

PAT 498/598 (Fall 2024)

Special Topics: Generative AI for Music and Audio Creation

Lecture 1: Introduction

Instructor: Hao-Wen Dong



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

Welcome! Tell Us about Yourself!

- Name
- Pronouns
- Major/year
- Have you ever used any GenAI tools? Which tools?

About Me

- Hao-Wen (**Herman**) Dong
- Pronouns: he/him
- Email: **hwdong@umich.edu**
- Office: **Stearns 131** (10–15 min walk to the north from Moore)
- Office hours: 3–4PM, Mondays & Wednesdays
- Research areas: Generative AI for music and audio creation



Learning Goals

- Gain an overview of the field of **AI music creation**
- Learn fundamental AI concepts and principles
- Learn representative AI tools for music creation
- Gain hands-on experience on creating music using AI tools

Generative AI for Music and Audio Creation

Generative AI for **Music** and Audio **Creation**

Generative AI for Music and **Audio Creation**

What is this course all about?

An introduction to generative AI and its applications to music and audio creation. Topics include **music generation**, **audio synthesis** and **assistive music creation tools**.

Music & AI

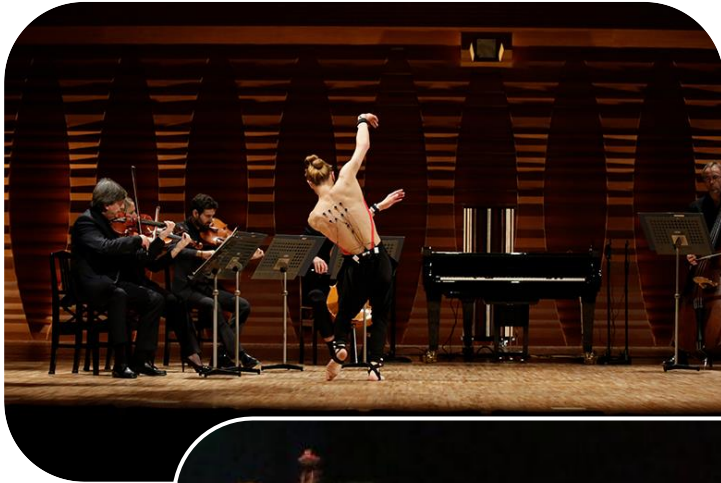
Music & Technology



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.

Music & AI

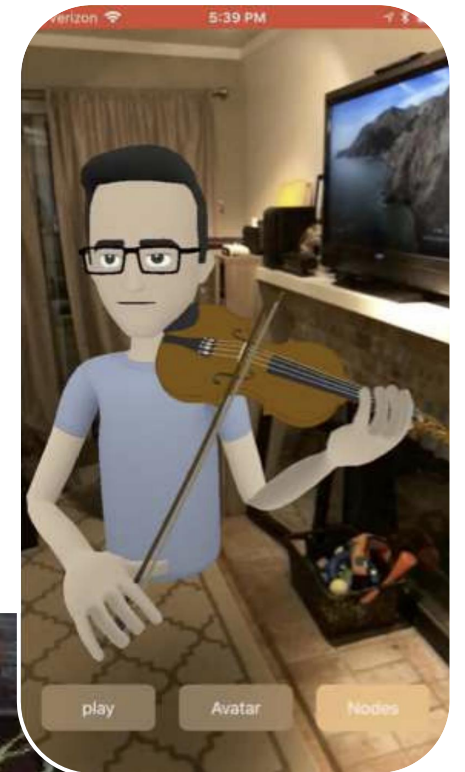
(Source: Yamaha)



(Source: Sankei Shimbun)



(Shlizerman et al., 2019)



(Source: Robot Gizmos)



(Source: NBC DFW)

Shlizerman et al., "Audio to Body Dynamics," *Proc. CVPR*, 2018.

