

A Gentle Introduction to AI, Language Modelling & Chatbots

Lecture Series in AI & Machine Learning

Dr. Donnacha Daly
Head of Bachelor in Artificial Intelligence & Machine Learning
Member of the Algorithmic Business (ABIZ) Research Lab

donnacha.daly@hslu.ch | [linkedin.com/in/donnacha](https://www.linkedin.com/in/donnacha) | hslu.ch/bachelor-ai | www.abiz.ch



https://twitter.com/hslu_aiml/status/1640285871692709892

Image created for this presentation with
www.midjourney.com

A. What's Going on?

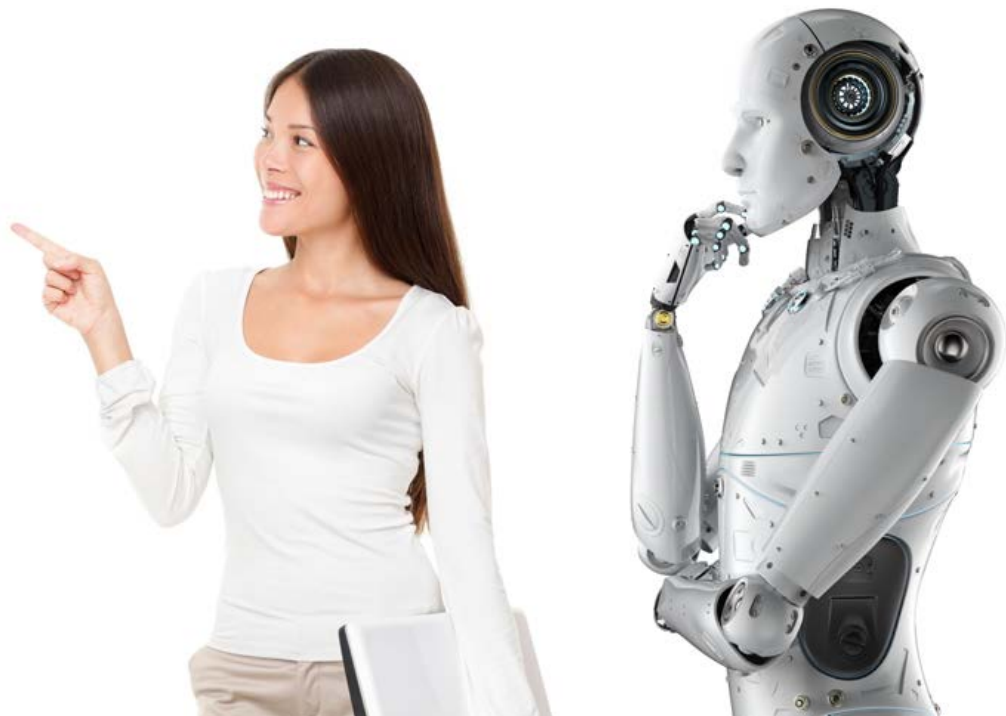
Language Models and ChatGPT

B. How does it Work?

Machine Learning & Transformers

C. What does it all mean?

Artificial (General) Intelligence: AI and AGI



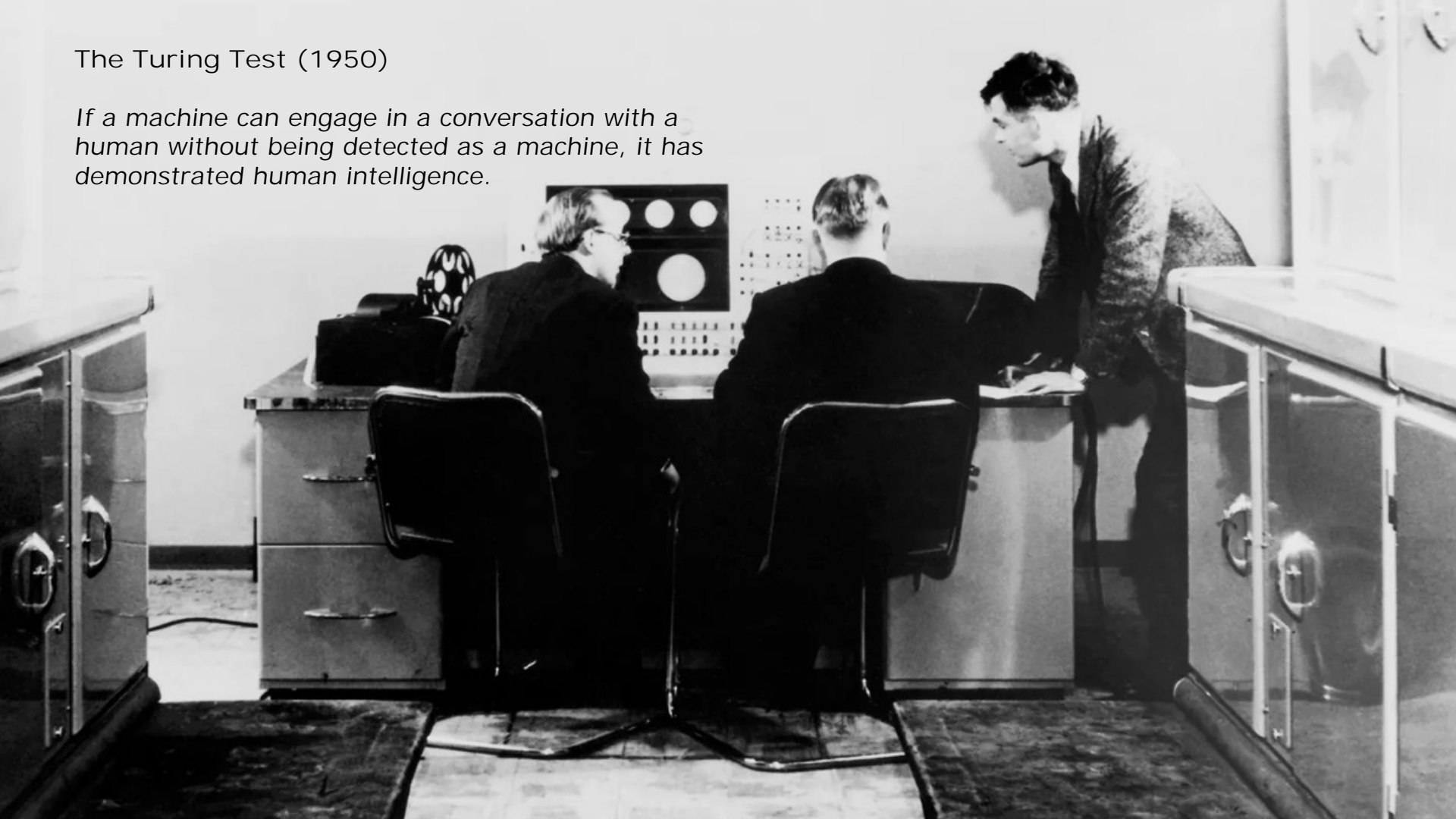
What's Going On?

