Oppenheimer Center for Entrepreneurship and Innovation

Nov 12, 2021

ARTIFICIAL INTELLIGENCE

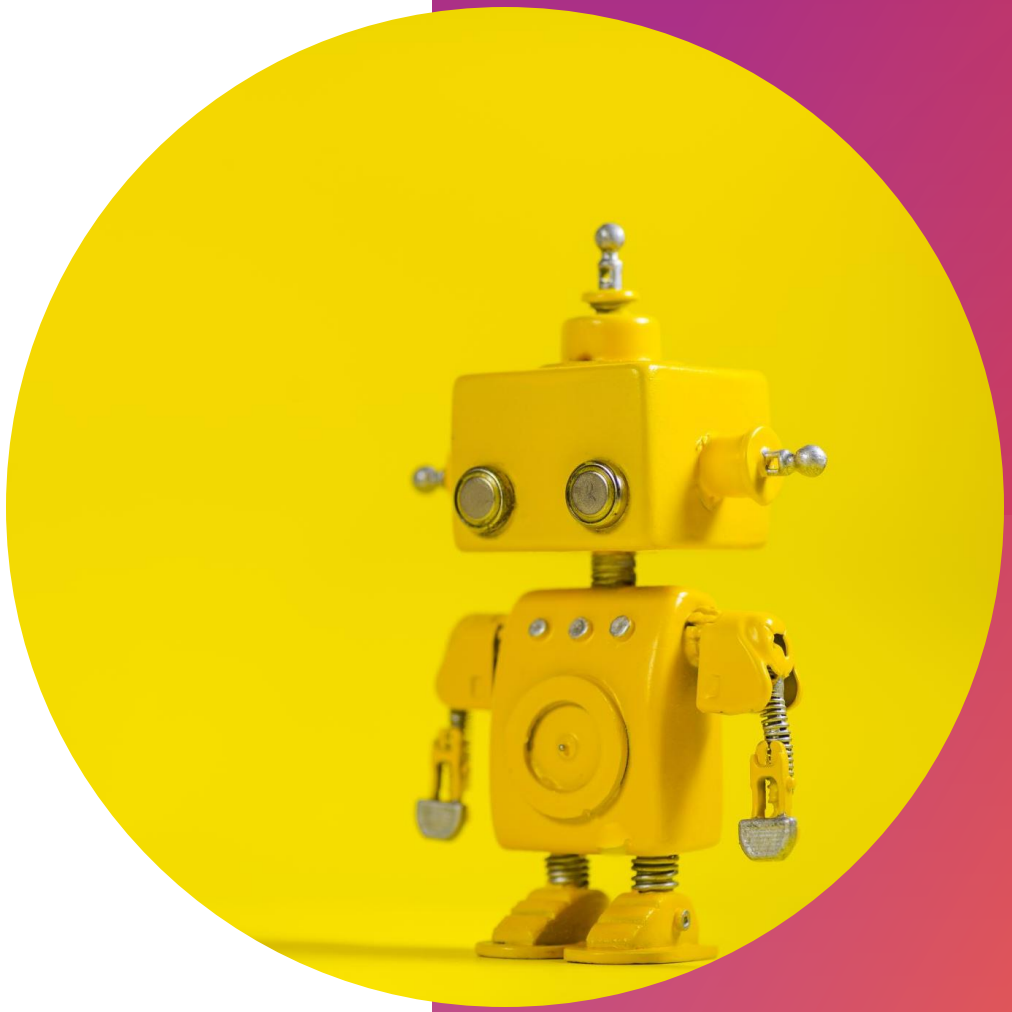
DR. SUGATA BANERJI

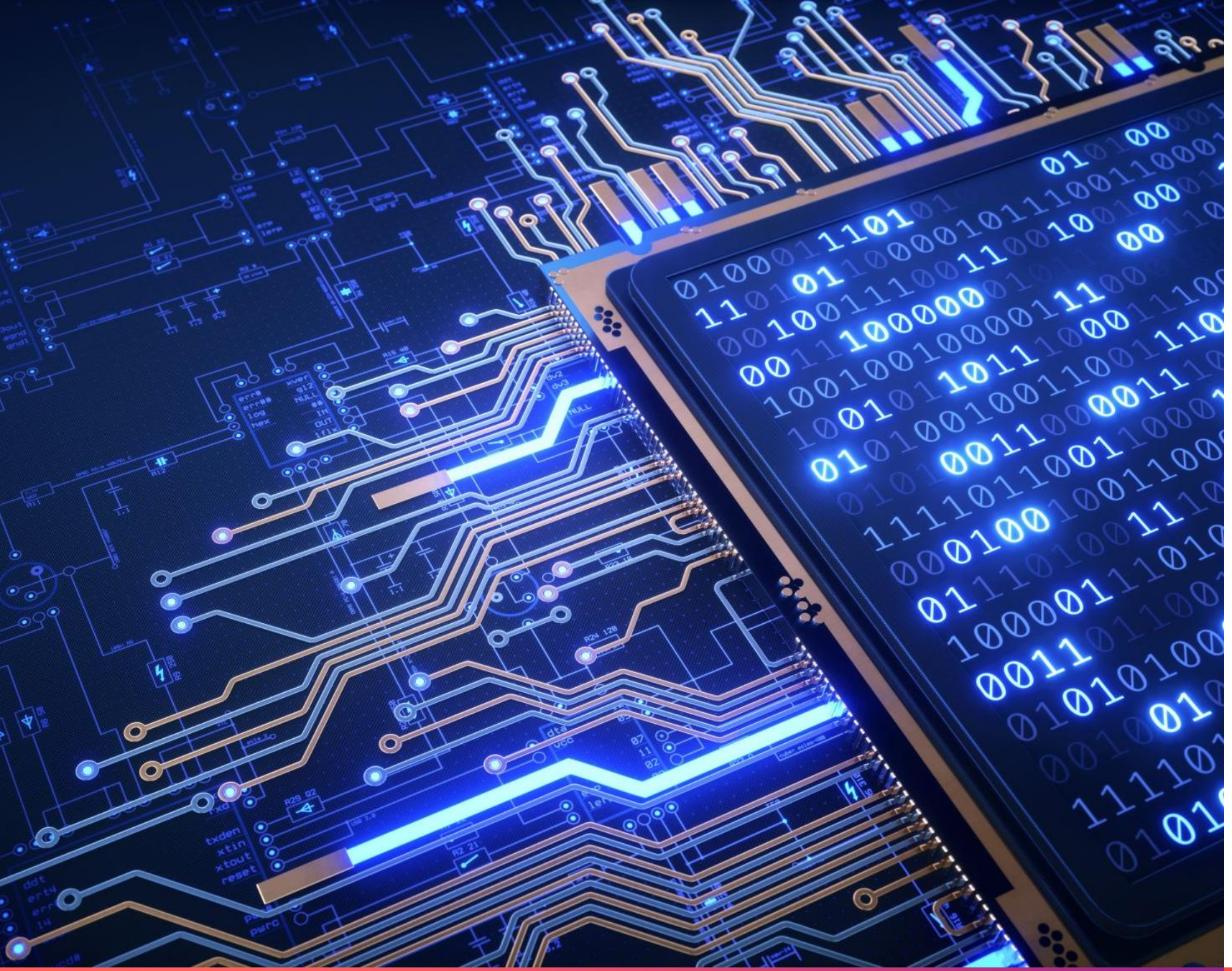
DR. SARA ZELENBERG

DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE AT LAKE FOREST COLLEGE

WHAT IS AI?

- **Artificial intelligence (AI)** is a subject that attempts to build software that makes decisions in line with intelligent beings, such as a human or an animal.
- It relies on computer programming.
- Most programming involves explicit commands.
- Some programming involves randomness and statistical techniques.





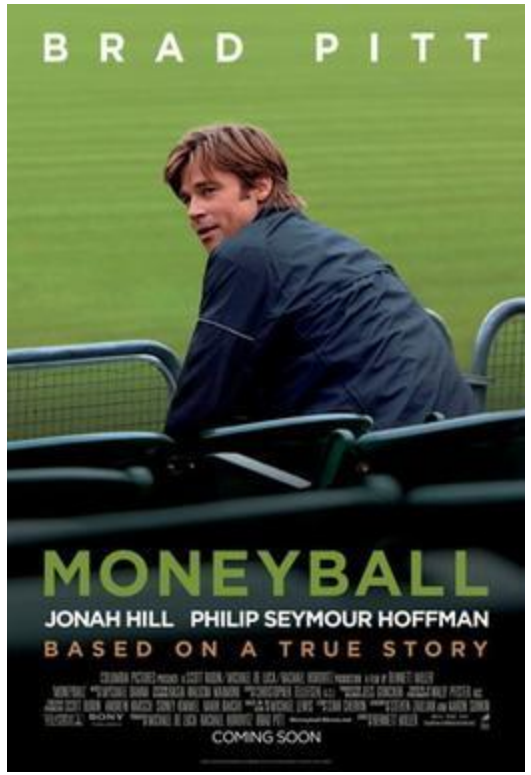
TYPES OF AI

- There are many different approaches in AI.
- Some attempt to capture high-level knowledge and program explicit commands using that knowledge.
- Some attempt to make representations of the world and write explicit programs that make plans in response.
- Some attempt to infer from large datasets using statistical techniques.

MACHINE LEARNING

- Machine learning (ML) is the hottest topic in AI currently.
- ML relies on statistical models applied to datasets to do one of two things::
  - Predict
    - Example: Oakland Athletics predictions of undervalued players (Moneyball)
  - Classify
    - Example: Is this a chihuahua or a muffin?

Taken from Medium article by Cristian Duget



Taken from Wikipedia

NEURAL NETWORKS

- The most popular technique in ML is currently neural networks.
- Neural networks attempt to build a model based on a dataset (as well as hyperparameters provided by the set of programmers).
    - This is an ansatz model that attempts to infer a pattern.
    - NOTE: Programmers generally don't know the pattern created until afterwards!

