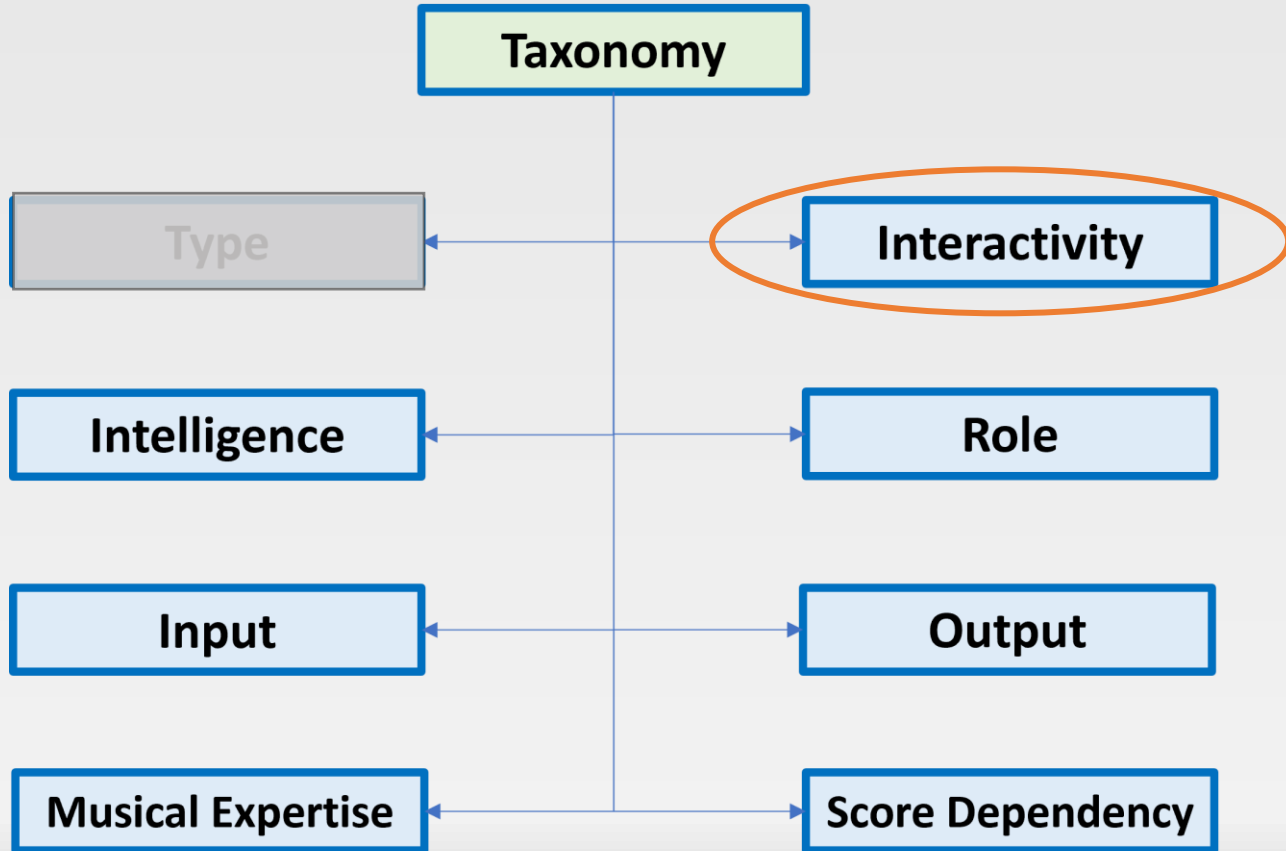
Dimensions



1. Type

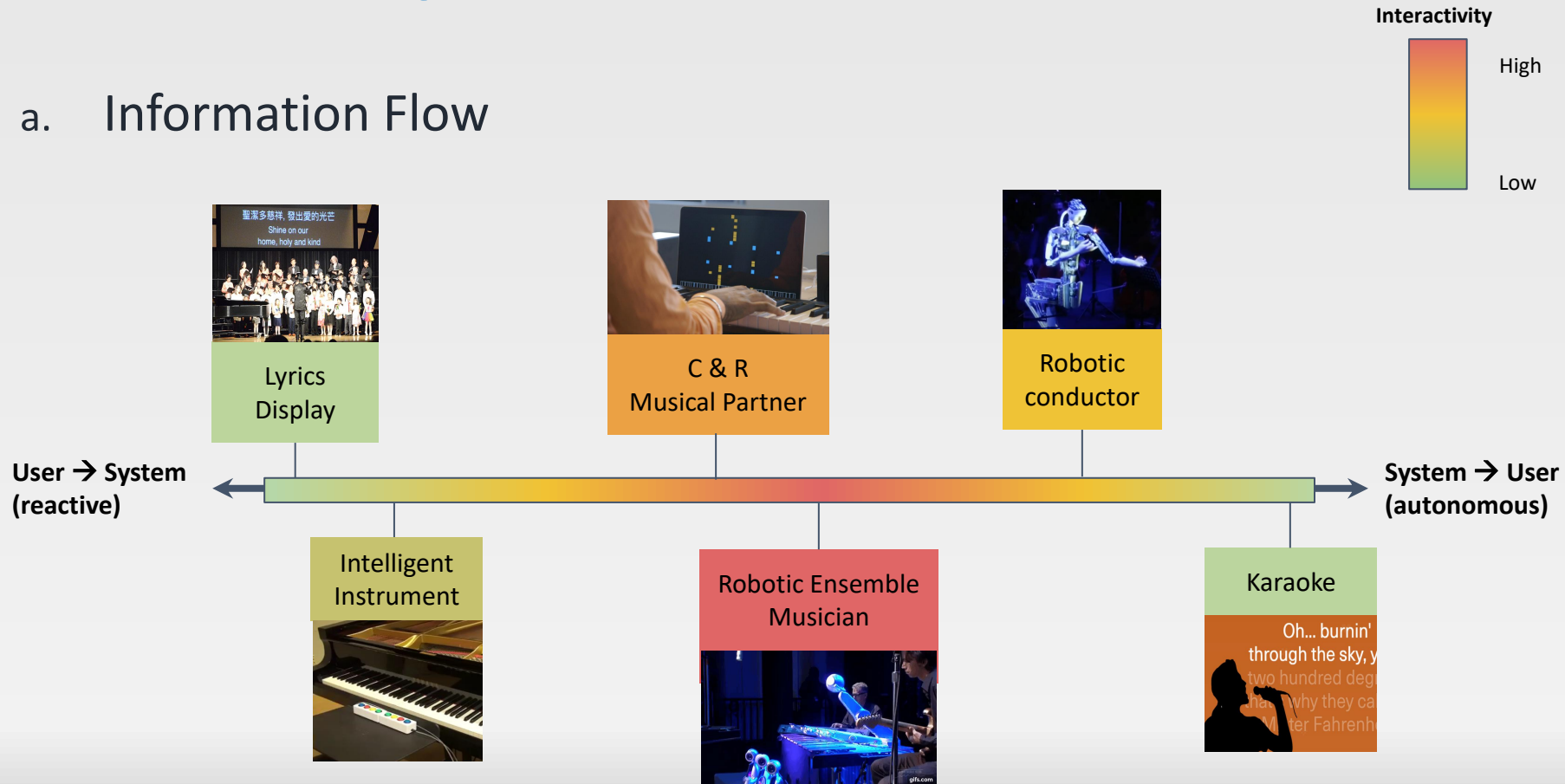


Dimensions



2. Interactivity

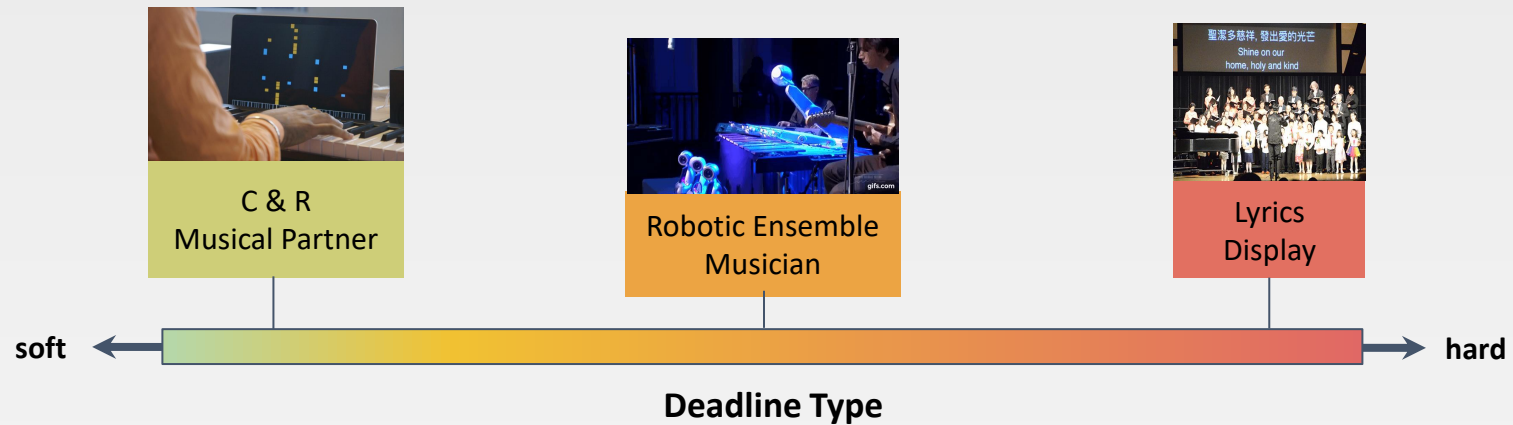
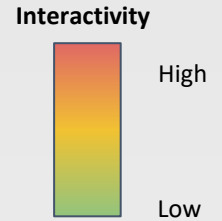
a. Information Flow



2. Interactivity

b. Responsiveness

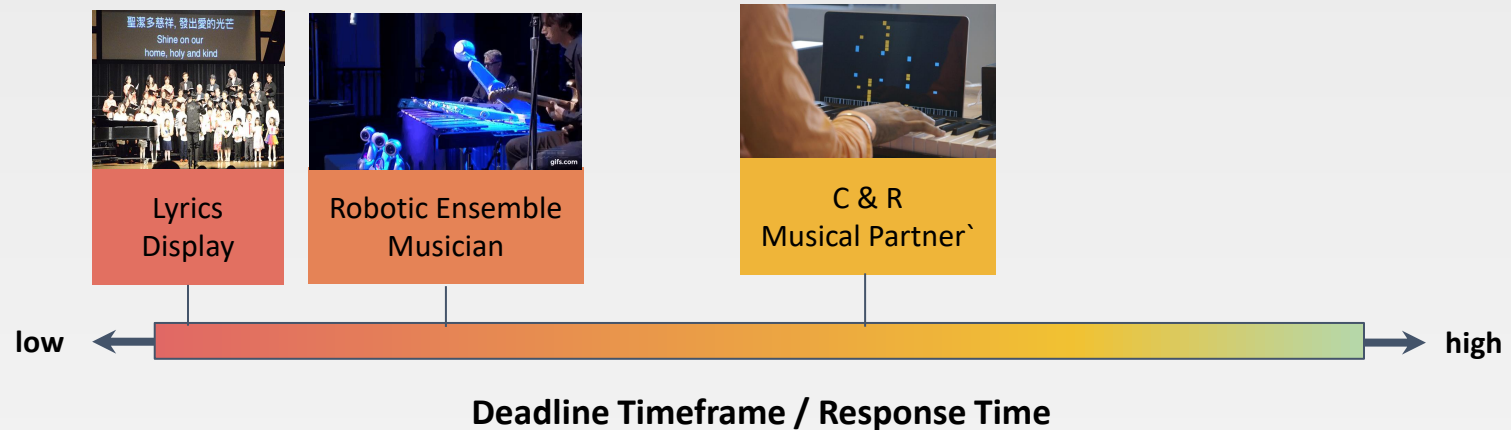
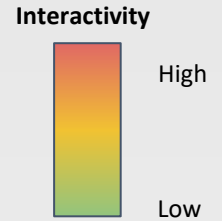
- For every **input** event, there is **response** within specified **deadline**



2. Interactivity

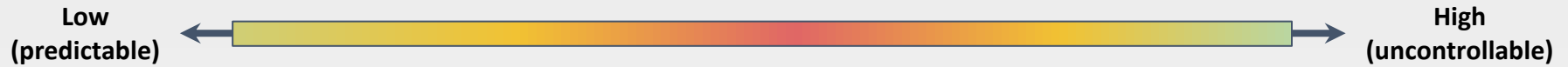
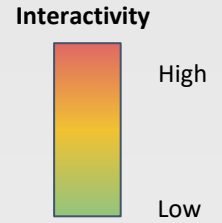
b. Responsiveness

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



2. Interactivity

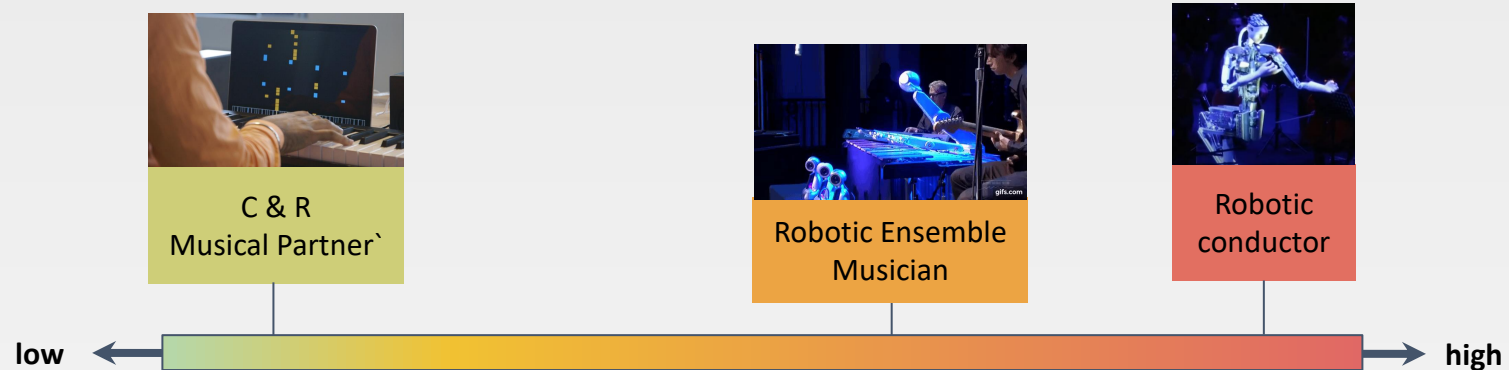
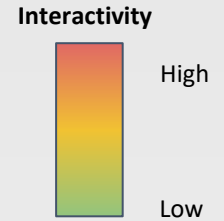
c. Stochasticity



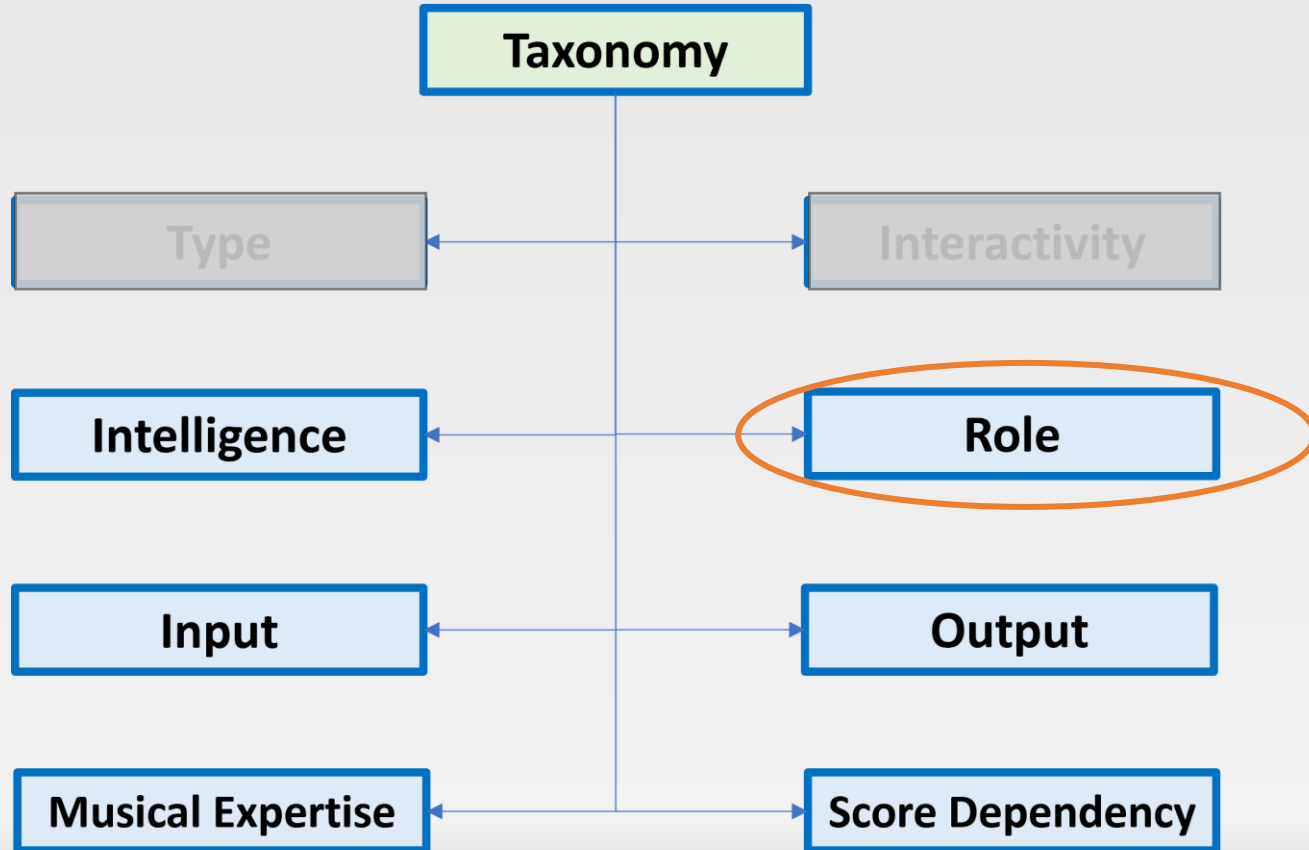
2. Interactivity

d. Synchronicity

- Percentage of concurrent task performance.



Dimensions



3. Role

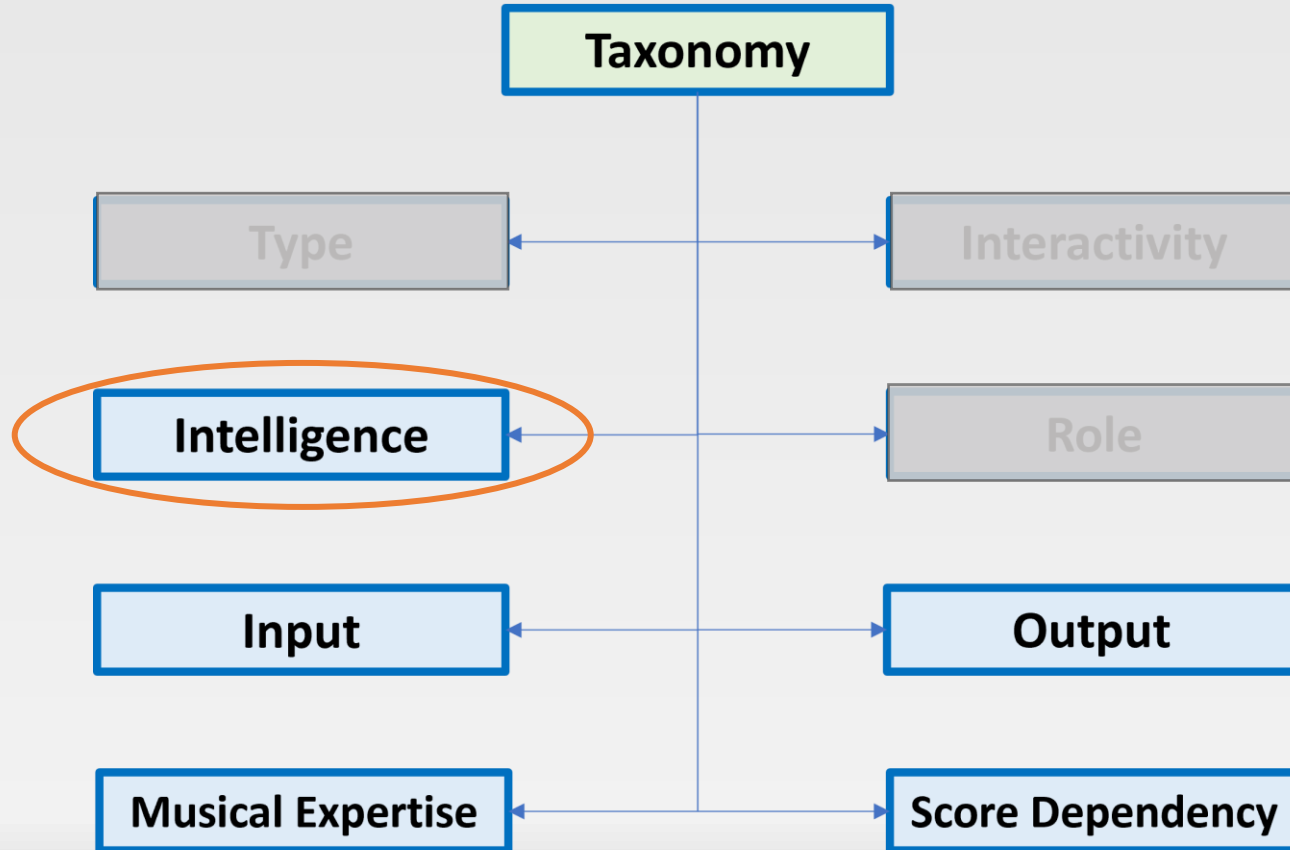
a. Agents

- Reactive
- Partners
- Leaders
- Autonomous

b. Tools

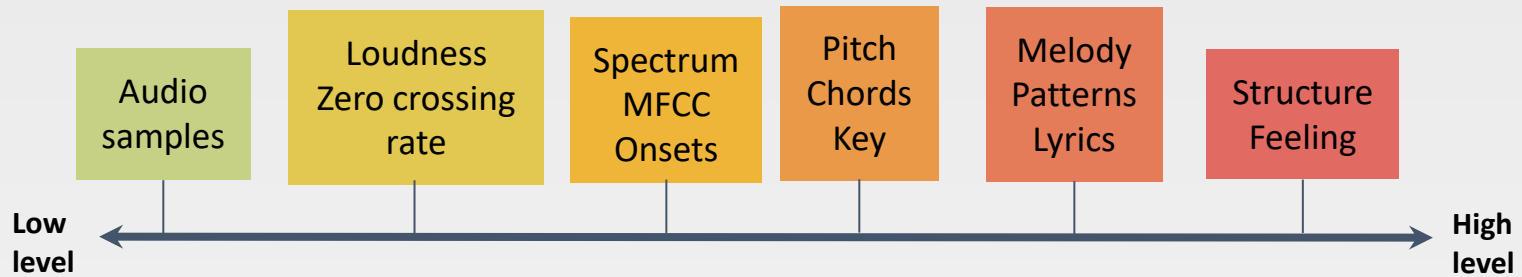
c. Instruments

Dimensions



3. Intelligence

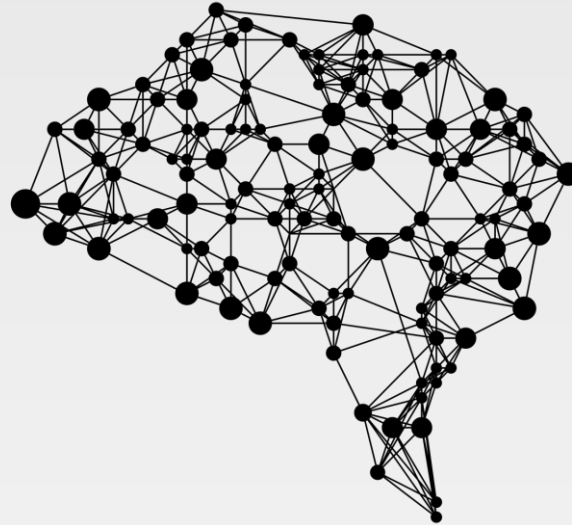
a. Perception



3. Intelligence

b. Cognition

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



3. Intelligence

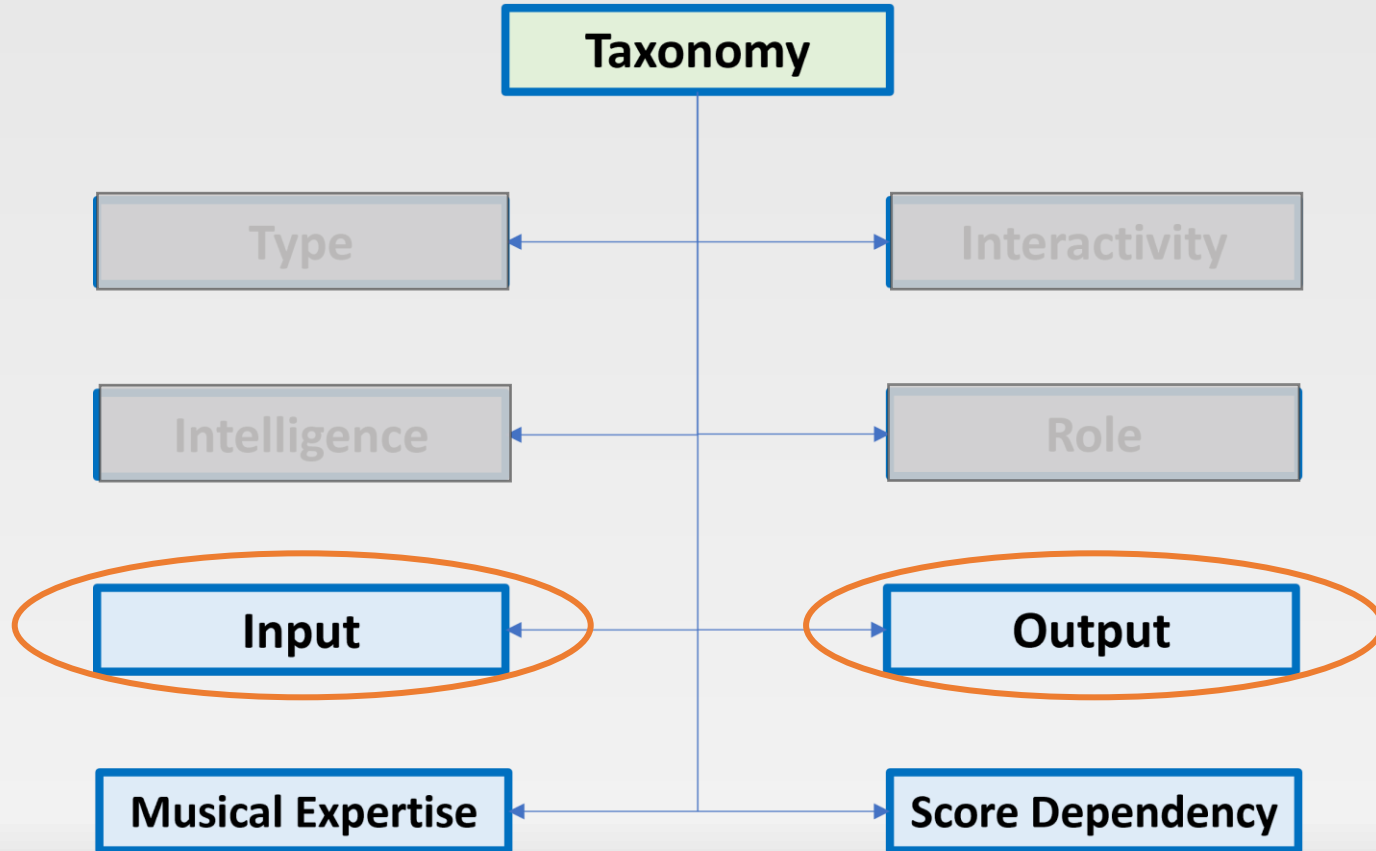
c. Expressivity

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

f *sfz*

poco rubato

Dimensions



5/6. Input/Output

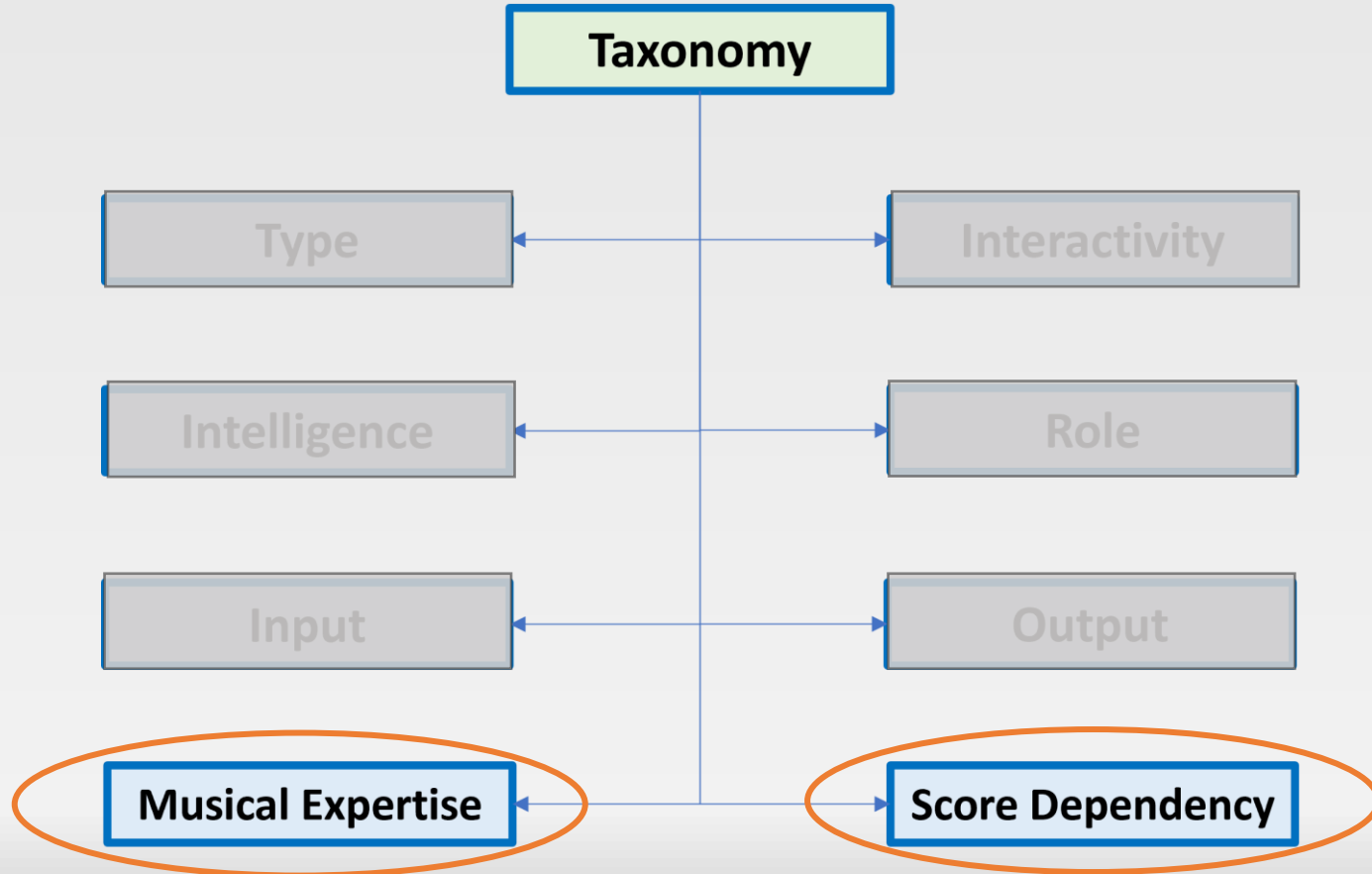
Input

- 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

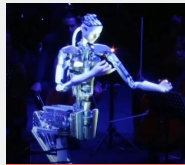


7. Musical Expertise

- 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



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)

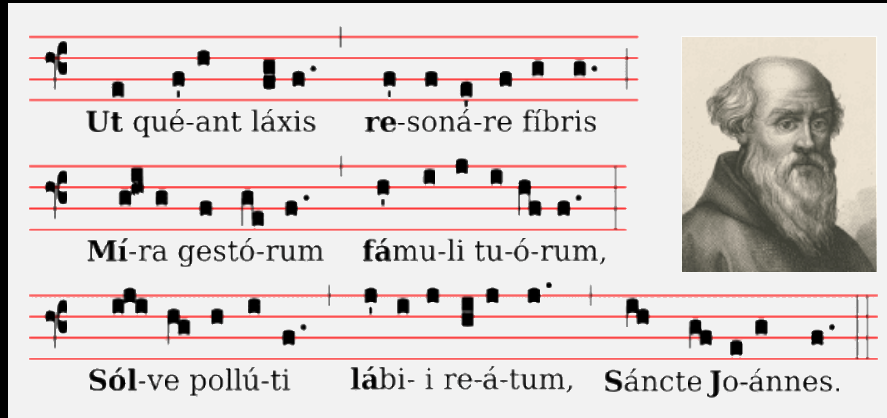
– Computer-Assisted Composition:

- Audio domain
- Symbolic domain

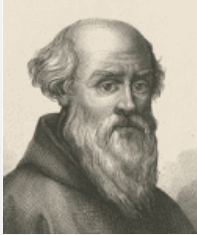
Historical Precedents

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



Ut qué-ant láxis re-soná-re fíbris
Mí-ra gestó-rum fámu-li tu-ó-rum,
Sól-ve pollú-ti lábi- i re-á-tum, Sáncte Jo-ánnes.



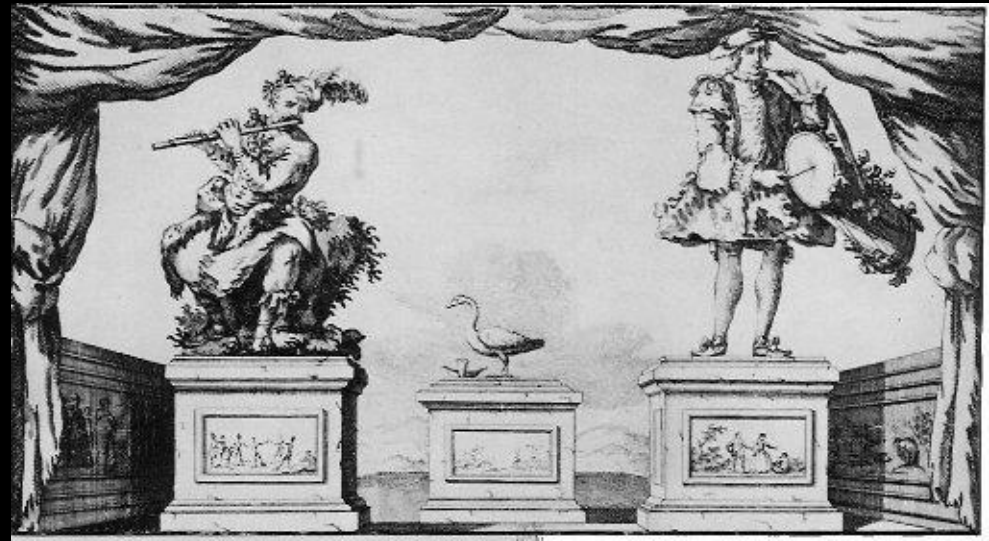
Early Automaton

- 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 ingenious 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 play different patterns.



