PAT 498/598 (Winter 2025)

Music & AI

Lecture 20: Discussions

Instructor: Hao-Wen Dong

SCHOOL OF MUSIC, THEATRE & DANCE
PERFORMING ARTS TECHNOLOGY
UNIVERSITY OF MICHIGAN

Course Evaluation

- Your feedback is highly appreciated!
- Enter at umich.bluera.com/umich

Project

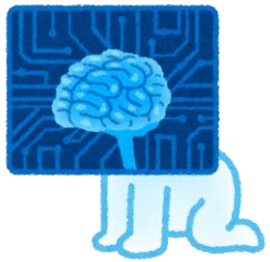
- **Presentation** in class on **Apr 21**
  - **10-min** presentation that summarizes your motivations, methods, results, analysis and discussions
  - You may follow any structure that best suits your narrative
- **Report** due at **11:59pm ET** on **Apr 28**
  - A **2 to 3-page** (excluding references) report that summarizes your motivations, methods, results, analysis and discussions
  - You may use any template
- **No late submissions!** Submit your work early and update it later.

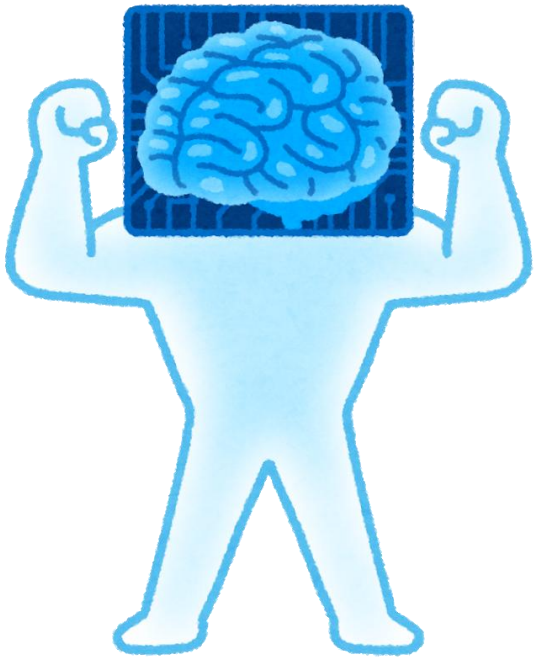
Project Rubrics

- **Presentation (15pt)**
  - Attendance (5pt)
  - Clarity (5pt)
  - Organization and presentation (5pt)
- **Report (15pt)**
  - Writing clarity (5pt)
  - Organization and presentation (5pt)
  - **Discussion (5pt)**
- **Results (15pt)**
  - System/experiment design (5pt)
  - Implementation (5pt)
  - **Experimental/analytic results (5pt)**

🤔 Is current music AI more capable than you thought?

- **Yes** or **No**?
- **Yes**, in what aspects?
- **No**, in what aspects?



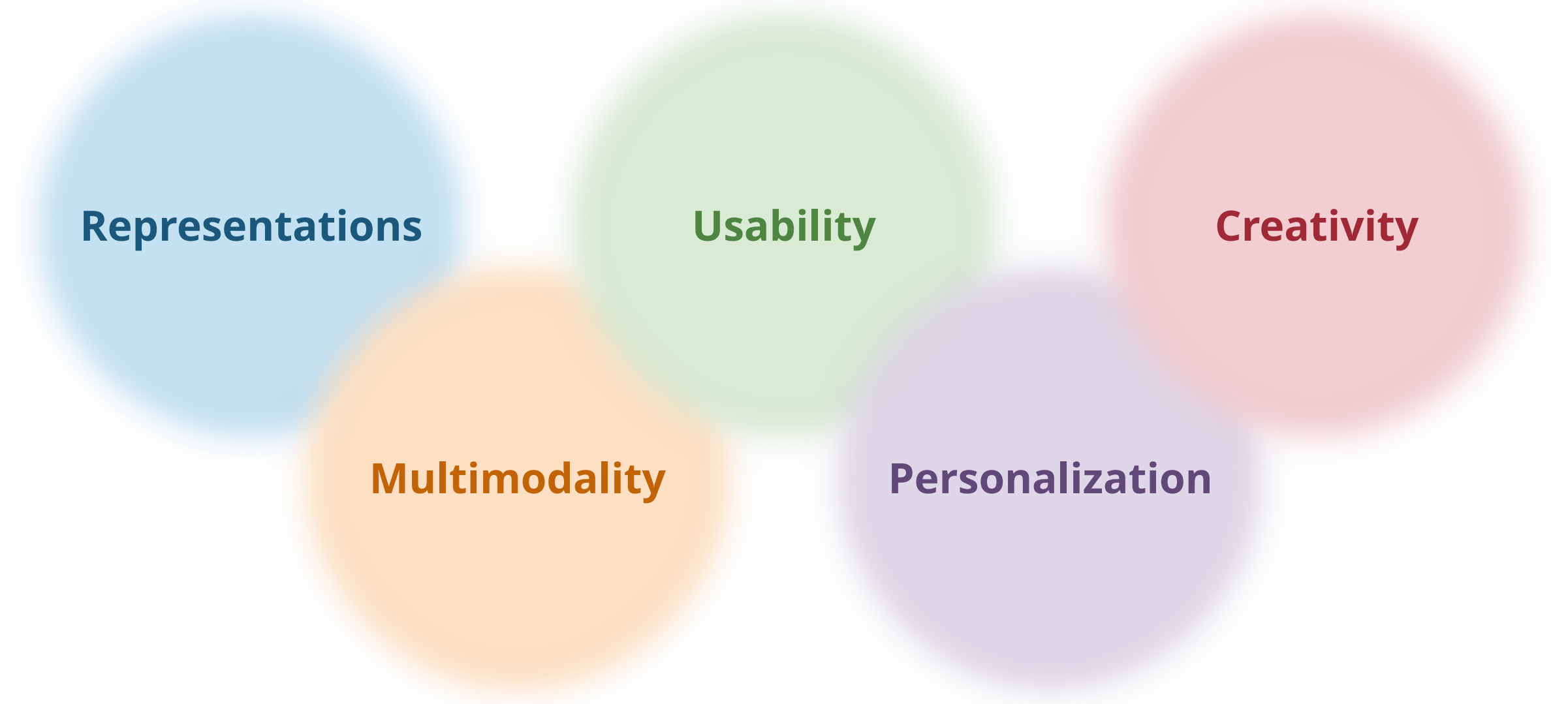


🤔 What AI music tools would you wish for?

- What AI music tools would you like to see **in the next 5 years**?
- What would you consider the **next AI milestone** for music?
- What is the **next grand challenge** of AI music?

Challenges of AI Music

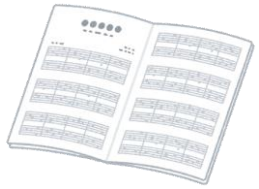
The Five Challenges



Challenge 1: Representations

How can we best represent music for machine learning?

Music Generation – Four Paradigms



Symbolic music generation

Audio-domain music generation

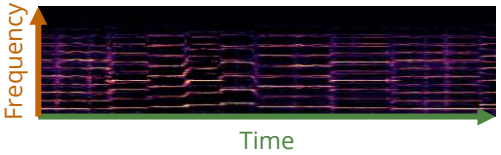
Text-based

Image-based

Time series-based

Image-based

```
Program_change_0,
Note_on_60, Time_shift_2, Note_off_60,
Note_on_60, Time_shift_2, Note_off_60,
Note_on_76, Time_shift_2, Note_off_67,
Note_on_67, Time_shift_2, Note_off_67,
...
```



MIDI

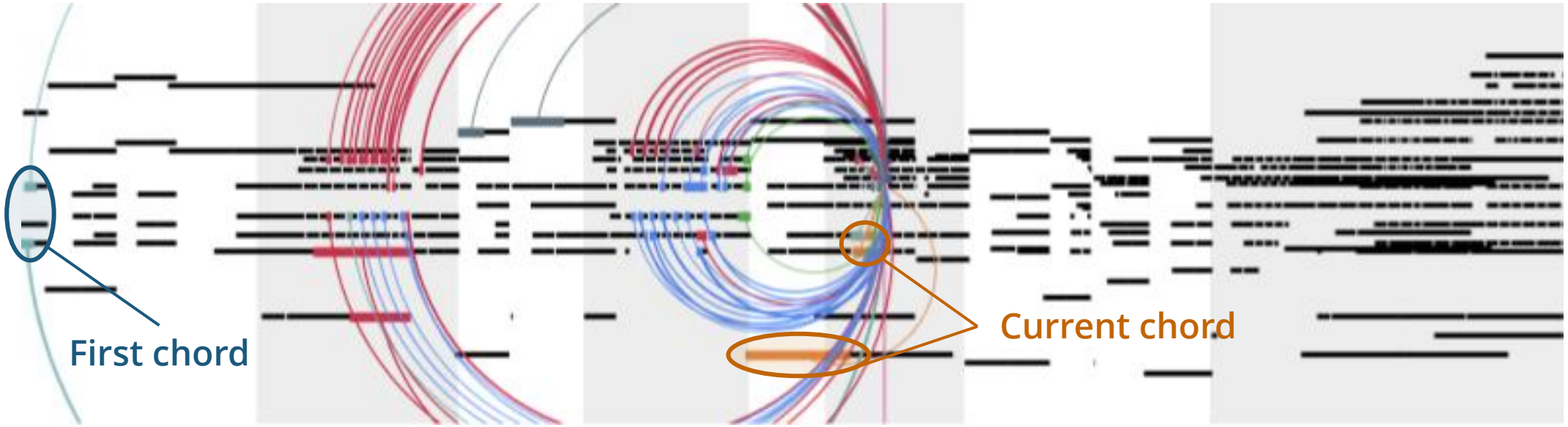
Piano roll

Waveform

Spectrogram

Visualizing Musical Self-attention (Huang et al., 2018)

(Each color represents an attention head)



(Source: Huang et al., 2018)

Cheng-Zhi Anna Huang, Ian Simon, and Monica Dinculescu, "Music Transformer: Generating Music with Long-Term Structure," *Magenta Blog*, December 13, 2018.

Systematically Analyzing Musical Self-attention

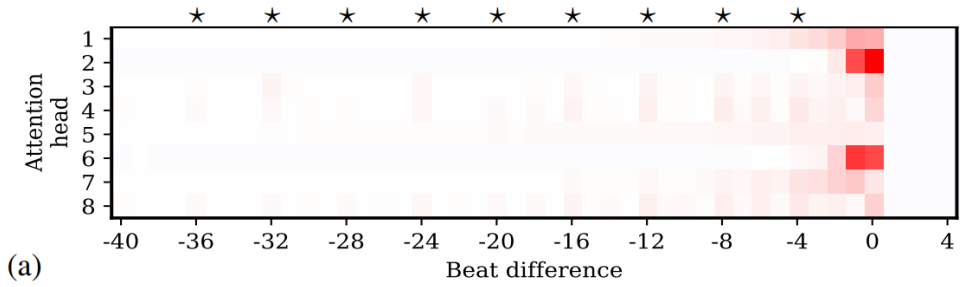
- We proposed two new quantities for measuring **mean relative attention**

$$\gamma_k^{(d)} = \frac{\sum_{\mathbf{x}\in\mathcal{D}} \sum_{s>t} a_{s,t}(\mathbf{x})\, \mathbb{1}_{x_t^{(d)} - x_s^{(d)} = k}}{\sum_{\mathbf{x}\in\mathcal{D}} \sum_{s>t} a_{s,t}(\mathbf{x})} \qquad \tilde{\gamma}_k^{(d)} = \gamma_k^{(d)} - \frac{\sum_{\mathbf{x}\in\mathcal{D}} \sum_{s>t} \mathbb{1}_{x_t^{(d)} - x_s^{(d)} = k}}{\sum_{\mathbf{x}\in\mathcal{D}} \sum_{s>t} 1}$$

- The MMT model attends more to notes

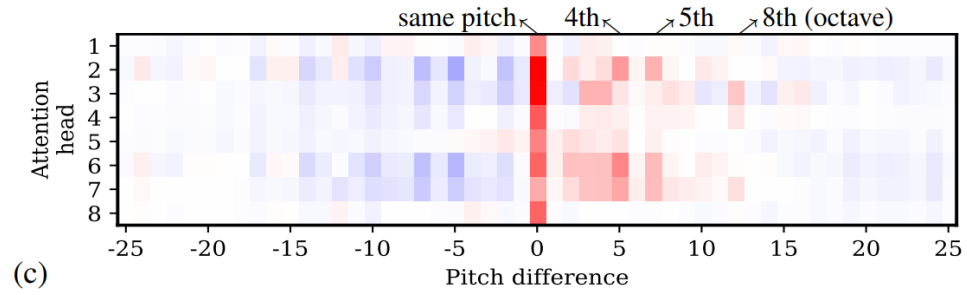
that are $4N$ beats away in the past

Positive and negative mean relative attention gain

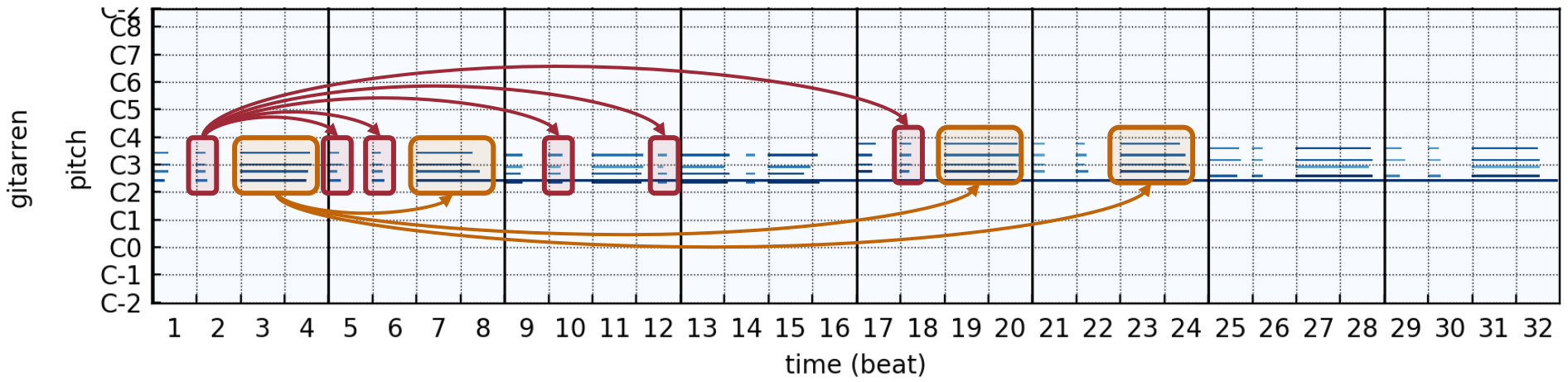


that has a pitch in an octave above which forms a consonant interval

Positive and negative mean relative attention gain



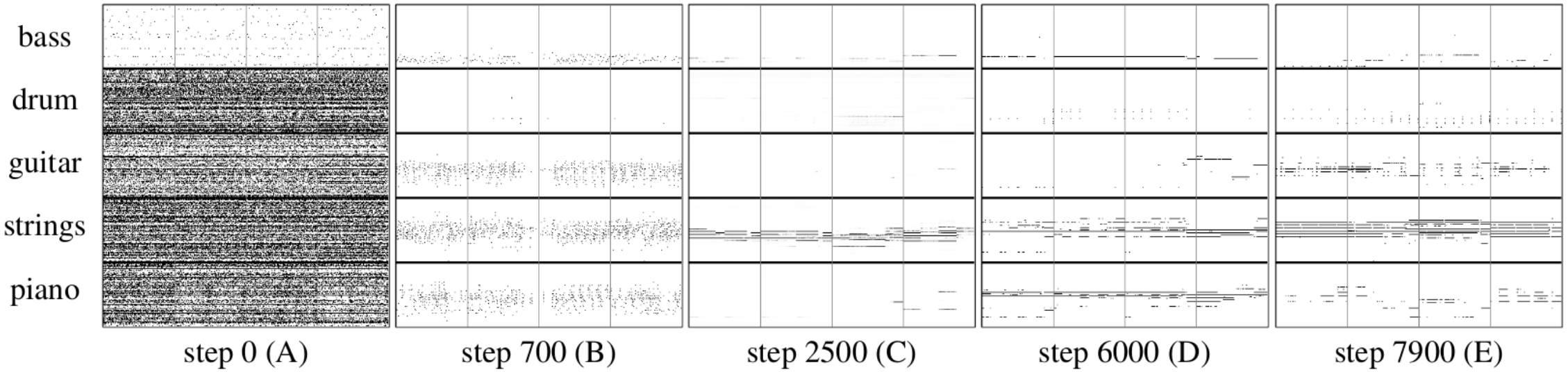
Why Piano Rolls?