CASE STUDY: PREGNANCY

- Some believe that pregnancy and other health information can be inferred by the heartbeat.
  - This is not a widely held belief.
- A research group, new to neural networks, attempted to write a program to predict if a woman was pregnant from heartbeat data.
- Their program was measured to be 95% accurate.
- Further analysis showed that 95% of their dataset was pregnant women and the model just guessed everyone was pregnant.
- We can control for this mistake, but we can't control for all possible "bad patterns" that seem to work in general.

NEURAL NETWORKS

CONCEPT

- Neural networks have inputs and outputs.
- Datasets used to make them have both inputs and outputs.
- These datasets are broken up into a training set and a testing set.
- The model is built with the training set and evaluated with the testing set.

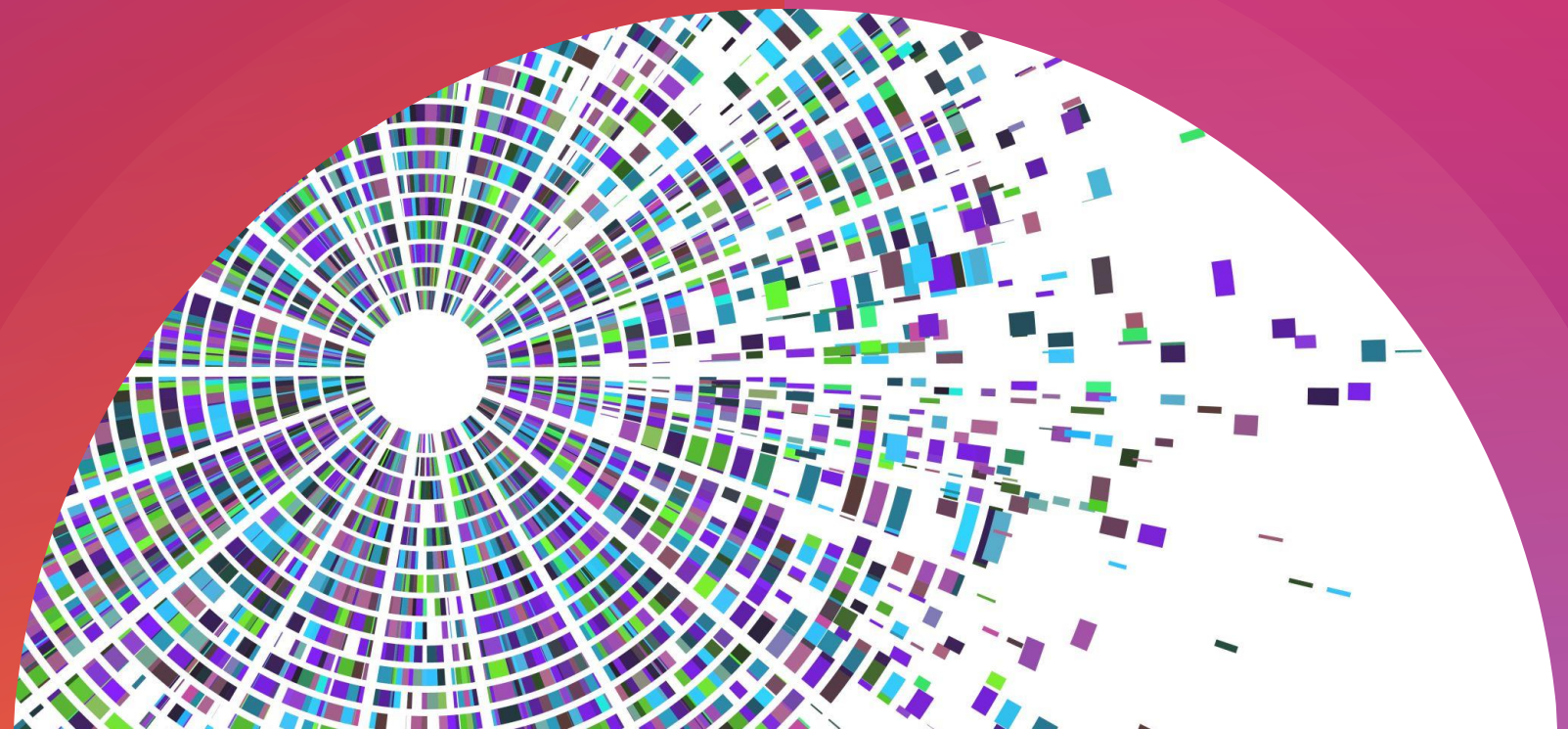
EXAMPLE

- Input: personal data; Output: Y/N Credit
- Dataset: historical data of 500,000 individuals and if they defaulted
- 400,000 to train, 100,000 to test.
- After training, the model predicts 87% of individuals in the testing set correctly.

These models are built from and evaluated with **correlation**

CORRELATION IS NOT CAUSATION!

WHAT MIGHT BE MITIGATING FACTORS IN THE EXAMPLE?



THE DATA WE START WITH MATTERS

- Tay was an AI twitter bot made by Microsoft
- It was meant to learn through interactions with humans
- A troll campaign led it to repeat hate speech
- Trolls used the repeat and Q/A features to train Tay to absorb offensive content into its dataset
- Through natural language processing, the bot regurgitated offensive content with some variation added in

THE DATA IS THE LEARNING

YOU CAN ONLY LEARN INFORMATION ONE CAN INFER FROM YOUR DATASET.

