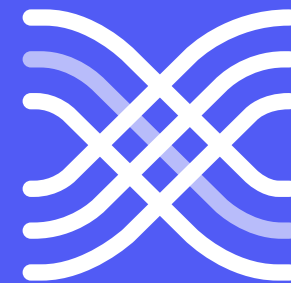# Fundamentals of AI

AI4Gov

Trusted AI for Transparent Public Governance fostering Democratic Values

Erik Novak

Department for Artificial Intelligence

Jožef Stefan Institute
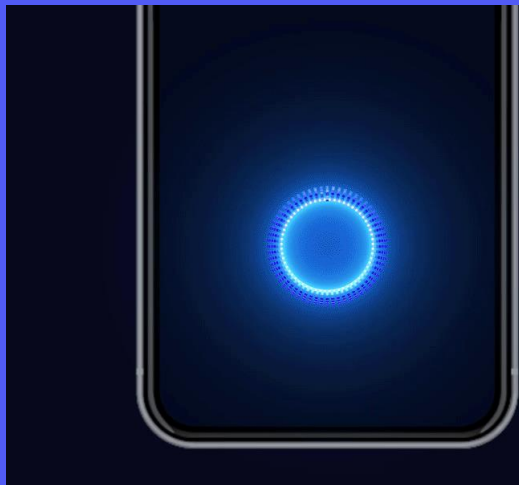
Ljubljana, Slovenia

Outline

- What is Artificial Intelligence?

- Where is it used?

- How does AI development look like?

- Categories of AI models
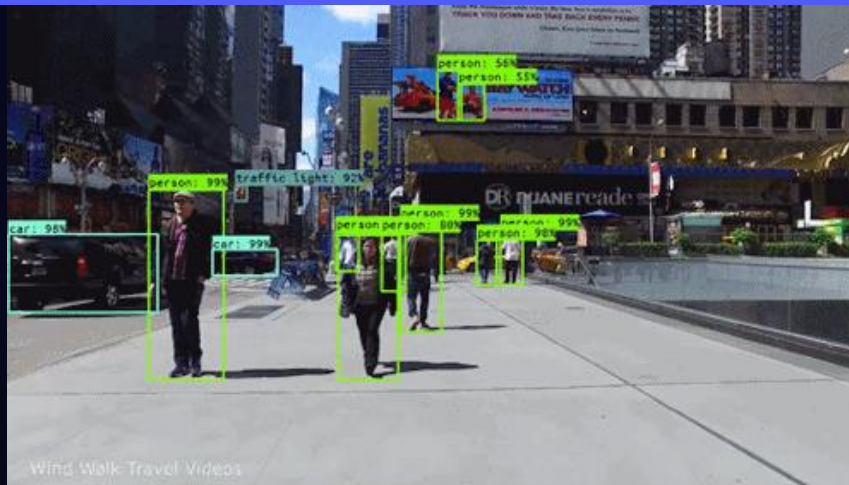
- Deep learning

- Problems with AI models

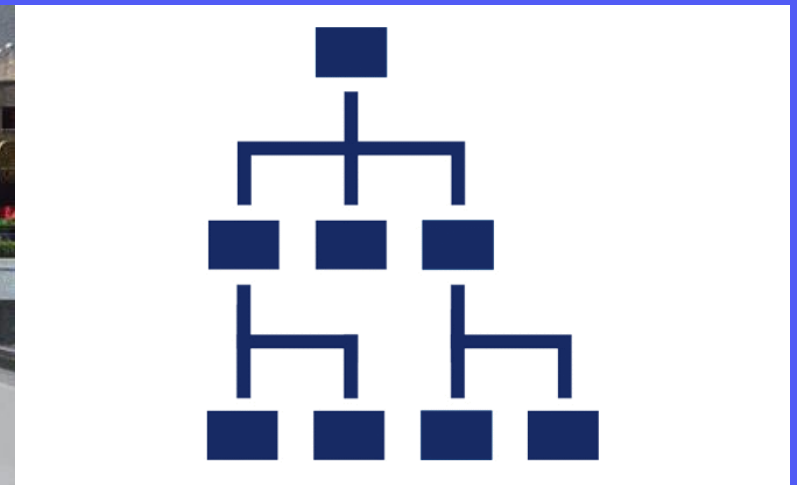# What is Artificial Intelligence (AI)?