Many musical patterns like melodies, chords, scales and arpeggios are translational invariant in the temporal and pitch axes

MuseGAN (Dong et al., 2018)

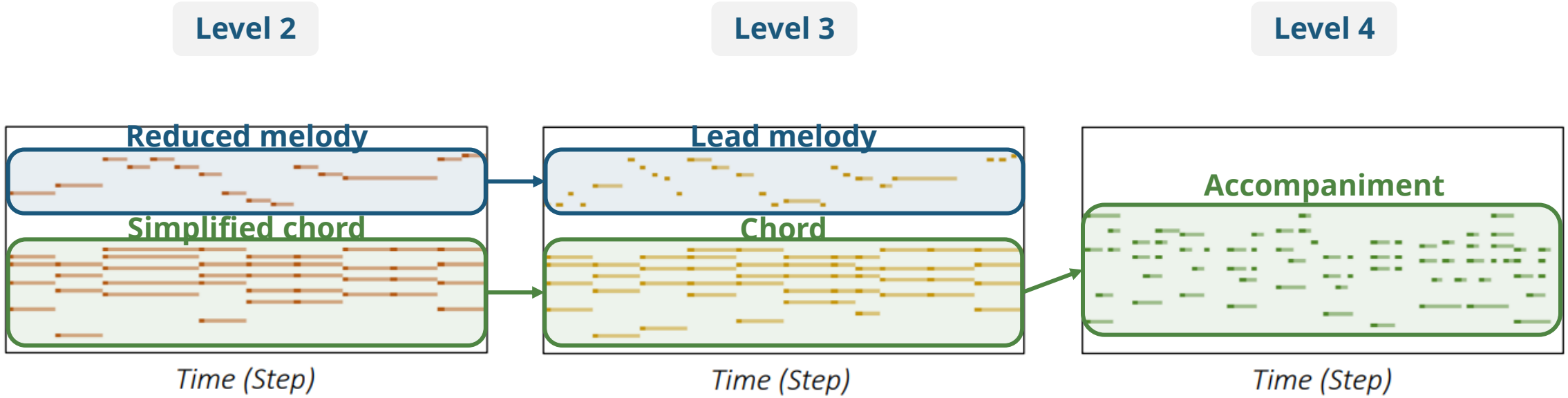
Examples of generated music



(Source: Dong et al., 2018)

Hao-Wen Dong, Wen-Yi Hsiao, Li-Chia Yang, and Yi-Hsuan Yang, "MuseGAN: Multi-track Sequential Generative Adversarial Networks for Symbolic Music Generation and Accompaniment," *AAAI*, 2018.

Example: Cascaded Diffusion Models (Wang et al., 2024)



(Source: Wang et al., 2024)

wholesonggen.github.io

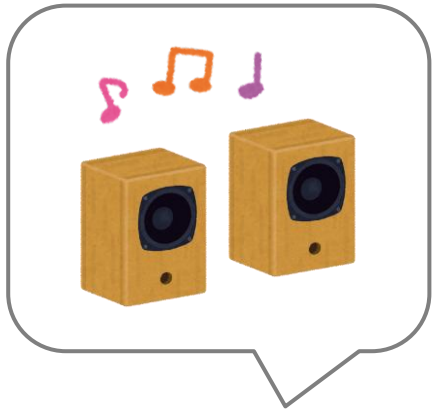
Ziyu Wang, Lejun Min, and Gus Xia, "Whole-Song Hierarchical Generation of Symbolic Music Using Cascaded Diffusion Models," *ICLR*, 2024.

Challenge 2: Multimodality

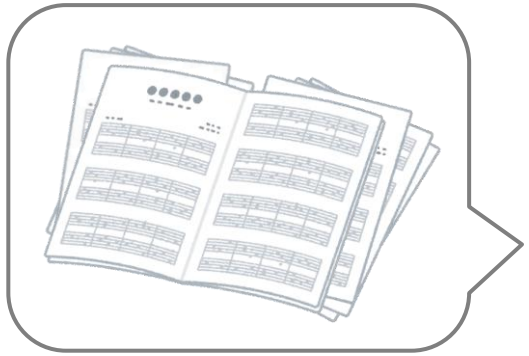
Can AI learn to create music by “listening to” music rather than “reading” music?

Human-inspired Machine Learning for Music & Audio

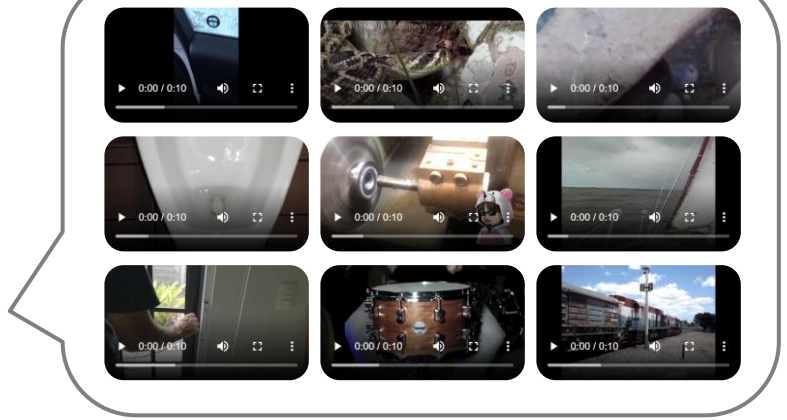
Learning from listening



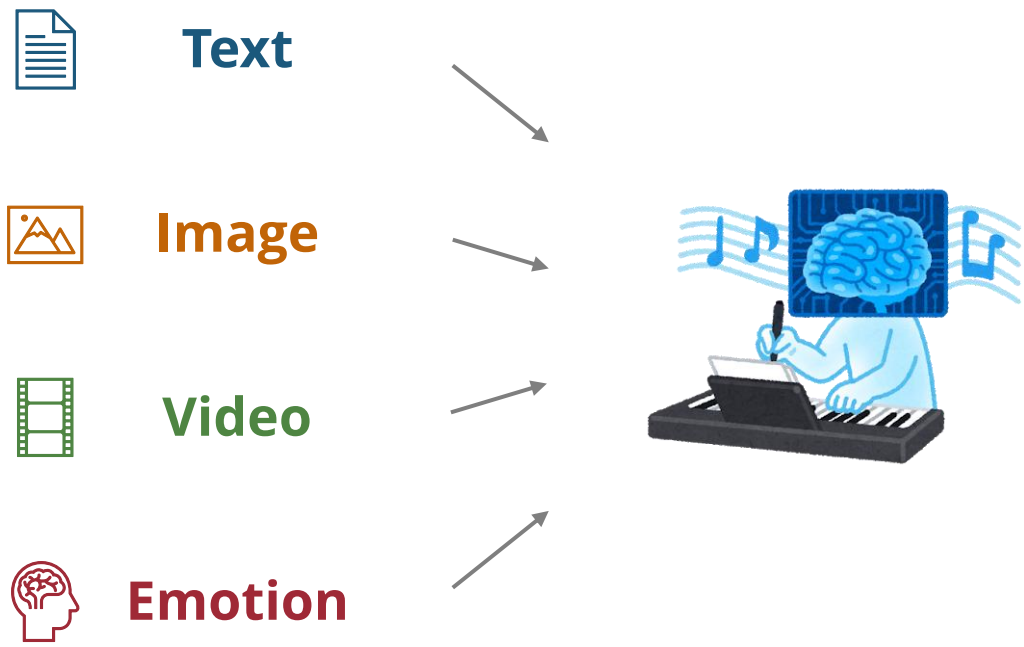
Learning from reading



Learning from watching



Multimodal Inputs for Generative Music AI



AI Creative Agents (2015)

youtu.be/DggF9m9xqik & github.com/DYCI2/Dicy2

Shimon: An Improvising Robotic Marimba Player (2021)

(Source: Robot Gizmos)

roboticgizmos.com/shimon-musical-robot-deep-learning/

Georgia Tech Center for Music Technology, "Shimon with the Aarhus Jazz Orchestra," *YouTube*, youtu.be/ZpTV1-acSU8, 2021.

Challenge 3: Usability

How can AI music tools be integrated into an artist's creative workflow?

WavJourney: Compositional Audio Creation (Liu et al., 2023)

Large language models
(GPT-4)

Pretrained generative audio models
(MusicGen, AudioLDM, Bark)

Instructions → **Audio Script** → **Audio**

Audio Type	Layout	ID	Character	Volume	Action	Content Description	Duration
Music	Background	1	N/A	-30	Begin	Dramatic orchestral news theme.	Auto
Speech	Foreground	N/A	Host	-15	N/A	Welcome to Mars News ...	Auto
Music	Background	1	N/A	N/A	End	N/A	Auto
Speech	Foreground	N/A	Host	-15	N/A	Now let's connect with our on-site reporter ...	Auto
Sound effect	Foreground	N/A	N/A	-35	N/A	Transition swoosh.	1
Sound effect	Background	2	N/A	-30	Begin	Background noise of busy engineering office.	Auto
Speech	Foreground	N/A	Reporter	-15	N/A	We're here at the headquarters of …	Auto
Speech	Foreground	N/A	Director	-15	N/A	Thank you, so it's a fantastic …	Auto
Speech	Foreground	N/A	Reporter	-15	N/A	This is truly an impressive feat …	Auto

Interactable intermediate outputs

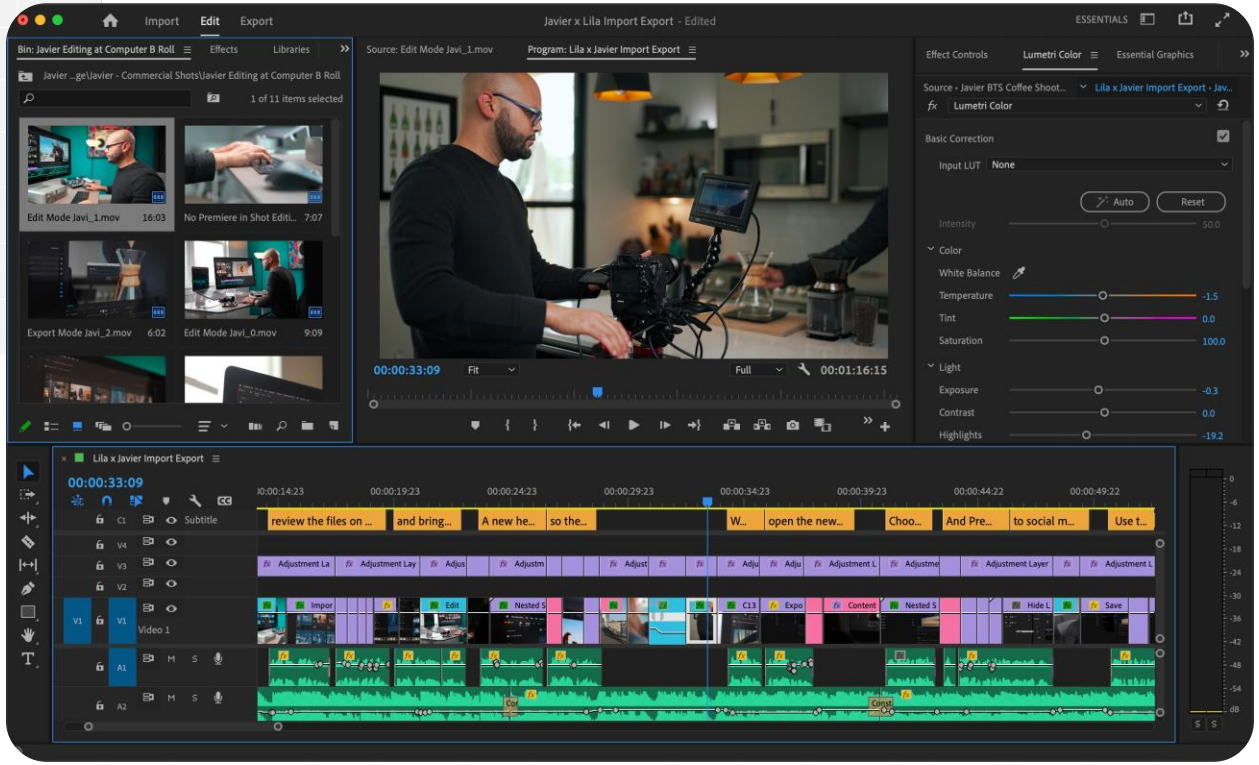
(Source: Liu et al., 2023)

Xubo Liu, Zhongkai Zhu, Haohe Liu, Yi Yuan, Meng Cui, Qiushi Huang, Jinhua Liang, Yin Cao, Qiuqiang Kong, Mark D. Plumbley, Wenwu Wanga, "WavJourney: Compositional Audio Creation with Large Language Models," *arXiv preprint arXiv:2307.14335*, 2023.

