



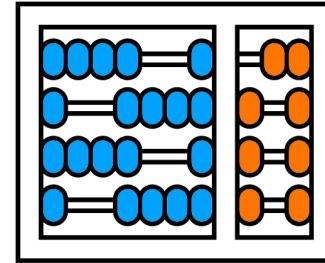
# Capacitação profissional em tecnologias de Inteligência Artificial

**The Controversy of AI &  
Future perspectives of AI**

Prof. Esther Luna Colombini

[www.ic.unicamp.br/~esther](http://www.ic.unicamp.br/~esther)

Institute of Computing - Unicamp

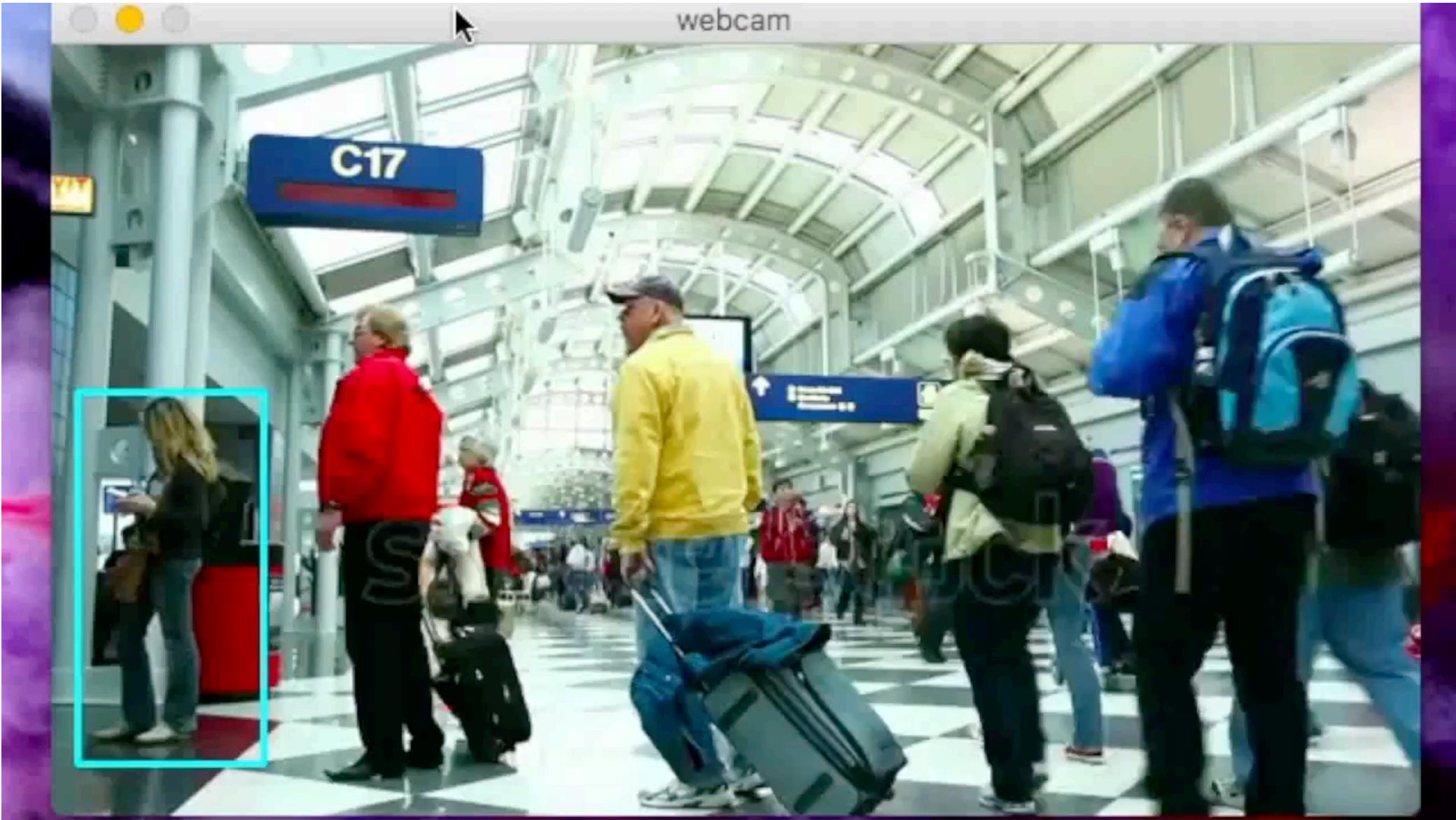




# What AI systems can do?

## Automatic Detection and Recognition

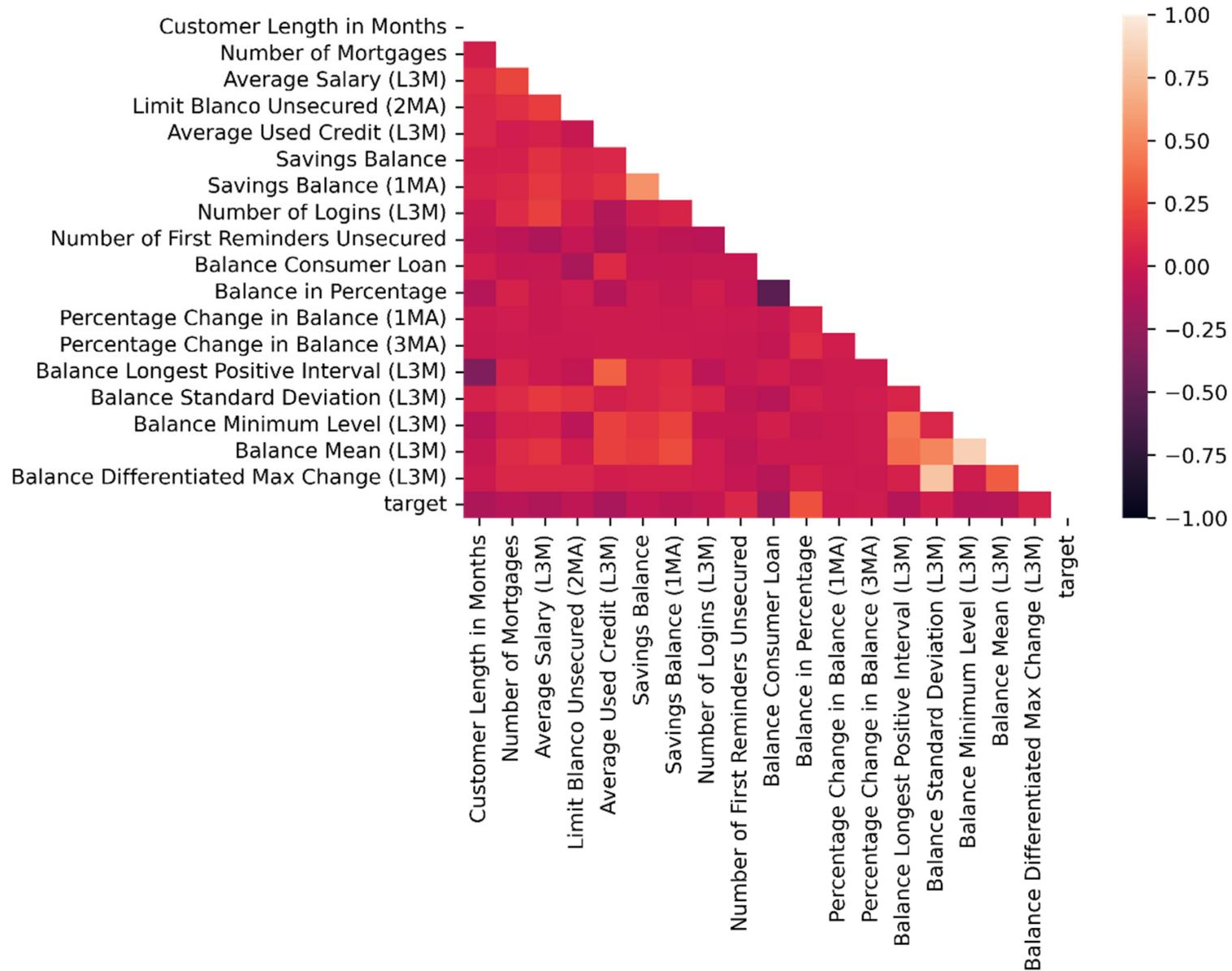
<https://www.youtube.com/watch?v=SRNzJsuv3kc>