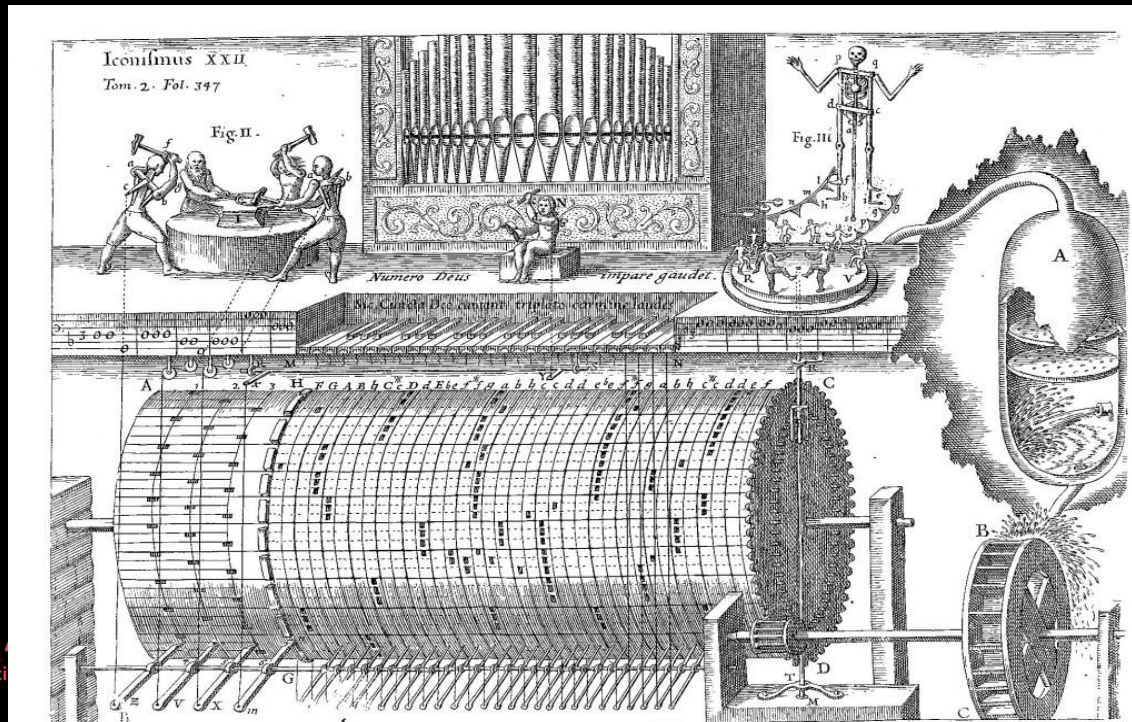
Early Automaton

- Around the 14th century, and with the development of physics hydrolic energy is supplemented with mecanical energy and steam powered systems.
- Automaton 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.

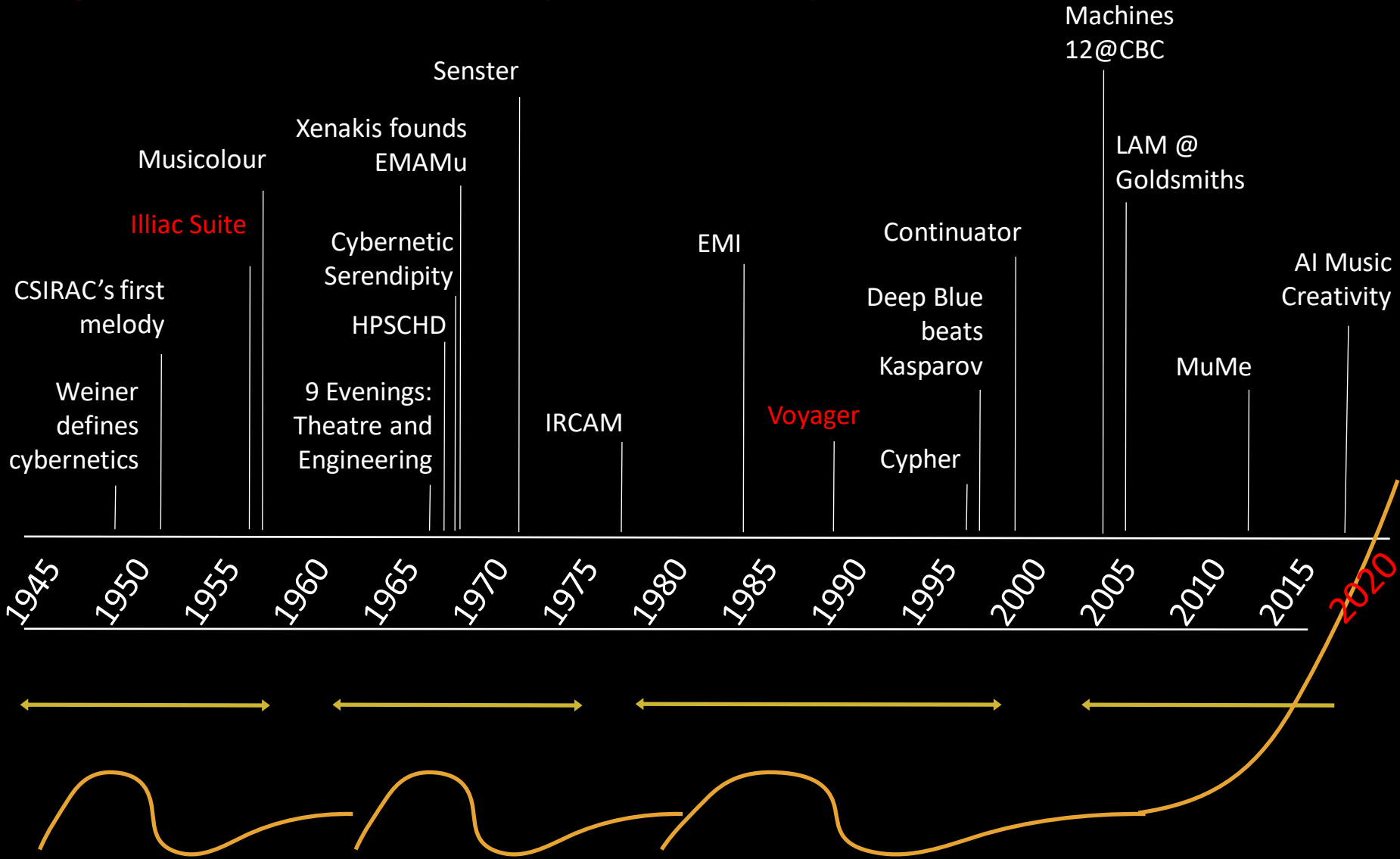


Early Automaton

- 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 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 reactivity and deliberativeness

Musical Agent



Journal of New Music Research

 **Routledge**
Taylor & Francis Group

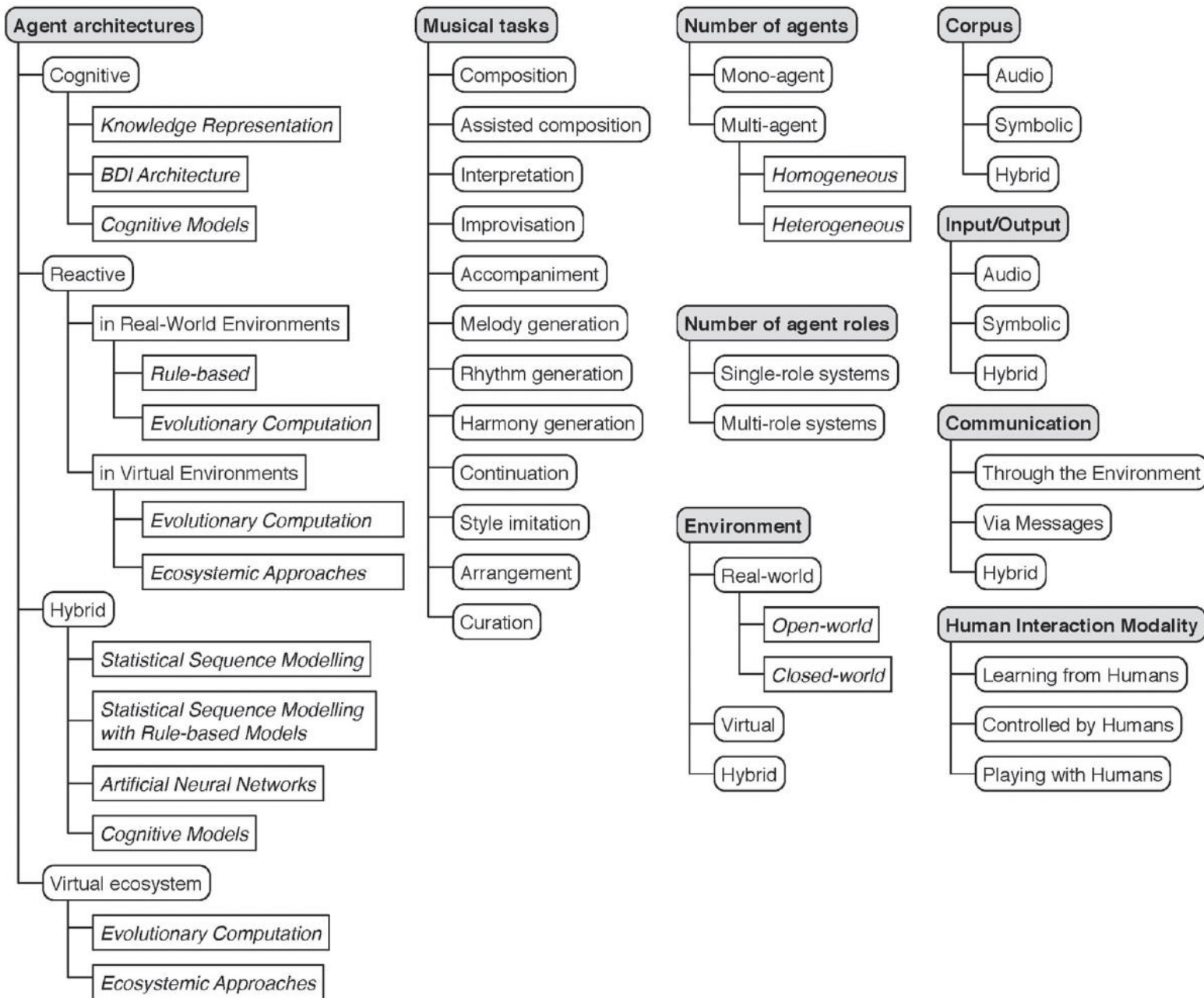
ISSN: 0929-8215 (Print) 1744-5027 (Online) Journal homepage: <http://www.tandfonline.com/loi/nnmr20>

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

Kivanç Tatar & Philippe Pasquier

To cite this article: Kivanç 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](https://doi.org/10.1080/09298215.2018.1511736)

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



#	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	

24	PIWeCS	Rule-based	Heterogenous 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.	

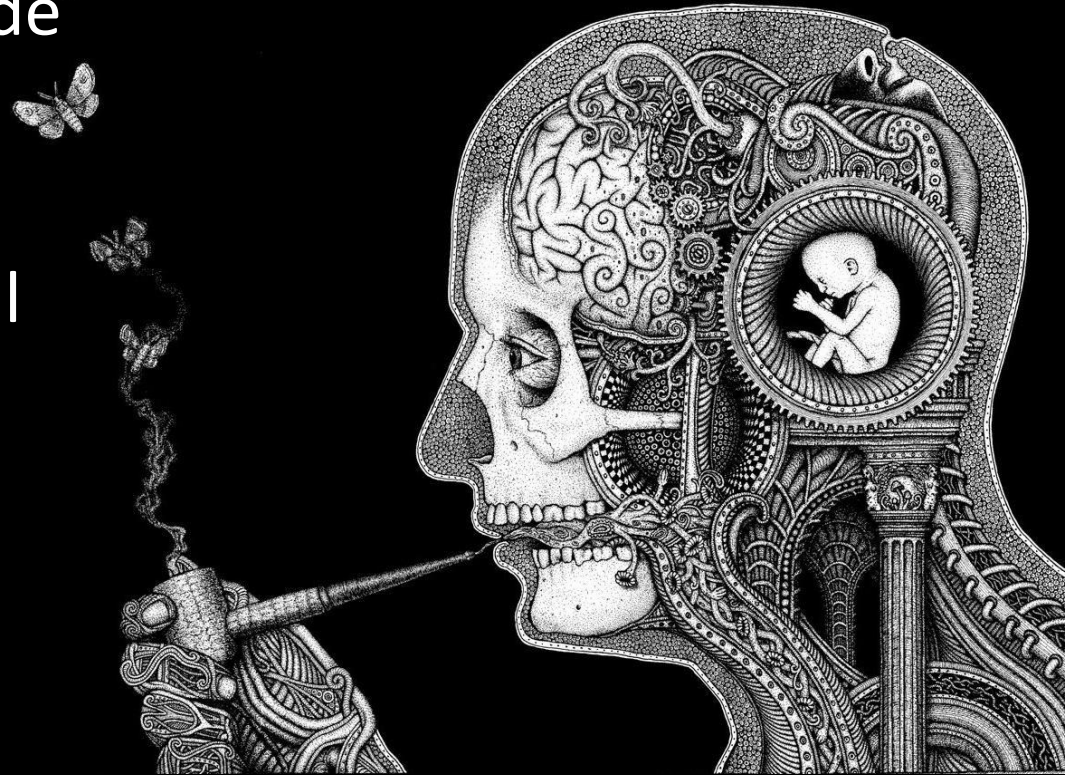
Table 1. Continued.

#	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	Real-world	–	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 Modelling	Multi-agent/ Heterogenous	Multi-role	Real-world	Hybrid	Hybrid	Hybrid	Env.	P+C+L	Improv., Style Im.		Not shared	✓

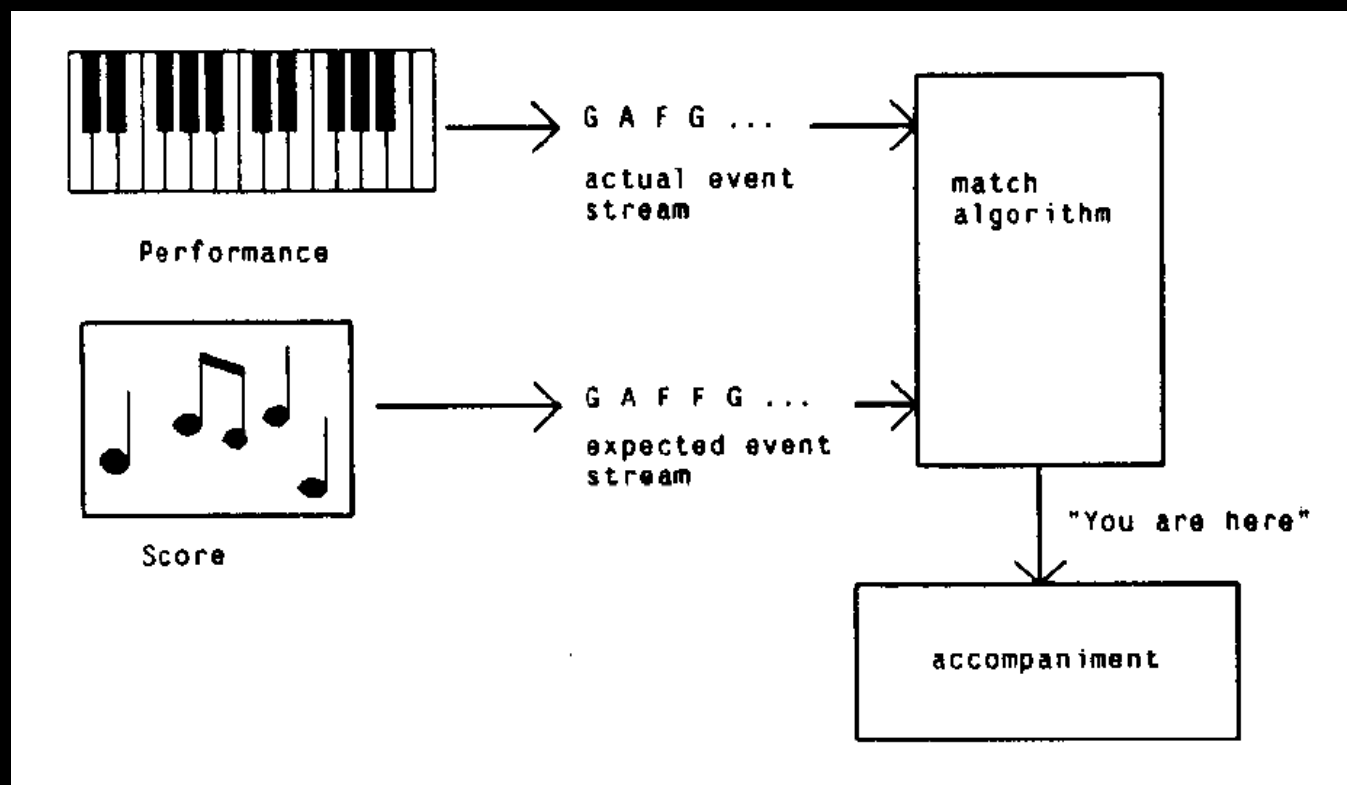
58	AO	Statistical Sequence Modelling	Mono-agent	Single-role	Real-world	Audio	Audio	Audio	Env.	P+C+L	Improv., Style Im.	Not shared	✓	
59	PyOracle	Statistical Sequence Modelling	Mono-agent	Single-role	Real-world	Audio	Audio	Audio	Env.	P	Improv., Style Im.	✓	Shared	✓
60	VMO	Statistical Sequence Modelling	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

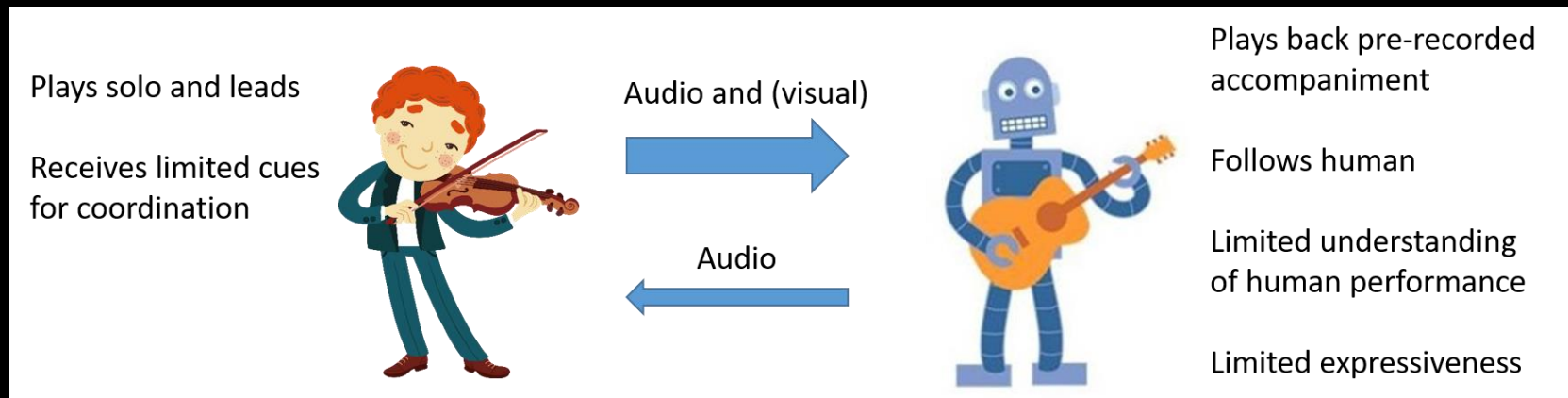


Early Accompaniment Agent



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

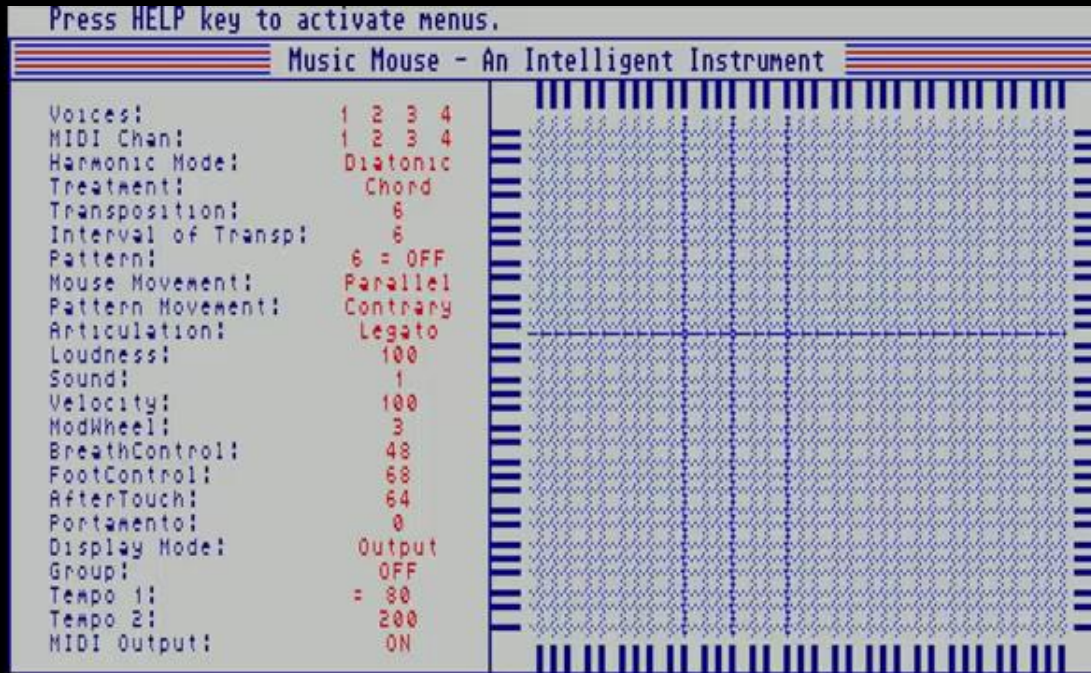
Lewis, George E. "Too Many Notes: Computers, Complexity and Culture in Voyager." *Leonardo Music Journal*, vol. 10, 2000, pp. 33-39.



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)

138

accel.-----♩.76

vln

ff

8

3

5

5

8

7.1

6.29.78

5.1

6.22.68

pf

7

8

ff

4

4

5

4

3

5

1

1

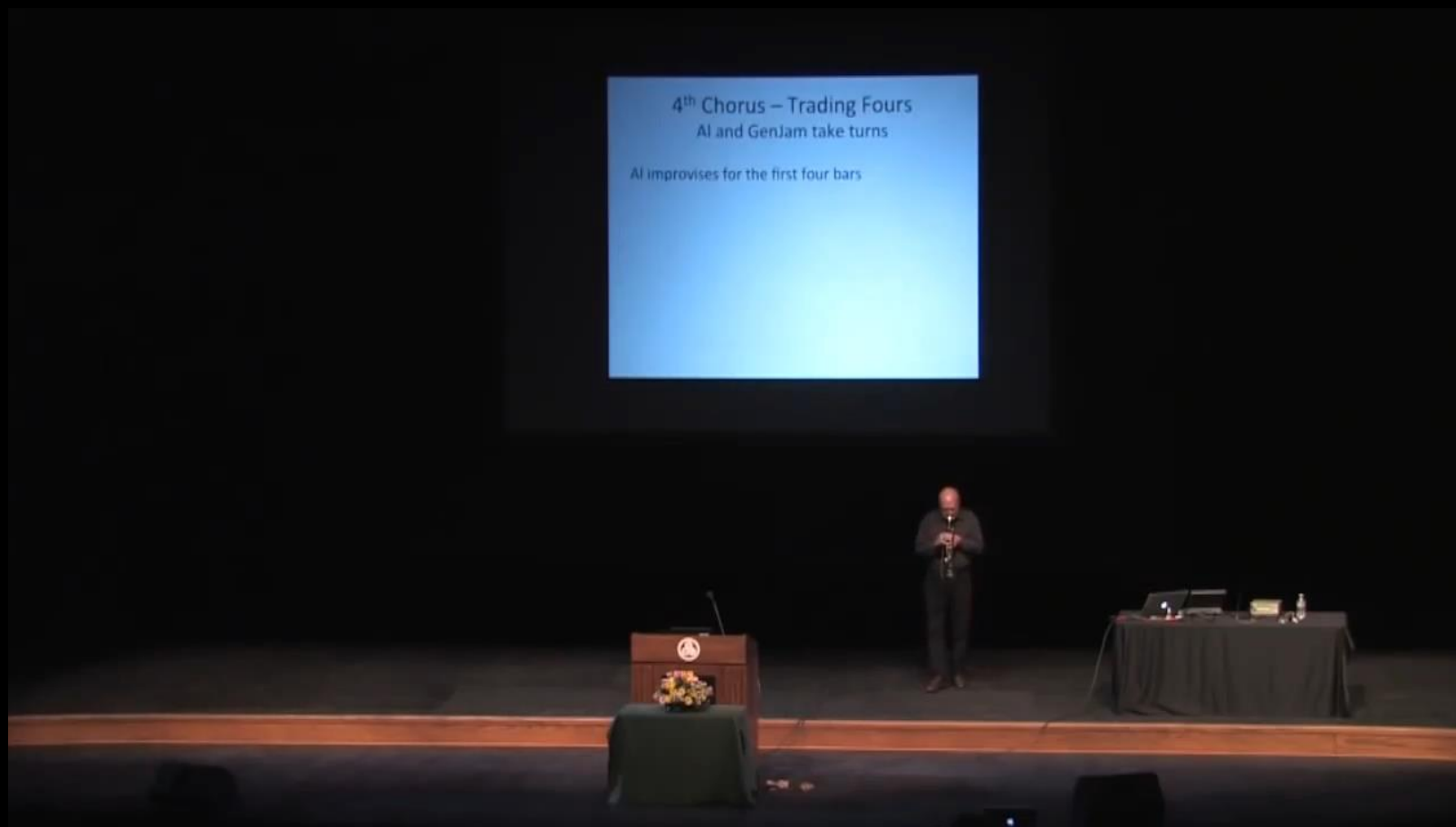
comp

State	35
Bank	2
Timbre	1
Shadow	pf
Spx	reverb

- 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

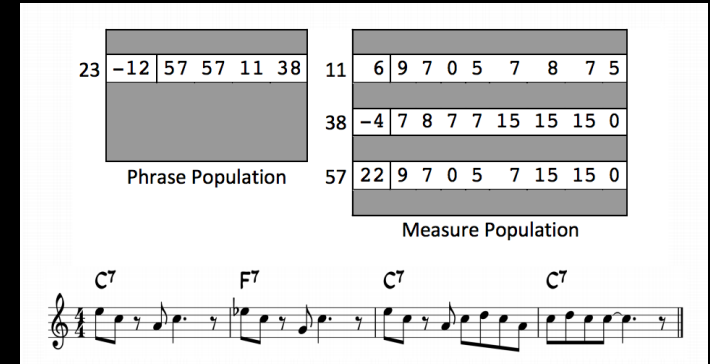
GenJam (Al Biles, 1993)

4th Chorus – Trading Fours
AI and GenJam take turns
AI improvises for the first four bars



GenJam

- In order to improvise Jazz solos, GenJam is co-evolving two populations of melodic ideas:
 - A **measure population** of 64 individuals: chromosomes are made of 8 genes that each map to an 8th notes. Each gene in a measure is encoded by four bits, with value 0 for rest, 15 being a hold, and 1-14 being the notes events that are mapped to an actual MIDI note through a set of scales that corresponds to the 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.



Continuator

(François Pachet, 2002)

BETWEEN OUR CONTINUATOR AND HANS ZIMMER

- 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)

The image displays the Kinetic Engine interface, which is divided into several functional panels. At the top center, the title "KINETIC ENGINE II" is shown with version information "v. 06.24.07" and "rae © 2005-2007". Below this, the word "CONDUCTOR" is prominently displayed.

System Overview and Parameters:

- Density:** A vertical bar on the left indicates a level that is "way too low". A "view auto" button is present.
- Pattern Check:** A toggle switch is currently set to "off".
- Counters:** Performance is 0 and Composition is 0.
- Parameters:** A central panel with a "view attraction" button contains several sliders: Confidence (0.27), Responsive (1.35), Social (1.7), Commitment (1.0), Mischievous (1.72), and Clumping (1.45). Each slider has a "reset" button. A "save agent parameters" button is at the bottom.
- Variation:** Amount is set to 1.0. Includes "view" and "auto" buttons, and a "Scarce" to "Often" scale.
- System Responsiveness:** A slider between "sluggish" and "tight". Includes a "view network" button.
- Probabilities:** Includes a "view" button and a grid for filter, delay, degrade, and tuning.
- Global Volume:** Includes a "stereo" button and a "reset" button.
- Receiving agents (group):** A row of six empty slots with a "jsui" button.

Agent Management:

On the right side, there are six agent control panels, each for an "agent#" (0 through 5). Each panel includes:

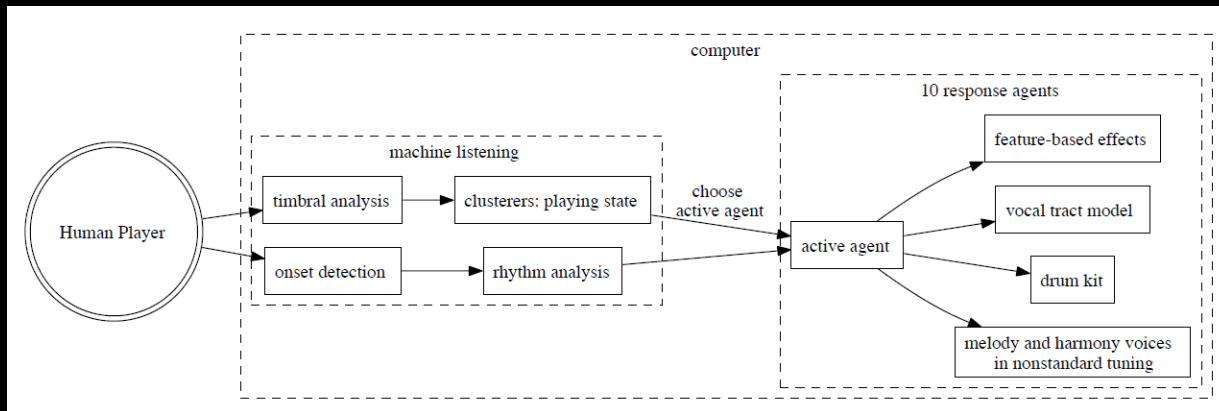
- Agent ID and name (e.g., agent# 0: 2669_agent, Argentina_CJ).
- Type (low, mid, or high).
- A "patterns" dropdown menu.
- "mute" and "view" buttons.
- An "active" button.
- Bond and hetro/poly options.

Audio and System Controls:

At the bottom left, there are controls for "Pulse" (set to 3), "Tempo" (148), "Metre" (16), and "Subdiv" (4). There are also "initialize system", "new section", "fade", "clock", and "help" buttons.

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

Coming Together / Beauty & Truth

arne eigenfeldt

The interface is titled "Coming Together / Beauty & Truth" by Arne Eigenfeldt. It features a top control bar with "Initialize", "Run", and a timer set to "00:00". Below this are dropdown menus for "One time" and "Output" (set to "Kontakt").

The main area is divided into several sections:

- Explorer One to Explorer Four:** Each explorer has a "nonCT check" section with "tempo" and "tala" values. Below are "goal" (actual, difference, misses), "tala belief" (others, certainty), "rhythm", "pitch" (range, p spacing, probVector, pitch goal, inertia), "pitch movement", "density" (actual), "ratings", and "nonCT movement" (volume). The instruments are "Jinghu Operaviolin", "Shakuachi", "Jinghu Operaviolin", and "Kantele".
- Quark One to Quark Four:** Each quark has a "Gesture Shape" (range), "window", "velocity", and "measures rested" section.
- Global Composition Variables:** Includes "interval spacing", "pitch avoidance", and "Wait time scaler" (Responsiveness).
- Active Variables:** Shows "chord tones" and "non-chord tones" on a piano roll, "Success 0.", and various feedback variables like "pitch goal", "pitch stabilized", "nonCTs stabilized", "measure lengths", "downbeat onsets", and "tala negotiated".
- Bottom Control Bar:** Features "Partial Analysis" buttons for each Explorer and Quark, and a "requested by" field.

Reflexive Looper (Pachet et al., 2013)



<https://www.youtube.com/watch?v=oquvn8GybR>


s

- 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.



PLAY NICE
MUSICAL COLLISIONS BETWEEN HUMANS & INTELLIGENT MACHINES

FRIDAY JULY 28, 2017

HUMANS
PEGGY LEE (CELLO) DAVID STOREN (VIDEO)
MATTHEW ARIARATNAM (PREPARED GUITAR)
ADRIAN VERDEJO (GUITAR) BARBARA ADLER (TEXT)
NATHAN MARSH (PREPARED GUITAR)

MACHINES
MUSEBOTS DESIGNED BY
ARNE EIGENFELDT MATTHEW HARRIGAN
PAUL PRACZBI YVES CANDAU

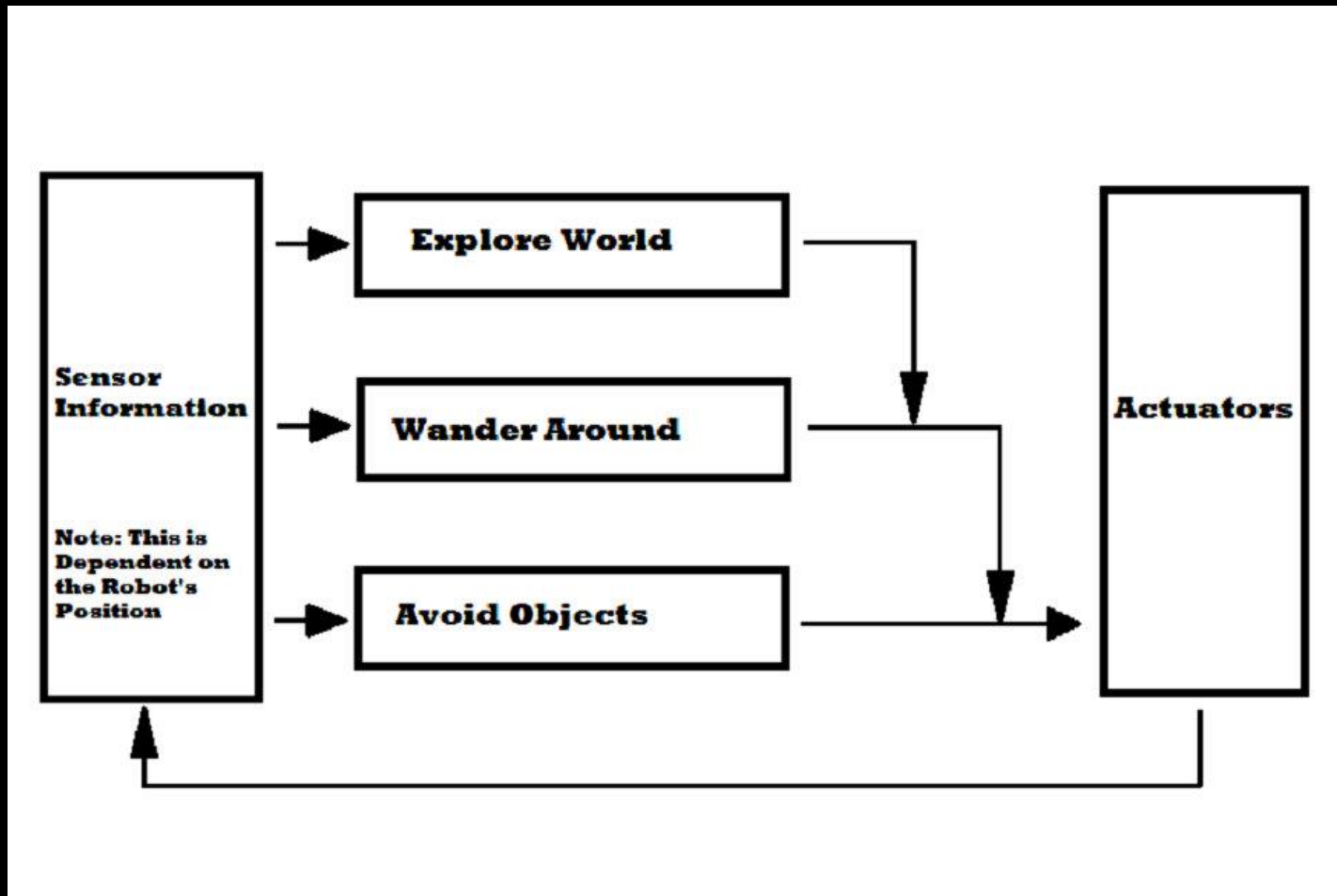
THE GOLD SAUCER STUDIO
211-207 W. HASTINGS, THE DOMINION BUILDING
DOORS 8:30 PM. SHOW 9:00 PM
\$10-\$15 AT THE DOOR
NO ONE TURNED AWAY FOR LACK OF FUNDS

WWW.MUSEBOTS.WEEBLY.COM
WWW.FACEBOOK.COM/MUSEBOTS

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 reactivity and deliberativeness