Integrating Generative AI into the Creative Workflow

Audio Type	Layout	ID	Character	Volume	Action	Content Description	Duration
Music	Background	1	N/A	-30	Begin	Dramatic orchestral news theme.	Auto
Speech	Foreground	N/A	Host	-15	N/A	Welcome to Mars News ...	Auto
Music	Background	1	N/A	N/A	End	N/A	
Speech	Foreground	N/A	Host	-15	N/A	Now let's connect with our on-site reporter ...	
Sound effect	Foreground	N/A	N/A	-35	N/A	Transition swoosh.	
Sound effect	Background	2	N/A	-30	Begin	Background noise of busy engineering office.	
Speech	Foreground	N/A	Reporter	-15	N/A	We're here at the headquarters of …	
Speech	Foreground	N/A	Director	-15	N/A	Thank you, so it's a fantastic …	
Speech	Foreground	N/A	Reporter	-15	N/A	This is truly an impressive feat …	

(Source: Liu et al., 2023)



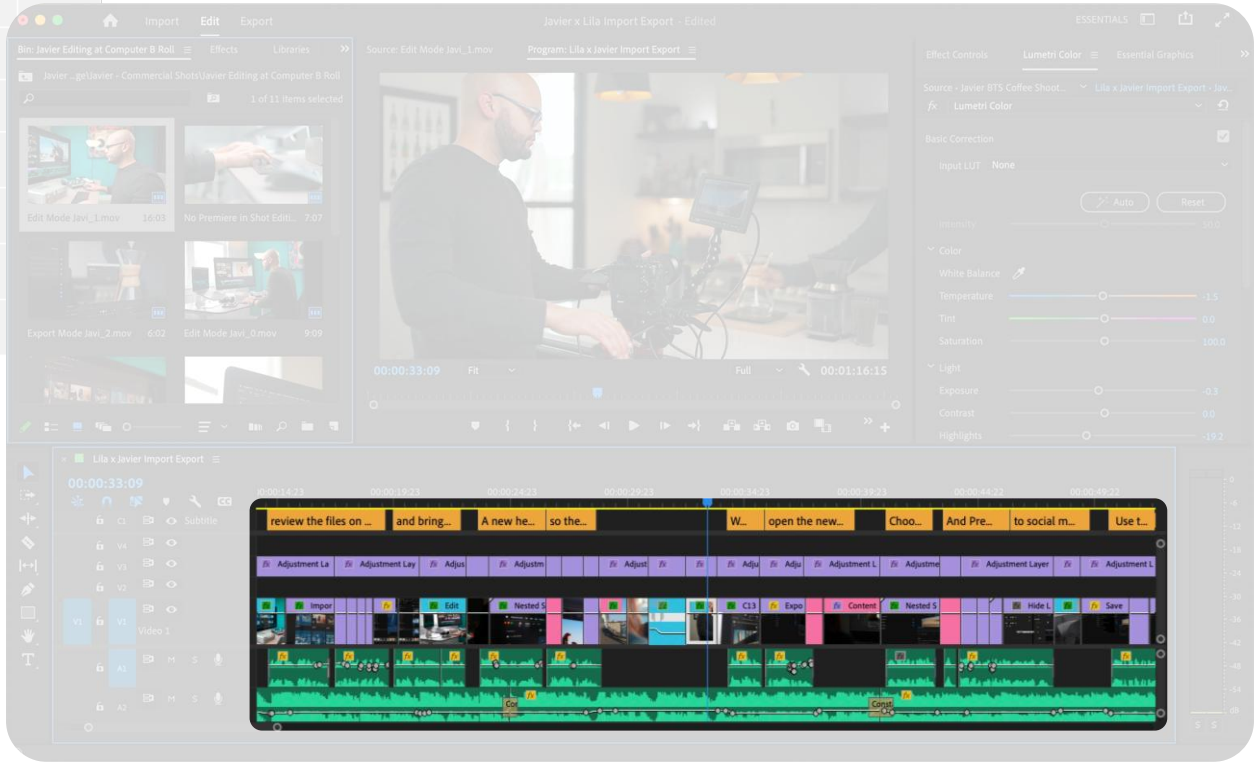
Xubo Liu, Zhongkai Zhu, Haohe Liu, Yi Yuan, Meng Cui, Qiushi Huang, Jinhua Liang, Yin Cao, Qiuqiang Kong, Mark D. Plumbley, Wenwu Wanga, "WavJourney: Compositional Audio Creation with Large Language Models," *arXiv preprint arXiv:2307.14335*, 2023.

Integrating Generative AI into the Creative Workflow

Audio Type	Layout	ID	Character	Volume	Action	Content Description	Duration
Music	Background	1	N/A	-30	Begin	Dramatic orchestral news theme.	Auto
Speech	Foreground	N/A	Host	-15	N/A	Welcome to Mars News ...	Auto
Music	Background	1	N/A	N/A	End	N/A	
Speech	Foreground	N/A	Host	-15	N/A	Now let's connect with our on-site reporter ...	
Sound effect	Foreground	N/A	N/A	-35	N/A	Transition swoosh.	
Sound effect	Background	2	N/A	-30	Begin	Background noise of busy engineering office.	
Speech	Foreground	N/A	Reporter	-15	N/A	We're here at the headquarters of …	
Speech	Foreground	N/A	Director	-15	N/A	Thank you, so it's a fantastic …	
Speech	Foreground	N/A	Reporter	-15	N/A	This is truly an impressive feat …	

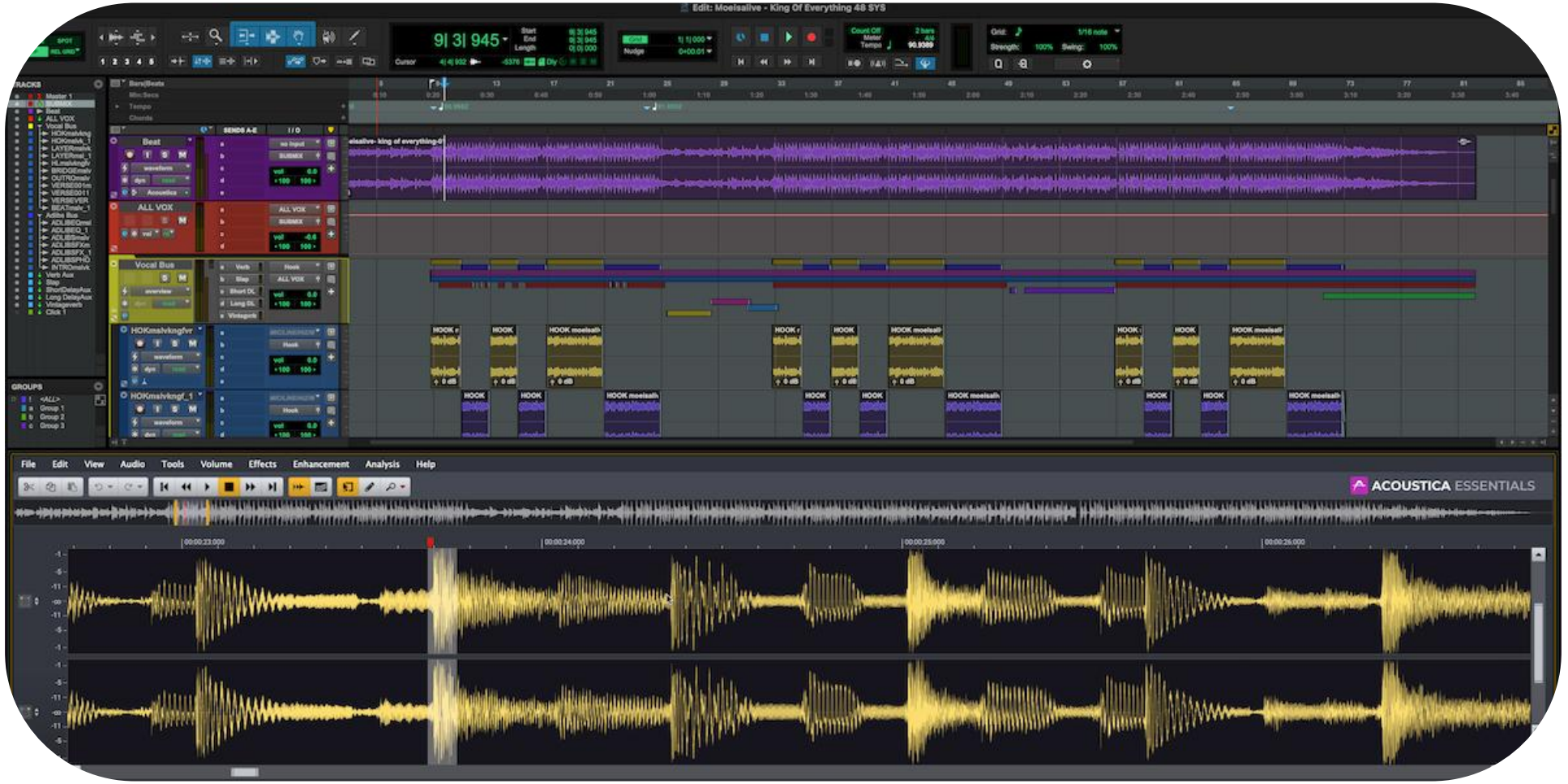
(Source: Liu et al., 2023)

Integration into professional creative workflow



Xubo Liu, Zhongkai Zhu, Haohe Liu, Yi Yuan, Meng Cui, Qiushi Huang, Jinhua Liang, Yin Cao, Qiuqiang Kong, Mark D. Plumbley, Wenwu Wanga, "WavJourney: Compositional Audio Creation with Large Language Models," *arXiv preprint arXiv:2307.14335*, 2023.

Integrating GenAI into the Music Creative Workflow



(Source: Avid)

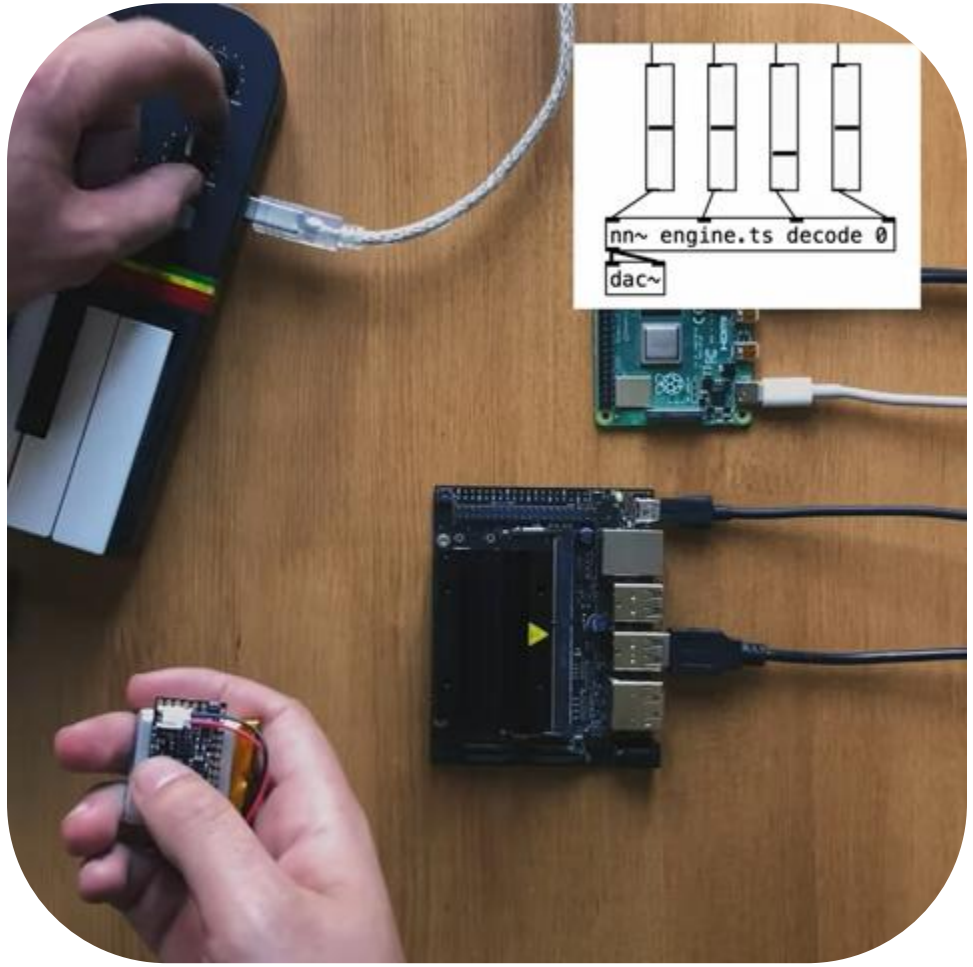
avid.com/pro-tools/whats-new

Integrating GenAI into the Music Creative Workflow

(Source: Avid)

avid.com/pro-tools

RAVE: Real-time Audio Synthesis (Caillon & Esling, 2022)



youtu.be/jAIRf4nGgYI

Antoine Caillon and Philippe Esling, "RAVE: A variational autoencoder for fast and high-quality neural audio synthesis," *arXiv preprint arXiv:2111.05011*, 2021.

Misusable Music Tools (Nao Tokui, 2024)

Throughout history, music and technology have often intertwined, with **new technologies being misused by artists** (turntables, etc).

– Nao Tokui, 2024

AI is more challenging to misuse because **it lacks a physical entity and operates as a black box**.

– Nao Tokui, 2024

(Source: Flintmi via Wikimedia Commons)

x.com/naotokui_en/status/1756180102344499338
Nao Tokui, “Surfing human creativity with AI — A user's guide,” *Lulu Press*, 2023.
Flintmi, CC BY 3.0, via Wikimedia Commons

Without **deviation from the norm**, progress is not possible.

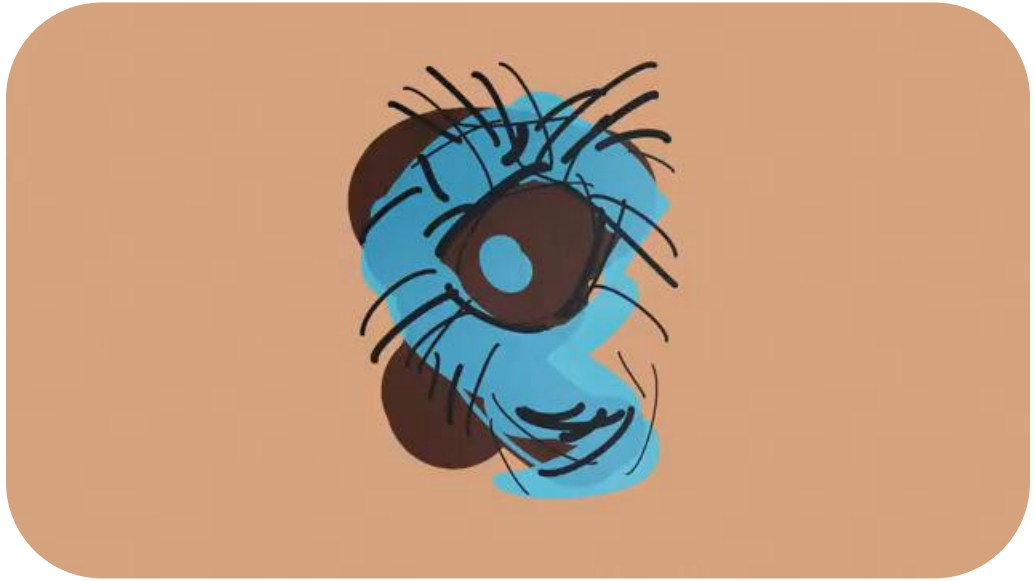
– Frank Zappa

Challenge 4: Personalization

How can we make “my personal AI music tools”?