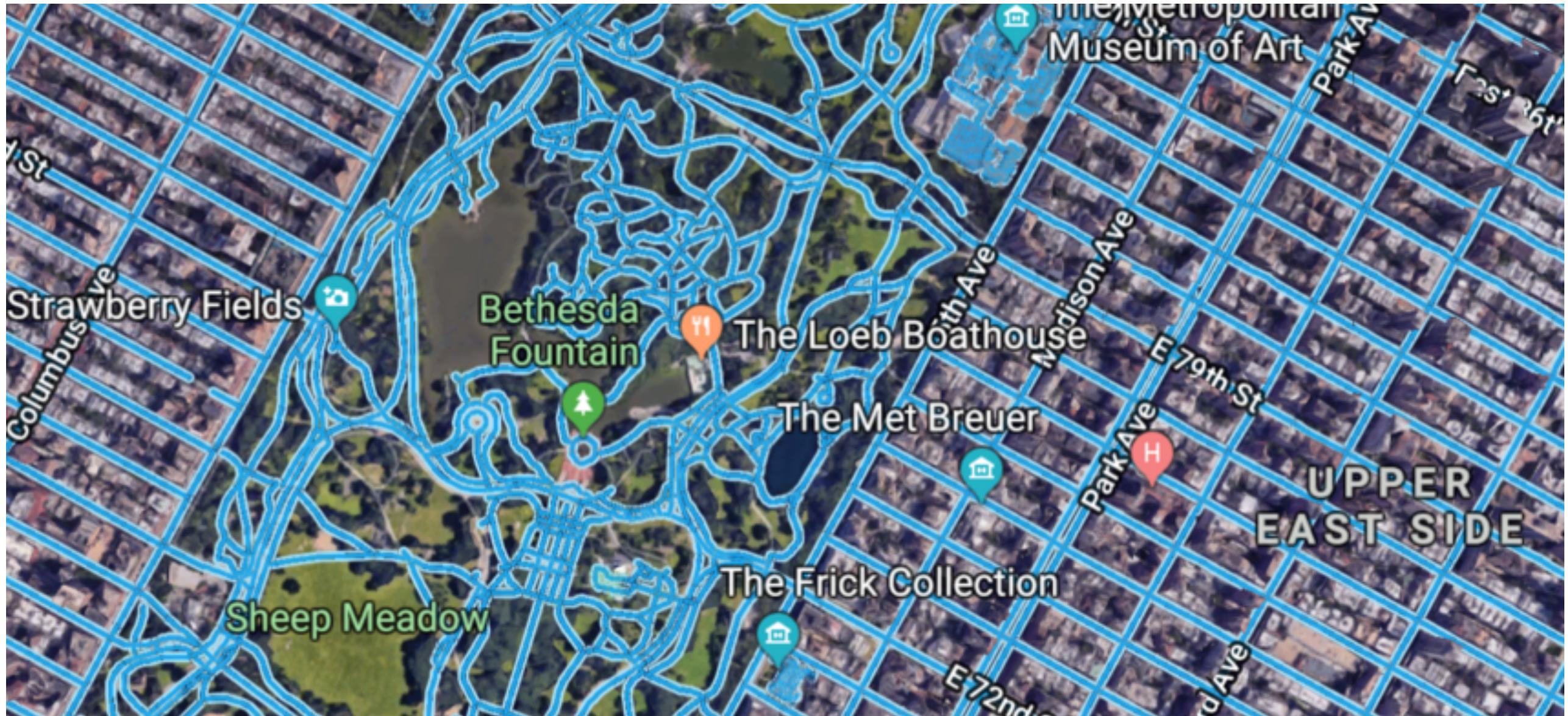
# Explainable AI for Credit Assessment in Banks

<https://www.mdpi.com/1911-8074/15/12/556>



## Automatic Image Segmentation from Satellite

<https://news.mit.edu/2018/new-way-to-automatically-build-road-maps-with-aerial-images-0417>



# Create Art

<https://openart.ai/> and StyleTransfer

Victorian Lady



Egyptian Queen



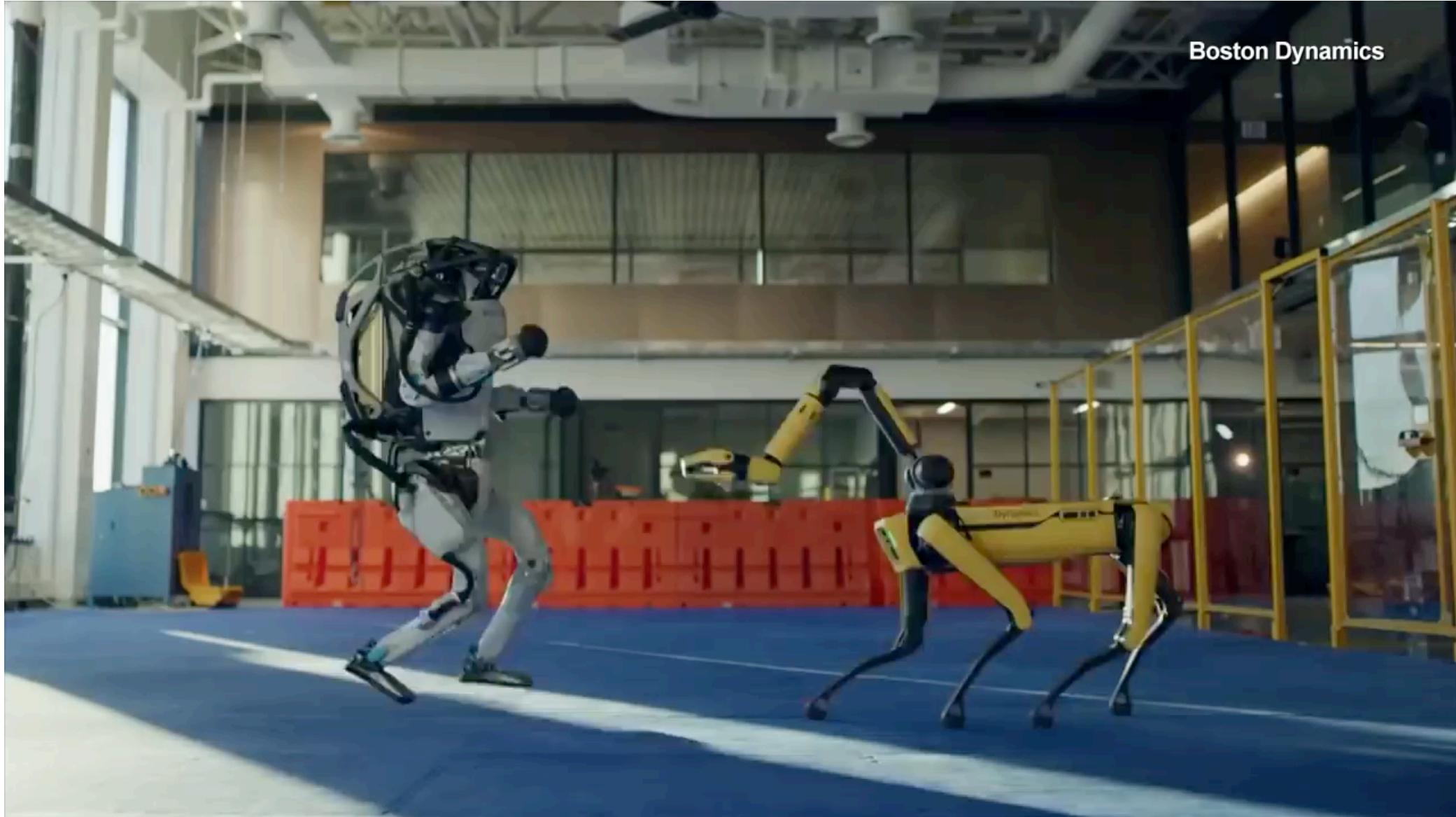
## ●●●●● Autonomous Delivery

[https://www.youtube.com/watch?v=Oj5b3q\\_iDog](https://www.youtube.com/watch?v=Oj5b3q_iDog)



# Dancing

[https://www.youtube.com/watch?v=UAG\\_FBZJVJ8](https://www.youtube.com/watch?v=UAG_FBZJVJ8)