Language Models and ChatGPT



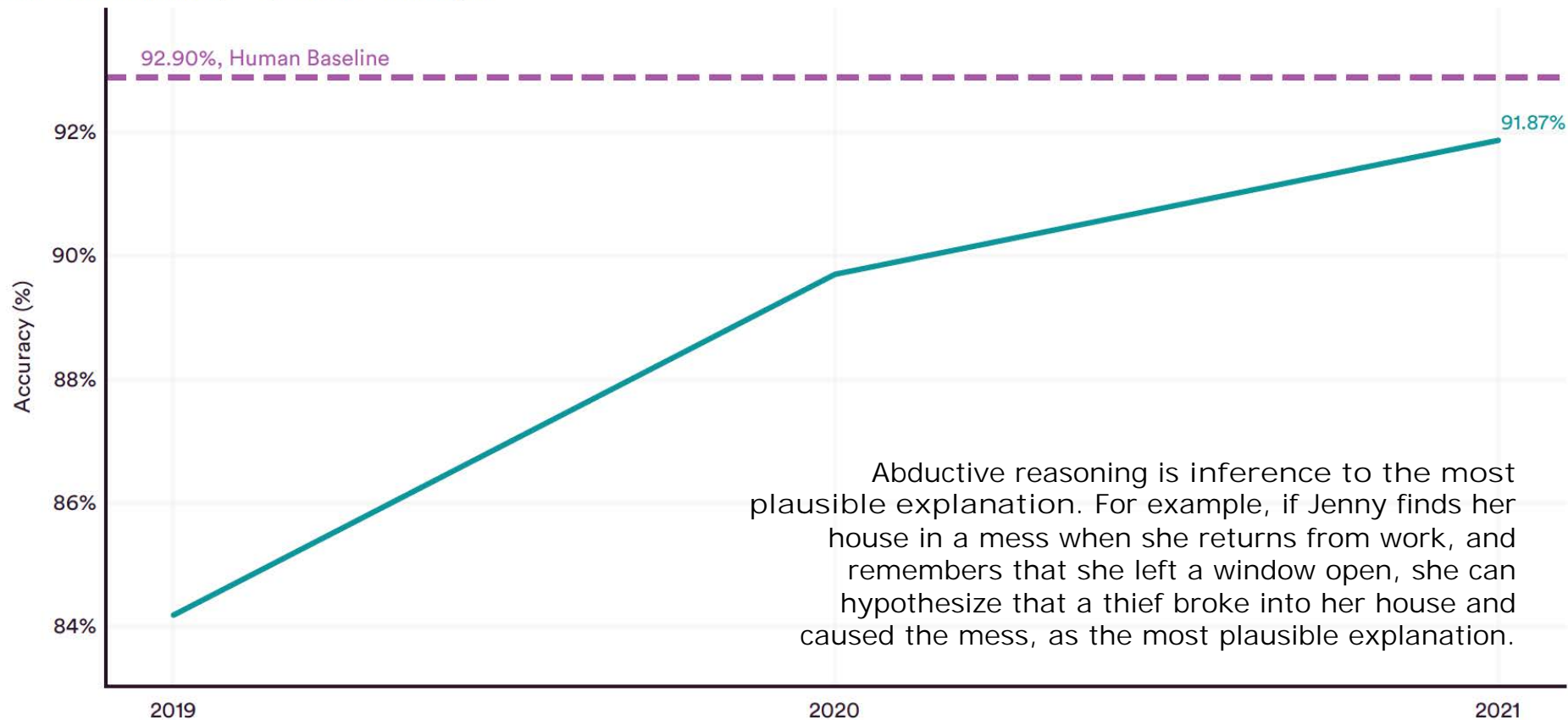
The Turing Test (1950)

If a machine can engage in a conversation with a human without being detected as a machine, it has demonstrated human intelligence.



