Symbolic & Sub-Symbolic AI: Co-exist or Combine?

Frank van Harmelen

Learning & Reasoning Lab

Vrije Universiteit Amsterdam

Symbolic & Sub-Symbolic AI: Compete or Combine?

Frank van Harmelen

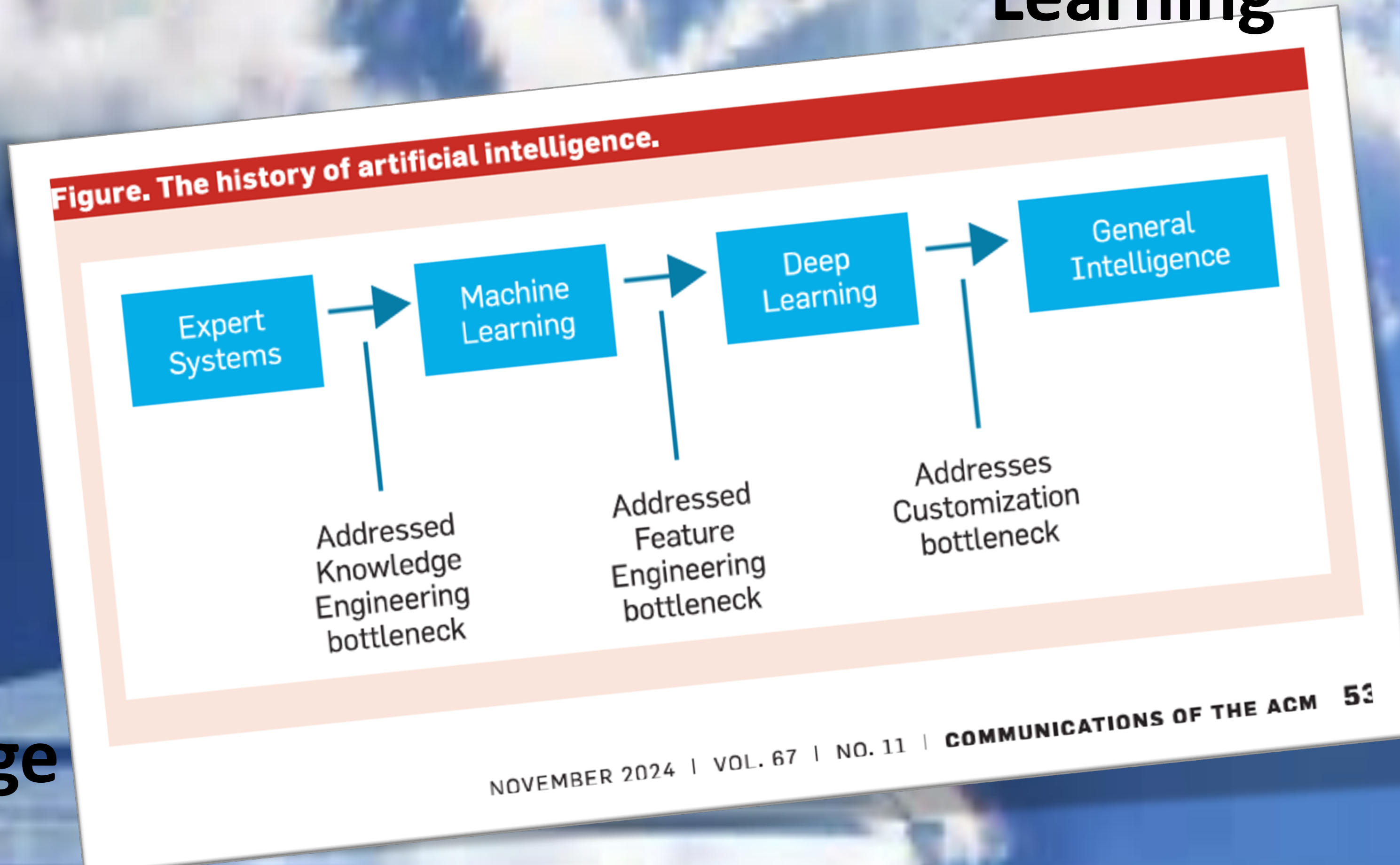
Learning & Reasoning Lab

Vrije Universiteit Amsterdam

Historical & current directions of travel in AI

History?

