TU Wien

2025

"What color am I?"

— Alex the Parrot, 2007

Alex the African Grey

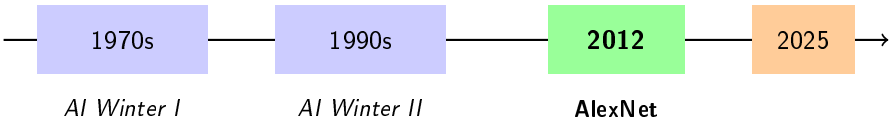
- 30 years of research (Pepperberg 1999)
- 150 words, 100+ objects categorized
- Understood zero (age 4 milestone)
- First non-human to **ask a question**

"Felt without abstracting"

The Stochastic Parrot

- (Bender et al. 2021) definition:
- "Stitches linguistic forms...
- ...without reference to meaning"
- 2023 AI Word of the Year

"Abstracts without feeling"



AlexNet Breakthrough (Krizhevsky et al. 2012):

- Error: 15.3% vs 26.2% (runner-up)
- Two GTX 580 GPUs
- Trained in a *bedroom*

Paradigm shift:

Rules $\rightarrow$ Learning

System	Power	
Human brain	**20 W**	
ChatGPT-3 operation	9 MW	$\times 450{,}000$
GPT-3 training	1,300 MWh	(130 households/year)
Brain simulation	2.7 GW	(nuclear plant)

Blue Brain Project (2013):
1.4 MW for 40 min $\rightarrow$ 1 second of 1% brain

500 million years

VS

70 years

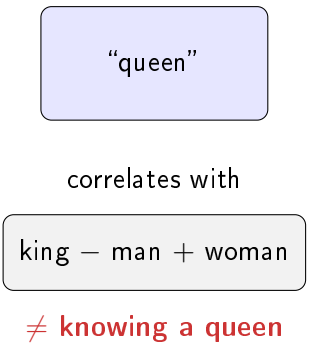
When AI reads "I'm sad":

- Sees correlation patterns
- Associates: tears, loss, certain words
- Never felt grief

Harnad's Analogy (Harnad 1990):

Like learning Chinese from a Chinese-Chinese dictionary

Plato's Cave:
Shadows $\neq$ Fire



How Humans Learn

- Pain creates permanent memory
- Hot stove $\rightarrow$ instant learning
- Genetic firmware (blind babies walk)
- Brain *hardware* adapts

Consequences $\rightarrow$ *Understanding*

How AI Learns

- Static datasets
- Reads about fire, never burned
- No interaction
- No stakes

"Swimming from a textbook"

Human Memory

Mom's name **cached forever**

Old classmate garbage-collected

Emotional tagging
Priority systems
Efficient forgetting

Transformer Memory

Fixed window, then gone
No permanent cache
Recompute every time

Library with index vs Read every book

Somatic Marker Hypothesis (Damasio 1994):

Emotions = computational shortcuts
Guide decisions before conscious thought

Phineas Gage (1848):

- Tamping iron through skull
- vmPFC destroyed
- IQ **intact** → life collapsed
- Couldn't decide what to eat

Iowa Gambling Task:
Body knew before mind did

Fear → Jump
Disgust → Recoil
Curiosity → Explore

AI has no gut feelings

	Human	AI
Power	20 W	Megawatts
Memory	Emotional tags	Context window
Learning	Embodied	Text-trained
Instincts	Pre-installed	From scratch
Decisions	Gut feelings	Statistics
Curiosity	Genuine	Pattern completion
Stakes	**Mortality**	None

Different in **kind**, not just degree

Alex asked:

"What color am I?"

→ Genuine curiosity
→ Embodied existence
→ **Actually wondering**

LLMs generate:

"What color am I?"

→ Pattern completion
→ Statistical likelihood
→ **Different**

Two kinds of parrot. Neither is wrong.

Know what you're working with.

Use it well.

- Bender, Emily M. et al. (2021). "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. FAccT '21. New York, NY, USA: Association for Computing Machinery, pp. 610–623. DOI: 10.1145/3442188.3445922.
- Damasio, Antonio R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: G.P. Putnam's Sons. ISBN: 978-0399138942.
- Harnad, Stevan (1990). "The Symbol Grounding Problem". In: *Physica D: Nonlinear Phenomena* 42.1–3, pp. 335–346. DOI: 10.1016/0167-2789(90)90087-6.
- Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton (2012). "ImageNet Classification with Deep Convolutional Neural Networks". In: *Advances in Neural Information Processing Systems*. Vol. 25. AlexNet paper. Curran Associates, Inc., pp. 1097–1105.

 Pepperberg, Irene M. (1999). *The Alex Studies: Cognitive and Communicative Abilities of Grey Parrots*. Cambridge, MA: Harvard University Press. ISBN: 978-0674008069.