

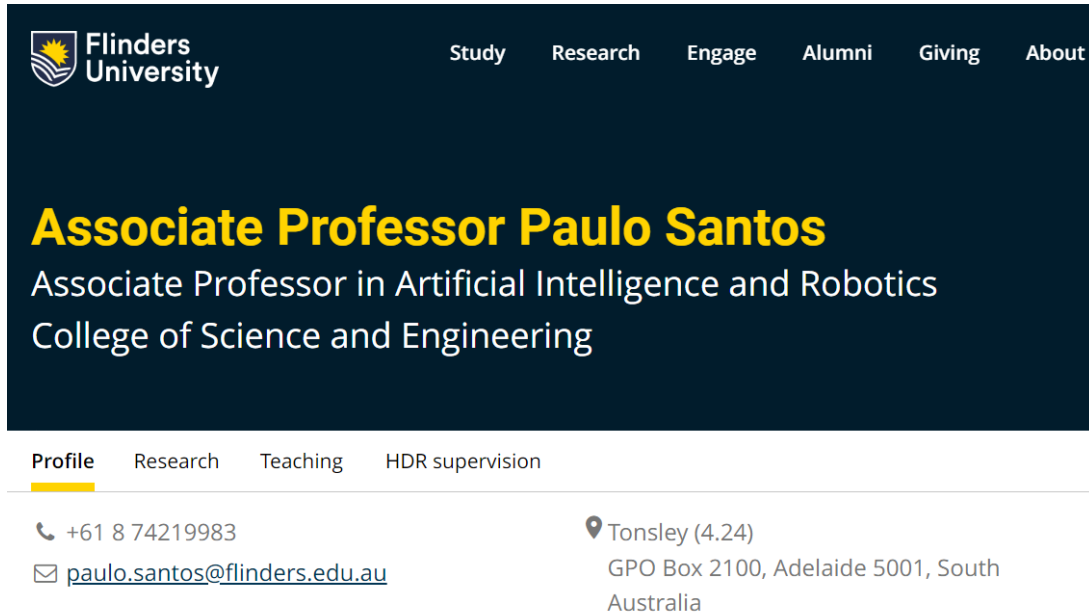


Ethical decision making under the AI Razor

Paulo E. Santos,
A/Prof of AI and Robotics,
College of Science and Engineering,
Flinders University



What I do.



The screenshot shows the Flinders University website profile for Associate Professor Paulo Santos. The header includes the Flinders University logo and navigation links: Study, Research, Engage, Alumni, Giving, and About. The main content area features the name 'Associate Professor Paulo Santos' in large yellow text, followed by his title 'Associate Professor in Artificial Intelligence and Robotics' and 'College of Science and Engineering' in white text. Below this is a secondary navigation bar with 'Profile', 'Research', 'Teaching', and 'HDR supervision', where 'Profile' is highlighted. Contact information includes a phone number (+61 8 74219983), an email address (paulo.santos@flinders.edu.au), and a location pin for Tonsley (4.24) with the address: GPO Box 2100, Adelaide 5001, South Australia.

Qualifications

PhD Imperial College, London

Master of Logic, Institute for Logic, Language and Computation, University of Amsterdam

MSc in Electrical and Electronics Engineering, University of São Paulo

Bachelor in Physics, Institute of Physics, University of São Paulo

- <https://www.flinders.edu.au/people/paulo.santos>

CNRS International Research Laboratory, Adelaide, Australia



About Us

IRL CROSSING is an International Research Laboratory

We are partnered with CNRS, University of Adelaide, IMT Atlantique, University of South Australia, Flinders University, Naval Group.

A French Australian Laboratory for Humans-Autonomous Agents Teaming

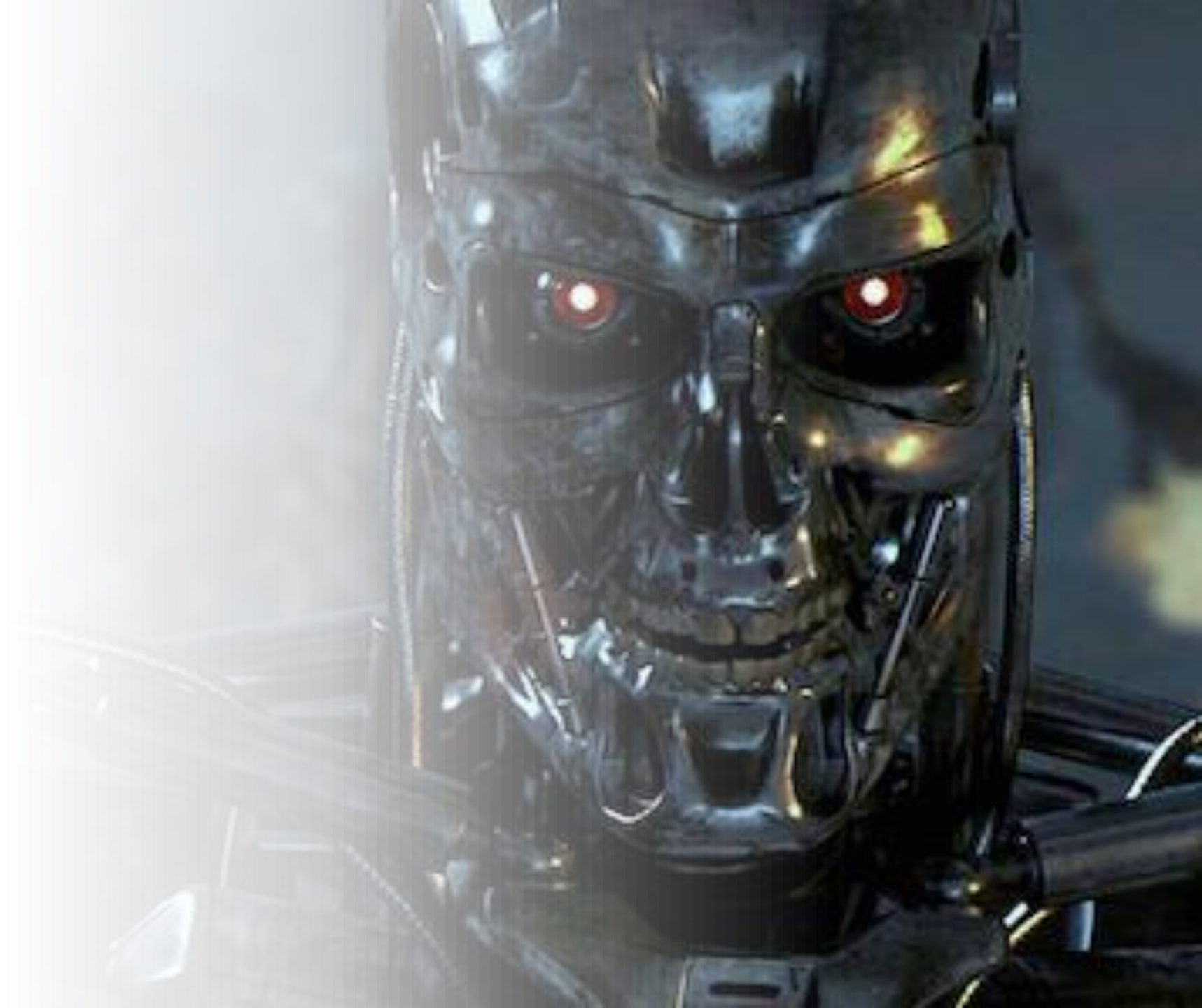




The evil AI...