Connectionist
Data
Statistics
Learning



Symbolic
Knowledge
Logic
Reasoning

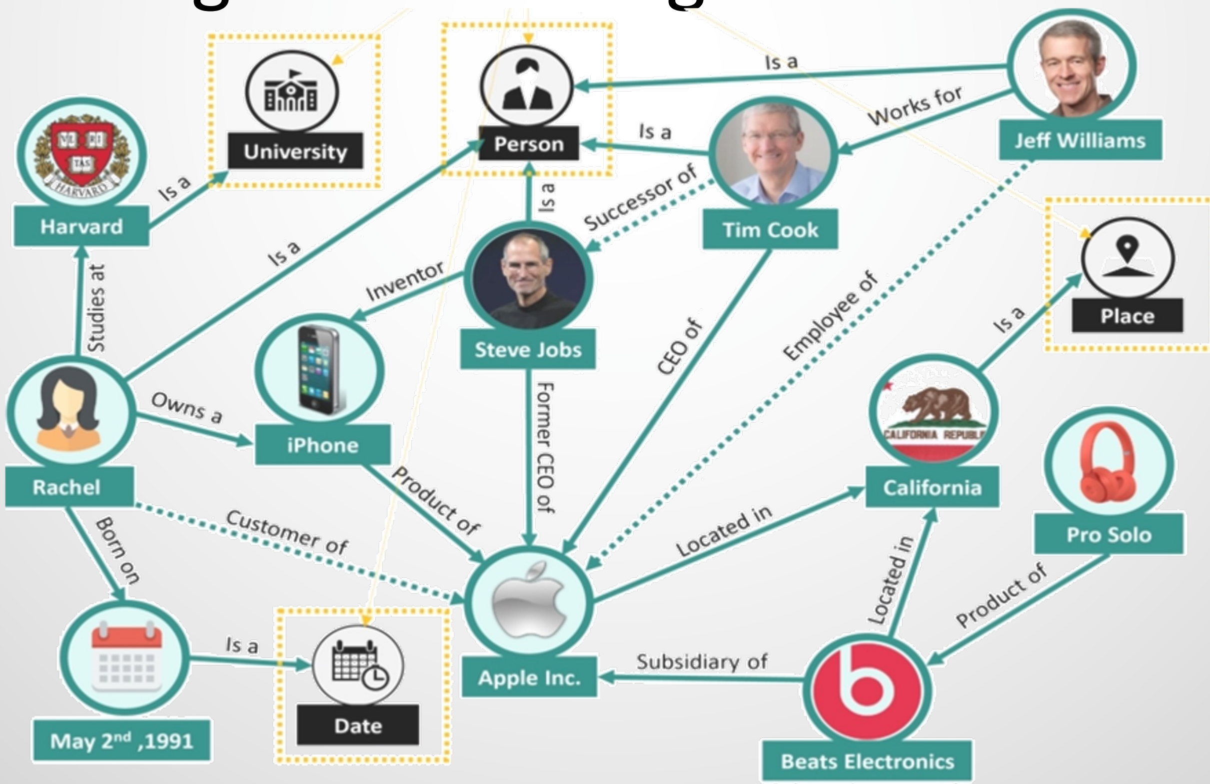
The AI pendulum!

Direction of travel in modern AI

1. Major *breakthroughs* in **machine learning**
2. Major *scale-up* in **knowledge representation**
3. **Complementary** strengths & weaknesses

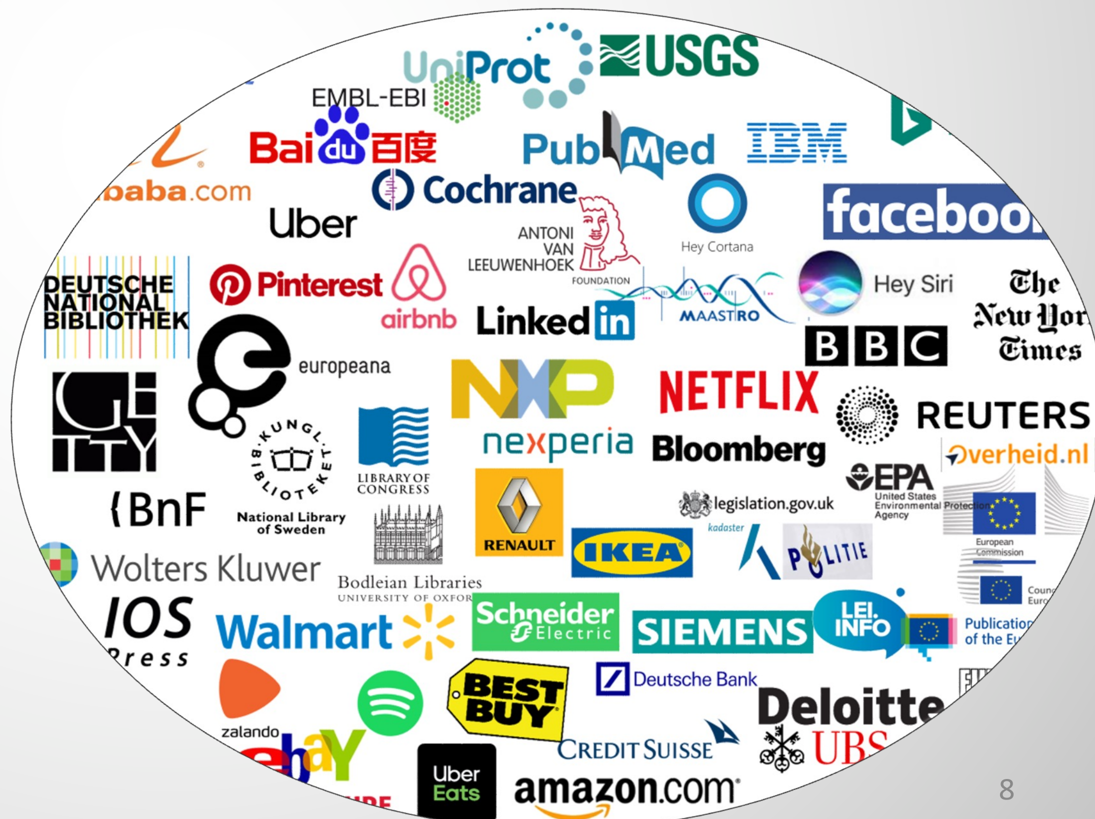
2. Major *scale-up* in **knowledge representation**