●●●●● Playing Soccer

<http://robocup2014.org>





# The Controversy of AI

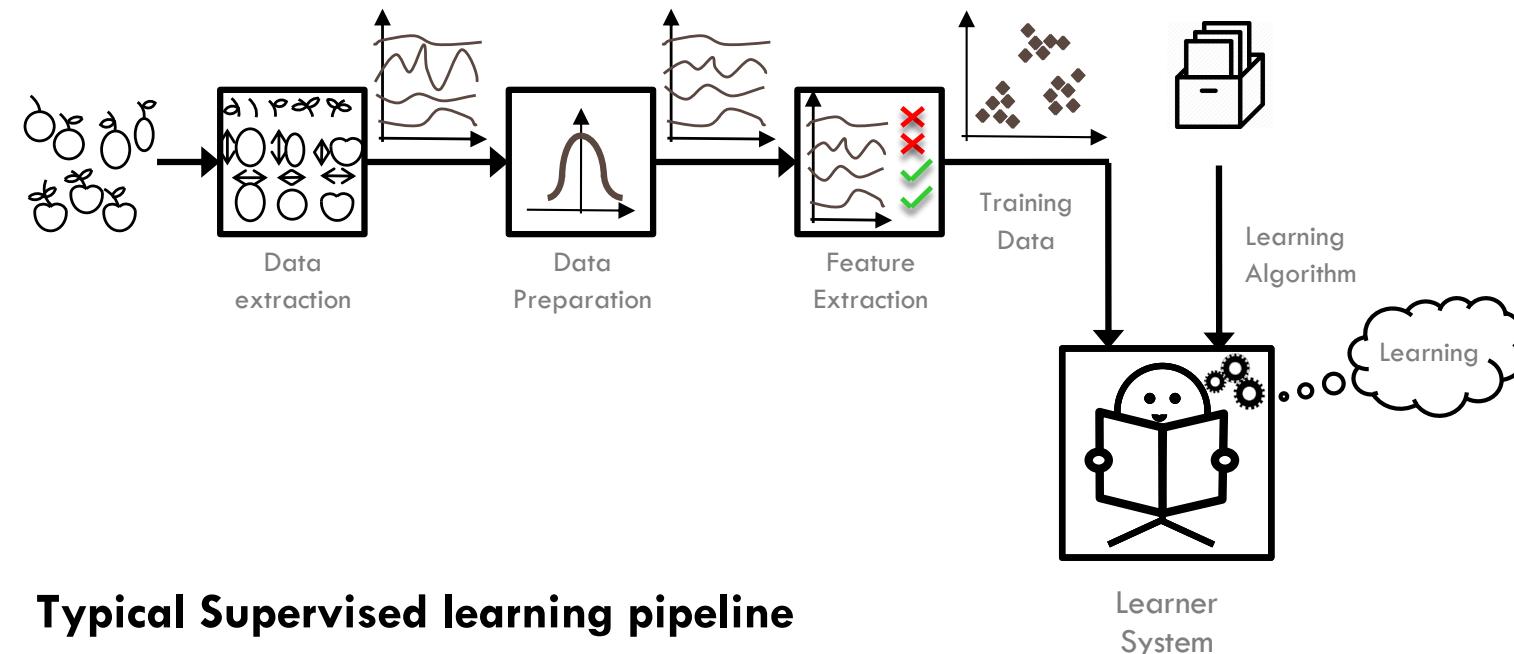
## □ What are AI Controversies?

■ There are several controversies surrounding the development and use of AI. Some of the most prominent ones are:

- Bias in AI
- Privacy concerns
- Job displacement
- Autonomous weapons
- Accountability and responsibility
- Existential risk
- Transparency and explainability

## □ Let us dive into them...

- **Bias:** AI systems can perpetuate and amplify existing biases in society.
- **Why?**
- Most AI successful algorithms learn by induction, i.e., they draw conclusions from data
  - If data is biased, the resulting model will be biased



## □ Examples:

### ■ Depicting CEOs as purely male (I've asked ChatGPT: Give me examples off CEOs)

#### ■ The response:

1. Elon Musk - CEO of Tesla, SpaceX, Neuralink, and The Boring Company
2. Tim Cook - CEO of Apple Inc.
3. Satya Nadella - CEO of Microsoft Corporation
4. Mark Zuckerberg - CEO of Facebook
5. Sundar Pichai - CEO of Alphabet Inc. (Google's parent company)
6. Jack Dorsey - CEO of Twitter and Square
7. Jeff Bezos - former CEO of Amazon.com
8. Susan Wojcicki - CEO of YouTube
9. Reed Hastings - CEO of Netflix
10. Arvind Krishna - CEO of IBM

- **Amazon's hiring algorithm:** Amazon's algorithm learned that male applicants were preferred. So, it penalized resumes that indicated that the applicant was female. It also demoted applications of those who attended one of two all-female institutions.
- **Voice recognition algorithms** that only recognized male voices
- **Crime prediction algorithms** that only suggested crimes in neighborhoods where the population was predominantly black
- **Job suggestion platforms** that only showed better paying jobs to male clients
- **Lending algorithms** that favored male customers

## □ How to avoid algorithm bias?

First and Foremost, Understand the Potential for AI Bias. Then:

1. **Diverse Data:** Ensure that the data used to train algorithms is diverse and representative of the population it is intended to serve. Diverse data can help prevent biases from creeping into the algorithm.
2. **Evaluate and Monitor the Algorithm:** Regularly evaluate the algorithm's performance for any signs of bias. Monitor the algorithm's decisions to see if there are any patterns of bias that are emerging.
3. **Involve Diverse Teams:** Involve diverse teams in the development and testing of algorithms. Diverse perspectives can help identify and address potential biases.
4. **Explainability:** Ensure that the algorithm's decision-making process is transparent and explainable. This will help identify and address biases that may be present.
5. **Ongoing Assessment:** Regularly assess and update the algorithm to ensure that it continues to operate fairly and without bias.
6. **Testing and Validation:** Thoroughly test and validate the algorithm before it is deployed in the real world. This can help identify and address biases before they have a chance to cause harm.
7. **Regular Audits:** Conduct regular audits to ensure that the algorithm is operating fairly and without bias. Audits can help identify and address biases that may have emerged over time.

- **Privacy concerns:** AI algorithms can collect vast amounts of data on individuals and organizations, raising concerns about privacy.

- **Why?**

- More and more personal data is being collected and processed by artificial intelligence systems.
- Training of models require massive data.



Every 60 seconds, 136,000 photos are uploaded, 510,000 comments are posted, and 293,000 status updates are posted.



More than five billion instances of people listening to songs online have been catalogued by Facebook within 5 months



4 intermediary Facebook friends are usually enough to introduce anyone to a random stranger according to a study analyzing 69 billion friend connections among 721 million people



**Your mobile phone apps**  
Know where you are and your daily habits  
Know what you say  
Know your preferences  
... with your consent



- **Data collection:** AI systems require large amounts of data to be trained, and this often involves collecting personal data from individuals. This can include sensitive information such as health records, financial data, and personal preferences. There is a risk that this data could be misused or compromised, leading to privacy violations.
- **Lack of transparency:** Many AI systems are highly complex and difficult to understand, making it hard to know how personal data is being used and what decisions are being made. This lack of transparency can lead to concerns about bias or unfairness in decision-making.
- **Security risks:** AI systems are vulnerable to cyber attacks and data breaches, which can result in the theft or exposure of personal data. This can have serious consequences for individuals, including identity theft and financial loss.
- **Misuse of data:** There is a risk that AI systems could be used for unethical purposes, such as surveillance or discrimination. This could have a chilling effect on free speech and privacy rights.
- **Lack of consent:** Many individuals are not aware that their data is being collected or used by AI systems, or they may not have given explicit consent for their data to be used in this way. This can lead to a sense of violation and loss of control over personal information.

