



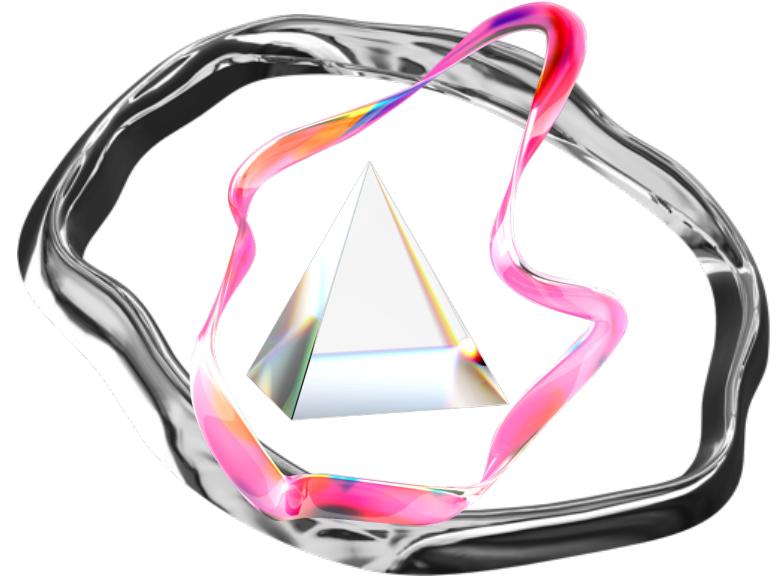
**AI + AI + AI**  
**AI MODULE X SCREENS**

enrique.encinas @aho.no

AIAIAINews

# 3 WEEKS

13th March -> 31st March



**HTTPS://GITHUB.COM/ENRIQUEKI/AIAIAI**  
+  
**SLACK**

# WEEK #3

22nd March -> 24th March



**AI X NEXT WEEK**

**:: AI ETHICS + SPECULATIVE AI DESIGN**

## **ARTIFICIAL INTELLIGENCE**

Any technique that allows computers to mimic human intelligence

### **MACHINE LEARNING**

Algorithms that use statistics to find patterns in large amounts of data

### **DEEP LEARNING**

ML methods based on artificial neural networks

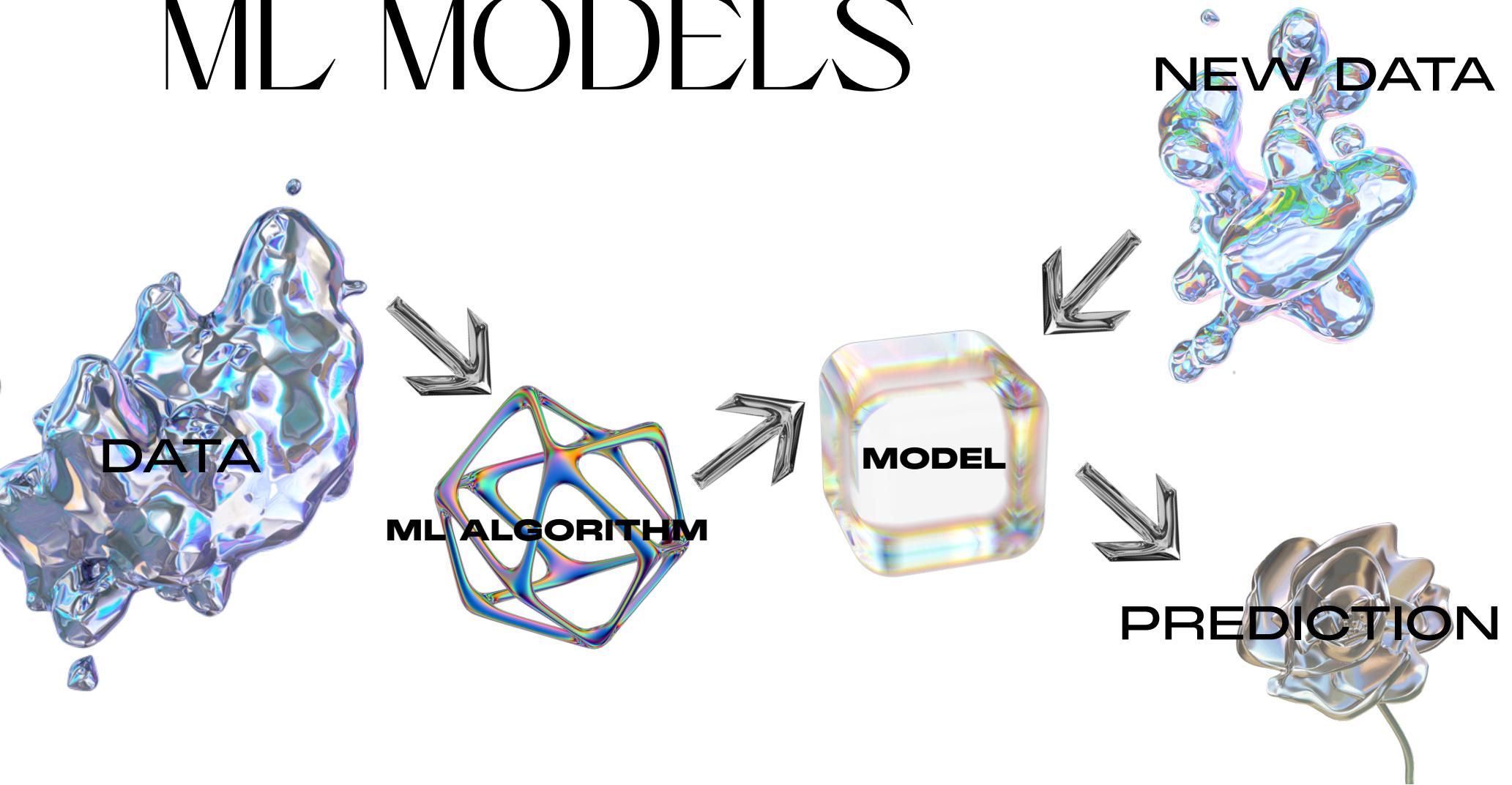
# AI WHAT?