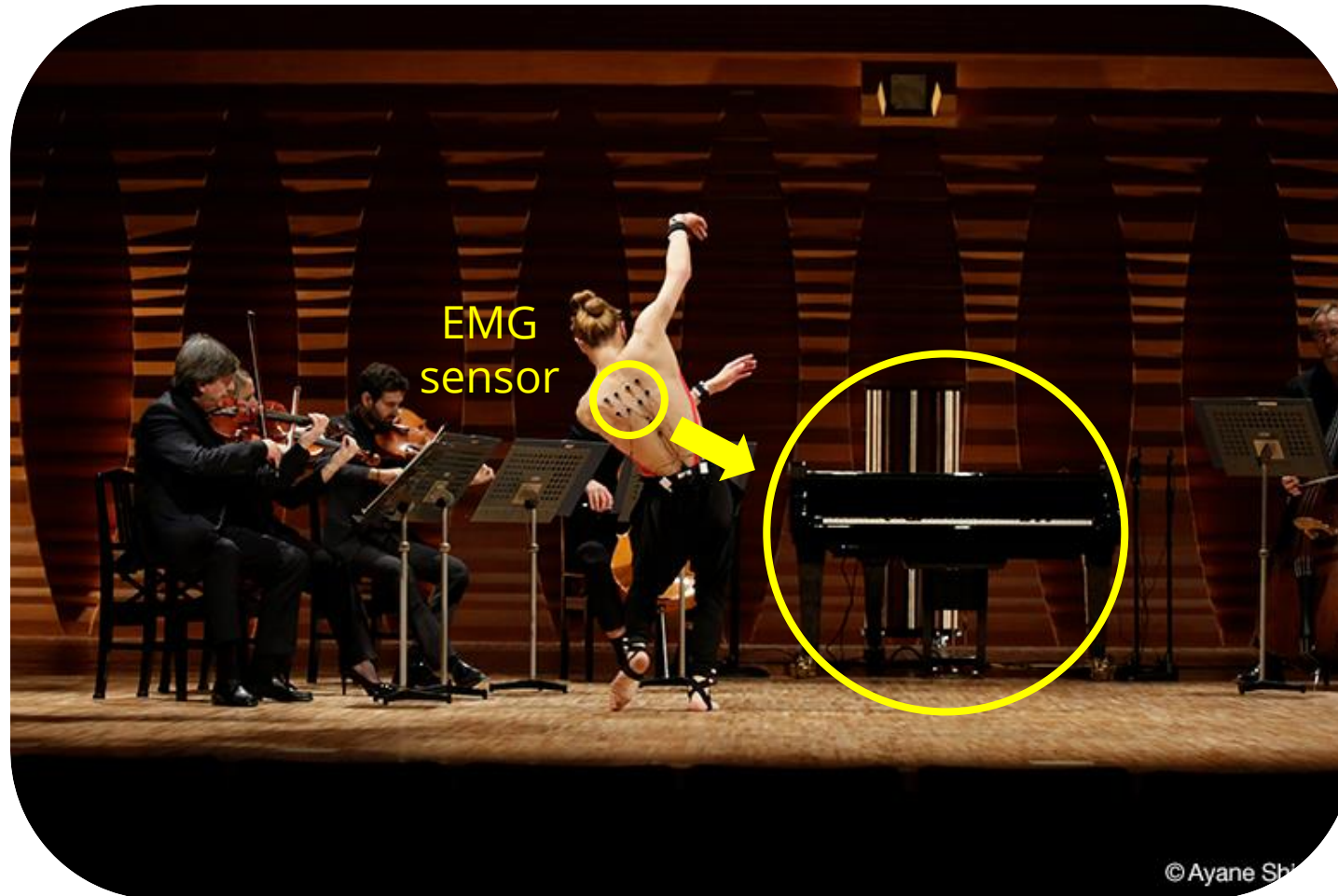
https://www.yamaha.com/en/news_release/2018/18013101/

<https://www.sankei.com/article/20240113-CQCOSQHJWFIYPJJKZDCITRTRVI/>

<https://www.roboticgizmos.com/shimon-musical-robot-deep-learning/>

<https://www.nbcdfw.com/entertainment/the-scene/how-verdigris-ensemble-is-using-ai-to-create-a-new-concert-experience/3366031/>

Music & AI



(Source: Yamaha)

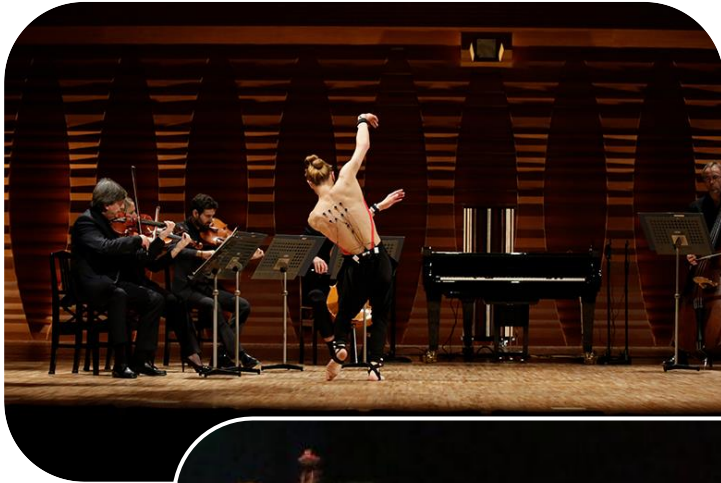
https://www.yamaha.com/en/news_release/2018/18013101/

Yamaha Global, "Yamaha Artificial Intelligence (AI) Transforms a Dancer into a Pianist - Short Version," *YouTube*, <https://youtu.be/21injmy1wsU>, 2018.