Subsumption Architecture

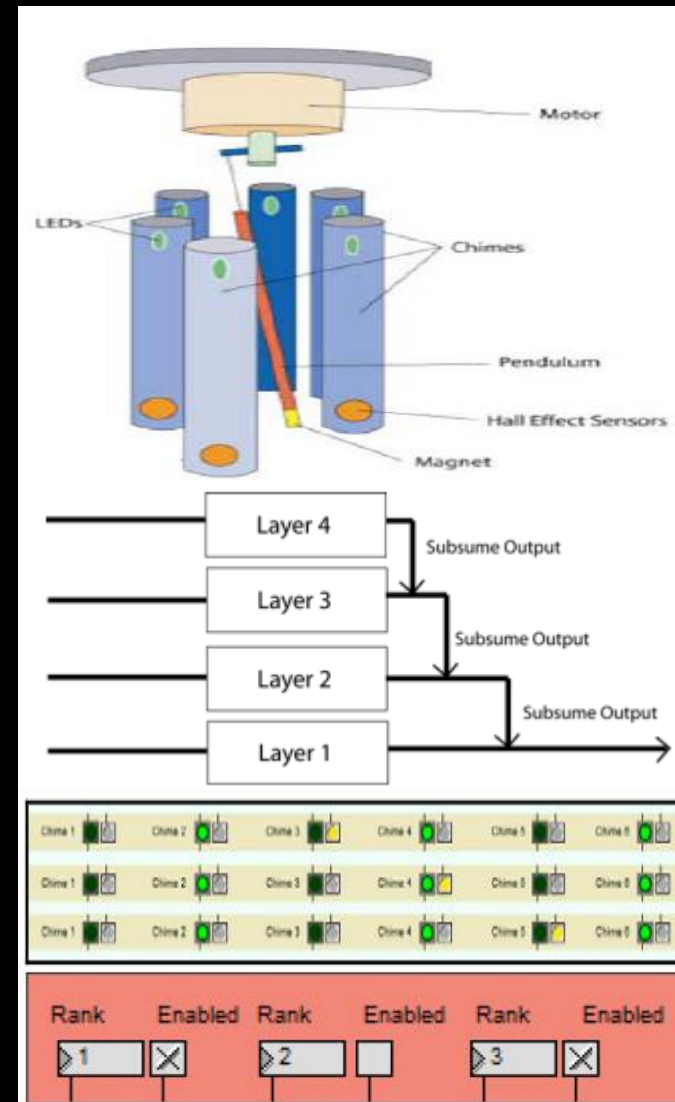


Musical Metacreation

With Aaron Levisohn
ACM ACE 2008

BeatBender: multi-agent rhythm generation

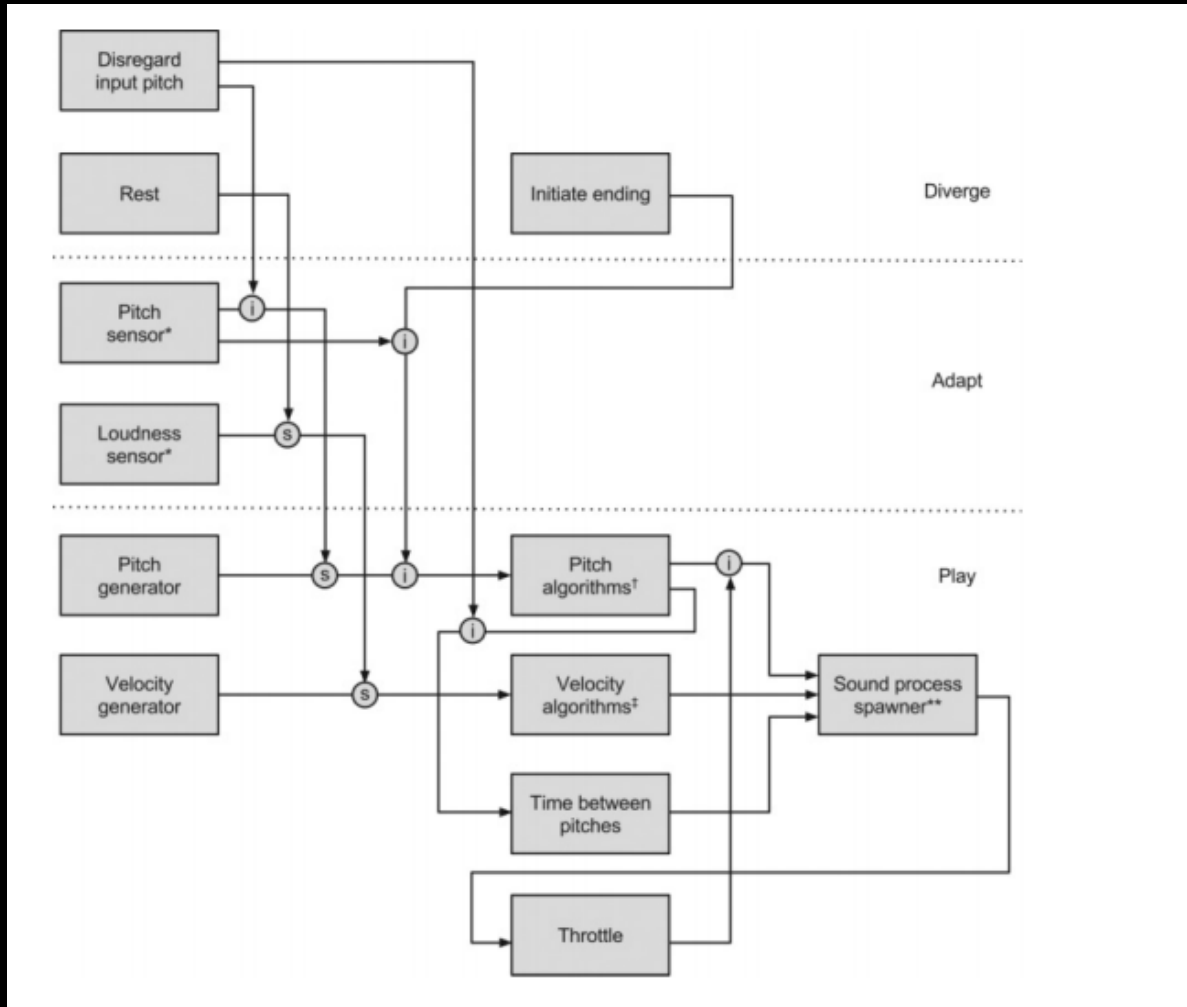
- Challenge: **non-corpus based** generation of rhythmic patterns
- Our approach:
 - Using reactive agents to create rhythmic patterns
 - Using **subsumption agent architecture**
- Experiments on a sample of 10+10 rhythms show that:
 - Humans prefer BeatBender rhythms over human composed ones
 - They 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.

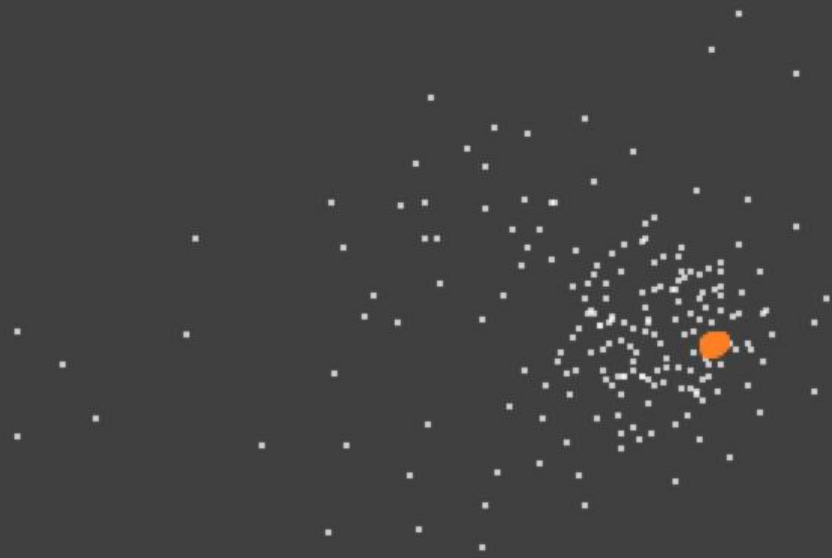
Musical agent – Odessa

- 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.

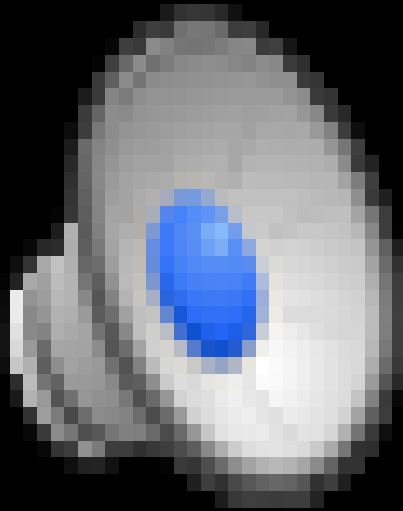


Reactive agents – Swarm Music

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



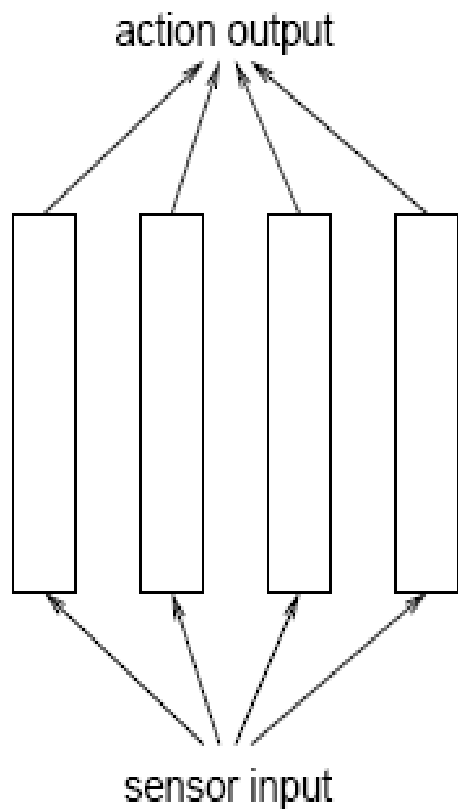
Musial agents – Porto Actors with Eargram



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

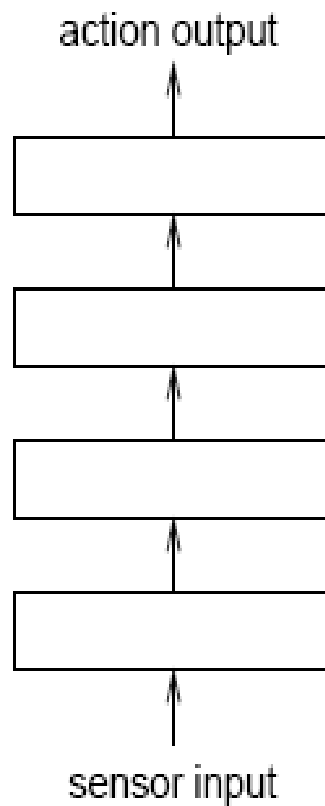
Layered Hybrid Architectures

Horizontal Layering

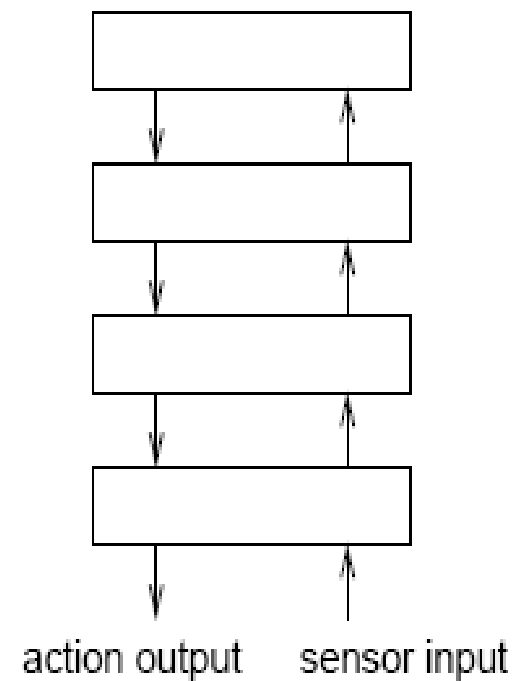


Vertical Layering

one-pass control



two-pass control



Generic Musical Agent Architecture

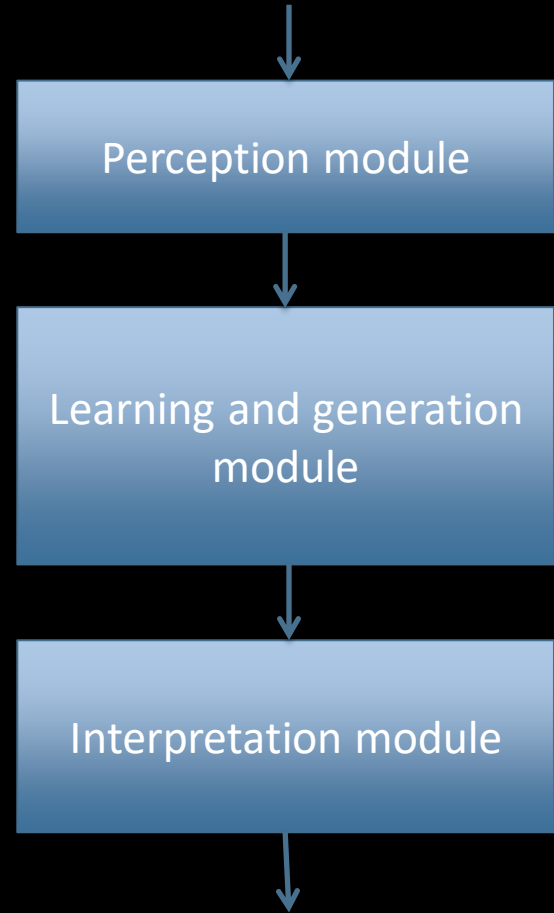
Input (audio or symbolic)

Perception module

Learning and generation
module

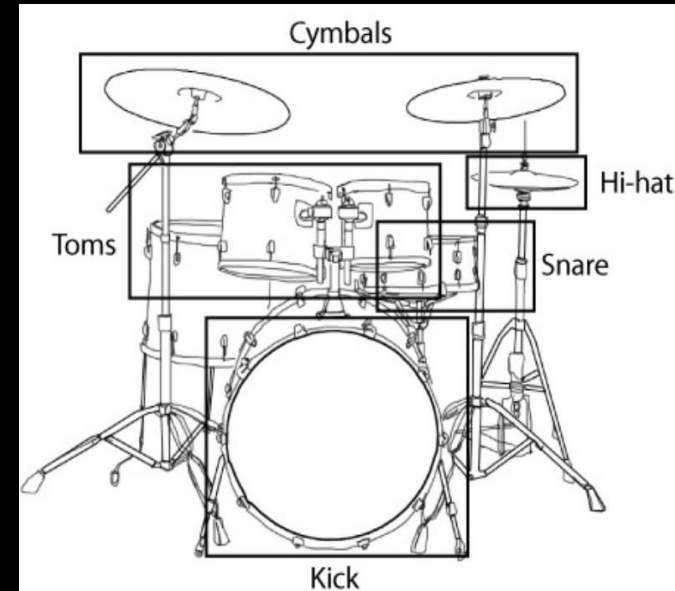
Interpretation module

Output (audio or symbolic)



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



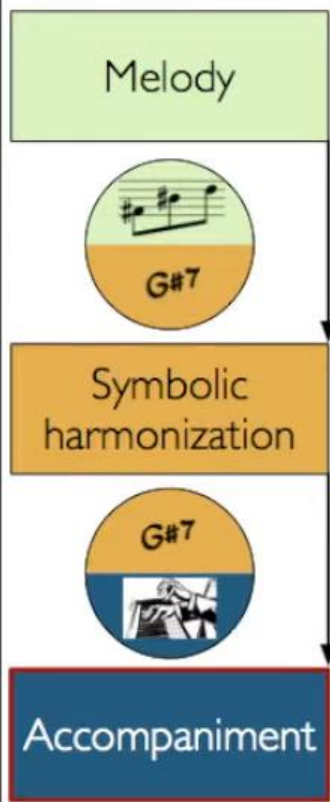
ImproteK, Jerome Nikka et al. 2012-ongoing

Automatic harmonization and arrangement using different small corpora

Examples: solo on

1) "Au Privave"

2) "Blues for Alice" 3) "J'aime pour la vie" (B. Lubat)
 (generated by the system)



I.1

"Au Privave"

 Song by
 Georges
 Brassens
 ("Le mauvais sujet
 repent")

I.2

"Au Privave"

 "Blues for
 Alice"

I.3

Set of jazz
 standards

 Set of jazz
 standards

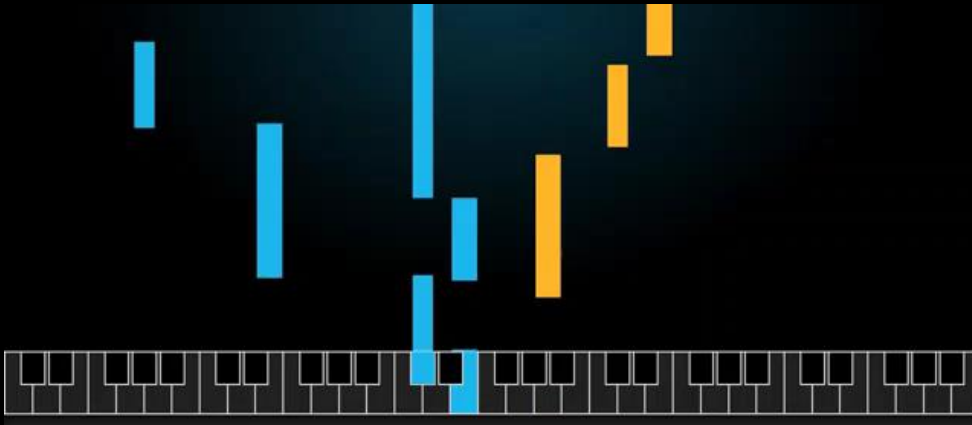
Experiments using an early MIDI version of ImproteK (2012)

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)

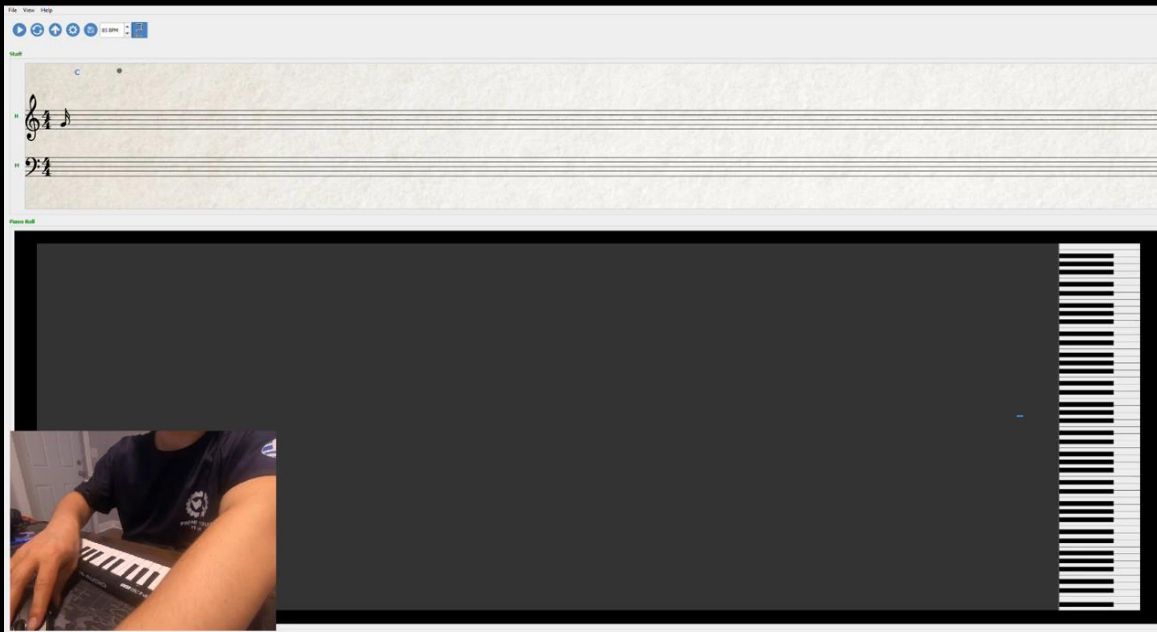


<https://www.youtube.com/watch?v=YRbOXAnUplk>

- 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

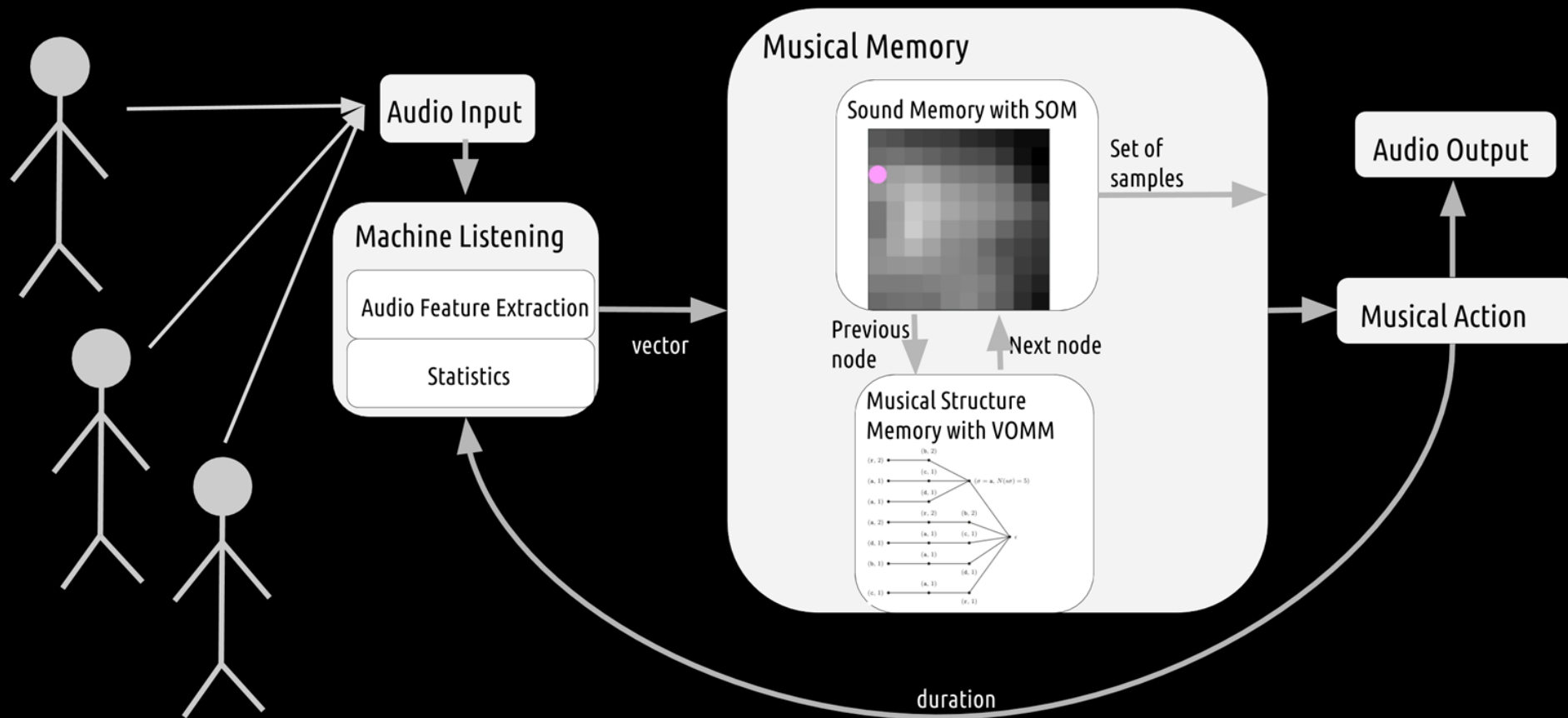
BachDuet (<https://bachduet.com/>), Benetatos & Duan, 2019.

*



- 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)

/MASOM-factor-v1_04/REVIVE-1

IN OUT

drop a folder here!

MASOM

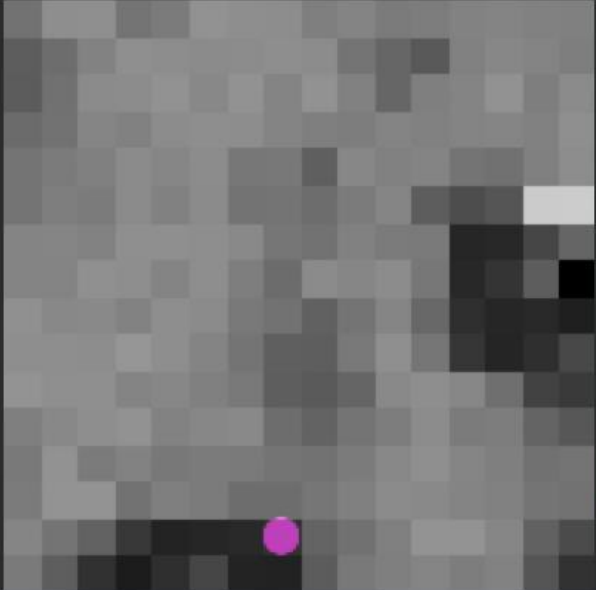
START

Gain: -96 dB
Fade out
▶ 500 ms

Self-Organizing Map

STATS

	Valence	Arousal
Mean	▶ 0.	▶ 0.
Std	▶ 0.	▶ 0.
Min	▶ 0.	▶ 0.
Max	▶ 0.	▶ 0.
NumData	▶ 0	



▶ 0.8 Congruence

141 Perceived node

231 Played node

0.73 Sample Length

Randomize-feature-vis-OFF

SOM dimensions:
a x a ▶ 16

MFCC_3-m... ▾

/MASOM-factor-v1_04/input

soundfile ▾
KIVANC-16.wav

0 : 0 : 0

Transpose 0 st

Open
Play
Loop

Audio On

DSP

Pan 0.00

Master Gain 0.00 dB

/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

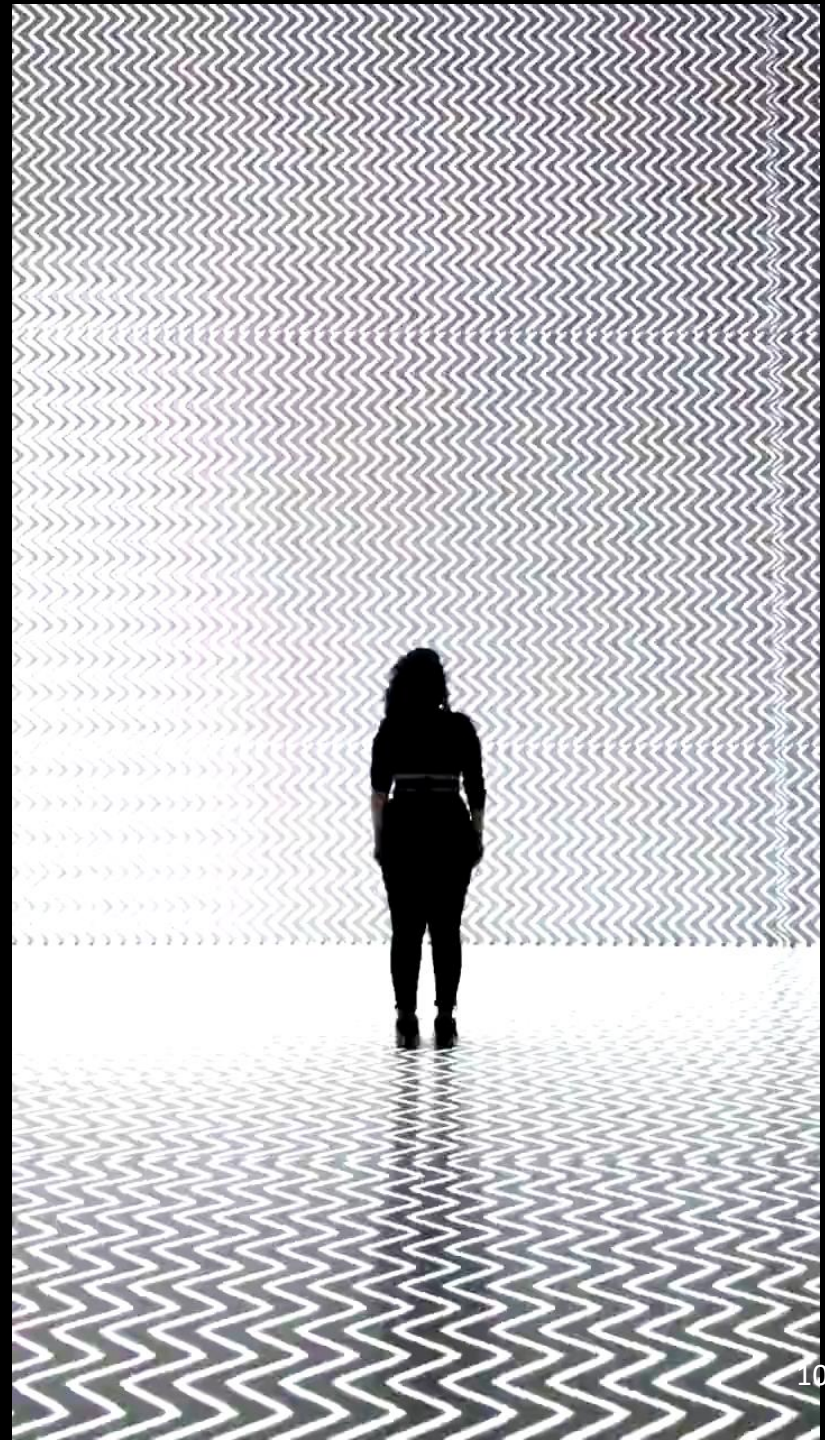
1 2

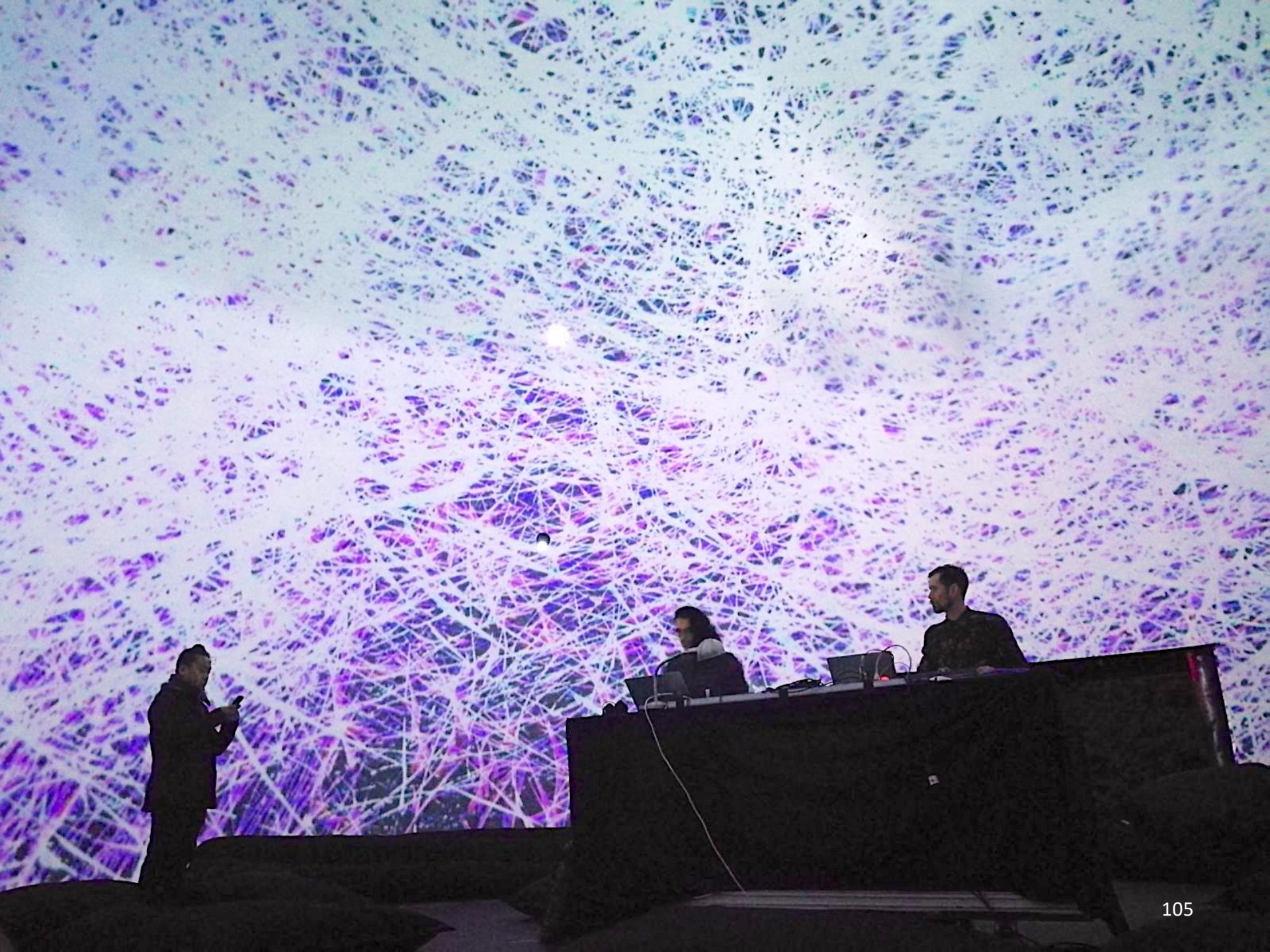
Pan 0.00

Master Gain ▶ 0. dB

iOTA

(collaboration with OUCHHH
and Audiofil)







SpireMuse: musical agent

End session 00:00:11

AGENT Mute drop the agent folder here!
Folder: guitar_acoustic_rhythm

INFLUENCES

Rhythmic Spectral Melodic Harmonic

STATE Mode: Shadowing
Select song to train FO
00_Funk2-108-Eb_comp_mic

INTERACTIVE MODE
Shadowing Mirroring Coupling

FEEDBACK
Thumbs up

IN OUT

End session 00:00:32

AGENT drop the agent folder here!
Folder: guitar_acoustic_solo

INFLUENCES

Rhythmic Spectral Melodic Harmonic

STATE Mode: Shadowing
Current song training the FO
00_BN1-129-Eb_solo_mic

NEGOTIATION
Go back Pause Change

FEEDBACK
Thumbs up

IN OUT

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

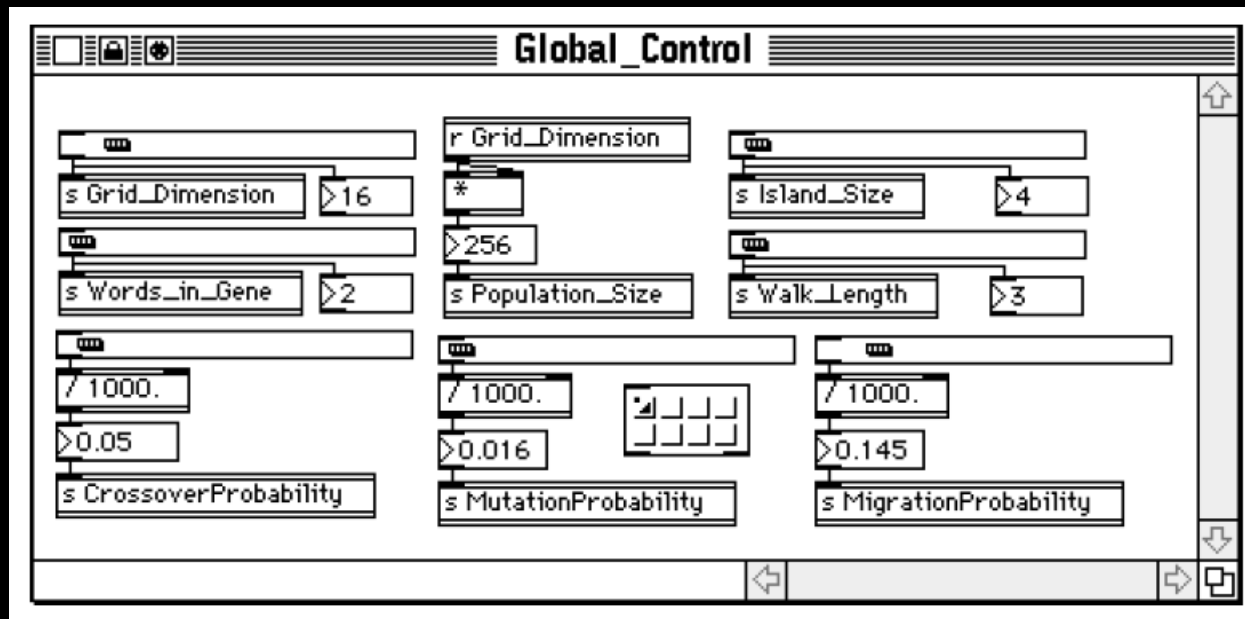
- Historical precedents
- Musical Agents:
 - Cognitive agents
 - Reactive agents
 - Hybrid agents
 - **Virtual Ecosystems (Artificial Life)**
- Computer-Assisted Composition:
 - Audio domain
 - Symbolic domain

(Evolutionary) ecosystem

- Mixing agents, ecosystems, and evolutionary computing



Purely sonic ecosystems



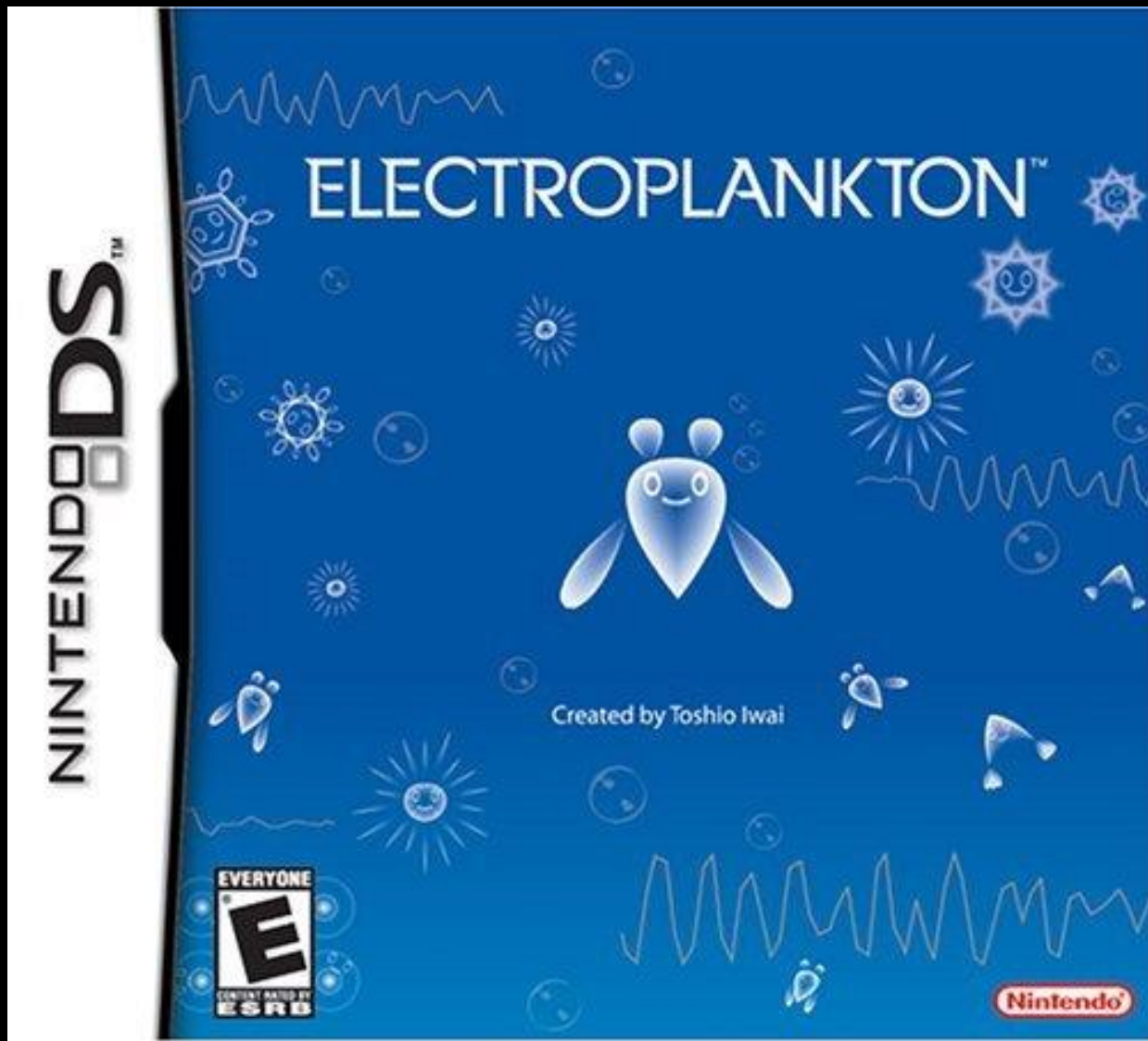
Sonomorphs: An application of genetic algorithms to the growth and development of musical organisms, Gary L. Nelson, 1993.