"Artificial intelligence is the **simulation of human intelligence with devices.**"

natural language processing

computer vision

expert systems
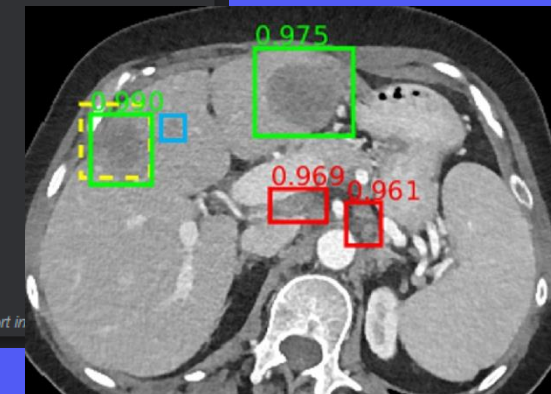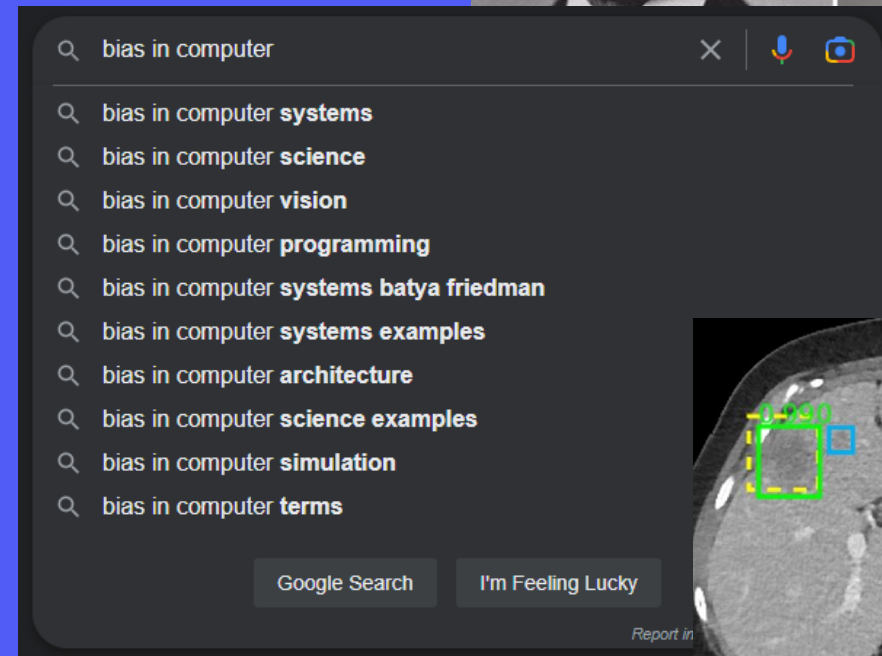
Where is AI used?

- Face recognition, friend recommendations (Facebook)
- Speech recognition (Siri, Cortana)
- Search recommendation (Google)
- Generation of 3D images (DeepFake)
- Product recommendations (Amazon)
- Self-driving vehicles (Tesla)