- **Knowledge graphs** as networks of relations between things
- Enables logical reasoning



2. Major *scale-up* in **knowledge representation**

- Efficient querying and reasoning
- Scales to 10^{10} relations between 10^9 things
- In use at major industries



Direction of travel in modern AI

1. Major *breakthroughs* in **machine learning**
2. Major *scale-up* in **knowledge representation**
3. **Complementary** strengths & weaknesses

	Symbolic	Statistical
Construction cost	Human effort	Data hunger
Scalable	Worse with more data	Worse with less data
Explainable	+	-
Generalisable	Performance cliff	Out of sample

Strengths & Weaknesses

	Symbolic	Statistical
Construction	Human effort	Data hunger
Scaleable	Worse with more	Worse with less
Explainable	+	-
Generalisable	Performance cliff	OOD data

- Knowledge acquisition bottleneck
  vs. data acquisition bottleneck
- Combinatorial explosion
  vs. sampling inefficiency
- Explainability
  vs. black box
- Brittleness
  vs. out-of-distribution data

Neuro-symbolic AI: a glimpse from the factory floor

Working hypothesis of neuro-symbolic AI

- Knowledge-driven & data-driven systems each have their strengths and weaknesses
- These are complementary
- We need combinations of both

This is not a new idea

"20 years"
[Artur d'Avila Garcez
& Luıs Lamb]

"30 years"
[Pascal Hitzler]

Context aware Machine Learning

flowe… cu…n?

P(cushion|chair) >> P(flower|chair)

“Given the context of chair, a cushion is much more likely than a flower”

Also: semantic loss-function

$\forall x, y$ cha… partOf$(y, x) \rightarrow$ cushion$(y) \vee$ armRest(y)

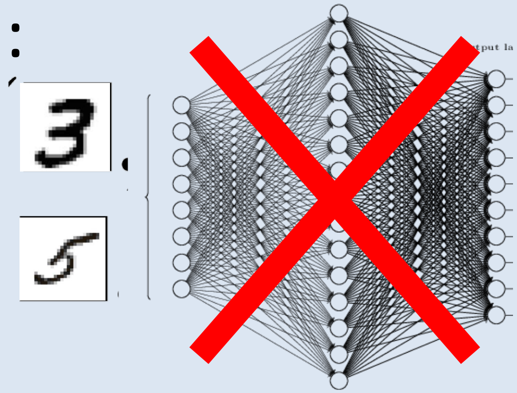
“Parts of a chair are: cushion and armrest”

See survey of 100+ systems in Von Rueden et al, *Learning, 2019*

Learning that is robust to change
(through intermediate abstractions)

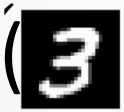
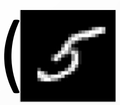
End-to-end:

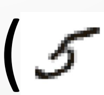
2x 784 inputs



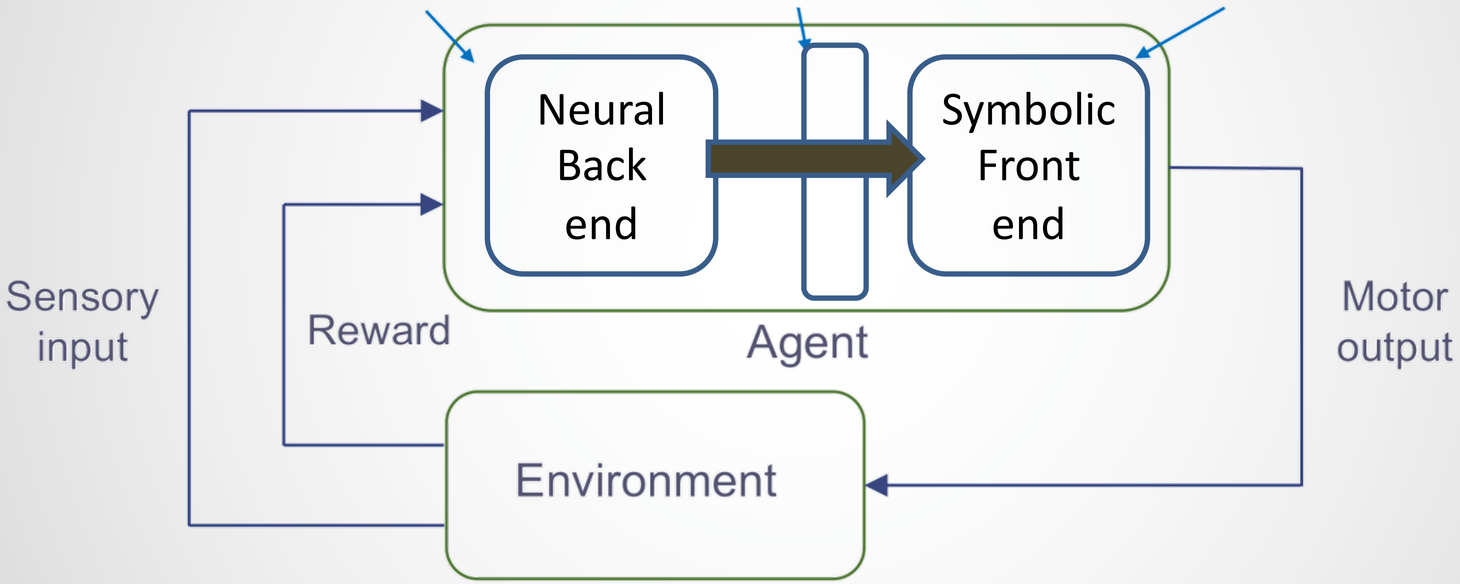
8

19 outputs

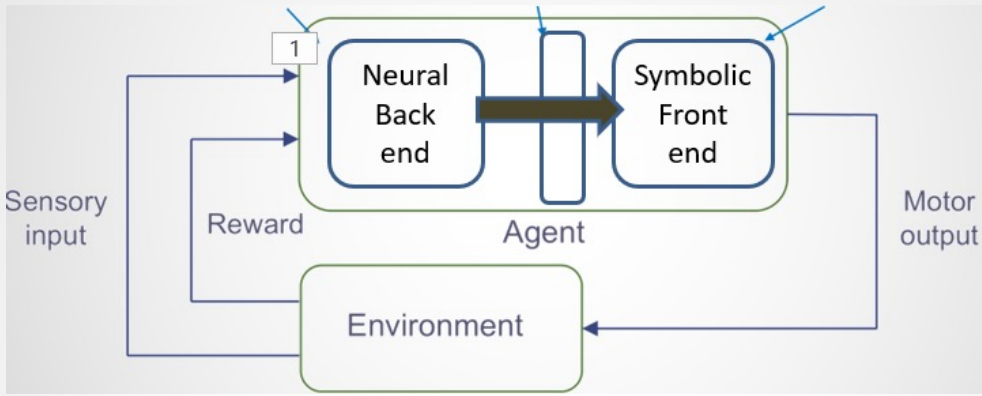
see()=3, see()=5, 3+5=8

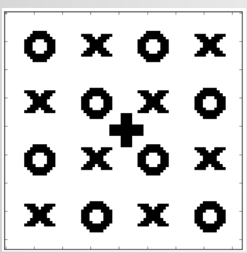
see()=3, **see**()=5, 3+5=8

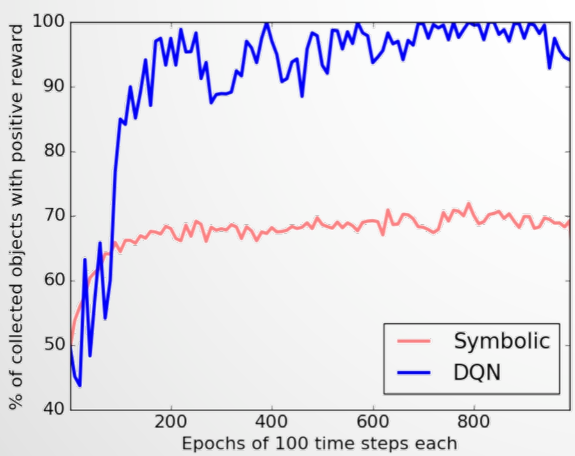
Learning that is robust to change
(through intermediate abstractions)