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



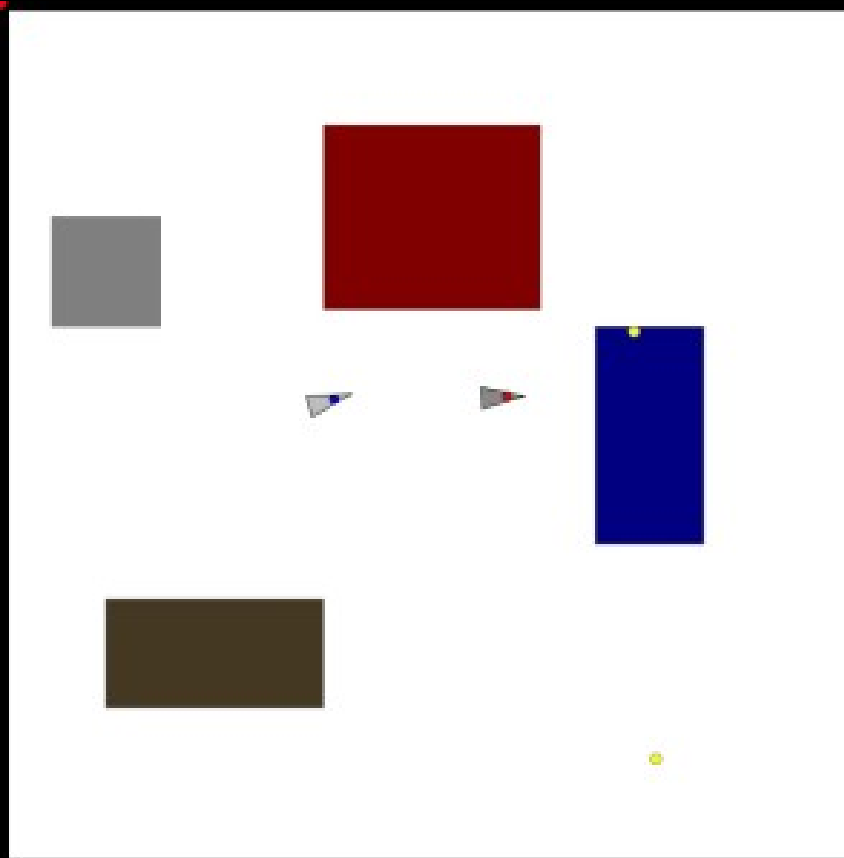
Listening Sky, Rodney Berry, Wasinee Rungsarityotin, Alan Dorin, 2001.



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** marry both approaches

Pros and Cons of musical agents

- Pros:



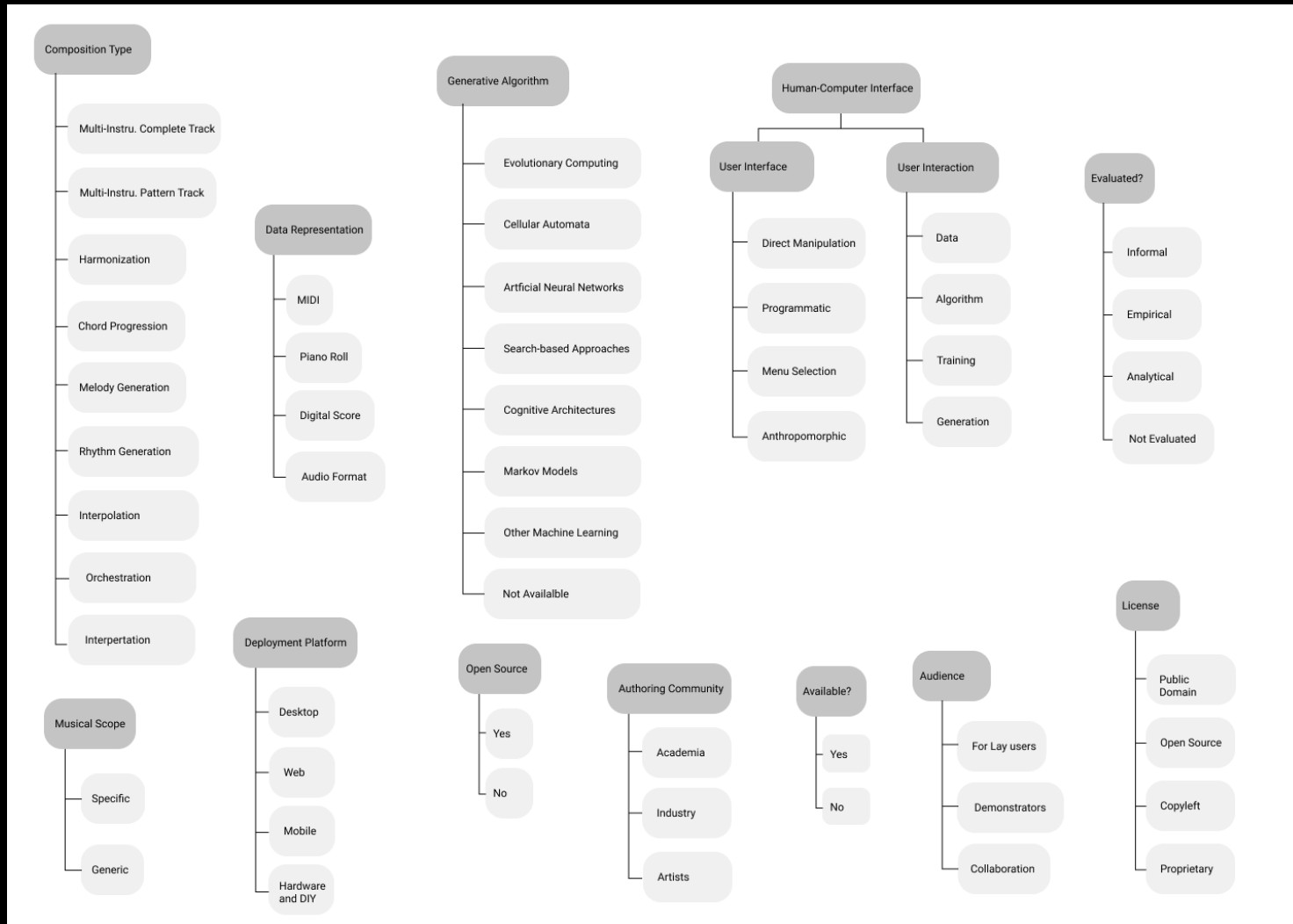
- Online
- Agents offer and anthropocentric conceptualization
- Modelling Flexibility

- 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

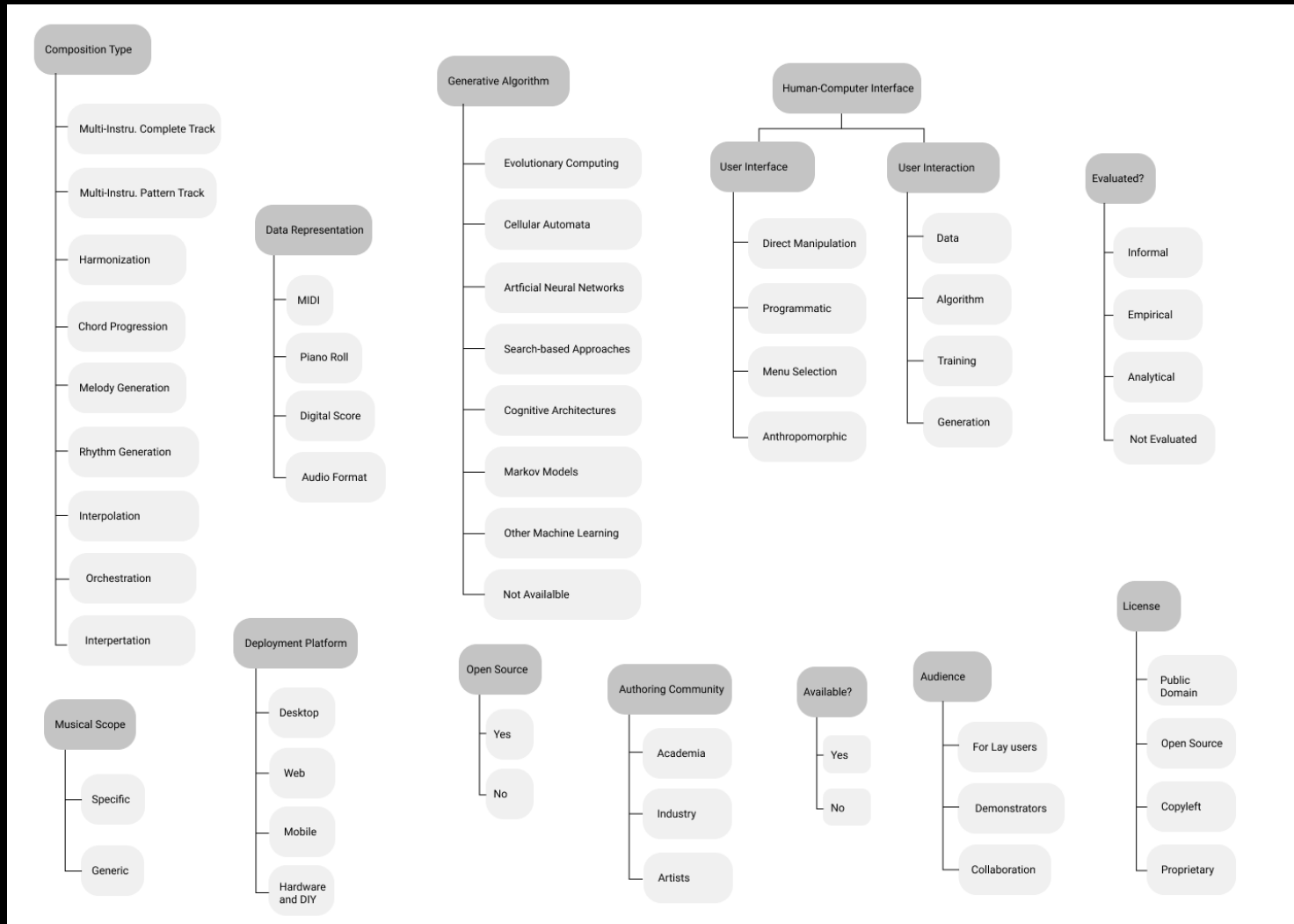
Taxonomy of CAC Systems



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	<i>Pending</i>
 Apollo	2018	Multi-Track	MIDI	Machine Learning	Amateurs + Pro	Web	<i>Pending</i>
 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](#).

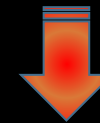
Try it at: <https://audiometaphor.ca/>

The screenshot shows the AuMe web interface. At the top, the logo 'AuMe' is displayed. Below it, a text input field contains the prompt 'A quenching rain drenched my burning head'. To the right of the input field, there is a 'Duration' slider set to 80s and a 'Length: 80s' label. Below the input field is a graph showing 'Value' on the y-axis (ranging from 0.0 to 1.0) and 'Time (s)' on the x-axis (ranging from 0 to 80). The graph displays two curves: 'Valence' (blue line) and 'Arousal' (red line). The Valence curve starts at 0.0, rises to approximately 0.5 at 40s, and then levels off at 1.0. The Arousal curve starts at 1.0, drops to approximately 0.1 at 40s, and then levels off at 0.0. Below the graph is a 'Generate Tracks' button. At the bottom of the interface, there is a multitrack sequence showing 'rain' and 'burning' tracks over time, and an audio waveform. A 'Play/Pause' button is located at the bottom center.

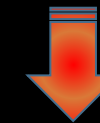
Prompt + duration



Valence and arousal
affect curves



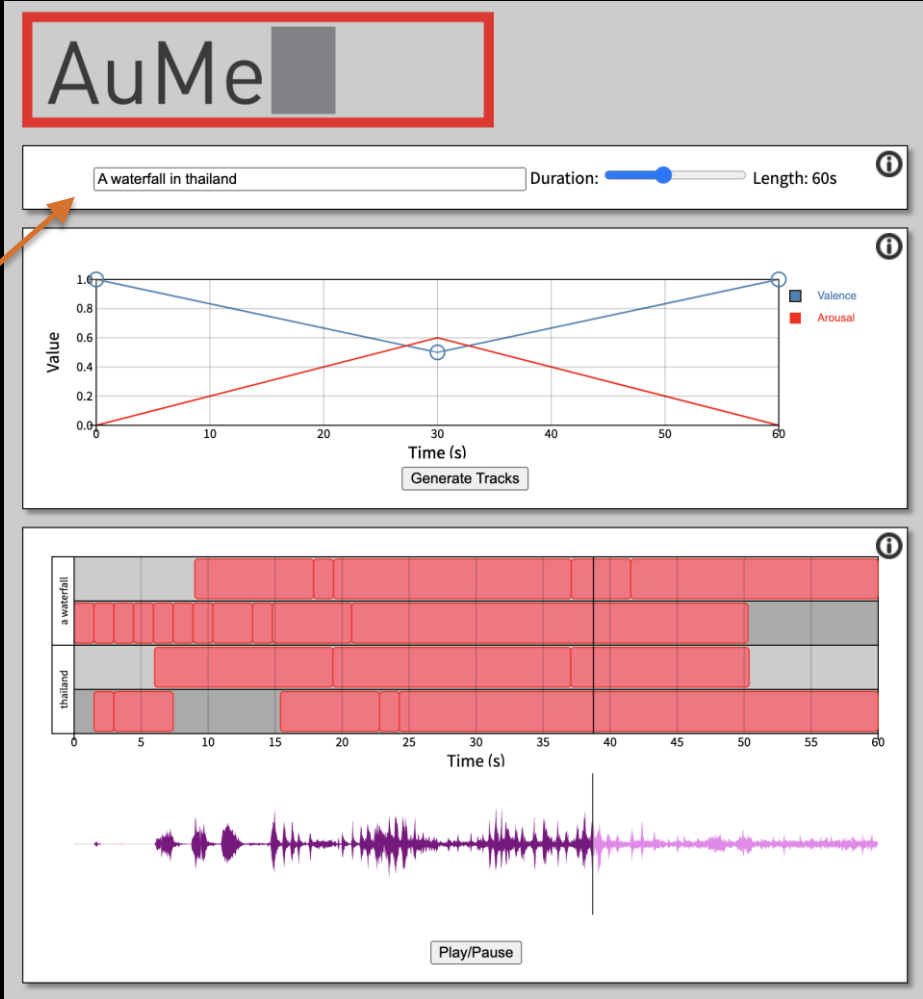
Multitrack sequence, mix (AAC)



Audio rendering (44k, stereo)

Audio Metaphor

A waterfall in Thailand



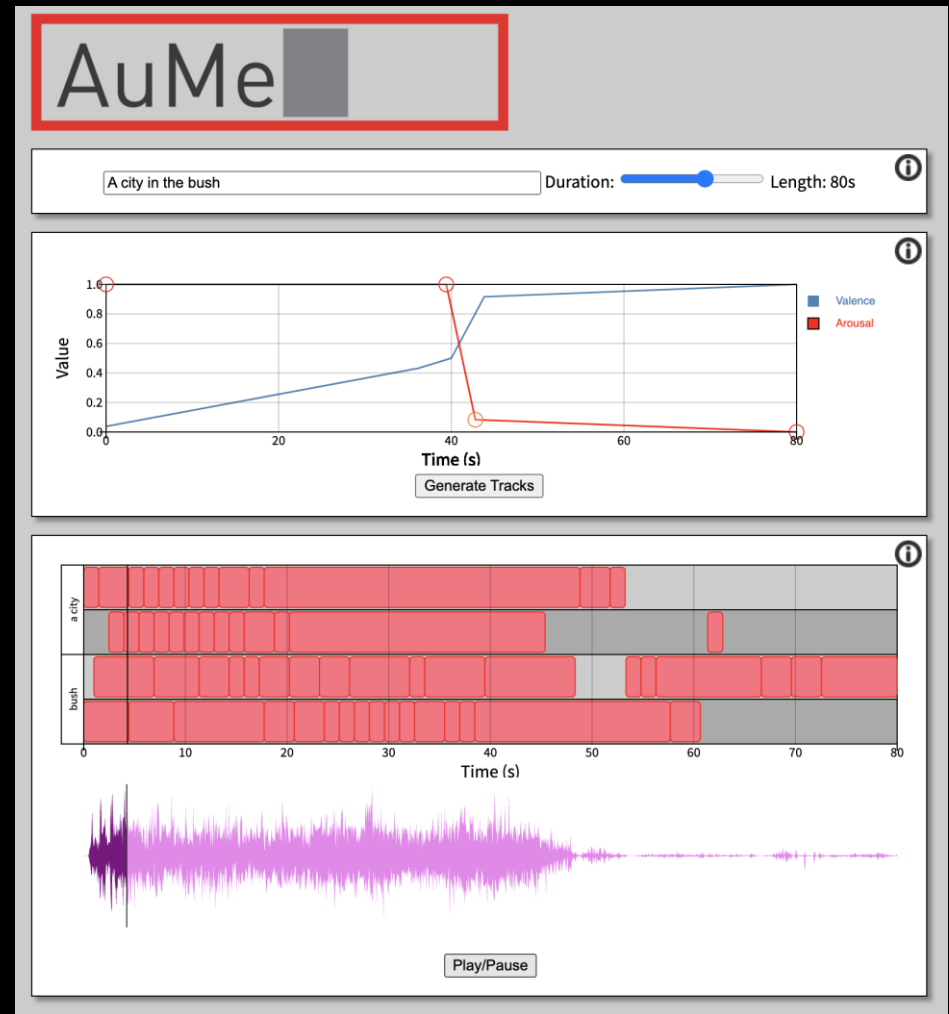
The screenshot shows the AuMe interface with the following components:

- Header:** 'AuMe' logo in a red-bordered box.
- Input:** A text field containing 'A waterfall in thailand' and a duration slider set to 60s.
- Line Graph:** A plot of Valence (blue line) and Arousal (red line) over 60 seconds. Valence starts at 1.0, dips to ~0.5 at 30s, and returns to 1.0. Arousal starts at 0.0, peaks at ~0.5 at 30s, and returns to 0.0. A 'Generate Tracks' button is below the graph.
- Heatmap:** A grid showing Valence (grey) and Arousal (red) for 'a waterfall' and 'thailand' segments over time. The 'a waterfall' segment is active from 10s to 60s, and 'thailand' is active from 0s to 50s.
- Waveform:** A purple audio waveform at the bottom with a 'Play/Pause' button.

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

Audio Metaphor

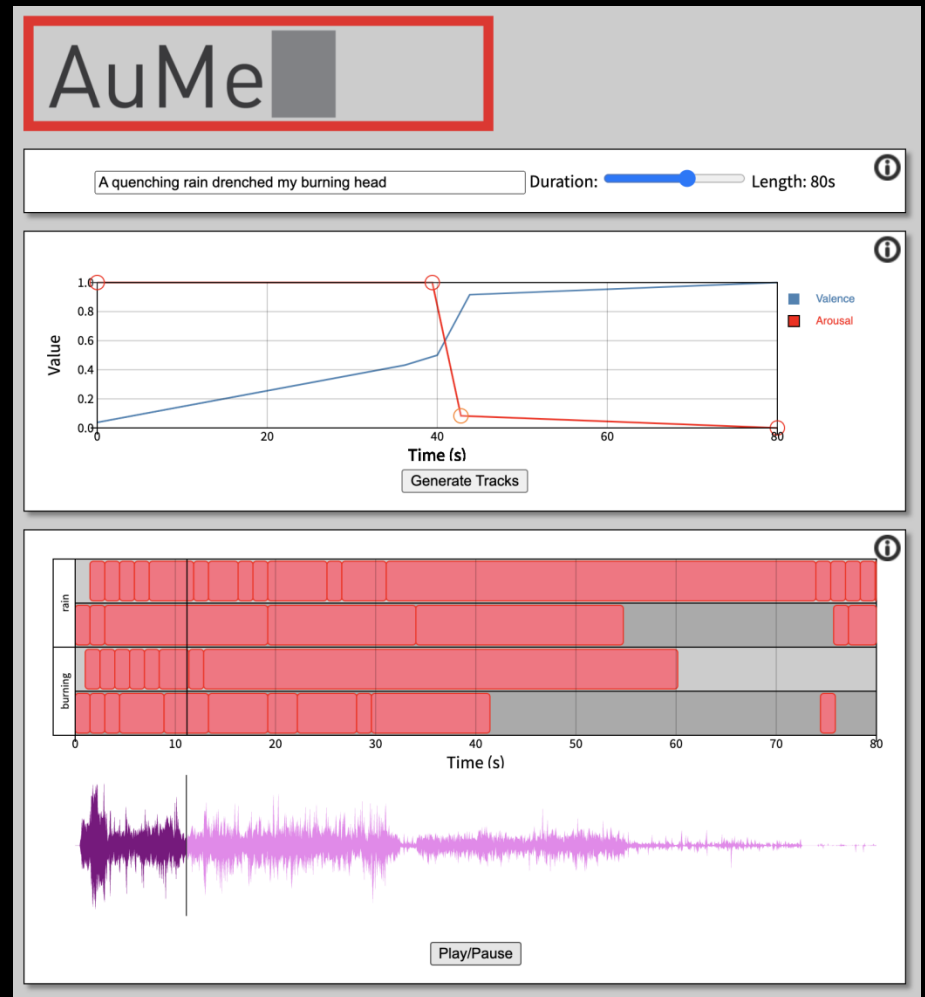
A city in the bush



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

Test

Day at the Mall

pop electronic tropical house relaxed KEY E

RENDER

BUY NOW

Keys

- gradient pad DIRECT
- pain synth pluck DIRECT
- prime synth pluck DIRECT
- ruby synth pluck DIRECT

Strings

- comb buzz synth bass DIRECT
- sub synth bass DIRECT

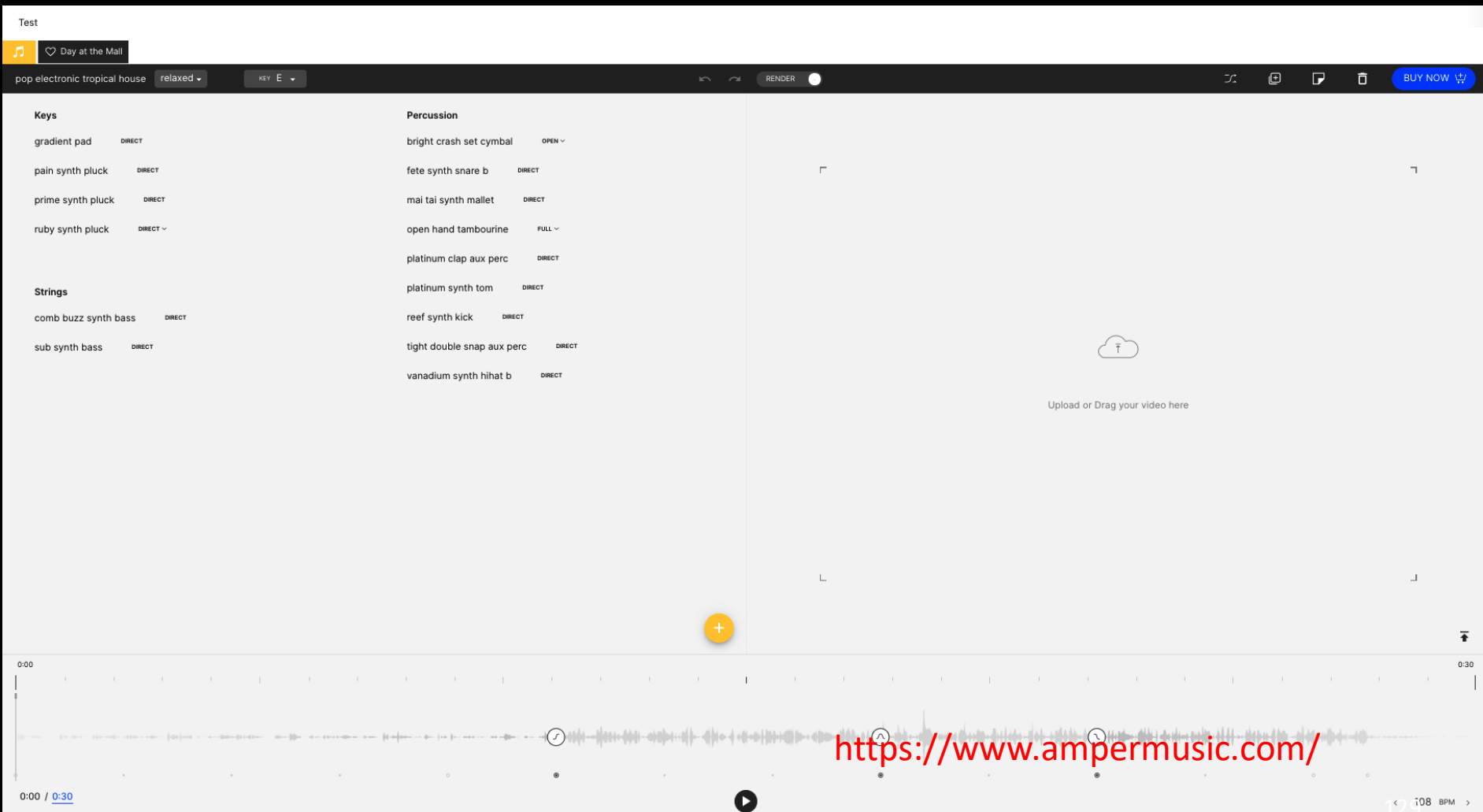
Percussion

- bright crash set cymbal OPEN
- fete synth snare b DIRECT
- mai tai synth mallet DIRECT
- open hand tambourine FULL
- platinum clap aux perc DIRECT
- platinum synth tom DIRECT
- reef synth kick DIRECT
- tight double snap aux perc DIRECT
- vanadium synth hihat b DIRECT

Upload or Drag your video here

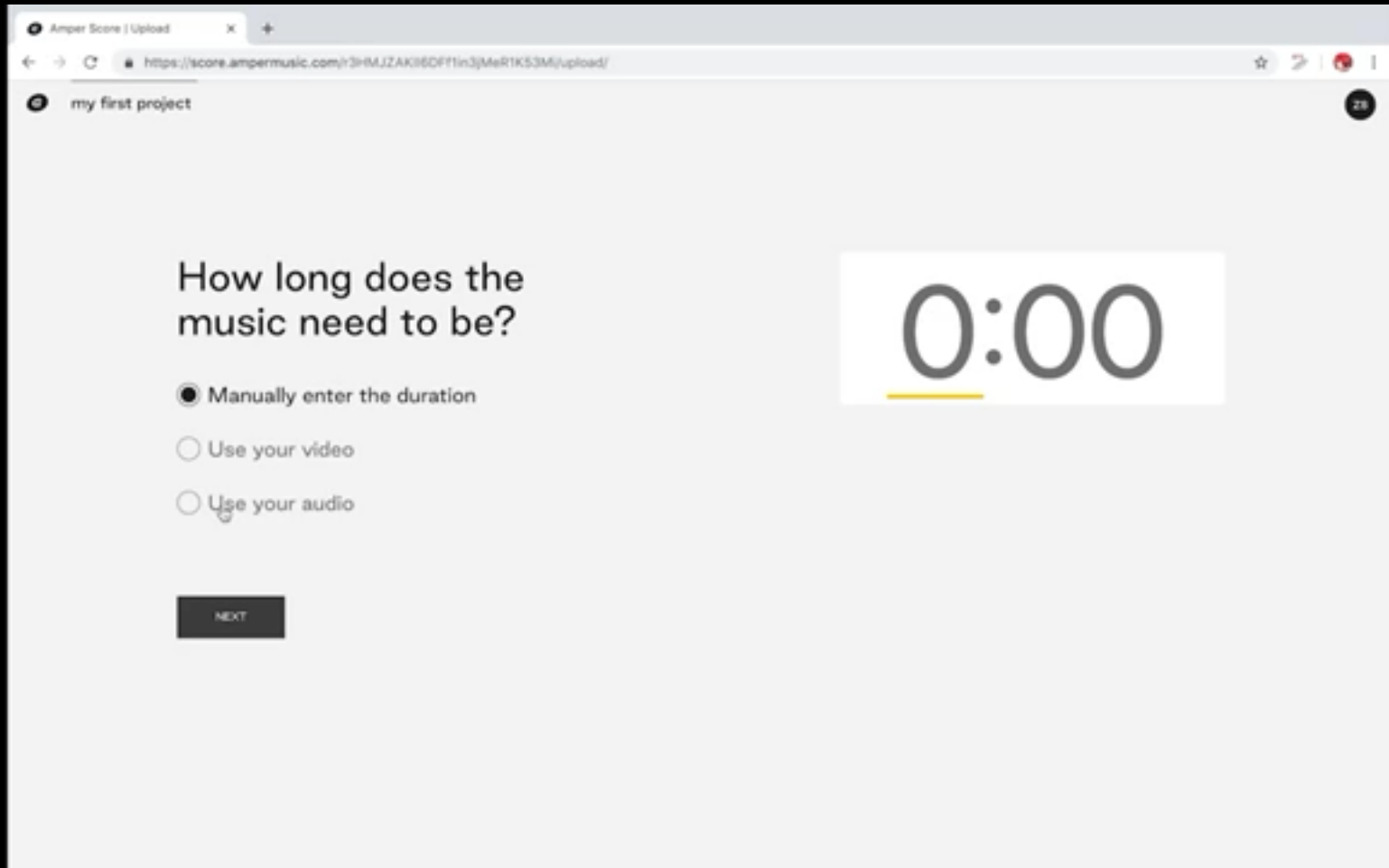
0:00 / 0:30

108 BPM



<https://www.ampermusic.com/>

Amper Score



Introducing Amper Score

Mubert

The screenshot displays the Mubert web interface for editing a stream. On the left, a sidebar contains navigation links: STREAMS, SAMPLES, USER, GUIDE, WHAT'S NEW, and COMMUNITY. The main content area is a form with the following fields and values:

- NAME: TEST
- KEY: CF
- SCALE: MINOR
- BPM: 128
- ACTIVITY: WORK X
- MOOD: ANGER
- WEATHER: ALL
- GENRE: HOUSE X, TECHNO X
- USER_TAGS: AFF_TECHHOUSE X, DEEP X
- GIPHY HASH: (empty)
- GIPHY AUTHOR: (empty)
- GIPHY LINK: (empty)

At the top right of the form, there are three buttons: ADD/EDIT SAMPLES, SAVE STREAM, and CLOSE. A tooltip above the SAVE STREAM button reads: "SAVE CHANGES IN THE STREAM AND FINISH EDITING". At the bottom of the interface, there is a white bar with a play button, a close button, the text "ANDY / MMS_FULL", and a volume control slider.

<https://beta.mubert.com/>

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

The screenshot shows the Mubert website interface in a browser window. The address bar displays "beta.mubert.com". The page is divided into several sections: "STREAMS", "SAMPLES", "USER", "GUIDE", "WHAT'S NEW", and "COMMUNITY". The "ALL STREAMS" section is active, displaying a table of music streams. A "CREATE STREAM" button is highlighted with a red circle. Below the table is a volume control bar.

	AUTHOR	NAME	BPM	KEY	SCALE	ACTIVITY	GENRE	MOOD	WEATHER
▶	A	DITCHSLAP	169	F#	MINOR	DRIVE, SPORT, SEX	ELECTRONICA	-	-
▶	A	BALKAN	118	E	MINOR	SNOKE	ELECTRONICA, BASS	SAD	RAIN
▶	A	OFFLINE	129	E	MINOR	DRIVE, WALK, SEX	TECHNO, ELECTRONICA	ANGER	RAIN
▶	KODGIK	TEST22	129	ALL	ALL	WORK	-	ALL	-
▶	ANDY	MMS_ANDY	128	D#	MINOR	ALL	ELECTRONICA	-	-
▶	ANDY	MMS_FULL	128	ALL	ALL	ALL	ELECTRONICA	-	-
▶	ANDY	COLORED	54	C	MAJOR	SEX	CHILL	-	-
▶	ANDY	C5	255	C	NONE	HIGH	HIP-HOP	-	-
▶	ANDY	8#1	159	A#	MINOR	SEX	ELECTRONICA	-	-
▶	ANDY	9#9	159	A#	MINOR	SEX	ELECTRONICA	-	-
▶	ANDY	ARPA	255	C	NONE	HIGH	HIP-HOP	-	-
▶	ANDY	CUBE DANCE BY JEAN MICHEL REICH	181	NONE	NONE	CREATE	HIP-HOP	-	-
▶	ANDY	MUSIC 4 ROBOTS PART II	255	C	MAJOR	ALL	HIP-HOP	-	-
▶	ANDY	C8	255	C	NONE	HIGH	HIP-HOP	-	-
▶	ANDY	C7	255	C	ALL	HIGH	HIP-HOP	-	-
▶	ANDY	MUSIC 4 ROBOTS PART III	255	C	MAJOR	ALL	HIP-HOP	-	-

How to Create a Generative Music Stream — Mubert Pro

Mubert

- Andy – High life



- Dancingteeth – Speed of Thought



- Ceadfra – Sweetye



- Ilberezkin – House of mub



AIVA

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

TITLE

SOURCE

INSTRUMENTATION

KEY

BPM

METER

CREATED

MODIFIED

DURATION

Start adding tracks by using the button below.

