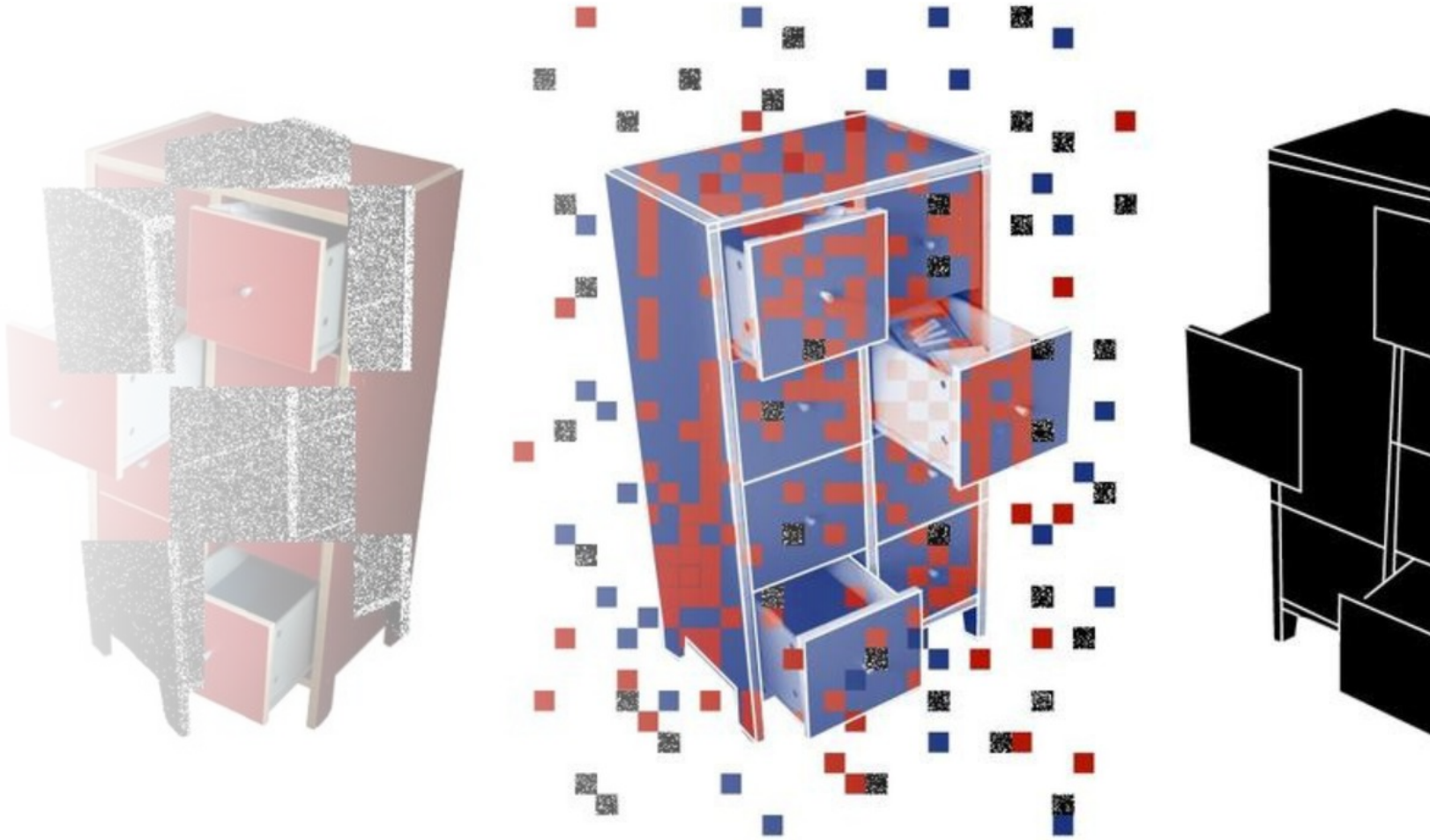
Artificial Intelligence

Week 1

Daniella Gáti



Plan for today

- Introduction to the Module
  - Including aims, structure, assessment brief
- What is (Artificial) Intelligence?
- An idea-centered history of AI
- Clarifying some essential concepts

Intro to the Module

Module aims

- To understand how AI technology works.
- To understand AI's potential, as well as its limitations, risks and ethical issues surrounding it.
- To be able to conceptualize and design an AI-driven solution to a real-world problem.
- To gain the technical skills necessary to integrate AI technology into creative digital artifacts.