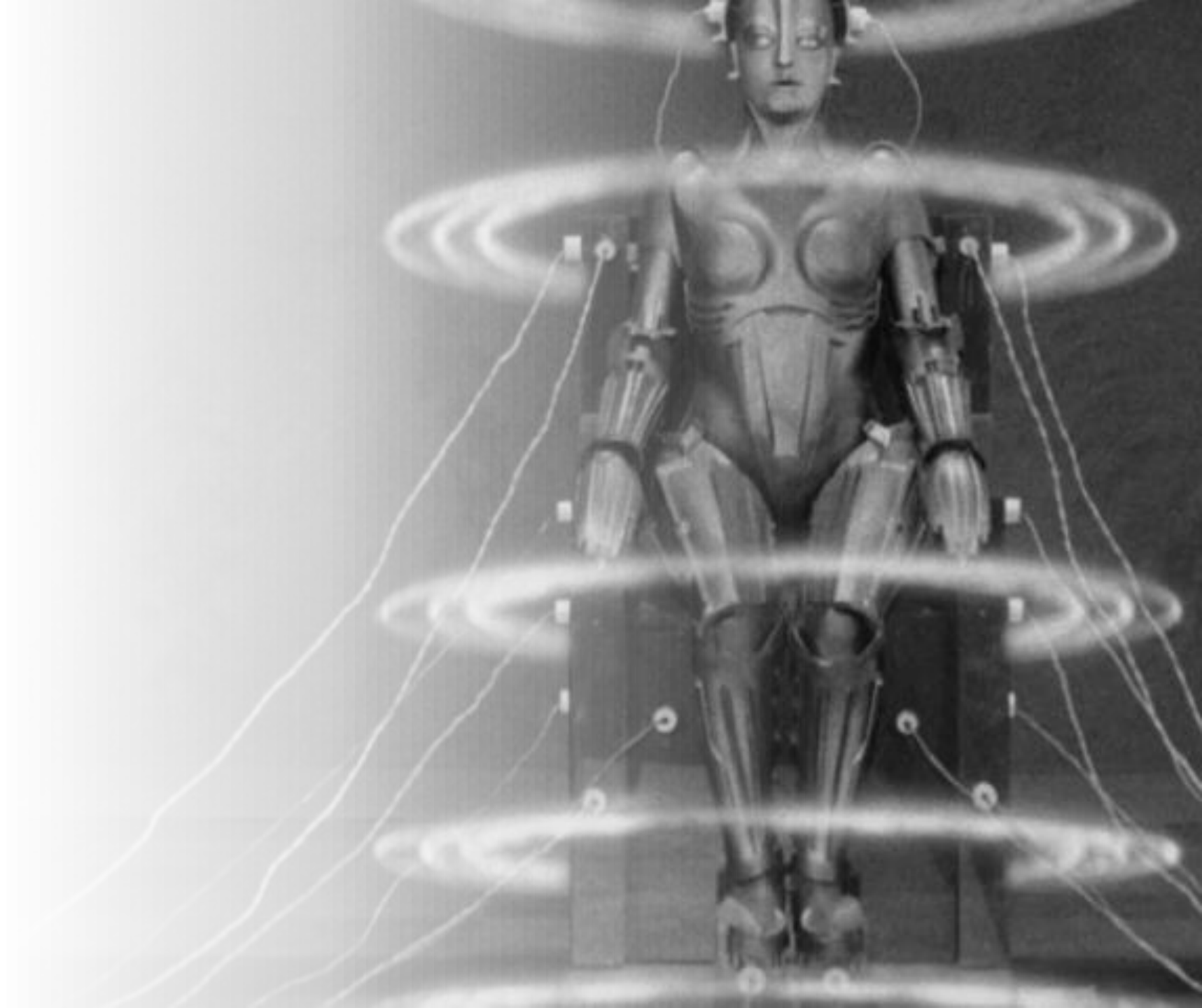
1984

- J. Cameron, *The Terminator*



1927

- Fritz Lang, Metropolis



ca. 300/280-270 BCE

'ΤΑΛΩΝ'. Silver didrachma from
Phaistos, Crete (ca. 300/280-270
BCE), obverse. (Cabinet des
Médailles, Paris)

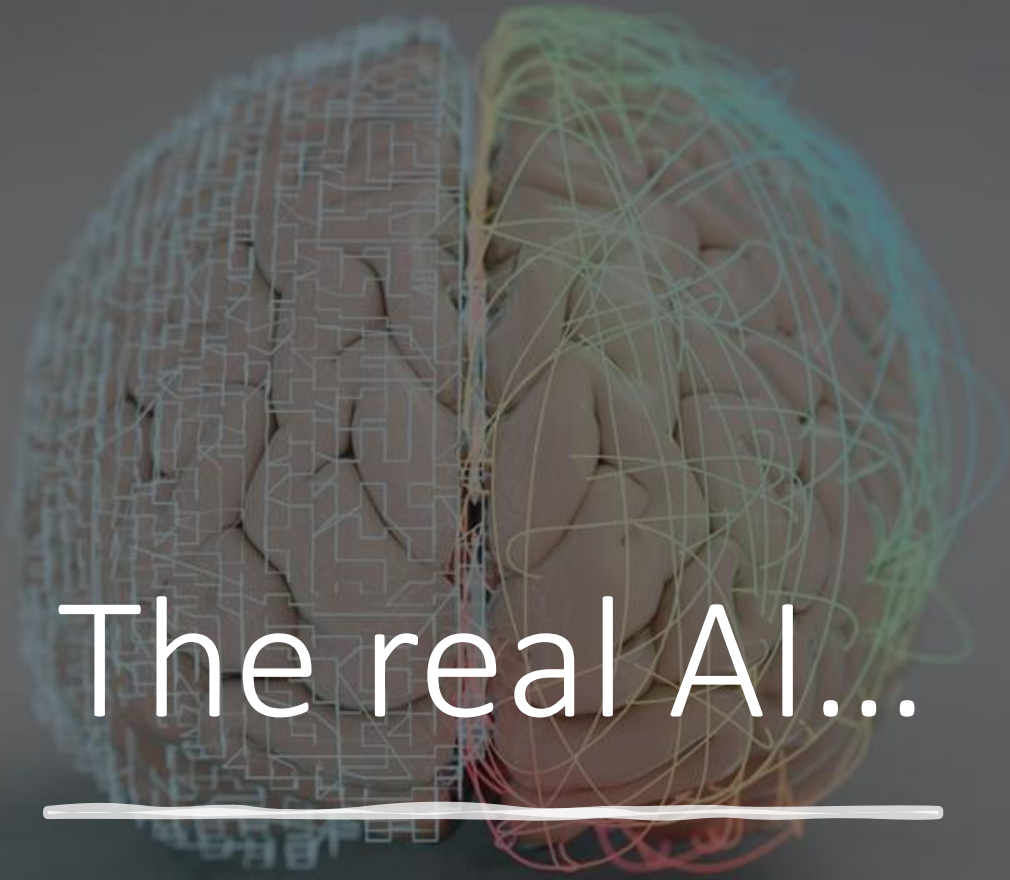




2023

- Chat.GPT





The real AI...