↑↑ SELECTED PRESET: JAZZ ↑↑

BACKING TRACK SONG

WITHOUT MELODY WITH MELODY

Ab Major **Slow** Jazz Ensemble < 0'30 1

KEY SIGNATURE PACING INSTRUMENTATION DURATION # OF COMPOSITIONS

Include Percussion (ensembles that support percussion) Include Tempo Variations

Auto
Slow 60 - 100 BPM
Medium 80 - 120 BPM
Fast 120 - 140 BPM

Cancel Create your track(s)

<https://creators.aiva.ai/>

New Folder

Create Track

AIVA

MY TRACKS

TITLE	PARAMETERS	DURATION	CREATION DATE
Rock Preset	-	-	Sep 13, 2019
Pop Preset	-	-	Sep 13, 2019
Jazz Preset	-	-	Sep 13, 2019
Fantasy Preset	-	-	Sep 13, 2019
Modern Cinematic Preset	-	-	Sep 13, 2019
This is a New Folder	-	-	Sep 17, 2019
This Track Has Been Renamed	Modern Cinematic, G# Minor, String Ensemble	0:44	Sep 17, 2019
New Composition #167	Sea Shanties, C Major, Small Pirate Band	2:06	Sep 17, 2019
New Composition #168	Pop, D Major, Piano & Strings	3:10	Sep 17, 2019
New Composition #169	Jazz, Bb Major, Lounge Band	2:22	Sep 17, 2019
New Composition #170	Fantasy, A Major, Symphonic Orchestra	1:08	Sep 17, 2019

New Folder Create Track

0:00 0:44 New Composition #166 Switch to Piano Roll

Introduction to Composing with AIVA

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



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



Home

Create

Library



Select style

Custom

Create without limits. Customize any style with your own composition and production instructions.

Composition

Warehouse Groove

Instruments

Acoustic Guitars

Drums

All Drums

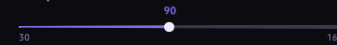
Mixing

Super Clean

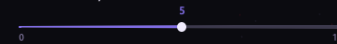
Sound Effects

Jungle

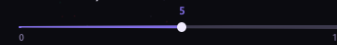
Tempo



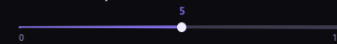
Lead Density



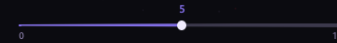
Bass Density



Chord Density



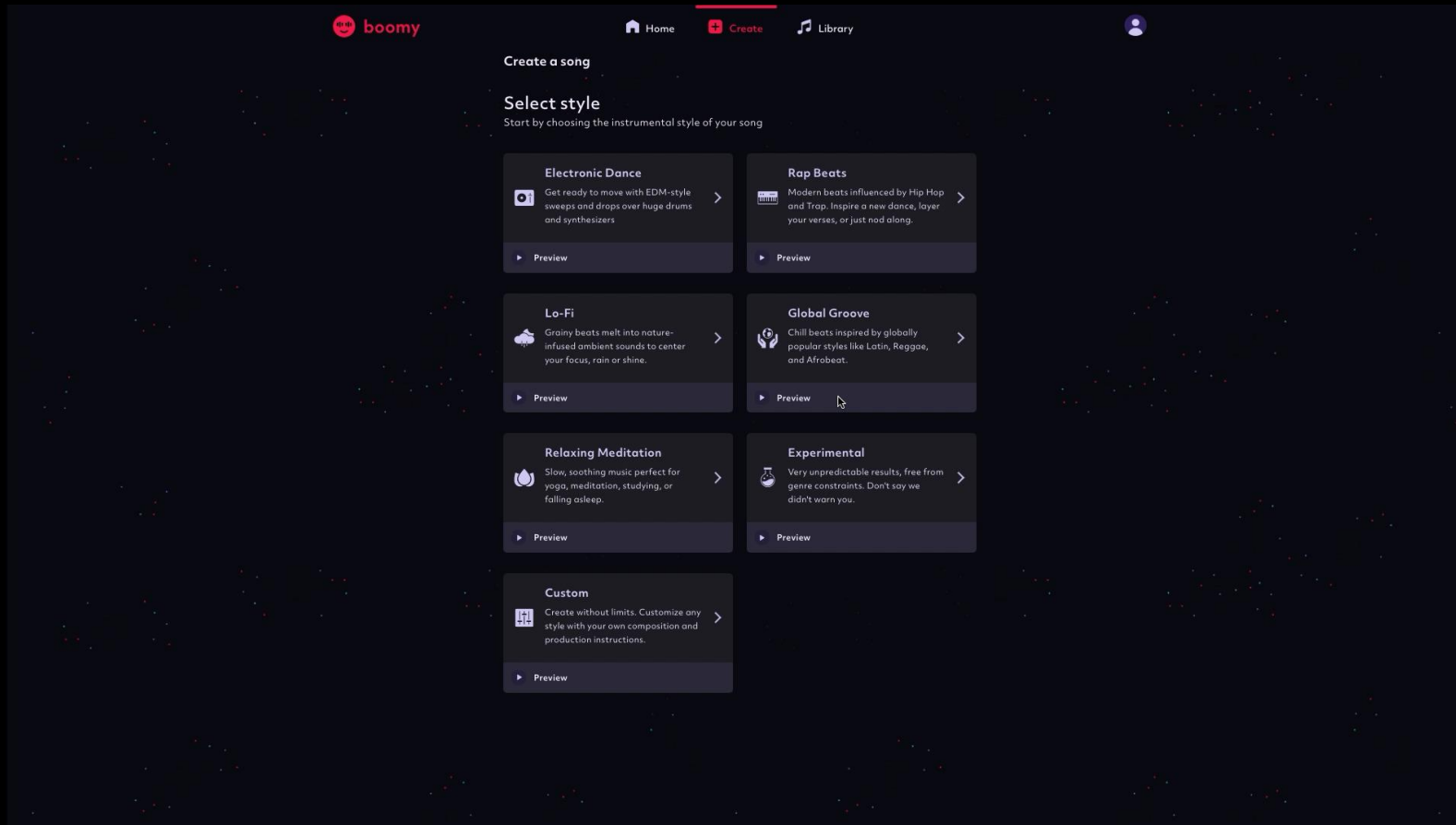
Drum Density



Create song

<https://boomy.com/style>

Boomy



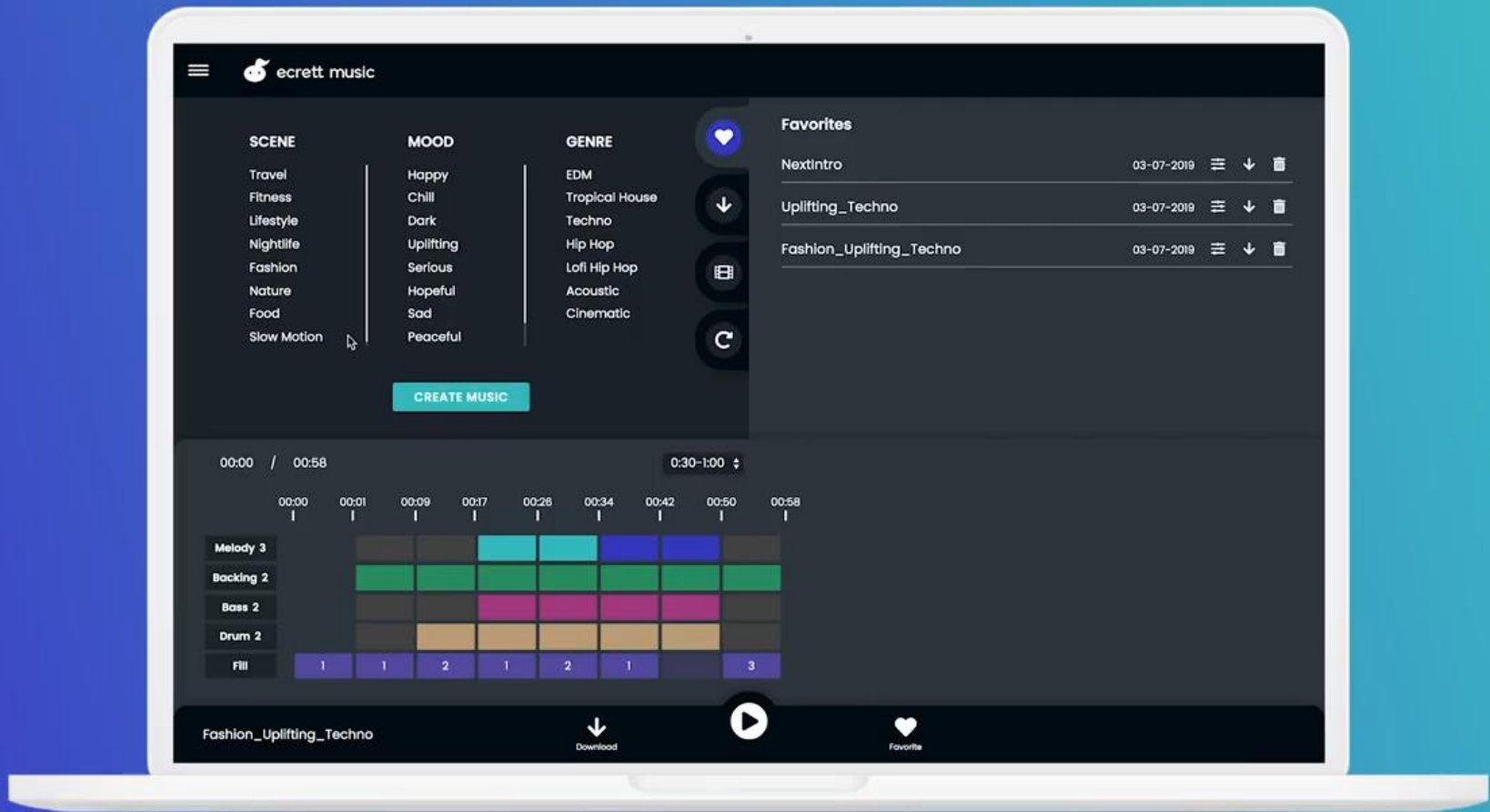
Ecrett

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

Ecrett

The screenshot shows the Ecrett Music web application interface. At the top left, there is a hamburger menu icon and the text "ecrett music". The main content area is divided into three columns for selecting music parameters: SCENE, MOOD, and GENRE. The SCENE column includes Adventure, Simulation, Puzzle, Village, Travel, Fitness (selected), and Lifestyle. The MOOD column includes Fantasy, SF, Chill (selected), Dark, Uplifting, Happy, and Serious. The GENRE column includes 8-bit, EDM, Tropical House, Techno, Hip Hop (selected), Lofi Hip Hop, and Acoustic. Below these columns are a "- clear all" button and a "CREATE MUSIC" button. On the right side, there is a "Favorites" section with a heart icon and the text "Please sign in to use this feature". Below the selection area, there is a progress bar showing "00:04 / 01:31" and a "Tempo & Volume" dropdown menu set to "0:30-1:00". The main visualization is a piano roll with a time axis from 00:00 to 01:31. The piano roll has five tracks: Melody 2 (cyan), Backing 1 (green), Bass 1 (magenta), Drum 1 (tan), and Fill (blue). The Fill track contains numerical values: 2, 1, 2, and 3. At the bottom of the piano roll, there is a "Fitness_Chill_HipHop" label, a "Download" button, a play button, and a "Favorite" button.

<https://ecrettmusic.com/>



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

[How to Create Royalty Free Music for YouTube Video | ecret music](#)

Soundraw

The screenshot displays the Soundraw web interface. At the top, there is a navigation bar with the Soundraw logo, a menu icon, and links for 'Create Music', 'Adobe Plugin', 'License', 'Pricing', 'Tutorial', 'FAQ', '日本語 / English', and a 'Subscribe' button. A '7-day free trial now!' banner is also present.

The main content area is divided into several sections:

- Video Theme:** A grid of 12 categories including VLOG, Talk BGM, Corporate, Advertisement, Games, and Jingle (marked 'Coming soon...').
- Mood:** A grid of 8 categories including Happy, Emotional, Exciting, Relaxing, Peaceful, and Dramatic.
- Length:** A dropdown menu currently set to '0:30'.
- Detail:** A panel with options for 'Select Tempo' (Slow, Normal, Fast), 'Select Instrument', and 'Select Genre'.
- Music List:** A list of 9 generated tracks, each with a play button, a waveform, and an 'Edit this Music' button. The tracks are: 1. Funk, BPM126, (0:32); 2. Cinematic, BPM100, (0:30); 3. Clap And Stomp, BPM126, (0:32); 4. Acoustic, BPM122, (0:31); 5. Ambient Piano, BPM82, (0:25); 6. Tech Wave, BPM74, (0:27); 7. Corporate, BPM112, (0:27); 8. Tokyo Night Pop, BPM103, (0:29); 9. Chill Beats, BPM80, (0:26).
- Download History:** A section for managing previously downloaded tracks.

At the bottom of the interface, there is a 'CREATE MUSIC' button and a 'Show Preview Video' button. A playback control bar is visible at the very bottom with 'Download', 'Keep', and 'Share' options.

<https://soundraw.io/play>

Soundraw

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

Soundraw

The screenshot displays the Soundraw.io interface. On the left, there are two columns of video theme thumbnails: 'Video Theme' (Web CM, Corporate, Motion graphic, Vlog) and 'Mood' (Happy, Exciting, Stylish, Emotional, Light, Dramatic, Relaxing, Peaceful). In the center, there are controls for 'Length' (set to 1:00) and 'Detail' (options: Select Tempo, Select Instrument, Select Genre). Below these are sliders for 'Length', 'BPM', 'Instruments', 'Key', and 'Volume'. A 'CREATE MUSIC' button is located below the detail controls. On the right, a 'Music List' is shown with 10 tracks, each with an 'Edit this Music' button and a heart icon. At the bottom, there is a video preview area with a filmstrip icon and a play button. A 'Live chat' button is in the bottom right corner.

MusicLM (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

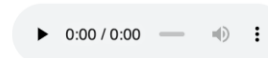
Abstract We introduce MusicLM, a model generating high-fidelity music from text descriptions such as "a calming violin melody backed by a distorted guitar riff". MusicLM casts the process of conditional music generation as a hierarchical sequence-to-sequence modeling task, and it generates music at 24 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.



SongR (2013)

The screenshot shows the SongR website interface. At the top center is the logo "SONGR BETA" with the text "Pick a genre and enter your prompt" below it. There are four genre selection boxes arranged in a 2x2 grid:

- Pop:** Features a text input field containing "Love you Gracia, you make my heart sing!" and a button labeled "I have my own lyrics".
- Hip hop:** Features a black and white image of a man rapping into a microphone.
- Cafe:** Features a black and white image of a woman singing into a microphone.
- Piano Rock:** Features a black and white image of a man playing piano and singing.

At the bottom of the interface, there is a small line of text: "By using SongR you agree to the [Terms & Conditions](#)
Built by RIFFIT © Copyright SongR 2023".

Pick a genre and enter your prompt

Pop

Hip hop

Edit Your Pop Lyrics!



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

Love you Gracia

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

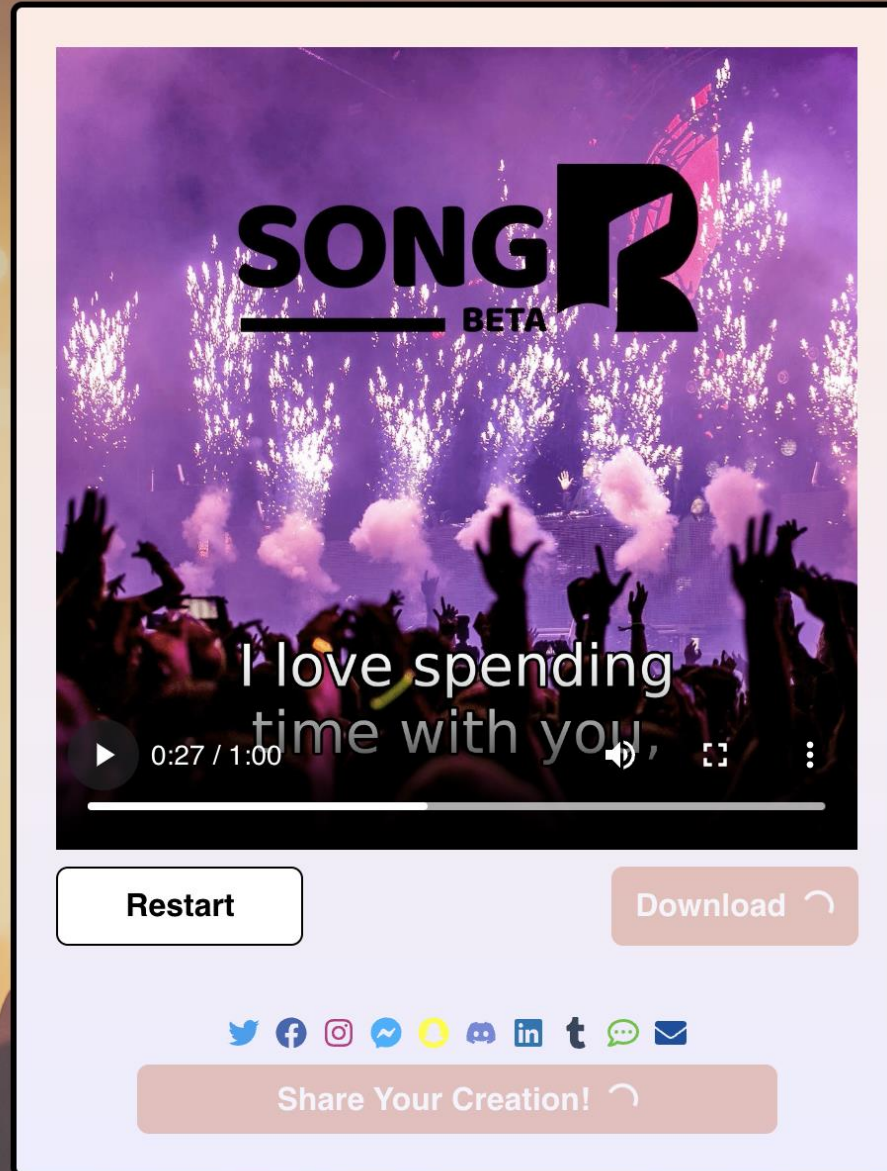
Natalie Kevin


Clear

Try Again

Render

By clicking "Render" you agree to abide by our [Terms & Conditions](#)



SONG
BETA 

I love spending
time with you,

0:27 / 1:00

Restart Download

Twitter Facebook Instagram WhatsApp Messenger LinkedIn Tumblr Email

Share Your Creation!

Riffusion (2023)

The screenshot displays the Riffusion website interface. At the top left is the Riffusion logo. A sidebar on the left contains three icons: a plus sign for 'Create', a magnifying glass for 'Explore', and a play button for 'My riffs'. The main content area features a large blue button labeled 'Generate a song' with a star icon. Below this is a search input field with the placeholder text 'Enter topic to sing about...' and a 'Riff' button. A small disclaimer below the input field reads: 'By continuing you agree to our [Terms of Service](#) and [Privacy Policy](#).' The 'Trending riffs' section is titled and contains three video thumbnails. Each thumbnail includes a play button icon, a view count, and a heart icon with a count. Below each thumbnail is a title, a user name, and a timestamp.

Generate a song ✨

Enter topic to sing about... **Riff**

By continuing you agree to our [Terms of Service](#) and [Privacy Policy](#).

Trending riffs

- T Groove of Shadows**
Tony Shaw • 5d ago
4287 views, 16 likes
- r Catchy Energy**
ringo • 3d ago
5232 views, 21 likes
- A Season of Wishes**
Amanda Martinez • 2d ago
2079 views, 22 likes

Functional music generation.

Halo: Relax, Focus, Meditate 4+

Relaxing Nature Soundscapes
Spliqs Intelligent Media Inc Apps
Designed for iPad

Free · Offers In-App Purchases

[View in Mac App Store ↗](#)

Screenshots iPad iPhone

Relaxing Audio That Adapts to You

Sustain Focus and Relaxed Creativity

Feel Better Quickly

Sleep Relax Meditate Focus

AI Explore

Sleep Relaxation Meditation Creative Focus

SETTINGS SOUNDSCAPE

BINEURAL BEATS 50%

CLOUDS 50%

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

Computer-assisted composition MIDI/Symbolic generation

CAC Systems examples

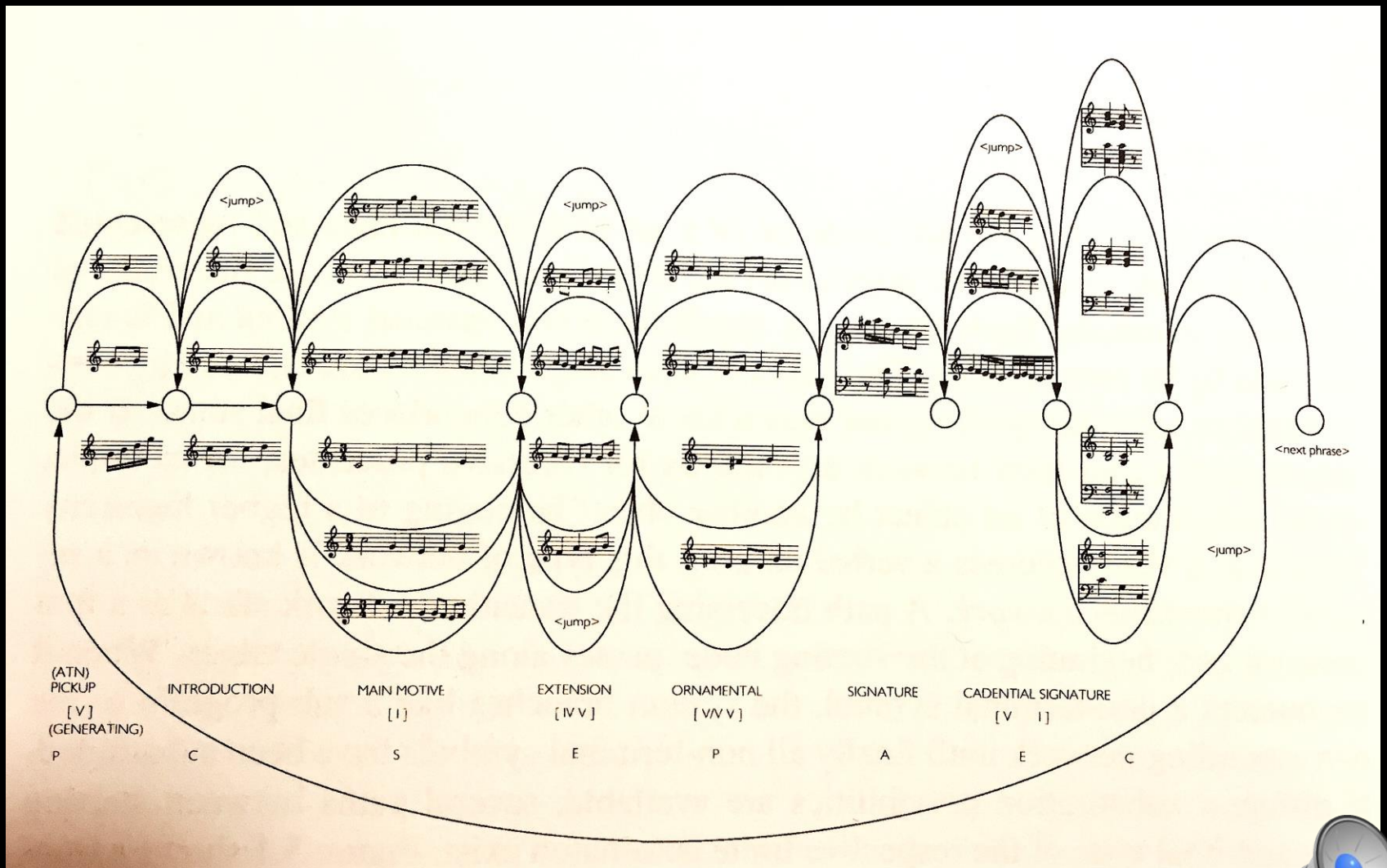
- **Audio output:**

- [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

The screenshot displays the Impro-Visor software interface for a 12-Bar Blues piece by Clifford Brown. The window title is "Impro-Visor: 12-Bar Blues". The menu bar includes File, Edit, Transpose, View, Play, Utilities, Window, Grammar: My, Preferences, and Help. The toolbar contains icons for file operations, playback, and generation. The playback controls show a location of 0:00 to 3:12, a looping button set to 2, a volume slider, a tempo of 180.0 BPM, and other settings like Transpose (0), Bars per Chorus (12), Tracker Delay (0), and Parallax (0). A "Textual Entry" field is present with a "Clear" button. Below the menu bar, a list of composers is shown: Clifford Brown, Dizzy Gillespie, Freddie Hubbard, Lee Morgan, Miles Davis, Tom Harrell, Bill Evans, Red Garland, and Charlie Parker. The main score area is titled "12-Bar Blues" and "Generated from grammars learned from solos of different players". It shows a musical score in 4/4 time with a key signature of one flat (Bb). The score is divided into three lines of four bars each. The first line (bars 1-4) has chords F13, Bb13, Bo7, and F13. The second line (bars 5-8) has chords Bb13, Bo7, F13, and D7#5#9. The third line (bars 9-12) has chords Gm9, C13b9, F13, D7#5#9, Gm9, and C13b9. A red vertical line is positioned at the start of the first bar, and a blue vertical line is at the start of the eighth bar. A "Style: swing" label is visible above the first bar.

Bob Keller et al. (2005-2020)

Impro-Visor

The screenshot displays the Impro-Visor software interface. At the top, there is a menu bar with options like File, Edit, Transpose, View, Play, Utilities, Window, Grammar, My Preferences, and Help. Below the menu is a toolbar with various icons for file operations and playback. A control panel includes a 'Play' button, a 'Stop' button, a 'Playback Location' indicator (0:00 to 0:12), a 'Looping' indicator (2), a 'Volume' slider, a 'Tempo (Beats per Minute)' slider (160.0), and buttons for 'Transpose', 'Solo/Chorus', 'Tasteful Delay', and 'Party Book'. The main workspace shows a 'Festival Entry' field and a 'Chorus 1' section. The musical score is displayed in two staves, featuring a jazz chord structure: Cm, Cm/Bb, Ab7, G7, Cm, Cm/Bb, Ab7, G7. The solo is written in 4/4 time, with notes and rests colored in green and blue. The first staff contains measures 1-4, and the second staff contains measures 5-8. The chords are labeled above the notes.

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**

Harmonic Progression Generator

The screenshot shows the 'Multi-Order Markov Process Chord Sequence Generation' web application. It is divided into several sections:

- Multi-Order Markov Process Chord Sequence Generation**: Includes the user's name 'arne eigenfeldt', a timestamp 'May 29 2012 00:37:39 AM', and an 'update' button.
- Analysis**: Features a 'clear database' button, a 'Song' dropdown menu set to 'Jobin_1', and statistics for '# of chains', '# of links', and '# of unique links' across '1st', '2nd', and '3rd' orders. The 'Markov Chain Length' is set to '3rd Order'.
- Targets**: Contains three sliders for 'bassline', 'complexity', and 'tension', each with a 'weight' bar on the right. A 'Clear / Reset' button is at the bottom.
- Generate Phrase**: Includes a 'Generate Phrase' button, a 'Bars' dropdown set to '16', a '# of chords' slider (8-16), and progress indicators for 'attempted: 11' and 'actual: 12'. A 'Progress' bar shows '1st chain'.
- Generated Vectors**: Three empty boxes for 'bassline', 'complexity', and 'tension'.
- Generated Chord Sequence**: A large empty box with a 'display' button and a vertical scrollbar.
- Playback**: Includes a speaker icon, 'Tempo' (120), 'Harmonic Rhythm' (4), 'Volumes' for piano, bass, and click, and playback controls like '<<', 'Play', and 'view Play'. A 'Bar Beat Tick' display shows '1 | 1 | 0.'.

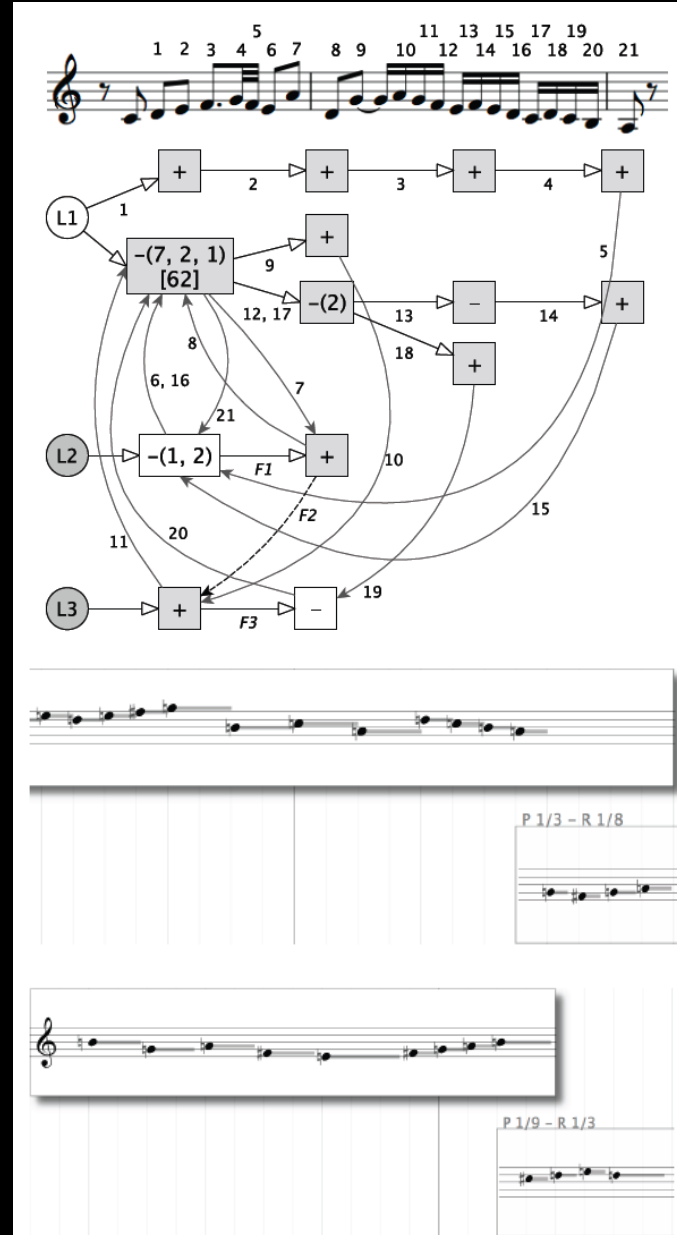
- It is doing **style imitation**, at human-competitive levels.
- The system is available online (**open source** and **free!**).
- It has been used by composers and the company **Teenage Engineering** (Sweden) when producing the Absolut Blank iPhone app.



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.

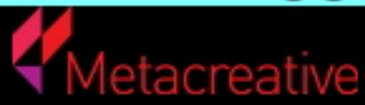
The screenshot displays a DAW interface with a piano roll for 'StyleMachine-Lite'. The piano roll shows a grid of notes across various instrument tracks. The Master track is highlighted in green. Below the piano roll are mixer controls for each instrument. A 'Help View' window is open in the bottom left, showing 'StyleMachine-Lite.maxpat' with a 'StyleMachine Lite' interface. This interface includes sliders for 'corpus', 'complexity', 'density', and 'length', a 'breakbeat (2001-06)' dropdown, a 'groove' dropdown set to 'Breaks - Brick Wall', and buttons for 'generate phrase' and 'generate all'. A 'Drop Files and Devices Here' area is visible on the right side of the piano roll.

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

Devices

StyleMachine-Lite.maxpat

StyleMachine Lite



corpus recalculating... complexity density length

breakbeat (2001-06)

clear

groove

Breaks - Brick Wall

generate phrase

done!
generate all

- Mbeat
- Ax1
- Ax2
- Bass
- RSyn
- MSyn
- Pads
- Keys
- Drone
- AxI

ss to
p-by-step
to learn

Drop an Inst

Style Machine

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

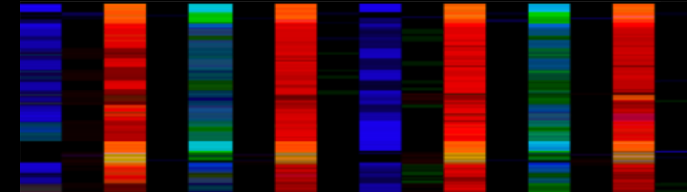
- Generative EDM

- How?

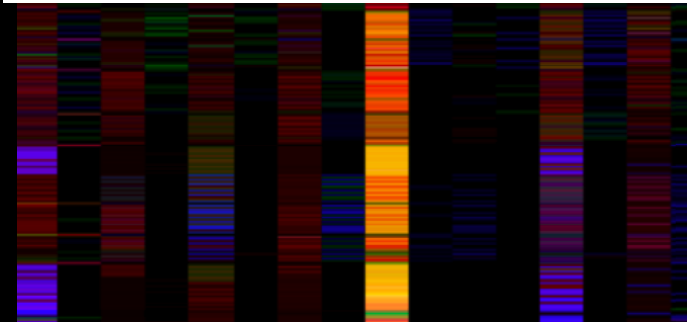
- **Manual analysis** of corpus by experts (composers, producers)
- **Our machine learning** algorithm
Genetic algorithm / VOMM

- Validation: ongoing!

- Confuses classifiers: pieces gets classified properly!
- Confuses listeners
- Public shows: Algorave ISEA, ...
- Album on ChordPunch (UK)



Three staves of musical notation in 4/4 time. The top staff is in treble clef, and the bottom two are in bass clef. The notation includes various note values, rests, and dynamic markings.



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



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

FolkRNN

folkRNN
generate a folk tune with a recurrent neural network

generate TUNE

Compose

MODEL
thesession.org (w/ :|:)

TEMPERATURE
1 624159

METER
4/4 C Major

INITIAL ABC
C2AB

FOLK RNN TUNE №70935

X:70935
M:4/4
K:Cmaj
C2EG cGEC | A, D02 F0E0 | C2EG cede | BGD F ECD |
C2EG cGEC | A, D02 F0A F | C2EG cGEG | AcBd c4 |
e2 | (3ede ecGc | e4 | (3ddd dcBd | eaeg ageg | AcBd cede |
e3e g0e0 | Ad | (3ddd adfd | e2ce dcBc | Adcb A2EG |

The RNN properties were thesession_ with _repeats with seed 954375 and temperature 1.
The prime ticks were M:4/4 K:Cmaj C 2 E G.
Generated on 10/21/2021, 4:02:46 PM.

HEAR IT

0:00 120 Download MIDI

SEE IT

<https://folkrrnn.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

The screenshot shows the folkRNN website in a browser window. The browser address bar shows 'folkRNN.org'. The page has a header with the title 'Folk RNN - Generate folk tunes with a recurrent neural network' and a sub-header 'The Machine Folk Session - An archive of folk music co-composed with machines'. On the left side, there is a form for generating a tune. The form includes a 'Compose' button, a 'MODEL' field with 'thesession.org (w/ d l :)', a 'TEMPERATURE' field with a slider set to 1 and a 'SEED' field with '528575', a 'METER' field with '4/4' and a 'MODE' field with 'C Major', and an 'INITIAL ABC' field with the text 'Enter start of tune in ABC notation'. On the right side, there is an 'ABOUT FOLK RNN' section. It contains a paragraph explaining the website's purpose, a quote from an article about AI in music, and a section for 'FREQUENTLY ASKED QUESTIONS' with a sub-section 'HOW MIGHT I CO-CREATE WITH FOLK-RNN?'. A yellow starburst badge in the top right corner of the content area reads 'Derri Lewis folkRNN composition competition winner'. The browser window also shows standard navigation buttons and window controls.

folkRNN
generate a folk tune with a recurrent neural network

PRESS TO GENERATE TUNE
Compose

MODEL
thesession.org (w/ d l :)

TEMPERATURE SEED
1 528575

METER MODE
4/4 C Major

INITIAL ABC
Enter start of tune in ABC notation

ABOUT FOLK RNN

This website lets you generate music using an artificial intelligence called a “recurrent neural network” (RNN). It’s called “folk-rnn” because the RNN is trained on transcriptions of folk music. Each press of the ‘compose’ button will create a new tune, shaped by your initial input. For example, raising ‘temperature’ will make the algorithm more adventurous. Or if a generated tune has a feature you like, you can copy that back into the ‘Initial ABC’ field and generate new tunes led by that feature.

Folk music is part of a rich cultural context that stretches back into the past, encompassing the real and the mythical, bound to the traditions of the culture in which it arises. Artificial intelligence, on the other hand, has no culture, no traditions. But it has shown great ability: beating grand masters at chess and Go, for example, or demonstrating uncanny wordplay skills when IBM Watson beat human competitors at Jeopardy. Could the power of AI be put to use to create music?
— ‘Machine folk’ music composed by AI shows technology’s creative side. *The Conversation*, March 2017

Why do this? As that article goes on to say, the original [folk-rnn](#) was developed, and its developers composed music using its successes and failures. This website aims to make that possible for everyone. It’s a tool anyone can use.