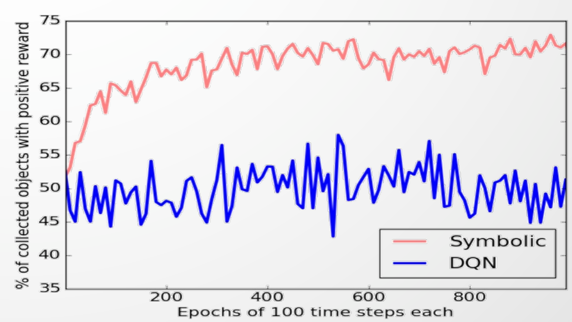


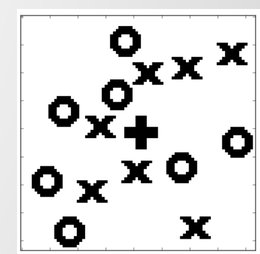
Learning that is robust to change
(through intermediate abstractions)







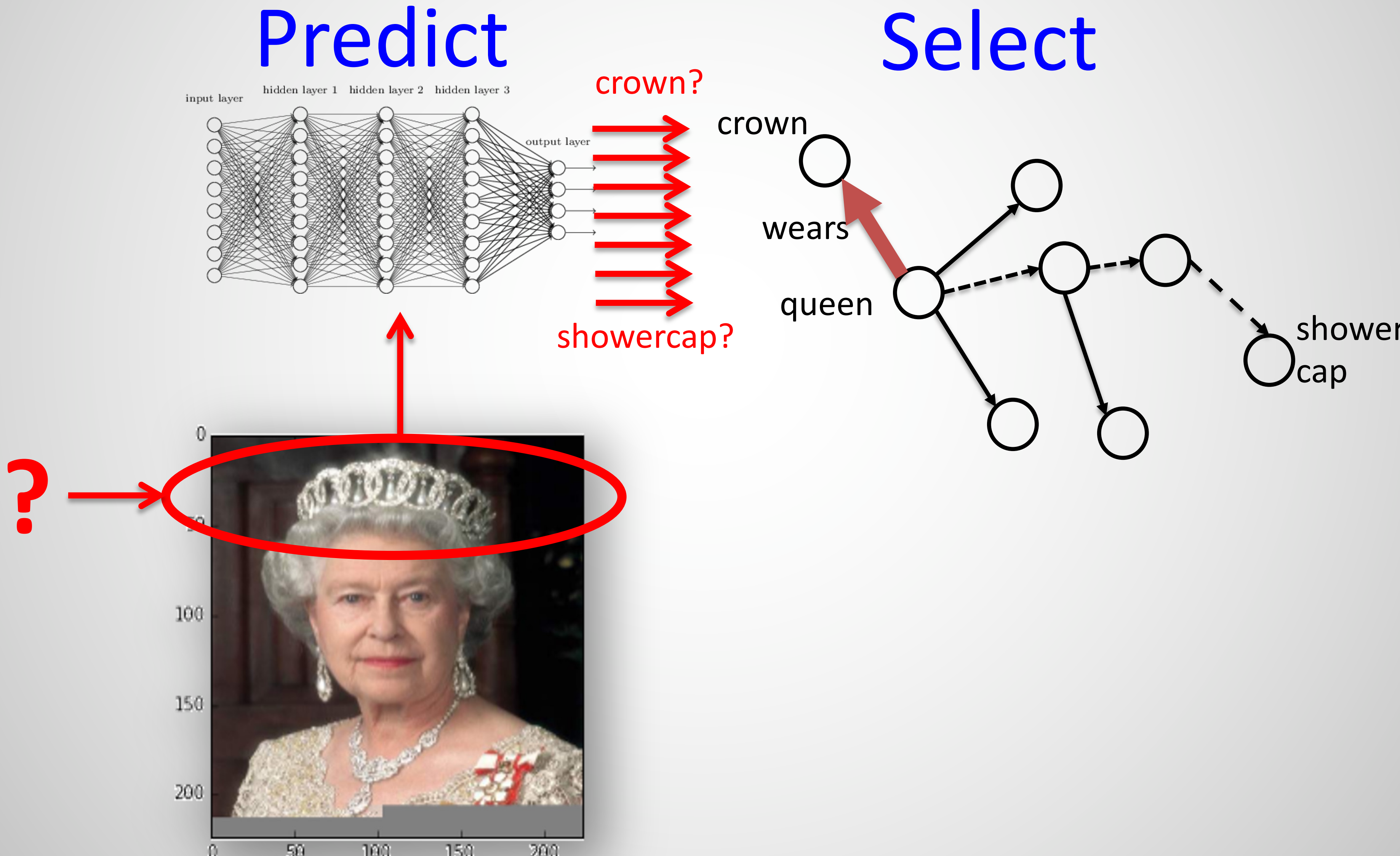




out-of-distribution data

```
class: 793
label: n04209133 shower cap
certainty: 99.7%
```