Music & AI

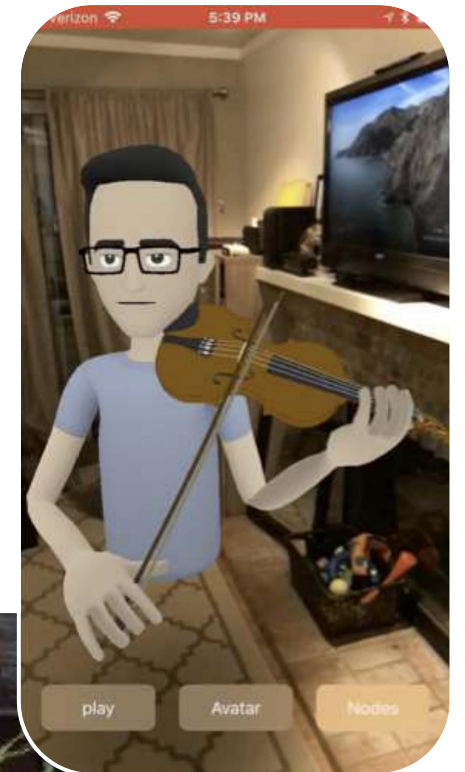
(Source: Yamaha)



(Source: Sankei Shimbun)



(Shlizerman et al., 2019)



(Source: Robot Gizmos)



(Source: NBC DFW)

Shlizerman et al., "Audio to Body Dynamics," *Proc. CVPR*, 2018.

https://www.yamaha.com/en/news_release/2018/18013101/

<https://www.sankei.com/article/20240113-CQCOSQHJWFIYPJJKZDCITRTRVI/>

<https://www.roboticgizmos.com/shimon-musical-robot-deep-learning/>

<https://www.nbcdfw.com/entertainment/the-scene/how-verdigris-ensemble-is-using-ai-to-create-a-new-concert-experience/3366031/>

Music & AI



(Source: Robot Gizmos)

<https://www.robotgizmos.com/shimon-musical-robot-deep-learning/>
Georgia Tech Center for Music Technology, "Shimon with the Aarhus Jazz Orchestra," *YouTube*, <https://youtu.be/ZpTV1-acSU8>, 2021.



Music & AI

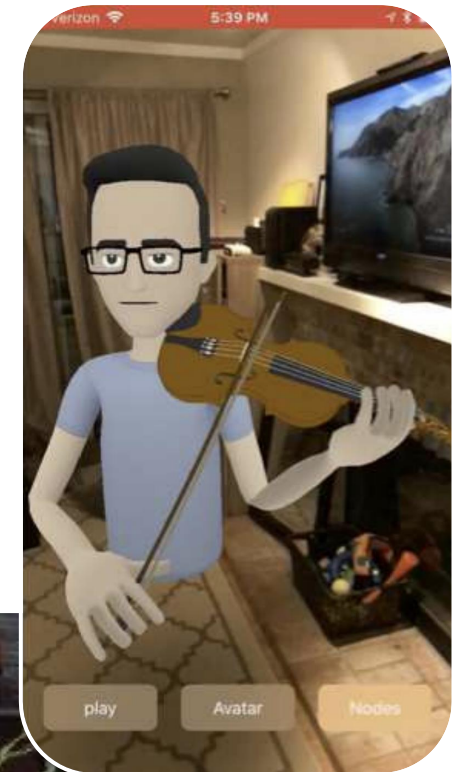
(Source: Yamaha)



(Source: Sankei Shimbun)



(Shlizerman et al., 2019)



(Source: Robot Gizmos)



(Source: NBC DFW)

Shlizerman et al., "Audio to Body Dynamics," *Proc. CVPR*, 2018.