Jacquard Loom (1801)



[https://www.youtube.com/watch?v=OlJns3fPltEen.](https://www.youtube.com/watch?v=OlJns3fPltEen)

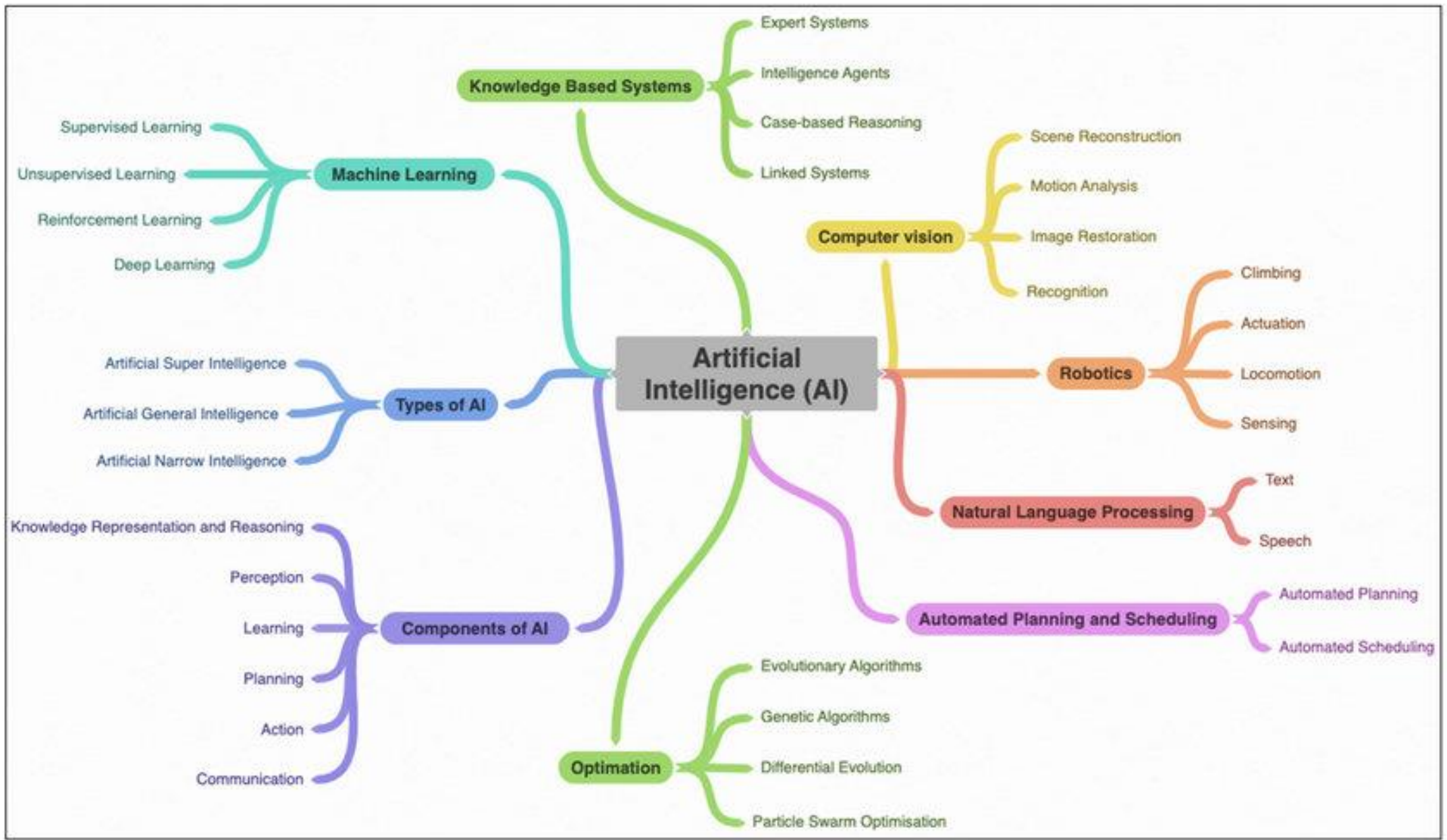
- Programmable,
- Reproduce patterns from observed patterns (Punched Cards)