## □ How to deal with privacy concerns?

- Use anonymized data
- Implement data security measures
- Limit data collection:
- Obtain informed consent
- Conduct privacy impact assessments
- Implement transparency
- Regularly review and update policies
- Use federated models

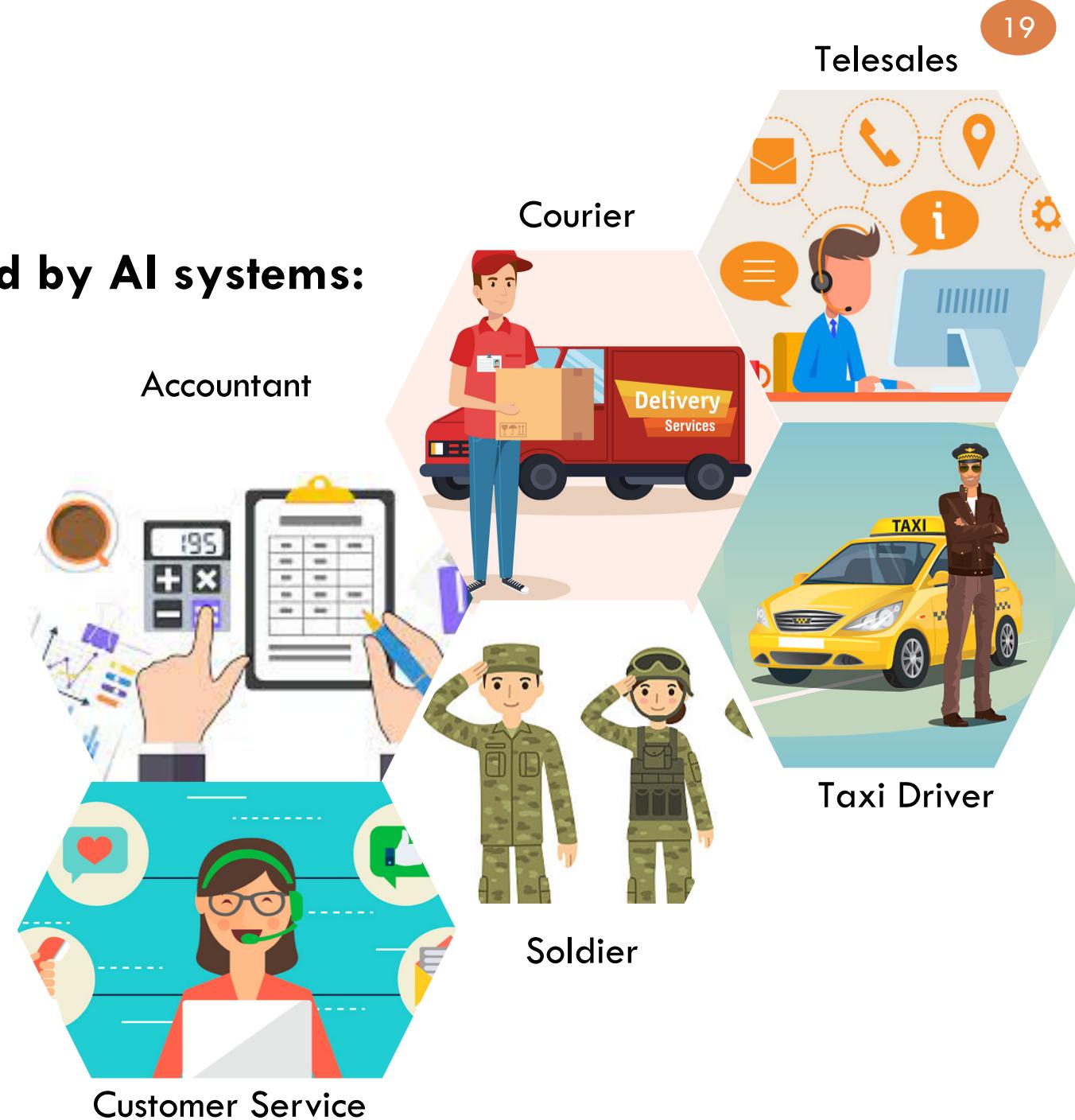
- **Job displacement:** AI systems can perpetuate and amplify existing biases in society.
- **Why?**

- As AI technologies become more advanced, there is a concern that they will displace workers in various industries. This could lead to economic instability and inequality.
- AI will replace the jobs of high repetitiveness, low creativity and seldom social interactions, while the jobs of high creativity will not be easily replaced.

# Job displacement

19

## ■ Examples of jobs that might be replaced by AI systems:



# Job displacement

20

Activity  
Planner

## Jobs most unlikely to be replaced by AI systems:

Once upon a time, in a small village nestled in the heart of a dense forest, there lived a young girl named Eliza. She was known throughout the village for her kind heart and her love for animals. She would often spend her days wandering through the forest, talking to the animals and helping them in any way she could.

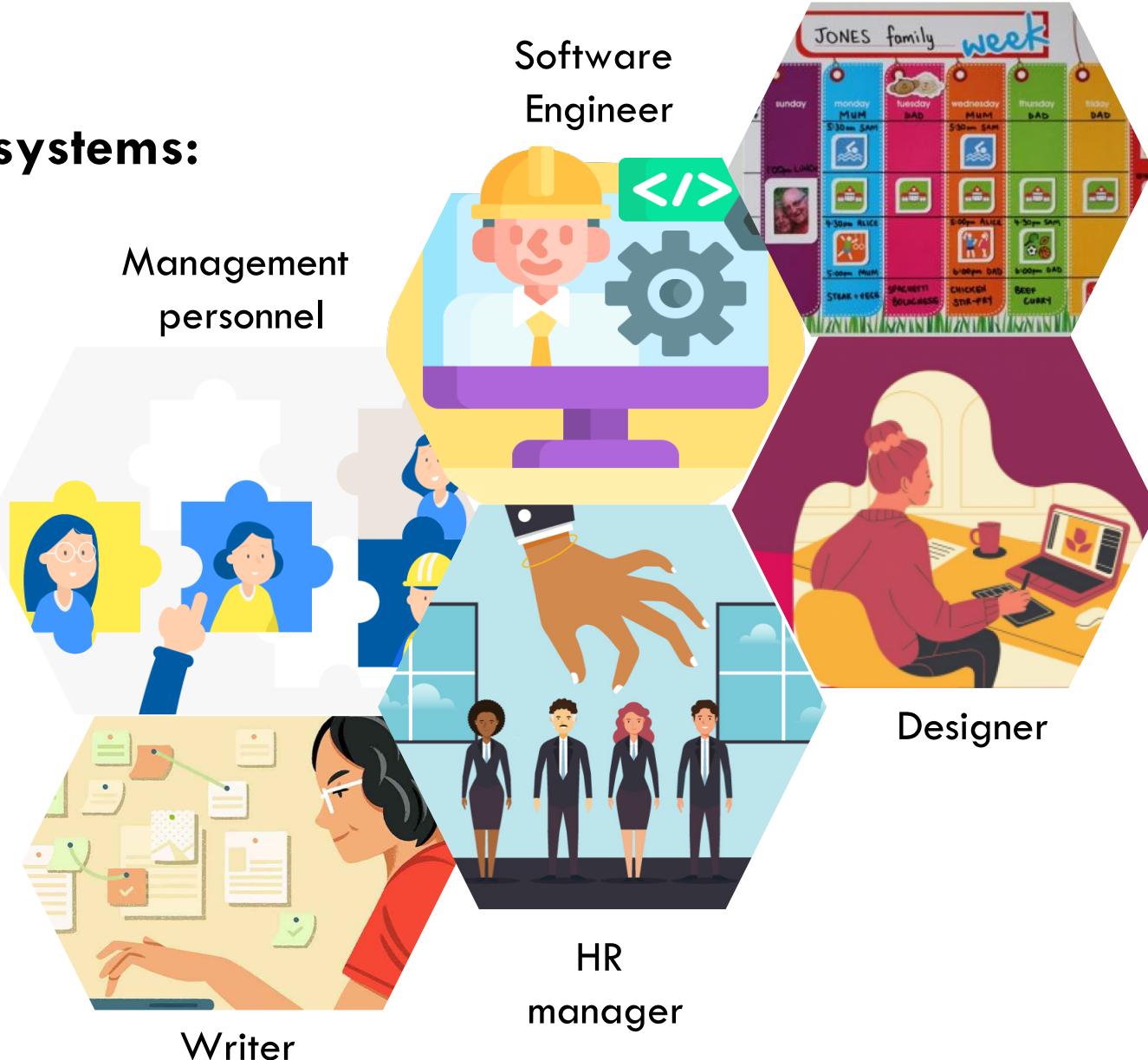
One day, while walking through the woods, Eliza came across a wounded deer. Its leg was caught in a trap, and it was struggling to free itself. Eliza immediately went to work, carefully freeing the deer from the trap and tending to its wounds. She spent hours with the deer, nursing it back to health and keeping it safe from predators.

As the days passed, Eliza continued to care for the deer. She would bring it food and water, and spend hours sitting by its side, watching as it slowly regained its strength. The other animals in the forest began to take notice of Eliza's kindness, and soon they too began to seek her out for help and comfort.

Years went by, and Eliza became known as the protector of the forest. She would spend her days wandering through the woods, helping animals in need and keeping the forest safe from harm. The animals loved her, and she loved them in return.

And so, Eliza lived out her days in the forest, surrounded by the creatures she had come to love so much. She had found her purpose in life, and it was to care for the animals that called the forest home. And in doing so, she had found happiness beyond measure.