FREQUENTLY ASKED QUESTIONS

HOW MIGHT I CO-CREATE WITH FOLK-RNN?

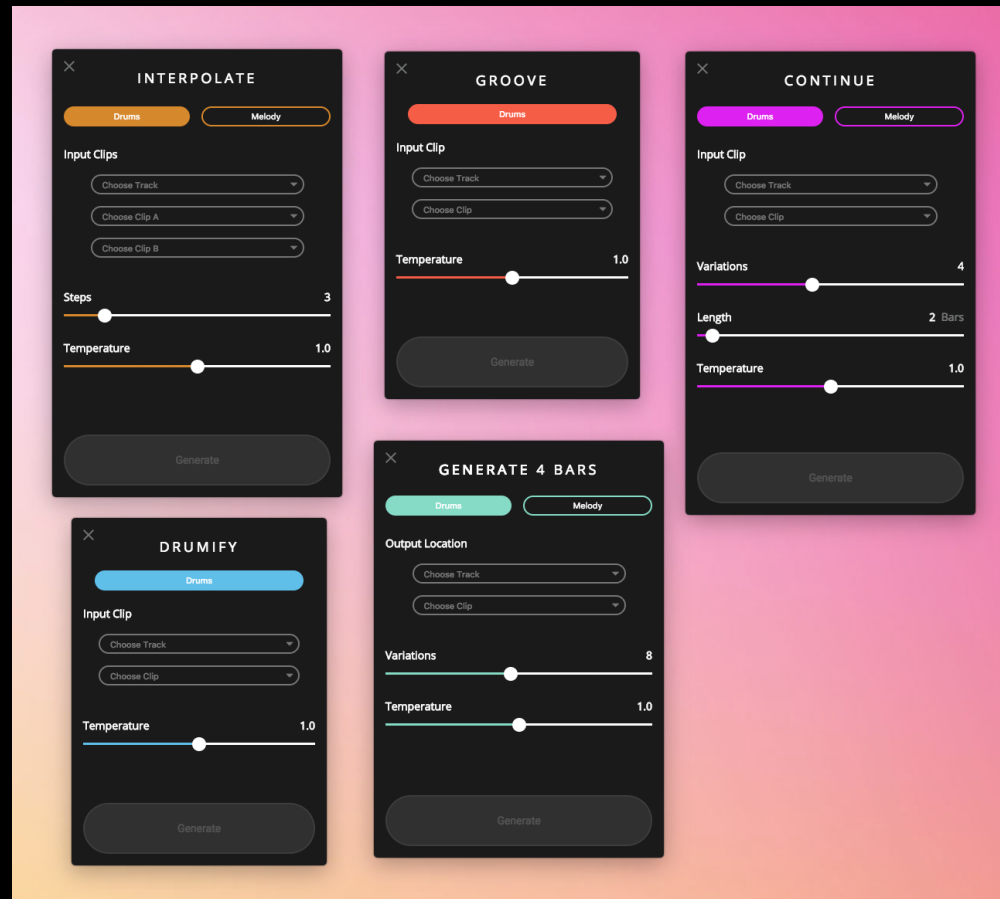
To get started, you might want to simply download a generated tune and import it into your composition app of choice. For each generated tune this site exports MIDI. The downloaded files have successfully been imported into e.g. Logic. It’s worth noting that site is not, and never will be, a composition app where you can then hand-edit the tunes generated by folk-rnn. That’s already well served elsewhere.

Explore the generation parameters. The ‘about’ section mentioned raising ‘temperature’. 1.0 is normal, 2.0 is more wild, and 0.5 more cautious. It also mentioned copying back into the ‘Initial ABC’ field features in the generated tune you like: to make this easier clicking on notes

*Derri Lewis
folkRNN
composition
competition
winner*

folkRNN.org [2018]

Magenta Studio



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

Magenta Studio

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

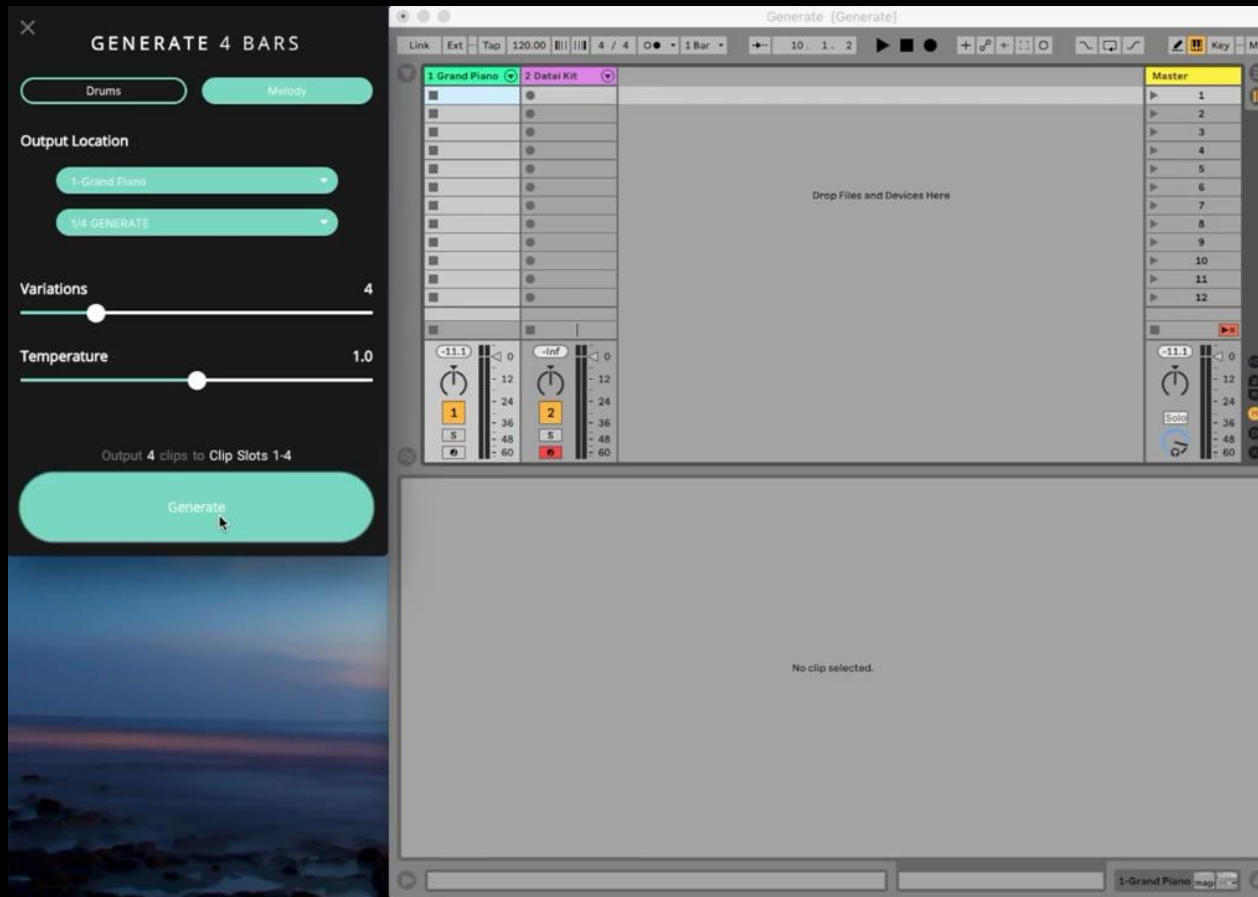
Roberts A, Engel J, Mann Y, Gillick J, Kayacik C, Nørly S, Dinculescu M, Radebaugh C, Hawthorne C, Eck D. Magenta studio: Augmenting creativity with deep learning in Ableton live.

Magenta Studio

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

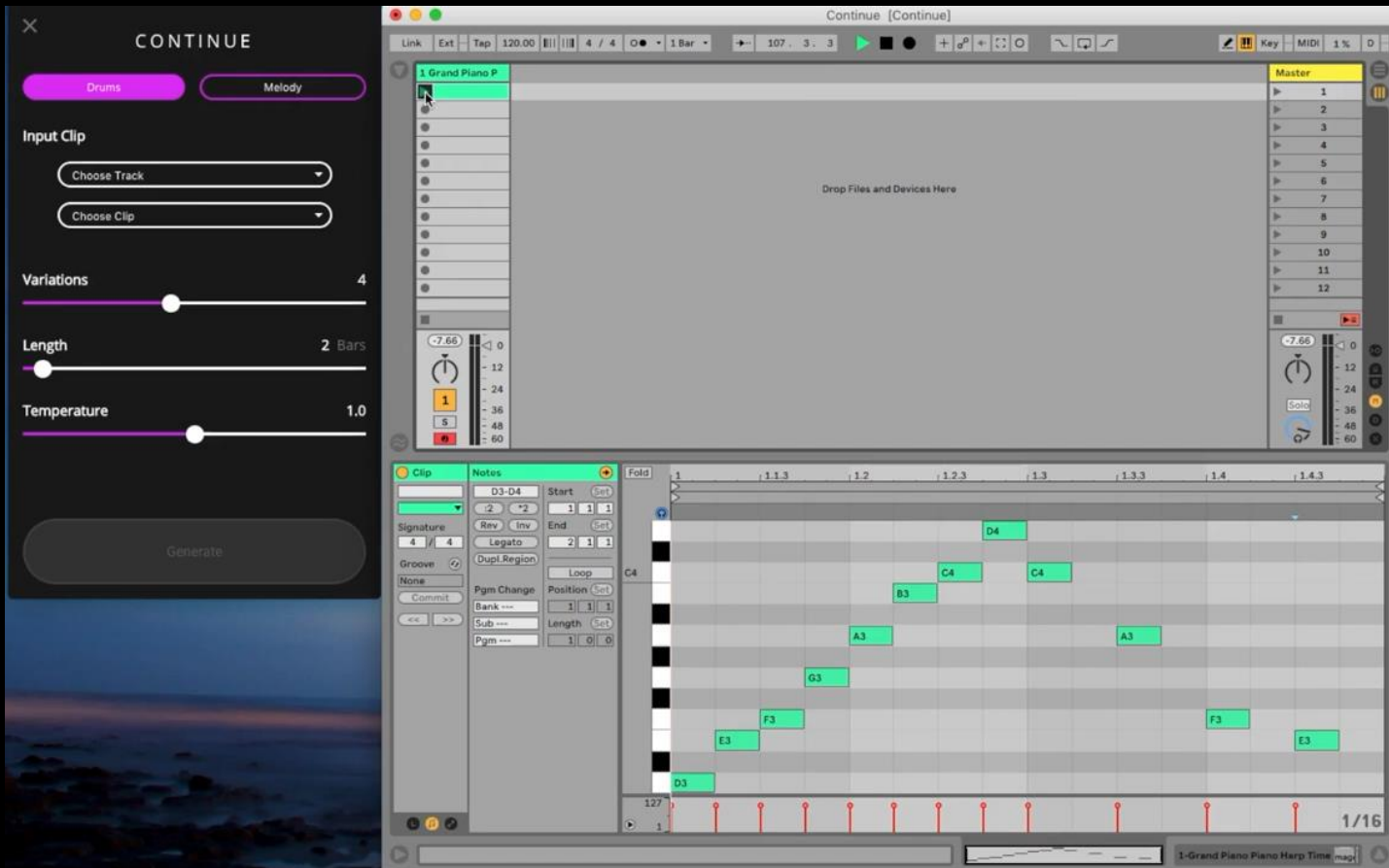


Magenta Studio



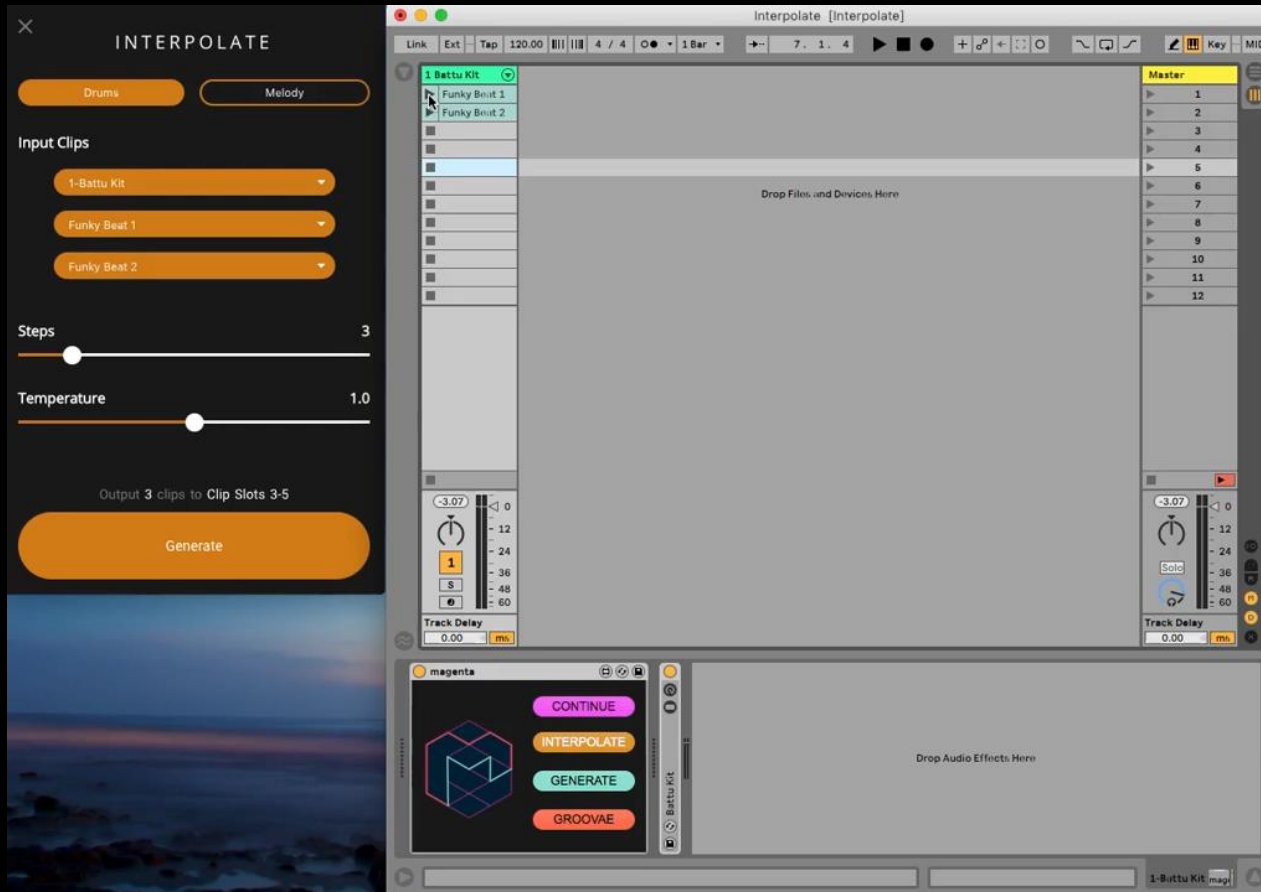
Magenta Studio: Generate Plugin

Magenta Studio



Magenta Studio: Continue Plugin

Magenta Studio



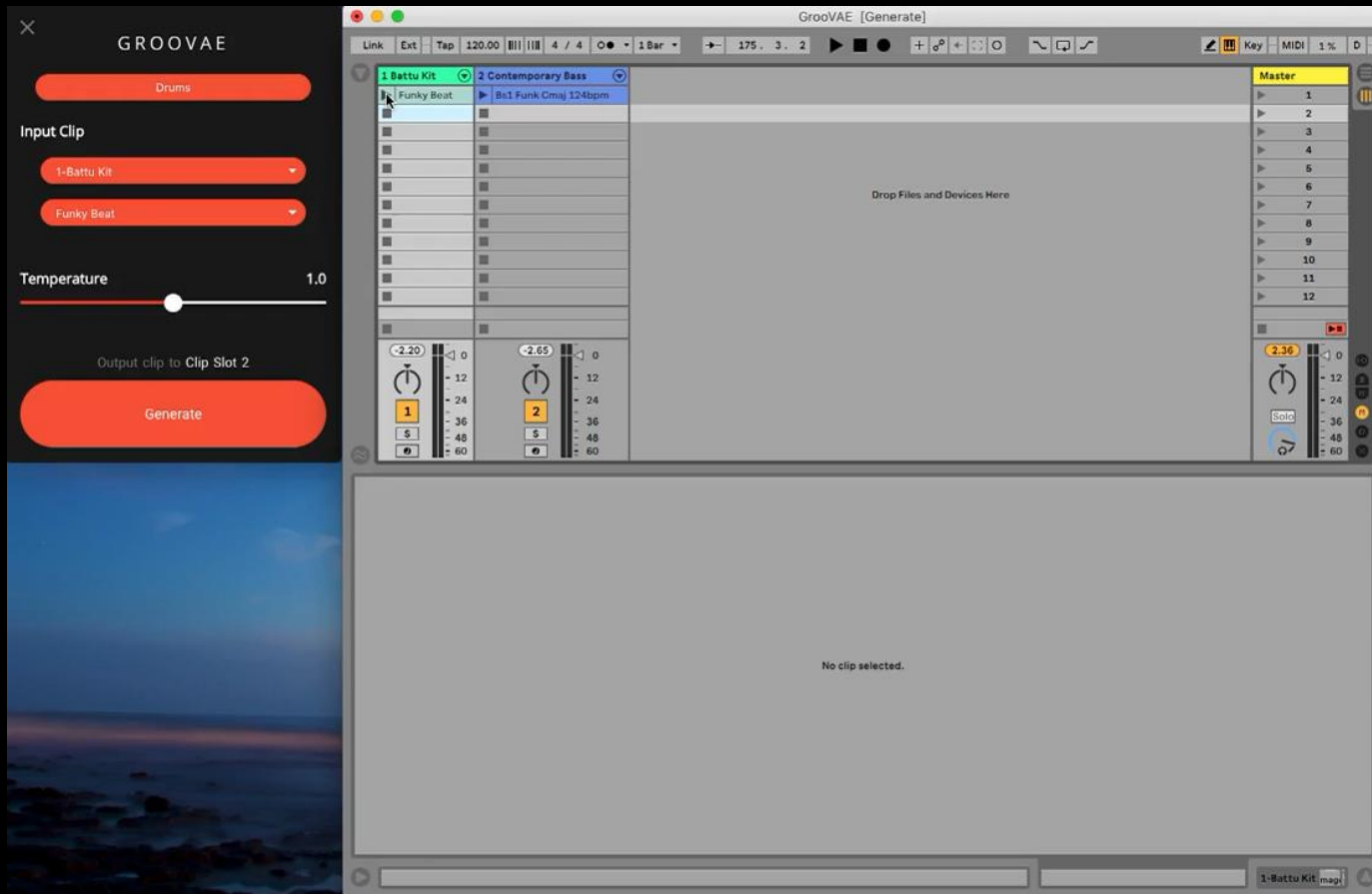
Magenta Studio: Interpolate Plugin

Magenta Studio



Magenta Studio: Drumify Plugin (Bass2Drum)

Magenta Studio



Magenta Studio: Groove Plugin

MuseNet

The screenshot displays the MuseNet web interface with the following settings:

- STYLE:** CHOPIN
- INTRO:** MOZART'S RONDO ALLA TURCA
- INSTRUMENTS:** PIANO (selected), STRINGS, WINDS, DRUMS, HARP, GUITAR, BASS
- NUMBER OF TOKENS:** 275 (indicated by a slider)

Below the settings is a section labeled "HIDE ADVANCED SETTINGS" and a visualization of the generated audio waveform. At the bottom, there are playback controls including "PLAY FROM START", "DOWNLOAD", "TWEET", and "RESET".

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

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 kind of music MuseNet generates. During training, we chose composer and

Flow Machines

The screenshot displays the Flow Machines app interface. At the top, it shows the time 15:59 on 9月2日 (木) and a battery level of 74%. The main header includes 'StylePalette : J-POP', 'J-002', and a 'Compose' button. Below this are navigation controls for '94' and 'G Maj', and a 'FLOW MACHINES' logo with 'ver 1.0.0 (202108271630)'. The central area features a piano roll with four tracks: GM7, F#7b13, Bm9, and D7. The piano roll shows red blocks representing notes and chords over time, with handwritten annotations in grey. To the left of the piano roll are controls for 'Harmony 1.0', 'Duration 0.5', 'Complex 0', and a list of songs (song 12-15) with an 'Edit' button. Below the piano roll is a 'Progression' section with buttons for GM7, F#7b13, Bm9, and D7, along with an 'Octave' control (-1, 0, 1). At the bottom, there is a 'CHORD' section with a 'Piano' instrument selection and a grid of chord buttons including Roman numerals (I, ii, iii, IV, V, vi, vii(dim)) and specific chords (G, Am, Bm, C, D, Em, F#dim, GM7, G7, G6, Gsus4, Gm). A 'MP3' toggle is also visible.

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

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 \(2016\)](#)
- [SKYGGGE feat Kiesza, Stromae and The Bionix - Hello Shadow \(Hello World Album, 2018\)](#)

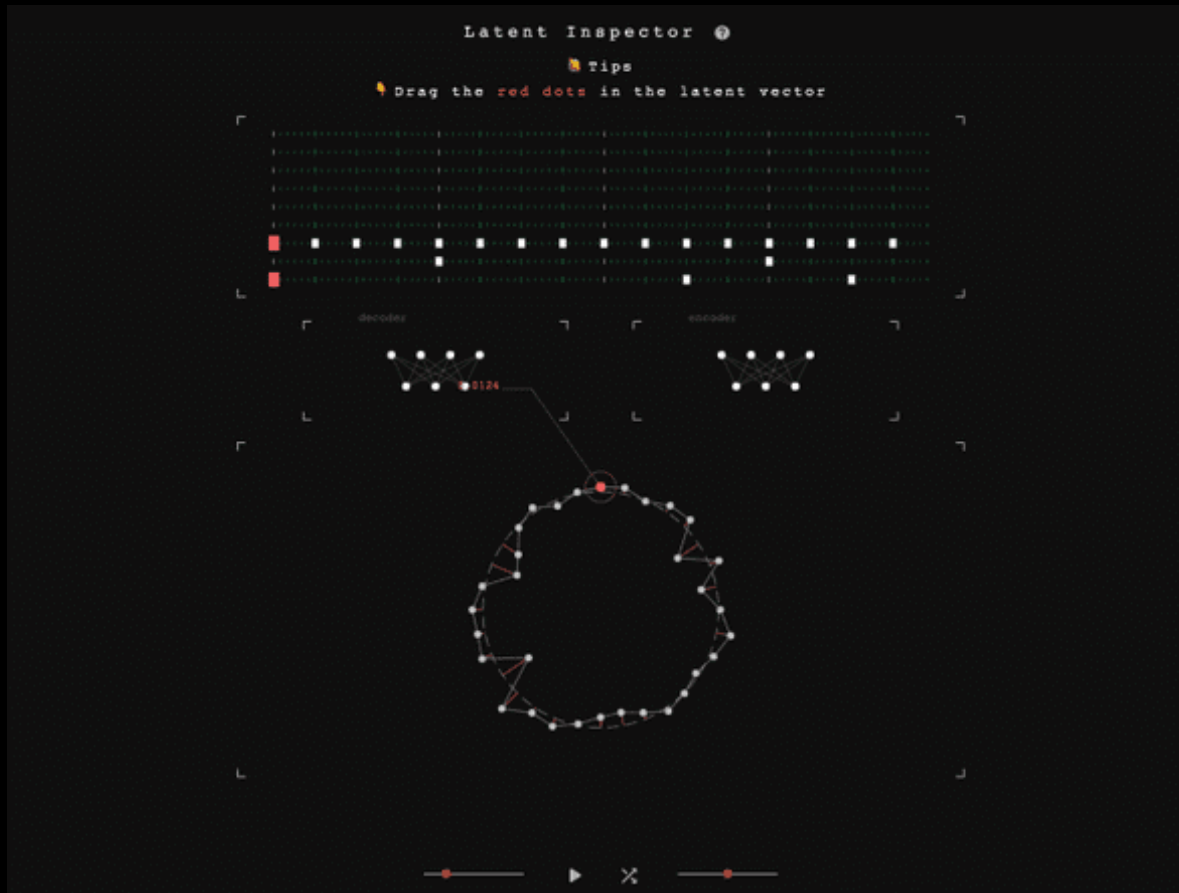


Flow Machines (2021)

FLOW
MACHINES

[Flow Machines Mobile Walkthrough | Flow Machines \(English\)](#)

DrumVAE



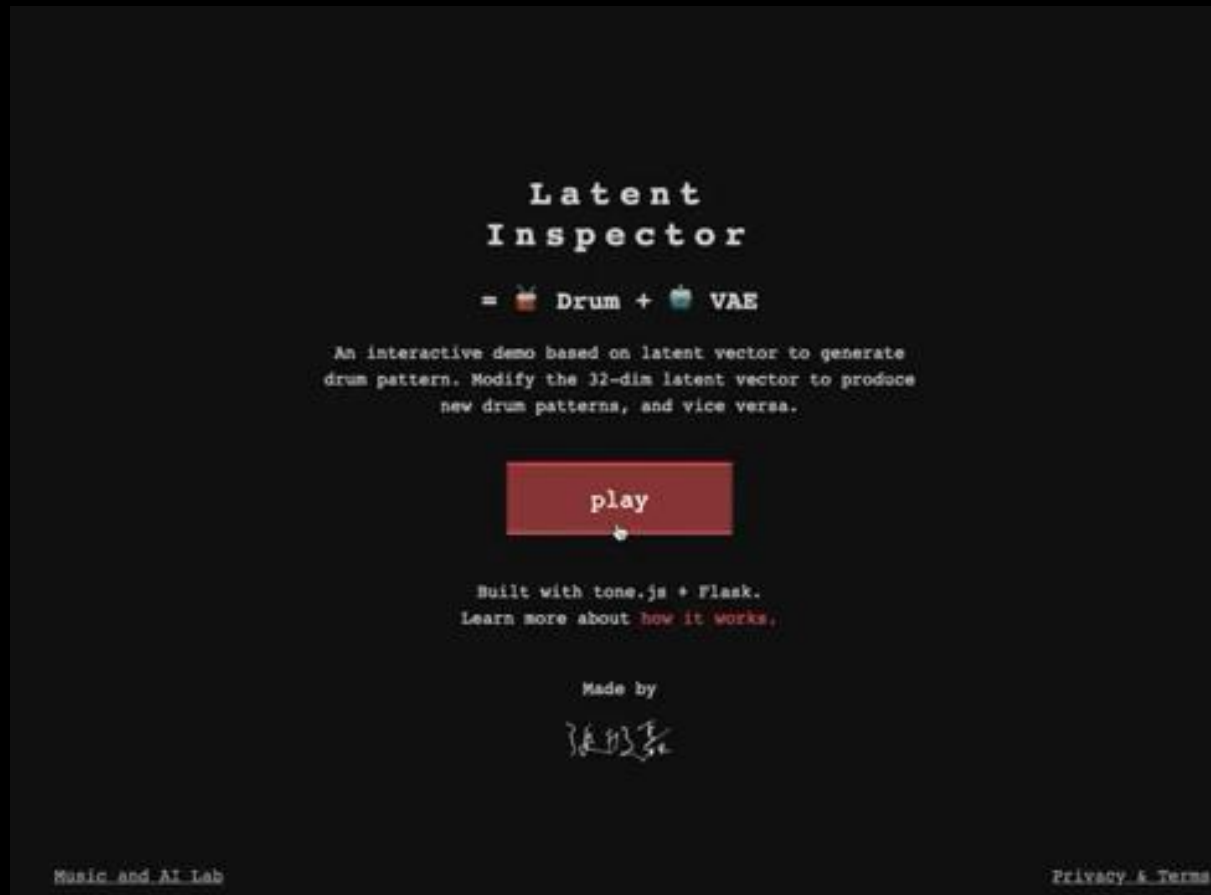
<https://vibertthio.com/drum-vae-client/>

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

Draw and Listen: A Sketch Based System for Music Inpainting

Christodoulos Benetatos, Zhiyao Duan, TISMIR 2022.

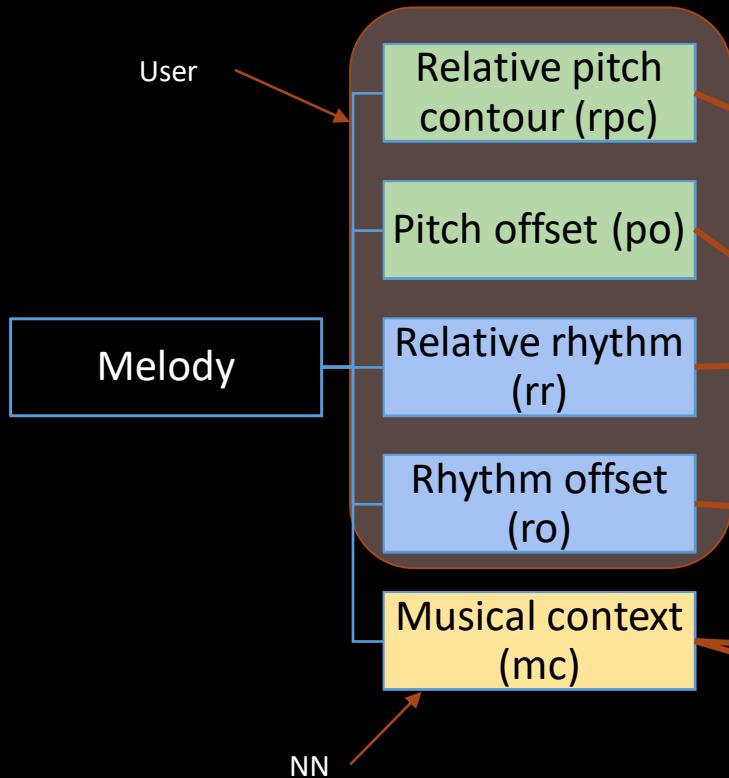
- Draw curves to guide inpainting
 - Intuitive
Non-musicians understand notions of low-high pitch and sparse-dense rhythm
 - No music theory knowledge required

The interface is divided into three main sections:

- Previous measure / Next measure:** Shows a musical staff with a missing central measure. A **Generate** button is below it.
- User's Input:** A graph showing a **pitch curve** (green line) and a **Note density curve** (blue line). Below the graph are a **pitch slider** and a **rhythm slider**, both with red knobs.
- Result:** Shows the final musical staff with the generated notes in the missing measure highlighted by a red oval. A **Listen** button is below it.

Draw and Listen: A Sketch Based System for Music Inpainting ⁺

Proposed Melody Decomposition

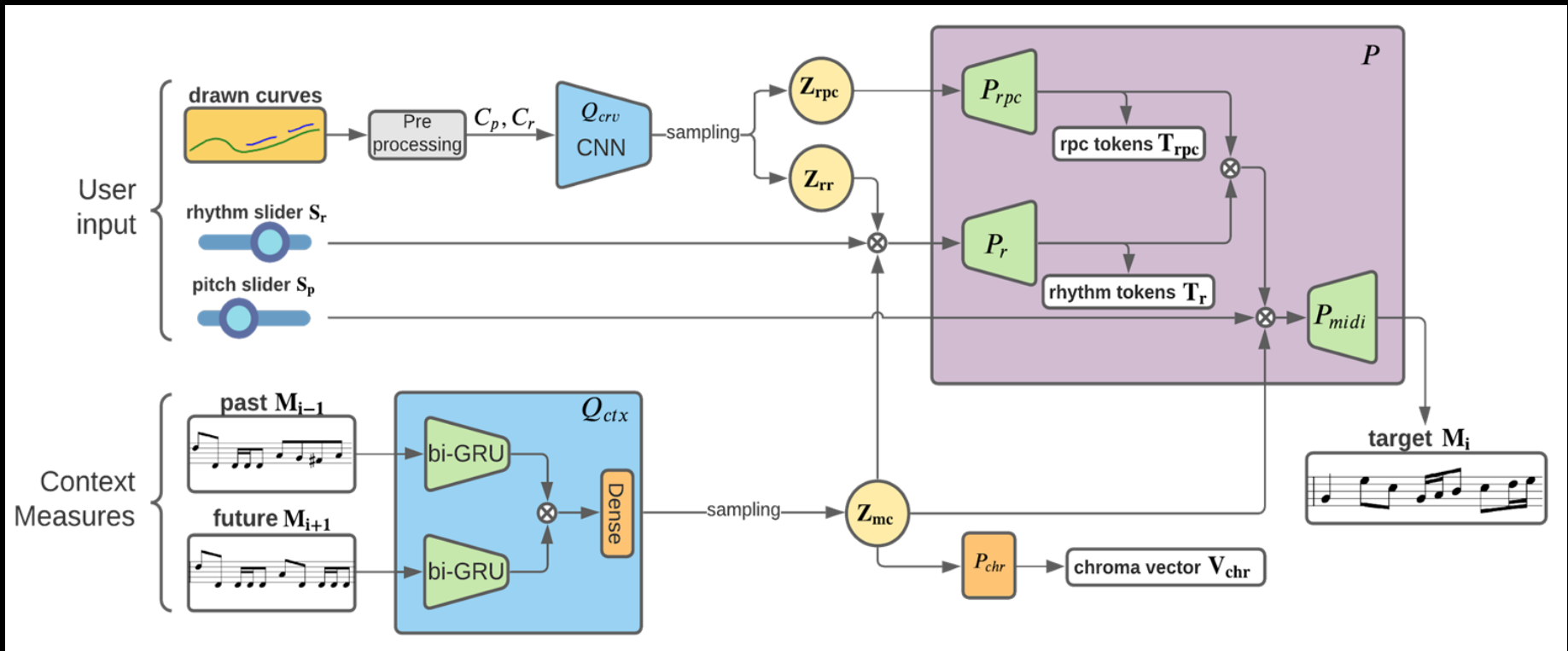


The interface is divided into three horizontal sections:

- Previous measure** and **Next measure**: Shows musical notation for the context. A **Generate** button is located below.
- User's Input**: A graphical area showing a green line representing the **pitch contour** and a blue line representing the **rhythm**. Below this are a **pitch slider** and a **rhythm slider**, both with red circular indicators.
- Result**: Shows the final generated musical notation. A red oval highlights the newly generated notes in the middle measure. A **Listen** button is located below.

Arrows from the decomposition diagram point to the green line (rpc), blue line (rr), sliders (po, ro), and the highlighted result (mc).

Model Architecture Based on Variational Auto-Encoders (VAE)



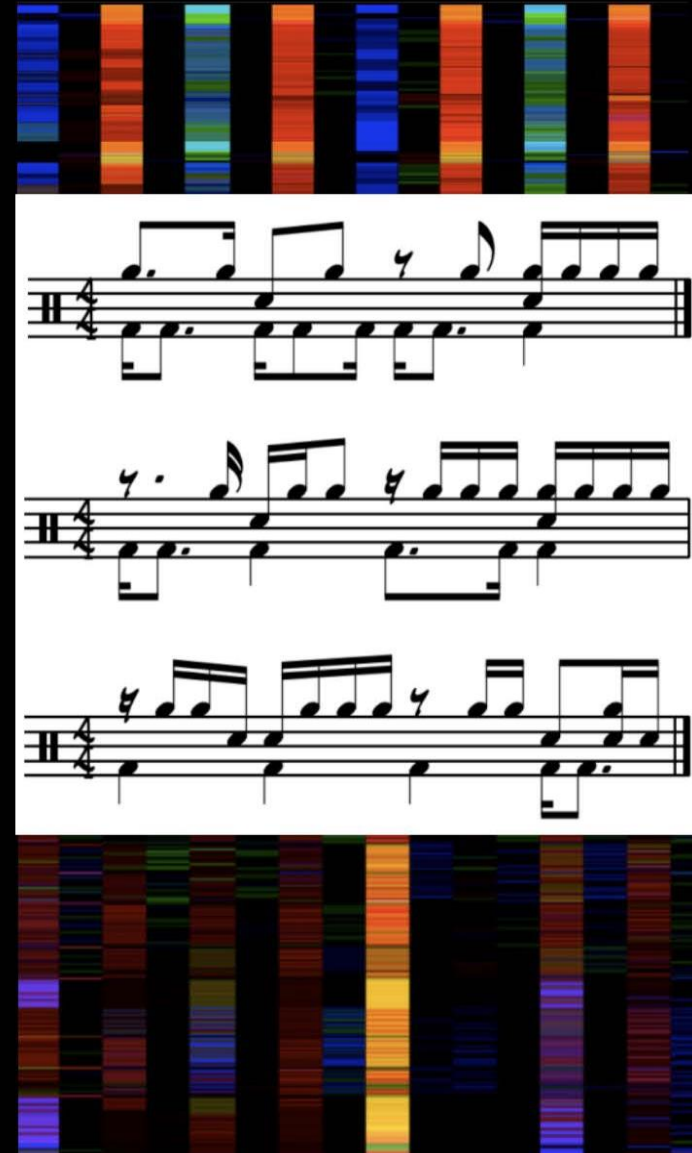
Multitrack Music Machine: MMM

- **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

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 Lakh Midi Dataset (existing)
 - Available on Zenodo (1k+ downloads)



MMM: Multitrack Music Machine

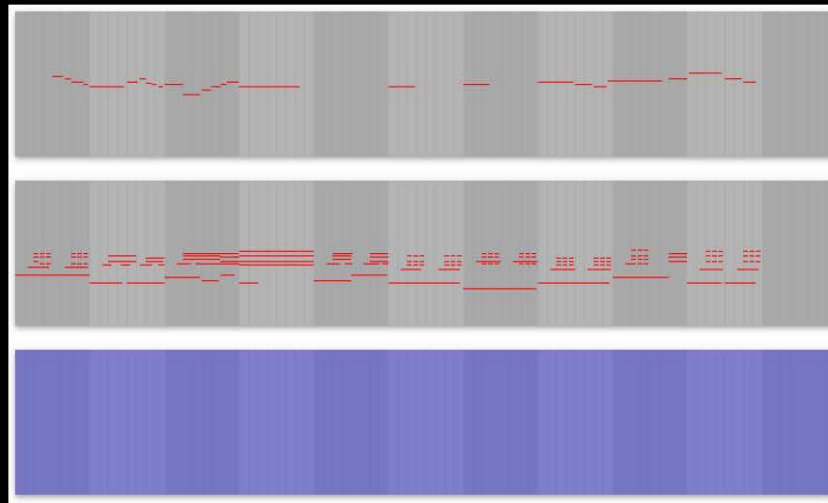
with Jeff Ens, Arxiv 2020,
ISMIR 2021

The interface features a top navigation bar with buttons for ADD MIDI, ADD TRACK, CLEAR TRACKS, GENERATE, PLAY, DOWNLOAD, and SETTINGS. Below this are five track control panels, each with a track name, instrument type, and piano roll visualization.

