HOGESCHOOL ROTTERDAM

IS TRUTH FINDING POSSIBLE WITH AI?

[HTML] **ChatGPT is bullshit**

MT Hicks, J Humphries, J Slater - Ethics and Information Technology, 2024 - Springer

… of **bullshit ChatGPT** … **ChatGPT**-generated text as **bullshit**, and flag up why it matters that – rather than thinking of its untrue claims as lies or hallucinations – we call **bullshit** on **ChatGPT**. …

☆ Save 99 Cite Cited by 518 Related articles All 22 versions

Hicks, M.T., Humphries, J. & Slater, J. ChatGPT is bullshit. *Ethics Inf Technol* **26**, 38 (2024). https://doi.org/10.1007/s10676-024-09775-5

HOGESCHOOL ROTTERDAM

AI anno 2026 cannot perform reliable "truth finding" and fails on complex, new, or local information. It is currently best used as a tool to augment human judgment, not replace it.

HOGESCHOOL ROTTERDAM

AI-Driven Truth Finding:

- **Retrieval-Augmented Generation (RAG):** reduces hallucinations??? by grounding LLM outputs in verified external data, rather than relying solely on the model's internal training data.
- **Hallucination Detection:** Researchers are developing tools to identify when an AI is making up information. For example, the CLATTER method guides LLMs through an explicit reasoning process to verify facts.
- **Trustworthy Search Engines:** New AI engines, such as "TrueGL," are being developed to assign truth scores (on a scale of 0 to 10) to search results, providing a visual gauge of credibility (e.g., green for high reliability, red for low).
- **Multi-Agent Systems:** Using multiple AI agents can improve fact-checking, as they can cross-reference information and cross-examine claims.

Current Challenges and Future Directions:

- **Context Dependence:** LLMs are more effective at checking national/international news than local or dynamic, real-time information.
- **Reducing "Careless Speech":** Research is focusing on curbing the tendency of LLMs to prioritize confident, polite-sounding, but false answers (a phenomenon sometimes called "careless speech").
- **Hybrid Models:** Combining LLMs with structured data (databases) and rule-based systems is critical for high-stakes environments like healthcare and finance.

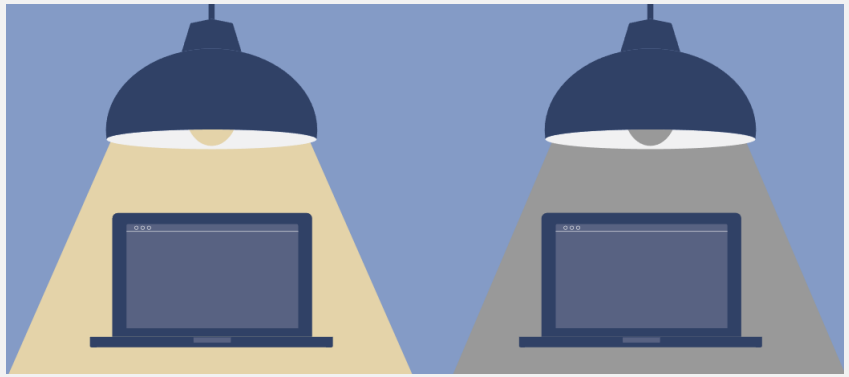
https://github.com/HR-DataLab-Healthcare/RESEARCH_SUPPORT/tree/main/WORKSHOPS/SAMEN_AAN_DE_SLAG_2026#readme

HOGESCHOOL ROTTERDAM

https://samenaandeslag.cyber-secure-te.src.surf-hosted.nl/

HOGESCHOOL ROTTERDAM

Schaduw-ICT in de vorm van AI



Hoe ga je daarmee om?

SURF: Cybersecurity

Schaduw-ict in onderwijs en onderzoek: wat moet je er als instelling mee? Latest update: 15 februari 2023

https://communities.surf.nl/cybersecurity/artikel/schaduw-ict-in-onderwijs-en-onderzoek-wat-moet-je-er-als-instelling-mee

In het onderwijs en onderzoek krijgen studenten, wetenschappers en docenten veel vrijheid in de manier waarop ze ict gebruiken.

Faculteiten en instituten werken autonoom, zodat ze snel kunnen inspringen op nieuwe ontwikkelingen. Mede daardoor ontstaat schaduw-ict.

Dit kan voor ict-afdelingen leiden tot een verlies van controle.

In het bijzonder voor de netwerk- en informatiebeveiliging zijn de risico's moeilijk te overzien en veel instellingen vragen zich af wat ze er mee aan moeten.