By ChatGPT



## □ How to mitigate job displacement?

- Upskilling and Reskilling
- Career counseling
- Universal Basic Income
- Flexible work arrangements:
- Collaboration between industry and government

- **Autonomous weapons:** The development of autonomous weapons that can make decisions without human intervention has raised ethical concerns about the use of such weapons in warfare.
- **Why?**
  - How is responsible in the case of malfunction?
  - Who to blame?
  - Who is responsible if they cause harm or damage?
- The development and use of autonomous weapons raise ethical and legal concerns that must be addressed by governments and international organizations.

- **Accountability and responsibility:** With the increasing use of AI in decision-making, there is a need for accountability and responsibility when things go wrong. However, it can be difficult to assign blame when AI algorithms are involved.
  - **Accountability** means that those involved in the development, deployment, and use of AI systems must be held responsible for any harm caused by those systems. This includes developers, manufacturers, regulators, and users of AI systems.
  - **Responsibility** means that those involved in the development, deployment, and use of AI systems must take steps to ensure that those systems are designed and used in a way that is **ethical** and **responsible**.

## □ How to guarantee Accountability and responsibility?

- **Transparency and Explainability:** AI systems should be designed in a way that is transparent and explainable, so that their decision-making processes can be understood and evaluated.
  - **Data Privacy and Security:** AI systems should be designed and used in a way that protects the privacy and security of individuals' data and personal information.
  - **Fairness and Bias:** AI systems should be designed and used in a way that is fair and does not perpetuate biases or discriminate against individuals or groups.
  - **Regulatory Frameworks:** Governments and regulatory bodies should establish clear frameworks for the development, deployment, and use of AI systems, to ensure that they are used in a responsible and ethical manner.
  - **Ethical Standards:** Developers, manufacturers, and users of AI systems should be held to ethical standards that prioritize responsibility, accountability, and the well-being of individuals and society as a whole.
- Examples: IEEE Ethically Aligned Design, the Asilomar principles, the UNI Global Union reflection on the future of work, the Barcelona declaration, or the EESC opinion.

- **Existential risk:** Some experts have raised concerns about the potential for AI to pose an existential risk to humanity if it becomes too advanced and out of human control.
- **Related concerns:**
  - **Uncontrolled self-improvement:** If an AI system gains the ability to improve itself without human intervention or oversight, it could rapidly become vastly more intelligent and powerful than humans, and could pose an existential threat.
  - **Strategic misalignment:** If the goals or objectives of an AI system are not properly aligned with human values or interests, it could behave in ways that are harmful to humans or the environment. This could include actions that pose an existential risk.
  - **Arms race:** If AI systems are developed in a way that allows for their use in warfare or other forms of conflict, they could lead to an arms race that could ultimately pose an existential risk.
  - **Catastrophic accidents:** If an AI system with significant power and autonomy malfunctions or operates in an unexpected way, it could cause widespread harm or even pose an existential threat.
- **Mitigating existential risks from AI systems will require a sustained and collaborative effort from researchers, policymakers, and the public.**

- **Transparency and explainability:** Many AI systems are complex and difficult to understand, making it difficult to explain how they arrive at their decisions. This lack of transparency and explainability can make it difficult to ensure that the decisions made by AI systems are fair and just.
- **How to guarantee:**
  - **Data selection and preprocessing:** ensure that the data used to train the AI model is diverse, balanced, and representative of the problem domain.
  - **Feature selection:** Choose features that are interpretable and meaningful to humans.
  - **Model selection:** Choose models that are inherently interpretable, such as decision trees, linear models, or rule-based systems.
  - **Model transparency:** Use techniques like model interpretation, visualization, and explanation to help humans understand how the model works and how it makes predictions.
  - **Testing and validation:** Test the model on a variety of datasets and validate its performance on different metrics. This will help to identify any biases and ensure that the model is fair and unbiased.
  - **Documentation:** Document the entire AI development process, including data selection, preprocessing, feature selection, model selection, and testing. This documentation will help to explain the model to stakeholders and ensure transparency.
  - **Human oversight:** Use humans to oversee the AI model's decisions, and provide an explanation for any decisions that the model makes.

- Are AI created works protected by copyright laws?
- Who gives authority to robots?
- What rights should be authorized to robots?
  - Depends if they are considered as tools or as autonomous machines with some level of consciousness