YACHT & Google Magenta

"**The band first took all 82 songs from their back catalog** and isolated each part, from bass lines to vocal melodies to drum rhythms; they then took those isolated parts and broke them up into four-bar loops. Then, **they put those loops into the machine learning model**, which **put out new melodies based on their old work**. **They did a similar process with lyrics**, **using their old songs plus other material they considered inspiring**. The final task was to pick lyrics and melodies that made sense, and pair them together to make a song."



youtu.be/_yz8QYzcfxI

YACHT, "YACHT — SCATTERHEAD (4K Lyric Video)", *YouTube*, July 26, 2019.
Megan Friedman, "Behind Magenta, the tech that rocked I/O," *The Keyword*, May 20, 2019.
Adam Roberts, "YACHT's new album is powered by ML + Artists", *Magenta Blog*, September 13, 2019.

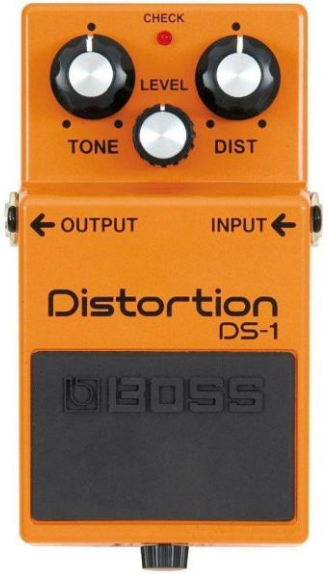
Ease of Personalization for Artists

- Through **finetuning our own models**
- Through **finetuning with live inputs**
- Python scripting vs friendly user interface

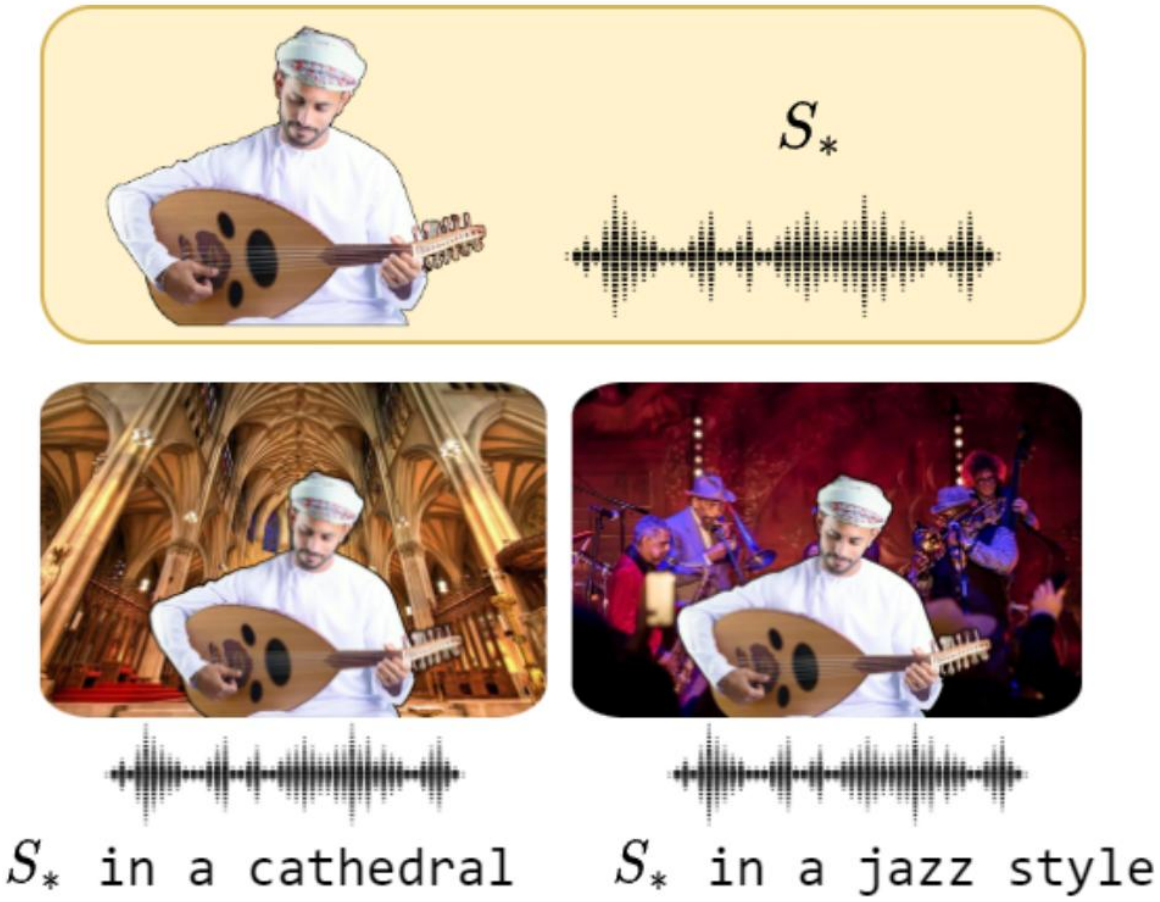
- **Can we do better?**

Overfitting vs Distortion

- Will **overfitting** be a new music expression, the "**distortion**" for AI music?



Personalized Text-to-Music Generation (Plitsis et al., 2024)



S* A disco song with a **S***

S* A recording of a **S*** song with a hip hop drum beat accompaniment

(Source: Plitsis et al., 2024)

Manos Plitsis, Theodoros Kouzelis, Georgios Paraskevopoulos, Vassilis Katsouros, and Yannis Panagakis, "Investigating Personalization Methods in Text to Music Generation," *ICASSP*, 2024.

Challenge 5: Creativity

Can AI ever be creative? How can AI augment human creativity?

Works of art make rules;
rules do not make works of art.

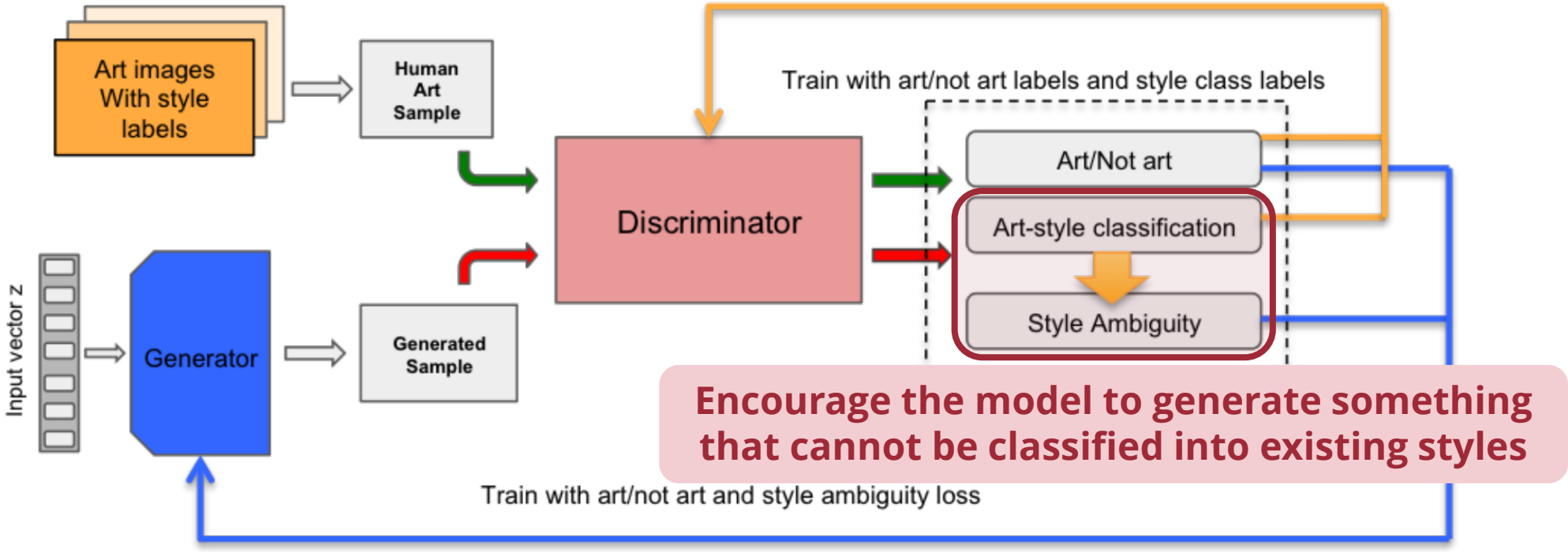
– Claude Debussy

The Curse of Machine Learning

- As the old saying goes, "**Artificial intelligence is only as good as the data it learns from**."
- Machine learning models are trained to approximate some distribution in its formal formulation.
- This seems to contradict the idea of creativity that requires **the ability to extrapolate** and **think out of the box**.

- **Can AI ever be creative?**

Creative Adversarial Network (Elgammal et al., 2017)

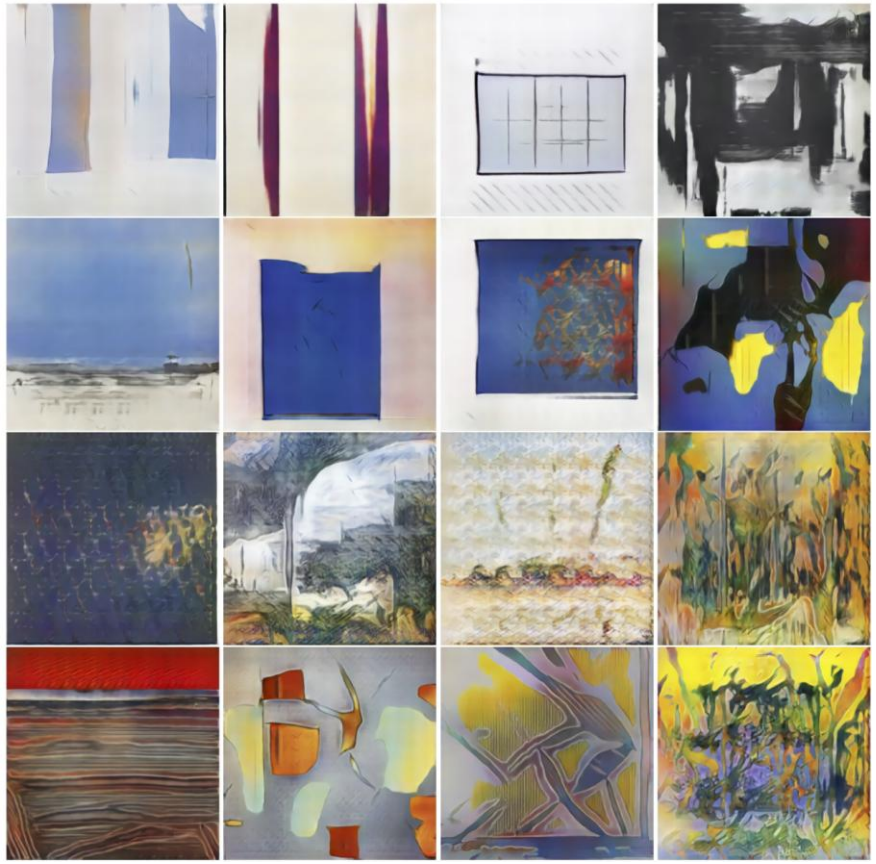


(Source: Elgammal et al., 2017)

Ahmed Elgammal, Bingchen Liu, Mohamed Elhoseiny, and Marian Mazzone, "CAN: Creative Adversarial Networks, Generating "Art" by Learning About Styles and Deviating from Style Norms," *ICCC*, 2017.

Creative Adversarial Network (Elgammal et al., 2017)

Example generated images



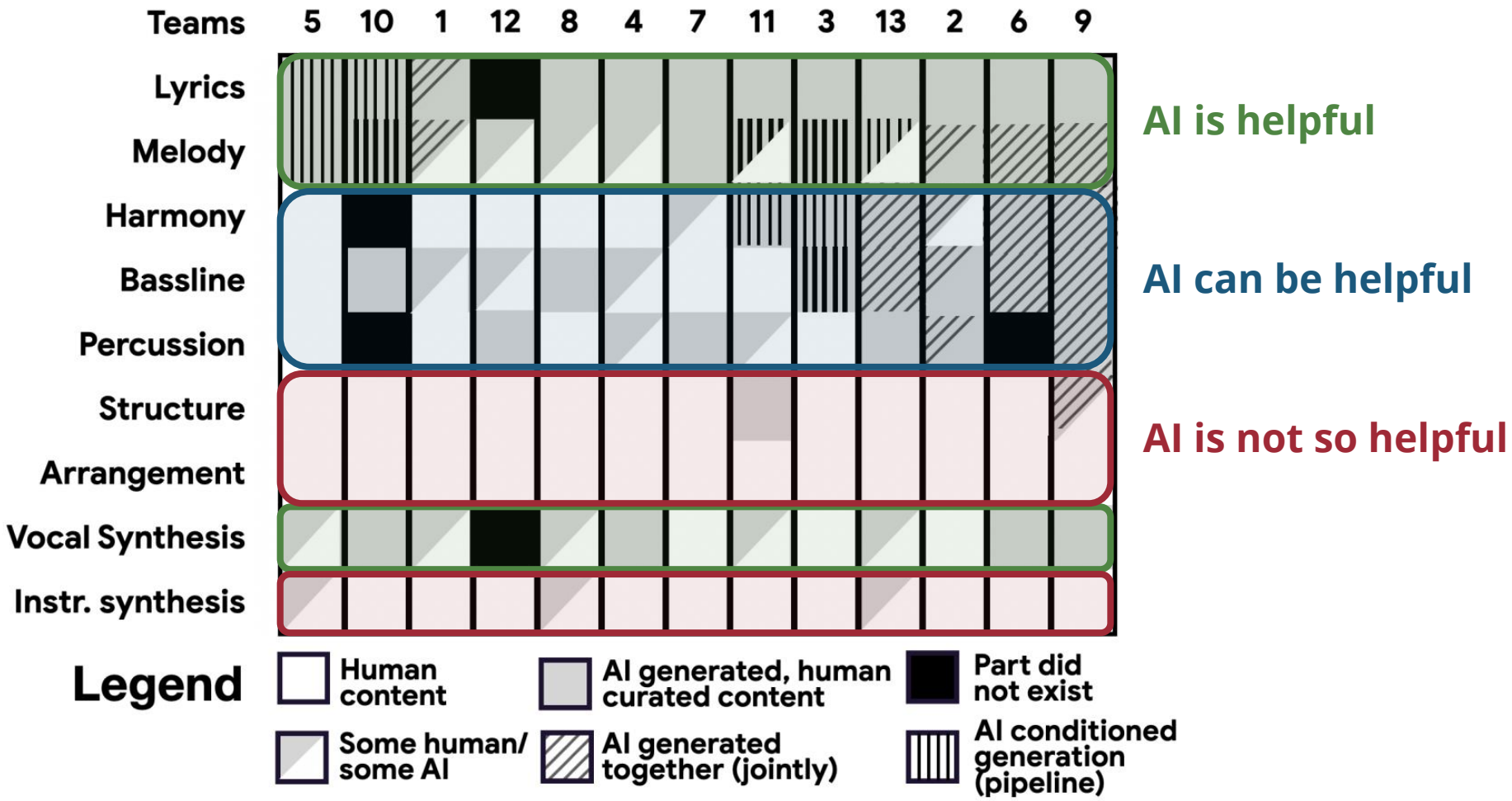
(Source: Elgammal et al., 2017)

Best samples

(Source: Elgammal et al., 2017)

Ahmed Elgammal, Bingchen Liu, Mohamed Elhoseiny, and Marian Mazzone, "CAN: Creative Adversarial Networks, Generating "Art" by Learning About Styles and Deviating from Style Norms," *ICCC*, 2017.