The background consists of a dense, abstract pattern of wavy, monochromatic lines in shades of gray. These lines create a sense of depth and motion, resembling ripples on water or complex architectural structures. The overall effect is organic and modern.

AIETHICS

<https://icarus.kumu.io/fluxus-landscape>

# ML MODELS



**Behind the scenes of TV's first deep fake comedy: 'None of it is illegal. Everything is silly'**



Fake it till you make it ... 'Stormzy' and 'Harry Kane'. Photograph: PA

Is that Harry Kane and Stormzy arguing over a brie? This new, CGI-assisted comedy is astonishing but is it ethically OK?

**There is no standard: investigation finds AI algorithm objectify women's bodies**

Guardian exclusive: AI tools rate photos of women as more sexually suggestive than those of men, especially if nipples, pregnant bellies or exercise is involved

• *This story was produced in partnership with the Pulitzer Center's AI Accountability Network*

by [Gianluca Mauro](#) and [Hilke Schellmann](#)

## *The People Onscreen Are Fake. The Disinformation Is Real.*

[Give this article](#) [Email](#) [Bookmark](#) [457](#)



By Adam Satariano and Paul Mozur

Adam Satariano, based in London, and Paul Mozur, based in Seoul, are tech correspondents who report internationally about online disinformation.

Feb 7, 2023

閱讀時間 中文版 Leer en español

In one video, a news anchor with perfectly combed dark hair and a stubbly beard outlined what he saw as the United States' shameful lack of action against gun violence.

60 MINUTES OVERTIME > ChatGPT and large language model bias

60 MINUTES OVERTIME

MARCH 8, 2023 / 7:00 P.M. / CBS NEWS



**60 Overtime**

<https://haveibeentrained.com/>

THEORY?

“A lot of our ideas about AI come from science fiction. Welcome to everything in Hollywood. It’s the Terminator. It’s Commander Data from Star Trek... It is the robots that take over the world and start to think like human beings. And it’s all totally imaginary. What we actually have is...narrow AI and narrow AI is just math. We’ve imbued computers with all of this magical thinking.”

Meredith Broussard. Author of Artificial Unintelligence.

“If you’re thinking about data and artificial intelligence, in many ways data is destiny. Data is what we’re using to teach machines how to learn different kinds of patterns. So if you have largely skewed data sets that are being used to train these systems, you can also have skewed results. So this is when you think of AI, it’s forward looking, but AI is based on data and data is a reflection of our history. The past dwells within our algorithms.”

Joy Buolamwini. Founder of the Algorithmic Justice League.