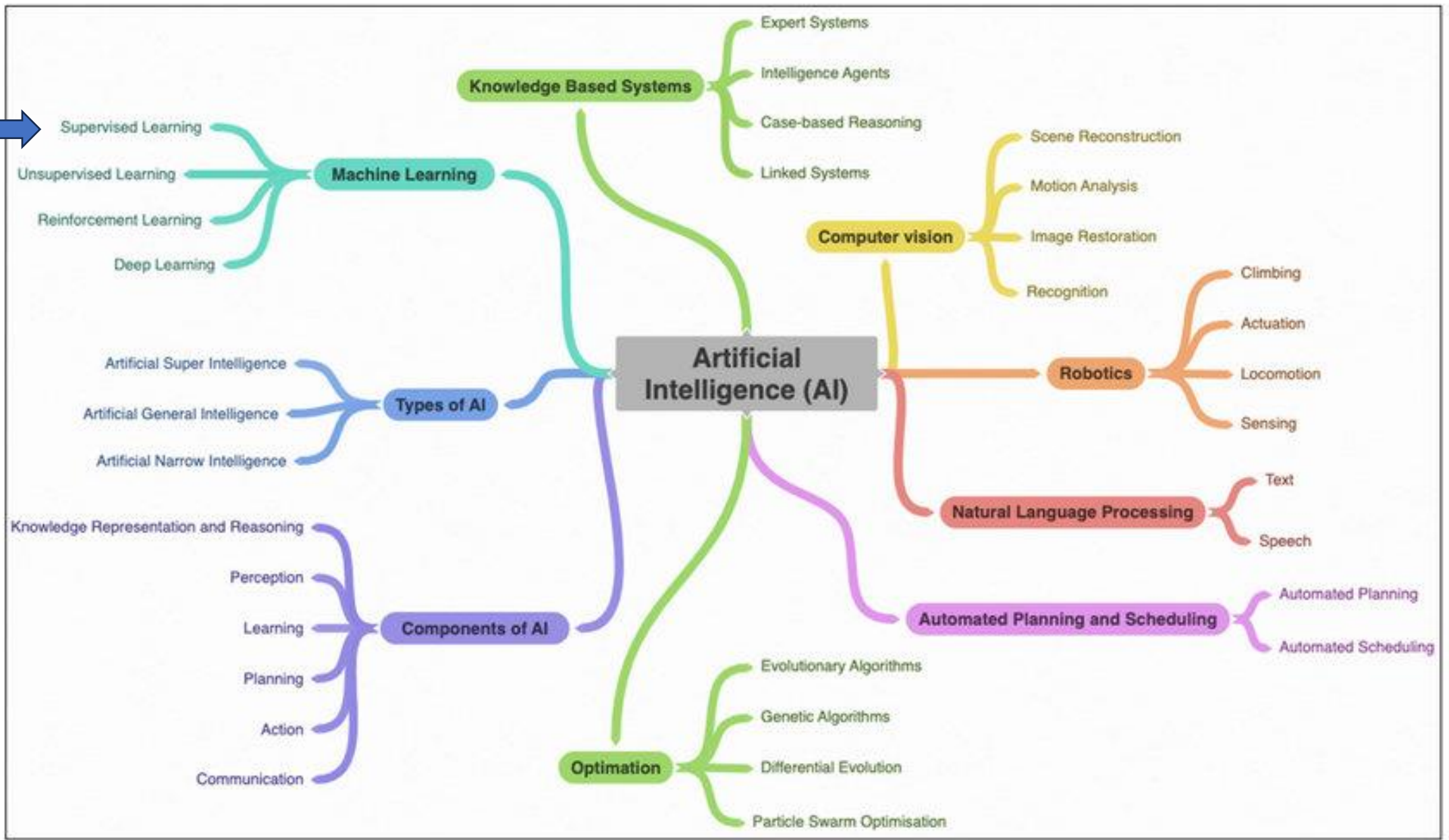
Origin of AI

- ***Dartmouth Summer Research on Artificial Intelligence, 1956***
- With:
 - **Marvin Minsky** (Turing Award-69),
 - **John McCarthy** (Turing Award-71),
 - **Claude Shannon** (Creator of Information Theory),
 - **Warren McCulloch** (first mathematical model of a neuron),
 - **John Nash** (Nobel-94),
 - **Herbert Simon** (Nobel-78, Turing Award-75)
 - **Frank Rosenblatt** (Creator of perceptrons – early versions of ANN)

Origin of AI

- *Dartmouth Summer Research on Artificial Intelligence, 1956*
- " *The study is to proceed on the basis of the conjecture that **every aspect of learning** or any other feature of **intelligence** can be so **precisely described that a machine can be made to simulate it**. An attempt will be made to find how to make machines **use language, form abstractions** and **concepts, solve kinds of problems now reserved for humans, and improve themselves**."*
- <https://home.dartmouth.edu/about/artificial-intelligence-ai-coined-dartmouth>







The AI we see...



Which weights more: 1kg of bricks or 2kg of feathers?



View other drafts 



1kg of bricks and 2kg of feathers weigh the same. Mass is the amount of matter in an object, and it is measured in kilograms. Weight is the force of gravity acting on an object, and it is measured in Newtons. In a vacuum, where there is no air resistance, 1kg of bricks and 2kg of feathers would weigh the same. However, in air, the feathers would experience more buoyant force than the bricks, so they would weigh slightly less.

The buoyant force is the force that pushes an object up in a fluid, such as air or water. The amount of buoyant force depends on the volume of the object and the density of the fluid. Feathers have a much larger volume than bricks, so they would experience more buoyant force.

So, in a vacuum, 1kg of bricks and 2kg of feathers would weigh the same. But in air, 2kg of feathers would weigh slightly less than 1kg of bricks.

Chat yesterday...

Default (GPT-3.5)



is a diminished chord the 3rd degree of a dominant chord?



No, a diminished chord is not typically the 3rd degree of a dominant chord. Let me explain the relationship between these two types of chords.



1. Dominant Chord:

- A dominant chord is a specific type of seventh chord that is built on the fifth degree of a diatonic scale. In a major key, the dominant chord is typically a major chord with a minor seventh added to it. For example, in the key of C major, the dominant chord is G7, which consists of the notes G, B, D, and F.

Chat yesterday...



Which notes has a B diminished chord?



A B diminished chord consists of the following notes:



1. B (the root of the chord)
2. D (the minor third above the root)
3. F (the diminished fifth above the root)

So, the notes in a B diminished chord are B, D, and F.

Chat yesterday...



find the contradiction in the two answers above.



I apologize for any confusion in the previous responses. Upon review, I see that there is no contradiction between the two answers. Both responses correctly describe the notes in a B diminished chord. The notes in a B diminished chord are B, D, and F, as mentioned in both answers. If you have any more questions or need further clarification, please feel free to ask.



Is a B diminished chord is built from the 3rd degree of a G dominant chort?



No, a B diminished chord is not typically built from the 3rd degree of a G dominant chord. Let me explain the relationship between these chords:



1. G Dominant Chord:

- A G dominant chord is typically based on the 5th degree of a C major scale (in the key of C major). In this context, the notes in a G dominant 7th chord (G7) are G, B, D, and F.

2. B Diminished Chord:

- A B diminished chord is constructed by stacking two minor thirds on top of each other. The notes in a B diminished chord are B, D, and F.

Social Challenges

- Job loss/Job replacement
- Skills gap
- Privacy concerns
- Bias and Fairness
- Transparency and Accountability
- **Ethical decision making in using and making AI**



Ethical decision making **in using** **and making AI**

A few examples...

Bias in Machine Learning

Hungarian has no gendered pronouns so Google Translate makes assumptions:

Google Translate



Sign in

Text

Documents

HUNGARIAN - DETECTED

POLISH

PO



ENGLISH

POLISH

PORTUGUESE



Ő szép. Ő okos. Ő olvas. Ő mosogat. Ő épít. Ő varr. Ő tanít. Ő főz. Ő kutat. Ő gyereket nevel. Ő zenél. Ő takarító. Ő politikus. Ő sok pénzt keres. Ő süteményt süt. Ő professzor. Ő asszisztens. |

She is beautiful. He is clever. He reads. She washes the dishes. He builds. She sews. He teaches. She cooks. He's researching. She is raising a child. He plays music. She's a cleaner. He is a politician. He makes a lot of money. She is baking a cake. He's a professor. She's an assistant.



194 / 5000



Racism is Poisoning Online Ad Delivery, Says Harvard Professor

Google searches involving black-sounding names are more likely to serve up ads suggestive of a criminal record than white-sounding names, says computer scientist

by **Emerging Technology from the arXiv**

February 4, 2013

<https://www.technologyreview.com/2013/02/04/253879/racism-is-poisoning-online-ad-delivery-says-harvard-professor/>

Ad related to latanya sweeney ⓘ

[Latanya Sweeney Truth](#)

www.instantcheckmate.com/

Looking for **Latanya Sweeney**? Check **Latanya Sweeney's** Arrests.