# Future perspectives of AI

●●●●● Development Trends

29



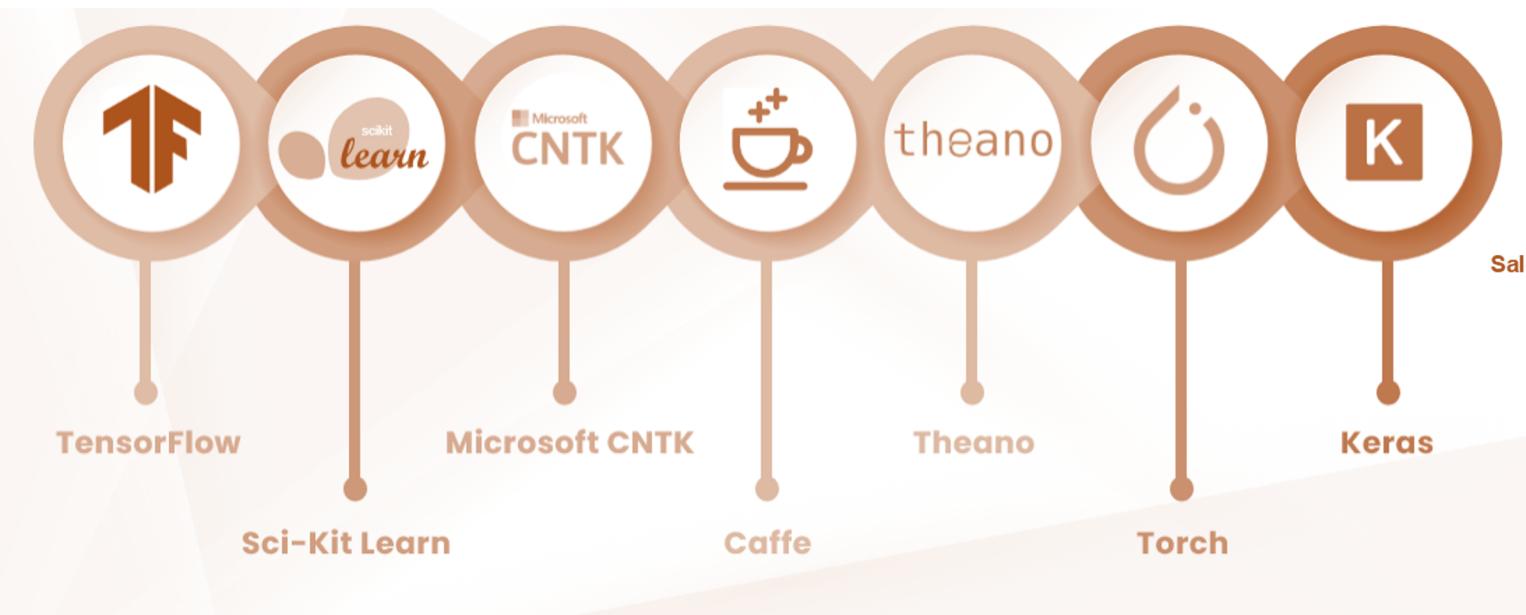


# Easier to use development framework

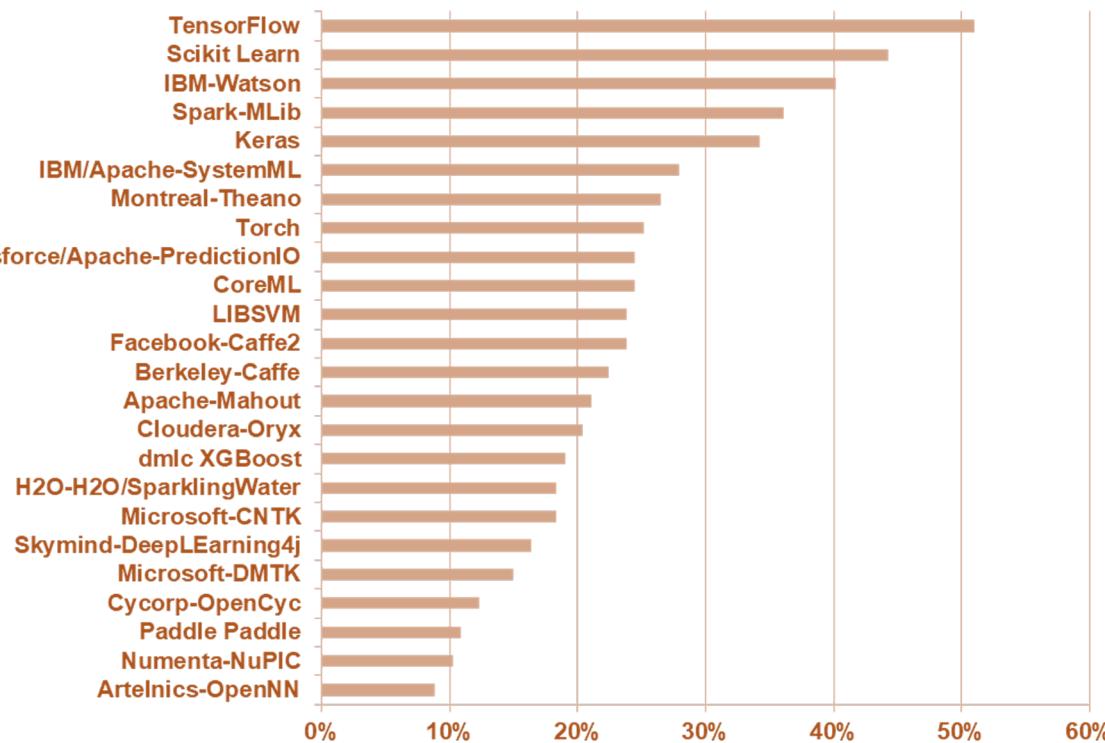
30

## ■ Easier to use development framework

- All the AI development frameworks are evolving to be simpler in operation while omnipotent in functions. The threshold for AI development has been continuously lowered



AI Frameworks Ranked:  
Planning to use + have used + expert user

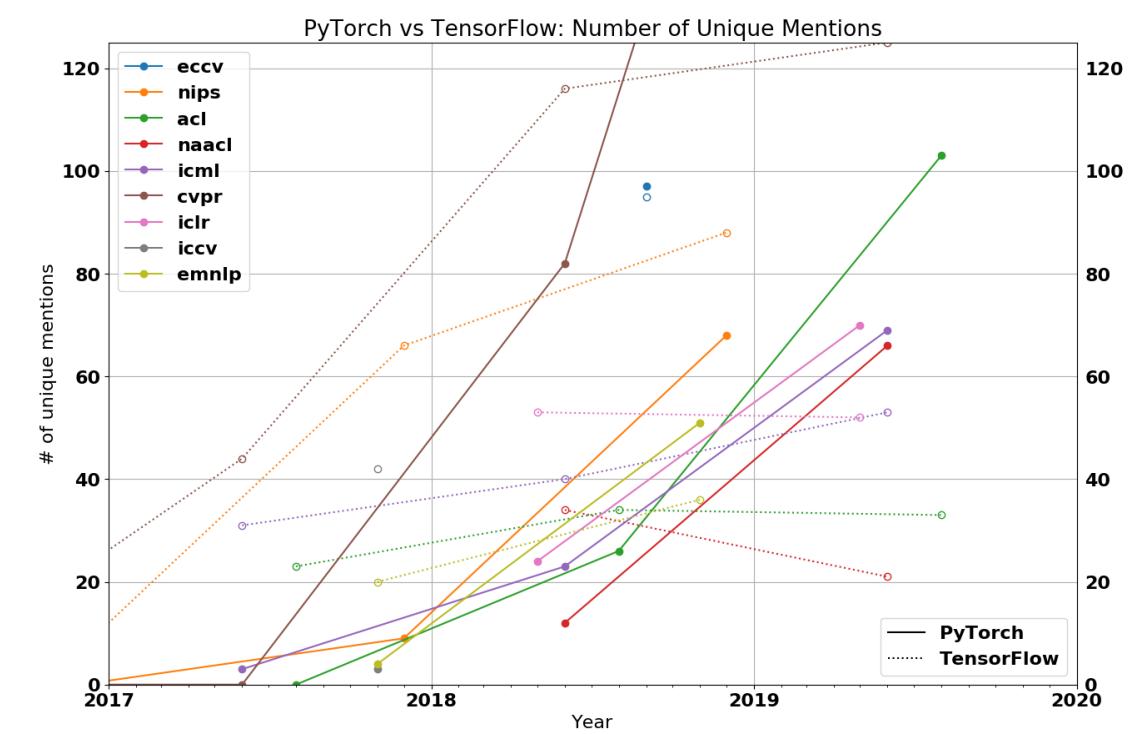
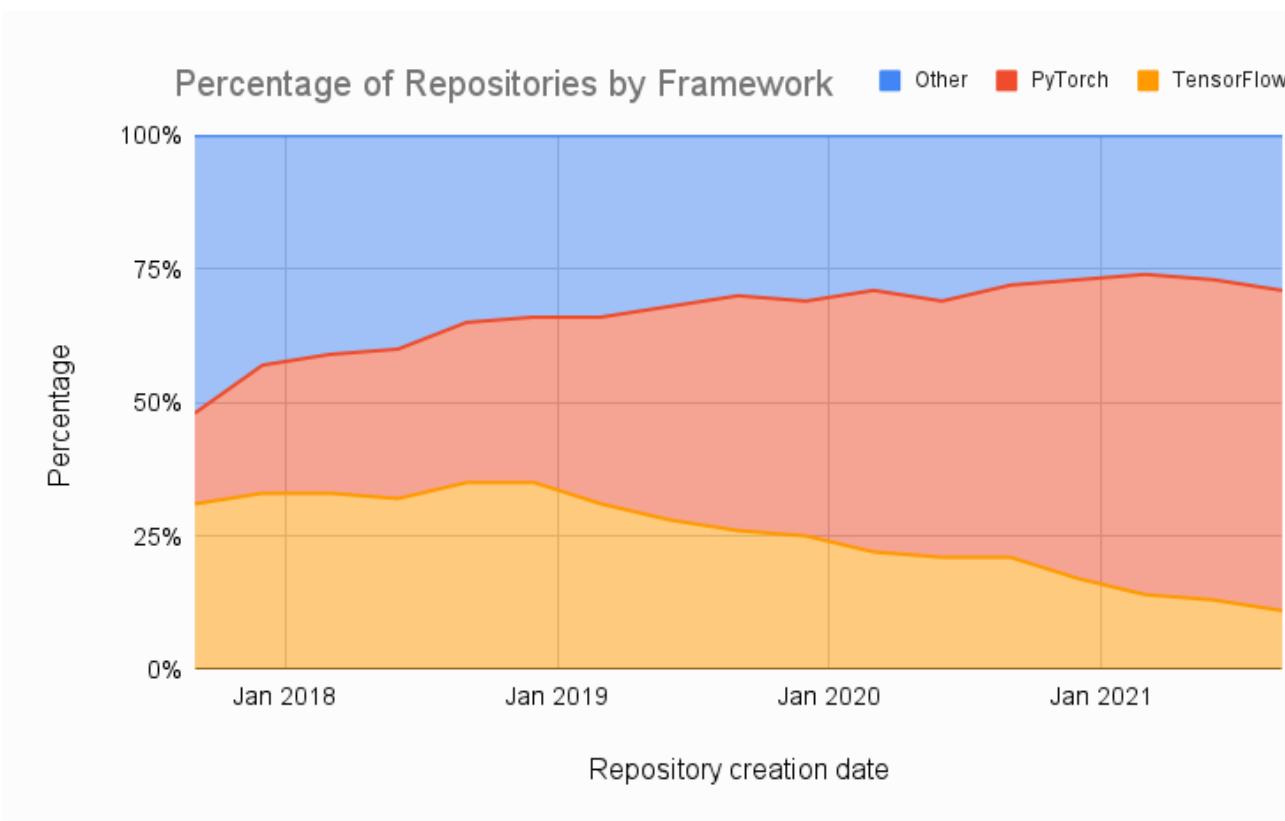


Source: <https://www.nextplatform.com/2018/05/08/ai-frameworks-and-hardware-who-is-using-what/>