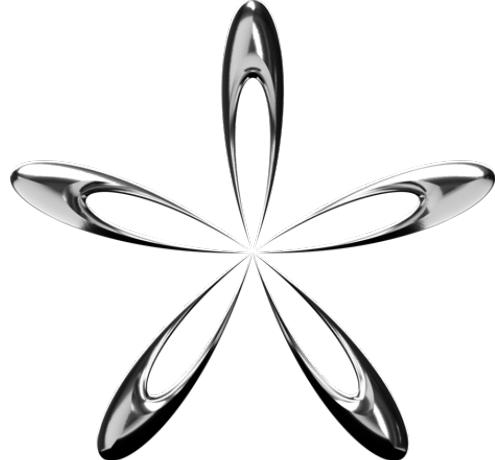
“Machine learning. It’s a scoring system that scores the probability of what you’re about to do. Are you going to pay back this loan? Are you going to get fired from this job? What worries me the most about AI or whatever you want to call it, algorithms, is power because it’s really all about who owns the f\*\*\*\*g code. The people who own the code then deploy it on other people and there is no symmetry there.”

Cathy O’Neill. Author of Weapons of Math Destruction.

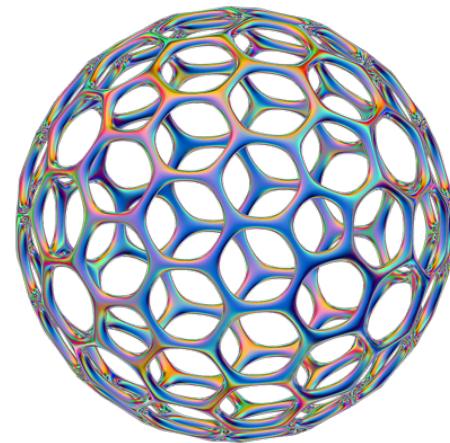
COMPLEXITY?

# TYPES OF ML

**SUPERVISED  
LEARNING**



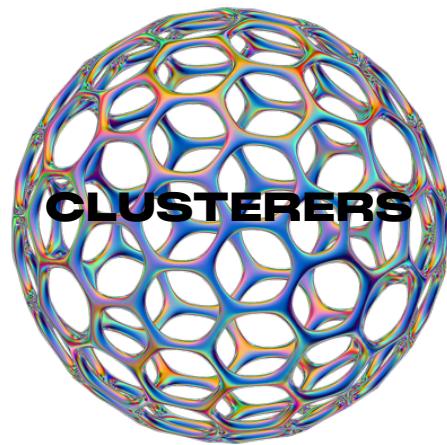
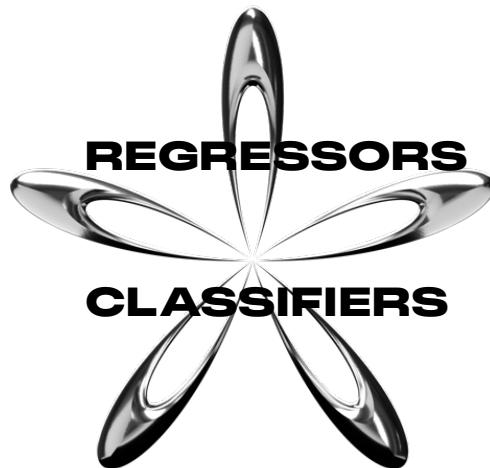
**UNSUPERVISED  
LEARNING**