- Help in factories and production lines
- Help in medicine
- Simulation of real scenarios (digital twin)
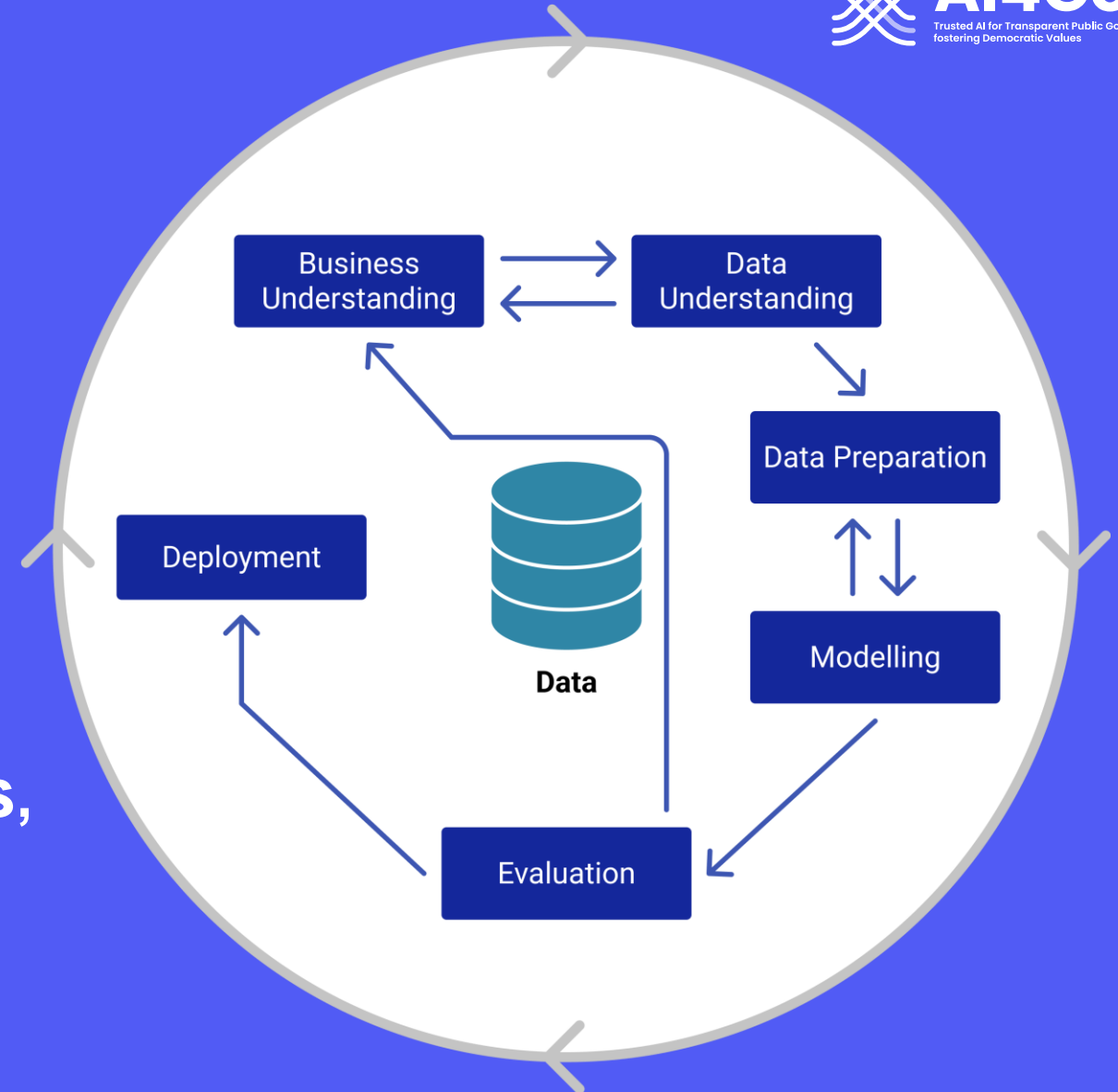
How does AI development look like?

Development cycle

How does AI development look like?