ABSTRACT OPEN QUESTIONS

- How can we thoroughly vet these algorithms and understand the patterns they find?
- How can we standardize the process of building these models to ensure some kind of optimality?
- How do we ensure our model works equally well on all kinds of data without bias?
- How do we build models that transfer to new situations?
- How do we program context?

WHERE NEURAL NETWORKS ARE BEING USED

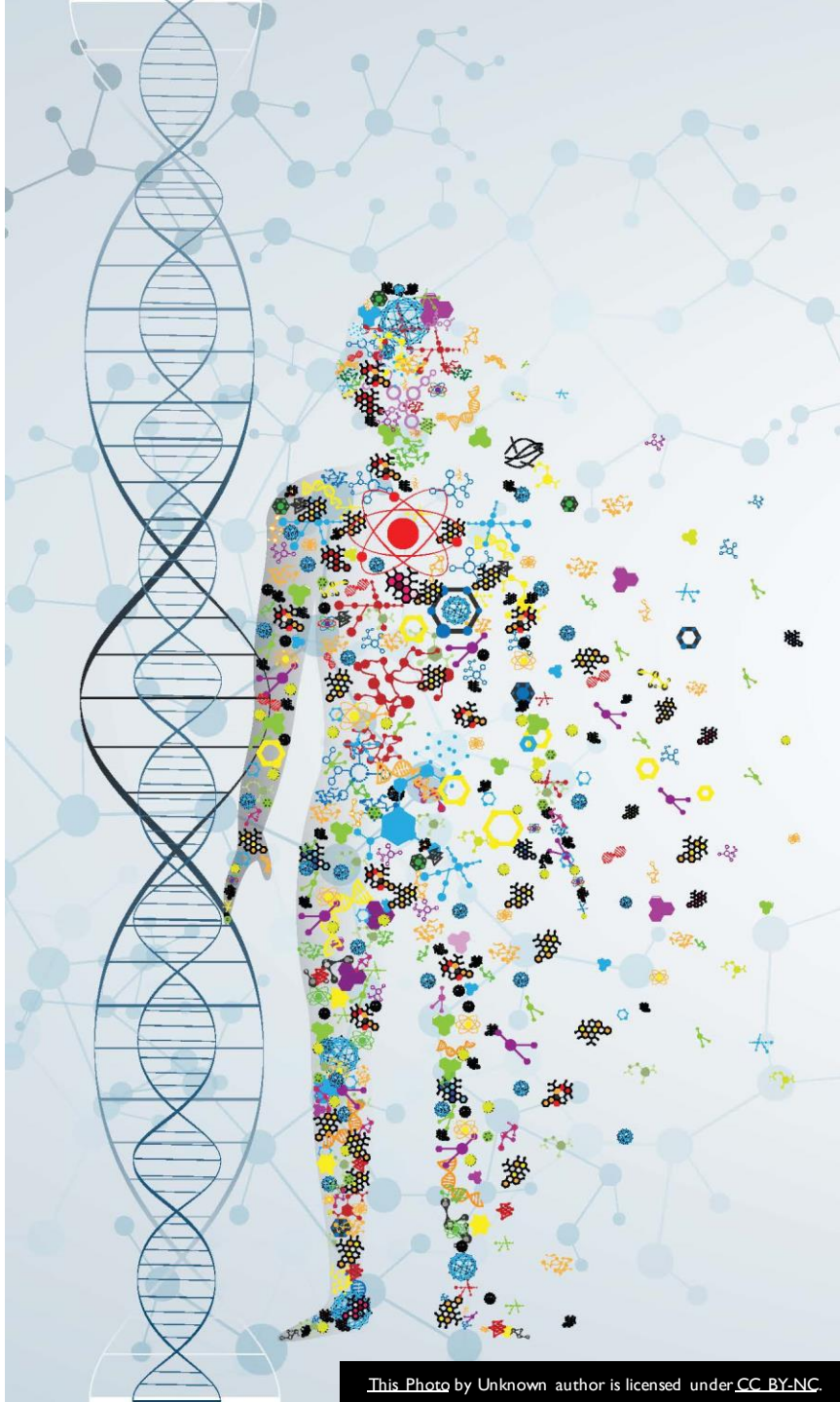
SOME APPLICATIONS

AI IN HEALTHCARE

Abnormality detection

Improved diagnosis

Personalized medicine



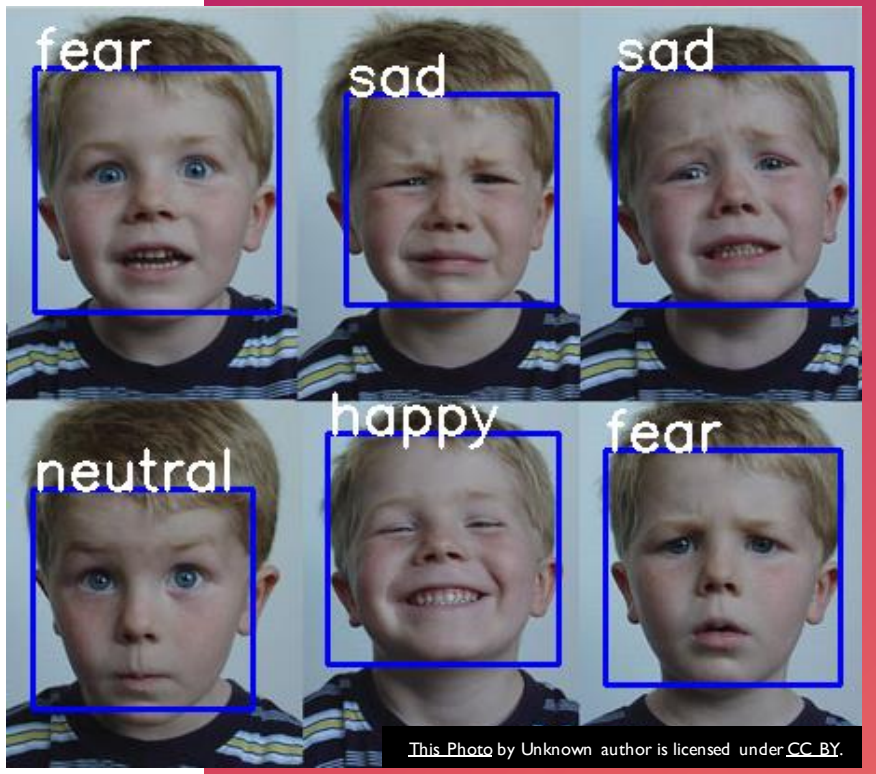
LAW ENFORCEMENT

Sentencing

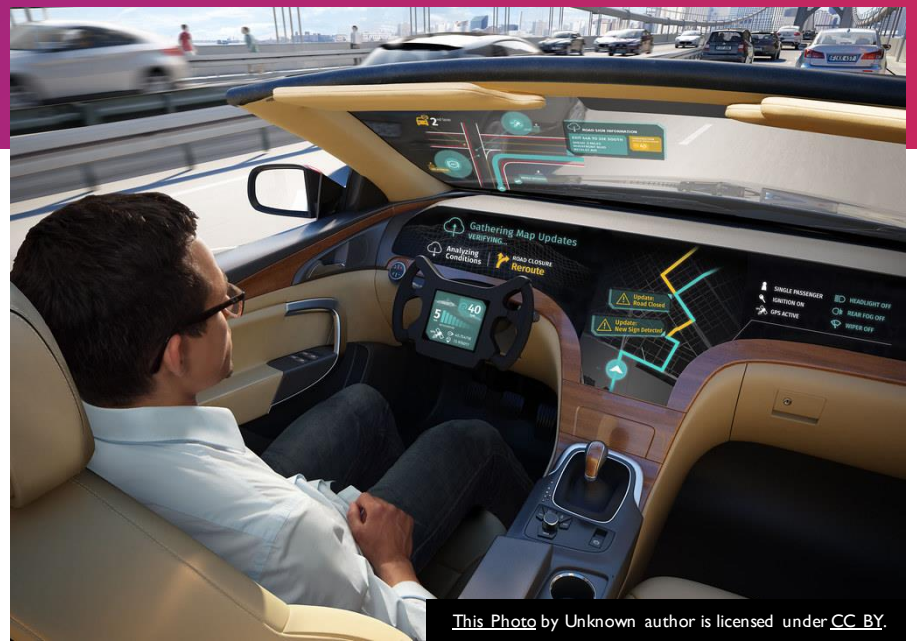
Tracking fugitives on CCTV systems

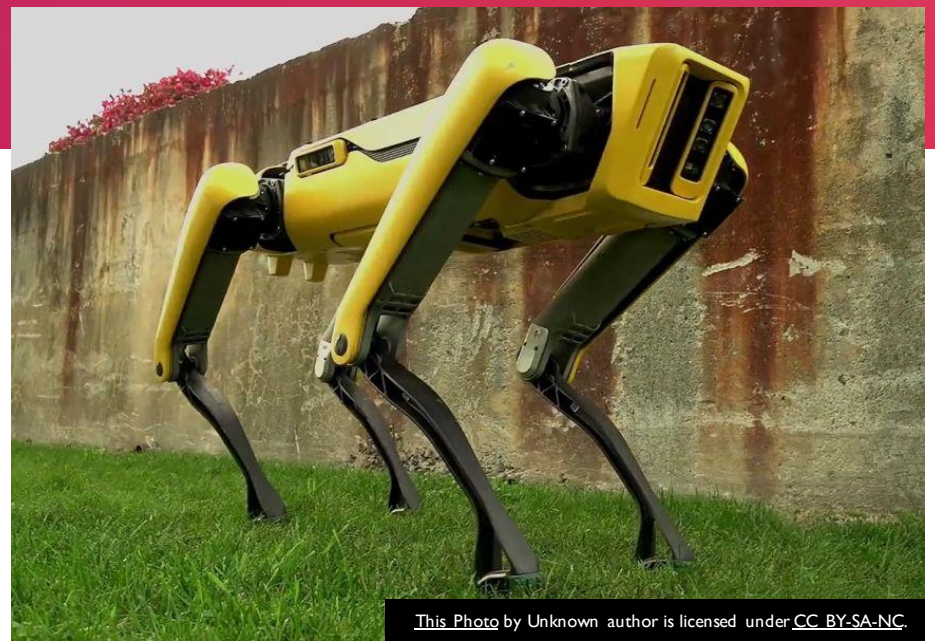
Emotion detection (interrogation)

