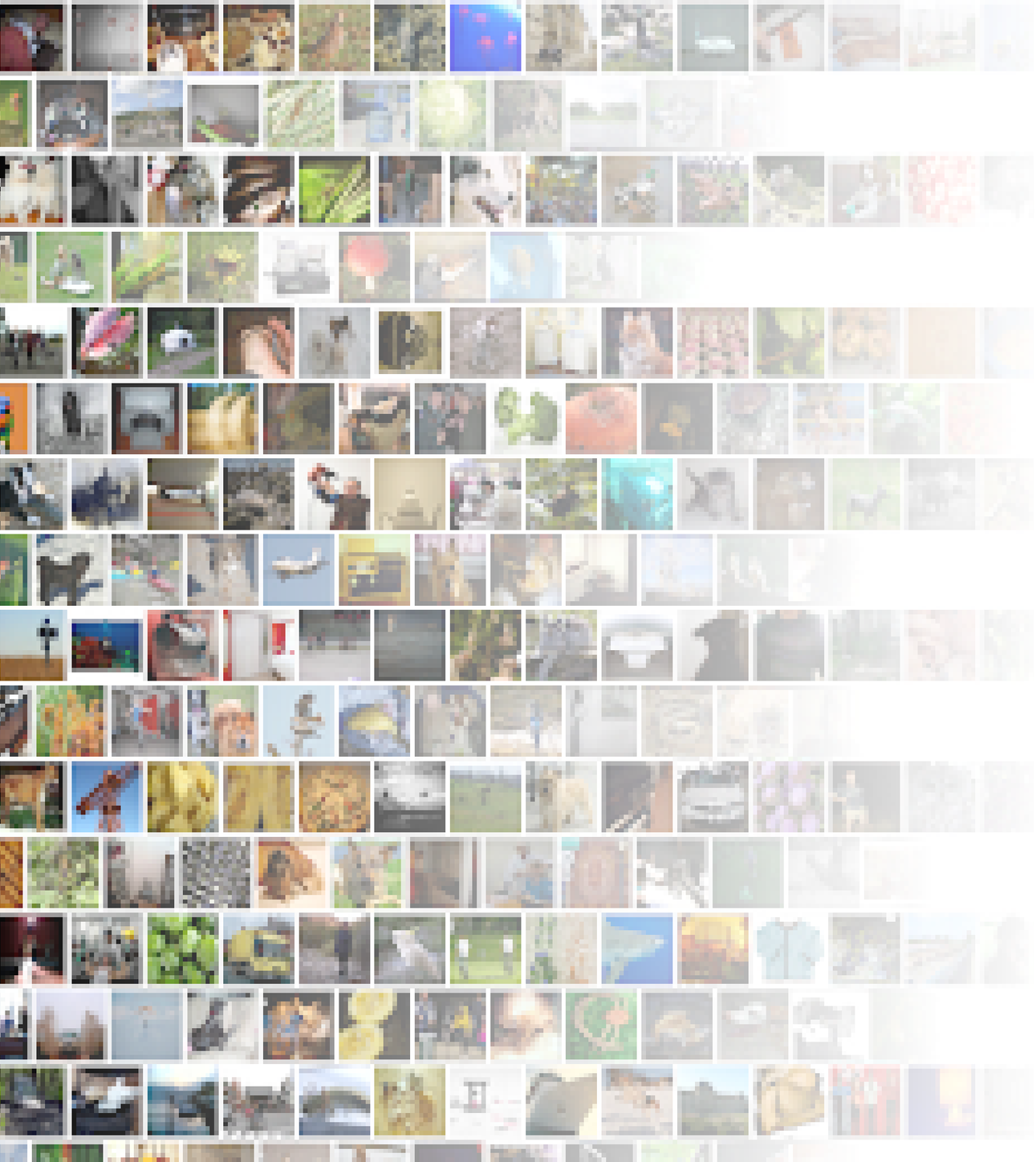


AI
Tutorial, EUREF 2023

mogren.one, olof.mogren@ri.se --- EUREF 2023

RI.SE

Who am I?

- Research director AI at RISE
- MSc in computer science
  Gothenburg University, 2007
- PhD in computer science
  Chalmers 2018

mogren.one, olof.mogren@ri.se

Research Institutes of Sweden

- Sweden’s public research institute
- Support for industry and public sector
- ~2800 people at 35 locations
- Extreme width of expertise

Divisions

- **Digital systems**
- Built environment
- Bioeconomy and health
- Materials and production
- Safety and transport

mogren_one, olof.mogren@ri.se

AI 2023

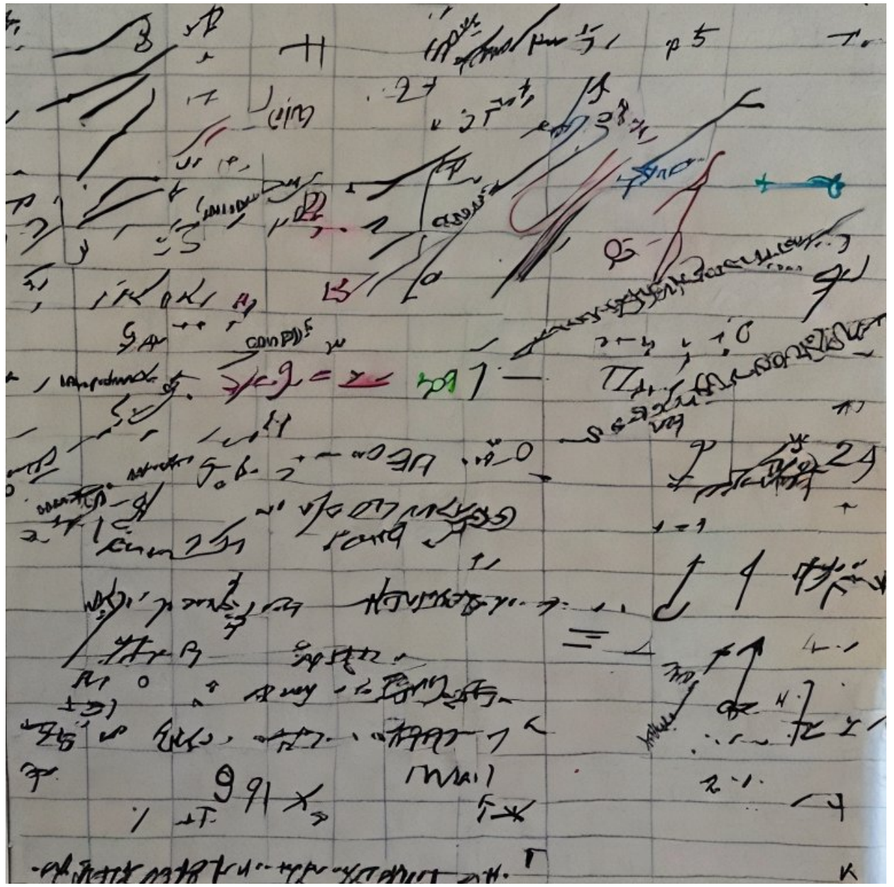
- Generative AI
  - ChatGPT
  - DALL-E
  - Etc
- Advanced cognition/analytics
  - Environment
  - Economy
  - Medicine
  - Self-driving vehicles