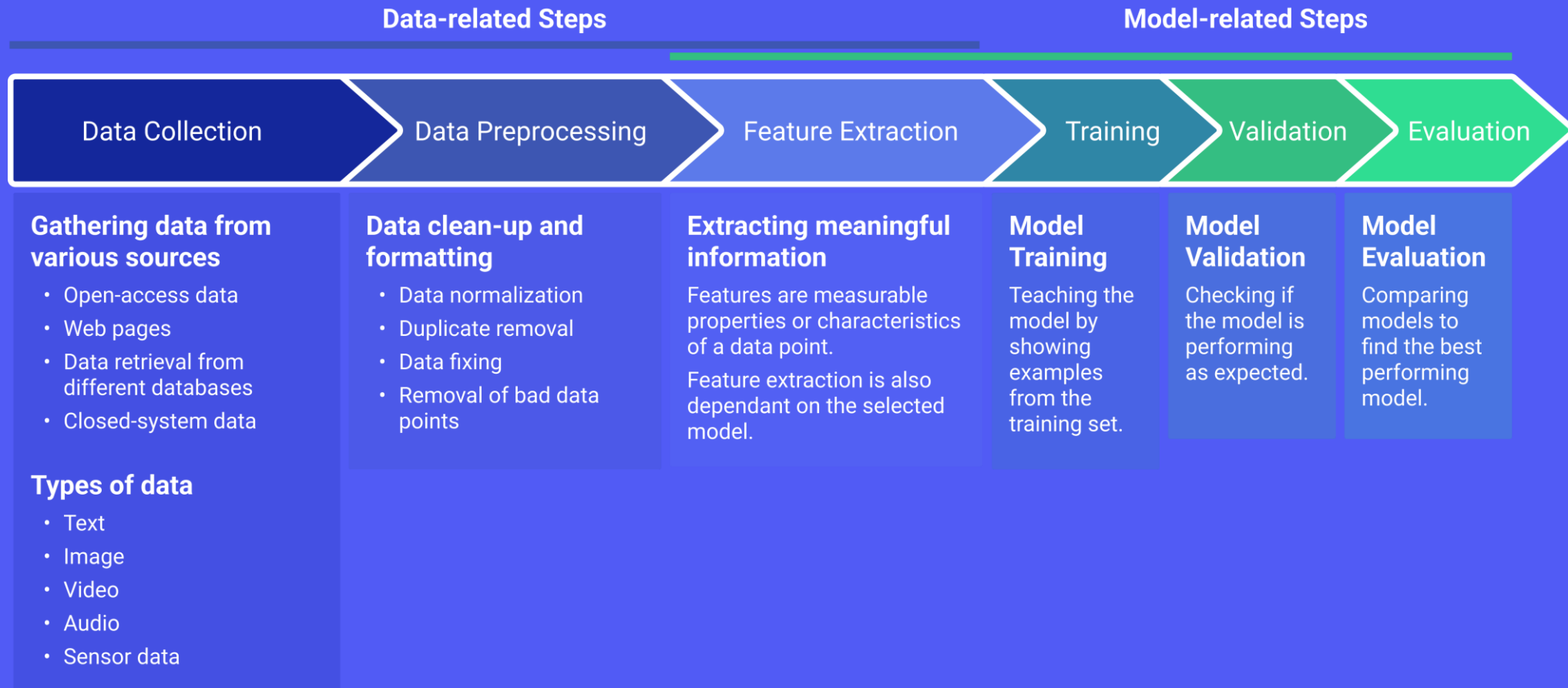Strong understanding of the problem

- What we want to solve?

- Why we want to solve?

- How do we want to solve it?

The development is continuous,
building on top of existing solutions

Model development timeline

How does AI development look like?

AI4Gov
Trusted AI for Transparent Public Governance
fostering Democratic Values

Data-related Steps   **Model-related Steps**

| Data Collection | Data Preprocessing | Feature Extraction | Training | Validation | Evaluation |

Gathering data from various sources
- Open-access data
- Web pages
- Data retrieval from different databases
- Closed-system data

Types of data
- Text
- Image
- Video
- Audio
- Sensor data

Data clean-up and formatting
- Data normalization
- Duplicate removal
- Data fixing
- Removal of bad data points

Extracting meaningful information

Features are measurable properties or characteristics of a data point.

Feature extraction is also dependant on the selected model.

Model Training

Teaching the model by showing examples from the training set.

Model Validation

Checking if the model is performing as expected.

Model Evaluation

Comparing models to find the best performing model.