	Assessed intended learning outcomes On successful completion of this assessment, you will be able to:
Knowledge and Understanding	1. Evidence an understanding of AI technology and its potential application. 2. Articulate the limitations, risks and ethical issues surrounding AI technology. 3. Demonstrate the knowledge required to implement AI technologies within a practical artifact.
Practical, Professional or Subject Specific Skills	4. Explore a real-world problem and respond with an AI driven solution. 5. Develop the capability to design and prototype an AI driven artifact. 6. Professionally document the process of developing an AI driven artifact. 7. Demonstrate an ability to generate ideas, find information and initiate activity independently.

The major aim of the assessment has two parts:

- To be able to create a digital artifact that incorporates (elements) of existing AI source code or libraries
- To be able to design AI-driven solutions that use AI productively while being critical of shortcomings or dangers

The deliverables are:

- A digital artifact that responds to a social problem (60%)
- A Professional Specification (40%)

But what kind of artifact?? And what problem?

Artifact: presumably something in Processing, p5.js, or Python, like an artwork or a visualization, or some other code that processes data.

Problem: It could be something like:

- How to help people decide in which bin to throw their trash?
- How to capture my hand for visualization in my Processing art piece?
- How to detect smiles on people's faces and change visualization based on that.

Don't worry, we'll spend plenty of time finding a problem and working on how to come up with an artifact!

And remember, you're not required to save the world: a prototype or demonstration will suffice!

Breaking down the brie

The practical project requires students to respond to a real-world problem with an AI driven solution, where you design, build, and test a tangible artifact that responds appropriately to the problem.

Technically, the project will:

- Develop an own, autonomous digital artifact
- Make use of student's own source code

25

- Make use of existing AI / machine learning codes and libraries, OR develop the student's own. It is <u>not</u> required to develop new AI applications; integrating existing ones into your work is sufficient.

Conceptually, the project will:

- Identify a real-world problem of some significance. This problem and its significance must be clearly identified and named (in the Professional Specification). The problem should be drawn from the topics considered in class, which could be creative (making art, playing a game), practical (accomplishing a task), environmental (helping people do something more sustainably), or it could be a problem of personal interest to you (for example, a program pretending to be human). You will receive ample guidance in finding a problem. Example problems include: how to capture the motion or shape of an object for representing it digitally in an artwork; how to draw humans expressing different emotions; how to detect the type of waste product.
- Design a (partial!) solution or response to the problem. You don't have to save the world, but your solution should be innovative (that is, not something that already exists) and it should make a contribution to solving or understanding the problem better. If the problem is highly complex, your artifact can be a prototype for how one *might* solve the problem. The main aim of your solution is to demonstrate your understanding of AI technologies.
- Consider the social, cultural, and ethical risks/limitations surrounding the use of AI to tackle your problem. These considerations should flow into the design.
- Acknowledge existing attempts/approaches to solving the problem, and other sources of inspiration, as well as codes/libraries/tools used.

The Brief, in plain English

- You could:
  - Create art that uses AI creatively and socially responsibly!
  - Create a tool that uses AI to help people manage waste
  - Create a tool that uses AI to help people manage time
  - Create a game that uses AI to animate the computer’s character or decide its moves
- Whatever you do, you must have a clear good justification for using AI and include social responsibility considerations into your design.

Examples from the real world

- AI adjusting stage lighting based on cues
- AI artists collective working specifically with, and on AI
- Google’s AutoDraw
- Color design by Khroma
- A play incorporating, and about, AI at the Young Vic
- Opening your phone with Face ID
- Digital voice assistants (language recognition)/ smart home devices
- And also, of course, work by our own students!

Assessment Deadline

Submission/Assessment Date	The submission deadline is **Tuesday 28th April 2026** by **no later than 16:00**. Any submission received after 16:00 (even if only by a few seconds) will be considered as late. **This is central university policy, and there is nothing I can do about it.** Please make sure to begin your upload well in advance of the deadline.

Schedule of sessions (subject to change)

1. History of AI vs. contemporary AI
2. Machine learning theory
3. Intro to PyTorch and Building a model (CNN)
4. Finishing building it + Deployment
5. Building a model: DCGAN
6. Building a model: LLM (or SLM)
7. Building a model: pose recognition / speech recognition (TBD)
8. Revisit brief and ideation
9. Project support
10. Project support
11. Project Support
12. Project support

Any questions?