AI

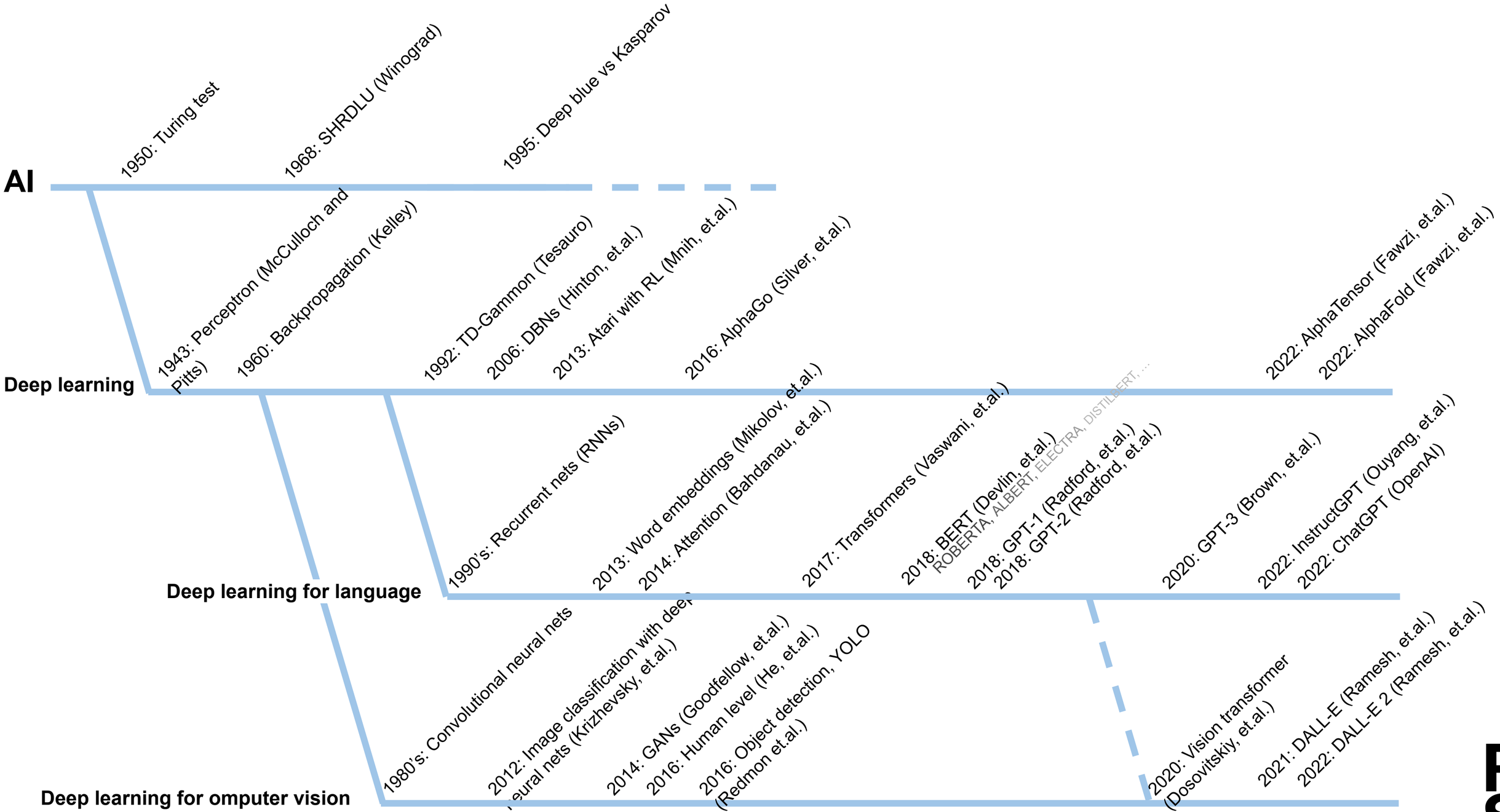
mogren.one, olof.mogren@ri.se

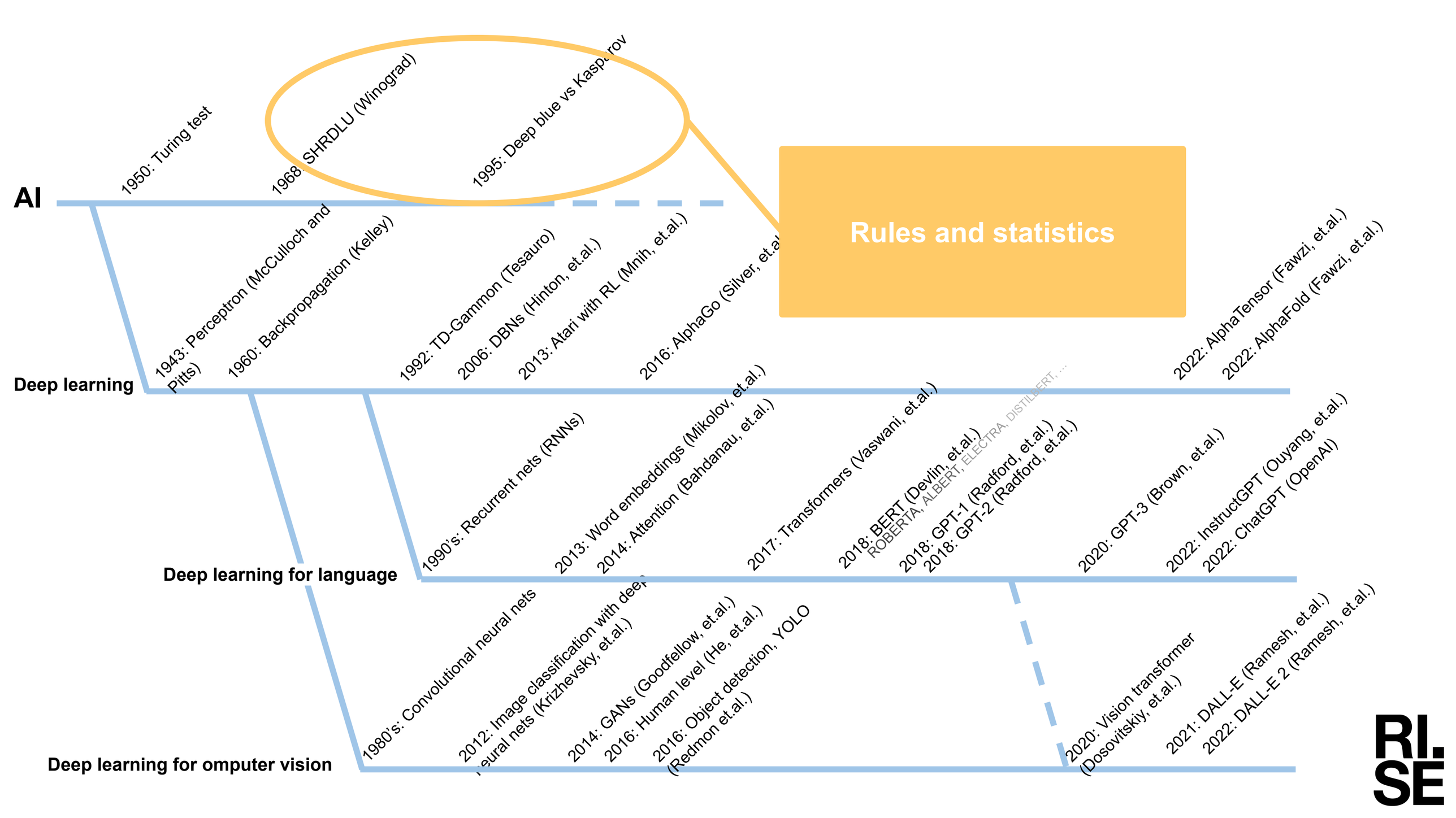
RI.SE

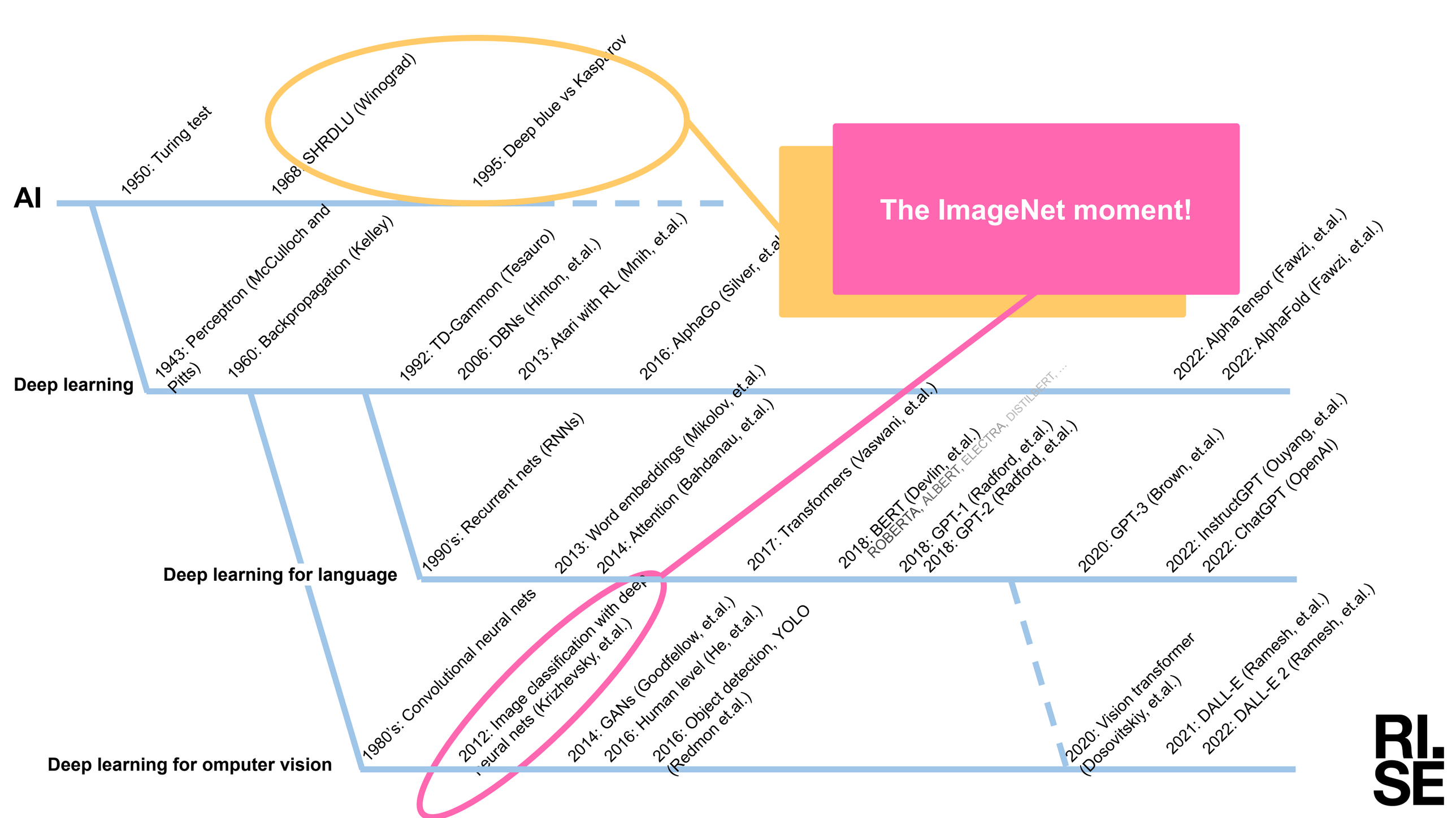
Also AI

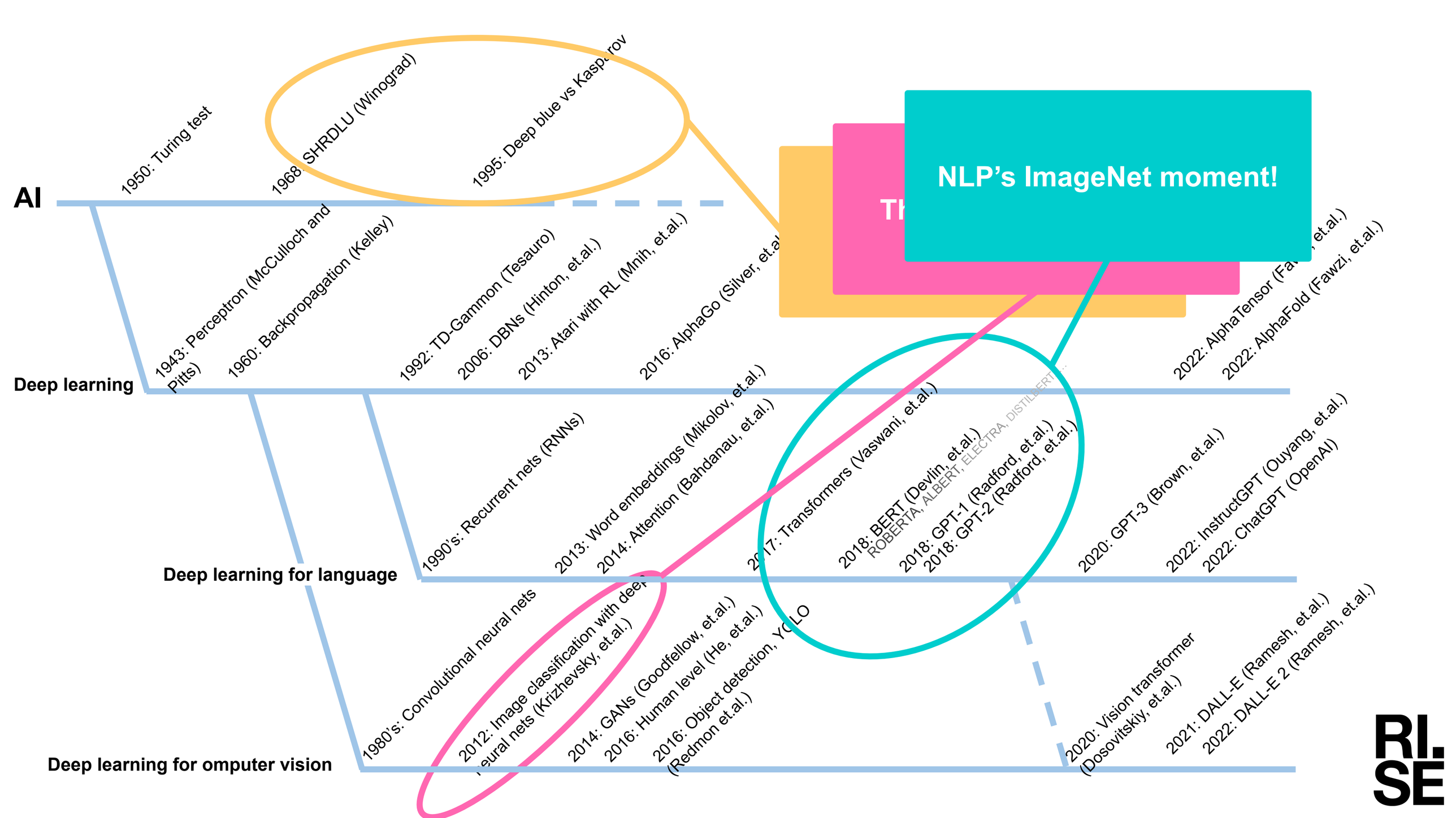


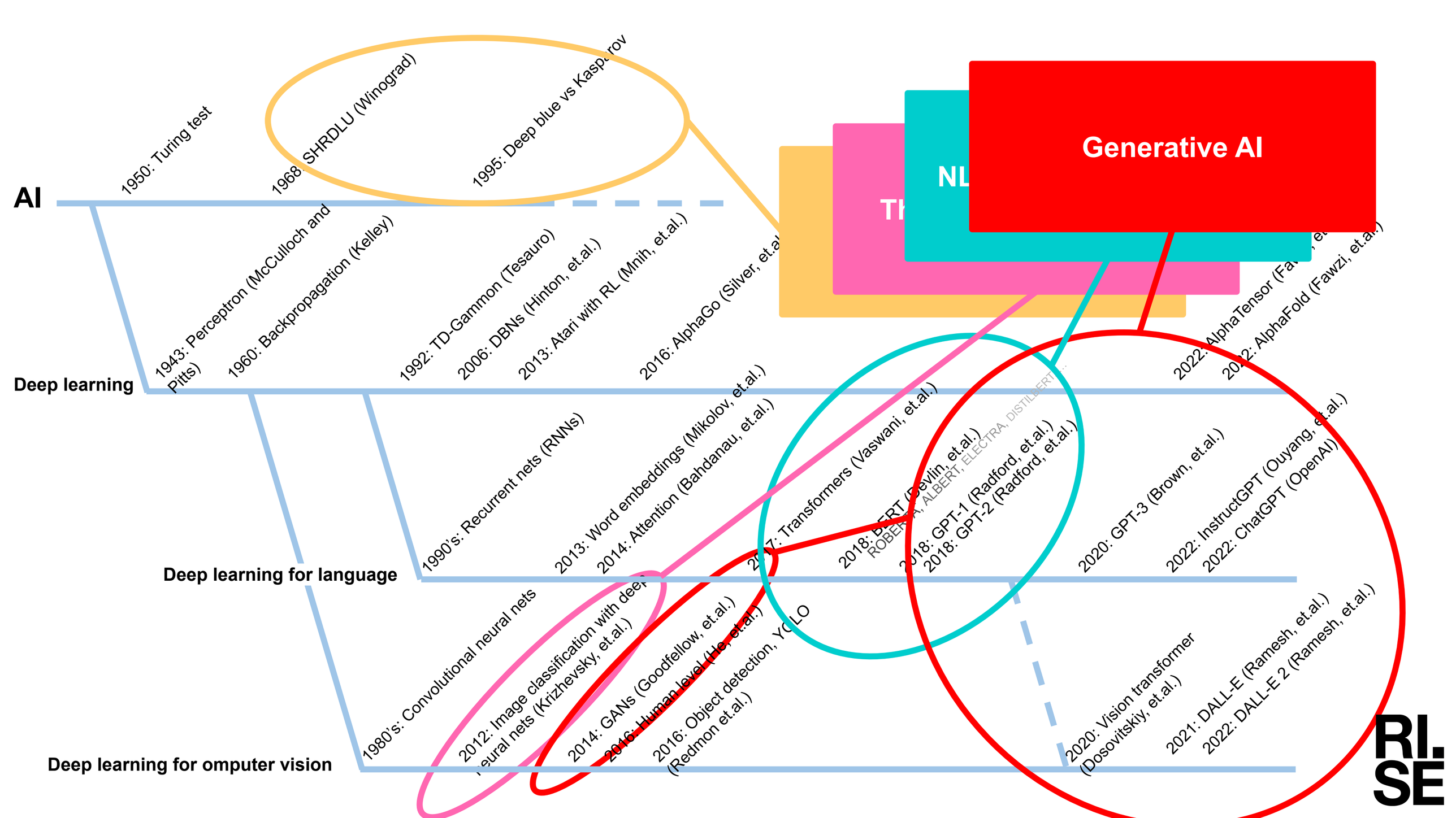
RI.
SE


AI
1950: Turing test
1968: SHRDLU (Winograd)
1995: Deep blue vs Kasparov
Deep learning
1943: Perceptron (McCulloch and Pitts)
1960: Backpropagation (Kelley)
1992: TD-Gammon (Tesauro)
2006: DBNs (Hinton, et.al.)
2013: Atari with RL (Mnih, et.al.)
2016: AlphaGo (Silver, et.al.)
2022: AlphaTensor (Fawzi, et.al.)
2022: AlphaFold (Fawzi, et.al.)
Deep learning for language
1990's: Recurrent nets (RNNs)
2013: Word embeddings (Mikolov, et.al.)
2014: Attention (Bahdanau, et.al.)
2017: Transformers (Vaswani, et.al.)
2018: BERT (Devlin, et.al.)
ROBERTA, ALBERT, ELECTRA, DISTILBERT, ...
2018: GPT-1 (Radford, et.al.)
2018: GPT-2 (Radford, et.al.)
2020: GPT-3 (Brown, et.al.)
2022: InstructGPT (Ouyang, et.al.)
2022: ChatGPT (OpenAI)
Deep learning for omputer vision
1980's: Convolutional neural nets
2012: Image classification with deep
eural nets (Krizhevsky, et.al.)
2014: GANs (Goodfellow, et.al.)
2016: Human level (He, et.al.)
2016: Object detection, YOLO
(Redmon et.al.)
2020: Vision transformer
(Dosovitskiy, et.al.)
2021: DALL-E (Ramesh, et.al.)
2022: DALL-E 2 (Ramesh, et.al.)


AI
1950: Turing test
1968: SHRDLU (Winograd)
1995: Deep blue vs Kasparov
Rules and statistics
Deep learning
1943: Perceptron (McCulloch and Pitts)
1960: Backpropagation (Kelley)
1992: TD-Gammon (Tesauro)
2006: DBNs (Hinton, et.al.)
2013: Atari with RL (Mnih, et.al.)
2016: AlphaGo (Silver, et.al.)
2022: AlphaTensor (Fawzi, et.al.)
2022: AlphaFold (Fawzi, et.al.)
Deep learning for language
1990's: Recurrent nets (RNNs)
2013: Word embeddings (Mikolov, et.al.)
2014: Attention (Bahdanau, et.al.)
2017: Transformers (Vaswani, et.al.)
2018: BERT (Devlin, et.al.)
ROBERTA, ALBERT, ELECTRA, DISTILBERT, ...
2018: GPT-1 (Radford, et.al.)
2018: GPT-2 (Radford, et.al.)
2020: GPT-3 (Brown, et.al.)
2022: InstructGPT (Ouyang, et.al.)
2022: ChatGPT (OpenAI)
Deep learning for omputer vision
1980's: Convolutional neural nets
2012: Image classification with deep
eural nets (Krizhevsky, et.al.)
2014: GANs (Goodfellow, et.al.)
2016: Human level (He, et.al.)
2016: Object detection, YOLO
(Redmon et.al.)
2020: Vision transformer
(Dosovitskiy, et.al.)
2021: DALL-E (Ramesh, et.al.)
2022: DALL-E 2 (Ramesh, et.al.)

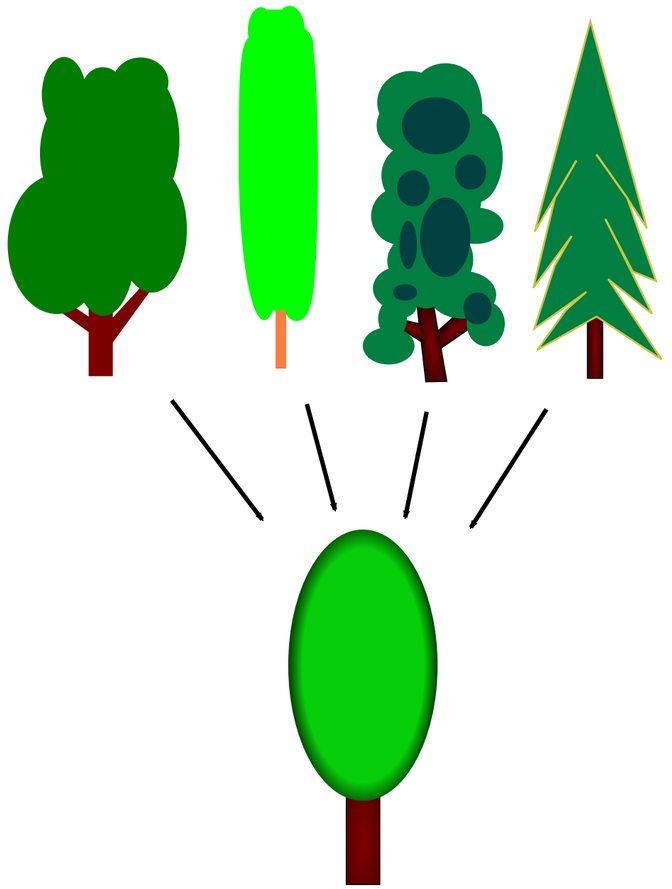

AI
1950: Turing test
1968: SHRDLU (Winograd)
1995: Deep blue vs Kasparov
The ImageNet moment!
Deep learning
1943: Perceptron (McCulloch and Pitts)
1960: Backpropagation (Kelley)
1992: TD-Gammon (Tesauro)
2006: DBNs (Hinton, et.al.)
2013: Atari with RL (Mnih, et.al.)
2016: AlphaGo (Silver, et.al.)
2022: AlphaTensor (Fawzi, et.al.)
2022: AlphaFold (Fawzi, et.al.)
Deep learning for language
1990's: Recurrent nets (RNNs)
2013: Word embeddings (Mikolov, et.al.)
2014: Attention (Bahdanau, et.al.)
2017: Transformers (Vaswani, et.al.)
2018: BERT (Devlin, et.al.)
ROBERTA, ALBERT, ELECTRA, DISTILBERT, ...
2018: GPT-1 (Radford, et.al.)
2018: GPT-2 (Radford, et.al.)
2020: GPT-3 (Brown, et.al.)
2022: InstructGPT (Ouyang, et.al.)
2022: ChatGPT (OpenAI)
Deep learning for omputer vision
1980's: Convolutional neural nets
2012: Image classification with deep
eural nets (Krizhevsky, et.al.)
2014: GANs (Goodfellow, et.al.)
2016: Human level (He, et.al.)
2016: Object detection, YOLO
(Redmon et.al.)
2020: Vision transformer
(Dosovitskiy, et.al.)
2021: DALL-E (Ramesh, et.al.)
2022: DALL-E 2 (Ramesh, et.al.)


AI
Deep learning
Deep learning for language
Deep learning for omputer vision
1950: Turing test
1968: SHRDLU (Winograd)
1995: Deep blue vs Kasparov
1943: Perceptron (McCulloch and Pitts)
1960: Backpropagation (Kelley)
1992: TD-Gammon (Tesauro)
2006: DBNs (Hinton, et.al.)
2013: Atari with RL (Mnih, et.al.)
2016: AlphaGo (Silver, et.al.)
2022: AlphaTensor (Fa
et.al.)
2022: AlphaFold (Fawzi, et.al.)
1990's: Recurrent nets (RNNs)
2013: Word embeddings (Mikolov, et.al.)
2014: Attention (Bahdanau, et.al.)
2017: Transformers (Vaswani, et.al.)
2018: BERT (Devlin, et.al.)
ROBERTA, ALBERT, ELECTRA, DISTILBERT...
2018: GPT-1 (Radford, et.al.)
2018: GPT-2 (Radford, et.al.)
2020: GPT-3 (Brown, et.al.)
2022: InstructGPT (Ouyang, et.al.)
2022: ChatGPT (OpenAI)
1980's: Convolutional neural nets
2012: Image classification with deep
eural nets (Krizhevsky, et.al.)
2014: GANs (Goodfellow, et.al.)
2016: Human level (He, et.al.)
2016: Object detection, YOLO
(Redmon et.al.)
2020: Vision transformer
(Dosovitskiy, et.al.)
2021: DALL-E (Ramesh, et.al.)
2022: DALL-E 2 (Ramesh, et.al.)
NLP's ImageNet moment!