**REINFORCEMENT  
LEARNING**

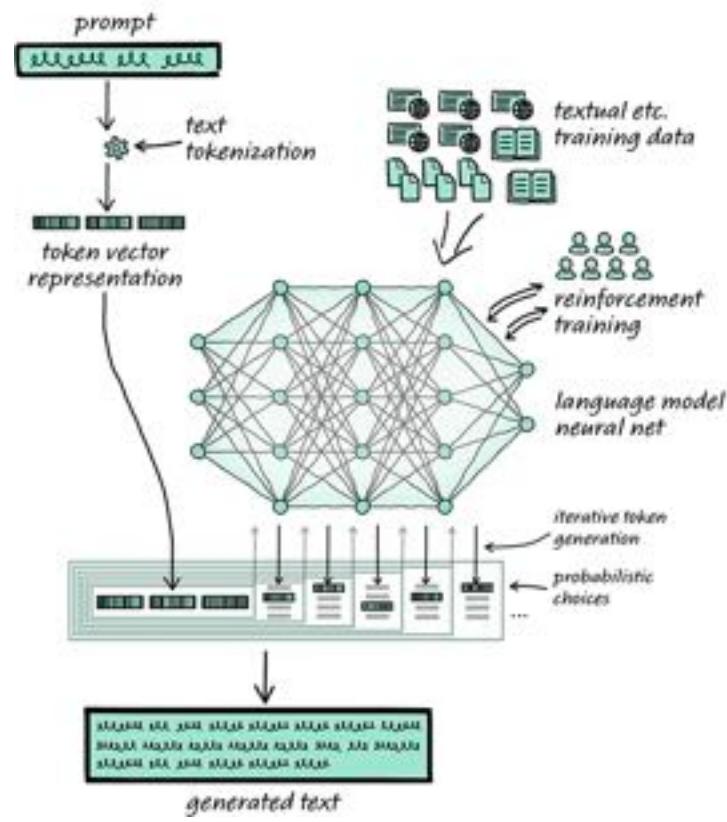


# USE CASES :



<https://becominghuman.ai/a-primer-of-29-interactions-for-ai-866164ab12f0>

ChatGPT



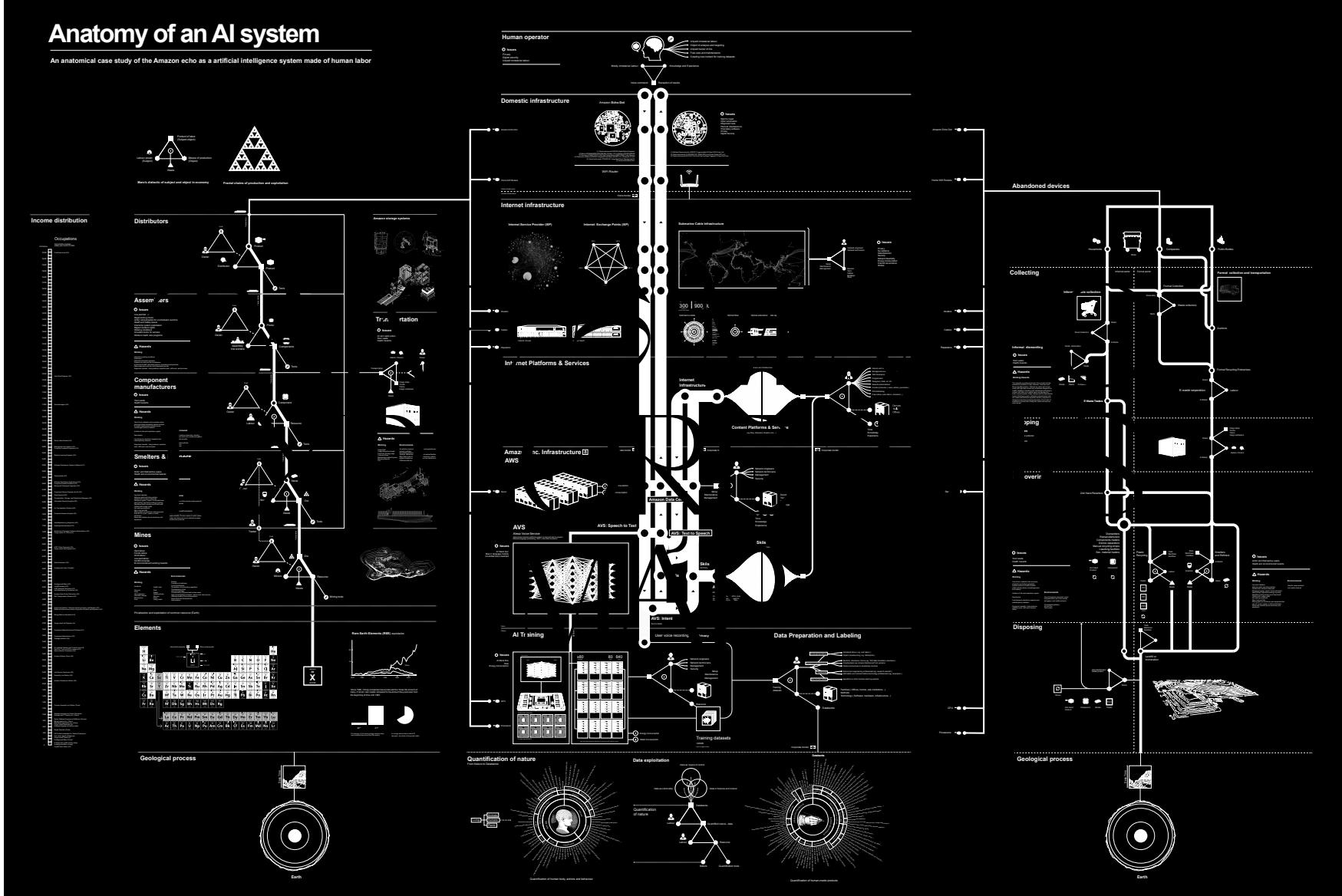
<https://writings.stephenwolfram.com/2023/01/wolframalpha-as-the-way-to-bring-computational-knowledge-superpowers-to-chatgpt/>