What is artificial intelligence?

- Extremely vague term
  - *Related to vagueness of “intelligence”*
- Shifting meaning across history
- I would like to differentiate between “the dream” of artificial intelligence, and contemporary technologies
- Next slides discuss these in turn (dream, intelligence, and history)
- Before we do that, though, can we name some examples? Past student works?

The “**dream**” of AI

- AGI: artificial general intelligence. Would be able to reason and think (and feel?) like a human
- The sci-fi robot dream (C3PO, etc.) (also a nightmare, see Terminator)
- “the singularity”, techno-utopianism
- Note: these are ideologies!

- Vs other goals and realities:
  - *Autonomous task execution: not able to think/feel, but able to outperform humans in tasks (e.g. safe driving)*
  - *Superpowered decision-making: using much more data/information than humans could ever use to make decisions*
  - *Seeing/recognizing/noticing: recognizing patterns humans cannot detect*

What makes something **intelligent**?

Process logic (or just process information—but what is information?) ?

Carry out tasks, such as finding things (text)?

Carry out tasks autonomously?

Make decisions? Make decisions “better” than humans?

“remember” –or be able to store its own actions, or data?

Feel emotions?

Understand? What does that even mean?

A brief **history** of "AI" (with shifting ideas)

- Turing, "On Computable Numbers, with an Application to the Entscheidungsproblem" (1936) – computers making decisions
  - *Turing test – the idea of appearing human-like*
- Dartmouth Summer Research Project on Artificial Intelligence (1956), incl. Claude Shannon, John McCarthy, Marvin Minsky, funded by US military funds – trying to code concepts and generalizations (logic)
- 1952: Christopher Strachey's Love Letters by MUC (creative computing) – plays with the idea of appearing human-like
- Frank Rosenblatt, perceptron: electronic device able to "sense" and "learn from" environment (1957) – sensing (and the idea of neural networks); abandoned in favor of logic à la Dartmouth
- Joseph Weizenbaum, ELIZA (1964-67) – attempt to refute the idea of appearing human-like
- John Searle, Chinese room experiment ("Minds, Brains, and Programs,"1980) – refuting the idea of appearing human-like
- 1980s: reigning paradigm is logic; "expert systems"
- 1980s/90s: moving away from language and logic, and towards probabilistic thinking = statistics: LeNet papers on digit recognition
- 1996: IBM's Deep Blue expert system defeats chess world champion Gary Kasparov; meanwhile, machine learning grows as a field
- 2015: AlphaGo, machine learning-based, defeats Go master Fan Hui.
- 2017: "Attention is all you need": transformer architecture of machine learning
- Nov 2022: ChatGPT launched – machine learning

Paradigms (or a summary of history)

- **Language** = heart of human thought
- **Logic** = formal mathematics as heart of intelligence
  - ***Deterministic* (GOFAI)**
- **Machine learning** = approximate human tasks using statistics
  - *stochastic*

- **The machine learning paradigm has won out**. It is superior in pretty much all applications, and it underlies all contemporary AI. But it is an approximation of human tasks, rather than something similar to them.
- But note: “good old fashioned algorithms” aren’t dead. They’re used a lot to correct for the errors of stochastic algorithms. **Always keep this in mind.**

TANGENT:

Understanding deterministic vs. stochastic algorithms, using ELIZA and MUC

ELIZA
Archaeology Project

Welcome to the ELIZA Archaeology Project. This research project brings together an interdisciplinary team of scholars, artists, and programmers to explore the history and ideas around the ELIZA chatbot.

ELIZA is the original and highly influential chatbot that launched the genre of human-computer interactions using text-based agents. It was created at MIT in the 1960s as part of Project MAC by its designer and programmer, Joseph Weizenbaum, the program not only allowed Wizenbaum to develop a mode of interaction with computers that is highly interactive, it also contributed to the way in which people were starting to conceptualize computers as having the capacity to usefully engage in conversation. You can try an accurate reimplementation of ELIZA developed as part of this project.

We plan to contextualize the program, offering its history and context as well as offering a detailed explanation of how the code works. This project will look at the culture of programming in which Weizenbaum was working and then explore his turn from ELIZA/DOCTOR, as he began to warn of the hazards of treating machines like humans. We will look at later works inspired by ELIZA and consider its influences on the way talking computer programs are represented in literature and film.