AI
Deep learning
Deep learning for language
Deep learning for omputer vision
1950: Turing test
1968: SHRDLU (Winograd)
1995: Deep blue vs Kasp rov
1943: Perceptron (McCulloch and Pitts)
1960: Backpropagation (Kelley)
1992: TD-Gammon (Tesauro)
2006: DBNs (Hinton, et.al.)
2013: Atari with RL (Mnih, et.al.)
2016: AlphaGo (Silver, et.al
2022: AlphaTensor (Fa
2022: AlphaFold (Fawzi, et.al.)
1990's: Recurrent nets (RNNs)
2013: Word embeddings (Mikolov, et.al.)
2014: Attention (Bahdanau, et.al.)
2017: Transformers (Vaswani, et.al.)
2018: BERT (Devlin, et.al.)
ROBERTA, ALBERT, ELECTRA, DISTILBERT...
2018: GPT-1 (Radford, et.al.)
2018: GPT-2 (Radford, et.al.)
2020: GPT-3 (Brown, et.al.)
2022: InstructGPT (Ouyang, et.al.)
2022: ChatGPT (OpenAI)
1980's: Convolutional neural nets
2012: Image classification with deep
eural nets (Krizhevsky, et.al.)
2014: GANs (Goodfellow, et.al.)
2016: Human level (He, et.al.)
2016: Object detection, YOLO
(Redmon et.al.)
2020: Vision transformer
(Dosovitskiy, et.al.)
2021: DALL-E (Ramesh, et.al.)
2022: DALL-E 2 (Ramesh, et.al.)
Generative AI
RI.
SE

Generalization

- Crucial to learning
- Allows to transfer knowledge to new situations



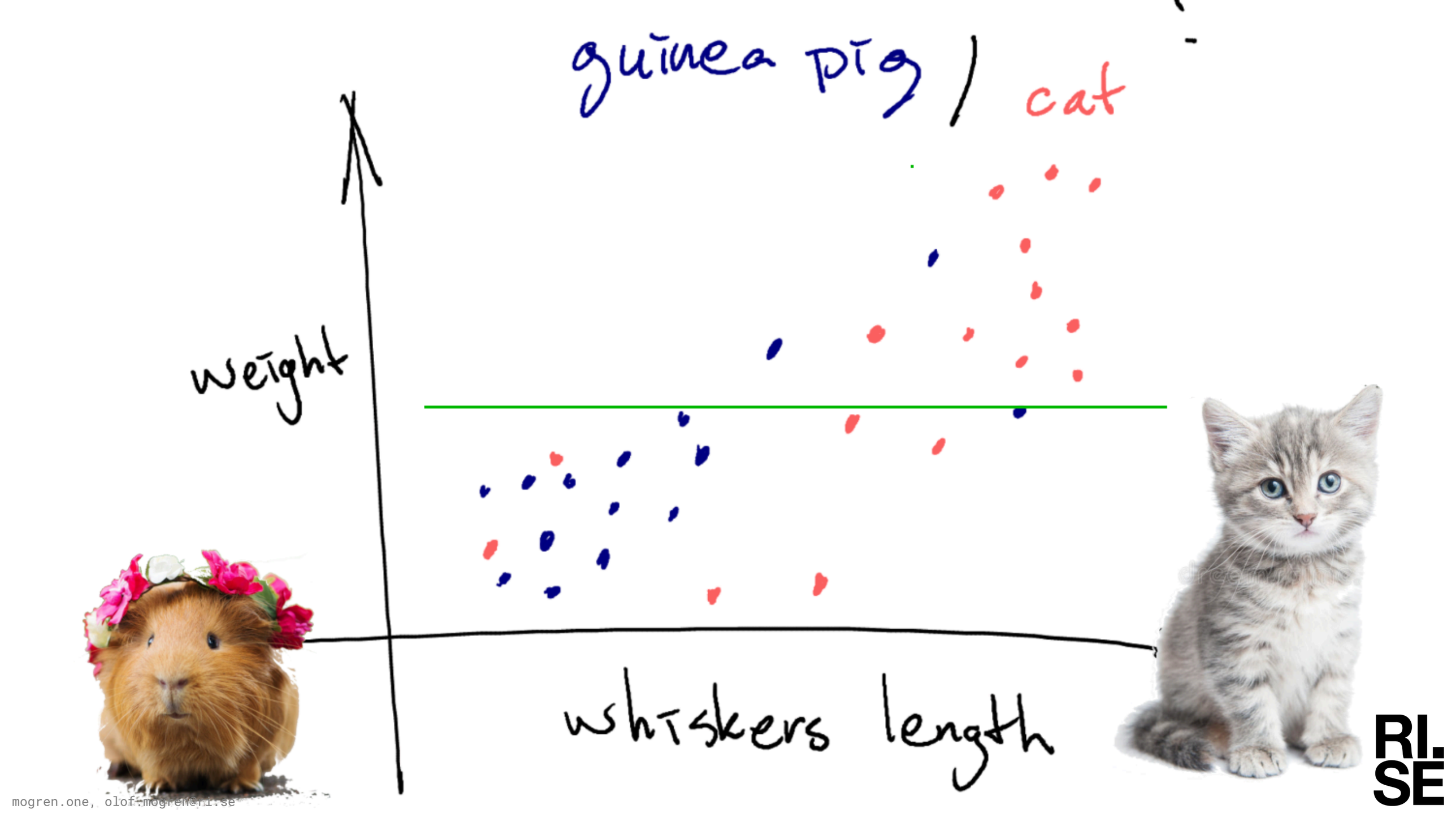
RI.
SE

guinea pig / cat

weight

whiskers length

RI.SE


guinea pig / cat
weight
whiskers length

guinea pig / cat

weight

whiskers length

RI.
SE

guinea pig / cat

weight

whiskers length

mogren.one, olof.mogren@ri.se

RI.
SE

guinea pig / cat

weight

whiskers length

RI.SE

cat / guinea pig

weight

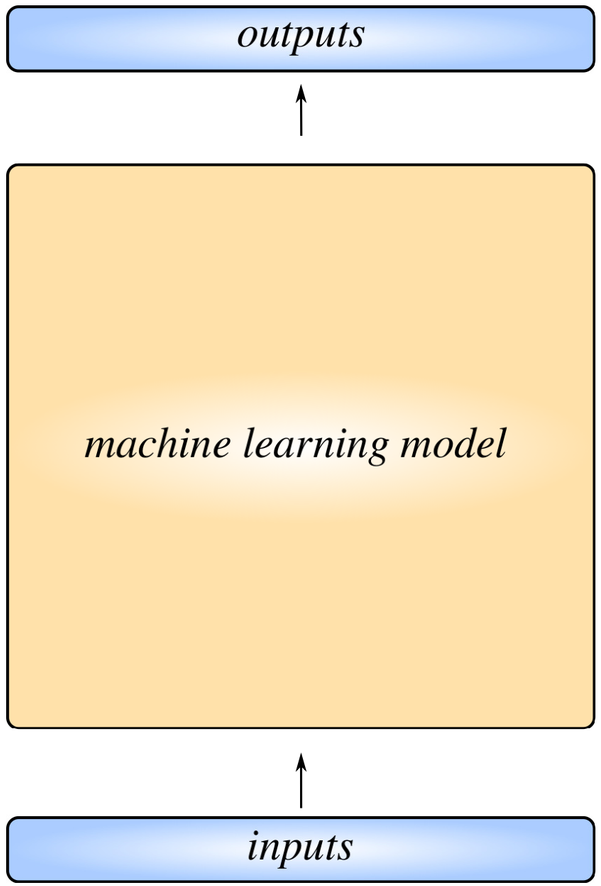
whiskers length

mogren.one, olof.mogren@ri.se

RI.SE

AI building blocks

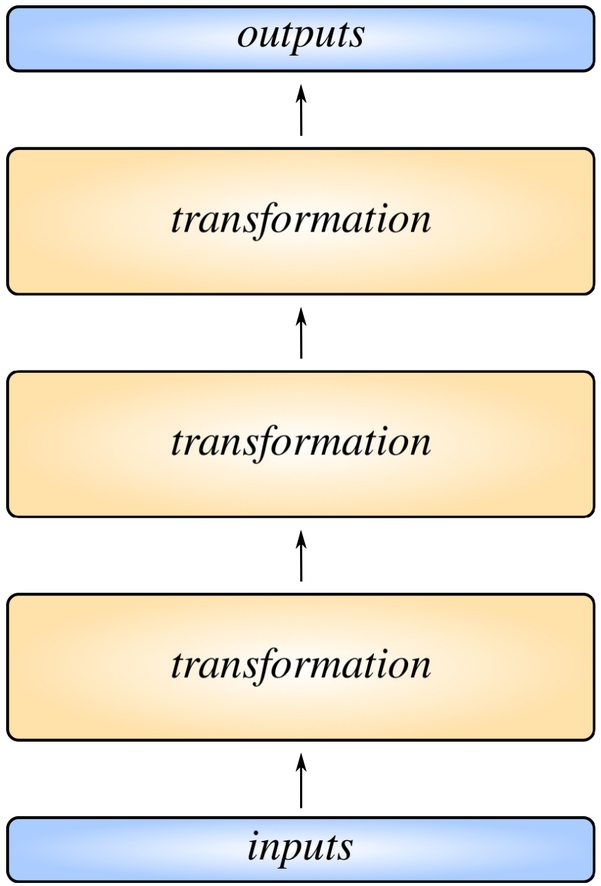
- Deep artificial neural networks
- Trained from data
  - Inputs
  - Outputs



Deep learning

Deep artificial neural networks

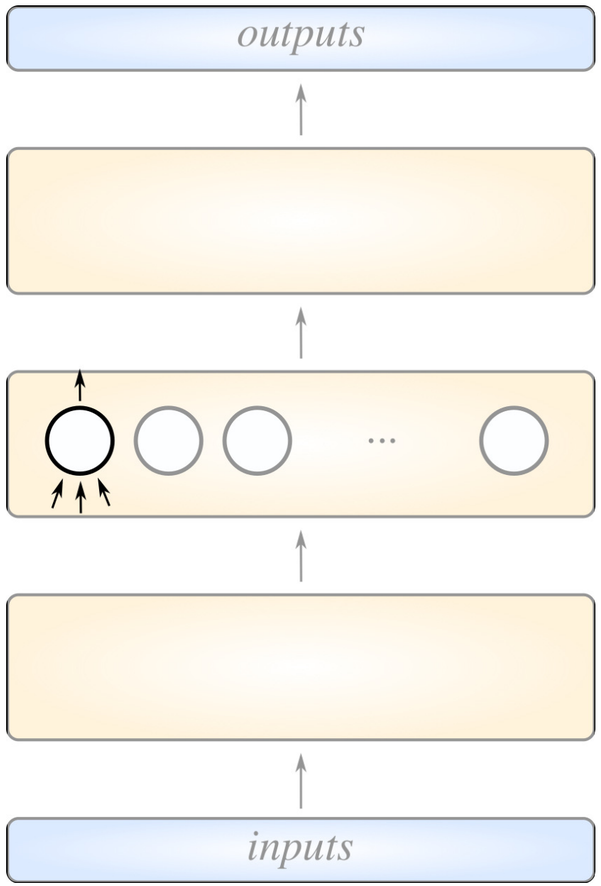
- Sequential transformations of data
- Trained together (backpropagation)
- Transformations are gradually changed



RI.
SE

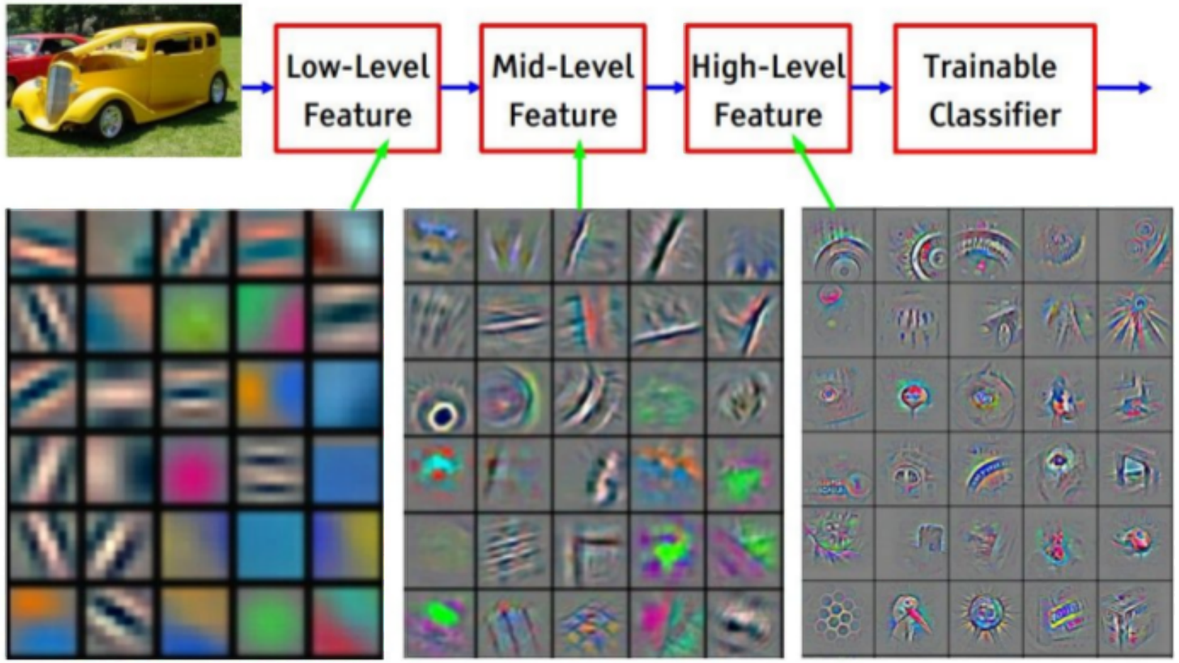
An artificial neuron

- Inspired by biological neurons
- Each layer in network:
  a collection of neurons
- Each neuron: $o = a(w * x + b)$
- w, b: trainable parameters
- A pattern as input,
  a value (scalar) as output

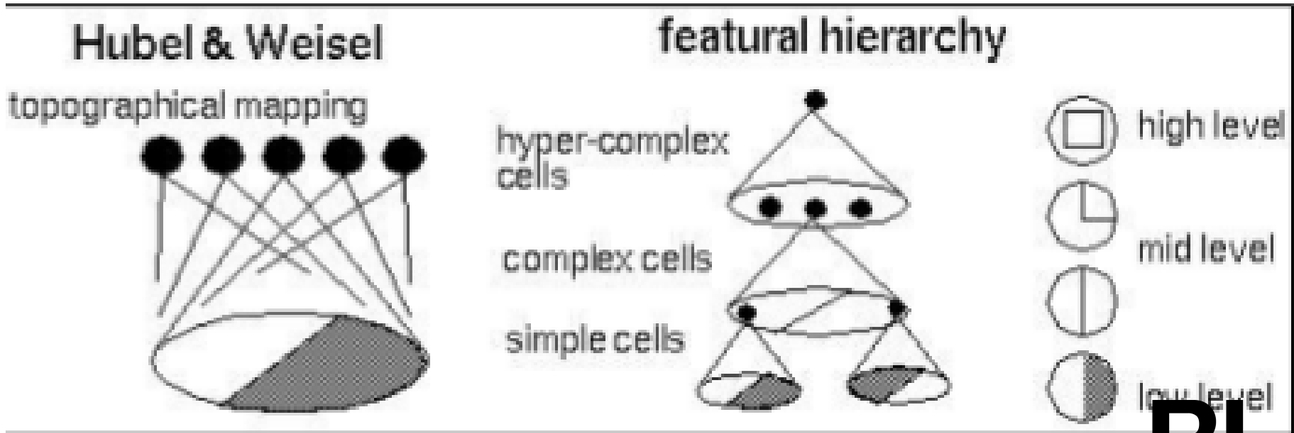


RI.
SE

Levels of abstractions



Feature visualization of convolutional net trained on ImageNet from [Zeiler & Fergus 2013]



RI.SE

Enough units -> universal approximation

Universal approximation theorem (Cybenko, 1989)

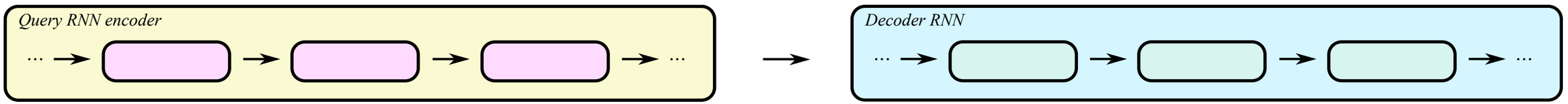
RI.
SE

AI: What is it good for?