Categories of AI models

Which model to use given the data?

Supervised learning

AI4Gov
Trusted AI for Transparent Public Governance
fostering Democratic Values

Categories of AI models

Use of **labelled data** to train algorithms
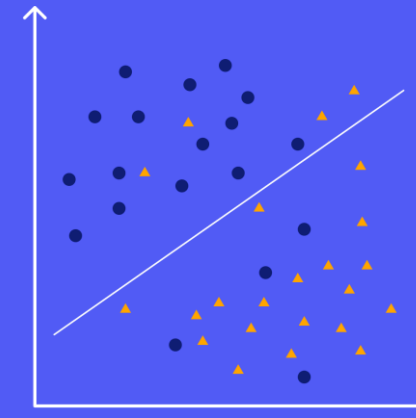to predict accurate outcomes
Example: email → spam or not spam
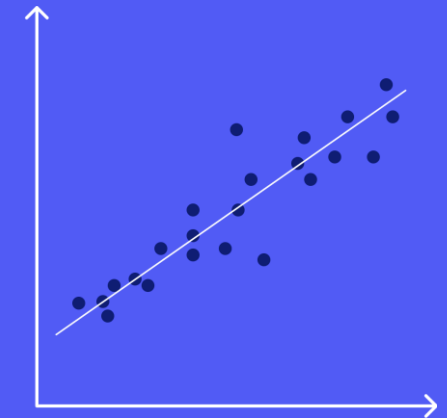

Figure: classification example

Figure: regression example

- Classification
- Regression
- Question Answering
- Named Entity Recognition
- Text Summarization
- Text Translation
- Etc.

To further elaborate on the geographical trends, North America `LOC` has procured more than 50% `PERCENT` of the global share in 2017 `DATE` and has been leading the regional landscape of AI `GPE` in the retail market. The U.S. `GPE` has a significant credit in the regional trends with over 65% `PERCENT` of investments (including M&As, private equity, and venture capital) in artificial intelligence technology. Additionally, the region is a huge hub for startups in tandem with the presence of tech titans, such as Google `ORG`, IBM `ORG`, and Microsoft `ORG`.
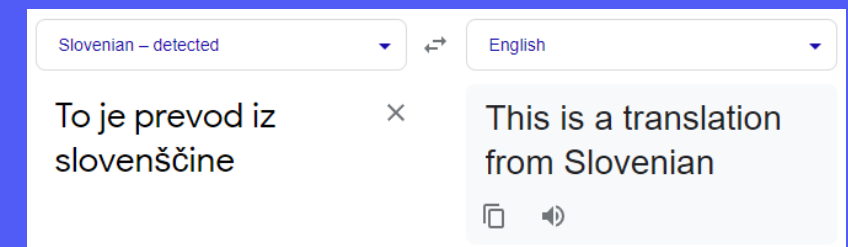
Figure: Spacy named entity recognition