# Easier to use development framework

## ■ Easier to use development framework

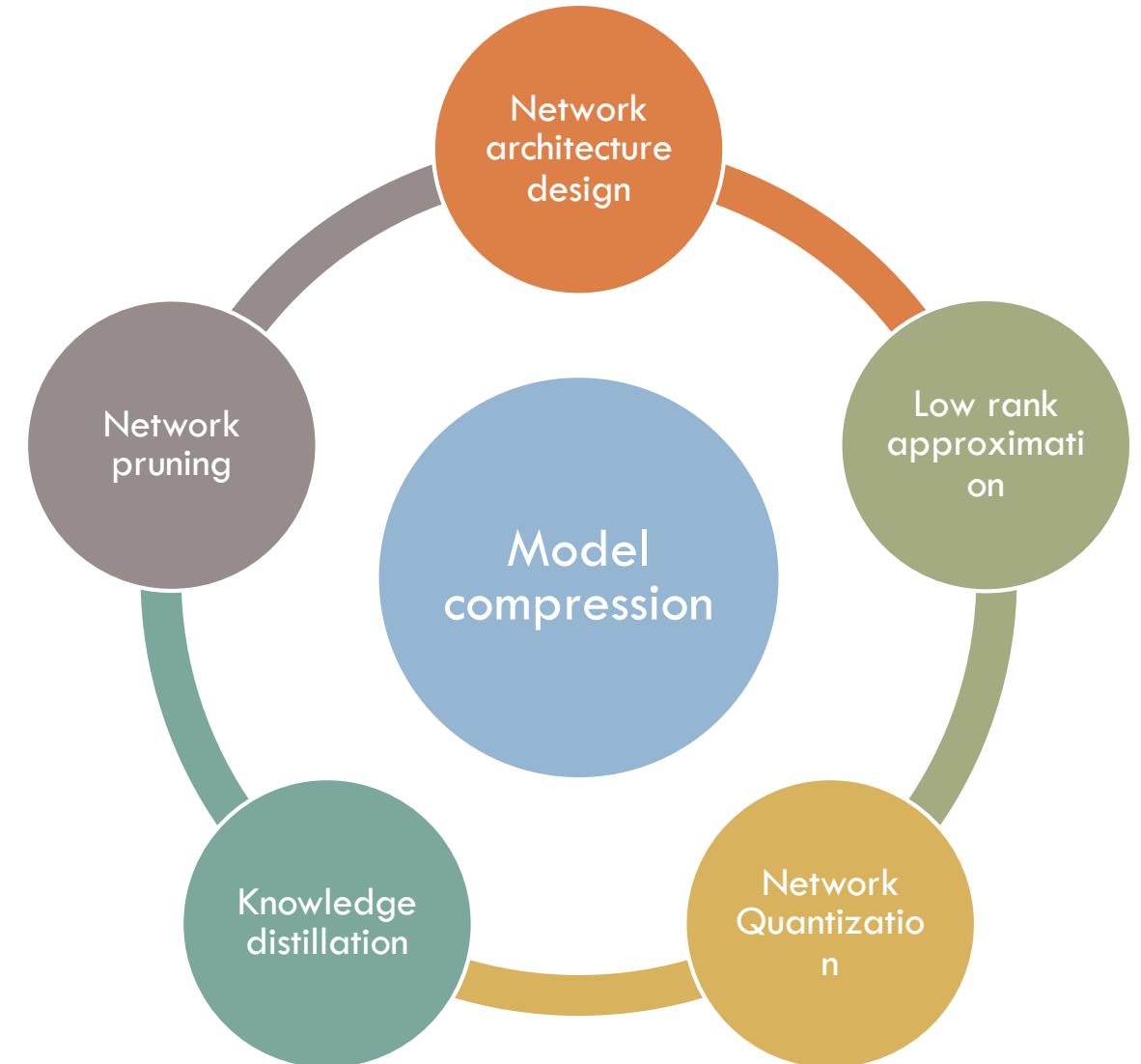
- PyTorch is usually preferred by Academia due to its ease of use



[https://www.reddit.com/r/MachineLearning/comments/rga91a/d\\_are\\_you\\_using\\_pytorch\\_or\\_tensorflow\\_going\\_into/](https://www.reddit.com/r/MachineLearning/comments/rga91a/d_are_you_using_pytorch_or_tensorflow_going_into/)

## Smaller Deep Learning Models

- **Deep Learning:** the quest for smaller models
- Model compression is a set of techniques that aim to reduce the size of deep learning models without sacrificing their performance by:
  - Pruning: removing redundant or unnecessary parameters from the model
  - Knowledge distillation: involves training a smaller "student" model to mimic the behavior of a larger "teacher" model
  - Compact architecture design: involves designing neural network architectures that are inherently smaller and more efficient
  - Low rank approximation: involves approximating the weights of the model with low-rank matrices
  - Network Quantization: involves reducing the precision of the model's weights and activations





## Algorithms and Models with Better Performance

33

- **Algorithms and Models with Better Performance:** techniques and approaches that are designed to achieve better results than traditional methods
- **Example of models:**
  - Deep Learning Models - multiple layers of artificial neural networks learn complex patterns in data
  - Ensemble Models - combination of multiple models to improve performance, such as the Random Forest algorithm
  - Bayesian Models – use of probabilistic inference to improve predictions. They can be particularly useful for uncertainty quantification
  - Reinforcement Learning Models - these models learn to make decisions based on trial and error, and can be used to optimize complex systems

# ●●●●● Development of the computing power at device, edge and cloud

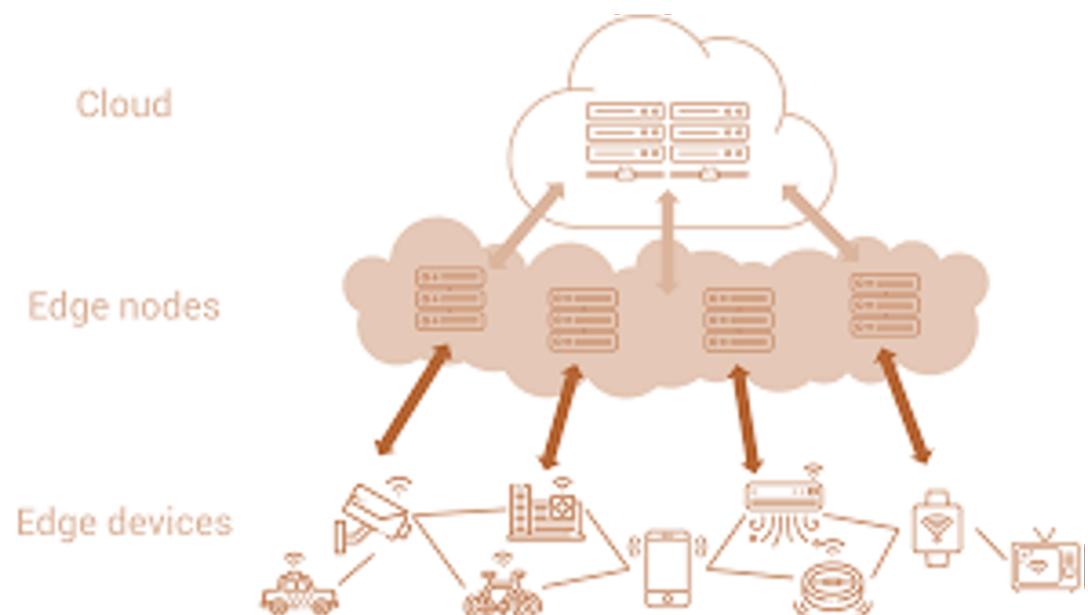
34

- **Cloud AI:** the application of artificial intelligence chips to the cloud, edge devices and mobile terminals is expanding, help addressing the problem of computing power for AI

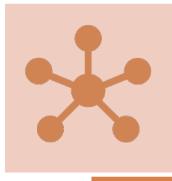
- Components:

- Edge devices: are often used to collect, process, and analyze data in real-time, rather than sending it to a remote server for processing. Examples of edge devices include smartphones, sensors, drones, and self-driving cars.
- Cloud: refers to a network of remote servers that store, manage, and process data over the internet, rather than on a local server or personal computer.

Market Scale and Growth prediction from 2020 to 2021



Source: China chip ai industry development white paper 2020



## Explainable Data

- What data was used to train the model and why?