How can AI Augment Human Creativity?



(Source: Huang et al., 2020)

Cheng-Zhi Anna Huang, Hendrik Vincent Koops, Ed Newton-Rex, Monica Dinculescu, and Carrie J. Cai, "AI Song Contest: Human-AI Co-Creation in Songwriting," *ISMIR*, 2020.

What is Creativity?

> Creativity is the ability to come up with ideas or artefacts that are **new**, **surprising** and **valuable**.
>
> – Margaret Boden, 2007

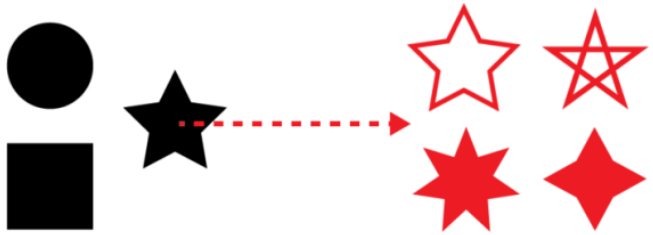
Three Types of Creativity (Boden, 2007)

Combinatorial Creativity

Combining existing ideas and things into something new

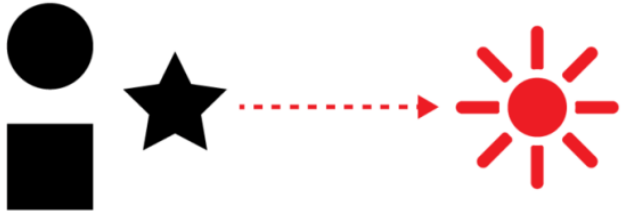
Exploratory Creativity

Exploring possibilities within a domain



Transformative Creativity

Radically new ideas that redefine domain and applicable rules



(Source: van Kuijk, 2023)

linkedin.com/pulse/creativity-ungraspable-mystic-proces-thus-can-ai-jasper-van-kuijk/

Margaret A. Boden, "Creativity in a nutshell," *Think*, 5(15):83–96, 2007.

AI is Good at Combinatory Creativity

Prompt: A Michigan space wolverine

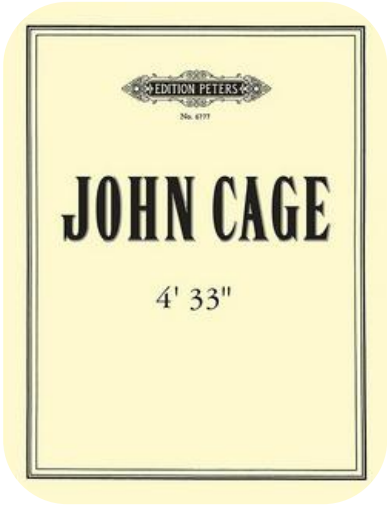
Generated by ImageFX

Can AI learn Transformative Creativity?

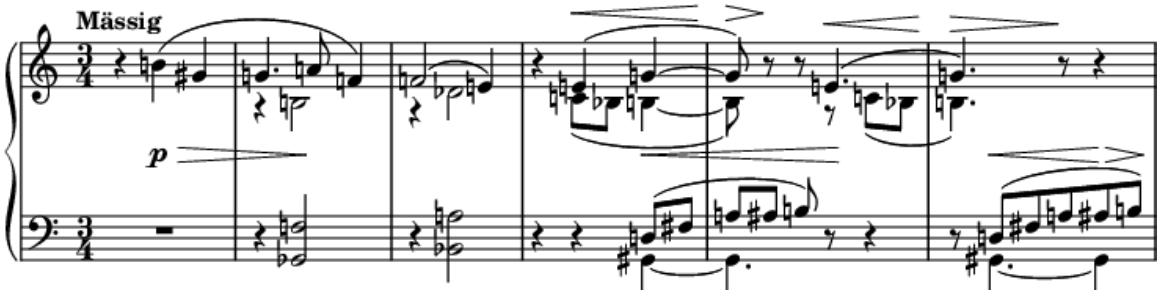
Duchamp & conceptual art

Picasso & cubism

Cage & modernist music



Schoenberg & atonality



John Cage's *4'33"* (1952)

They missed the point. **There's no such thing as silence**. What they thought was silence, because they didn't know how to listen, was full of accidental sounds. You could hear the wind stirring outside during the first movement. During the second, raindrops began pattering the roof, and during the third the people themselves made all kinds of interesting sounds as they talked or walked out.

– John Cage, on the premiere of 4'33", 1952

Reading: David Cope on Emily Howell (2020)

- Tim Adams, "David Cope: 'You pushed the button and out came hundreds and thousands of sonatas'," *The Observer*, *The Guardians,* July 10, 2010.

"**People tell me they don't hear soul in the music**," he says. "When they do that, I pull out a page of notes and ask them to show me where the soul is. We like to think that what we hear is soul, but I think audience members put themselves down a lot in that respect. **The feelings that we get from listening to music are something we produce, it's not there in the notes. It comes from emotional insight in each of us, the music is just the trigger.**"

– David Cope, 2010

Creativity vs Art

Creativity is allowing yourself to **make mistakes**.
Art is knowing **which ones to keep**.

– Scott Adams

Reading: Can Computer Create Arts?

youtu.be/HPMCWtoC_rM

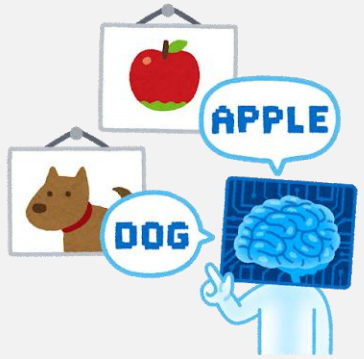
The Five Challenges

- **Representations**: How can we best represent music for machine learning?
- **Multimodality**: Can AI learn to create music by “listening to” music rather than “reading” music?
- **Usability**: How can AI music tools be integrated into an artist’s creative workflow?
- **Personalization**: How can we make “my personal AI music tools”?
- **Creativity**: Can AI ever be creative? How can AI augment human creativity?

What is AI?

(Recap) What is Artificial Intelligence?

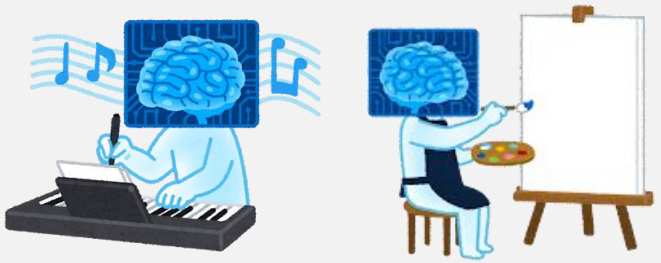
Systems that think like humans



Systems that think rationally

Systems that act like humans

Systems that act rationally



🤔 What is your expectation of a Real AI?

- **Thinking vs Acting**
  - An AI **needs to know how to think**
  - An AI **doesn't need to know how to think**

- **Human vs AI**
  - An AI **needs to behave like a human**
  - An AI **doesn't need to behave like a human**

🤔 What is your expectation of a Real AI?

- **Thinking vs Acting**
  - An AI **needs to know how to think** **(Votes: 10)**
  - An AI **doesn't need to know how to think** **(Votes: 3)**

- **Human vs AI**
  - An AI **needs to behave like a human** **(Votes: 0)**
  - An AI **doesn't need to behave like a human** **(Votes: 13)**

Copyright

Purpose of Copyright

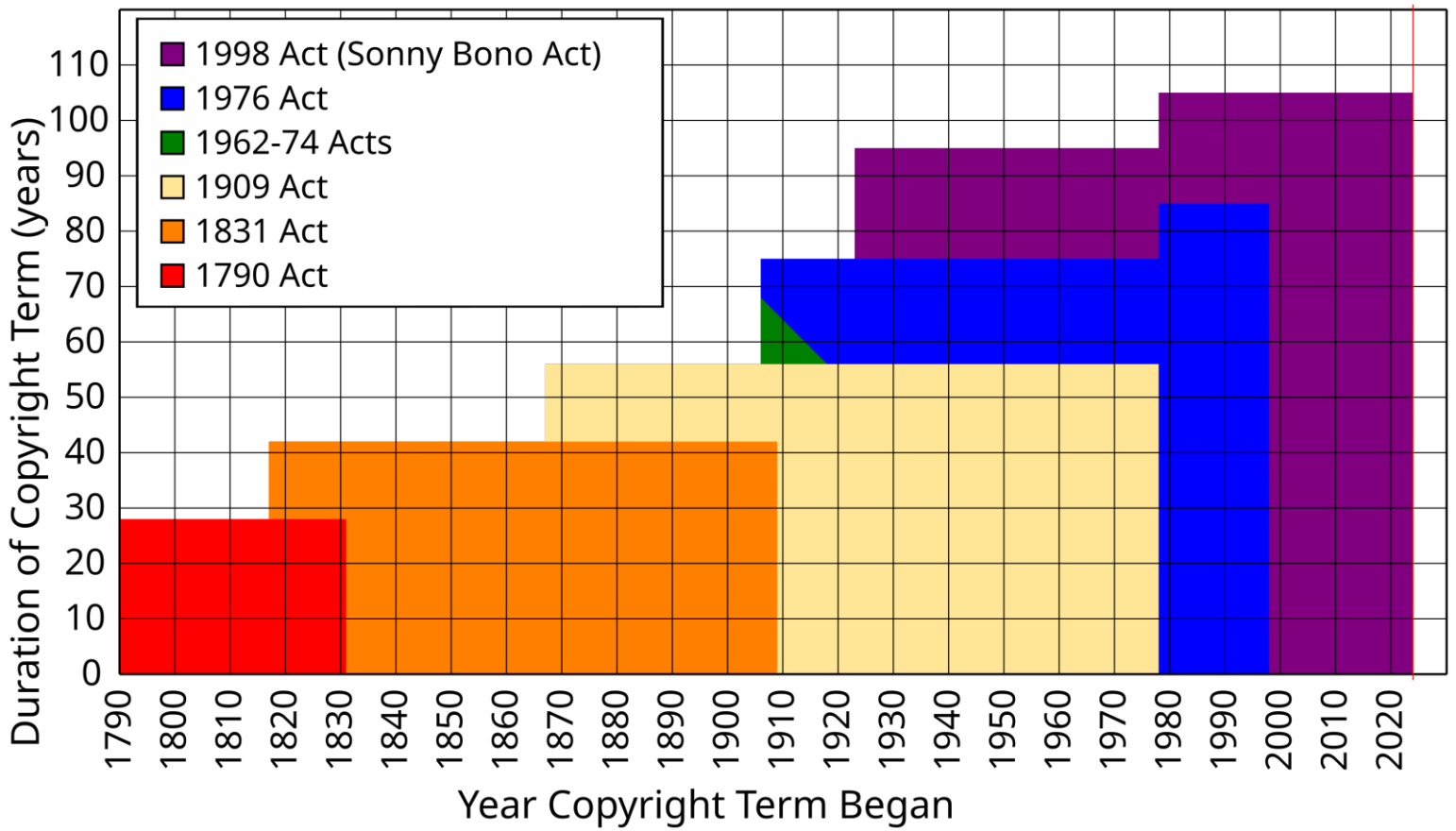
- According to Copyright Alliance:
  - "The primary purpose of copyright is to **induce and reward authors**, through the provision of property rights, to create new works and to make those works available to the public to enjoy."
  - "The theory is that by granting certain exclusive rights to creators that allow these creators to protect their creative works against theft, **creators receive the benefit of economic rewards** and **the public receives the benefit of the creative works** that might not otherwise be created or disseminated."

copyrightalliance.org/education/copyright-law-explained/copyright-basics/purpose-of-copyright/

Paul McCartney on AI & Copyright

youtu.be/DpebXMTlYNo

Expansion of Copyright Law



(Source: Tom Bell via Wikimedia Commons)

🤔 Purpose of Copyright

- What about **meme culture**?
- Think about **portrait vs photography**
  - While photography displace realistic portraits, it also **frees portrait painting into new art forms** like impressionism and abstract arts
  - Will AI music **raise the collective standards of music** for the public?

Fairly Trained: L Certification

- All of the **training data used for the model(s)** being certified must fall into one of the following categories:
  - Be **explicitly provided to the model developer for the purposes of being used as training data**, according to a contractual agreement with a party that has the rights required to enter such an agreement
  - Be available under an **open license** appropriate to the use-case
  - Be in the **public domain globally**
  - Be **fully owned by the model developer**

Fairly Trained Certified

fairlytrained.org

Fairly Trained Certified Companies


beatoven.ai

fairlytrained.org/certified-models

Fairly Trained Certified Products & Models

fairlytrained.org/certified-models

🤔 Music Gen AI & Copyright

- How does AI-generated music compare to other industry-disrupting technologies like **recording** and **music sampling**?
- Does model training justify **fair use**?
- Is a music Gen AI model **a derivative of its training dataset**?
- Can a music Gen AI model **reproduce a song** in its training dataset?

Reading: How Will the Law Handle Generative AI?