Slovenian – detected   English

To je prevod iz slovenščine   This is a translation from Slovenian

Figure: google translate

Unsupervised learning

AI4Gov
Trusted AI for Transparent Public Governance fostering Democratic Values

Categories of AI models

Use of algorithms to analyse and cluster **unlabelled data;** identify the data's distribution

Example: group the survey's answers based on their content

- Data clustering
- Dimensionality reduction
- Anomaly detection
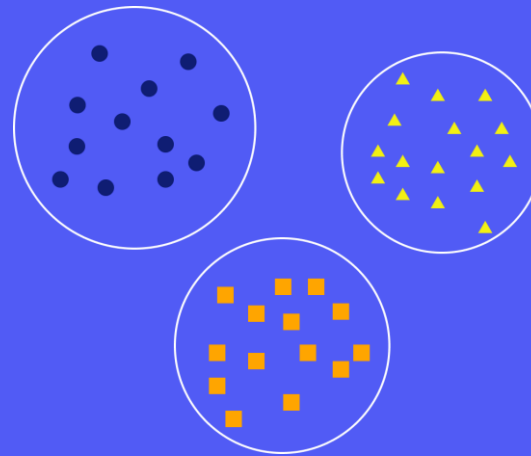- Meta-learning
- Etc.
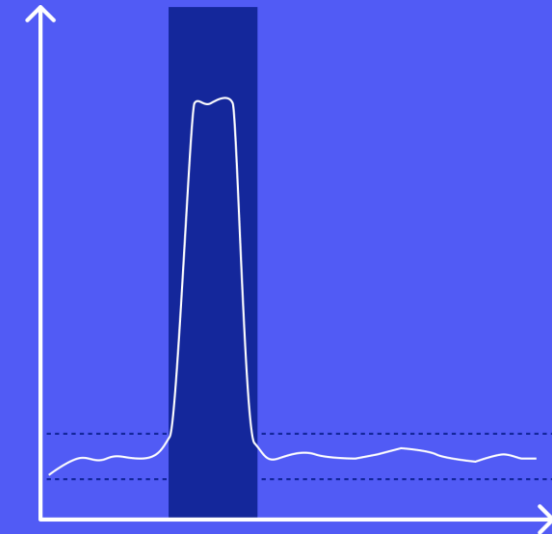
Figure: data clustering example

Figure: anomaly detection example

Semi-supervised learning

Categories of AI models

Use of both **labelled and unlabelled data** to train algorithms to solve problems

- Labelled data is used for initial ground predictions
- Unlabelled data to learn the shape of the larger data distribution

Example: given a small number of labelled spam and not spam emails, create a spam filter

- Active learning
- Reinforcement learning
- Generative adversarial networks (GAN)
- Etc.

Deep learning

- A subset of machine learning based on **artificial neural networks**
  - Simulating the behaviour of the human brain

- **Eliminating some of the data pre-processing** that is typically involved with machine learning (namely data pre-processing & feature extraction)

- Neural networks push the **state-of-the-art in AI**
  - Language models (ChatGPT), image generation (midjourney)
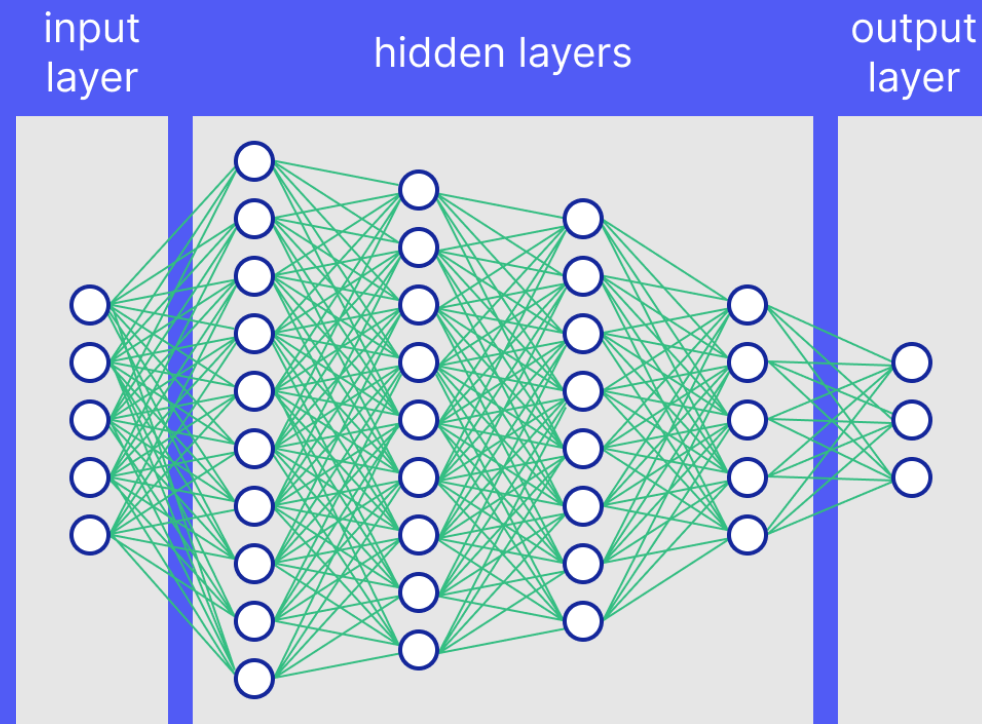
Deep learning