Currently, the research and development of AI is primarily in the hands of big technology corporations—who have made astonishing profits off the technology. The U.S. government has taken a mostly hands-off approach to influencing or regulating AI development. With the increasing use of commercial AI applications in the areas of healthcare, law enforcement, employment, and elections, however, there is a growing public awareness of the potential risks of leaving the use of AI technology unmonitored.

Cathy O'Neill. Author of Weapons of Math Destruction.

# Anatomy of an AI system

An anatomical case study of the Amazon echo as a artificial intelligence system made of human labor



HISTORY?

## History of machine learning

- 1957 — Perceptron algorithm (implemented as a circuit!)
- 1959 — Arthur Samuel wrote a learning-based checkers program that could defeat him
- 1969 — Minsky and Papert's book *Perceptrons* (limitations of linear models)
- 1980s — Some foundational ideas
  - ▶ Connectionist psychologists explored neural models of cognition
  - ▶ 1984 — Leslie Valiant formalized the problem of learning as PAC learning
  - ▶ 1988 — Backpropagation (re-)discovered by Geoffrey Hinton and colleagues
  - ▶ 1988 — Judea Pearl's book *Probabilistic Reasoning in Intelligent Systems* introduced Bayesian networks

Intro ML (UofT)

CSC311-Lec1

17

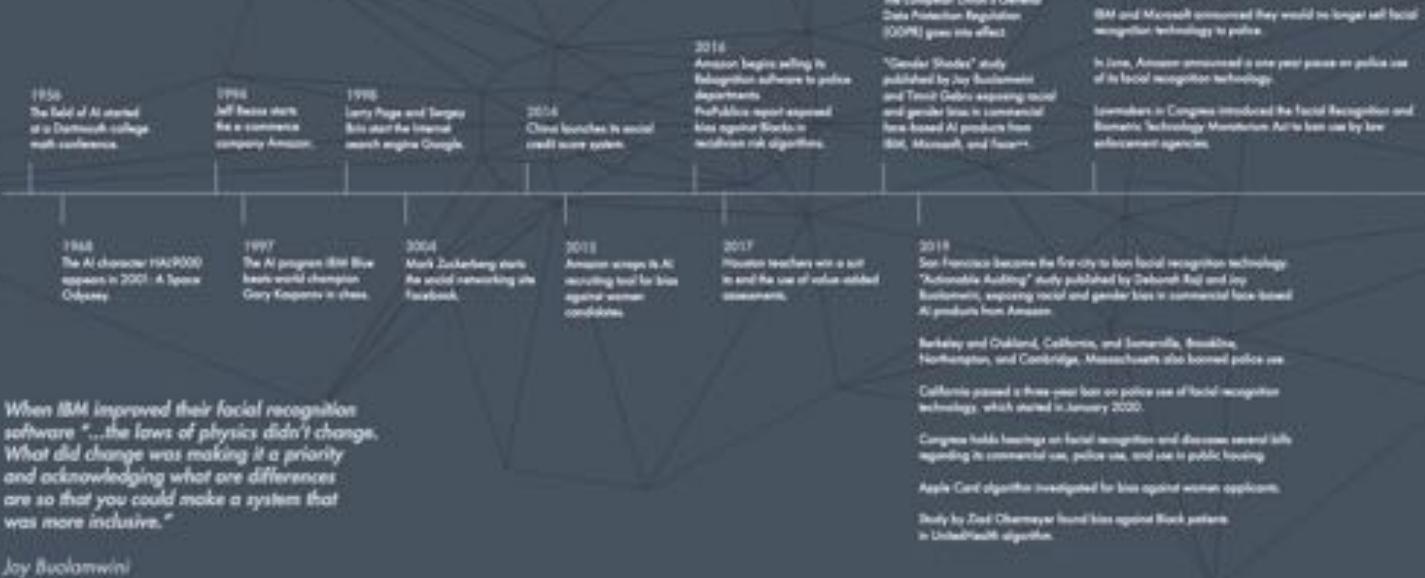
## History of machine learning

- 1990s — the “AI Winter”, a time of pessimism and low funding
- But looking back, the '90s were also sort of a golden age for ML research
  - ▶ Markov chain Monte Carlo
  - ▶ variational inference
  - ▶ kernels and support vector machines
  - ▶ boosting
  - ▶ convolutional networks
  - ▶ reinforcement learning
- 2000s — applied AI fields (vision, NLP, etc.) adopted ML
- 2010s — deep learning
  - ▶ 2010–2012 — neural nets smashed previous records in speech-to-text and object recognition
  - ▶ increasing adoption by the tech industry
  - ▶ 2016 — AlphaGo defeated the human Go champion
  - ▶ 2018–now — generating photorealistic images and videos
  - ▶ 2020 — GPT3 language model