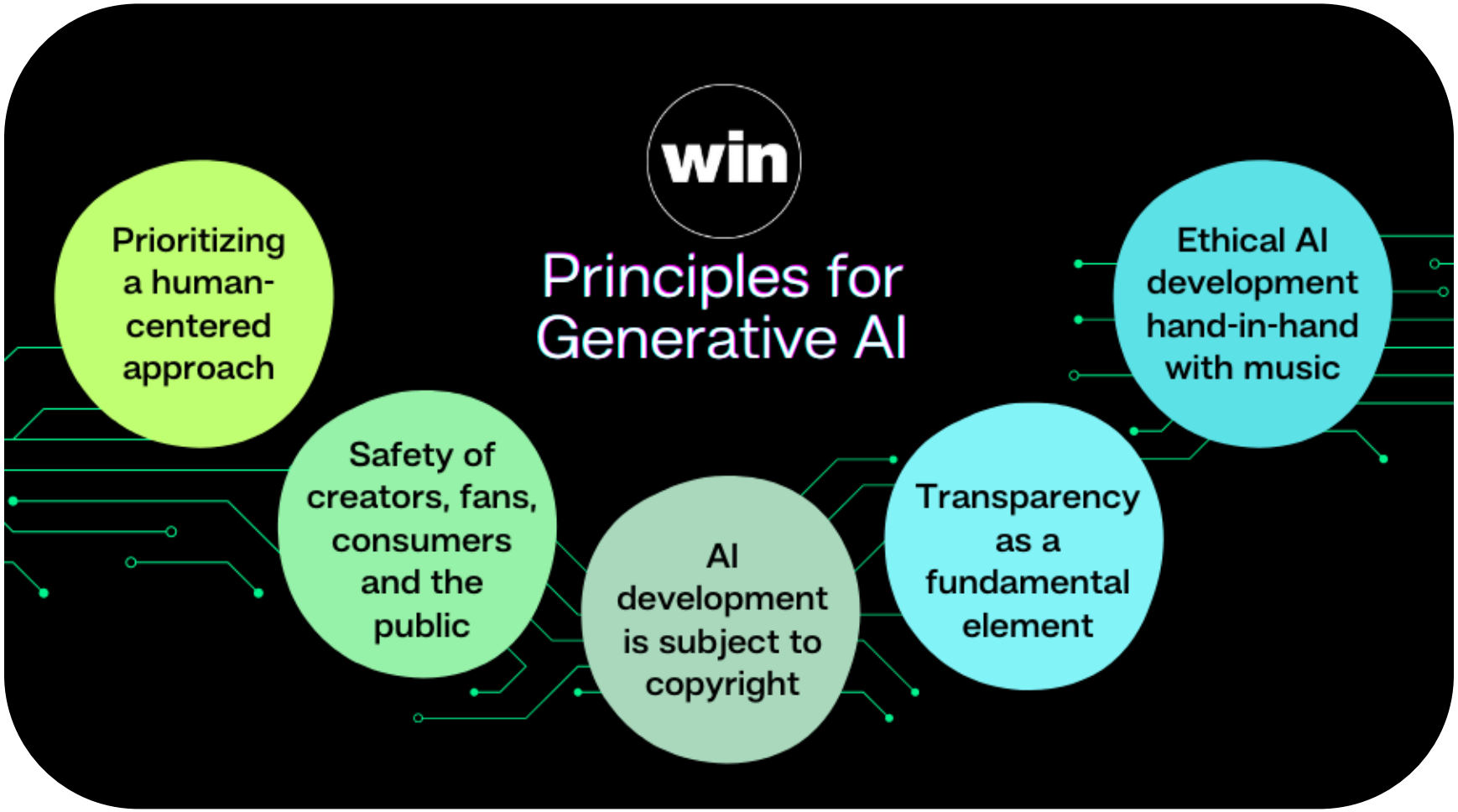
youtu.be/VxJXPWFoYoc

Ethical Considerations

Ethics vs Laws

> Ethics is doing **more than the law requires** and **less than the law allows**.
>
> – Michael Josephson

Principles for GenAI (Worldwide Independent Network)



winformusic.org/ai-principles/

Ethical Implications of Music GenAI (Barnett, 2023)

- **Loss of agency and authorship**
- **Creativity stifling**
  - The repetitive nature of the music generation and that by limiting the creative output to possibilities of the model may result in a similar bound on human creativity
- **Predominance of western bias**
- **Copyright infringement**
- **Cultural appropriation**
  - Generative models make it easier to use content from marginalized cultures without any accompanying investment in or engagement from the community

Julia Barnett, "The Ethical Implications of Generative Audio Models: A Systematic Literature Review," *AIES*, 2023.

Ethical Implications of Speech GenAI (Barnett, 2023)

- **Phishing and fraud**
- **Misinformation and deepfakes**
- **Security and privacy**
  - The potential for risk to security and privacy of individuals as a result of speech generative models, especially when they only require small segments of training data to produce a realistic voice of a targeted speaker
- **Non-consensual use of biometric data**

Julia Barnett, "The Ethical Implications of Generative Audio Models: A Systematic Literature Review," *AIES*, 2023.

Homework 1: Real of Fake!?

AI Drake: “Heart on My Sleeve” (Ghostwriter977, 2023)

youtu.be/S2qxvg9NNPM

This is NOT Morgan Freeman (2021)

youtu.be/oxXpB9pSETo

This is NOT Morgan Freeman (2021)

youtu.be/F4G6GNFz0O8

Reading: Ethical Implications

- Andre Holzapfel, Bob L. Sturm, and Mark Coeckelbergh "Ethical Dimensions of Music Information Retrieval Technology," *TISMIR*, 1(1):44–55, 2018.
- Rujing Huang, Bob L. T. Sturm, and Andre Holzapfel, "De-centering the West: East Asian Philosophies and the Ethics of Applying Artificial Intelligence to Music," *ISMIR*, 2021.
- Rujing Stacy Huang, Andre Holzapfel, Bob L. T. Sturm, and Anna-Kaisa Kaila, "Beyond Diverse Datasets: Responsible MIR, Interdisciplinarity, and the Fractured Worlds of Music," *TISMIR*, 6(1):43–59, 2023.
- Julia Barnett, "The Ethical Implications of Generative Audio Models: A Systematic Literature Review," *AIES*, 2023.

Reading: Diversity in Music Information Retrieval

- Xavier Serra, Martin Clayton, and Barış Bozkurt, "Computational Approaches for Analysis of Non-Western Music Traditions," *ISMIR Tutorials*, 2018.
- Georgina Born, "Diversifying MIR: Knowledge and Real-World Challenges, and New Interdisciplinary Futures," *TISMIR*, 3(1):193–204 2020.
- Lorenzo Porcaro, Carlos Castillo, and Emilia Gómez "Diversity by Design in Music Recommender Systems," *TISMIR*, 4(1):114–126, 2021.

Environmental Concerns

GPU Energy Consumption

V100

250W

A100

400W

RTX 5090

575W

Multiple GPUs

(Source: Supermicro)

2-6kW

~1kW

~1.5kW

~2kW

Training BLOOM (A 176B Parameter LLM) (Luccioni et al., 2022)

Total training time	118 days, 5 hours, 41 min
Total number of GPU hours	1,082,990 hours
Total energy used	433,196 kWh
GPU models used	Nvidia A100 80GB
Carbon intensity of the energy grid	57 gCO_2eq/kWh

(Source: Luccioni et al., 2023)

Average household power usage:
10.5 MWh / year

A household for 40+ years
(or **40+ households for a year**)

Model name	Number of parameters	Datacenter PUE	Carbon intensity of grid used	Energy consumption	CO_2eq emissions	CO_2eq emissions × PUE
GPT-3	175B	1.1	429 gCO_2eq/kWh	1,287 MWh	*502 tonnes*	552 tonnes
Gopher	280B	1.08	330 gCO_2eq/kWh	*1,066 MWh*	*352 tonnes*	380 tonnes
OPT	175B	*1.09* [2]	*231gCO_2eq/kWh*	*324 MWh*	70 tonnes	*76.3 tonnes* [3]
BLOOM	176B	1.2	57 gCO_2eq/kWh	433 MWh	25 tonnes	30 tonnes

(Source: Luccioni et al., 2023)

Alexandra Sasha Luccioni, Sylvain Viguier, and Anne-Laure Ligozat, "Estimating the Carbon Footprint of BLOOM, a 176B Parameter Language Model," *JMLR*, 24(253):1–15, 2023.
eia.gov/energyexplained/use-of-energy/electricity-use-in-homes.php

The Growing GPU Needs

Tom's Hardware

First in-depth look at Elon Musk's 100,000 GPU AI cluster — xAI Colossus reveals its secrets

YouTuber ServeTheHome was granted access to the Supermicro servers within the 100,000 GPU beast, showing off several facets of the supercomputer...

Oct 28, 2024

Tom's Hardware

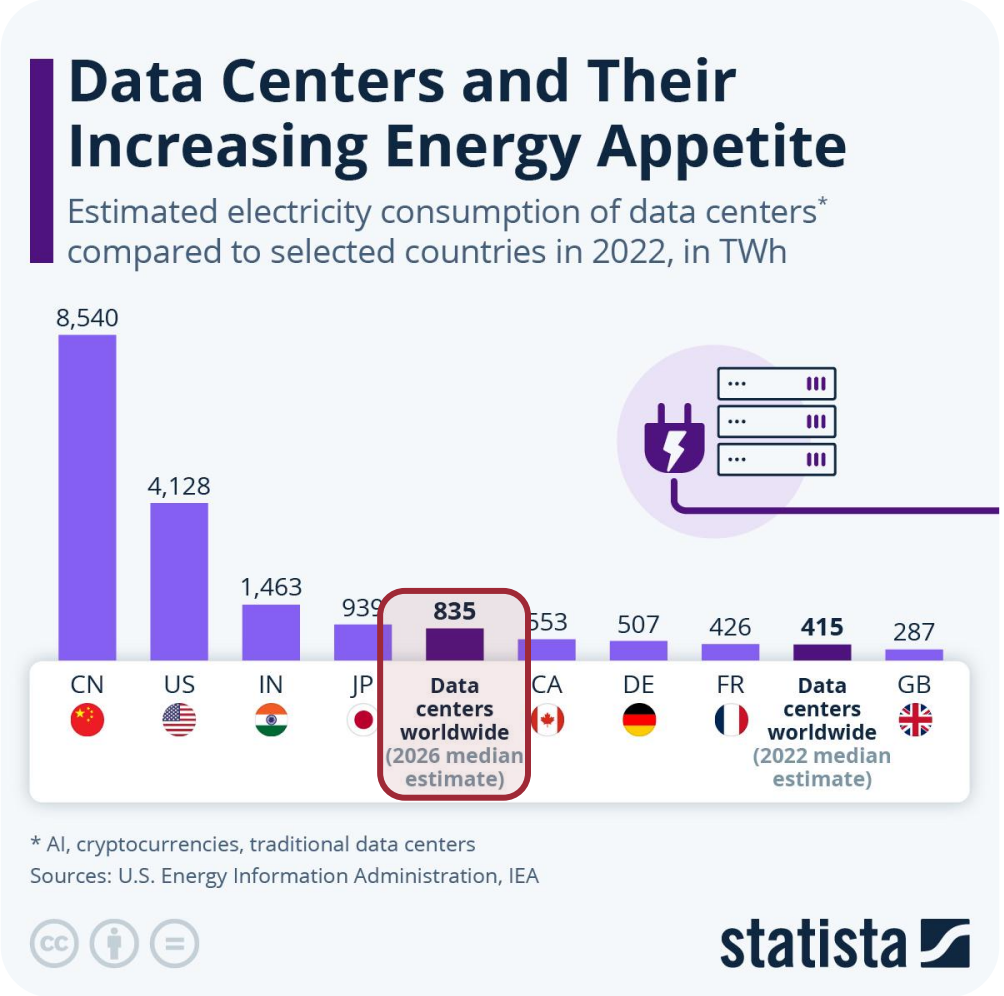
Meta is using more than 100,000 Nvidia H100 AI GPUs to train Llama-4 — Mark Zuckerberg says that Llama 4 is being trained on a cluster "bigger than anything that I've seen"

Meta is using more than 100,000 Nvidia H100 AI GPUs to train Llama-4 — Mark Zuckerberg says that Llama 4 is being trained on a cluster "bigger...

Oct 31, 2024

tomshardware.com/desktops/servers/first-in-depth-look-at-elon-musks-100-000-gpu-ai-cluster-xai-colossus-reveals-its-secrets
tomshardware.com/tech-industry/artificial-intelligence/meta-is-using-more-than-100-000-nvidia-h100-ai-gpus-to-train-llama-4-mark-zuckerberg-says-that-llama-4-is-being-trained-on-a-cluster-bigger-than-anything-that-ive-seen

Increasing Energy Consumption of Data Centers

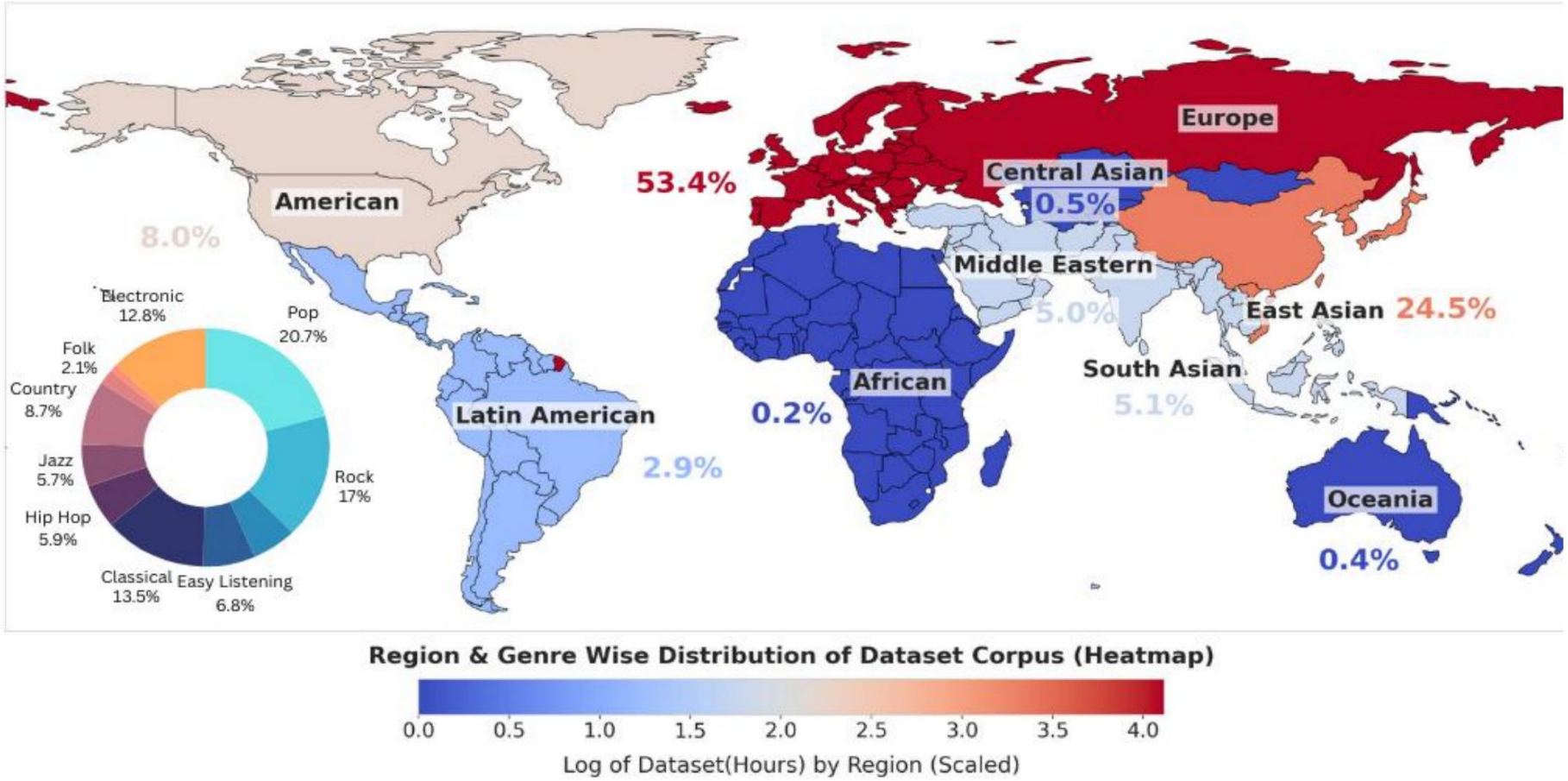


(Source: Statista)

statista.com/chart/32689/estimated-electricity-consumption-of-data-centers-compared-to-selected-countries/