ABDUCTIVE NATURAL LANGUAGE INFERENCE (aNLI): ACCURACY

Source: Allen Institute for AI, 2021 | Chart: 2022 AI Index Report



November 2022: ChatGPT Explodes onto the Scene

ChatGPT Sprints to One Million Users

Time it took for selected online services to reach one million users



*one million backers **one million nights booked ***one million downloads

Source: Company announcements via Business Insider/LinkedIn

statista

ChatGPT Passed the HSLU
Machine Learning Exam!



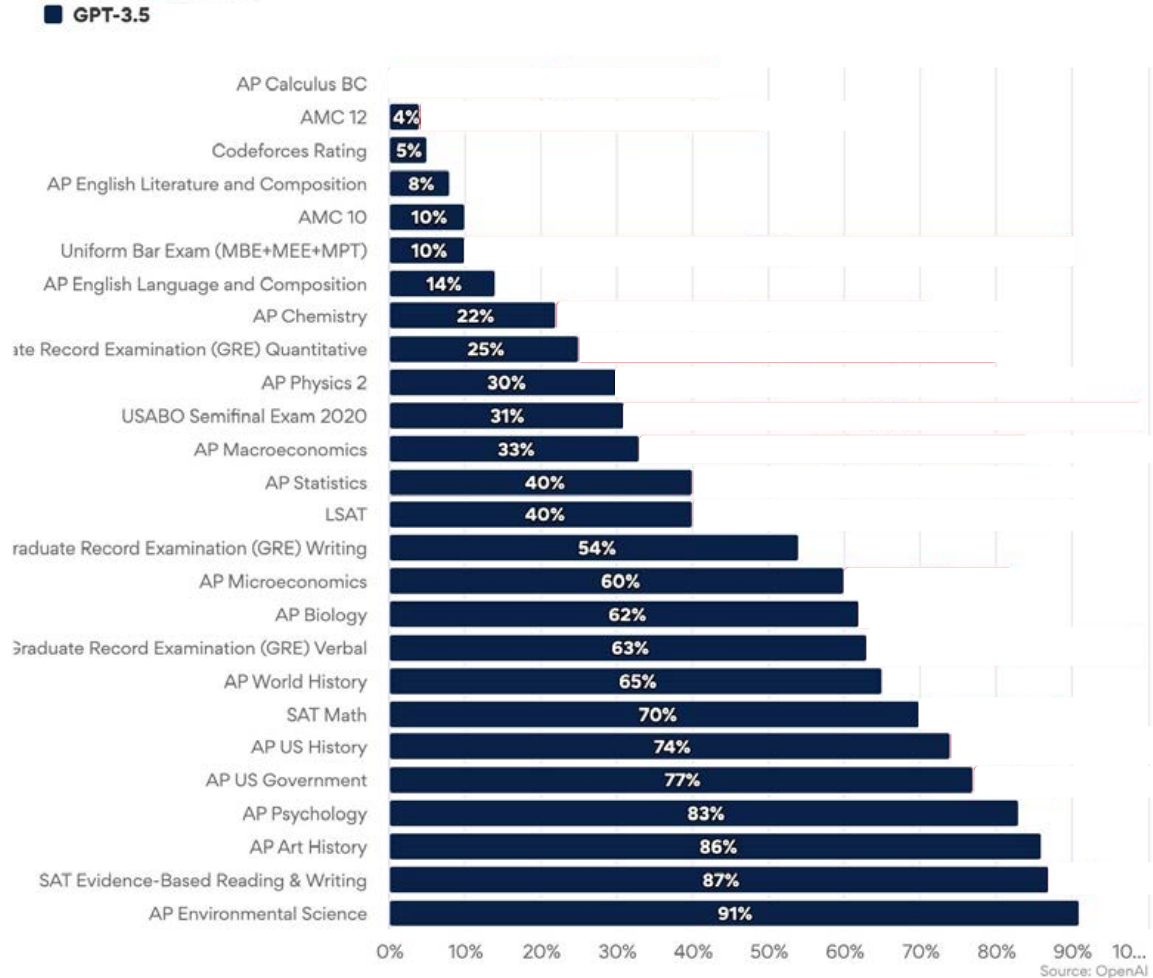
Bojan Tunguz ✓
@tunguz

Machine Learning at Nvidia. Kaggle
Quadruple Grandmaster. ...

This is insane. This is what I've been alluding to for months now. This is an epochal transformative technology that will soon touch - and radically transform - ALL knowledge work. If most of your work involves sitting in front of a computer, you will be disrupted very, very soon.