- inst_0 Electric Piano 2:** SOLO, MUTE, IGNORE, AUTOREG, DENSITY slider.
- inst_1 Acoustic Guitar (nylon):** SOLO, MUTE, IGNORE, AUTOREG, DENSITY slider.
- inst_2 Tenor Sax:** SOLO, MUTE, IGNORE, AUTOREG, DENSITY slider.
- inst_3 Fretless Bass:** SOLO, MUTE, IGNORE, AUTOREG, DENSITY slider.
- inst_4 drum_0:** SOLO, MUTE, IGNORE, AUTOREG, DENSITY slider.

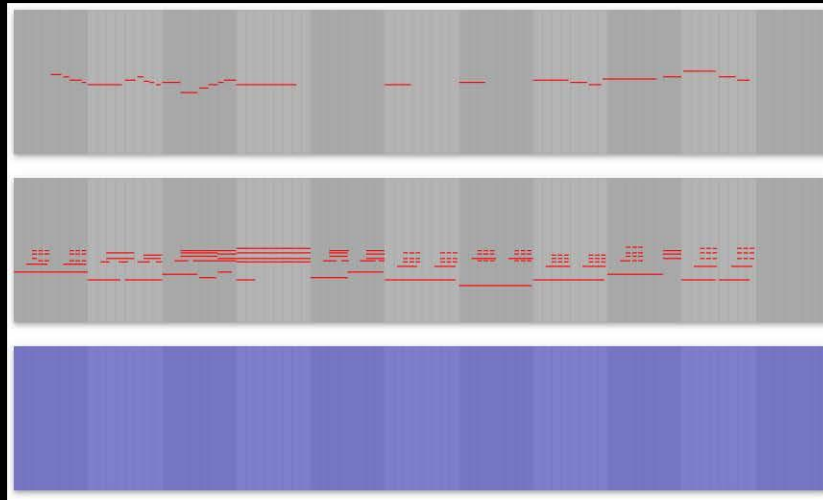
Each track's piano roll shows red horizontal lines representing notes over time. The tracks are arranged vertically, with inst_0 at the top and inst_4 at the bottom.

Flexible Generation (track in-filling)



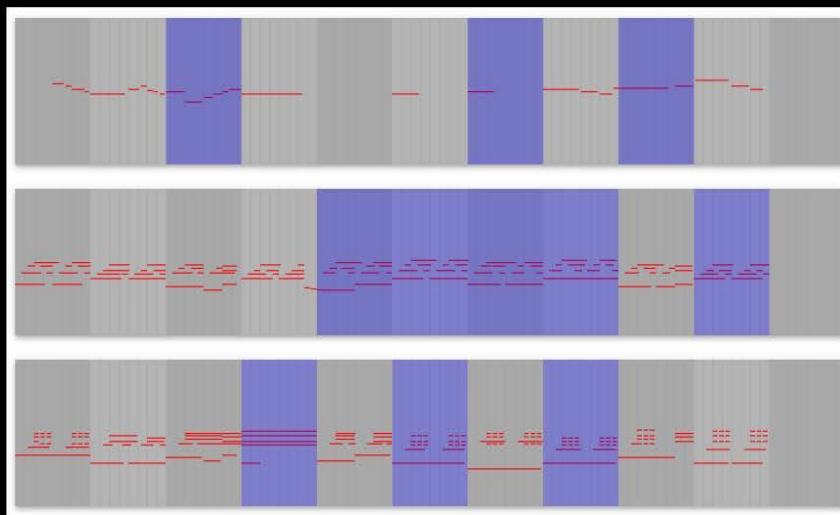
Generate a new track (shown in blue) conditioned on a set of tracks.

Flexible Generation (track in-filling)



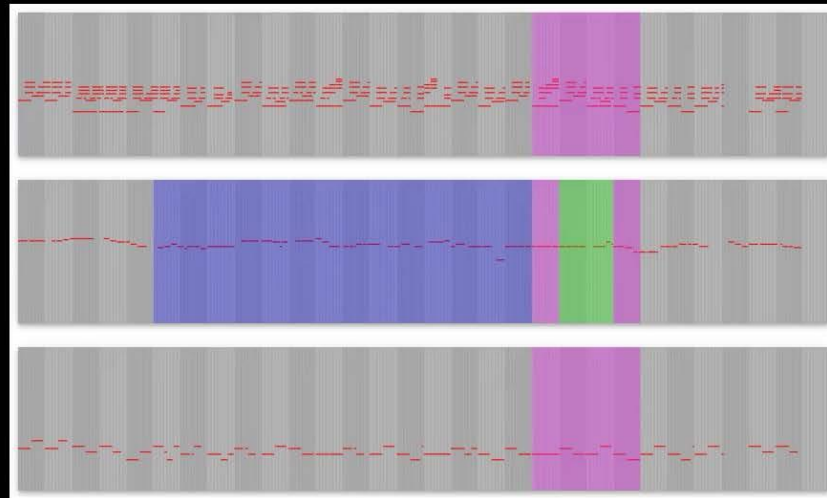
RE-Generate a track (shown in blue) conditioned on a set of tracks.

Generation (Bar in-filling)



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

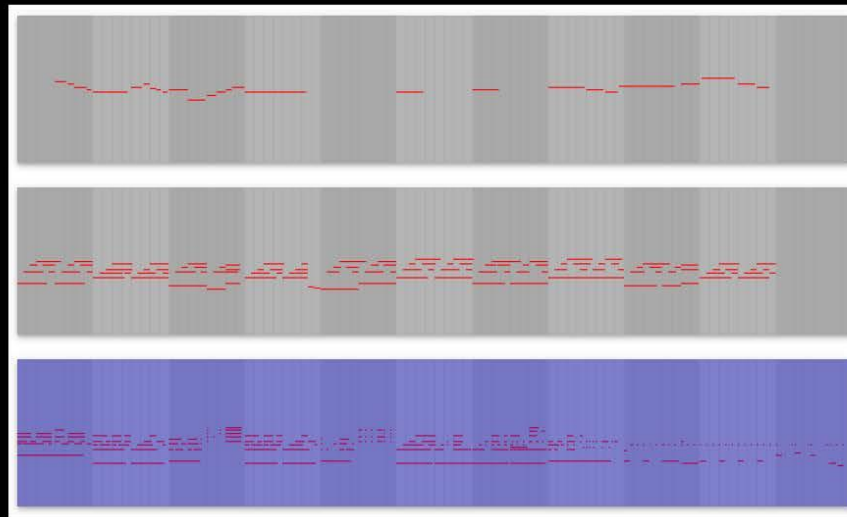
Flexible Generation



17 bars
generated

Generation of longer sections

Flexible Generation (controls)



High Density

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

MMM Generation Control

- 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/>

The screenshot displays the MMM web interface. At the top, there are control buttons: "ADD TRACK", "ADD MIDI", "CLEAR TRACKS", "MIDI EXAMPLE NONE", "GENERATE", "TEMPERATURE 0.975", "TEMPO 108", and "NBARS 4". Below these are three track controls. The first track, "drum_0", is active and shows a blue piano roll with a red dashed line. Its controls include "SOLO" (off), "MUTE" (off), "RESAMPLE" (checked), and "DENSITY" (a slider with a tooltip that says "Enable density control."). The second track, "Electric Guitar (clean)", is inactive (greyed out) and shows a grey piano roll with a red dashed line. Its controls include "SOLO" (off), "MUTE" (off), "RESAMPLE" (off), and "DENSITY" (a slider). The third track, "Electric Bass (finger)", is also inactive and shows a grey piano roll with a red dashed line. Its controls include "SOLO" (off), "MUTE" (off), "RESAMPLE" (off), and "DENSITY" (a slider).

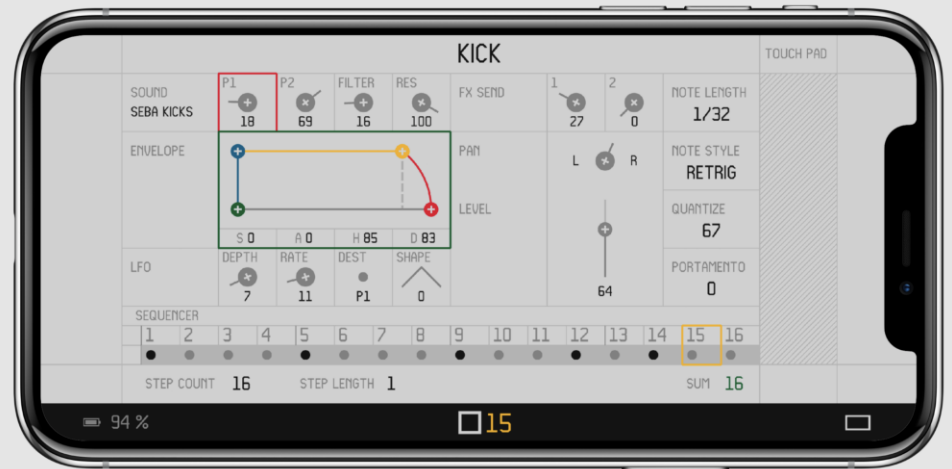
Applications

Integrating MMM in existing interfaces

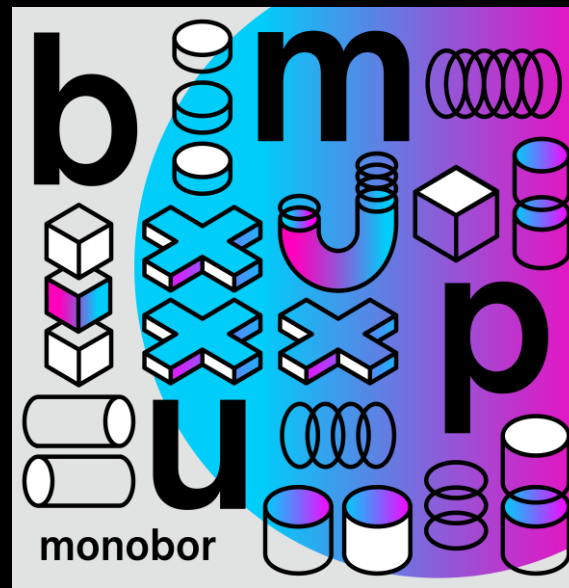
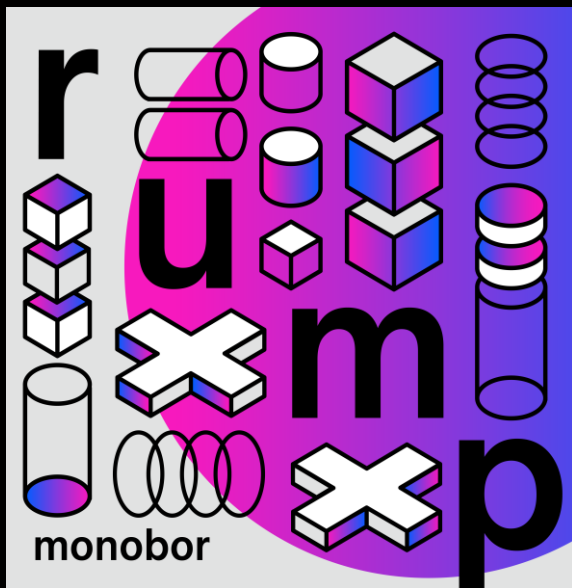
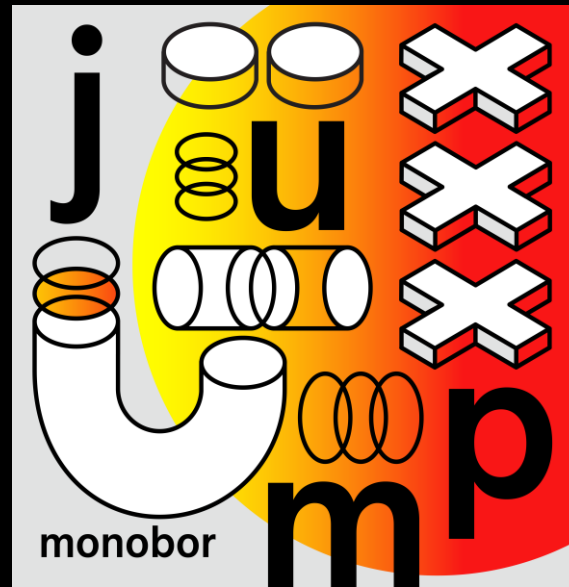
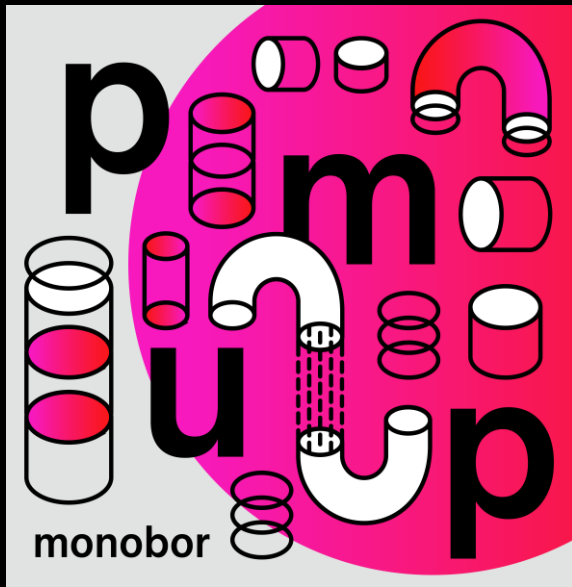
The screenshot displays the ELIAS Composer's Studio interface, which is organized into six vertical columns representing different musical elements: Drums, Bass, Melody, Synths, Harmony, and Counter Melody. Each column has a distinct color and contains a 'Reverb' and 'Settings' section, followed by a 'Key' control set to 'C' and 's'/'m' buttons. Below these are track slots for audio files, numbered 1 through 13. The 'Synths' column shows a sequence of 'Space Synths' files, while the 'Counter Melody' column includes 'Space CtrMelody' files and a 'SILENCE.wav' file.

Track	Drums	Bass	Melody	Synths	Harmony	Counter Melody
1	Space Drums_01.wav	Space Bass_01.wav	Space Melody_01.wav		Space Harmony_01.wav	
2						Space CtrMelody_01.wav
3				Space Synths_01.wav		
4		Space Bass_02.wav	Space Melody_02.wav			
5	Space Drums_02.wav				Space Harmony_02.wav	Space CtrMelody_02.wav
6			Space Melody_03.wav			
7		Space Bass_03.wav		Space Synths_02.wav		Space CtrMelody_03.wav
8	Space Drums_03.wav					
9		Space Bass_04.wav	Space Melody_04.wav			
10	Space Drums_04.wav				Space Harmony_03.wav	Space CtrMelody_04.wav
11		Space Bass_05.wav	Space Melody_05.wav			SILENCE.wav
12	Space Drums_05.wav			Space Synths_03.wav		
13					Space Harmony_04.wav	

Integrating to Synthesizers



Making and Releasing Music



Calliope: Online System

With Renaud Bougueng Tchameube,
Jeff Ens, SMC 2022, ICCG 2022,
C&C2022.

The screenshot displays the Calliope online system interface. At the top, the title "Calliope" is shown in a red header. Below the header, the session name "SMC_22_Pump" is visible. The "Selected File" section shows "1_1289_t_1.00_.mid" with an "ADD FILE(S)" button. A "Global AI Settings" gear icon is also present. The main area features a music score with three staves: "Ch. 3 electric bass (pick)" and "Ch. 2 electric piano 2". A large red QR code is overlaid on the score. To the right, a "GENERATE" button is shown with a "How many?" dropdown set to "2". Below this is a list of files: "4_1289_t_1.00_.mid", "3_1289_t_1.00_.mid", "2_1289_t_1.00_.mid", "1_1289_t_1.00_.mid" (highlighted), and "pump_up_variation (1).mid". The bottom control bar includes a file name "1_1289_t_1.00_.mid", a "MIDI Streaming Port" dropdown set to "None", a "Soundfont" dropdown set to "sgm-plus", a "Tempo" slider at 120, and a volume slider at 100.

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 Tchameube,
Jeff Ens, SMC 2022, ICCG 2022,
C&C2022.

The screenshot displays the Calliope online system interface. At the top left, the logo "CALLIOPE" is visible. The top right navigation bar includes "Home", "Update Profile", and "Logout renadmin". Below this, the session name "classical-session" is shown with a "New Session" button. A "Global AI Settings" gear icon is also present.

The main workspace features two piano tracks, both labeled "Ch. 1" and "acoustic grand piano". Each track has a MIDI piano roll editor with blue and red notes respectively. Below each piano roll are four sliders: "Density" (set to 0.00), "Min Polyphony" (set to 1), "Max Polyphony" (set to 4), and "Min Note Length" (set to 4). A "Temperature" slider is located at the bottom center, set to 1.00. To the right of the temperature slider is a "Track to compare against" dropdown menu and a "Rank" button.

At the bottom left, the file name "K333 Piano Sonata n13 3mov - 32bars.mid" is displayed. Below it are playback controls (play, stop, next, previous, volume) and a "No. of samples" field set to 1. A "Generate" button is located to the right of the sample count. At the bottom right, there are "Streaming Port" (set to None), "Soundfont" (set to sgm-plus), "Tempo" (set to 165), and "Volume" (set to 100) controls.

Apollo

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

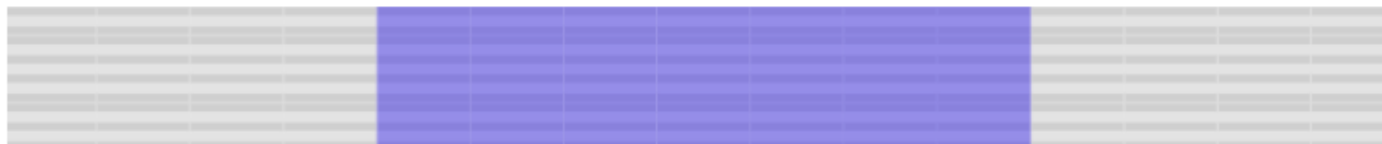
- Apollo EP – Hard Disk Fever (Original)
- Apollo EP – Hard Disk Fever (Melody variation)
- Apollo EP – Hard Disk Fever (New Brass section)



Bar Selection & Local Settings

Ch. 9 **acoustic grand piano** ▾

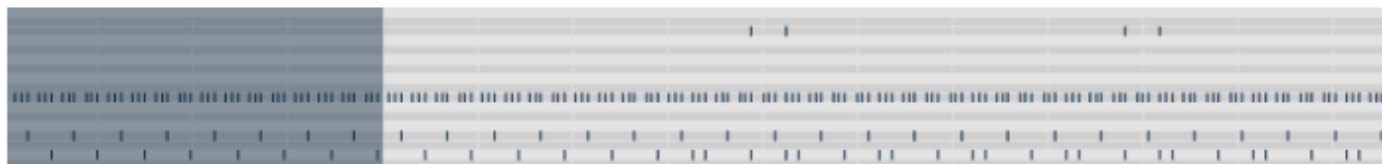
MUTE SOLO AUTOREG



Density 0.00 Min Polyphony 1 Max Polyphony 4 Min Note Length 1 Max Note Length 4

Ch. 10 **standard kit** ▾

MUTE SOLO AUTOREG



Density 0.00 Min Note Length 1 Max Note Length 4

Ch. 3 **electric bass (finger)** ▾

MUTE SOLO AUTOREG



Density 0.00 Min Polyphony 1 Max Polyphony 4 Min Note Length 1 Max Note Length 4

Global Settings

Craziness / entropy

Global Settings

Temperature

1.00

Polyphony Hard Limit

6

Variation engine

Percentage

100

Model Dimensions

4

Sampling control

Tracks per Step

4

Bars per Step

2

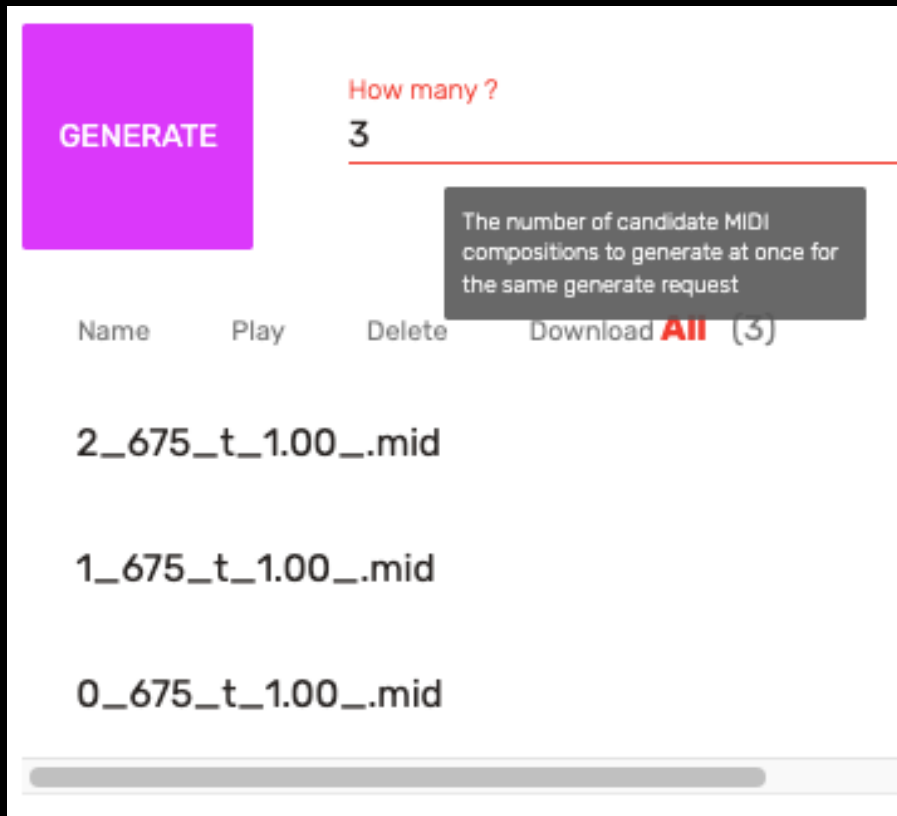
Max Steps

0

Tempo

120

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, ICC 2018. Best Paper Award*

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/>


The screenshot displays the MMM4Live software interface, which is designed for live performance and recording. The interface is divided into several key sections:

- Top Bar:** Shows the current tempo (125.00), time signature (4/4), and transport controls (play, stop, record, solo, mute).
- Left Panel:** A 'Collections' sidebar with categories like Sounds, Drums, Instruments, and MIDI Effects. It also includes a 'Places' section for organizing projects and samples.
- Mixer (Center):** A multi-track mixer with 8 tracks: 1 Drums, 2 Bass, 3 Melo, 4 Aux, 5 Strings, 6 Percussi, 7 Aux 2, and 8 MKI1 Holl. Each track has a MIDI From dropdown, a Monitor section with In/Auto/Off buttons, and an Audio To dropdown. Below the mixer are volume faders and send controls for each track.
- Central Control Panel (interface):** A prominent 'GENERATE!' button is at the top. Below it are tabs for '5 Strings', '6 Percussions', and '7 Aux 2'. The current instrument is set to 'Acoustic Grand Piano'. A 'generate' button is followed by 'selected bars' and a 'density' knob set to 6. A bar selection bar is visible below. Further down, there are 'conservative', 'craziness 1.', and 'experimental' options, and a 'percentage of generated steps: 100' setting. At the bottom of this panel are 'tracks per step: 1' and 'bars per step: 2' controls.
- Right Panel:** A master section with a 'Cue Out' dropdown set to 1/2, and a 'Master Out' dropdown also set to 1/2. It includes volume faders and send controls for the master.
- Bottom Panel:** A piano roll for track 5 (Strings) showing a sequence of notes (A#1-G3) over 16 bars. The notes are represented by yellow rectangles on a grid. Below the piano roll is a MIDI piano roll showing a sequence of red notes.

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 (?).

[View this email in your browser.](#)

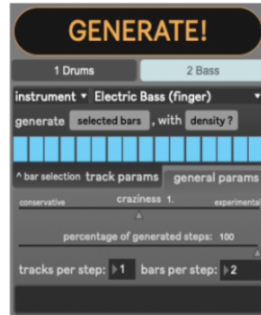


METACREATION
Lab for Creative Artificial Intelligence

MMM4Live Beta Testing Invitation

We invite you to join the closed beta testing group of MMM4Live, a plugin for Ableton Live under development by the Metacreation Lab, as we believe that you are interested in music AI and computer-assisted composition.

The plugin offers an interface to [the MMM model](#) an algorithm capable of generating multi-track MIDI patterns and sequences, for use within Ableton Live 10 or 11. Trained on more music than a human can possibly listen to, the system is the state of the art in music AI research.





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)



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 instruct for melody, harmony and rhythmic generation of new

41
42
43
44
45

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?

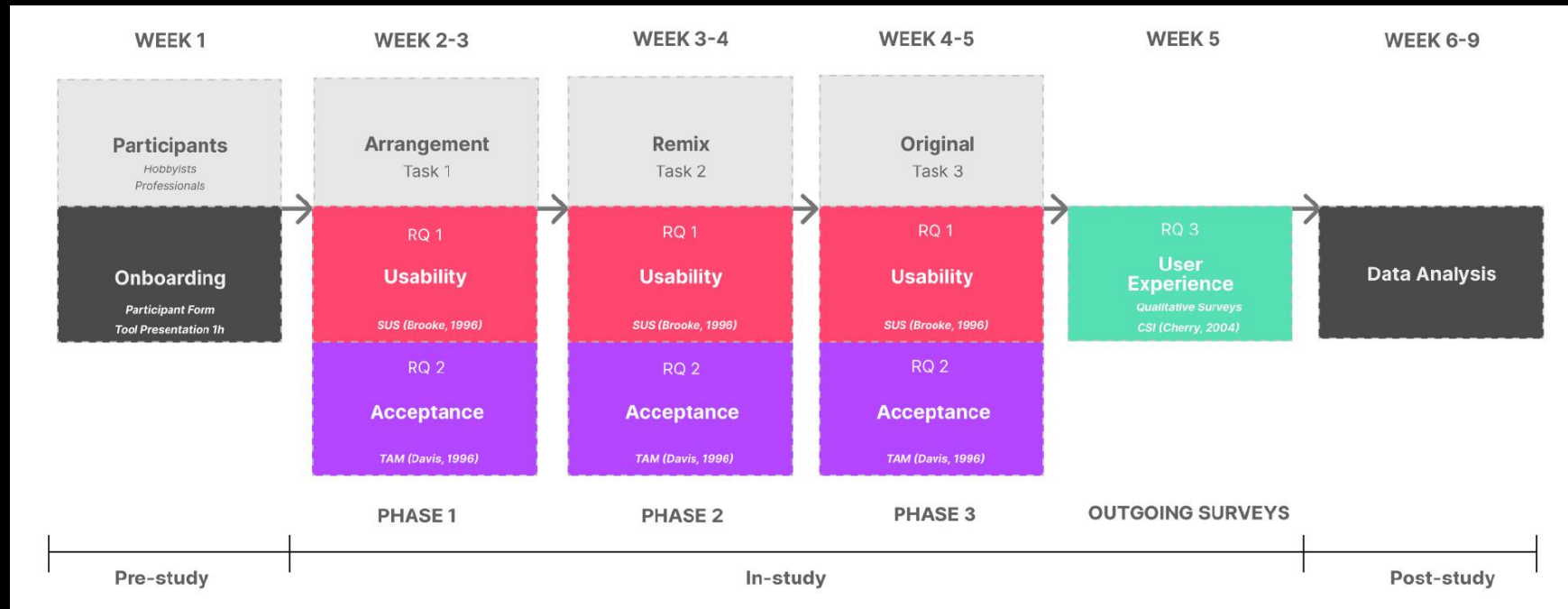


Usability and Acceptability Evaluation (CuBase)



Study Protocol

With Renaud Bougueng Tchameube, 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)

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



Active — Credit Eligible
Open for Enrollment

Generative Art and Computational Creativity

Simon Fraser University
Philippe Pasquier with special guest Arne Eigenfeldt



[VIEW](#)



Active — Credit Eligible
Open for Enrollment (Program Exclusive)

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

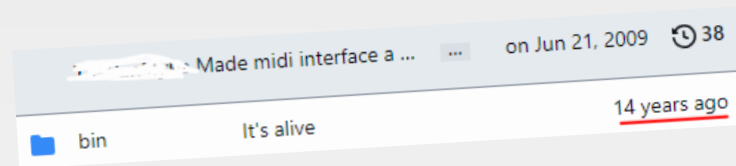
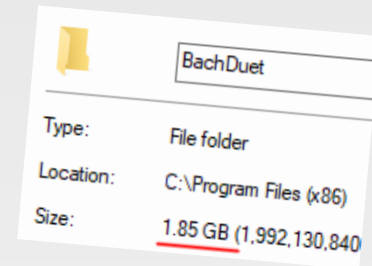
Yongyi Zhang*, Christodoulos Benetatos*, Zhiyao Duan

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



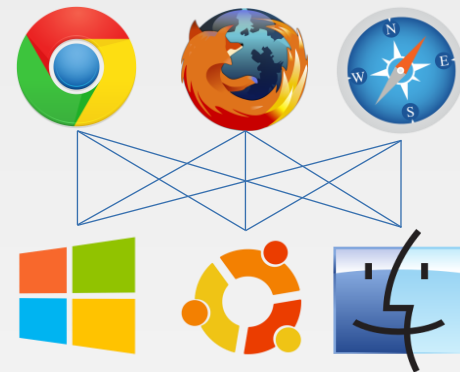
Installation

OS Support

We currently only support Windows 64-bit.

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 ...

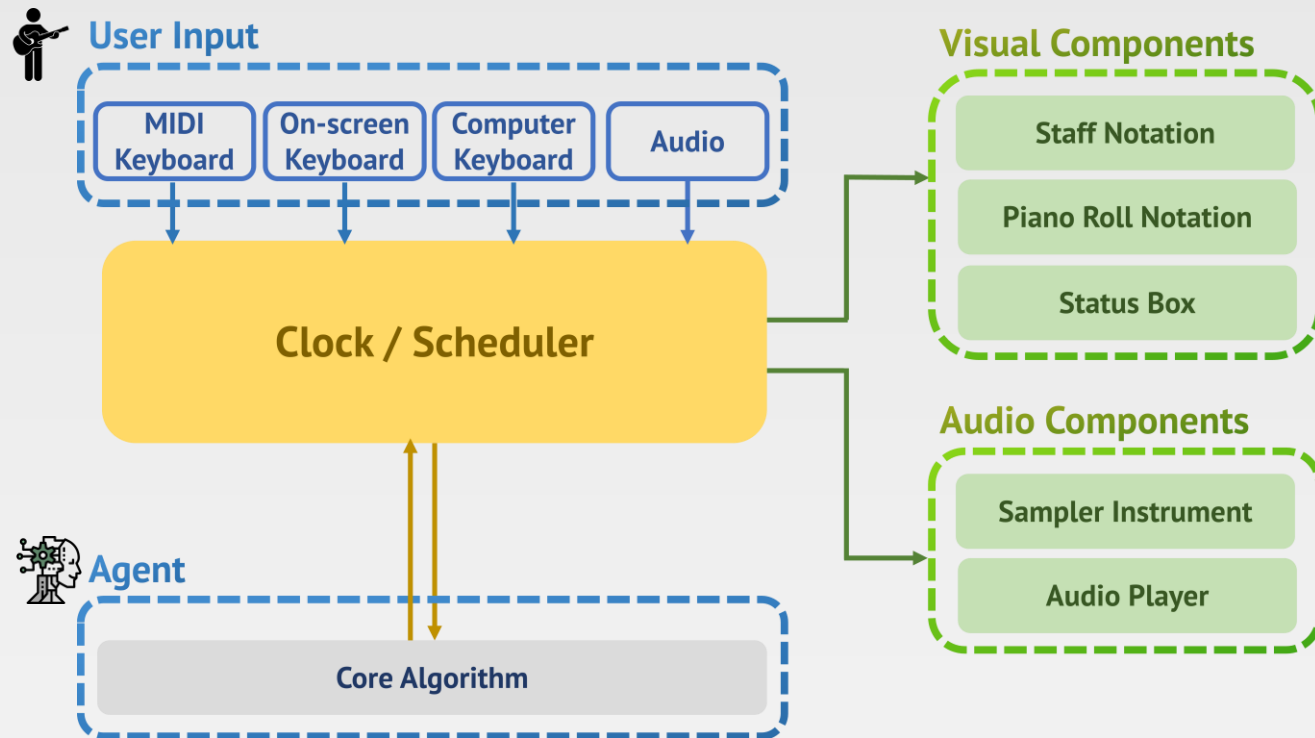
- Promote the development of **web** musical systems
 - Cons
 - Knowledge of web programming is required (JavaScript, CSS, HTML)



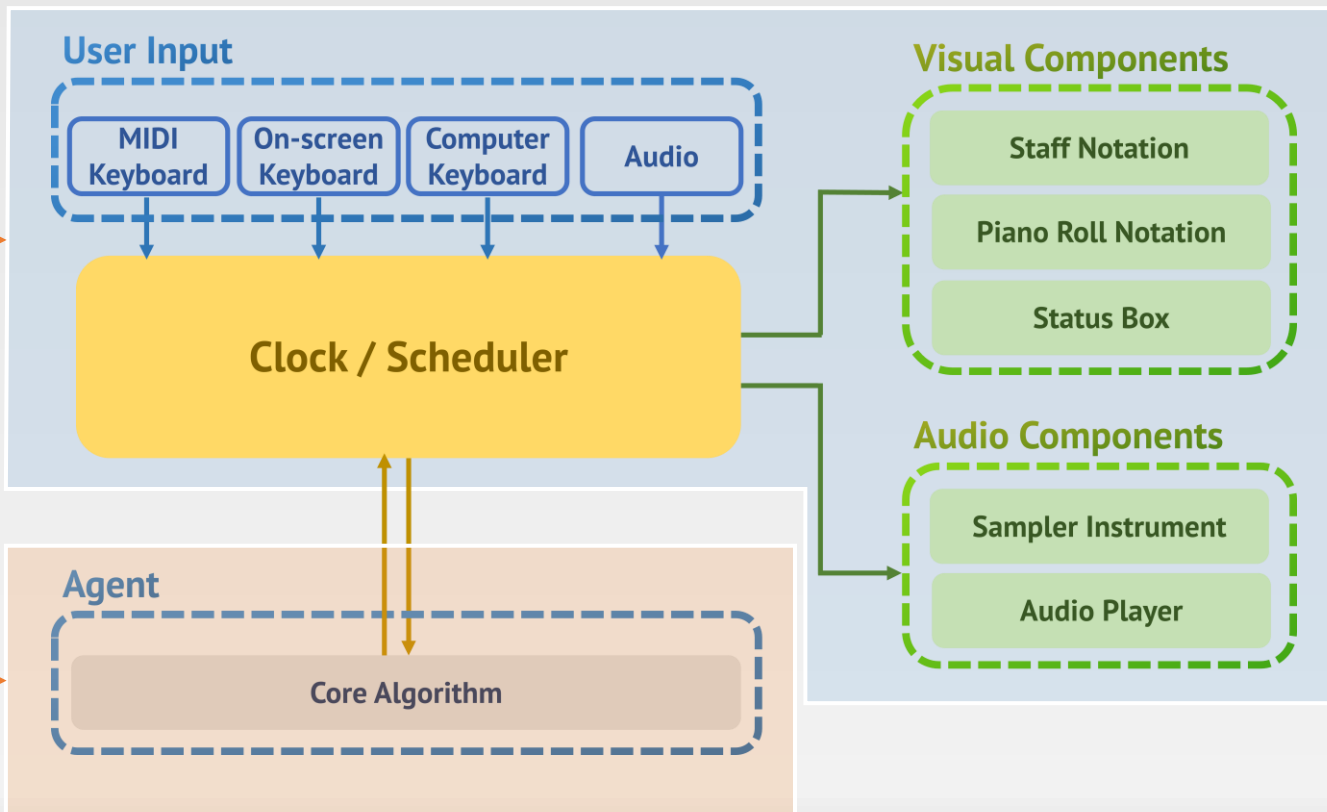
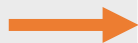
Euterpe's Goal