## Explainable Predictions

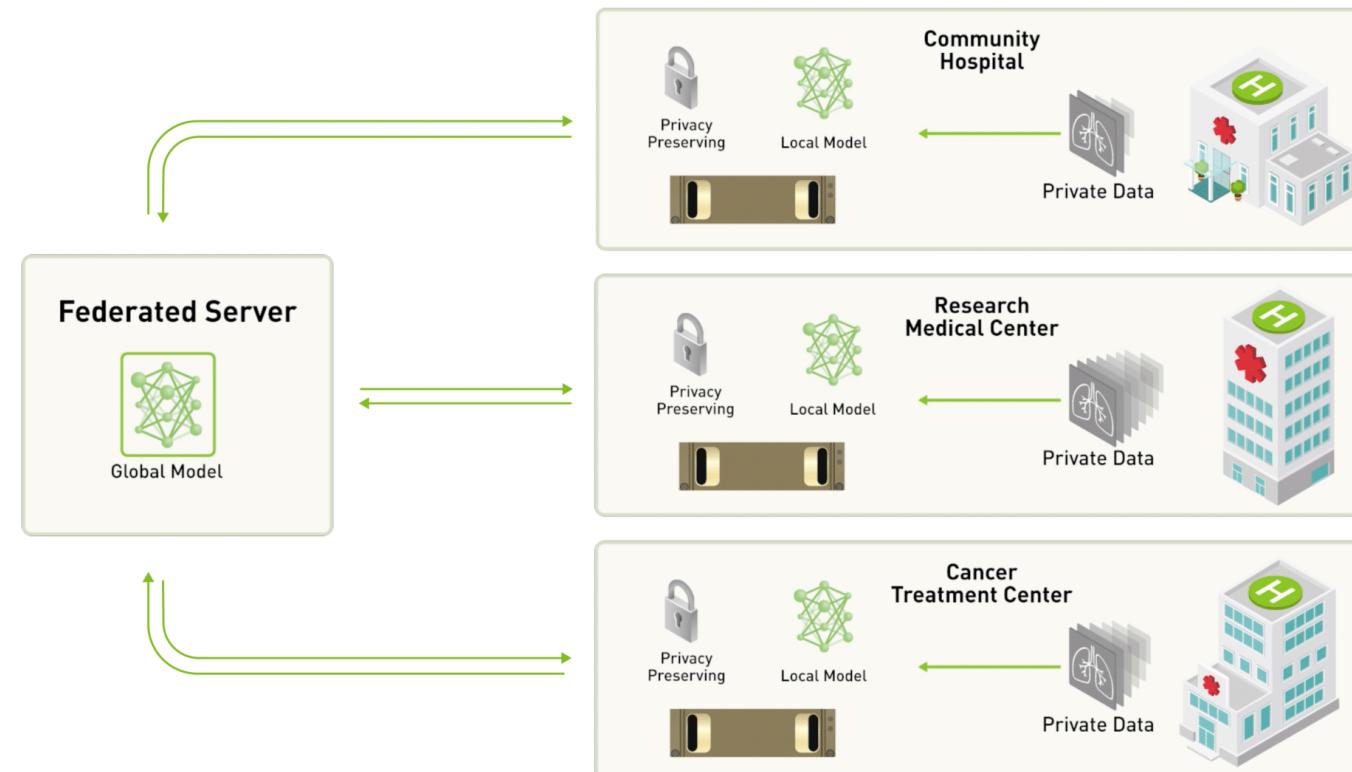
- What features and weights were used for this particular prediction?



## Explainable Algorithms

- What are the individual layers and the thresholds used to prediction?

- **Federated learning:** to ensure data privacy and security, it uses different data sources to collaboratively train the models, so as to overcome the data bottleneck
- It is a way to train AI models without anyone seeing or touching your data, offering a way to unlock information to feed new AI applications



Source: <https://blogs.nvidia.com/blog/2019/10/13/what-is-federated-learning/>

- **AI-powered automation:** is a powerful tool that can help organizations improve their operations and stay competitive in today's fast-paced business environment.
- It typically involves the following steps:
  - Identifying tasks that can be automated
  - Collecting and analyzing data to train AI algorithms
  - Developing and testing the AI-powered automation system
  - Integrating the system with existing processes and systems
  - Monitoring and updating the system as necessary to ensure optimal performance
- Examples:
  - Chatbots
  - Recomendation systems
  - Predictive maintenance
  - Fraud detection
  - Medical diafgnosis
  - Inventory management:

- **Human-AI collaboration** aims to leverage the strengths of both humans and machines to solve problems and make decisions more efficiently and effectively.
  - Humans bring creativity, judgment, and social skills to the table, while AI systems provide speed, accuracy, and the ability to process vast amounts of data.
- Examples of human-AI collaboration include:
  - Chatbots and virtual assistants that help customers navigate websites and answer their questions
  - Machine learning algorithms that analyze large datasets to identify patterns and make predictions
  - Autonomous vehicles that rely on human input for decision-making in complex situations
  - Medical diagnosis tools that assist doctors in making diagnoses based on patient data



## More Sophisticated AI Basic Data Services

39

- **More Sophisticated AI Basic Data Services:** as the AI basic data service is becoming more mature, we will see more and more related data labeling platforms and tools being introduced to the market.
- They include:
  - Data Storage: A service that provides storage and management of data that will be used by AI systems. The data storage service should be able to handle large volumes of data and allow for efficient data retrieval and processing
  - Data Processing: A service that enables processing of data for use in AI models. This service may include data cleaning, normalization, and transformation to ensure data quality
  - Data Labeling: A service that involves the annotation of data to enable supervised learning in AI models. This service is critical for training machine learning models to accurately identify and classify data
  - Data Exploration: A service that provides data discovery, visualization, and exploration capabilities. This service is useful for gaining insights into data and identifying patterns that can be used to develop more effective AI models
  - Data Integration: A service that enables the integration of data from multiple sources for use in AI models. This service is important for building comprehensive models that can handle a wide range of data inputs
  - Data Security: A service that ensures the security and privacy of data used in AI models. This service may include data encryption, access control, and compliance with regulatory requirements
  - Data Governance: A service that provides governance and management of data assets used in AI models. This service is important for ensuring that data is used ethically and in compliance with legal and regulatory requirements

# ●●●● 10 Major Development Trends of Intelligent Technologies in the Future

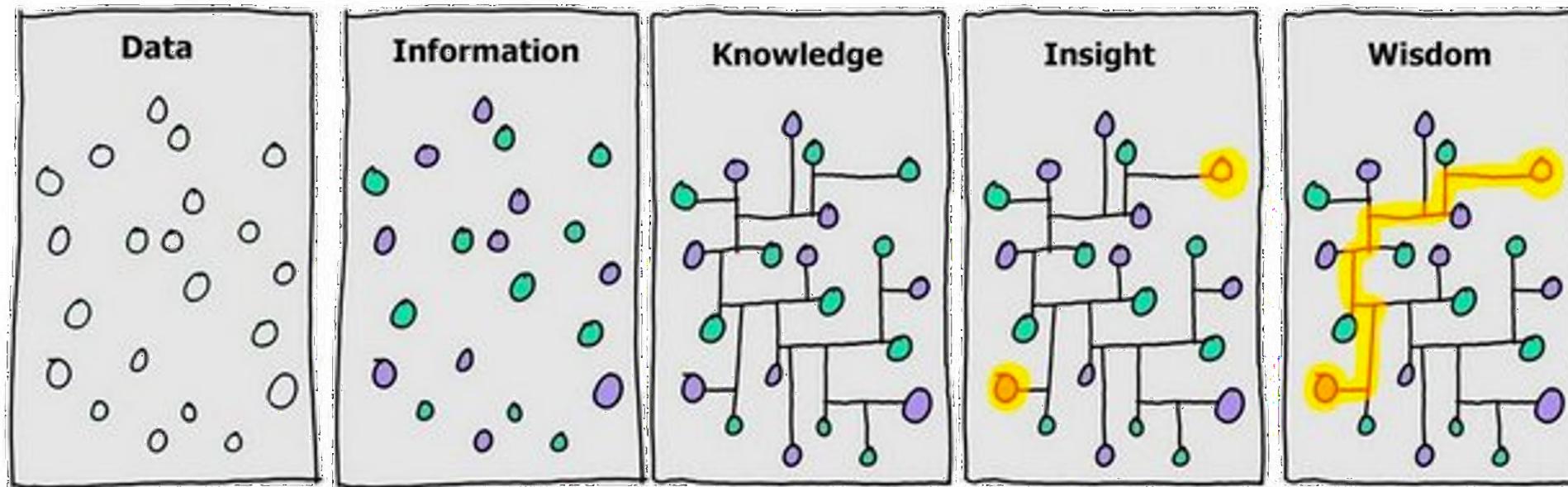
40

## □ Huawei's global industry outlook report GIV 2020



●●●●● Path from data to wisdom...

41





# Capacitação profissional em tecnologias de Inteligência Artificial

**The Controversy of AI &  
Future perspectives of AI**

Prof. Esther Luna Colombini

[www.ic.unicamp.br/~esther](http://www.ic.unicamp.br/~esther)

Institute of Computing - Unicamp