- A subset of machine learning based on **artificial neural networks**
  - Simulating the behaviour of the human brain

- **Eliminating some of the data pre-processing** that is typically involved with machine learning (namely data pre-processing & feature extraction)

- Neural networks push the **state-of-the-art in AI**
  - Language models (ChatGPT), image generation (midjourney)

Figure: created using midjourney using the prompt "Politicians sitting around a round table talking with each other. Above the table, levitating planet Earth. Isometric perspective. In the style of a drawn illustration"
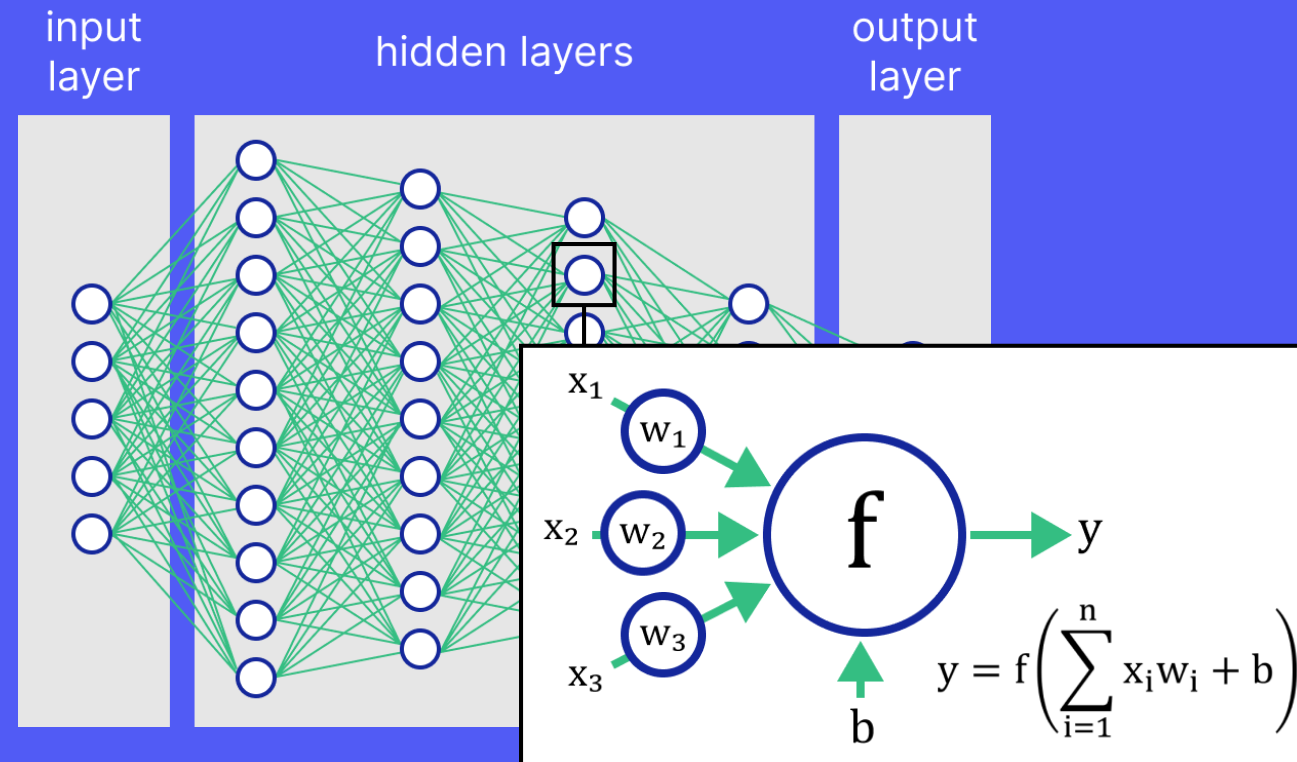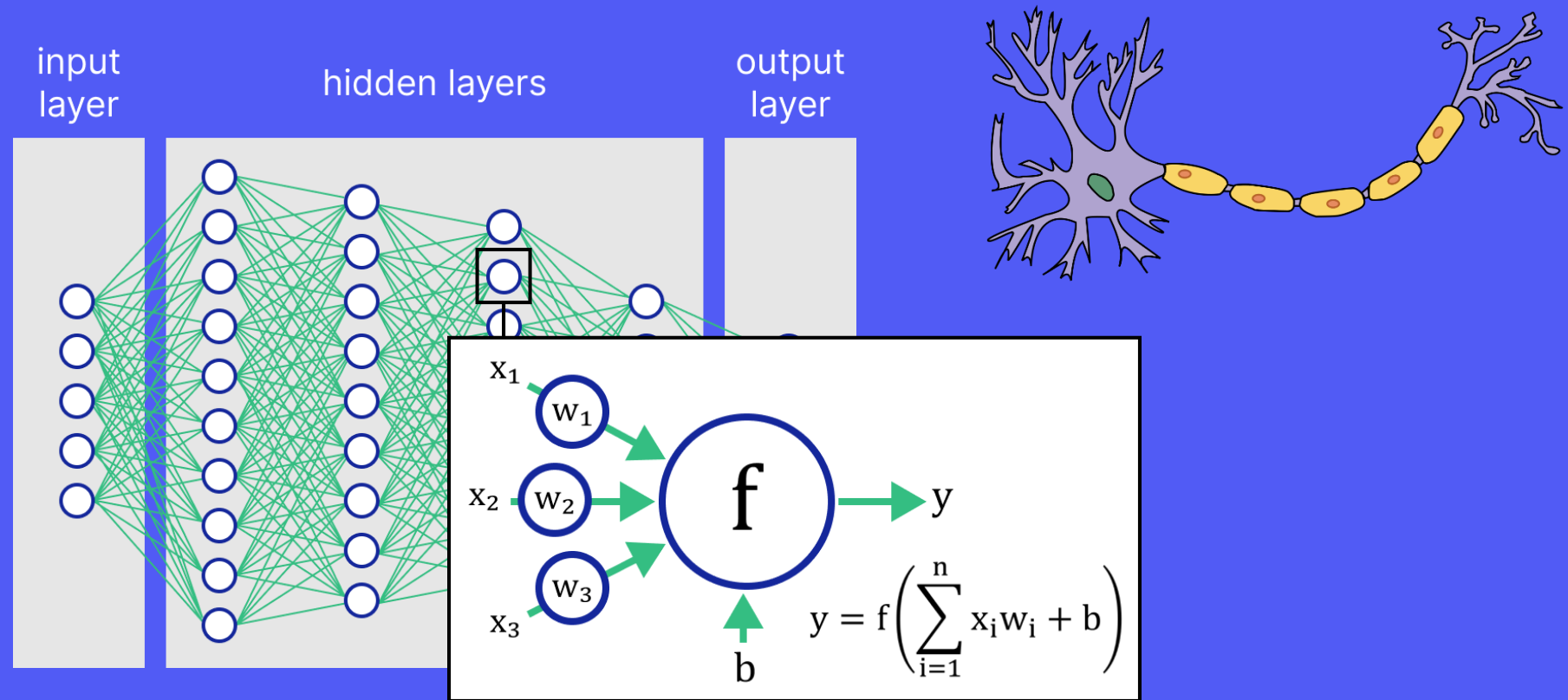
Neural networks