# ONE HISTORY OF AI AMONG MANY

# ONE HISTORY OF AI AMONG MANY

This timeline provides a snapshot overview of the rapid development of AI and the growing movement to expose AI bias and invasive surveillance.



CODING WARS

EDUCATIONAL DISCUSSION GUIDE

# ALGORITHMIC JUSTICE

- That people have agency and control over how they interact with an AI system, which requires knowing how systems are used and what their potential harms might be.
- That AI secure affirmative consent from people, such as with an opt-in request, on how or whether they interact with an AI system. No one should be penalized for opting out.
- To respect human life, dignity, and rights, AI requires centering justice by prohibiting certain corporate and government uses.
- Meaningful transparency. For an AI system to demonstrate meaningful transparency it must provide an explanation of how the system works, how it was designed, and for what specific purpose—as well as its known limitations.
- Because it is constantly evolving, AI requires continuous oversight by independent third parties and laws that require companies and government agencies deploying AI to meet minimum reporting requirements.
- That AI provides people who have been harmed with access to remedy, meaning that there is a working pathway for people to contest and correct a harmful decision made by artificial intelligence.

<https://www.ajl.org/>

Individual Harms		Collective / Societal Harms
Illegal	Unfair	
<b>Loss of Opportunity</b>		
<b>Employment Discrimination</b> E.g. Filtering job candidates by race or genetic/health information	E.g. Filtering candidates by work proximity leads to excluding minorities	<b>Differential Access to Job Opportunities</b>
<b>Insurance &amp; Social Benefit Discrimination</b> E.g. Higher termination rate for benefit eligibility by religious group	E.g. Increasing auto insurance prices for night-shift workers	<b>Differential Access to Insurance &amp; Benefits</b>
<b>Housing Discrimination</b> E.g. Landlord relies on search results suggesting criminal history by race	E.g. Matching algorithm less likely to provide suitable housing for minorities	<b>Differential Access to Housing</b>
<b>Education Discrimination</b> E.g. Denial of opportunity for a student in a certain ability category	E.g. Presenting only ads on for-profit colleges to low-income individuals	<b>Differential Access to Education</b>
<b>Economic Loss</b>		
<b>Credit Discrimination</b> E.g. Denying credit to all residents in specified neighborhoods ("redlining")	E.g. Not presenting certain credit offers to members of certain groups	<b>Differential Access to Credit</b>
<b>Differential Pricing of Goods and Services</b> E.g. Raising online prices based on membership in a protected class	E.g. Presenting product discounts based on "ethnic affinity"	<b>Differential Access to Goods and Services</b>
	<b>Narrowing of Choice</b> E.g. Presenting ads based solely on past "clicks"	<b>Narrowing of Choice for Groups</b>
<b>Social Detriment</b>		
<b>Network Bubbles</b> E.g. Limited exposure to opportunity or evaluation based on "who you know"	E.g. Filter Bubbles E.g. Algorithms that promote only familiar news and information	
<b>Dignitary Harms</b> E.g. Emotional distress due to bias or a decision based on incorrect data	<b>Stereotype Reinforcement</b> E.g. Assumption that computed decisions are inherently unbiased	
<b>Constraints of Bias</b> E.g. Constrained conceptions of career prospects based on search results	<b>Confirmation Bias</b> E.g. All-male image search results for "CEO," all-female results for "teacher"	
<b>Loss of Liberty</b>		
<b>Constraints of Suspicion</b> E.g. Emotional, dignitary, and social impacts of increased surveillance	<b>Increased Surveillance</b> E.g. Use of "predictive policing" to police minority neighborhoods more	
<b>Individual Incarceration</b> E.g. Use of "recidivism scores" to determine prison sentence length (legal status uncertain)	<b>Disproportionate Incarceration</b> E.g. Incarceration of groups at higher rates based on historic policing data	

The image above is an excerpt from "Distilling the Harms of Automated Decision-Making" by the Future of Privacy Forum Report.