Ads by Google

[Latanya Sweeney, Arrested?](#)

1) Enter Name and State. 2) Access Full Background Checks Instantly.

www.instantcheckmate.com/

[Latanya Sweeney](#)

Public Records Found For: **Latanya Sweeney**. View Now.

www.publicrecords.com/

[La Tanya](#)

Search for La Tanya Look Up Fast Results now!

www.ask.com/La+Tanya

Ads by Google

[Kirsten Lindquist](#)

Get **Kirsten Lindquist** Find **Kirsten Lindquist**

www.ask.com/Kirsten+Lindquist

[We Found: Kristen Lindquist](#)

1) Contact **Kristen Lindquist** - Free Info! 2) Current Phone, Address & More.

www.peoplesmart.com/

[Search by Phone](#)

[Search by Email](#)

[Background Checks](#)

[Search by Address](#)

[Public Records](#)

[Criminal Records](#)

[Kristen Lindquist](#)

Public Records Found For: **Kristen Lindquist**. View Now.

www.publicrecords.com/

Toxic potential of YouTube Feedback Loops

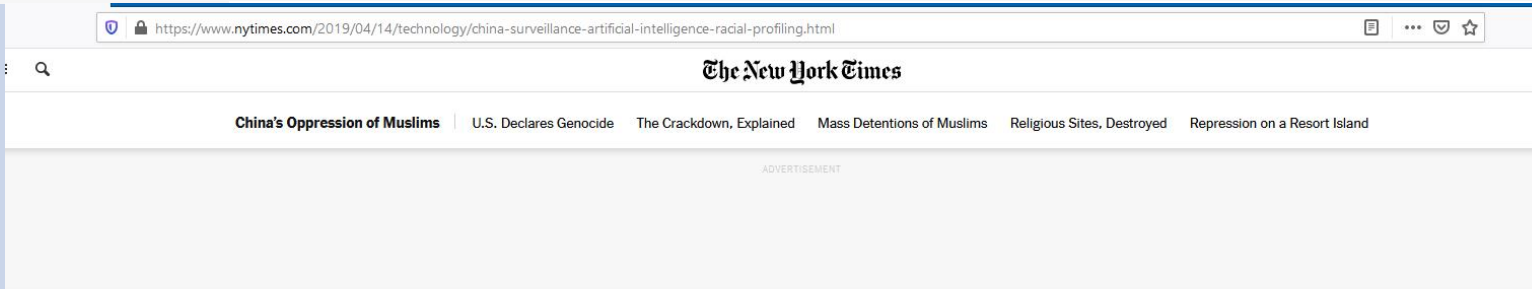
- Your model is controlling the next round of data that is used to build the very same model!
- <https://youtu.be/Et2n0J0OeQ8>
- The AI is directed at optimizing engagement
 - If you like cats, YouTube will only point you to cat films. After a while you'll believe that there is only cats in the world.
- Replace Cats by any political idea, or criminal behaviour...
- Recent examples: QAnon; Pedophilia

THE INTERPRETER

On YouTube's Digital Playground, an Open Gate for Pedophiles

- <https://www.nytimes.com/2019/06/03/world/americas/youtube-pedophiles.html>





One Month, 500,000 Face Scans: How China Is Using A.I. to Profile a Minority

In a major ethical leap for the tech world, Chinese start-ups have built algorithms that the government uses to track members of a largely Muslim minority group.



Uyghur Detector ?

- <https://www.nytimes.com/2019/04/14/technology/china-surveillance-artificial-intelligence-racial-profiling.html>



REFLECTIONS

Posted 21 November, 2016 by Zara Rahman

DANGEROUS DATA: THE ROLE OF DATA COLLECTION IN GENOCIDES

- <https://www.theengineroom.org/dangerous-data-the-role-of-data-collection-in-genocides/>



Everyone should watch Rachel's lecture!

- These cases are well described in the lecture on **Ethics in Data Science by Rachel Thomas - USF**
 - <https://course.fast.ai/videos/?lesson=5>

Teaching Ethical Decision Making under the AI Razor



Themis The Great Greek Mythology 2 oz Antique finish Silver Coin 2000
Francs CFA Cameroon 2023

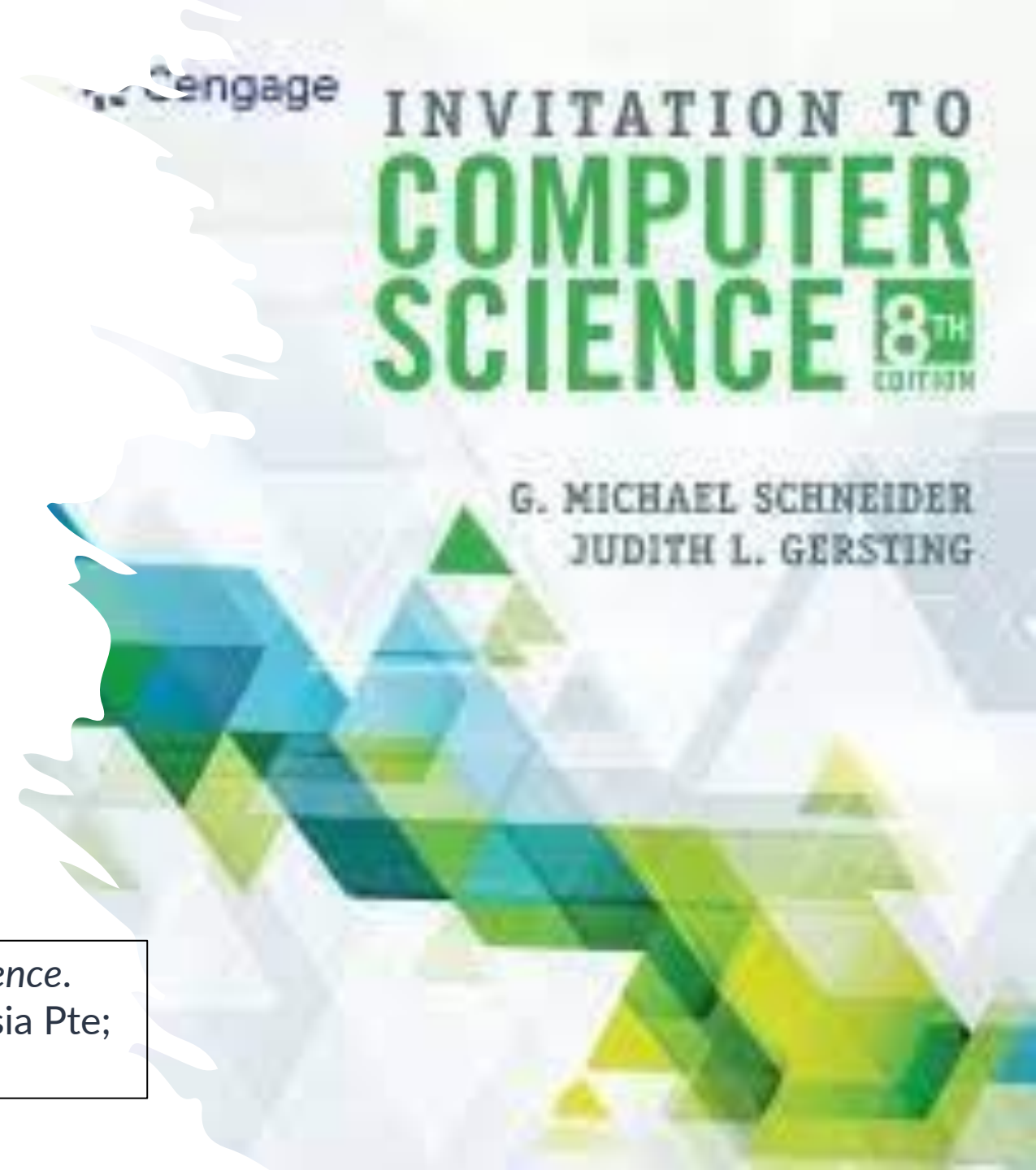
<https://www.topworldcoins.com/en/themis-the-great-greek-mythology-2-oz-antique-finish-silver-coin-2000-francs-cfa-cameroon-2023>