Deep learning

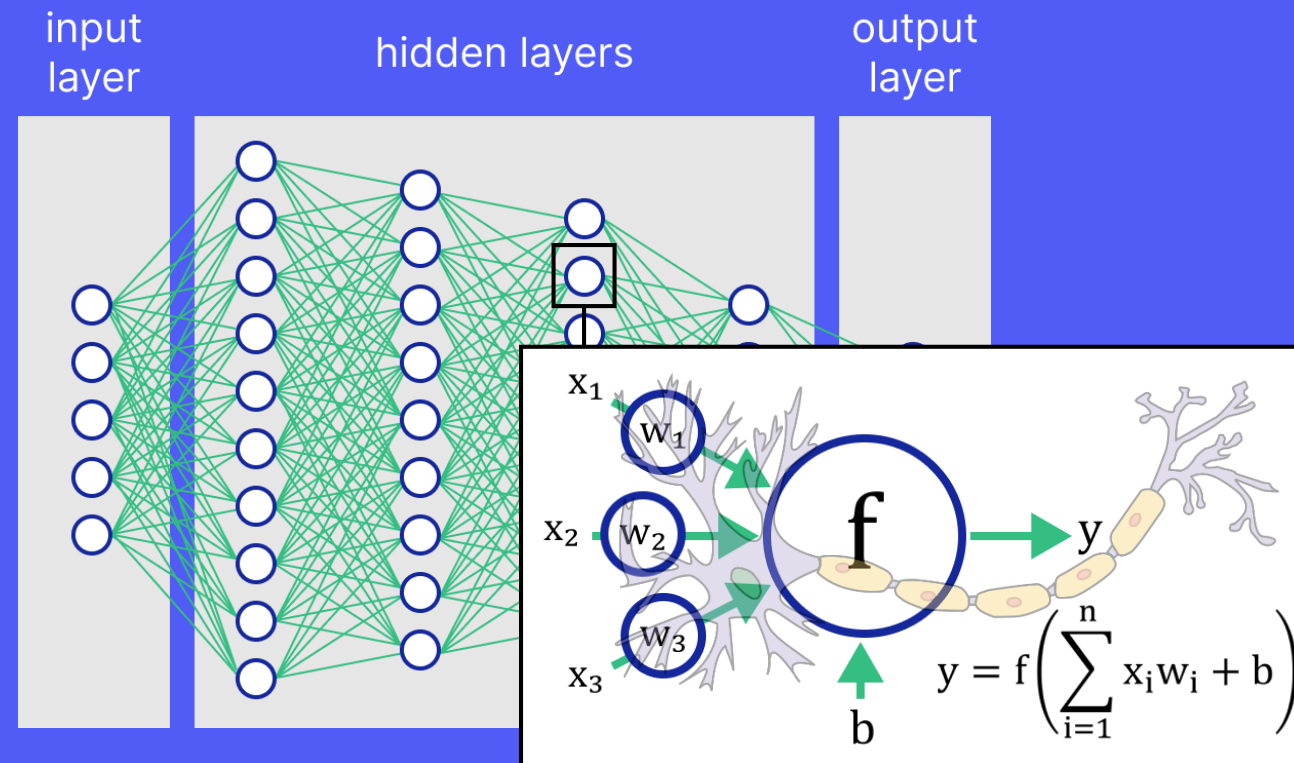# Neural networks (cont.)

Deep learning

Neural networks (cont.)

Deep learning

input layer

hidden layers

output layer

$$y = f\left(\sum_{i=1}^{n} x_i w_i + b\right)$$

Neural networks (cont.)

Deep learning

input layer

hidden layers

output layer

$$y = f\left(\sum_{i=1}^{n} x_i w_i + b\right)$$

Neural networks (cont.)

Deep learning

Problems with AI models

AI4Gov
Trusted AI for Transparent Public Governance
fostering Democratic Values

- **Black box models (neural networks).** We generally do not know how/why such predictions were made.

Input → Black Box Model → Output
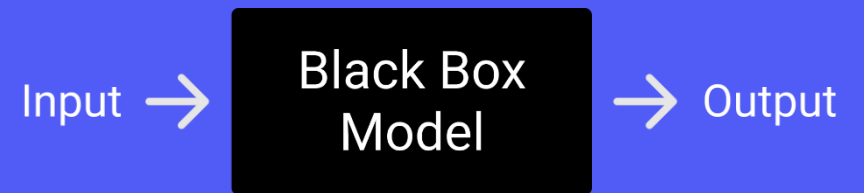
- **AI models can make mistakes.** AI models can make bad predictions due to various reasons (insufficient data, problems when training a model, etc.).

- **AI models incorporate bias.** Each model has it's own internal structure, and were trained using the provided data. Both can incorporate some notion of bias.

THANK YOU FOR YOUR ATTENTION
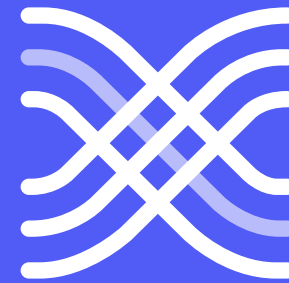
AI4Gov

Trusted AI for Transparent Public Governance fostering Democratic Values

Erik Novak

Email: erik.novak@ijs.si

Department for Artificial Intelligence

Jožef Stefan Institute

Ljubljana, Slovenia