https://www.yamaha.com/en/news_release/2018/18013101/

<https://www.sankei.com/article/20240113-CQCOSQHJWFIYPJJKZDCITRTRVI/>

<https://www.roboticgizmos.com/shimon-musical-robot-deep-learning/>

<https://www.nbcdfw.com/entertainment/the-scene/how-verdigris-ensemble-is-using-ai-to-create-a-new-concert-experience/3366031/>

Music & AI



Shlizerman et al., "Audio to Body Dynamics," *Proc. CVPR*, 2018.

https://www.yamaha.com/en/news_release/2018/18013101/

<https://www.sankei.com/article/20240113-CQCOSQHJWFIYPJJKZDCITRTRVI/>

(Source: Sankei Shimbun)

<https://www.roboticgizmos.com/shimon-musical-robot-deep-learning/>

<https://www.nbcdfw.com/entertainment/the-scene/how-verdigris-ensemble-is-using-ai-to-create-a-new-concert-experience/3366031/>

第4楽章



Music & AI

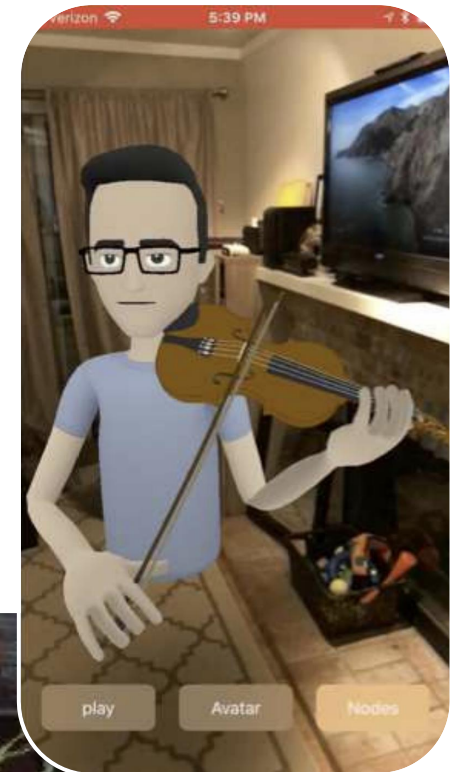
(Source: Yamaha)



(Source: Sankei Shimbun)



(Shlizerman et al., 2019)



(Source: Robot Gizmos)



(Source: NBC DFW)

Shlizerman et al., "Audio to Body Dynamics," *Proc. CVPR*, 2018.

https://www.yamaha.com/en/news_release/2018/18013101/

<https://www.sankei.com/article/20240113-CQCOSQHJWFIYPJJKZDCITRTRVI/>

<https://www.roboticgizmos.com/shimon-musical-robot-deep-learning/>

<https://www.nbcdfw.com/entertainment/the-scene/how-verdigris-ensemble-is-using-ai-to-create-a-new-concert-experience/3366031/>

Music & AI



(Source: NBC DFW)

Music & AI

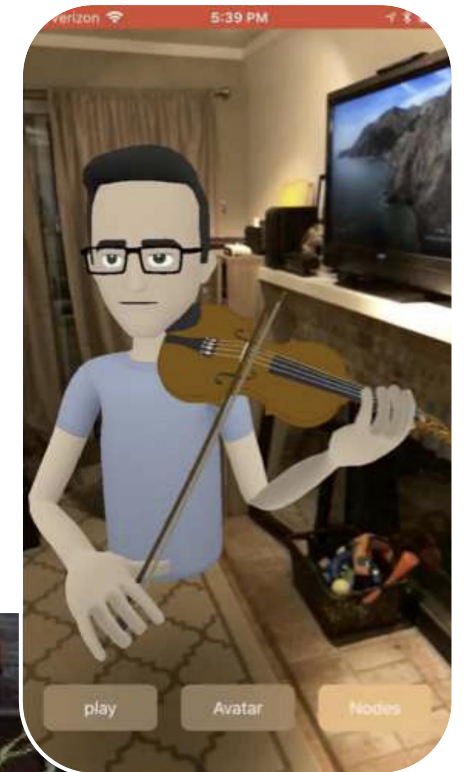
(Source: Yamaha)



(Source: Sankei Shimbun)



(Shlizerman et al., 2019)



(Source: Robot Gizmos)



(Source: NBC DFW)

Shlizerman et al., "Audio to Body Dynamics," *Proc. CVPR*, 2018.

https://www.yamaha.com/en/news_release/2018/18013101/

<https://www.sankei.com/article/20240113-CQCOSQHJWFIYPJJKZDCITRTRVI/>

<https://www.roboticgizmos.com/shimon-musical-robot-deep-learning/>

<https://www.nbcdfw.com/entertainment/the-scene/how-verdigris-ensemble-is-using-ai-to-create-a-new-concert-experience/3366031/>