Beveiligingsincidenten en datalekken melden (bijgewerkt op 31 augustus 2023)

Als school probeer je de persoonsgegevens die je bewaart zo goed mogelijk te beveiligen. Maar soms gaat er is iets mis. Je hebt bijvoorbeeld zelf geen toegang meer tot je gegevens of buitenstaanders krijgen onbedoeld toegang. Dan is er sprake van een beveiligingsincident. Heeft het incident gevolgen voor de privacy van leerlingen of medewerkers? Dan spreken we van een datalek. Voor het melden van datalekken gelden andere regels dan voor andere beveiligingsincidenten. Volgens de Algemene Verordening Gegevensbescherming (AVG) ben je als school verplicht datalekken direct te melden bij de Autoriteit Persoonsgegevens.

https://aanpakibp.kennisnet.nl/beveiligingsincidenten-en-datalekken-melden/

Shadow IT creates the possibility that organizations may run afoul of regulations such as PCI-DSS, GDPR, HIPAA, SOX and others, exposing them to severe penalties and fines. It can also lead to an increase in the likelihood of data breaches when IT and security operations lose control over the software and applications used in an environment.

https://www.forbes.com/sites/forbestechcouncil/2022/07/19/how-shadow-it-can-keep-compliance-efforts-in-the-dark/

Wat is schaduw-ICT [shadow IT / gray IT]

Schaduw-ICT omvat ongeautoriseerde *hardware, software of diensten die voor "zakelijke" doeleinde (onderwijs en/of onderzoek) worden ingeregeld, ingevoerd en/of gebruikt zonder uitdrukkelijke goedkeuring of medeweten van de organisatie / systeembeheerders en/of technische staf.*

Omdat schaduw-IT niet wordt meegenomen in assetmanagement en evenmin aansluiten bij het AVG-compliance beleid, vormen ze een veiligheid en/of compliance risico.

Dit kan leiden tot het lekken van gevoelige gegevens (datalekken)of de verspreiding van malware binnen de organisatie.

https://www.ncsc.gov.uk/guidance/shadow-it

Interest in AI applications and features increased shadow IT. ChatGPT claimed #1 spot

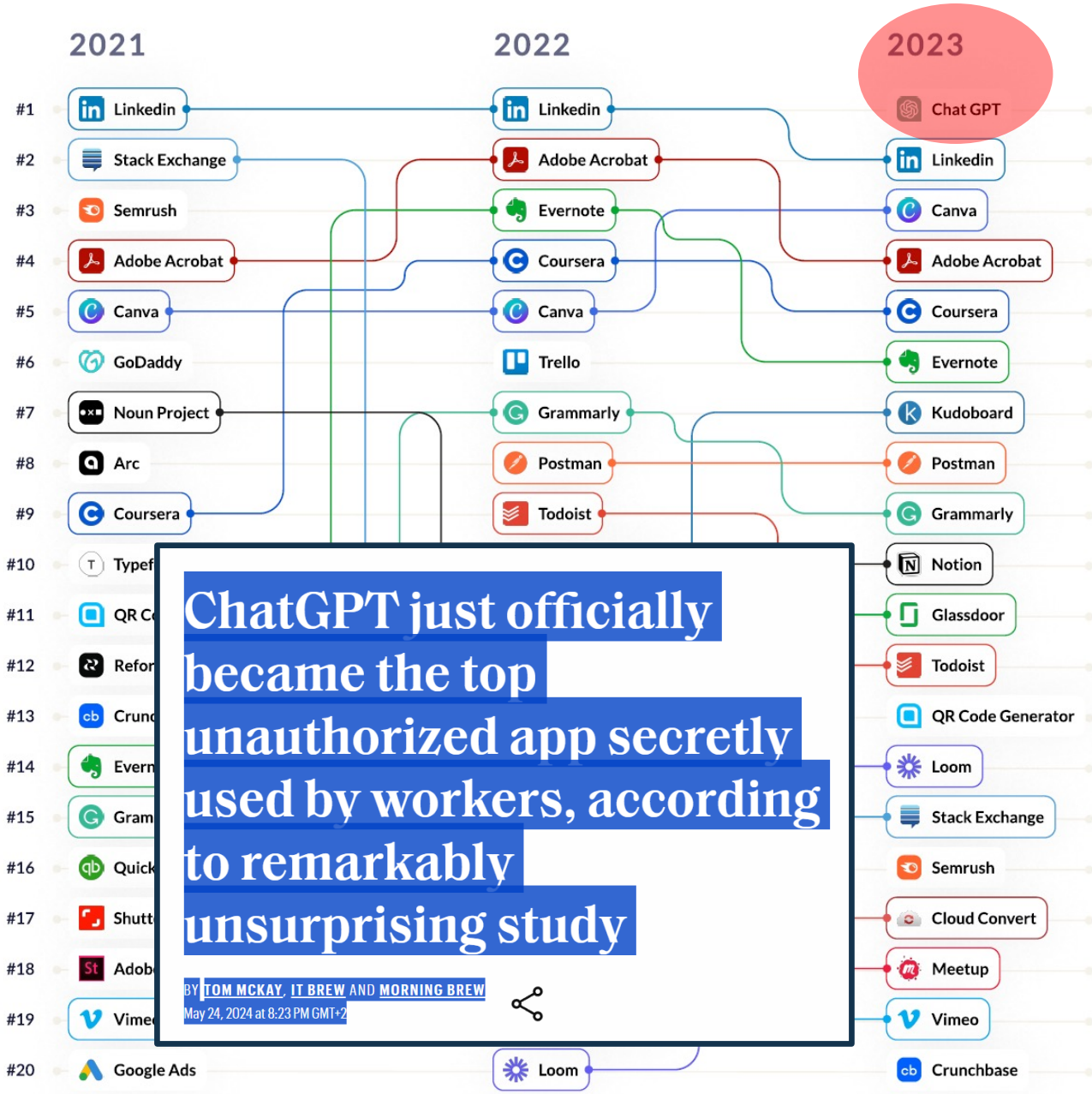
https://www.itbrew.com/stories/2024/05/22/chatgpt-is-the-number-one-offender-in-shadow-it-report-finds