Drunk driving detection



ROBOT VISION

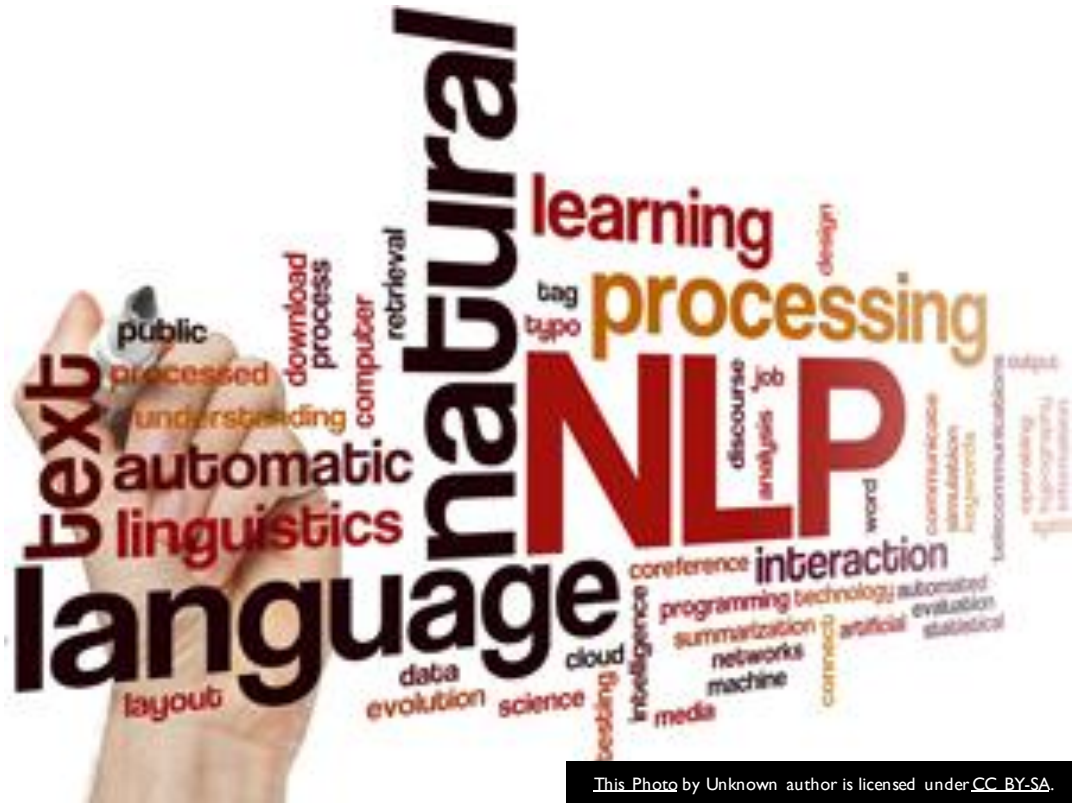




- Self-driving cars are a reality now
- Robots are watching for undesirable behavior in Singapore

NATURAL LANGUAGE PROCESSING

- Digital Virtual Assistants
- Text analysis
  - Author recognition
  - Tweet analysis
  - Text summarization
  - Translation
  - Chatbots and phone bots



GENERATIVE NETWORKS

- AI can generate content
  - Images
  - Videos
  - Music
- AI can also modify content
  - Add/ reduce noise
  - Increase/ decrease sharpness
  - Age/ deage