- Computer vision
  - Classification
  - Segmentation
  - Object detection
  - Generating images
- Machine listening

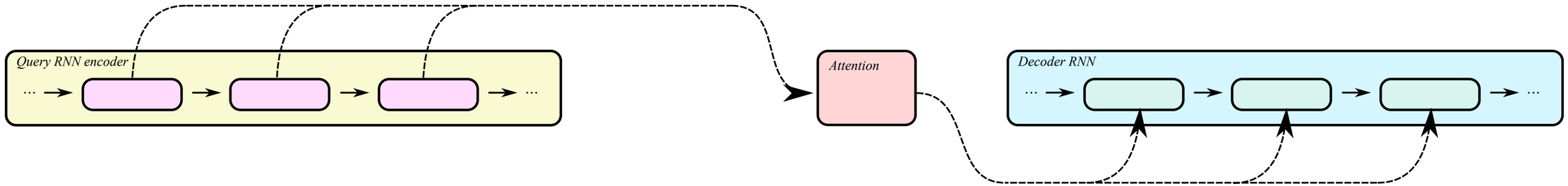
- Natural language processing
  - Translation
  - Summarization
  - Generating text
  - Chatting
- Sensor modelling

Deep learning for language



RI.
SE

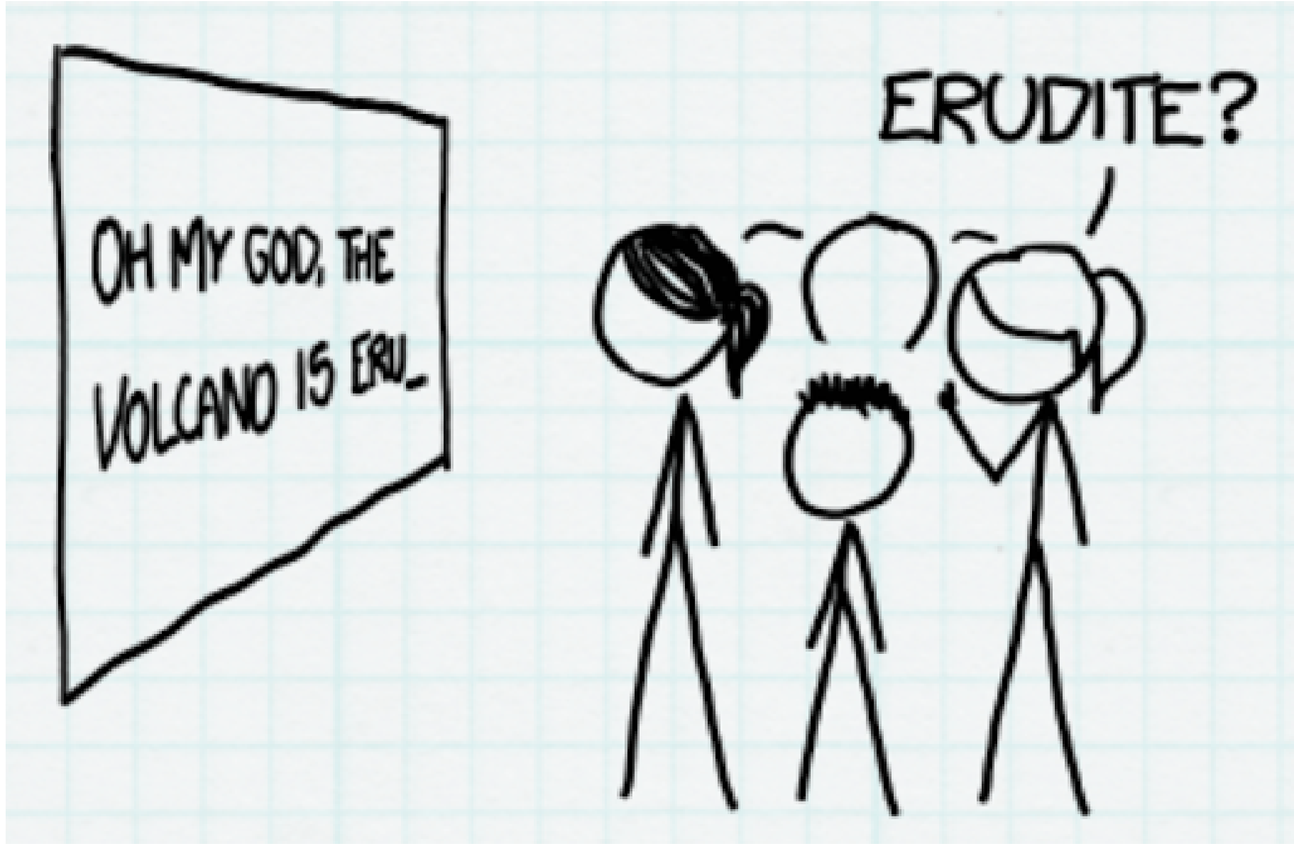
Attention mechanisms



RI.
SE

Language models

- Likelihood for a text
- Recently:
  - translate
  - generate
  - classify
  - etc



RI. SE

Generating language
Large language models

- Deep neural networks scale well with compute and data

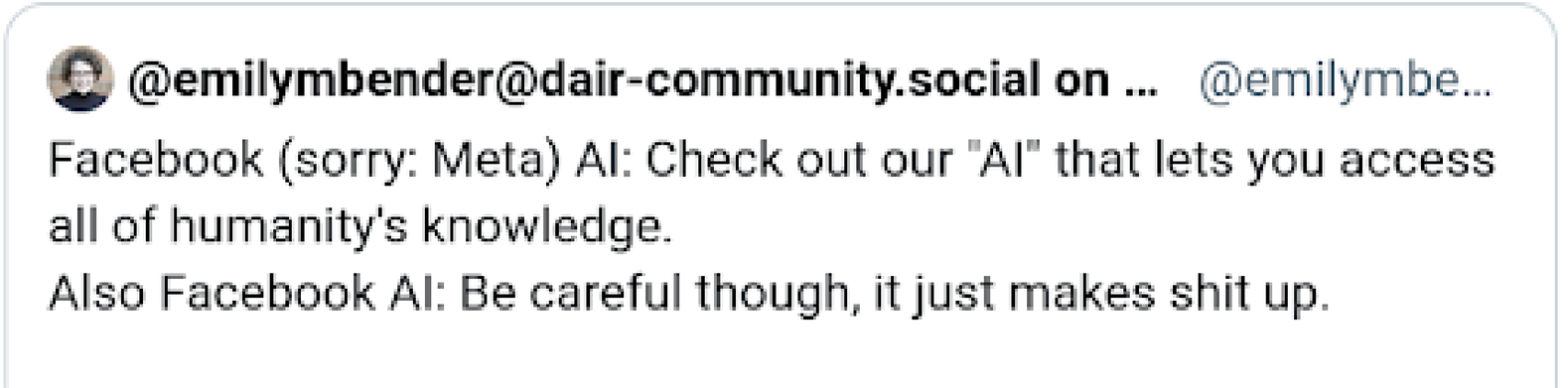
RI.
SE

Chatbots history

- ELIZA (Weizenbaum, 1960)
  - Pattern matching, templates
  - No contextualization
- Tay (Microsoft, 2016)
  - Praised Hitler
  - Taken offline within 16h

RI.SE

Chatbots now

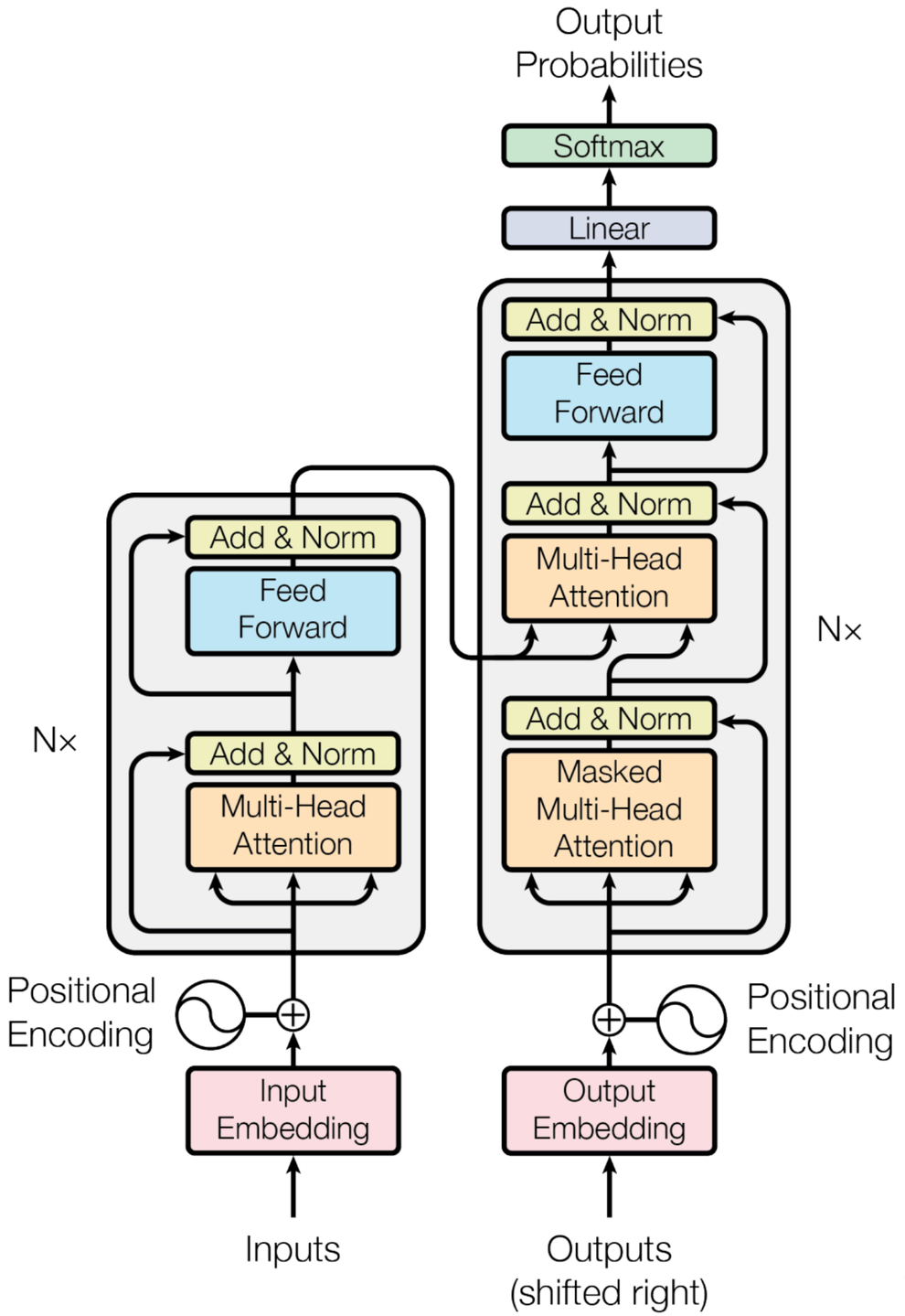


- Galactica (Meta AI Research, 2022; Taken offline within 72h)
- ChatGPT (OpenAI, 2022; Online since 2022-11-22)
- BARD (Google, 2023), YouChat (you.com 2022)

RI.
SE

Transformers

- Sequenceto sequence (encoder/decoder)
- Parallel computation, scaling
- Scaling training data
- Scaling models
- State-of-the-art in machine translation



Vaswani, et al., 2017

GPT
Generative pretrained Transformer

- Using decoder part of a Transformer
- Unsupervised pretraining
  - Next word prediction
- Fine-tuning for various tasks
- Can take inputs, “prompts”
- Generate answers as text

RI.
SE

Model	Training data			Parameters	Context (tokens)	Languages
GPT-1 (Radford, 2018)	BooksCorpus	~4.6 GiB	7 000 books	117M	512	English
GPT-2 (Radford, 2018)	BooksCorpus	~4.6 GiB	7 000 books	1,5B	1024	English
GPT-3 (Brown, et.al., 2020)	Common Crawl and Wikipedia	~380 TiB+ ~9 GiB		175B	2048	> 90 different languages
InstructGPT: (Ouyang, et.al., 2022)	GPT-3+supervised-questions/ answers+human-in-the-loop			175B		> 90 different languages
ChatGPT (2022)	“GPT-3.5”+supervised-questions/answers+human-in-the-loop			175B	4 096 or 8 192*	> 90 different languages
GPT-4 (expected 2023)				175B or more*		

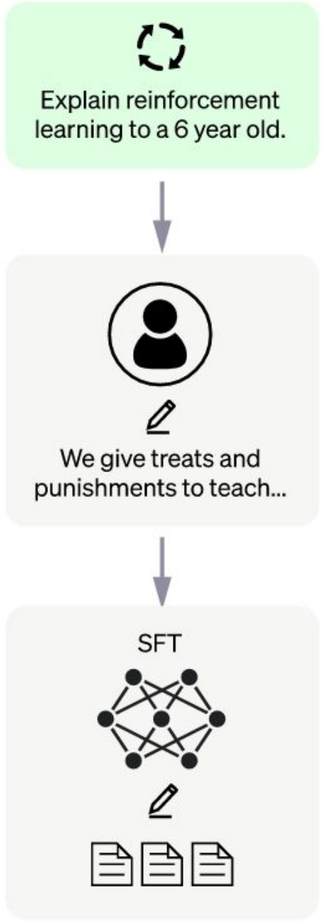
** unverified estimates.*

Olof Mogren, PhD
http://mogren.one
olof.mogren@ri.se

Step 1

Collect demonstration data and train a supervised policy.

A prompt is sampled from our prompt dataset.



A labeler demonstrates the desired output behavior.

This data is used to fine-tune GPT-3.5 with supervised learning.

https://openai.com/blog/chatgpt/

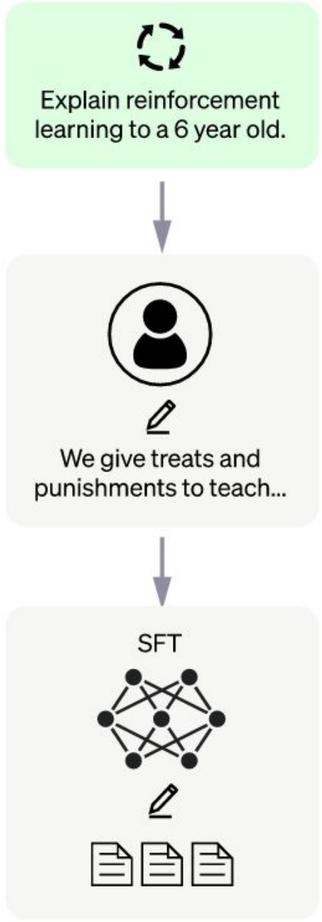
Step 1

Collect demonstration data and train a supervised policy.

A prompt is sampled from our prompt dataset.

A labeler demonstrates the desired output behavior.

This data is used to fine-tune GPT-3.5 with supervised learning.



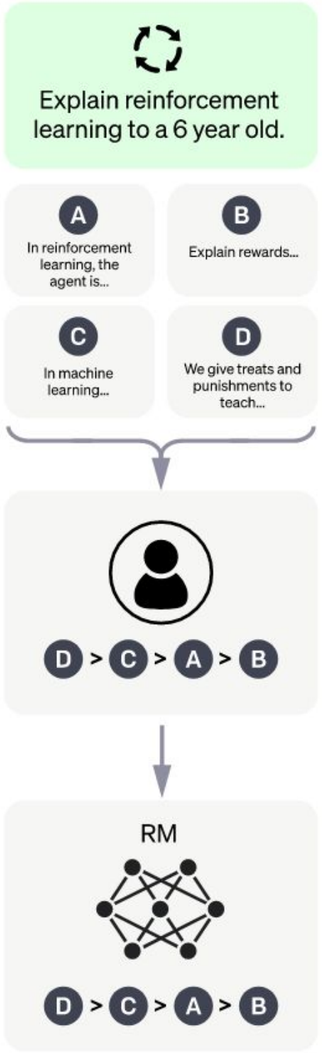
Step 2

Collect comparison data and train a reward model.

A prompt and several model outputs are sampled.

A labeler ranks the outputs from best to worst.