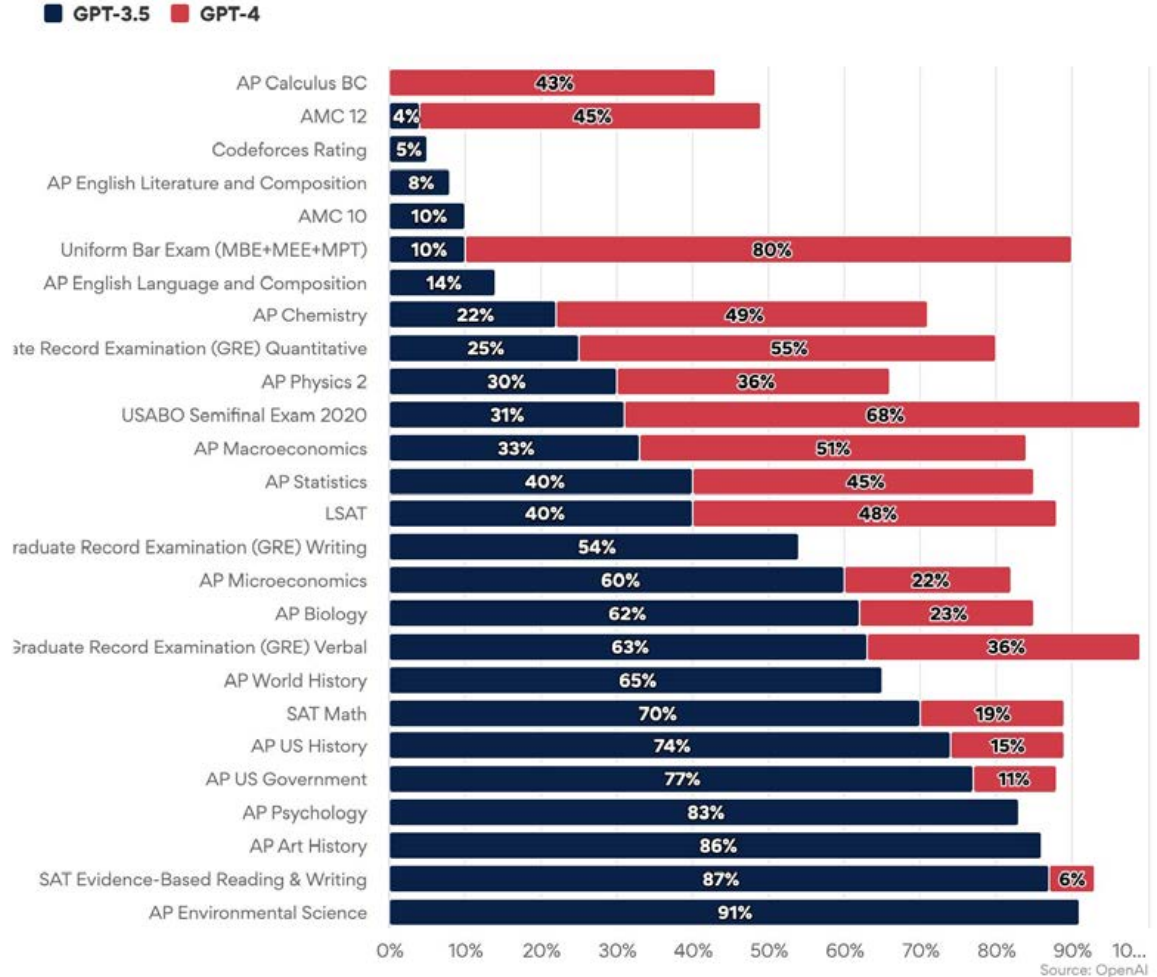
In fact,
ChatGPT has
passed many
exams with
grades in the
top quantile

Lower Bound Percentile Exam Results (Ordered By GPT-3.5 Performance)



GPT-4 was released in March 2023

Lower Bound Percentile Exam Results (Ordered By GPT-3.5 Performance)



How Does it Work?

Machine Learning & Transformers



Let's start with a simple language model....

Word w	Roll dice	Probability $p(w)$
a	2	1/36
apple	3-4	5/36
ate	5-6	9/36
green	7-8	11/36
man	9-10	7/36
the	11-12	3/36



20 Word Example

man green green ate green man apple ate ate the green ate man green green apple ate green the man

We normally sample from large vocabularies...

Word w	Probability $p(w)$
...	...
ability	0.000353%
ablaze	0.000254%
able	0.002118%
ably	0.000155%
...	...



20 Word Example

cajoles laugh bold bluish brim zephyr for scuba flit obfuscate whimsical armless of yawn jovial am came
living some glib

What about context? Some words are more likely to follow others...

Word $w(n-1)$	Word $w(n)$	Probability $p(w(n) w(n-1))$
...
lemon	cake	0.89%
lemon	juice	1.10%
lemon	jacket	0.07%
lemon	follows	0.002%
...

Bigram Model

The probability of a word depends only on the previous word

(also known as 1st order Markov model)

Completion Example

Smoked salmon tastes great with a squeeze of lemon _____

Basic requirements of a language model....