Cultural Concerns

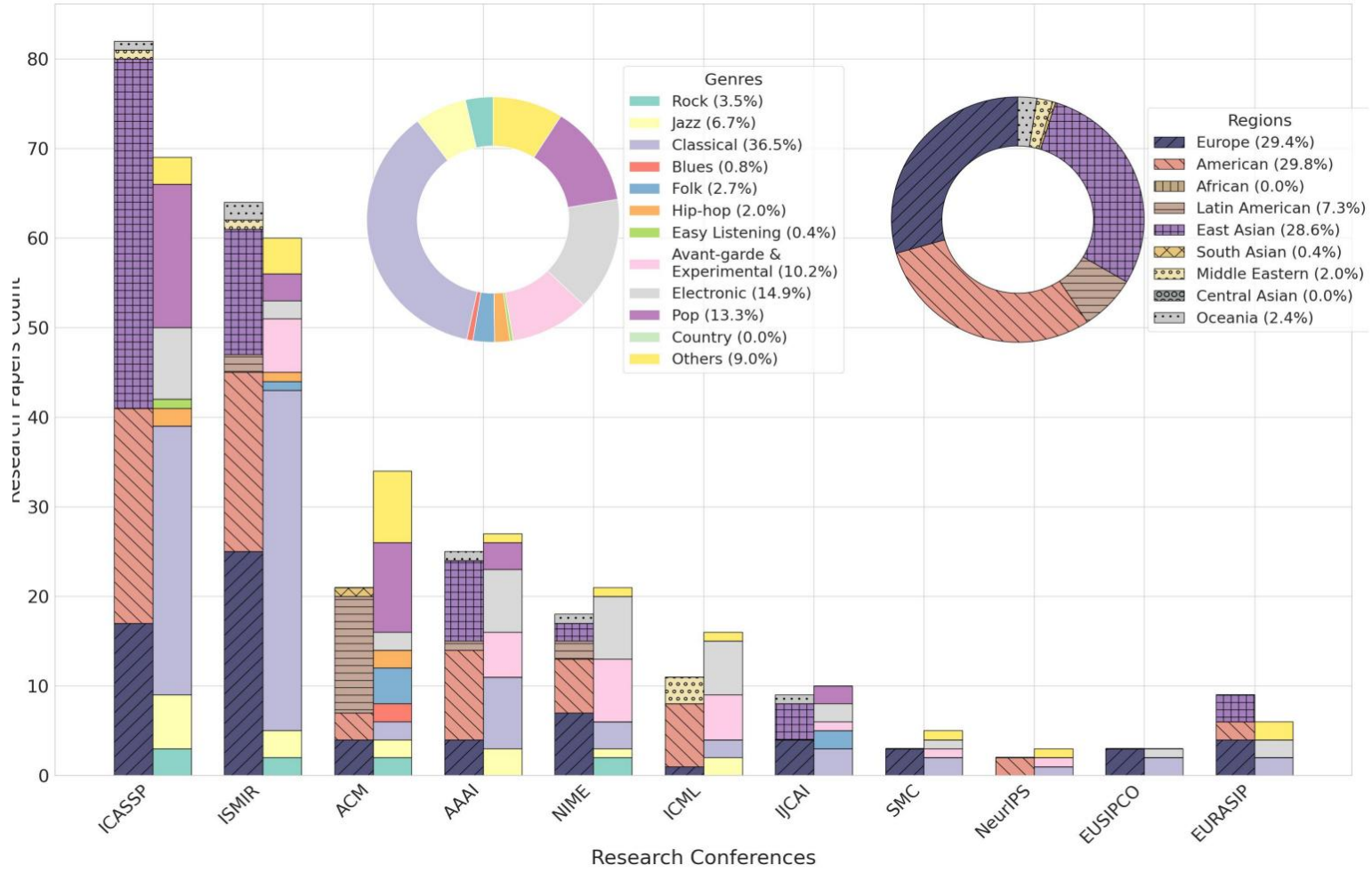
Global Divide in AI Music Datasets (Mehta et al., 2024)



(Source: Mehta et al., 2024)

Atharva Mehta, Shivam Chauhan, and Monojit Choudhury, "Missing Melodies: AI Music Generation and its "Nearly" Complete Omission of the Global South," *arXiv preprint arXiv:2412.04100*, 2024.

Global Divide in AI Music Research (Mehta et al., 2024)



(Source: Mehta et al., 2024)

Atharva Mehta, Shivam Chauhan, and Monojit Choudhury, “Missing Melodies: AI Music Generation and its "Nearly" Complete Omission of the Global South,” *arXiv preprint arXiv:2412.04100*, 2024.

Implications

- **Limiting Global South creativity**
  - Limits the potential for Global South music genres to evolve and adapt in the digital age
- **Economic disparities**
  - Worsens the economic disparities within the music industry
- **Reinforcement of existing biases**
  - Reinforces existing cultural biases, perpetuating a cycle where Global South music is viewed as less important or less valuable
- **Cultural erosion**
  - Genres such as Hindustani folk and traditional Arabic Maqam that represent centuries-old traditions, philosophies, and artistic expressions could fade from the public mind

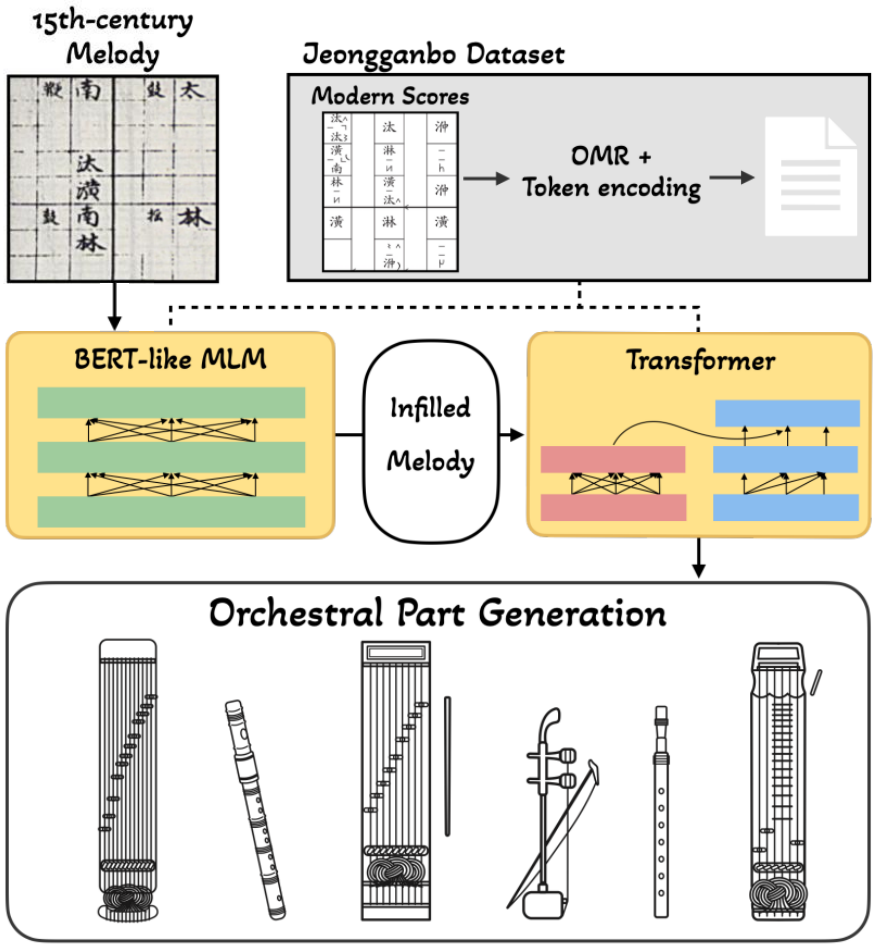
Atharva Mehta, Shivam Chauhan, and Monojit Choudhury, "Missing Melodies: AI Music Generation and its "Nearly" Complete Omission of the Global South," *arXiv preprint arXiv:2412.04100*, 2024.

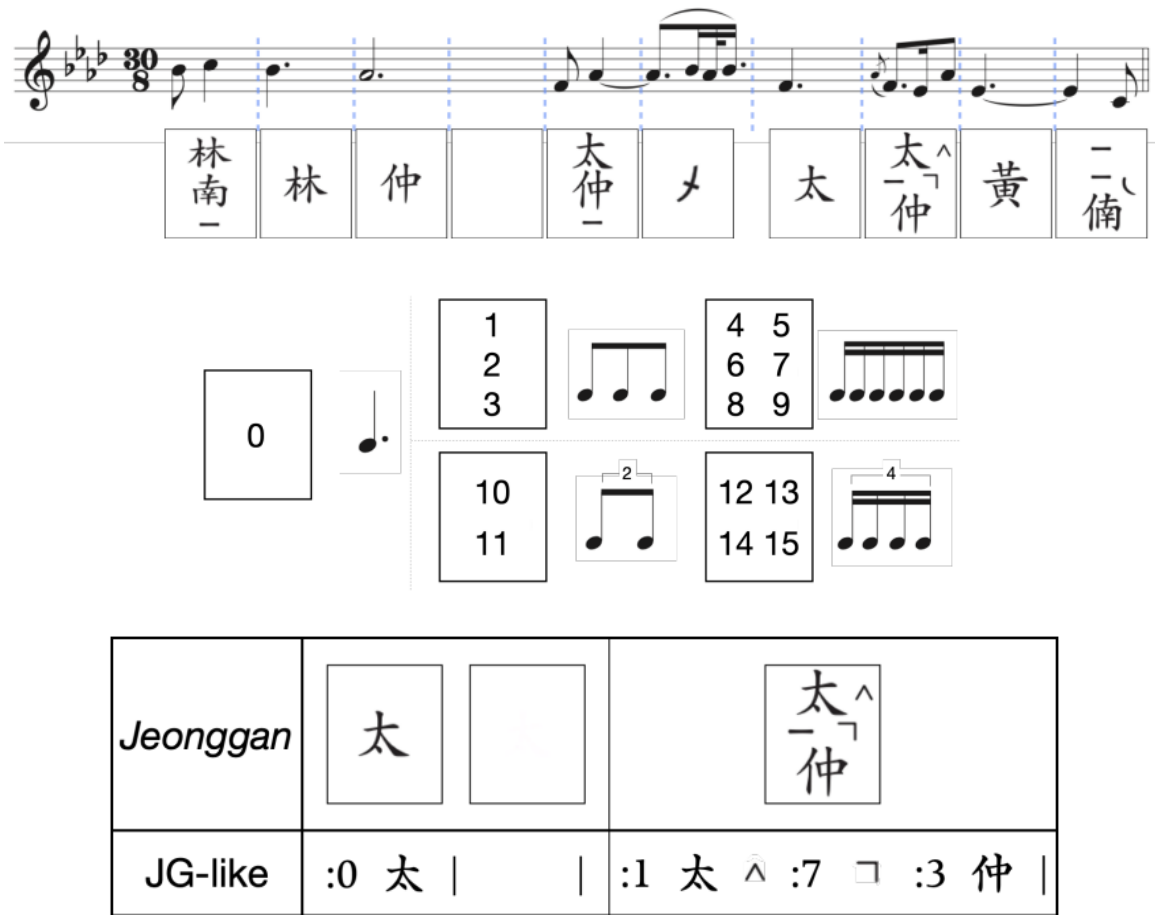
Recommendations

- **Explicit mention** of genres and model limitations
- **Avoid generation** when uncertainty exists
- Investing on **inclusive datasets**
- **Transfer learning** for underrepresented styles
- **Inclusive evaluation**

Atharva Mehta, Shivam Chauhan, and Monojit Choudhury, “Missing Melodies: AI Music Generation and its "Nearly" Complete Omission of the Global South,” *arXiv preprint arXiv:2412.04100*, 2024.

Reviving Korean Court Music with AI (Han et al., 2024)





(Source: Han et al., 2024)

Danbinaerin Han, Mark Gotham, Dongmin Kim, Hannah Park, Sihun Lee, and Dasaem Jeong, "Six Dragons Fly Again: Reviving 15th-Century Korean Court Music with Transformers and Novel Encoding," *ISMIR*, 2024.

Six Dragons Fly Again (2024)

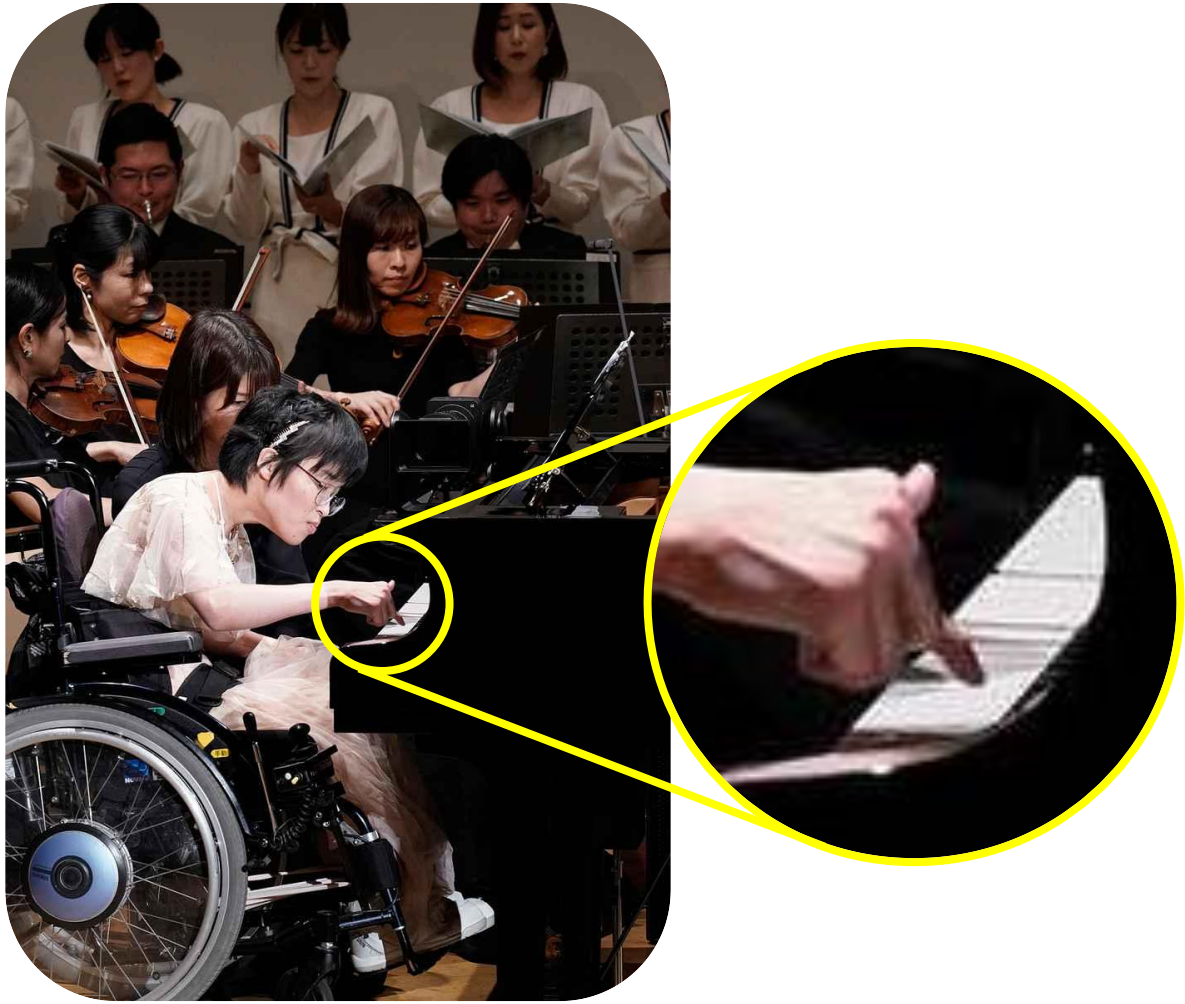
youtu.be/7zS1FSG7dcg

Let’s Talk about Something Positive!