This data is used to train our reward model.



https://openai.com/blog/chatgpt/

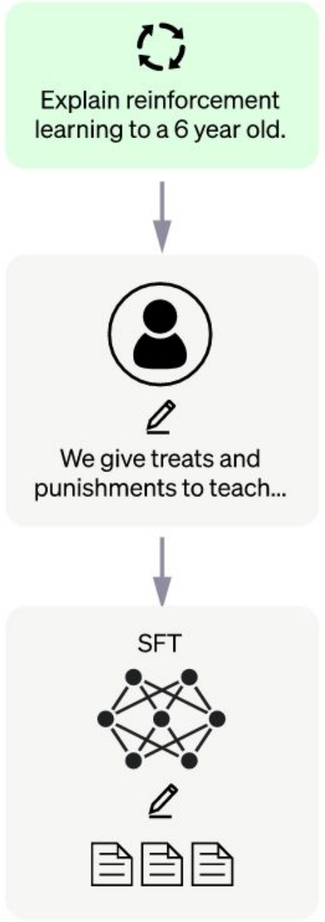
Step 1

Collect demonstration data and train a supervised policy.

A prompt is sampled from our prompt dataset.

A labeler demonstrates the desired output behavior.

This data is used to fine-tune GPT-3.5 with supervised learning.



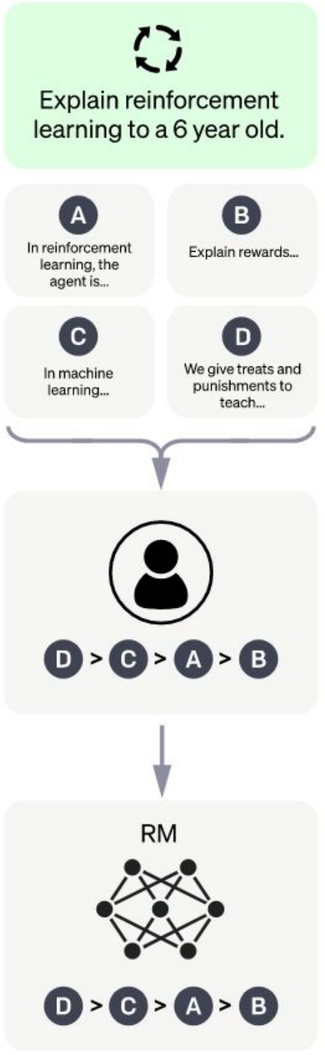
Step 2

Collect comparison data and train a reward model.

A prompt and several model outputs are sampled.

A labeler ranks the outputs from best to worst.

This data is used to train our reward model.



Step 3

Optimize a policy against the reward model using the PPO reinforcement learning algorithm.

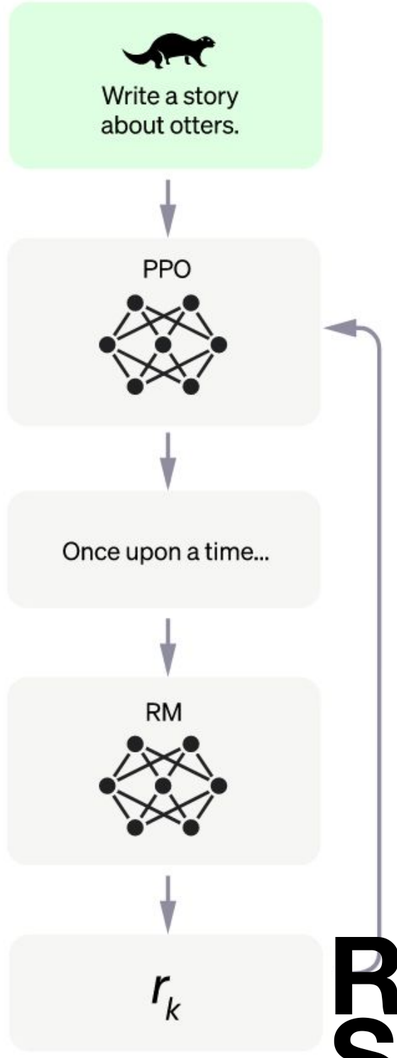
A new prompt is sampled from the dataset.

The PPO model is initialized from the supervised policy.

The policy generates an output.

The reward model calculates a reward for the output.

The reward is used to update the policy using PPO.



https://openai.com/blog/chatgpt/

RI. SE

Features (according to OpenAI)

- Can respond to follow-up questions
- Can admit mistakes
- Can challenge incorrect premises
- Can reject inappropriate requests

RI.
SE

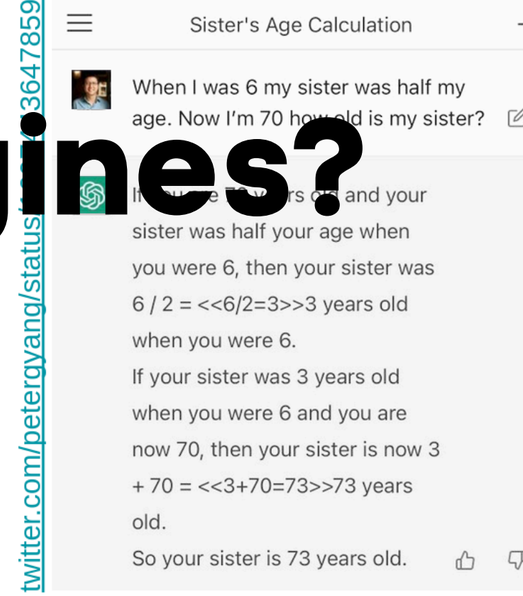
Content moderation

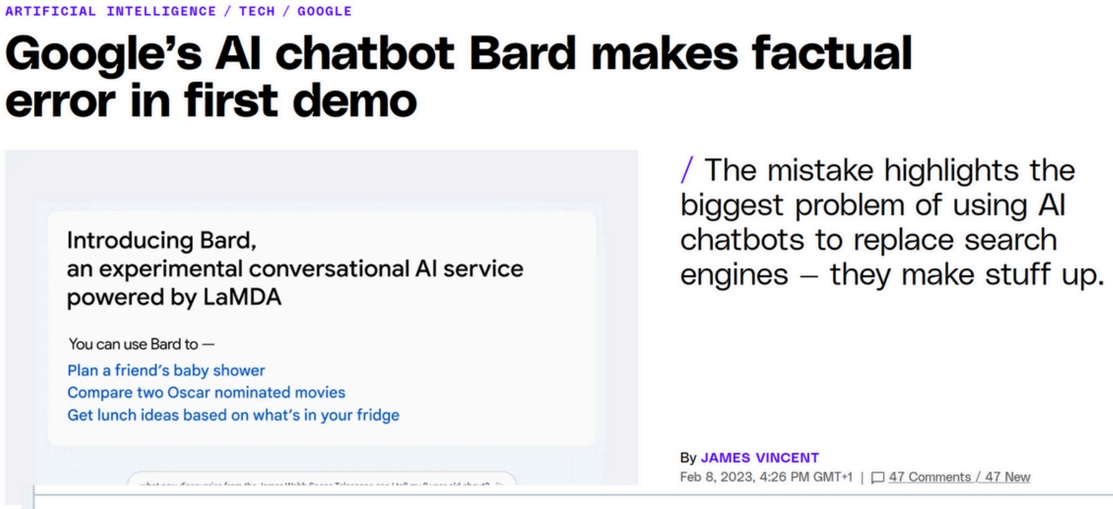
- “reductions in harmful and untruthful outputs” (RLHF)
- “Content moderation tooling”
  - Classifies content for
    - Violence
    - Self-harm
    - Hate
    - Sexual
  - Will decline to respond if question or answer is classified

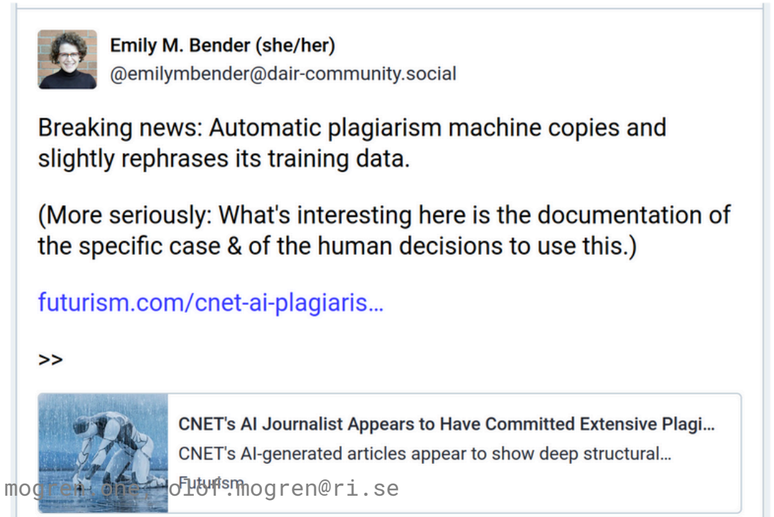
RI.SE

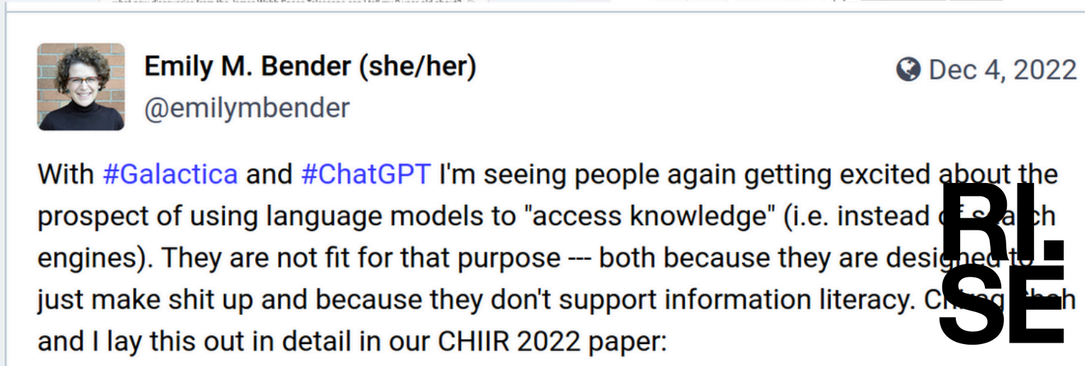
Can this replace search engines?

- “Know-it-all” - incorrect but confident
- Over confidence in answers
- Information litteracy
  - ChatGPT makes things up
  - Even references may be made up
- LLMs are already being used in search





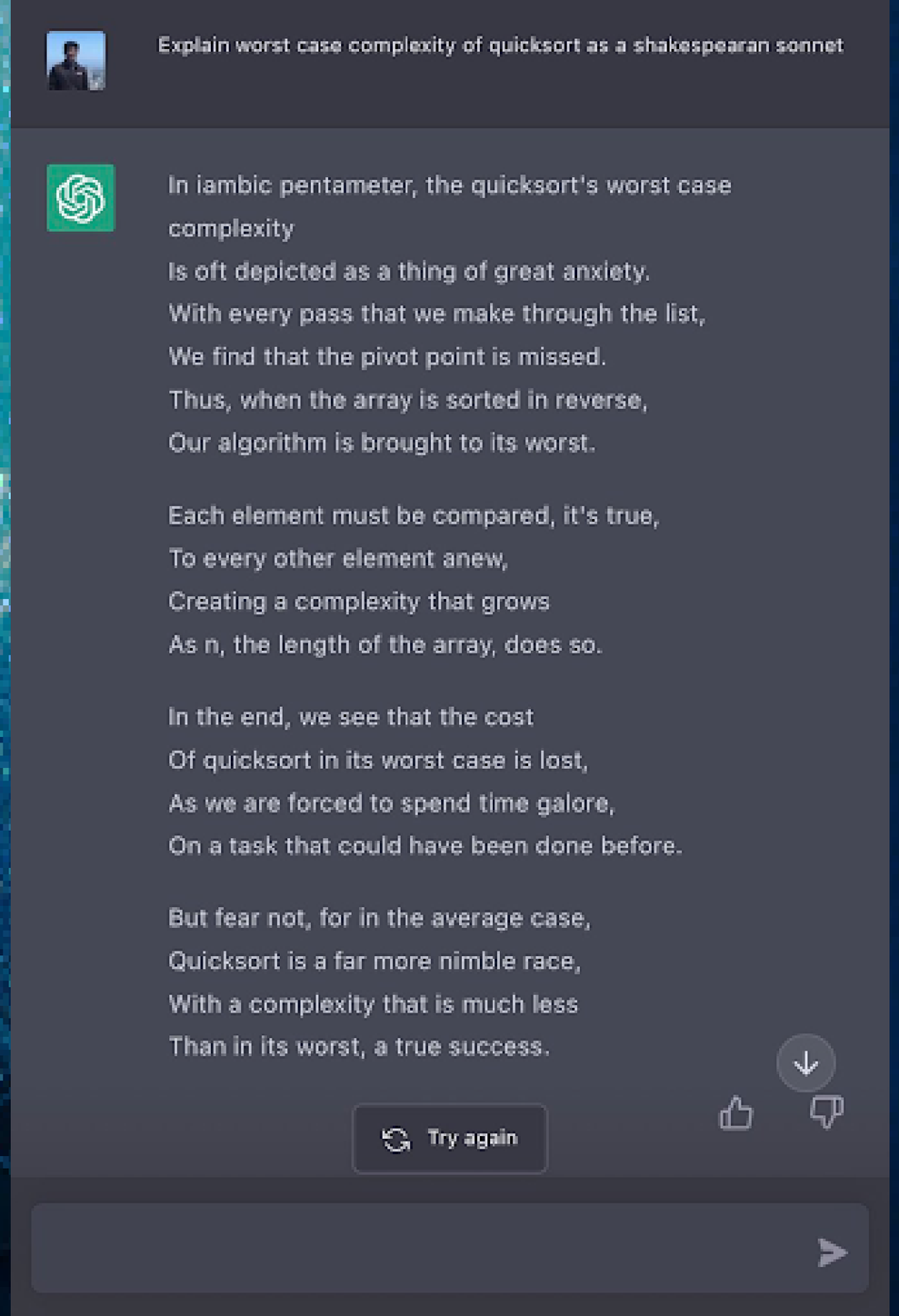




RI.SE

Why is this cool?

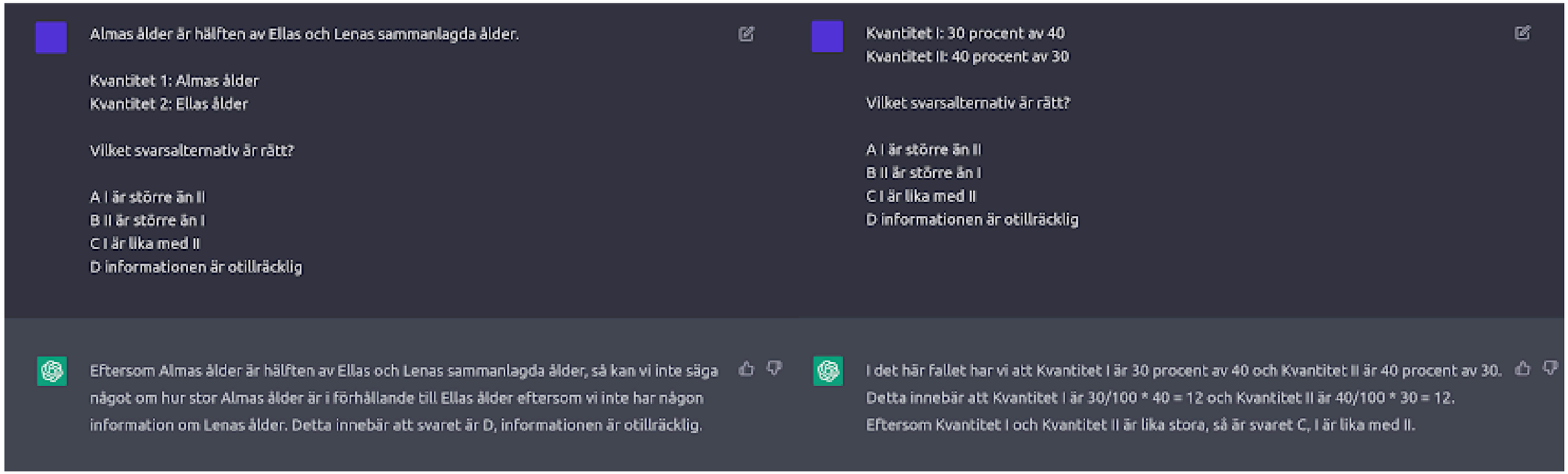
- Impressive engineering job
- Impressive fluency
- Scaling - laguage perception and abstraction ability increase with larger data and model
- Context
- Tracks a long sequence in memory
- Abstraction; finds the important
- Finds lines of thoughts
- Can combine content and aspects