- **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

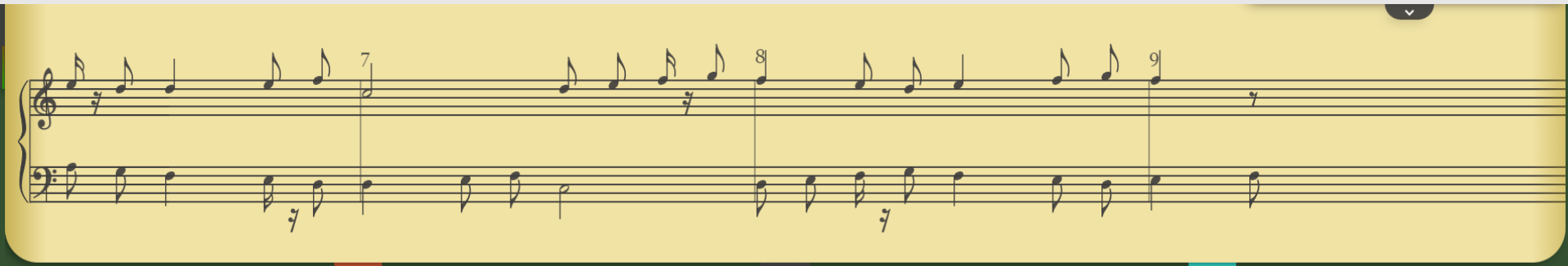
- Modular
 - **Separate** the **Agent code** from the peripheral components
- Configuration files
 - Allow app setup and **customization without** writing **JavaScript**

Visual Components

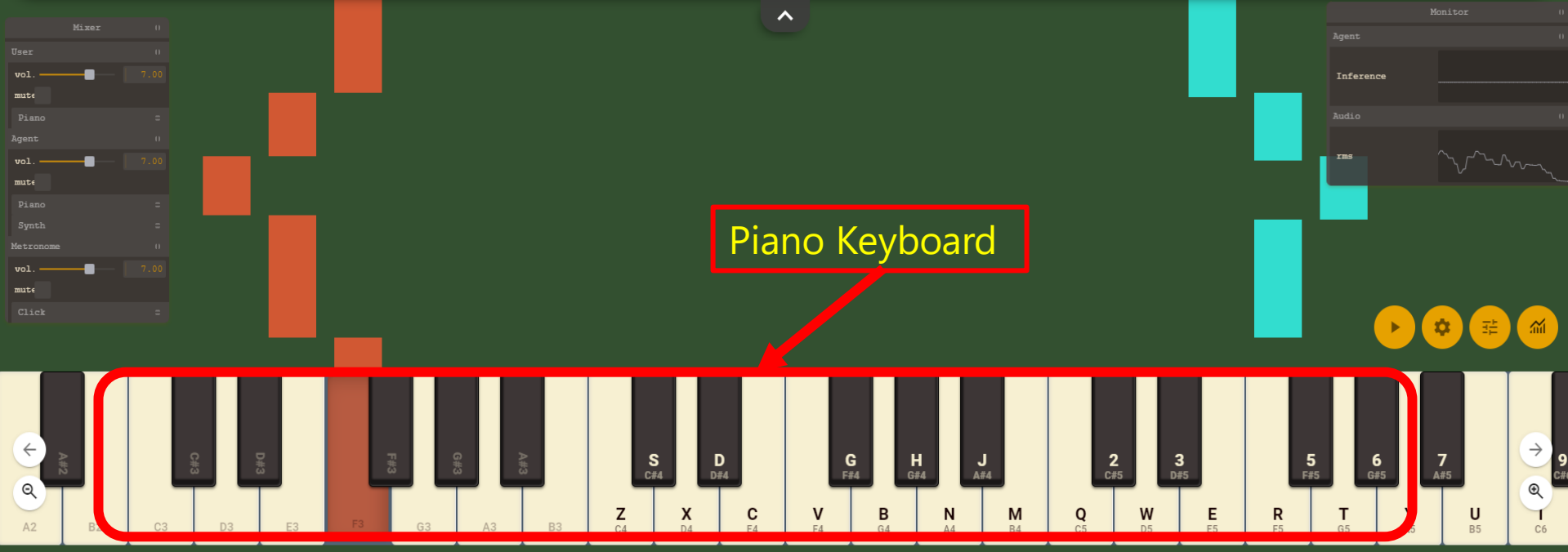
The image displays a music software interface with several key components:

- Score Widget:** A red-bordered box at the top highlights a musical score with two staves (treble and bass clef) and measures numbered 7, 8, and 9. A red arrow points from this box to a label.
- Mixer:** A vertical panel on the left side contains controls for various tracks: Mixer, User, Piano, Agent, Piano, Synth, and Metronome. Each track has a volume slider set to 7.00 and a mute button.
- Monitor:** A panel on the right side shows monitoring options for Agent, Inference, and Audio, along with an RMS waveform graph.
- Keyboard:** A digital piano keyboard is visible at the bottom, with keys labeled with letters (A-Z) and numbers (2-7) and their corresponding musical notes (e.g., A#2, C#3, F#3, G#3, A#3, S C#4, D D#4, G F#4, H G#4, J A#4, 2 C#5, 3 D#5, 5 F#5, 6 G#5, 7 A#5).
- Score Widget Label:** A red-bordered box with the text "Score Widget" is positioned in the center, with a red arrow pointing to the score area.

Visual Components



A musical score for piano, showing two staves (treble and bass clef) with notes and rests. The score is displayed on a yellow background. The treble clef staff contains a melody of eighth and quarter notes, while the bass clef staff provides a harmonic accompaniment with chords and single notes. Measure numbers 7, 8, and 9 are visible at the top of the staves.



A digital piano keyboard interface with various controls and visual elements. The keyboard is shown with keys labeled with notes (A#2, C#3, D#3, F#3, G#3, A#3, S C#4, D D#4, G F#4, H G#4, J A#4, 2 C#5, 3 D#5, 5 F#5, 6 G#5, 7 A#5) and letters (A2, B2, C3, D3, E3, F3, G3, A3, B3, Z C4, X D4, C F4, V F4, B G4, N A4, M B4, Q C5, W D5, E F5, R F5, T G5, U B5, I C6). A red box highlights the keyboard area, and a red arrow points to it from a text box labeled "Piano Keyboard".

Mixer controls on the left include:

- Mixer (0)
- User (0)
- vol. 7.00
- mute
- Piano (0)
- Agent (0)
- vol. 7.00
- mute
- Piano (0)
- Synth (0)
- Metronome (0)
- vol. 7.00
- mute
- Click (0)

Monitor controls on the right include:

- Monitor (0)
- Agent (0)
- Inference
- Audio (0)
- rms

Navigation and utility icons at the bottom right include play, settings, and zoom controls.

Visual Components

The image shows a screenshot of a music software interface with a dark green background. At the bottom is a piano keyboard with keys labeled with letters and notes. Two red boxes with arrows point to specific areas of the PianoRoll:

- PianoRoll (agent):** A red box on the left points to a vertical column of orange rectangular notes on the piano roll.
- PianoRoll (user):** A red box on the right points to a vertical column of cyan rectangular notes on the piano roll.

Other visible components include a Mixer panel on the left with sliders for User, Piano, Agent, Piano, Synth, Metronome, and Click. A Monitor panel on the right shows Agent, Inference, Audio, and rms levels. A frequency spectrum analyzer is visible in the top left corner.

Configuration

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

Visual Components



A digital piano interface is shown on a dark green background. At the bottom, there is a piano keyboard with keys labeled with letters and notes (e.g., A2, B2, C3, D3, E3, F3, G3, A3, B3, Z C4, X D4, C E4, V F4, B G4, N A4, M B4, Q C5, W D5, E E5, R F5, T G5, Y A5, U B5, I C6). On the left side, there is an 'Audio Mixer' panel with a red border. It contains several tracks: 'Mixer', 'User', 'Piano', 'Agent', 'Synth', and 'Metronome'. Each track has a volume slider set to 7.00 and a mute button. On the right side, there is a 'Variable Monitor' panel with a red border. It shows an 'Agent' section, an 'Inference' section with a progress bar, an 'Audio' section with a waveform, and an 'rms' section with a line graph. A red arrow points from the 'Audio Mixer' panel to the 'Variable Monitor' panel. In the bottom right corner, there are four yellow circular icons: a play button, a settings gear, a list icon, and a waveform icon.

Audio Mixer

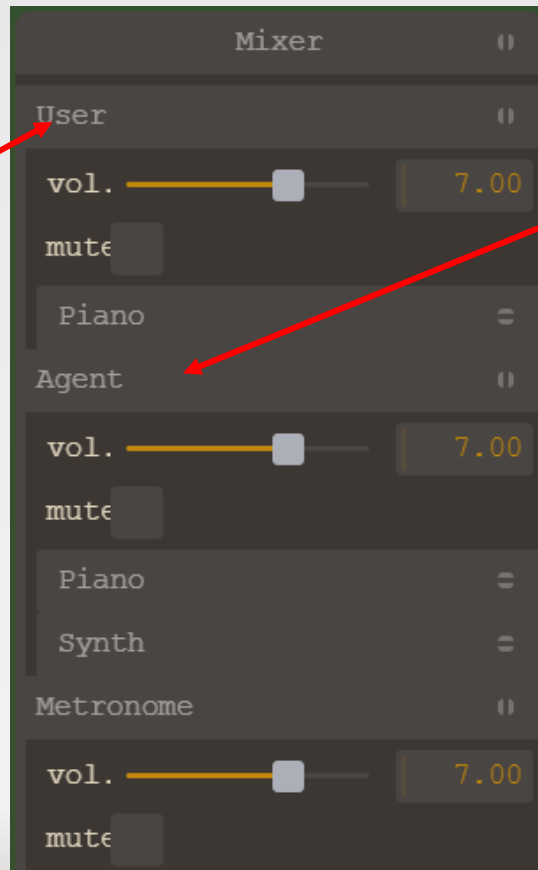
Variable Monitor

A detailed view of the 'Variable Monitor' panel. It has a dark background and a white border. The title 'Monitor' is at the top. Below it, there are three sections: 'Agent' with a small icon, 'Inference' with a horizontal progress bar, and 'Audio' with a waveform plot. At the bottom, there is an 'rms' section with a line graph showing a fluctuating signal.

Configuration

```

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



```

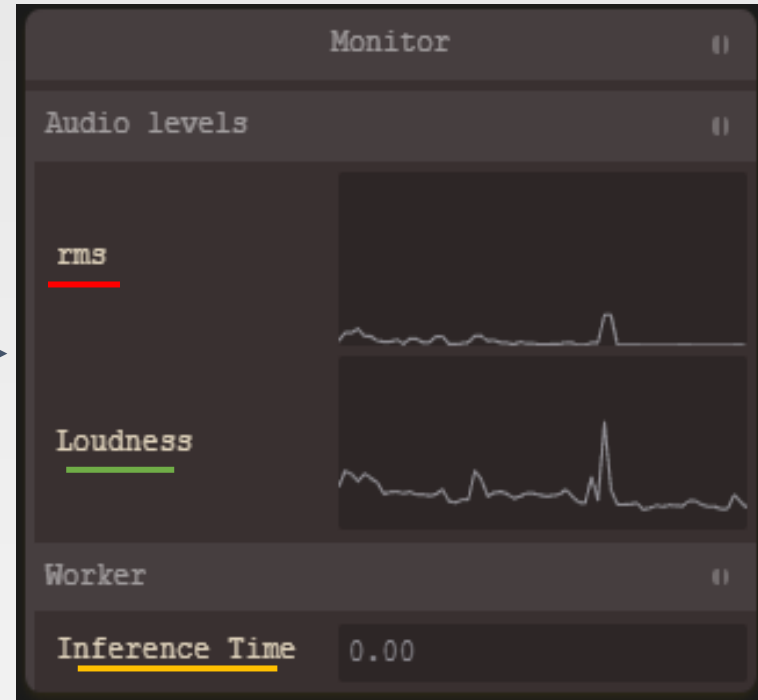
players:
  agent:
    label: 'Agent'
    mute: false
    volume: 5
    instruments:
      - id: "piano"
        label: "Piano"
        mute: false
        volume: 7
        default: true
      - id: "synth"
        label: "Synth"
        mute: false
        volume: 8
  
```

Configuration

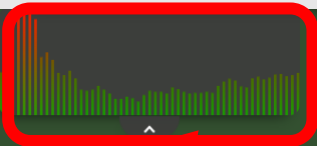
```

1  monitor:
2    title: "Monitor"
3    structure:
4      - label: "Audio levels"
5        parameters:
6          - id: 0 # id's must be unique
7            label: "rms" # choose any name
8            interval: 50 # in ms
9            graph: true
10           min: 0
11           max: 0.2
12          - id: 1
13            label: "Loudness"
14            interval: 50 # in ms
15            graph: true
16            min: 0
17            max: 100
18      - label: "Worker"
19        parameters:
20          - id: 2
21            label: "Inference Time"
22            interval: 100 # in ms
23            graph: false
24            min: 0
25            max: 30

```



Visual Components



Audio Spectrum



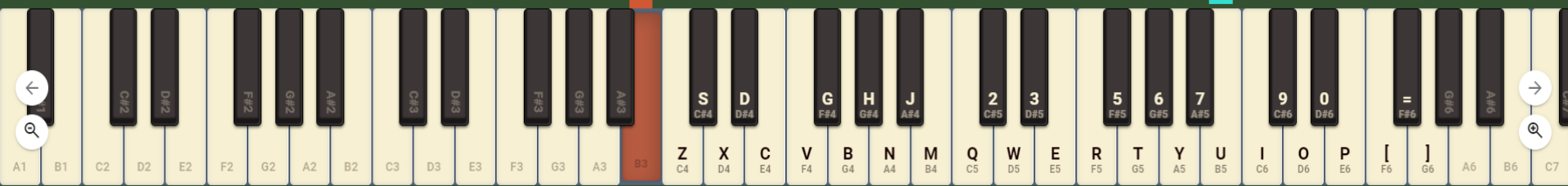
Chroma Vector

Mixer

- User
 - vol. 7.00
 - mute
- Piano
 - vol. 7.00
 - mute
- Agent
 - vol. 7.00
 - mute
- Piano
 - vol. 7.00
 - mute
- Synth
 - vol. 7.00
 - mute
- Metronome
 - vol. 7.00
 - mute
- Click
 - vol. 7.00
 - mute

Monitor

- Agent
- Inference
- Audio
 - rms



A digital piano keyboard interface. The keyboard is shown with various keys labeled with notes (e.g., C#2, D#2, F#2, G#2, A#2, C#3, D#3, F#3, G#3, A#3, B3, C4, D4, E4, F4, G4, A4, B4, C5, D5, E5, F5, G5, A5, B5, C6, D6, E6, F6, G6, A6, B6, C7). The interface includes a mixer on the left, a monitor on the right, and various control buttons at the bottom (play, settings, volume, and zoom).

The image shows a music application interface. At the top, there are two musical staves with notes. Below the staves is a piano roll with red rectangular blocks representing notes on a keyboard. At the bottom, a portion of a piano keyboard is visible, with keys labeled with notes like D#2, F#2, G#2, A#2, C#3, D#3, F#3, G#3, A#3, S, C#4, D#4, G, F#4, G#4, J, A#4, 2, C#5, D#5, 3, F#5, G#5, 6, A#5, 7, 9, C#6, D#6, 0, F#6, and G#6. A settings overlay is centered on the screen, titled "Settings" in orange. It has a close button (X) in the top right corner. The settings are organized into sections: "Clock" with a BPM (Max: Infinity) slider set to 100; "MIDI" with a search input field containing the text "Type here to search for MIDI device"; and "Agent Parameters" with three vertical sliders for "Randomness", "Delay", and "Pitch Shift". The "Delay" slider has a blue dot. A red rectangular box highlights the word "Settings" in yellow text. In the bottom right corner of the application, there are two yellow circular buttons: a play button and a settings gear icon. On the right side of the interface, there is a "Monitor" panel with "Agent" and "Inference" sections.

Settings



Clock

BPM (Max: Infinity)

100



MIDI

Type here to search for MIDI device

Agent Parameters



Randomness



Delay



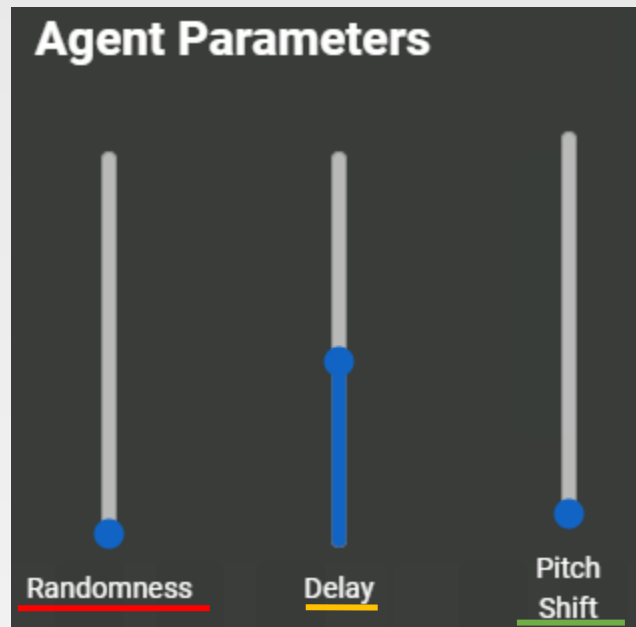
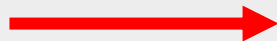
Pitch Shift

Settings



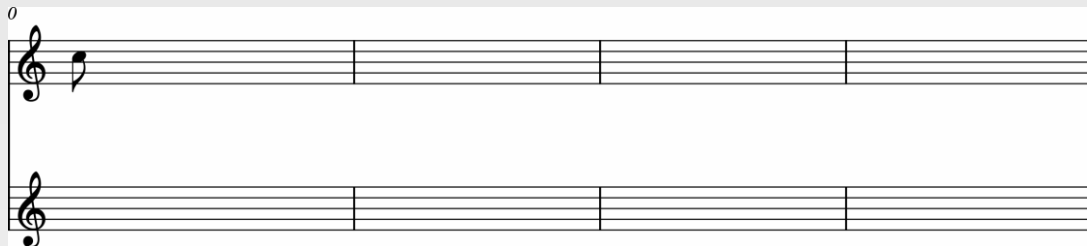
Configuration

```
settingsModal:  
  sliders:  
    - id: 1  
      label: "Randomness"  
      value: 0  
      min: 0  
      max: 10  
    - 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

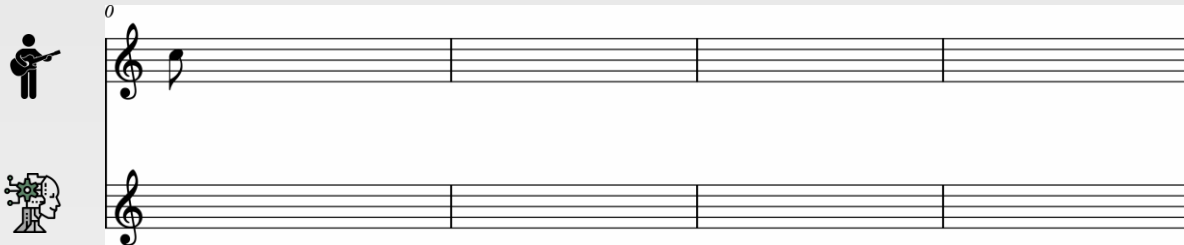


Diagram illustrating the Call & Response paradigm. It shows two staves of music. The top staff is marked with a '0' at the beginning and contains a single quarter note. The bottom staff is empty, representing the response.

- Simultaneous

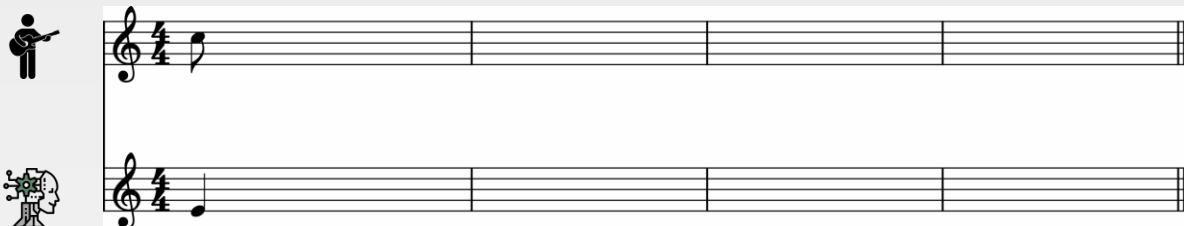
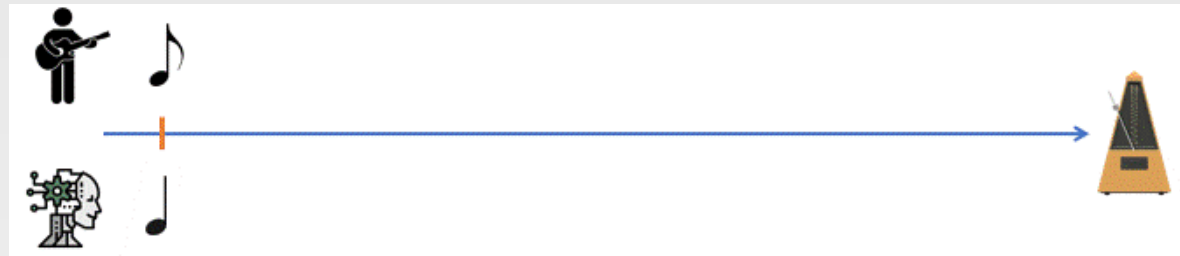


Diagram illustrating the Simultaneous paradigm. It shows two staves of music. The top staff starts with a treble clef and a 4/4 time signature, followed by a quarter note. The bottom staff also starts with a treble clef and a 4/4 time signature, followed by a quarter note. Both staves end with a double bar line.

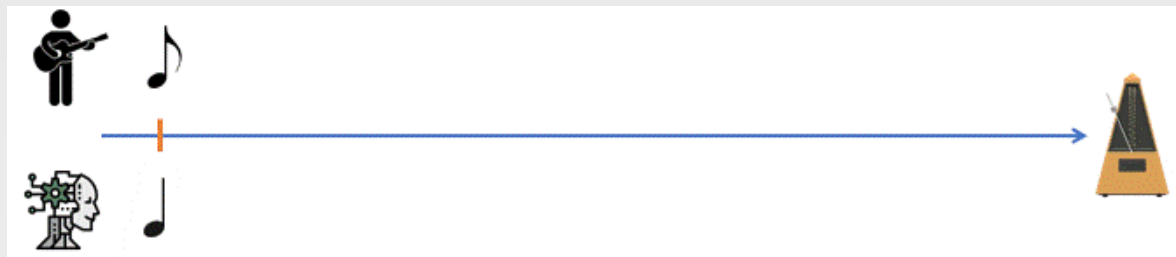
Paradigms of Music Interaction

- Grid-based



Paradigms of Music Interaction

- Grid-based



- Event-based



Configuration

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

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

```
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

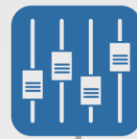


```
function hook(event){  
  // your code  
}
```

Euterpe Lifecycle

Initialization

Interaction



tick

Audio
buffer

UI
event

Audio
buffer

tick

MIDI
event

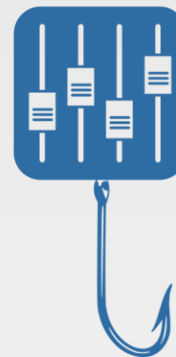
Agent - Hooks

- `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



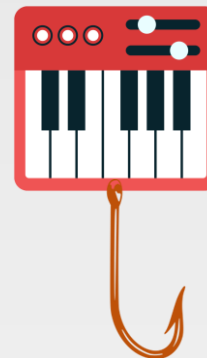
Agent - Hooks

- `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

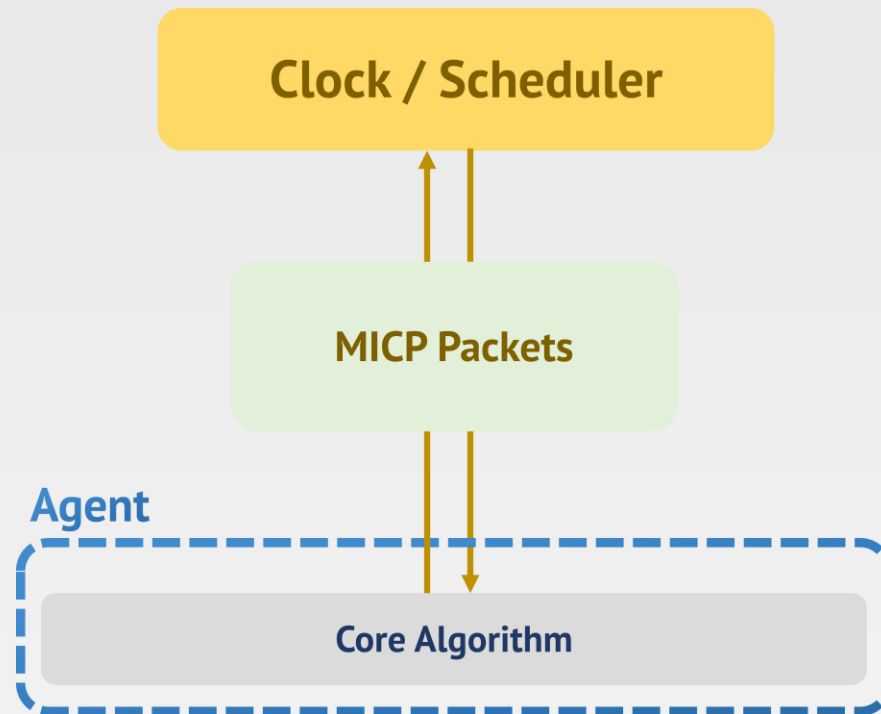


Agent - Hooks

- `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 :

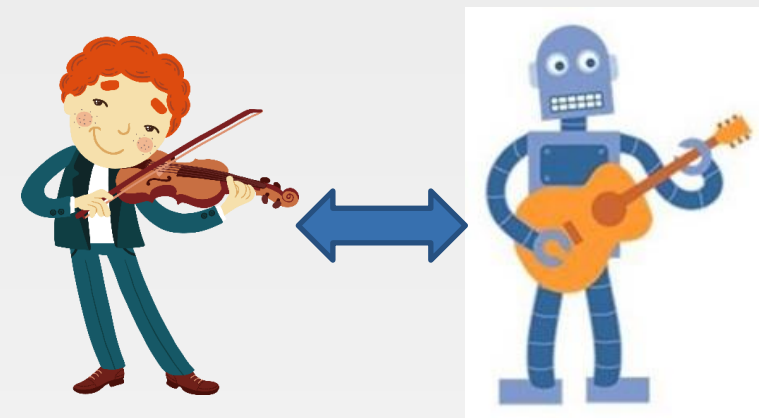
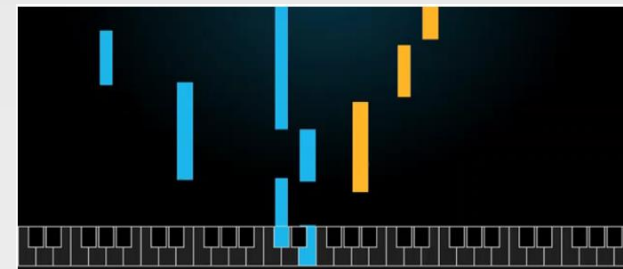
<https://xribene.github.io/>

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

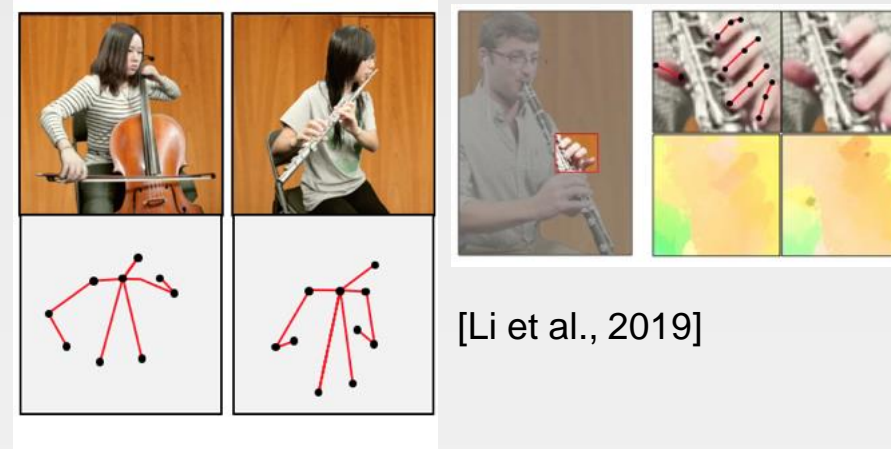


Incorporating the Visual Modality

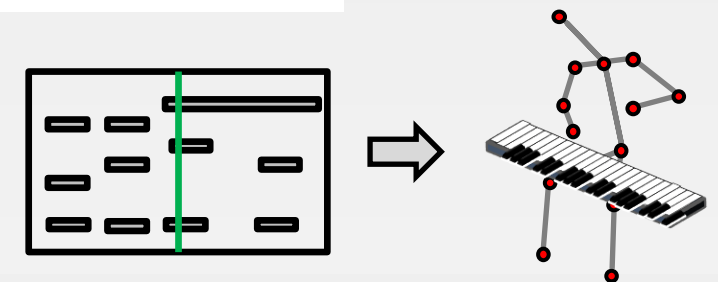
- Music performance is audiovisual in nature
 - Visual performance is important in musical expression
 - Musicians use visual cues to coordinate on timing, dynamics and intention
- Need algorithms to **analyze** various aspects of visual performance
 - Instrument recognition, body movement, facial expression, 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

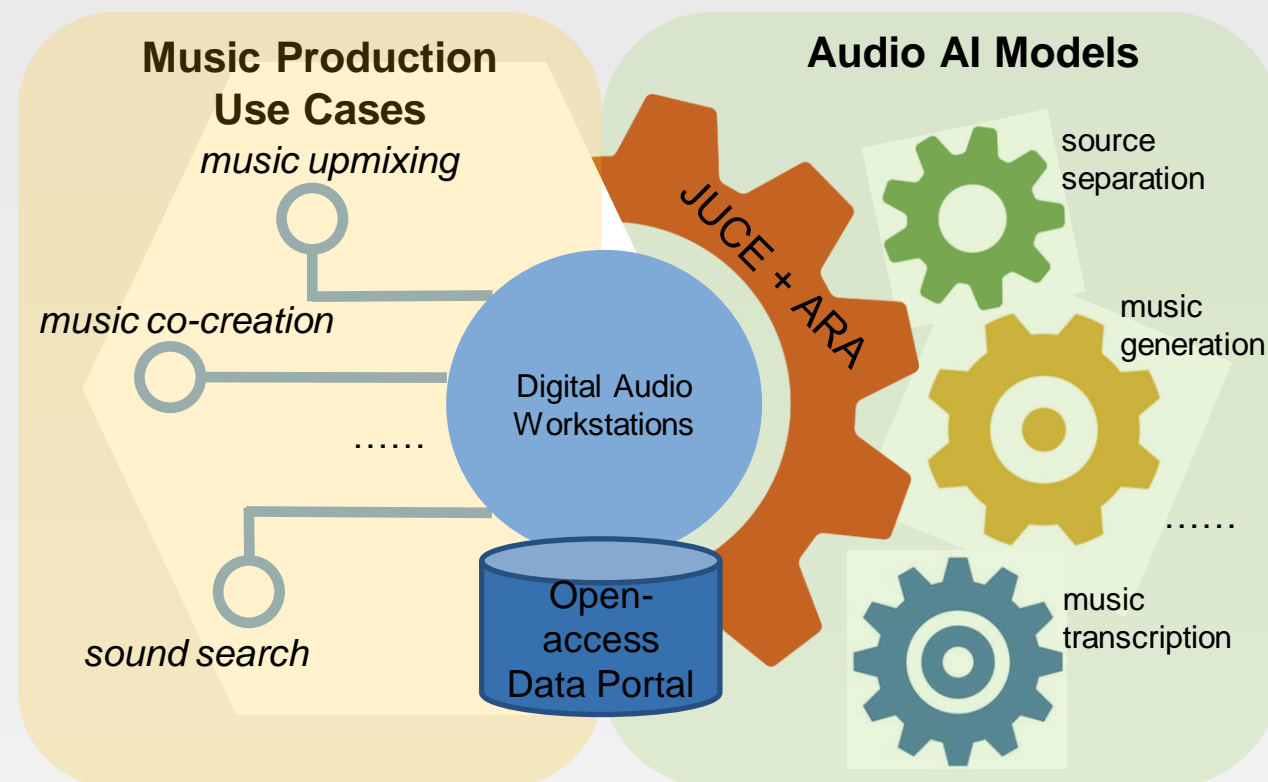
- Music theory is hard
- Music instruments are not easy to learn
- Many existing CAMM systems seem to be even harder to interact with
 - Require music background + software literacy
- Lower the barrier to entry?
 - Use music AI to augment users' capabilities

The image shows a three-panel interface for generating music. The top panel, labeled 'Previous measure' and 'Next measure', displays a musical staff with a treble clef and a key signature of one flat. The first measure contains a sequence of notes, and the second measure is empty. A 'Generate' button is centered below the staff. The middle panel, labeled 'User's Input', shows a green line graph representing pitch and a blue horizontal line representing rhythm. Below the graph are two sliders: a 'pitch slider' and a 'rhythm slider', each with a red circular knob. The bottom panel, labeled 'Result', shows the same musical staff as the top panel, but with the second measure filled with a complex sequence of notes. A red oval highlights this second measure. A 'Listen' button is centered below the staff.

Draw&Listen [Benetatos & Duan, 2022]

Bridging Music AI and Music Production

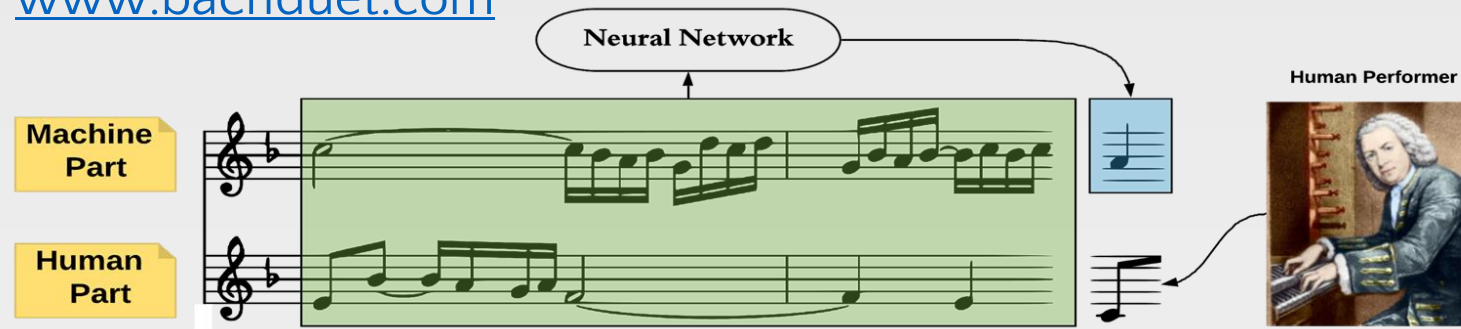
- Music AI models are released on a weekly basis, but most of them only stay in the labs
- Tools for music making (e.g., notation software, DAWs) have limited AI functionalities
- Idea: Build (open-source) tools to bridge the gap
 - Euterpe [Zang et al., 2023]
 - Hosted, Asynchronous, Remote Processing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



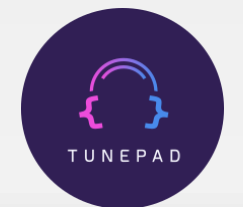
TEAMuP NSF project (2022-2026)

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.”



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 Sarmiento, 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



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., & Wagdy, 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 International 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)