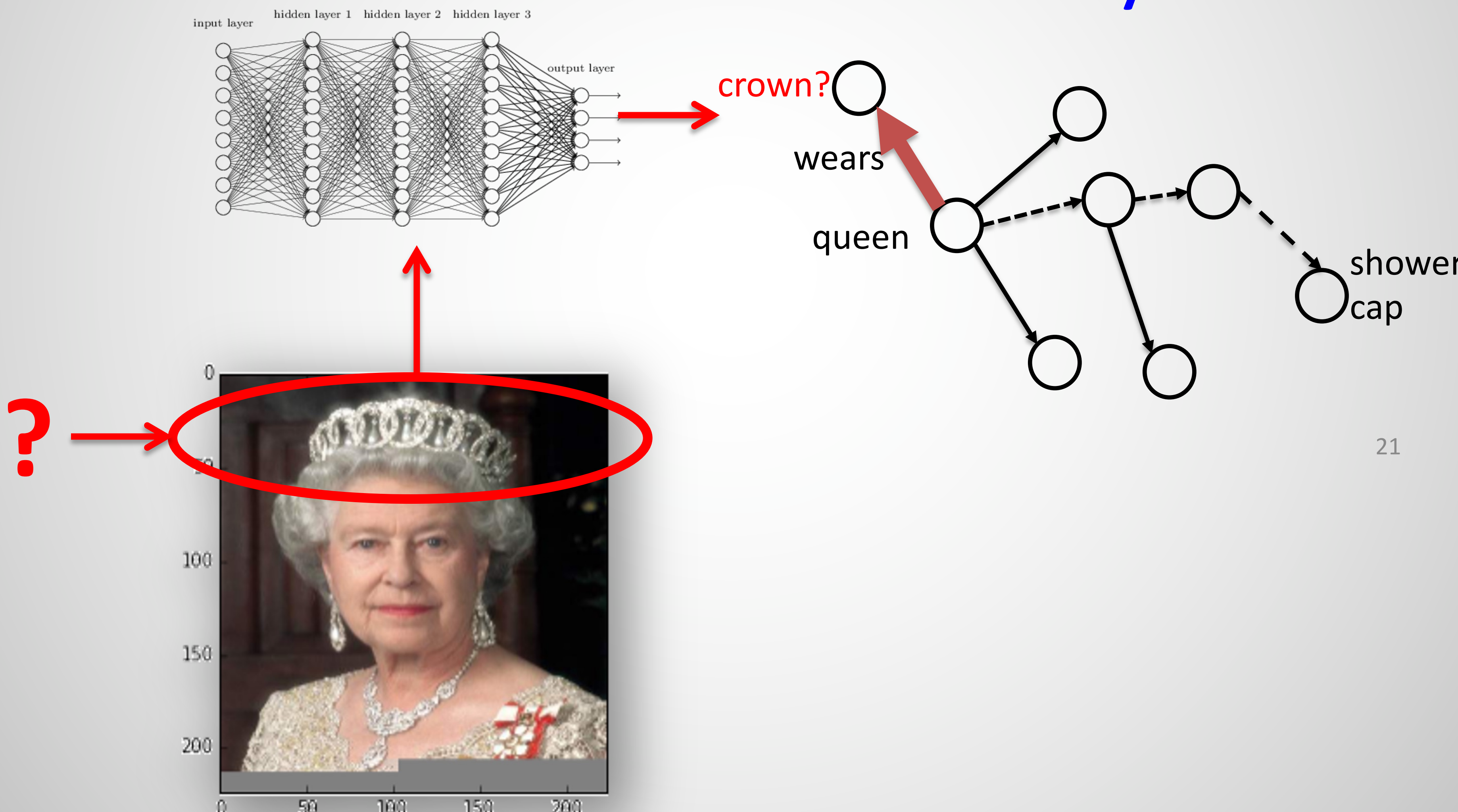
Robust for out-of-distribution data
(by using background knowledge)