mogren.one, olof.mogren@ri.se

RI.SE

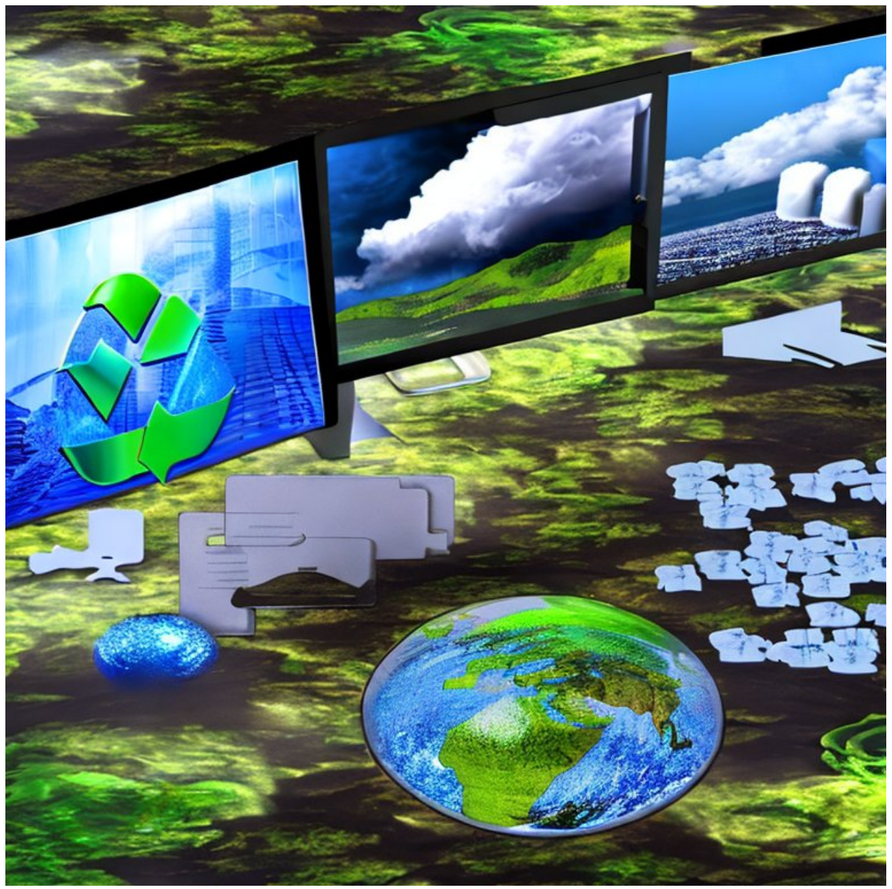
Swedish university entrance exam



Thank you

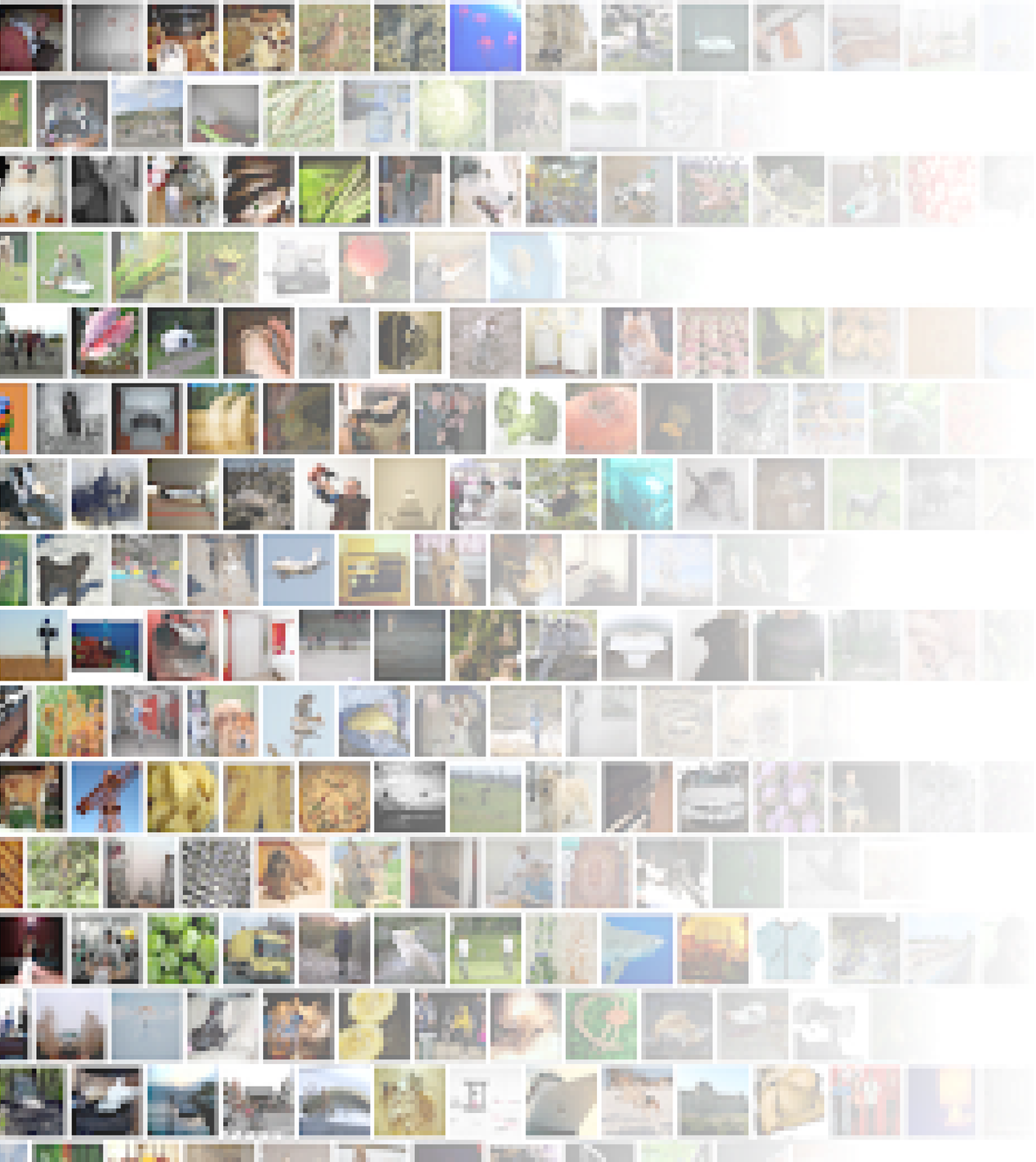
Next hour

- Computer vision
- Machine learning for environmental analysis



mogren.one, olof.mogren@ri.se

RI.
SE



AI
for the environment

mogren.one, olof.mogren@ri.se

RI.
SE

Environment data

- Remote sensing
  - Satellite imagery
  - Drone footage
- Sound recordings
  - Passive acoustic devices
- Video data
  - Camera traps
- Other sensors
  - Water flow sensors
  - Atmospheric measurements

Remote sensing

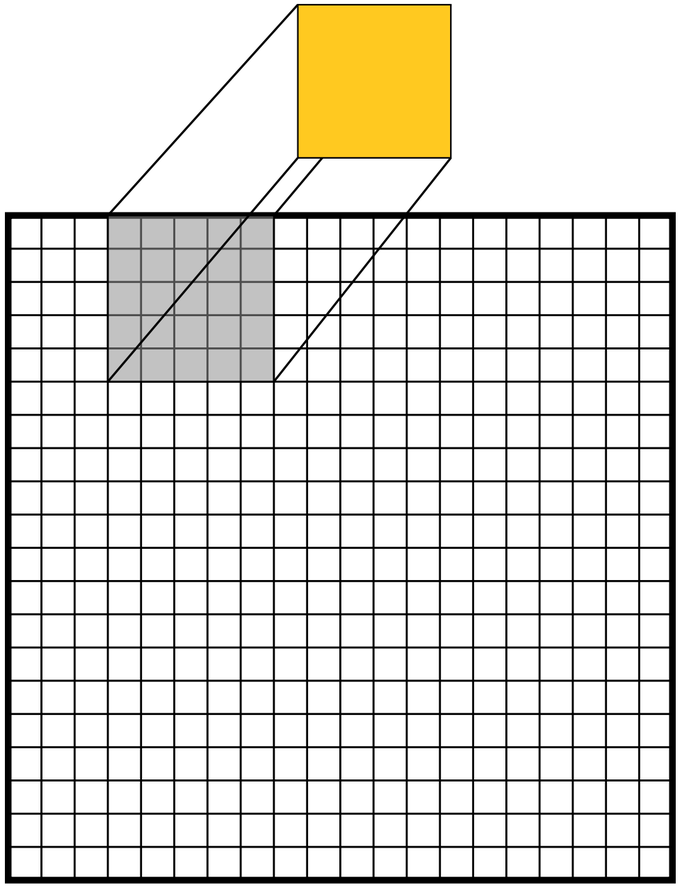
- Computer vision tasks,
  - Classification
  - Semantic segmentation
  - Object detection
  - Regression

Stable Diffusion 2.1: Sattelite in earth orbit.

mogren.one olof.mogren@ri.se

RI.SE

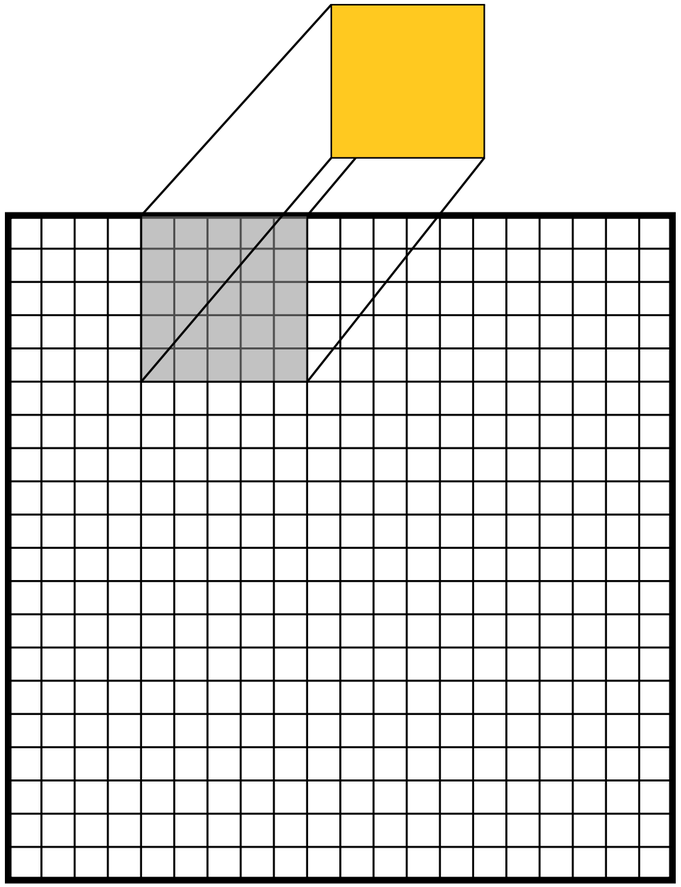
Deep learning
for images



- Convolutional neural networks
- Weight 'filters' slides across input
- Retains spatial dimensions

RI.
SE

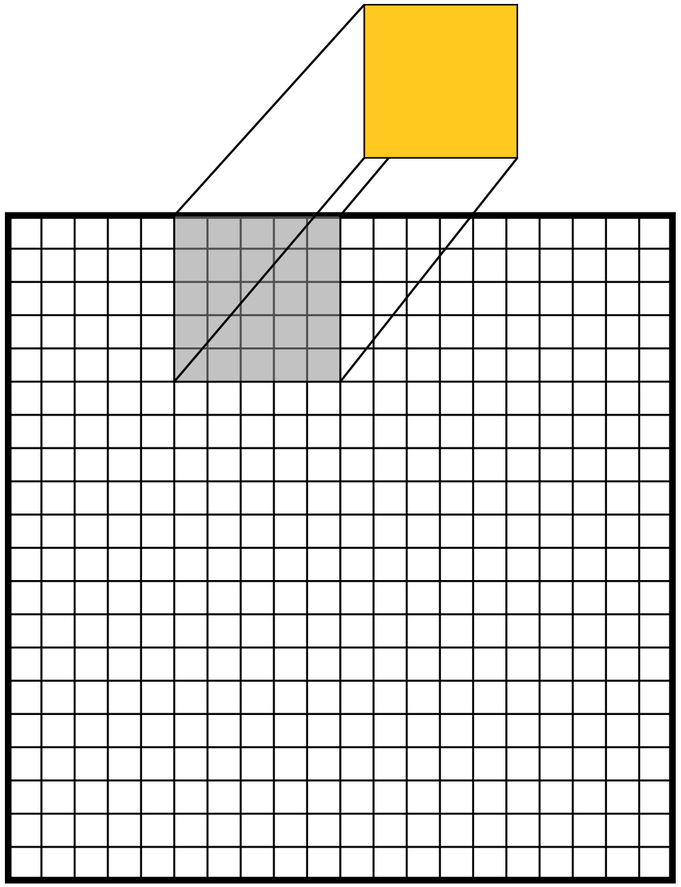
Deep learning
for images



- Convolutional neural networks
- Weight 'filters' slides across input
- Retains spatial dimensions

RI.
SE

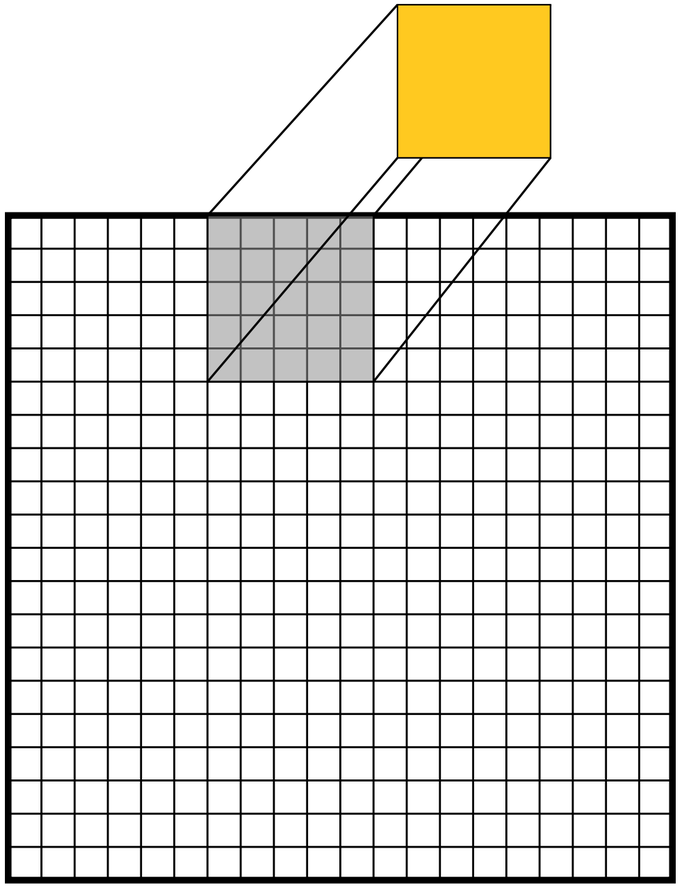
Deep learning
for images



- Convolutional neural networks
- Weight 'filters' slides across input
- Retains spatial dimensions