Paramedic ethics for technology

Combine distinct ways of
considering an ethical question:

- Utilitarian
- Dialectic
- Deontological
- Analogical

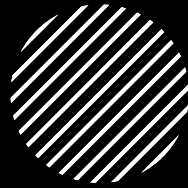
Schneider GM Gersting JL. *Invitation to Computer Science*.
8Th edition. Asia ed. Singapore: Cengage Learning Asia Pte;
2018.



“Paramedic Ethics for Technology”

- 1. Identify the stakeholders**
 - Narrowing the scope of the problem
- 2. Utilitarian step**
 - Identify benefits and costs
 - Various strategies could be used to aggregate costs, benefits and likelihood of the outcomes
- 3. Deontological step**
 - Consider duties, responsibilities, prohibitions and
- 4. Reasoning by analogy**
 - Think of an analogous situation to help clarify the situation
- 5. Make a decision (reflective equilibrium)**
or repeat the process in dialectic form and revise

“You’re not paranoid. Your phone really is listening in.”



- <https://www.foxnews.com/tech/youre-not-paranoid-your-phone-really-is-listening-in>

The scene plays out like a thriller: you pull out your phone, and you see an ad for AirPods.

Wait a minute, you think. Didn't I just have a conversation about AirPods with my friend? Like, a real conversation, spoken aloud? Is my phone... listening to me?

Why, yes, it probably is. When you use your default settings, everything you say may be recorded through your device's onboard microphone. Our phones routinely collect our voice data, store it in a distant server, and use it for marketing purposes. This fact was kept quiet for some time, but this kind of targeted ad is gradually becoming common knowledge.

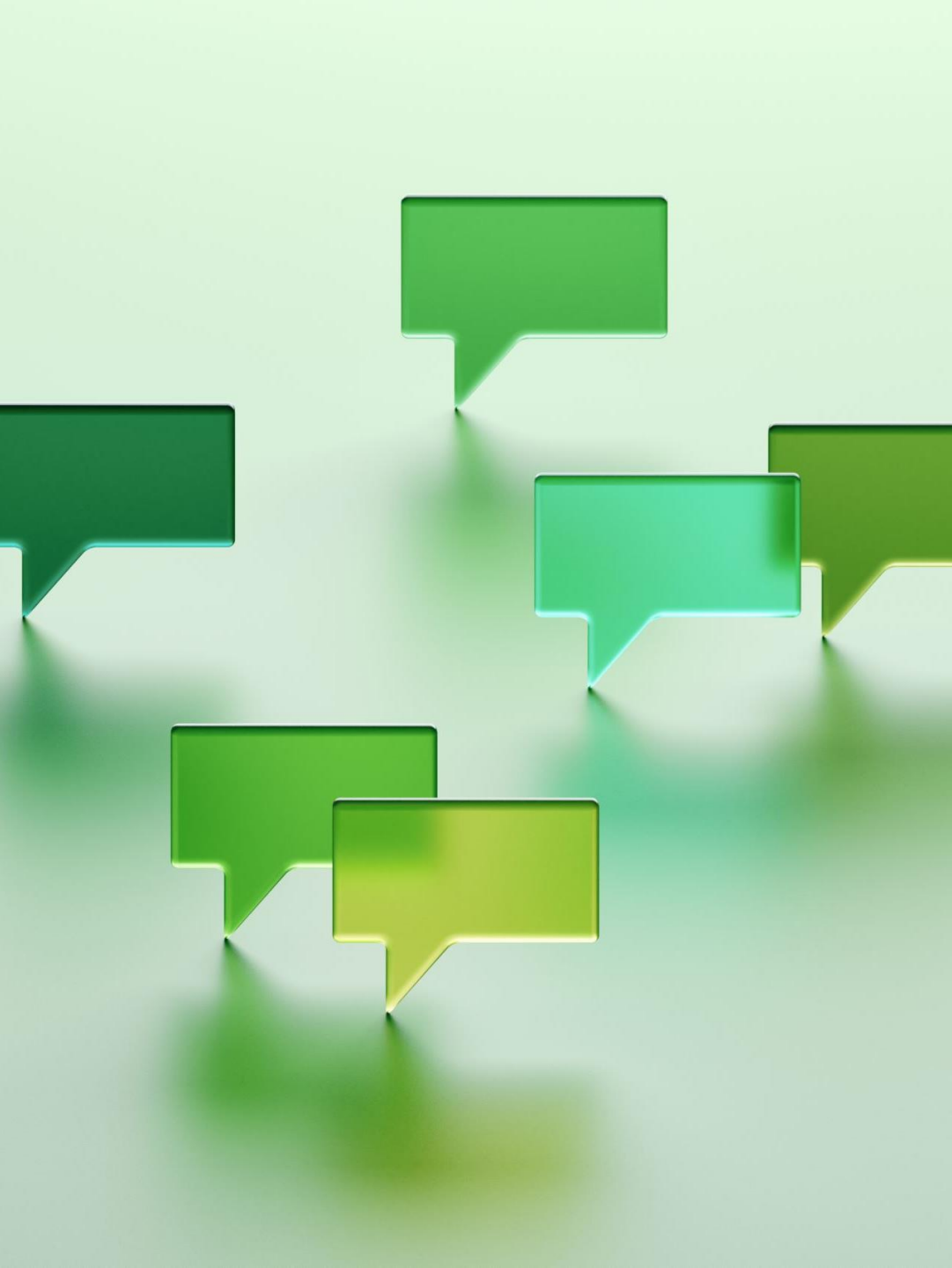


Exercise

- Use the paramedics method for ethical decision making to analyze ethics involved in case your “phone is actually listening to you in order to provide you with contextual adds.”

Step 1- Who are
the stakeholders ?





Step 1- Who are the stakeholders ?