- To generate a probability distribution over possible next words (learn this from data!)
- This probability distribution should be **context dependent**
- We should be able to sample from this distribution to create text
- RNNs, LSTMs, GRUs, **Transformers**

Attention Is All You Need

Ashish Vaswani* Google Brain avaswani@google.com	Noam Shazeer* Google Brain noam@google.com	Niki Parmar* Google Research nikip@google.com	Jakob Uszkoreit* Google Research usz@google.com
Llion Jones* Google Research llion@google.com	Aidan N. Gomez* † University of Toronto aidan@cs.toronto.edu	Łukasz Kaiser* Google Brain lukasz.kaiser@google.com	
Illia Polosukhin* † illia.polosukhin@gmail.com			

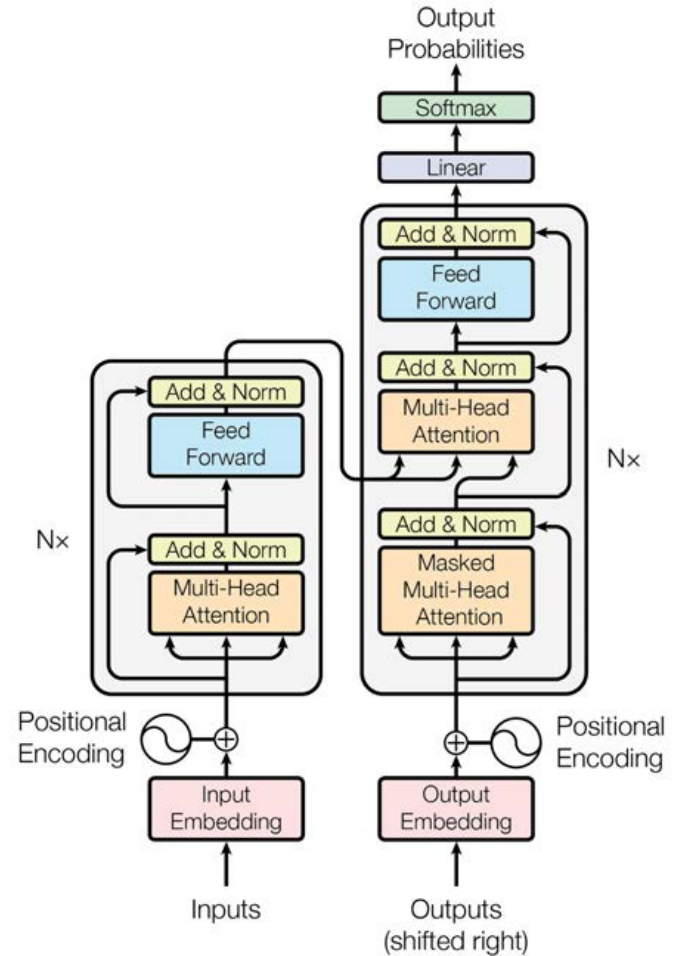
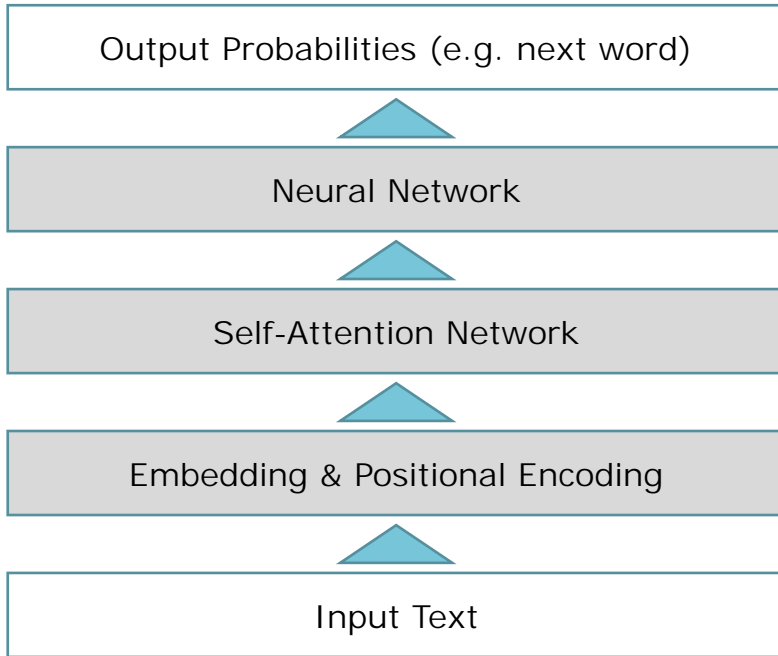
Abstract

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new single-model state-of-the-art BLEU score of 41.8 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the best models from the literature. We show that the Transformer generalizes well to other tasks by applying it successfully to English constituency parsing both with large and limited training data.

Self-Attention can be used to encode context

smoked	→	smoked	salmon	tastes	great	→	[0.70 , 0.20 , 0.07 , 0.03]
salmon	→	smoked	salmon	tastes	great	→	[0.70 , 0.80 , 0.07 , 0.03]
tastes	→	smoked	salmon	tastes	great	→	[0.70 , 0.20 , 0.75 , 0.03]
great	→	smoked	salmon	tastes	great	→	[0.01 , 0.02 , 0.07 , 0.90]

The Transformer Model



Model Pre-Training: Unsupervised Learning

Used to train the Transformer



I	am	definitely	going	Mask	Mask	Mask	Mask	Mask
I	am	definitely	going	to	Mask	Mask	Mask	Mask
I	am	definitely	going	to	Ireland	Mask	Mask	Mask
I	am	definitely	going	to	Ireland	on	Mask	Mask
I	am	definitely	going	to	Ireland	on	my	Mask
I	am	definitely	going	to	Ireland	on	my	holidays