RI.
SE

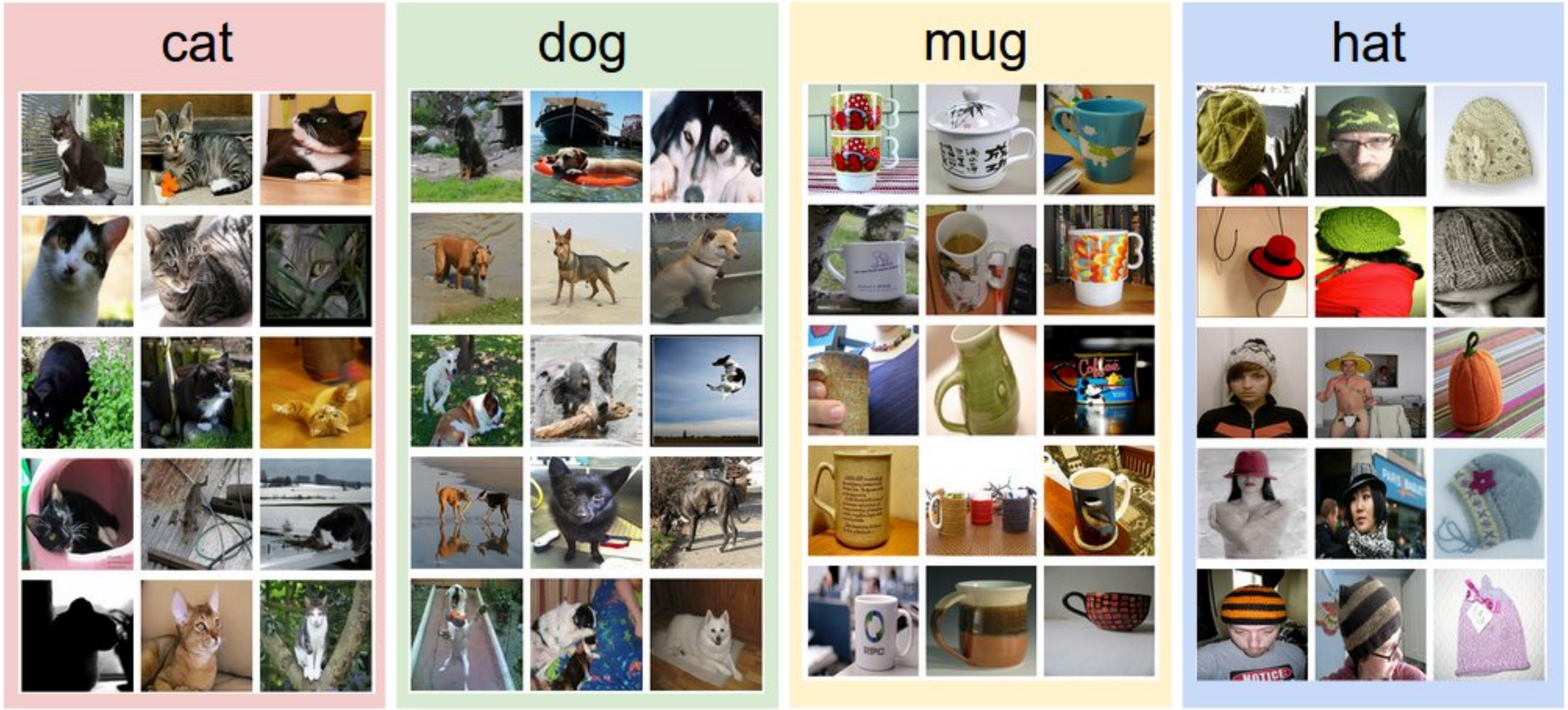
Deep learning
for images



- Convolutional neural networks
- Weight 'filters' slides across input
- Retains spatial dimensions

RI.
SE

Image classification



RI.
SE

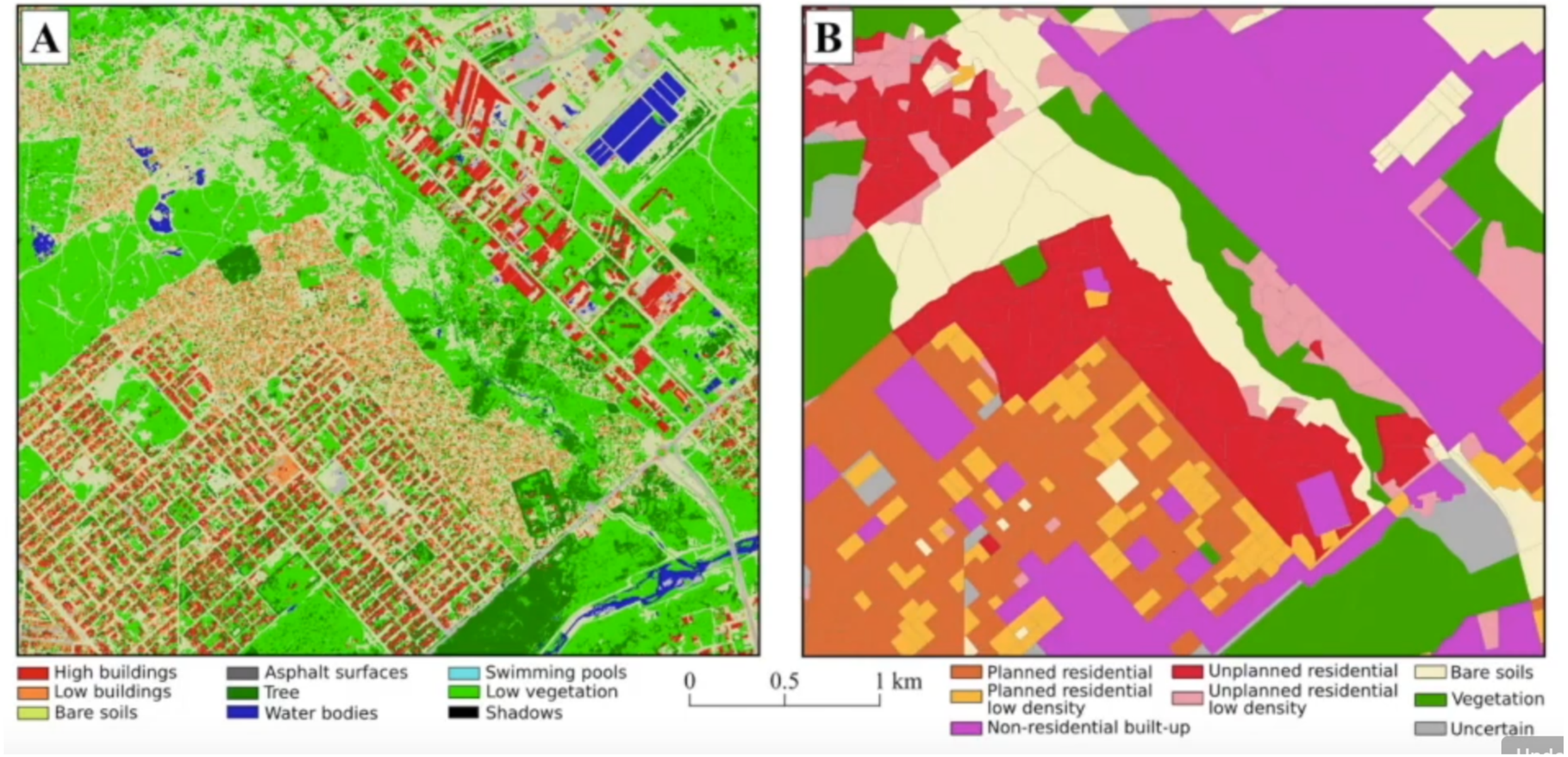
Semantic segmentation

RI.
SE

Generative AI

RI.
SE

Land cover and land use



WORKFLOW

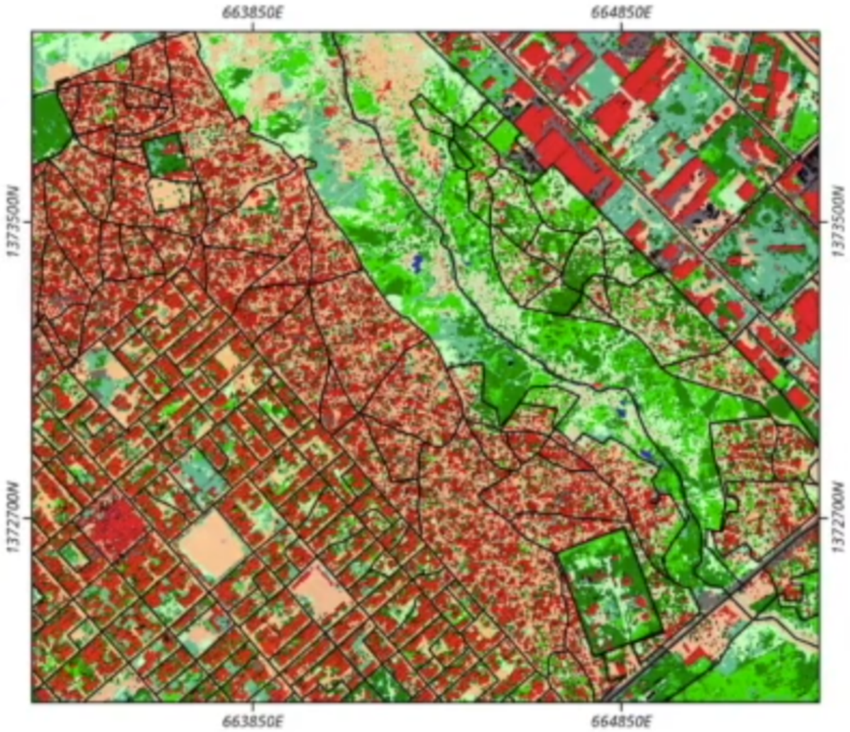
Image

↓

Land cover map

↓

Land use map
at street block level



WORKFLOW

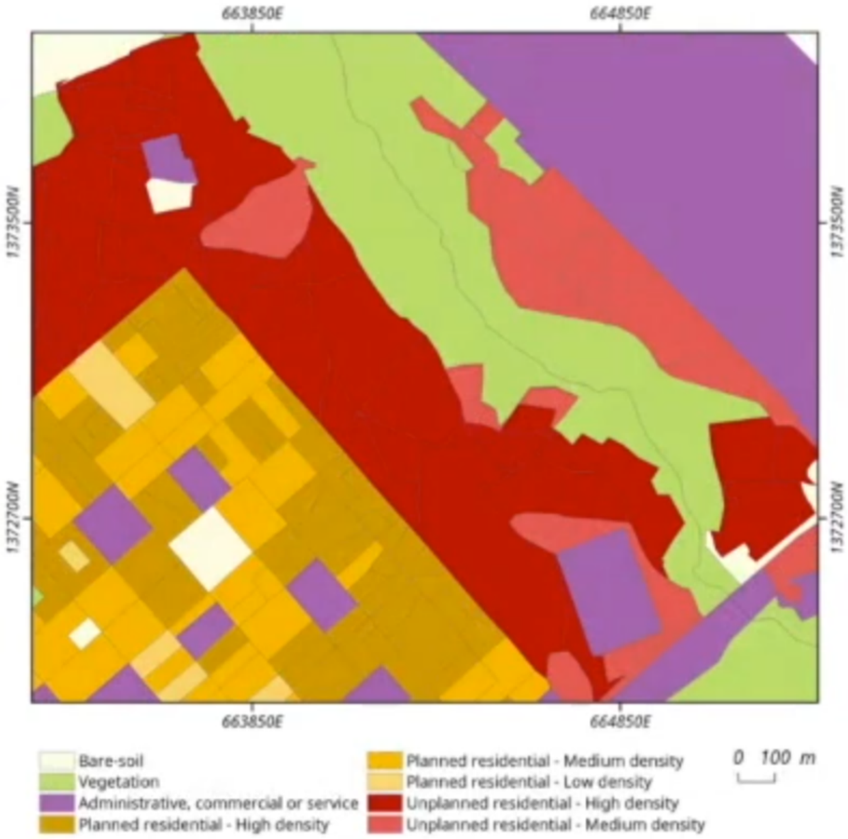
Image

↓

Land cover map

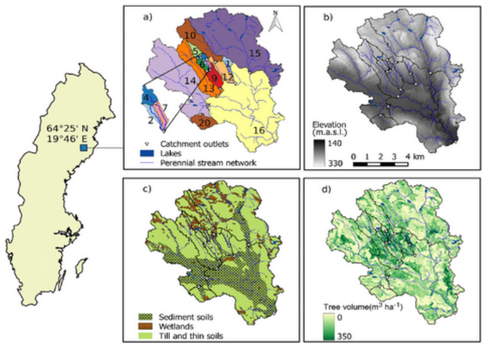
↓

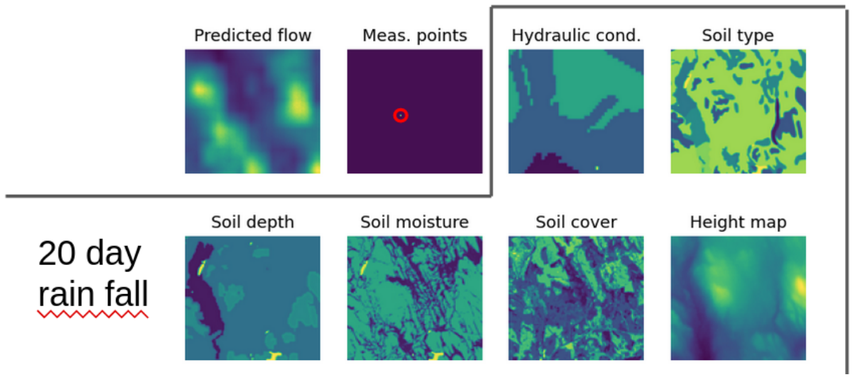
Land use map
at street block level



Stream flow forecasting

- Collaboration with University of São Paulo
- Dense predictions of water flow





Pirinen, Mogren, Västerdal, 2023

Active vision

- UAVs in environmental applications
- Constraints
  - Flight time
  - Data storage
  - Cost and difficulty for humans
- Active collection
  - Choose where to fly
    for optimal data collection

Example: How to efficiently use a UAV for tree health monitoring, to ensure an accurate inference while simultaneously reducing flight time?

Pirinen et al., 2023

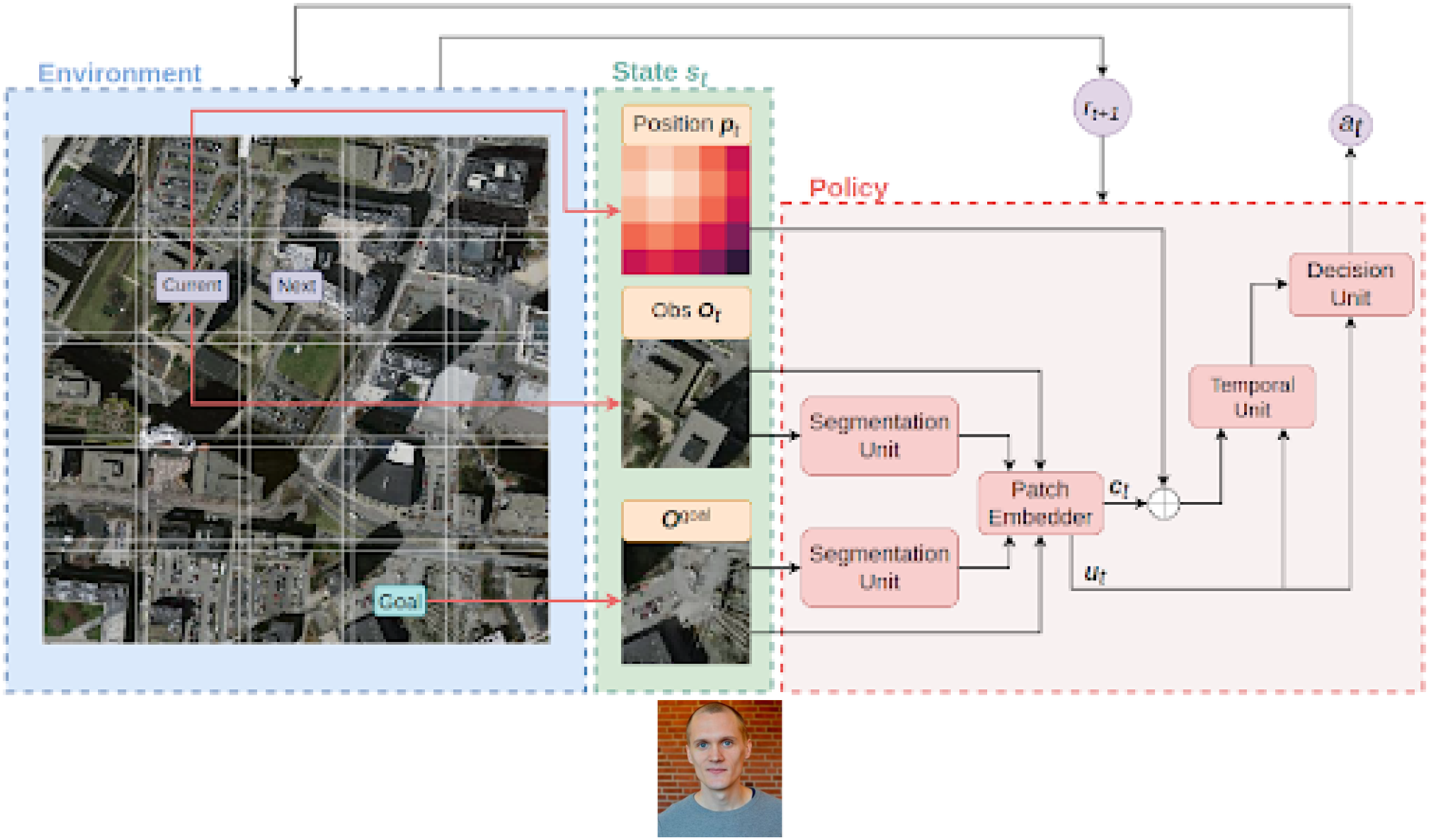
RI.
SE

Active vision

Pirinen et al., 2023

RI.
SE

Active vision



Pirinen et al., 2023

RI.
SE

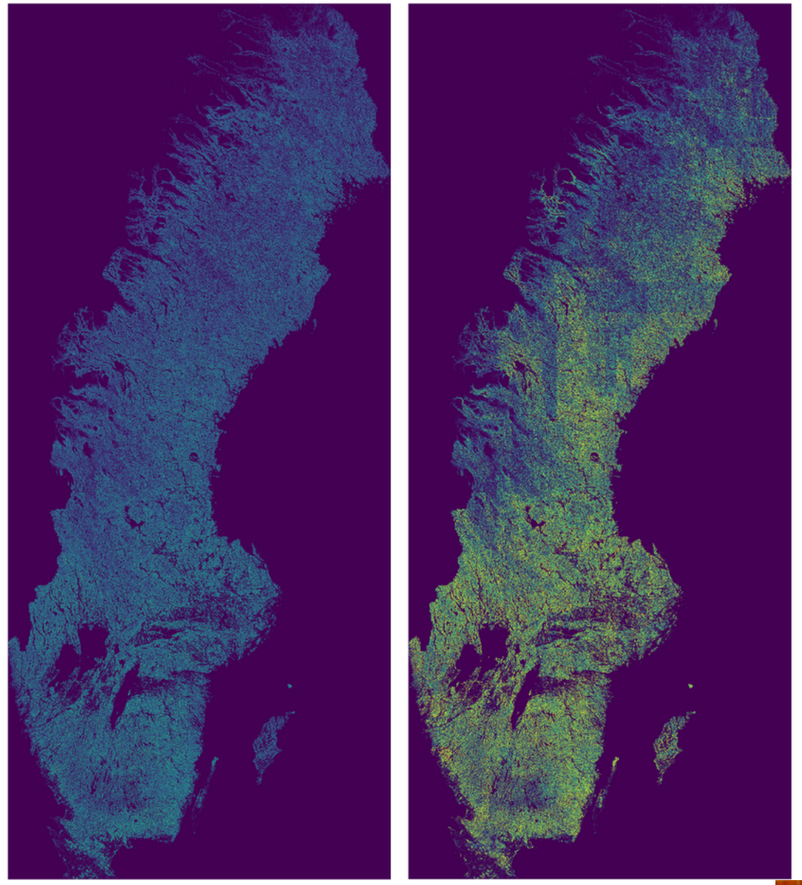
Coffee berry disease

- Master project
  - Help detect infected plants
  - Highly dependent on climate change and factors such as rainfall, humidity, and temperature
- Limited data
  - Few raw images **and** few annotations