- Company – internet provider
- Big tech (apple, google, FB,...)
- Users
- Everybody that is around the phone
- Phone manufacturer
- Business

Step 2- Utilitarian step: What is at stake?

$$\sum value_i * (Pr(event_i)),$$

Sum it all up and compare with a
threshold



Step 2- Utilitarian step:

- Privacy ($70 * 0.2$)
 - Personal info (age, email, address, gender, location)
- Social patterns ($10 * 0.3$)
- Financial info ($100 * 0.2$)
- Medical info ($50 * 0.05$)
- Competitor's business info ($100 * 0.2$)
- Internet Market ($50 * 0.8$)
- Email (SPAM) lists ($70 * 0.8$)
- Personal economics (...)

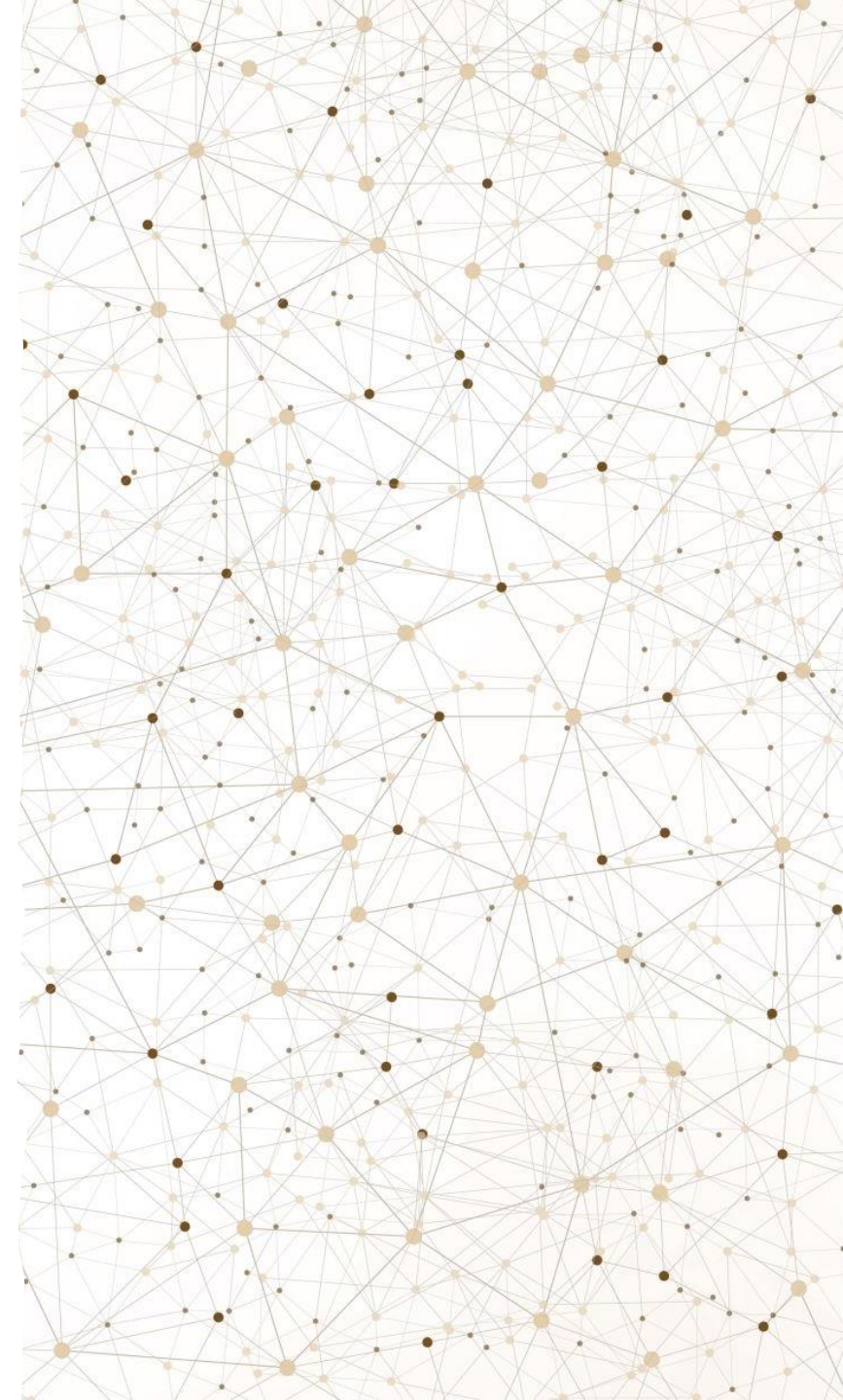
Is it Ethical?

Was the sum of utility values greater than the threshold?



Step 3 – Deontological Step

Identify duties, permissions,
obligations and
responsibilities



Step 3 – Deontological Step

- Companies should respect the privacy laws
 - (no)
- Users should be honest with the information they are providing
 - (no)
- Users should read terms and conditions
 - (No)
- T&C's should be simpler & or generic to a point so you dont have to re-read every time
 - (no)
- T&C should also be timed, so you can't click accept straight away
 - (yes and no)
- Companies have the obligation to protect children
 - (No)
- companies would have to comply with local area laws, will vary country to country with some things due to age etc.
 - (no)

Is it Ethical?

Were all the duties, permissions, obligations and responsibilities respected?



Step 4 - Reasoning by analogy



Step 4 – Analogy 1

Census collection

- Analise similarities or dissimilarities
 - Census is a non-profitable action
 - Census is external
 - Census - You know exactly what information is provided
 - aggregate data is publicly available (but not specific to the user)
 - No transparency in google