Language model training requires a lot of data and a lot of computation

Words

10^0	"Hello World"
10^1	One Tweet (pre-Elon)
10^2	"Blank Space" Chorus
10^3	Wikipedia: "Fermi Estimation"
10^4	Wikipedia: "Taylor Swift"
10^5	GPT-3 Paper + Appendices
10^6	One Year of The New Yorker
10^7	Encyclopaedia Britannica
10^8	Reddit Posts per Month
10^9	English Wikipedia
10^{10}	WhatsApp Messages per Hour
10^{11}	Published Books per Year
10^{12}	Large Language Models



Researchers estimate that the training of GPT-3

- consumed 1.3 GWh of energy
- led to emissions of more than 550 tons of CO₂ equivalent

We don't have figures for Chat-GPT (GPT3.5) or GPT-4

Model Training: Supervised Learning (3-Step plus a content filter)

Step 1

Collect demonstration data and train a supervised policy.

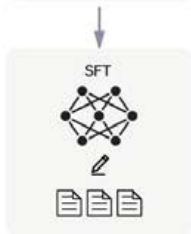
A prompt is sampled from our prompt dataset.



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.

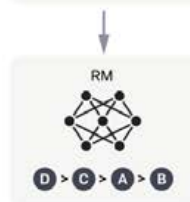
A prompt and several model outputs are sampled.



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.



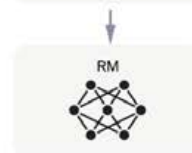
The PPO model is initialized from the supervised policy.



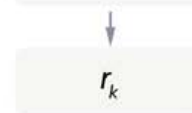
The policy generates an output.



The reward model calculates a reward for the output.

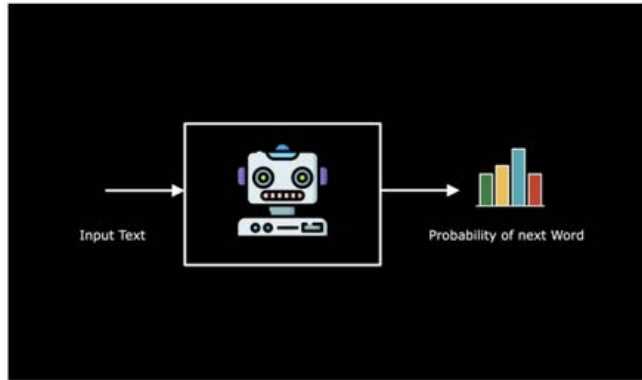


The reward is used to update the policy using PPO.



Want to know more? (auf Deutsch!)

So funktionieren ChatGPT und Co.: Unser Experten-Video



ChatGPT denkt nicht selbstständig. Das Sprachmodell arbeitet – wie nahezu alle KI-Systeme – mit Wahrscheinlichkeiten.