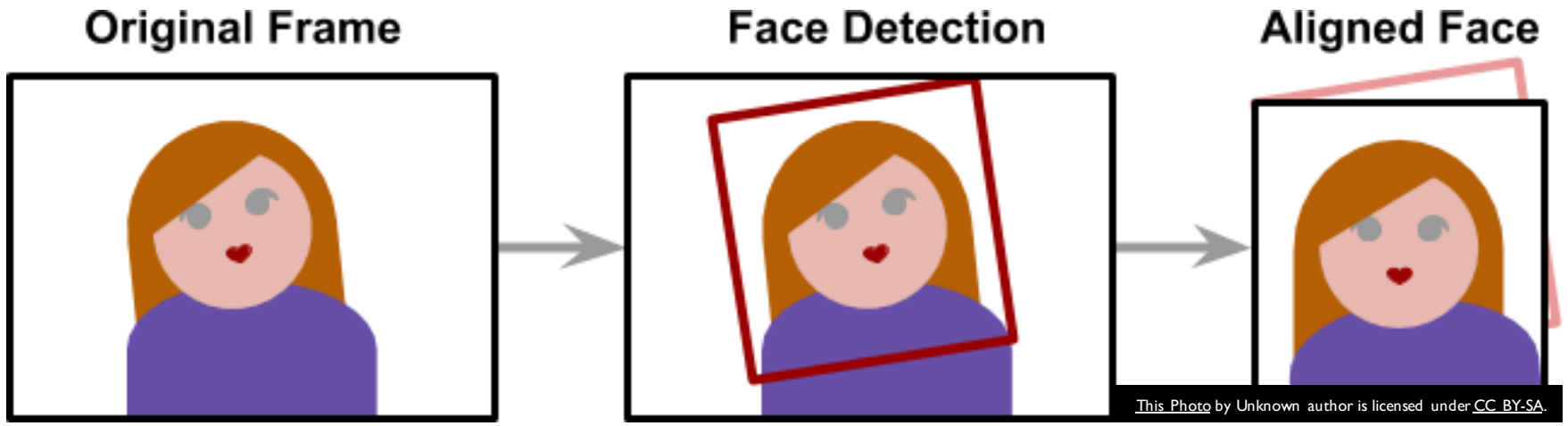
Image: R.A. Gonsalves

Image: Disney

Image: NVIDIA Research

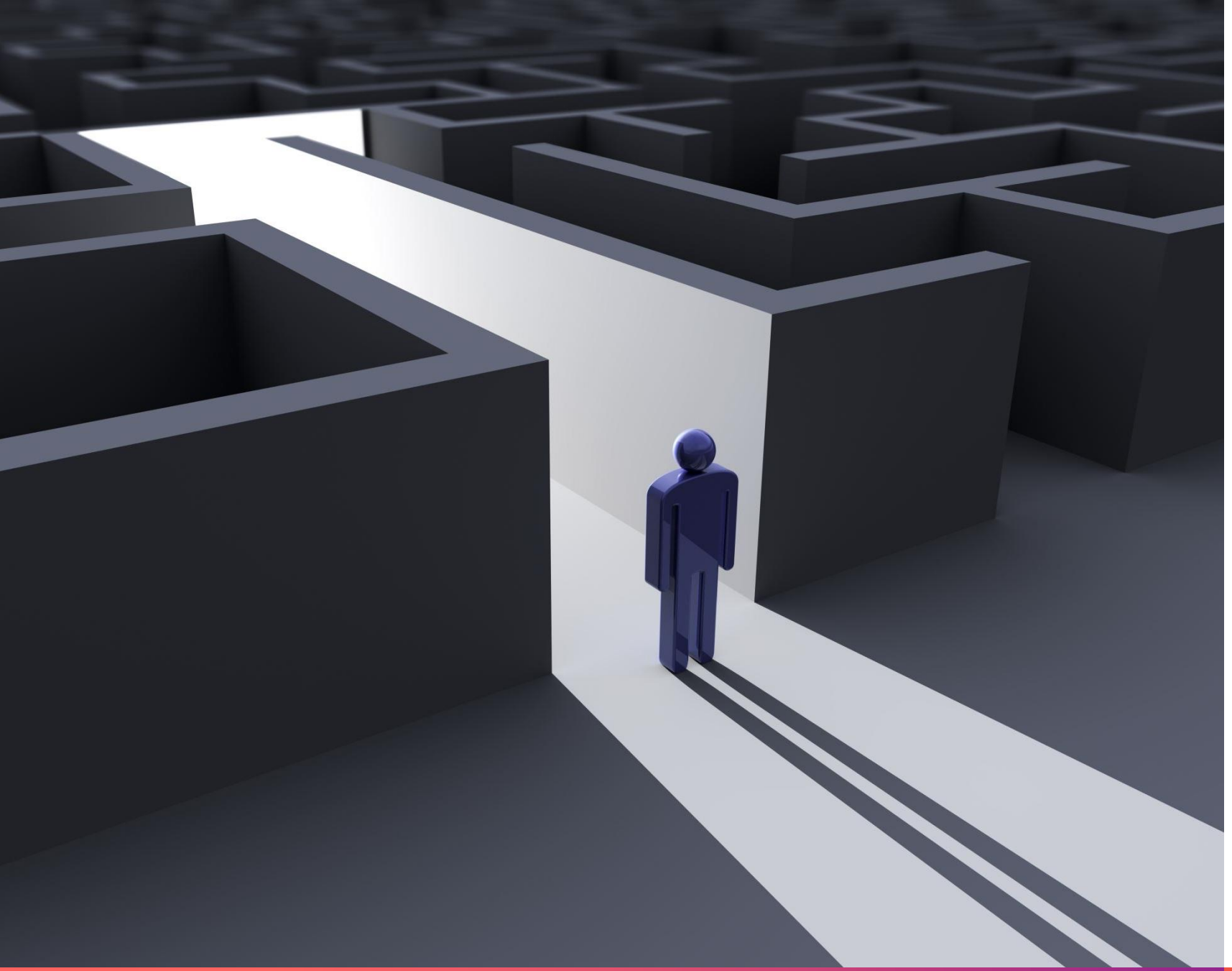
SOME CONCERNS OF RESEARCHERS AND PROGRAMMERS

ETHICAL CONCERNS



- Deepfakes
- Databases of faces (corporate and governmental)
- Profiles of you for ad targeting

ISSUES OF CONSENT



POOR MECHANISMS

- Fallacious belief in biological determinism
  - Criminality from face
- Cultural ignorance; a belief in universal qualities without supporting evidence
  - Emotion detection
- No evidence that inputs relate to outputs
  - Pregnancy and heartbeat
- A belief that historical data is accurate when evidence exists to the contrary
  - Automated sentencing

INTENSIFYING ESTABLISHED PREJUDICES

- Colorism
- Racism
- Sexism
- Western-centrism
- We are building AI on top of already existing data with these problems!