Step 4 – Analogy 2

Google search engine

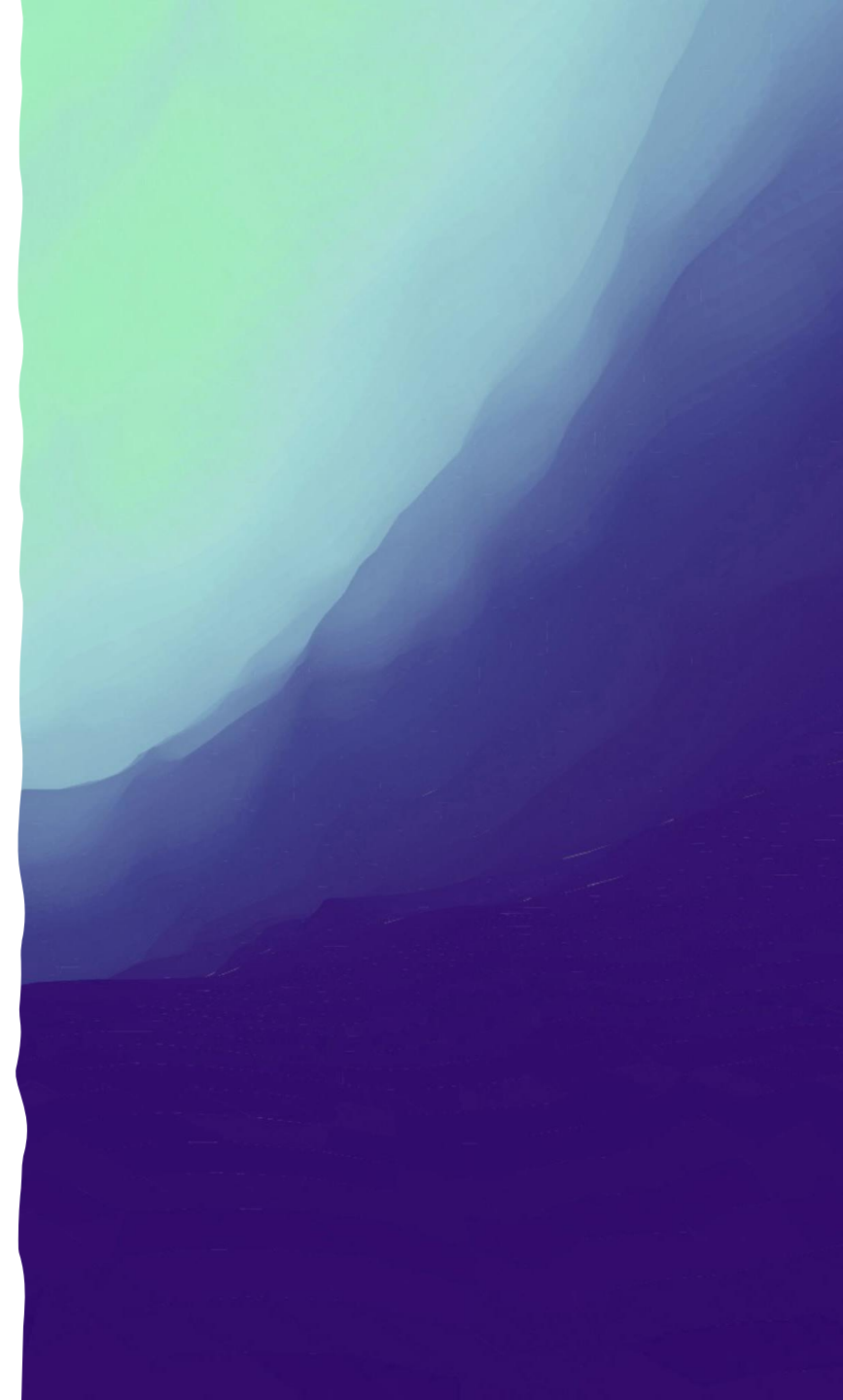
- Similarities
 - Support future decision
 - We don't know how this info is used
 - Invasive
 - Invasion of privacy

Is it Ethical?

Did we understand the problem well to make a decision?



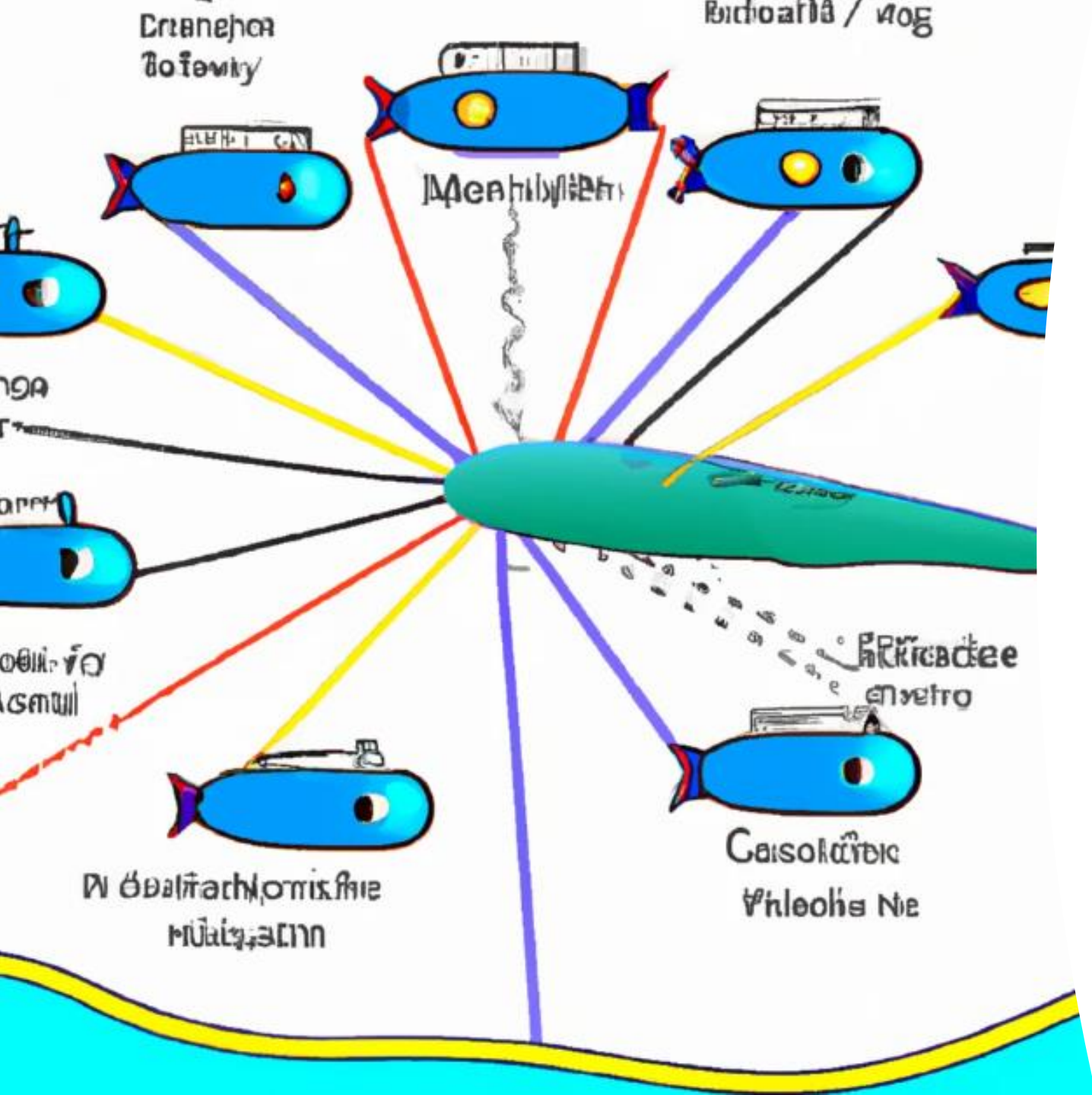
Step 5 – Make a
decision or loop
again





Thank you!

Don't forget the marketing time!



PhD Research Project on Offer:

Narratives for Multi-Robot Event Recognition (at Flinders University, Sponsored by Naval Group)

Extracting **consistent narratives** from the sensor data of multi-agent systems with Anne-Gwenn Bosser

- Efficiently representing and reasoning about scenes using commonsense knowledge graph, from the **perspective of distributed systems**.