LUCERNE SCHOOL OF COMPUTER SCIENCE AND
INFORMATION TECHNOLOGY

Room 7.56

Campus Zug-Rotkreuz

Suurstoffi 1

6343 Rotkreuz

RESEARCH

Lecturer

Co-Head Algorithmic Business Research Lab

+41 41 757 68 25

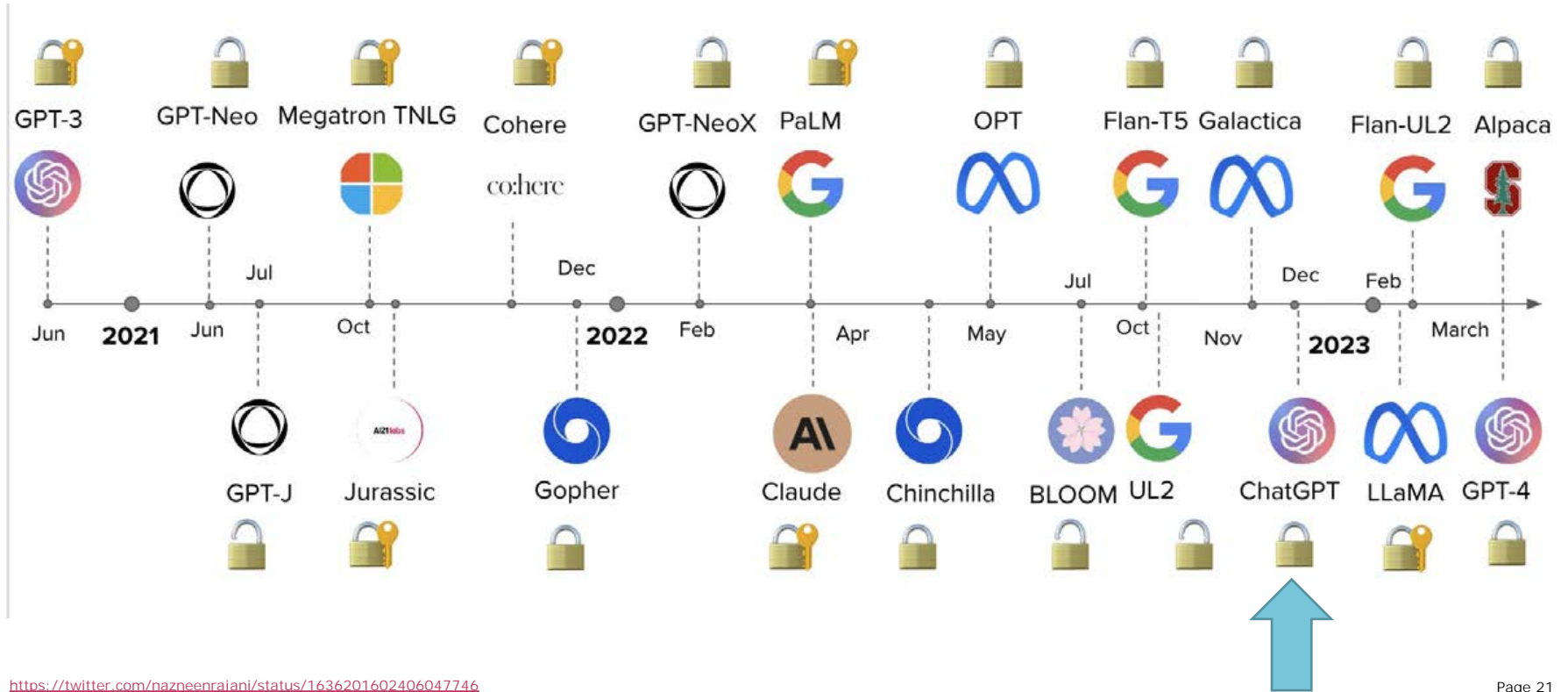
marc.pouly@hslu.ch

What does it all mean?

Artificial Intelligence (AI) and
Artificial General Intelligence (AGI)



Large Language Models since GPT-3



Latest Transformer Models are Multi-Modal



Example of GPT-4 visual input:

User: What is funny about this image? Describe it panel by panel.



Source: <https://www.reddit.com/r/hmmm/comments/ubab5v/hmmm/>

GPT-4: The image shows a package for a "Lightning Cable" adapter with three panels.

Panel 1: A smartphone with a VGA connector (a large, blue, 15-pin connector typically used for computer monitors) plugged into its charging port.

Panel 2: The package for the "Lightning Cable" adapter with a picture of a VGA connector on it.

Panel 3: A close-up of the VGA connector with a small Lightning connector (used for charging iPhones and other Apple devices) at the end.

The humor in this image comes from the absurdity of plugging a large, outdated VGA connector into a small, modern smartphone charging port.

Some people believe that LLMs are the first step towards a general AI

AI pioneer Geoff Hinton: "Deep learning is going to be able to do everything"

Thirty years ago, Hinton's belief in neural networks was contrarian. Now it's hard to find anyone who disagrees, he says.

By Karen Hao

November 3, 2020



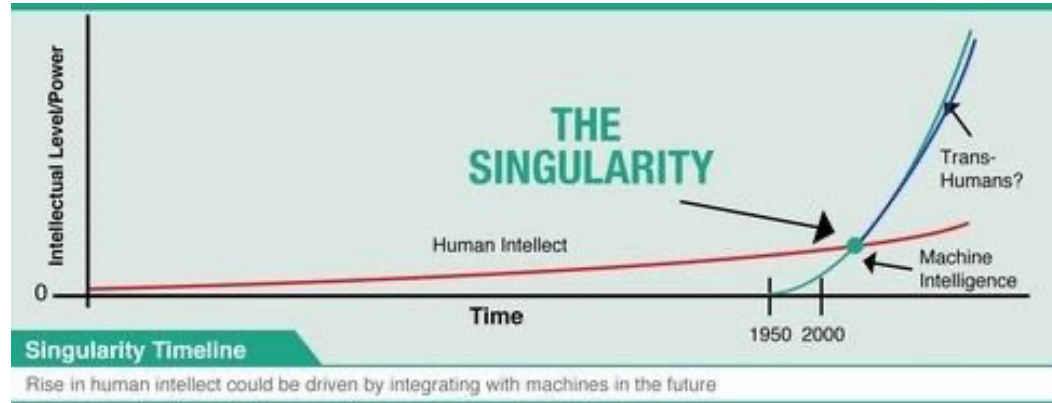
NOAH BERGER / AP

MIT Technology Review  @techreview · Nov 26

"I do believe deep learning is going to be able to do everything, but I do think there's going to have to be quite a few conceptual breakthroughs."

<https://twitter.com/markopoljarvi/status/1638224173343318016>
<https://www.technologyreview.com/2020/11/03/1011616/ai-godfather-geoffrey-hinton-deep-learning-will-do-everything>

The Singularity



Ray Kurzweil, a Director of Engineering at Google and pioneer of concepts related to the AI Singularity, forecasts that computers will have human level intelligence by 2029

We are wrong about AI in many ways

- X Narrow intelligence is on a continuum with general intelligence: **No!**
- X Easy things are hard and hard things are easy (e.g. picking up a box v beating a grandmaster at chess): **No!**
- X Machine learning resembles learning in humans (pursuit of a goal, understanding data): **No!**
- X Intelligence is all in the brain (the body is not central to cognitive activities): **No!**

Why AI is Harder Than We Think

Melanie Mitchell
Santa Fe Institute
Santa Fe, NM, USA

Abstract

Since its beginning in the 1950s, the field of artificial intelligence has cycled several times between periods of optimistic predictions and massive investment (“AI spring”) and periods of disappointment, loss of confidence, and reduced funding (“AI winter”). Even with today’s seemingly fast pace of AI breakthroughs, the development of long-promised technologies such as self-driving cars, housekeeping robots, and conversational companions has turned out to be much harder than many people expected. One reason for these repeating cycles is our limited understanding of the nature and complexity of intelligence itself. In this paper I describe four fallacies in common assumptions made by AI researchers, which can lead to overconfident predictions about the field. I conclude by discussing the open questions spurred by these fallacies, including the age-old challenge of imbuing machines with humanlike common sense.