## Worksheet: UX of AI challenges

1. Explainability - How will we help our user understand certain outcomes?	2. Managing expectations - How will we establish realistic expectations?	3. Graceful failure & accountability - How will we design for trust in case of failure?
4. User feedback - How will your user provide feedback to the system?	5. User autonomy - How will the user be able to customize their experience?	6. Data privacy & security - How will you collect, store, and handle data?
7. Computational translation - How will you turn needs into parameters?	8. Bias & inclusivity - How will you prevent bias and guard inclusivity?	9. Ethics & (un)intended consequences - How will you look out for negative and positive impact?
10. Which other (design) challenges do you foresee?		

<https://medium.com/aixdesign/ux-challenges-for-ai-ml-products-3-3-value-alignment-e0b6b22ad9>

# *AI & Ethics: Collaborative Activities for Designers*



OUR FOCUS



## Ethics By Design and Ethics of Use Approaches for Artificial Intelligence

Version 1.0  
25 November 2021

[https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/ethics-by-design-and-ethics-of-use-approaches-for-artificial-intelligence\\_he\\_en.pdf](https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/ethics-by-design-and-ethics-of-use-approaches-for-artificial-intelligence_he_en.pdf)

# ETHICAL PRINCIPLES

1. Respect for Human Agency: human beings must be respected to make their own decisions and carry out their own actions. Respect for human agency encapsulates three more specific principles, which define fundamental human rights: autonomy, dignity and freedom.
2. Privacy and Data governance: people have the right to privacy and data protection and these should be respected at all times;
3. Fairness: people should be given equal rights and opportunities and should not be advantaged or disadvantaged undeservedly;
4. Individual, Social and Environmental Well-being: AI systems should contribute to, and not harm, individual, social and environmental wellbeing;
5. Transparency: the purpose, inputs and operations of AI programs should be knowable and understandable to its stakeholders;
6. Accountability and Oversight: humans should be able to understand, supervise and control the design and operation of AI based systems, and the actors involved in their development or operation should take responsibility for the way that these applications function and for the resulting consequences.

[https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/ethics-by-design-and-ethics-of-use-approaches-for-artificial-intelligence\\_he\\_en.pdf](https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/ethics-by-design-and-ethics-of-use-approaches-for-artificial-intelligence_he_en.pdf)

## Annex I Checklist: Specification of Objectives against Ethical Requirements

This checklist is a supporting tool and does not constitute an exhaustive list of all ethics requirement that may be applicable to the development of each specific AI system. It has to be used in conjunction with Part 1-5 of the current guidelines and applied to a degree matching the type of AI system and the research being proposed (from basic to precompetitive).

Specification of Objectives against Ethical Requirements	Yes	No (how potential risks will be mitigated?)
<b>Respect For Human Agency</b>		
End-users and others affected by the AI system are not deprived of abilities to make all decisions about their own lives; have basic freedoms taken away from them.		
End-users and others affected by the AI system are not subordinated, coerced, deceived, manipulated, objectified or dehumanized; nor is attachment or addiction to the system and its operations being stimulated.		
The system does not autonomously make decisions about vital issues that are normally decided by humans by means of free personal choices or collective deliberations or similarly significantly affects individuals.		
The system is designed in a way that give system operators and, as much as possible, end-users the ability to control, direct and intervene in basic operations of the system (when relevant)		
<b>Privacy &amp; Data Governance</b>		
The system processes data in line with the requirements for lawfulness, fairness and transparency set in the national and EU data protection legal framework and the reasonable expectations of the data subjects.		
Technical and organisational measures are in place to safeguard the rights of data subjects (through measures such as anonymization, pseudonymisation, encryption, and aggregation).		
There are security measures in place to prevent data breaches and leakages (such as mechanisms for logging data access and data modification).		
<b>Fairness</b>		
The system is designed to avoid algorithmic bias, in input data, modelling and algorithm design.		
The system is designed to avoid historical and selection bias in data collection, representation and measurement bias in algorithmic training.		

The background consists of a dense, abstract pattern of wavy, monochromatic lines in shades of gray and white, creating a sense of depth and motion. The lines are more concentrated in the center and spread out towards the edges.