Construct justifications

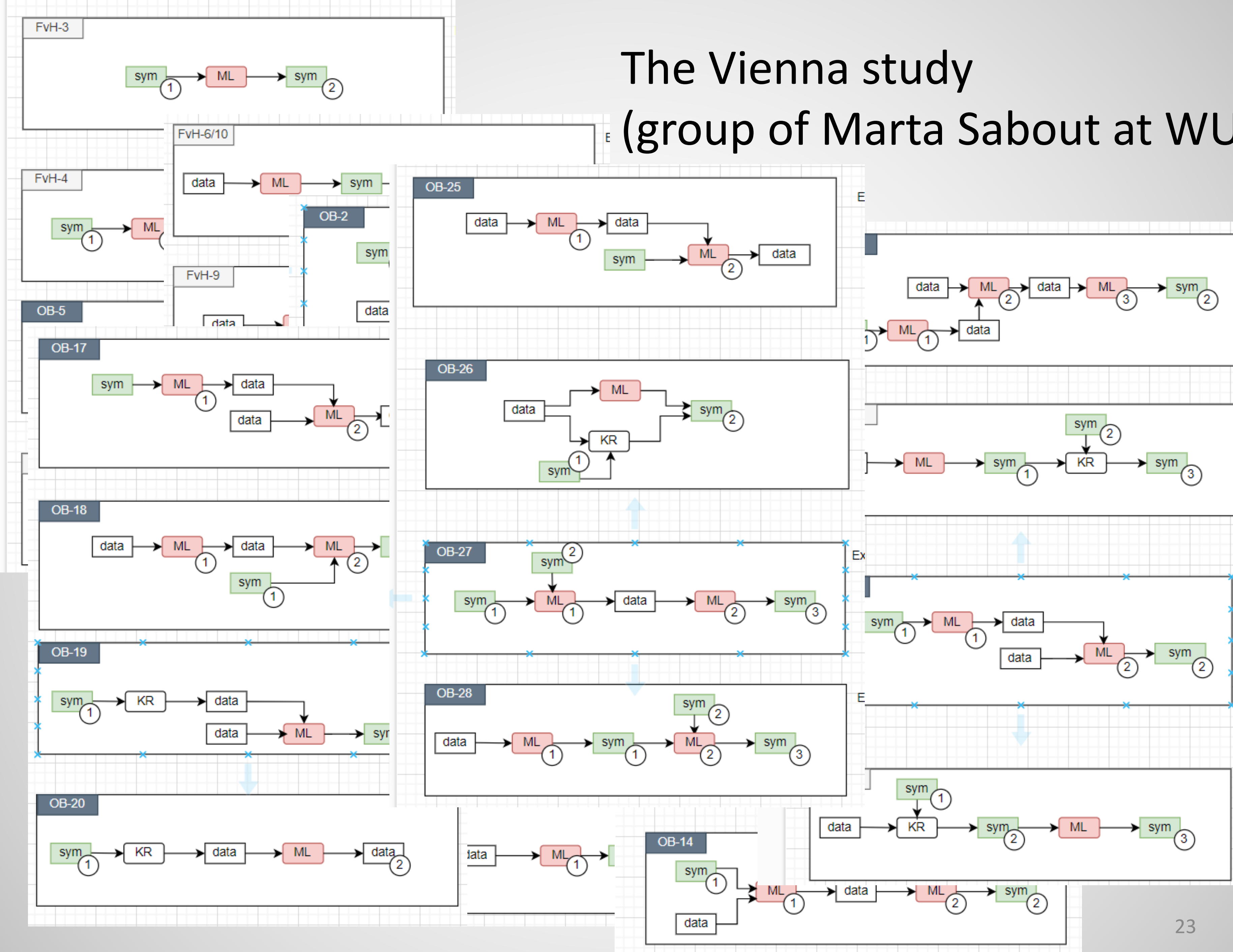
Predict

Justify



Etc…

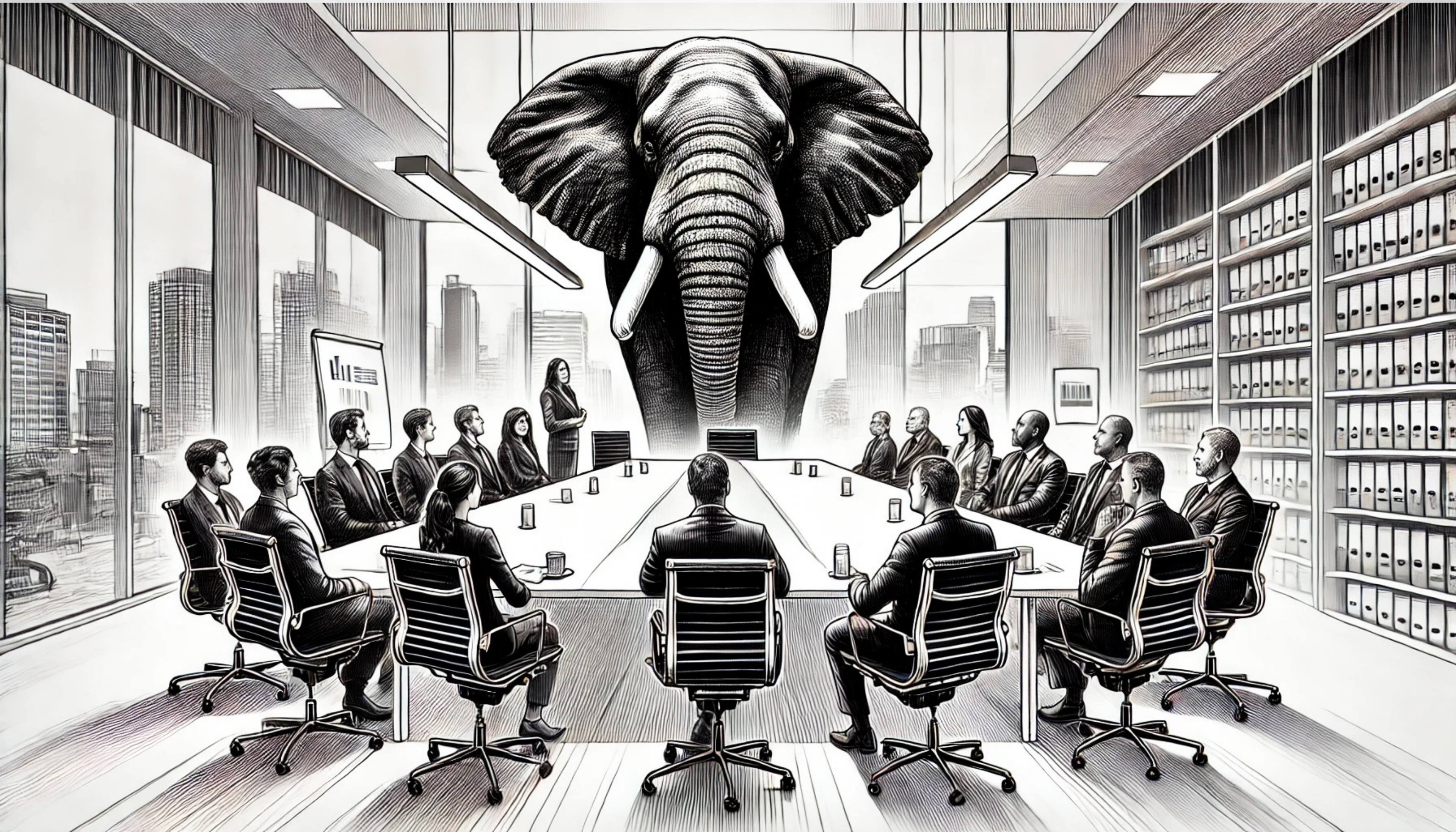
The Vienna study (group of Marta Sabout at WU)



So the neuro-symbolic paradigm enables AI systems that are

- Explainable
- Guaranteed safe
- Aware of context
- Robust against data drift
- Robust against out-of-distribution data
- Effective on a weak signal from small data

The LLM in the room…



The LLM in the room would say:

1. ~~LLMs are the new paradigm in AI~~
2. LLMs don’t need any symbols
   (they have purely neural representations)
3. Therefore we can happily stay
   in the neural paradigm

LLMs are *language* models…

… they are not *world* models

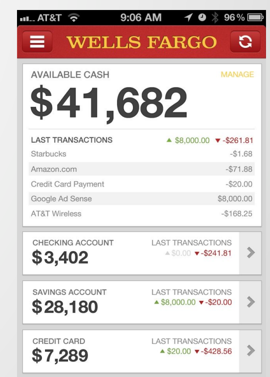
Representational hypothesis in cognitive science:

"cognitive processes are mediated by
internal representations of information.
The mind operates by creating, storing, and manipulating
mental representations of the external world
to guide perception, reasoning,
decision-making, and action."

Neuro-symbolic synthesis:

AI systems will need world models,
and these world models will be

Partially subsymbolic/neural/distributed
(& reasoning = network activations)



Partially symbolic/propositional/discrete

(& reasoning = symbol manipulations)

But:

reasoning about the world ≠ reproducing linguistic patterns

Symbolic & Sub-Symbolic AI: ~~Compete or~~ Combine

Frank van Harmelen

Learning & Reasoning Lab

Vrije Universiteit Amsterdam