mogren.one, olof.mogren@ri.se

RI.SE

Wetland estimation



AI-teknik testas för att identifiera våtmarker

2022-10-11 NYHET Nu testas AI-teknik för att komplettera befintliga kartverktyg till Sveriges våtmarksarbete. Med de nya kartorna kan myndigheter sedan prioritera vilka våtmarker som behöver åtgärder. Tekniken testas just nu på uppdrag av Naturvårdsverket.

Naturvårdsverket har tidigare uppmärksammat att det behövs mer detaljerade kartbilder över våtmarkernas naturtyper i hela Sverige. Det behövs bättre kunskap om var våtmarker finns, deras status samt vilka natur- och kulturvärden de har. Endast hälften av våtmarksarealen är inventerad i Sverige och det är enbart våtmarker i skyddade områden, som till exempel naturreservat, som är inventerade.

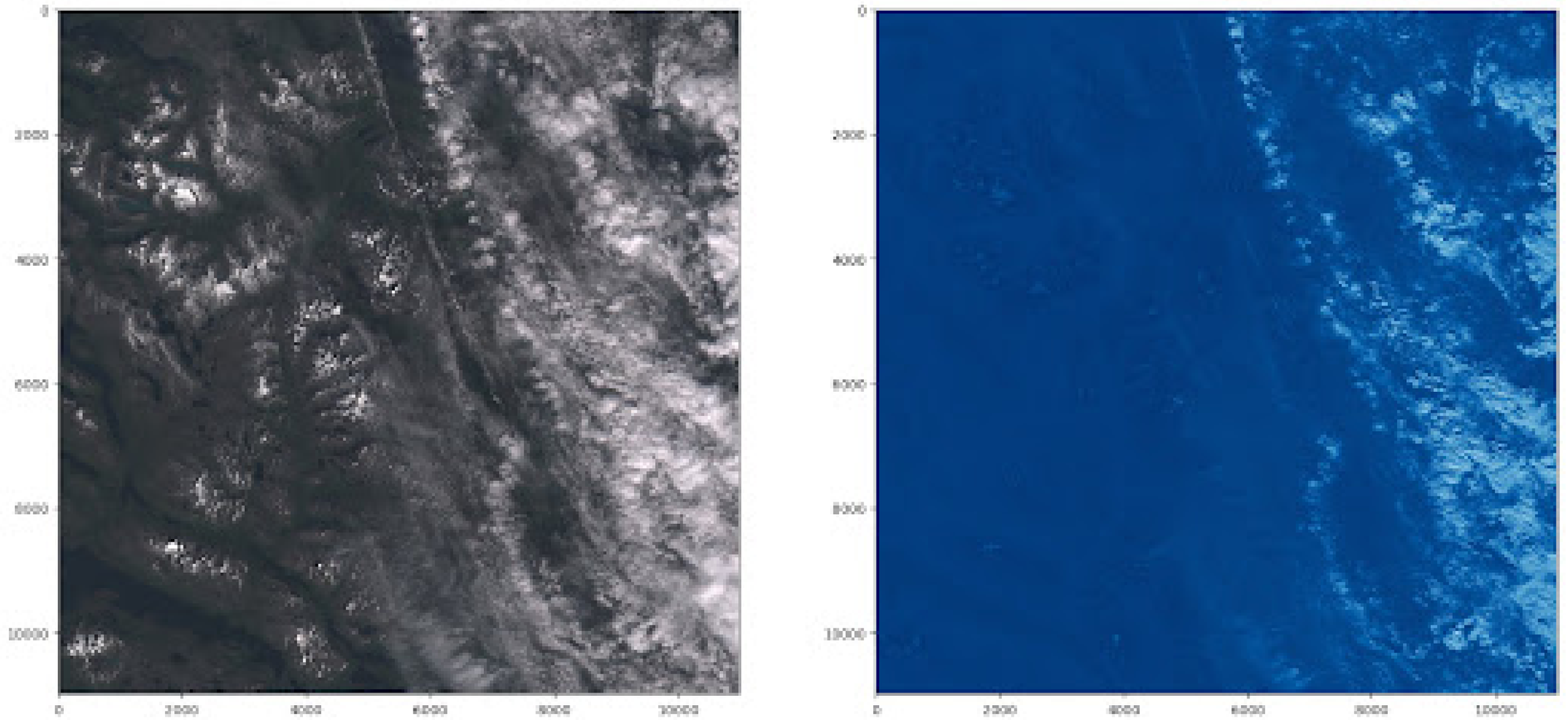
– Vilken status som våtmarkerna har utanför de skyddade områdena är kunskap som både vi och länsstyrelserna behöver för att prioritera arbetet där det gör mest nytta. Bland annat behövs fler åtgärder för att stärka våtmarker med viktiga naturtyper för både biologisk mångfald och klimat, som till exempel rikkärr och högmossar, förklarar **Matti Ermold**, handläggare på Naturvårdsverket.

Naturvårdsverket har lämnat ett uppdrag åt forskningsinstitutet RISE att undersöka om det går att komplettera befintliga digitala kartverktyg. De nya bilderna ska identifiera våtmarkernas naturtyper utanför skyddade områden i Sverige. Med hjälp av AI-teknik ska projektet analysera och beräkna datamängder från andra kända kartor samt rita upp och testa träffsäkerheten.

https://www.naturvardsverket.se/om-oss/aktuellt/nyheter-och-pressmeddelanden/ai-teknik-testas-for-att-identifiera-vatmarker/

Pirinen et al., 2023

Cloud thickness estimation

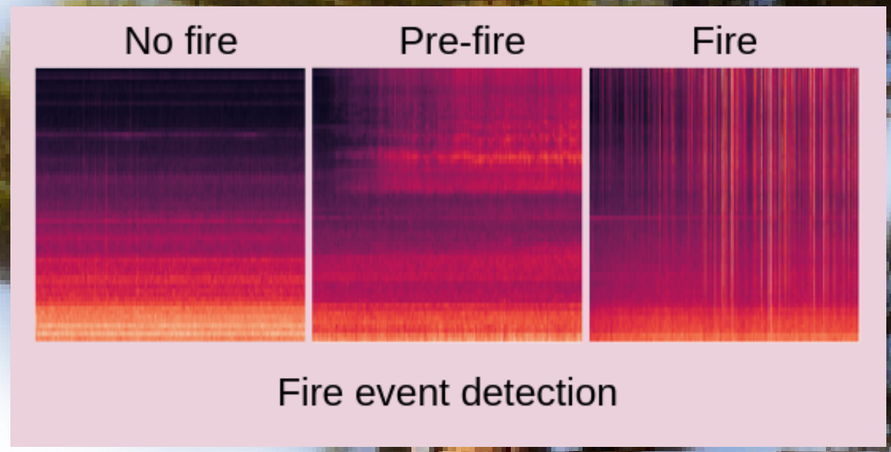


Pirinen et al., 2023

RI.
SE

Machine listening

- Environment
  - Weather events
  - Biodiversity markers
- Fire detection
  - Good results in lab settings
- General
  - Sound source separation
  - PhD project

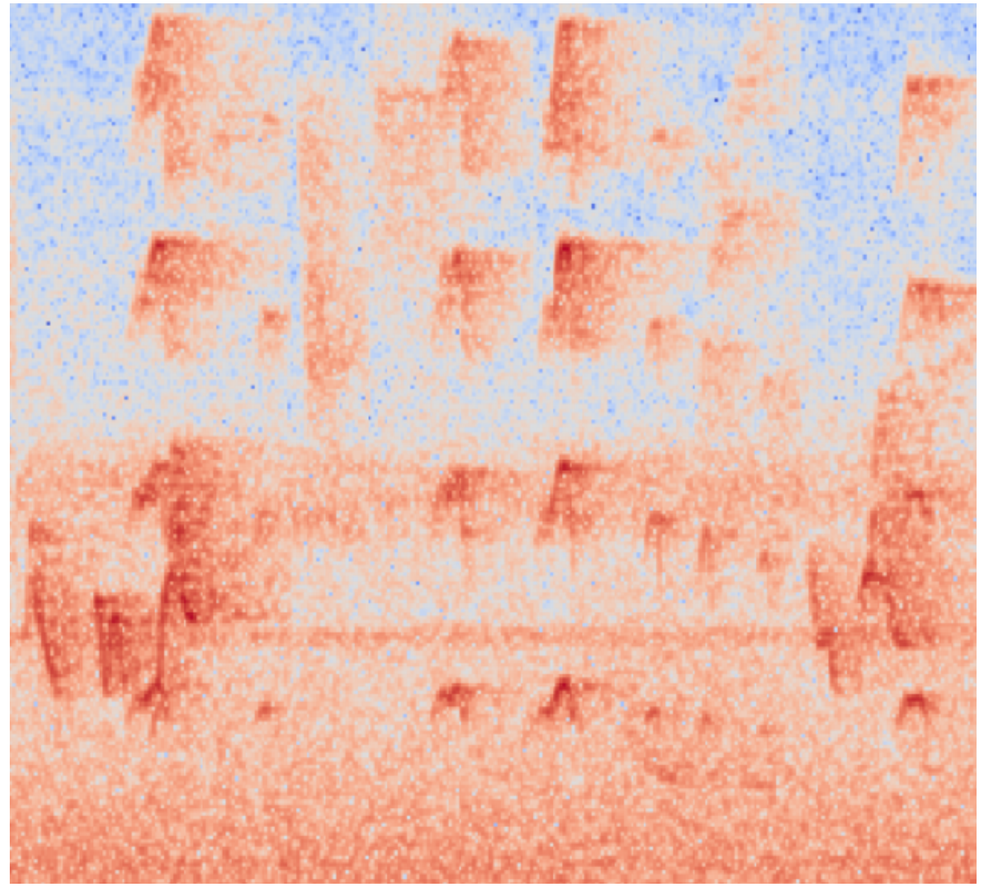




mogren.one, olof.mogren@ri.se

Bird species identification

- Modelling sound using spectrograms and convolutional neural networks
- Altitude information improved results
- Location information future work



RI.
SE

Sound event detection

- Large amounts of data
- Labour intensive annotation
- Few-shot learning

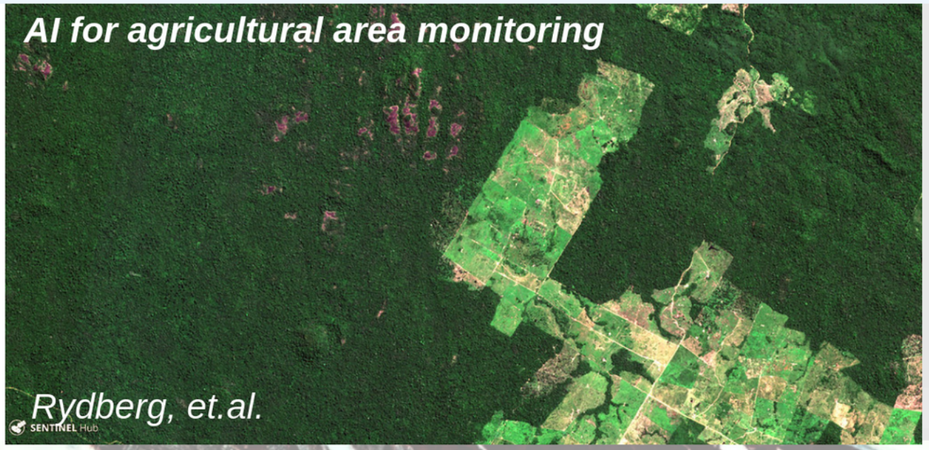
mogren.one, olof.mogren@ri.se

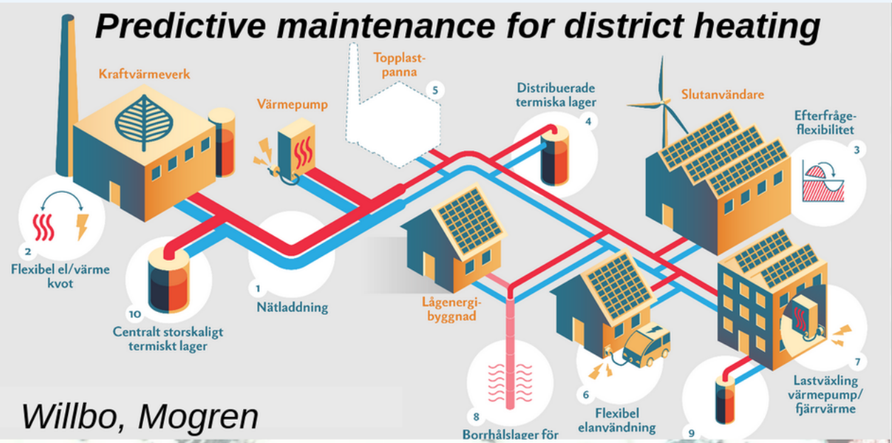
RI. SE

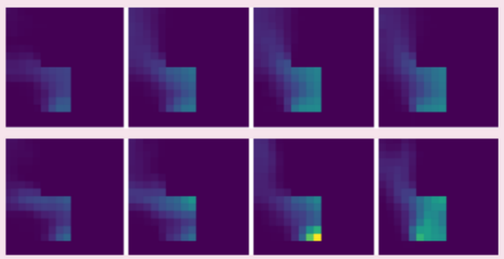
More related activities at RISE

AI for circular business models

Listo Zec, Mogren







CHALMERS UNIVERSITY OF TECHNOLOGY

Physics simulations, Lam & Mogren

*Martinsson, Listo Zec, Gillblad, **Mogren.** Adversarial representation learning for synthetic replacement of private attributes. IEEE Big Data 2021 https://arxiv.org/abs/2006.08039, 2020.*

*Onoszko, Karlsson, **Mogren**, Listo Zec. Decentralized federated learning of deep neural networks on non-iid data. 2021 FLUPDC workshop at ICLM, https://arxiv.org/abs/2107.08517*

*Martinsson, J., Schliep, A., Eliasson, B., **Mogren, O.**, Blood glucose prediction with variance estimation using recurrent neural networks.Journal of Healthcare Informatics Research. 2020. http://mogren.one/publications/2019/blood/*

RI. SE

So, how do I start?

RI.
SE

Getting started

- Study your data
- Identify target test set
- Investigate what kind of model would fit
- Find existing model or develop custom
- Evaluate on test set; compare to baseline

RI.
SE

RISE Learing machines seminars

Every thursday at 15:00

- 2023-05-25: Puzhao Zhang, KTH
- More talks after summer break!

mogren@chalmers.se, olof.mogren@ri.se

Thank you

mogren.one, olof.mogren@ri.se