# HANDS ON ETHICS

# THE ETHICS/DESIGN ASSESSMENT



45.min - In pairs, use the Checklist here:

[https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/ethics-by-design-and-ethics-of-use-approaches-for-artificial-intelligence\\_he\\_en.pdf](https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/ethics-by-design-and-ethics-of-use-approaches-for-artificial-intelligence_he_en.pdf)

To evaluate the case described here:

<https://www.wired.com/story/welfare-state-algorithms/>

30.min - Find another pair and share your work.

The background consists of a dense, abstract pattern of wavy, monochromatic lines in shades of gray and white, creating a sense of depth and motion. The lines are more concentrated in the center and spread out towards the edges.

**WATCH PARTY!**

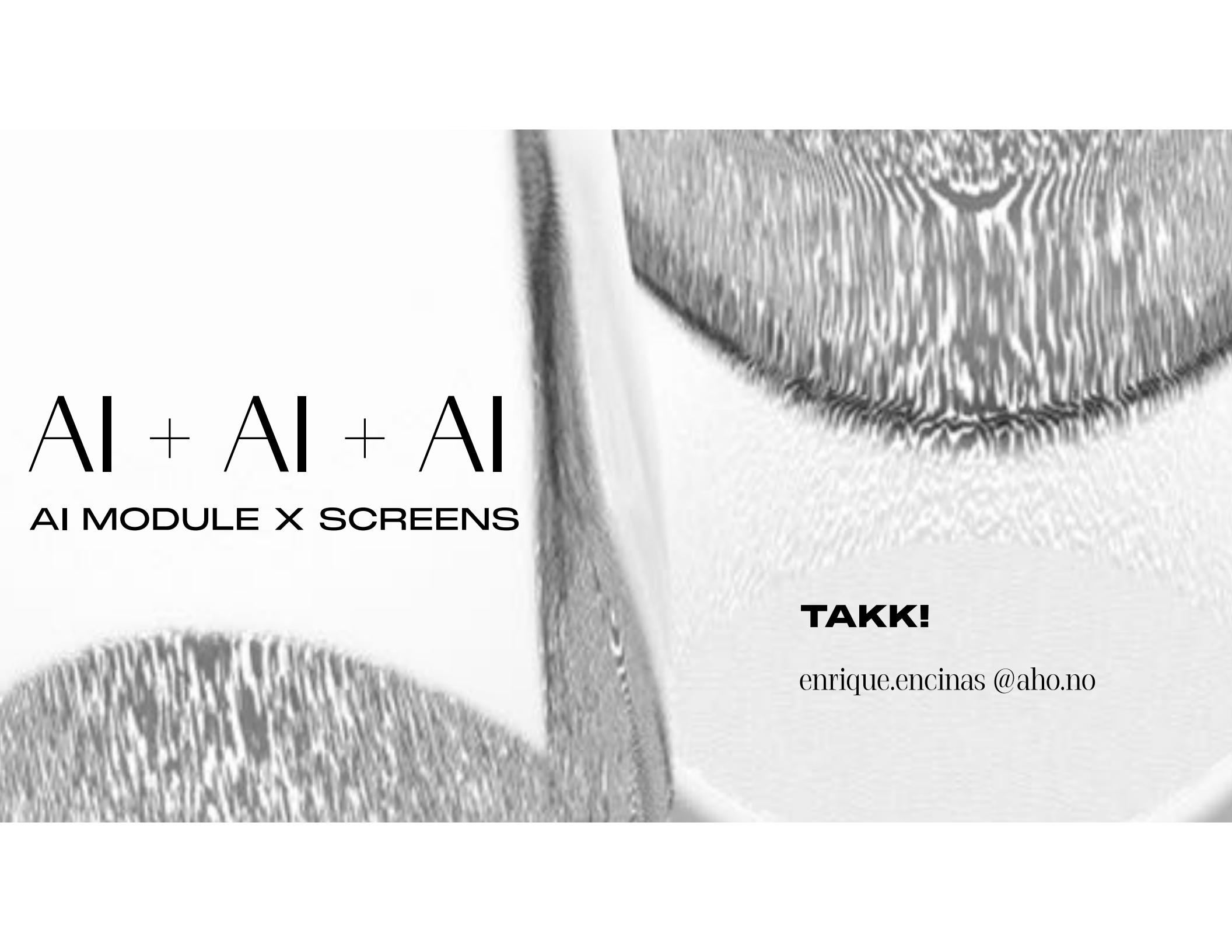
# C O D E D   B I A S



A SHALINI KANTAYYA FILM



WMM A WOMEN MAKE MOVIES RELEASE



**AI + AI + AI**  
**AI MODULE X SCREENS**

**TAKK!**

enrique.encinas @aho.no