Robotic Drum Prosthesis (2014)

youtu.be/ntrlHw6f4E4

Anyone’s Ninth (2023)



(Source: Sankei Shimbun)

sankei.com/article/20240113-CQCOSQHJWFIYPJJKZDCITRTRVI/


第4楽章

Course Promos

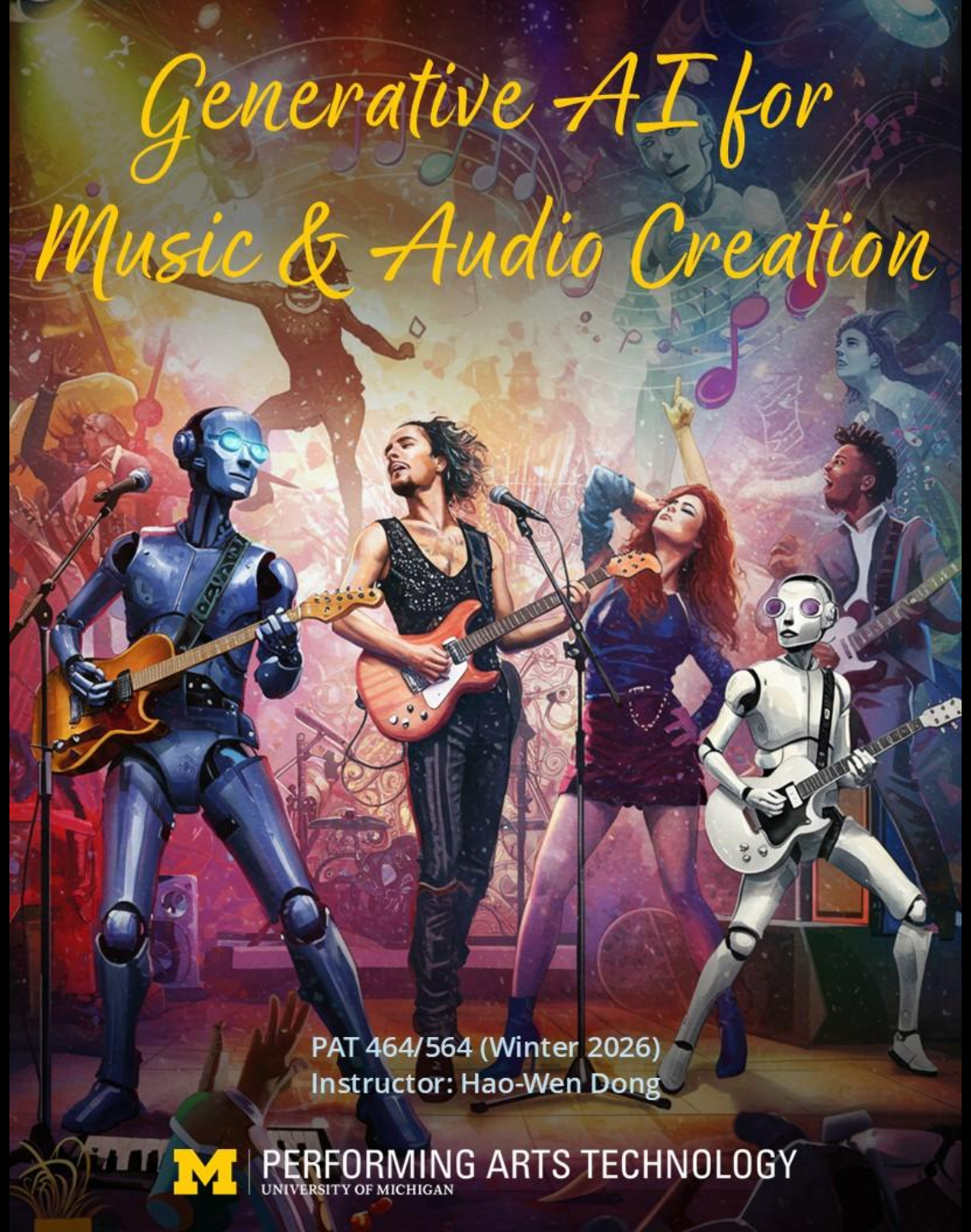
INTERMEDIA AI MUSIC PRACTICE

A NEW FALL 2025 COURSE OPEN TO ALL MAJORS
TAUGHT BY DR. JULIE ZHU IN PERFORMING ARTS TECHNOLOGY (PAT)
MON WED 8:30-10AM @BURTON MEMORIAL TOWER 506
GET EXPOSED TO THE NEWEST (AND CONSTANTLY CHANGING) LANDSCAPE OF MUSIC AI TOOLS
FROM ENTIRE-SONG TEXT-PROMPTS TO A BEAT TRACK PLUG-IN, FROM A LYRIC HELPER TO VOICE-TRANSFER SOFTWARE
THE CLASS WILL PRIORITIZE ARTISTIC PROCESS AND PROJECTS
WHERE AI IS A TOOL AND NOT A FAST-TRACK REPLACEMENT FOR CREATION

Fall 2025
Intermedia AI Music Practice
(PAT 205)

Prioritize artistic process and projects

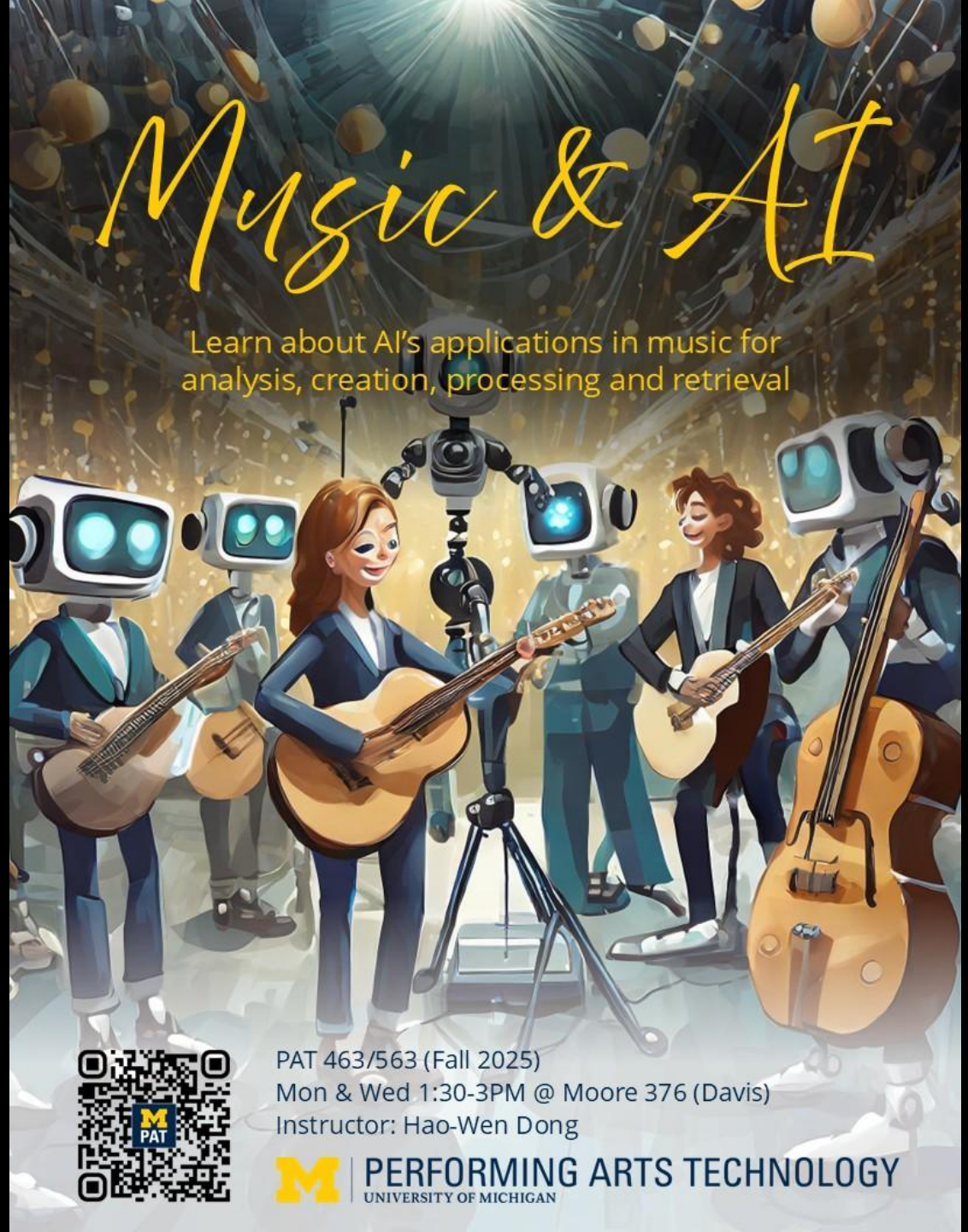
where **AI is a tool** and
NOT a fast-track replacement for creation



Winter 2026
Generative AI for Music and Audio Creation
(PAT 464/564)

Dive deeper into GenAI for music & audio

If you want to learn more about all the latest music and audio generation models



Fall 2025
Music and AI
(PAT 463/563)

Help spread the word!

If you've enjoyed it 😄

Final Thoughts

Music & Technology

Hildegard Dodel, Public domain, via Wikimedia Commons.
Taken at Hamamatsu Museum of Musical Instruments, August 2019.
yan, CC BY-SA 4.0, via Wikimedia Commons.

Use Cases of AI for Music & Audio

(Source: The Denver Post)

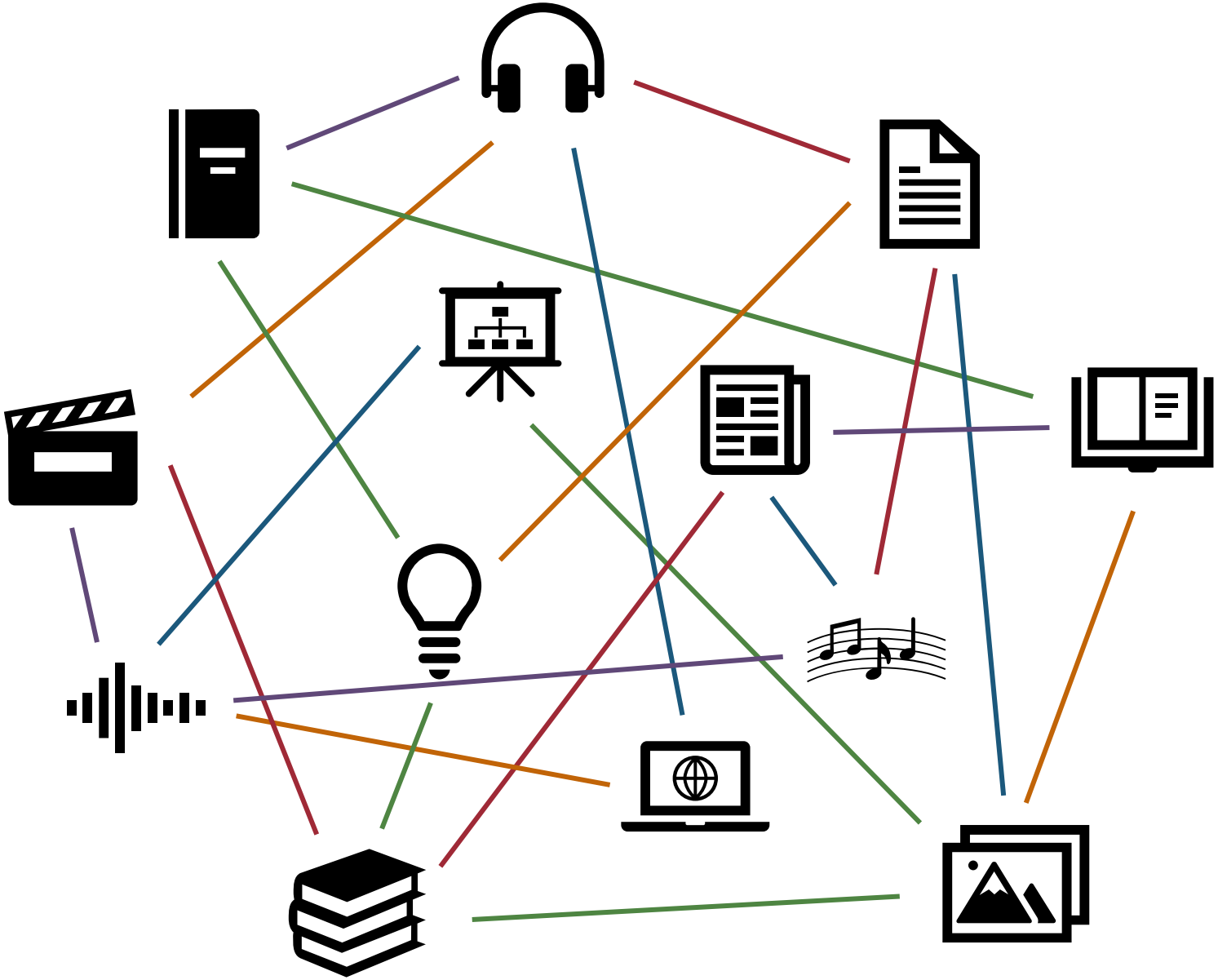
(Source: UploadVR)

(Source: Descript)

(Source: Daily Bruin)

(Source: Wikimedia Commons)

Universitaetsmedizin, CC BY-SA 4.0, via Wikimedia Commons
uploadvr.com/iron-man-vr-breaks-free-from-cords-load-screens-on-quest-2
descript.com/blog/article/what-is-the-best-audio-interface-for-recording-a-podcast
denverpost.com/2019/08/02/colorado-symphony-movie-scores-harry-potter-star-wars
dailybruin.com/2023/08/04/theater-review-the-musical-les-misrables-offers-stellar-displays-and-impassioned-vocals



Teachers

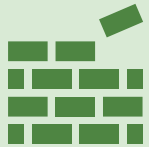
organize knowledge

Researchers

create knowledge

Engineers

apply knowledge



Artists

challenge knowledge