Music & AI



(Shlizerman et al., 2019)



Course Logistics

Course Website

- **Main website:** hermandong.com/teaching/pat498_598_fall2024
 - Syllabus, schedule, lecture slides, code examples, etc.
- **Piazza:** Announcements, Q&A
- **Gradescope:** Assignment submission, grading, regrade requests
- **Canvas:** Recordings



Prerequisites

- **Prior coding experience** is recommended!
 - If you've taken any programming course, you should be fine
 - You should be comfortable reading code written by others
 - You should be able to write a nontrivial project using some programming language

Tentative Schedule

Generative AI Background

Week	Date	Lecture
1	Aug 26	Introduction
Background		
	Aug 28	┆ AI & machine learning fundamentals
2	Sep 2	┆ No Class (Labor Day)
	Sep 4	┆ Deep learning fundamentals I
3	Sep 9	┆ Deep learning fundamentals II
	Sep 11	┆ Language models - RNNs, LSTMs & transformers
4	Sep 16	┆ Generative adversarial nets & diffusion models
	Sep 18	┆ Music & audio processing fundamentals

Symbolic Music Generation

Week	Date	Lecture
Symbolic Music Generation		
5	Sep 23	┆ Melody generation
	Sep 25	┆ Harmony & chord progression generation
6	Sep 30	┆ Polyphonic music generation
	Oct 2	┆ Multitrack music generation
7	Oct 7	┆ Multimodal music generation I
	Oct 9	┆ Multimodal music generation II
8	Oct 14	No Class (Fall Study Break)

Audio Synthesis

Week	Date	Lecture
Audio Synthesis		
	Oct 16	┆ Time-domain audio synthesis I
9	Oct 21	┆ Time-domain audio synthesis II
	Oct 23	┆ Frequency-domain audio synthesis I
10	Oct 28	┆ Frequency-domain audio synthesis II
	Oct 30	┆ Multimodal audio synthesis I
11	Nov 4	┆ Multimodal audio synthesis II
	Nov 6	Project pitch & discussion
12	Nov 11	No Class (Travel)
	Nov 13	No Class (Travel)

Assistive Music Creation Tools

Week	Date	Lecture
Assistive Music Creation Tools		
13	Nov 18	┆ Neural audio effects
	Nov 20	┆ Auto-mixing
14	Nov 25	┆ Live performance & interactive systems
	Nov 27	No Class (Thanksgiving)
15	Dec 2	Discussions — ethical concerns & copyright issues
	Dec 4	Review
16	Dec 9	Project presentation



Assignments

- Programming exercises to get you familiar with music AI tools
- Due at **11:59pm ET** on the date specified
- Late submissions: **3 points deducted per day**

Project

- An **open-ended group project** (group size: 2–3)
 - A major component of this course
 - Focused on **creative & artistic use of AI tools**
- Proposal is due on **November 8** (*tentative*)
- Final presentation is scheduled on **December 9**
 - Let me know as soon as possible if you can't make it
- Final report is due on **December 15** (*tentative*)
- Late submissions: **Not Accepted**
 - You can always submit your work early and update it later

Grading

Assignments	40%	Project	60%
┆ Assignment 1	10%	┆ Proposal	10%
┆ Assignment 2	10%	┆ Final report	20%
┆ Assignment 3	10%	┆ Presentation	30%
┆ Assignment 4	10%		

Optional Reading

- *“AI Song Contest: Human-AI Co-Creation in Songwriting”* by Cheng-Zhi Anna Huang, Hendrik Vincent Koops, Ed Newton-Rex, Monica Dinculescu, and Carrie J Cai ([paper](#))
- Course slides for *“Deep Learning for Music Analysis and Generation”* by Yi-Hsuan Yang ([slides](#))
- *“Intelligent Music Production”* by Brecht De Man, Ryan Stables, and Joshua D. Reiss ([Amazon](#))
- *“Fundamentals of Music Processing”* by Meinard Müller ([notebooks](#)) ([Amazon](#))