What was ELIZA?

A computer program that wrote text responses that were designed to look like a psychotherapist.

Written in 1966, Joseph Weizenbaum, MIT.

Purpose: highlighting the distinctions between humans and machines.

What happened?

People treated ELIZA as if ”she” was human.

They knew full well she wasn’t.

It didn’t matter.

This is known as **the ELIZA effect**: when people treat an object that *looks like* it exhibits superior intelligence as if it really possessed superior intelligence—even though they know it doesn’t.

Try the Eliza emulator:

- Go to: https://sites.google.com/view/elizaarchaeology/try-eliza?authuser=0
- Play around with the emulator for 5-10 minutes.
- Then we'll discuss:
  - How does this feel?
  - Does it work?
  - How does it work?

Here’s the program reimplemented in Processing

- https://github.com/codeanticode/eliza/blob/master/examples/ElizaChat/ElizaChat.pde

- You need to add the library Eliza first.

MUC Love Letters

Love Letters, 1952/3

Designed by Christoper Strachey in 1952/3 using the Manchester University Computer (MUC), a Ferranti Mark 1

First computer-written ”poems” or first ever computer art

Idea: demonstrate productive vs. thinking capabilities of computer (/joke)

Love Letters reimplementations

- Try out Nick Montfort's reimplementation of Love Letters:
- https://nickm.com/memslam/love_letters.html
- We'll do this together.
- (Then, for yourself, play around with this emulator:
- https://www.gingerbeardman.com/loveletter/)
- How does the algorithm work?

Contemporary AI algorithms

on the contrary...

- Do not process logic
- All rely on machine learning = statistics instead
- Ⓟ they are able to give “educated guesses” to any query

UNDERSTANDING CONTEMPORARY AI

Do we want to agree on what we think is intelligence?

- One candidate: Situated, embodied, affective, value-driven decision-making
- Not equal to information processing, calculating, or solving (even stochastic) equations
- How does this relate to machine learning (=AI)?

AI does not reflect, consider, weigh, analyze information or data, or come up with decisions based on values or priorities. It does one thing only: it creates an output whose probability is maximal given the training data.

Gathering hype words

MACHINE LEARNING

- Focus is not on intelligence, but on performance: the task is well executed
- Mitchell, paraphrased: A machine is said to learn with respect to a task if its performance at the task improves with experience.
- Break this down:
  - *Performance:* we get a result that aligns with what we want (not proof of intelligence, think of Searle)
  - *Task:* not general – though may look like it! E.g. LLMs (discuss!)
  - *Experience* = iterative encounter with data (this is called learning)

Recap on statistics / basic intuition on learning

- Let’s focus on learning first:
- Computer is given a clear task ☾ attempts to execute task ☾ is given some feedback from the data on its performance ☾ adjusts its task execution
- The key element is the feedback, which can arrive in different forms, but in all its forms it involves a **cost that the algorithm is trying to minimize**
- It adjusts its task execution so that the next time, its cost is smaller
- Ⓟ the quality of the algorithm’s performance depends on what criteria we give it in terms of the cost function
- Discuss examples: identify cats, produce language that looks human, weather forecast
- Clarify types:
  - *Prediction*
  - *Classification*
  - *Generation*
  - *Recommendation*

How does this differ from classic inferential statistics?

- Classic inferential statistics: we calculate the overall minimum of the cost function all at once, rather than doing it step by step
- So why does machine learning do it step by step?
  - *Because the data is so massive, computationally it's difficult/impossible*
  - *Because the cost function is so complex, and not always differentiable*
  - *Because there are literally millions of parameters, so computationally, this is almost necessity.*
- So it's really practical reasons, but otherwise it's the "same old" as statistics.
- On the surface, machine learning looks like human learning; but in fact it's just the same as solving some equations in classic statistics.
- Ⓟ it makes no sense whatsoever to think of these algorithms as intelligent

Last words for today

- All contemporary AI is machine learning
- Machine learning doesn’t really learn, it isn’t intelligence
- It is stochastic, not deterministic or logical (there is no reasoning there)
- It is fundamentally dependent on how we set up the task, the measure of performance, as well as the data

Next week

- We’ll discuss the theory of machine learning a bit more technically, and in detail
- We’ll review some essential statistics and math