Melanie Mitchell is a professor of computer science at Portland State University. Her major work has been in the areas of analogical reasoning, complex systems, genetic algorithms and cellular automata.

Language models are “next word predictors”

- Large language models (LLMs) **hallucinate**
- There are mistakes in the generated text that are semantically or syntactically **plausible**
- But they can in fact be incorrect or **nonsensical**
- In short, you can't trust what the machine is telling you.

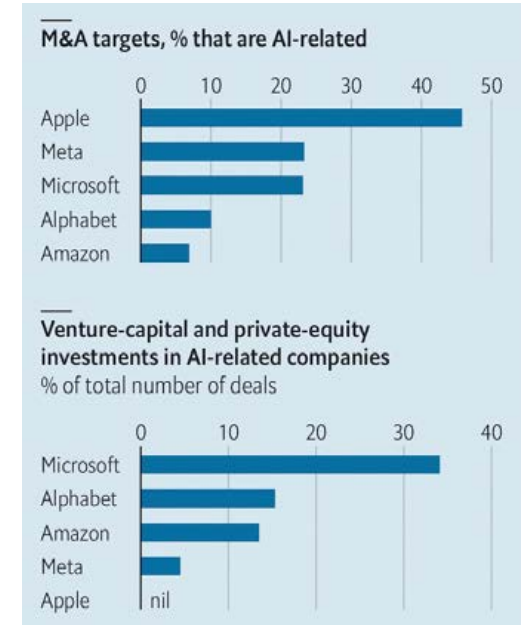
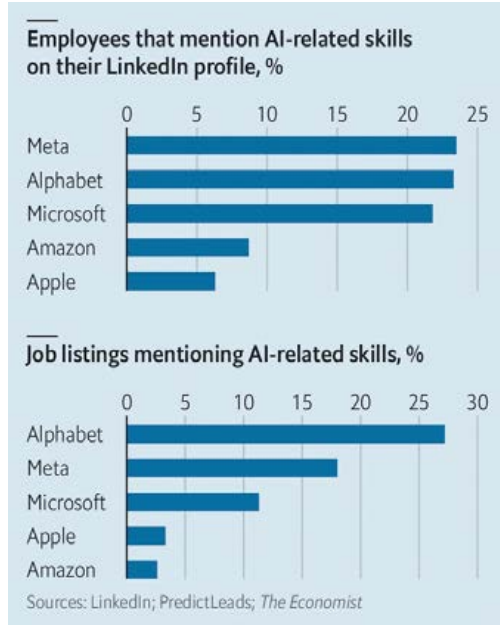
“Large language models have no idea of the underlying reality that language describes” – Yann Lecun, Chief AI Scientist at Facebook

<https://spectrum.ieee.org/ai-hallucination>



Image created for this presentation with www.bing.com/images/create.
Source https://twitter.com/hslu_aiml/status/163839496594100234

Nevertheless, Big Tech is Laser Focused on AI



AI is a big topic in Luzern! See www.lac2.org

AI INTERVIEWS



January 18, 2023

These new technologies in AI and Data Science will define how we work and live in the future

LAC² Interview with Dr. Patricia Feubl, Head of the Competence Center for Communication and Marketing Technologies, HSLU Lucerne School of Business.

... for
... AI
... nitoring and

... sch & Till

Schöpe, Co-Founders at Alpacayo Solutions, located in Sarnen.

AI BREAKFASTS



AI for Manufacturing and Industry

Christoph Natsch, Alpacayo & CSM

Raphael Von Thiesen, Amt für Wirtschaft und Arbeit, Kanton Zürich

AI EVENTS



AI/ML Student-Industry Dialog Event

Wednesday, 23 November 2022 @ 10:00 AM
Multi-Pass ID: 9534-949694 (Shareable ID)

Christoph Natsch, Alpacayo & CSM

Anastasia Natsch, Alpacayo & CSM

Thomas Brändli, HSLU Luzern

... partnered with Hochschule Luzern - AI, Student-Industry event, ... see more



Lucerne AI & Cognitive Community

A bridge to future AI value generation

In Summary

- A. **Progress** in large language models such as ChatGPT has taken the world by storm: Transformer models bring us close to human levels of performance
- B. These **technologies** work by ingesting massive amounts of data, and training a neural network to learn the statistics of contextual relationships within this data
- C. Business and society are rapidly changing because of these developments. We can expect a **tsunami of change** in the coming years

Thank You!
Questions?



Thank You! Any Questions?



Dr. Donnacha Daly

Head of Bachelor Program in Artificial Intelligence & Machine Learning
Member of the Algorithmic Business (ABIZ) Research Lab

Lucerne University of Applied Sciences and Arts
School of Computer Science & Information Technology

donnacha.daly@hslu.ch | hslu.ch/bachelor-ai | www.abiz.ch