Academic Integrity

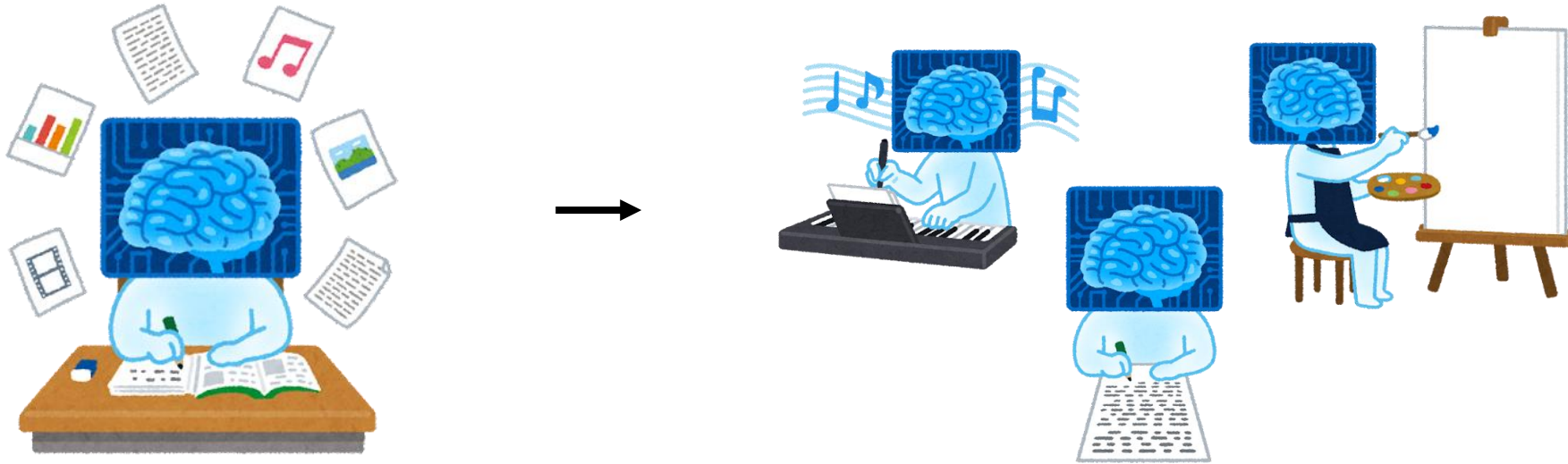
- Plagiarism and cheating violate SMTD's Academic Code of Conduct. **All plagiarism, cheating and other academic misconduct cases will be reported to SMTD's Office of Academic and Student Affairs.**
- **All assignments must be completed on your own.** You are welcome to exchange ideas with your peers, but this should be in the form of concepts and discussion, not in the form of writing and code.
- You must **provide proper citations/references for any external resources** you use in your writing and code.

| Any Questions on the Syllabus?



Generative AI

What is Generative AI?



Generative AI is AI capable of generating **text, images, music** or other **media**.

Analytic AI vs Generative AI



Generative AI for Visual Arts



(Source: Cosmopolitan)

First Prize in Digital arts
at Colorado State Fair
Fine Arts Competition

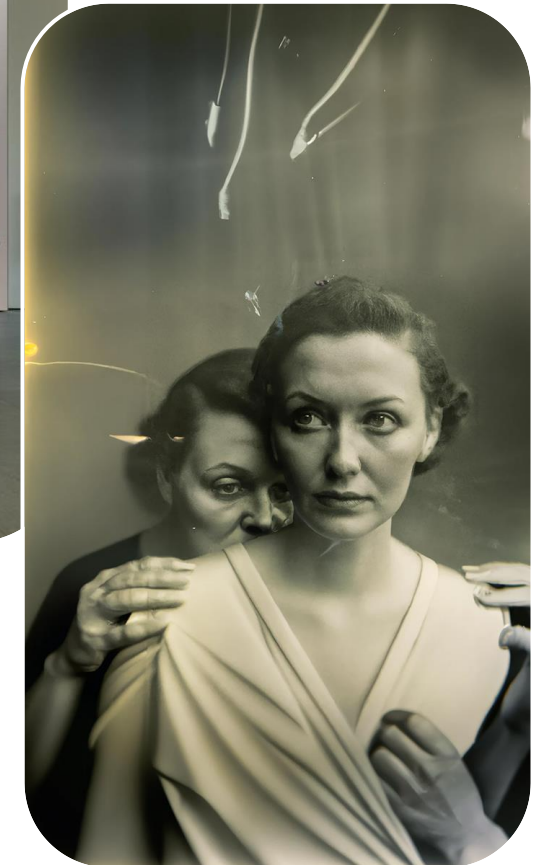


(Source: CNN Business)

(Source: MoMA Magazine)



Sony World
Photography Award in
Creative Open Category



(Source: CNN)

Gloria Liu, "[The World's Smartest Artificial Intelligence Just Made Its First Magazine Cover](#)," *Cosmopolitan*, June 21, 2022.
Rachel Metz, "[AI won an art contest, and artists are furious](#)," *CNN Business*, September 3, 2022.
Refik Anadol, "[Refik Anadol on AI, Algorithms, and the Machine as Witness](#)," *MoMA Magazine*, December 20, 2022.
Lianne Kolirin, "[Artist rejects photo prize after AI-generated image wins award](#)," *CNN*, April 18, 2023.