AI IS NOT "OBJECTIVE"

- AI cannot be objective; it relies on datasets that are often biased.
- We need researchers competent in societal prejudice in order to mitigate its effects in AI
- Remember, AI happens on a much larger scale than any one human!

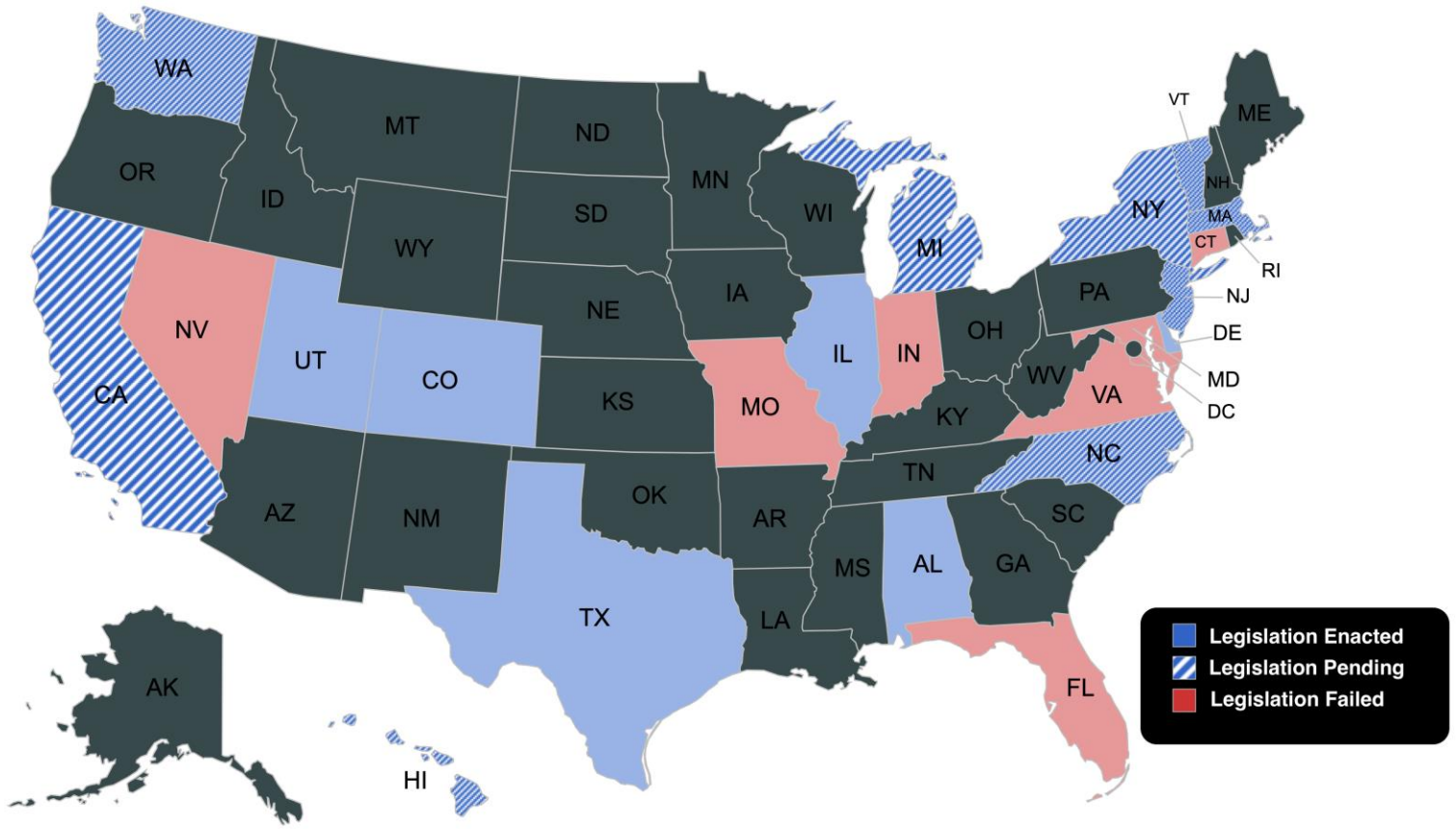
Skin cancer

AI skin cancer diagnoses risk being less accurate for dark skin – study

Studies suggest image recognition technology can classify skin cancers as successfully as humans. Posed by model. Photograph: ChesiireCat/Getty Images/iStockphoto

Nicola Davis *Science correspondent*

STATES WHERE SOME AI LAWS EXIST



Created with mapchart.net

THANK YOU!

- Thanks to the Oppenheimer Center for Entrepreneurship and Innovation for organizing and inviting us to speak.
- Additional resources can be found in the QR code.
- Questions?