State of the Art – Text-to-Music Synthesis

Prompt: relaxing and smooth jazz played in a stylish cafe



Prompt: delightful country music with acoustic guitars



Prompt: cinematic and suspenseful orchestral music



huggingface.co/spaces/facebook/MusicGen



State of the Art – Video Generation



State of the Art – Music + Video Generation



Video **Runway Gen-2**
Music **MusicGen**



Use Cases of Generative AI for Music & Audio

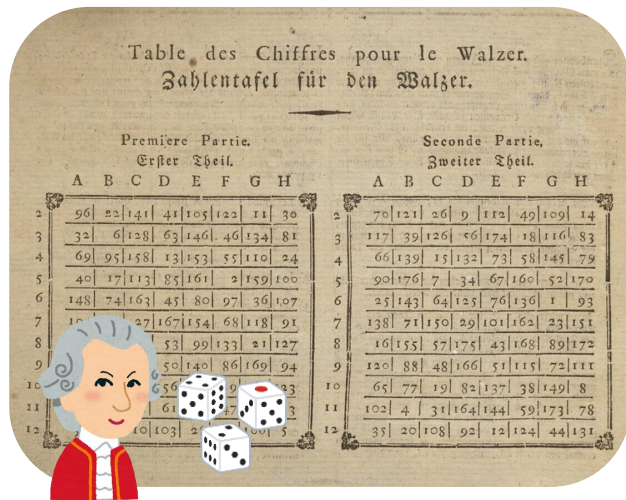


Universitaetsmedizin, [CC BY-SA 4.0](https://www.wikimedia.org/licenses/by-sa/4.0/), via Wikimedia Commons
<https://www.uploadvr.com/iron-man-vr-breaks-free-from-cords-load-screens-on-quest-2/>
<https://www.descript.com/blog/article/what-is-the-best-audio-interface-for-recording-a-podcast>
<https://www.denverpost.com/2019/08/02/colorado-symphony-movie-scores-har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

AI Music in the Early Days

The Early Days

Musical Dice Game (1792)



(Source: gbrachetta)

gbrachetta.github.io/Musical-Dice/

ILLIAC Suite (1957)



(Source: Illinois Distributed Museum)

Emily Howell (2003)



(Source: The Guardian)

<https://gbrachetta.github.io/Musical-Dice/>
<https://distributedmuseum.illinois.edu/exhibit/illiac-suite/>
<https://www.theguardian.com/technology/2010/jul/11/david-cope-computer-composer>

Musical Dice Game (1792)


Table des Chiffres pour le Walzer.
Zahlentafel für den Walzer.

Premiere Partie.
Erster Theil.

	A	B	C	D	E	F	G	H
2	96	22	141	41	105	122	11	30
3	32	6	128	63	146	46	134	81
4	69	95	158	13	153	55	110	24
5	40	17	113	85	161	2	159	100
6	48	74	163	45	80	97	36	107
7	24	157	27	167	154	68	118	91
8	2	60	171	53	99	133	21	127
9	9	84	114	50	140	86	169	94
10	98	142	42	156	75	129	62	123
11	3	87	165	61	135	47	147	33
12	9	9	9	93	28	37	106	5

Seconde Partie.
Zweiter Theil.

	A	B	C	D	E	F	G	H
2	70	121	26	9	112	49	109	14
3	117	39	126	56	174	18	116	83
4	66	139	15	132	73	58	145	79
5	90	176	7	34	67	160	52	170
6	25	143	64	125	76	136	1	93
7	138	71	150	29	101	162	23	151
8	16	155	57	175	43	168	89	172
9	120	88	48	166	51	115	72	111
10	65	77	19	82	137	38	149	8
11	102	4	31	164	144	59	173	78
12	35	20	108	92	12	124	44	131



(Source: gbrachetta)

gbrachetta.github.io/Musical-Dice/

Let's try it out!



Lejaren Hiller – ILLIAC Suite for String Quartet (1957)

ILLIAC SUITE FOR STRING QUARTET

3

I. EXPERIMENT NO. 1

L.A. HILLER, JR. AND L.M. ISAACSON

PRESTO

VIOLIN I

VIOLIN II

VIOLA

CELLO

(A)



youtu.be/n0njBFLQSk8 &
music.arts.uci.edu/abauer/3.1/scores/Hiller_Illiac_Suite.pdf

Emily Howell (by David Cope) – Prelude (2007)



Reading: David Cope on Emily Howell

- [David Cope: 'You pushed the button and out came hundreds and thousands of sonatas'](#)
 - Interview by Tim Adams published on *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.**"*



What is Artificial Intelligence?

What is Artificial Intelligence?

AI is the study of how to make computers **do things at which, at the moment, people are better.**

– Elaine Rich and Kevin Knight, 1991

1997



(Source: Britannica)

2016



(Source: The Guardian)

20??



(Source: SC2HL)

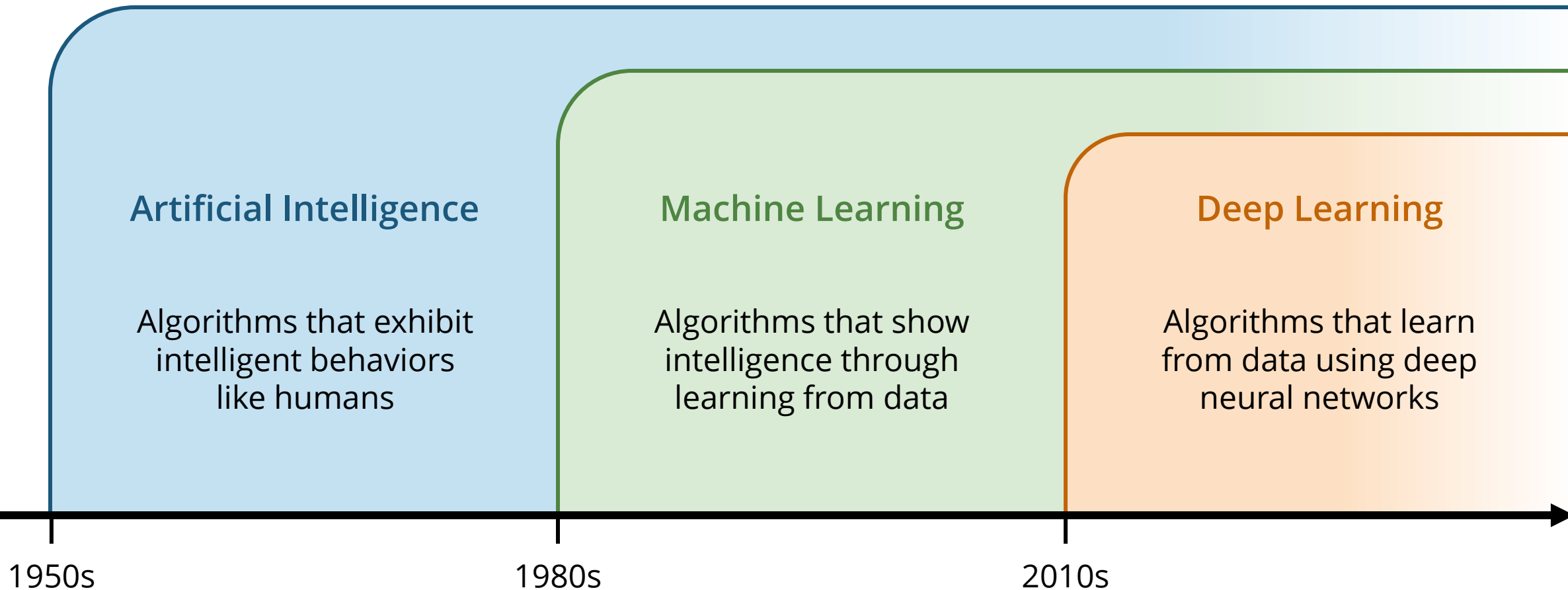
Elaine Rich and Kevin Knight, *Artificial Intelligence*. United Kingdom: McGraw-Hill, 1991.

<https://www.britannica.com/topic/Deep-Blue>

<https://www.theguardian.com/technology/2016/mar/15/alphago-what-does-google-advanced-software-go-next>

https://www.youtube.com/watch?v=PFMRDm_H9Sg

AI vs ML vs DL



Building Blocks of Modern AI Systems



Data

×



Model

×



Use Case

Key difference from
traditional AI