KEY TAKEAWAYS

- ChatGPT has jumped to the top of the shadow IT chart as employees continue to adopt Artificial Intelligence.
- As the innovators and early adopters within a company continue to seek out AI-native applications (like ChatGPT and Grammarly) and AI solutions (like those offered by Canva and Evernote) for unmet needs, organizations should be developing a cohesive AI strategy.
- Nearly every application here offers, or will likely offer, some type of AI functionality; #5, Coursera, saw signups for AI courses every minute in 2023, on average. They also demonstrate the continued strength of the PLG go-to-market motion, with most offering free signup variants.
- Outside of AI trends, use of LinkedIn stayed consistent during a time of increased revenue pressure and insecurity in the job market. Trello fell off the shadow IT chart in 2022, as more companies purchase Atlassian's suite of products.

HOGESCHOOL ROTTERDAM

MOST POPULAR SHADOW IT



Aanpak Generatieve AI Hogeschool Rotterdam {HR}

Generatieve AI pilots

Via pilots wordt beproefd hoe generatieve AI op een verantwoorde en veilige manier kan worden ingezet.

Voorwaarden Gen-AI gebruik

https://www.surf.nl/files/2021-09/waardenwijzer_def.pdf

1. **Bewustzijn**

   Surf waardenwijzer

   is bij alle organisatieonderdelen bekend.

2. **Verantwoording**

   Medewerkers verantwoorden zich expliciet over de risico's

   bij de inzet van Gen-AI.

3. **Scholing**

   De HR Academie biedt scholing aan om de dialoog over het

   waardenkader + waardenwijzer te faciliteren.

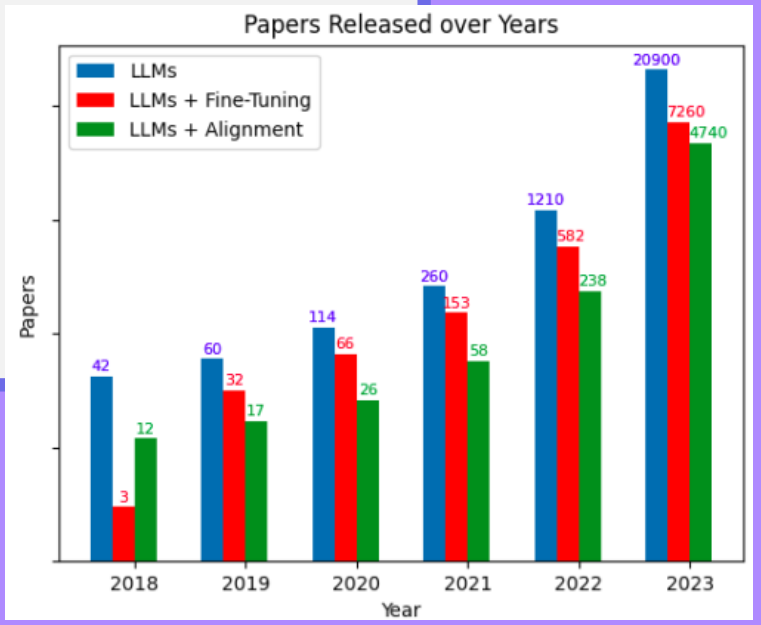
4. **Implementatie**

   Voordat IT-tooling wordt ingezet, wordt per opleiding een

   verantwoordelijke aangesteld die scholing heeft gevolgd.

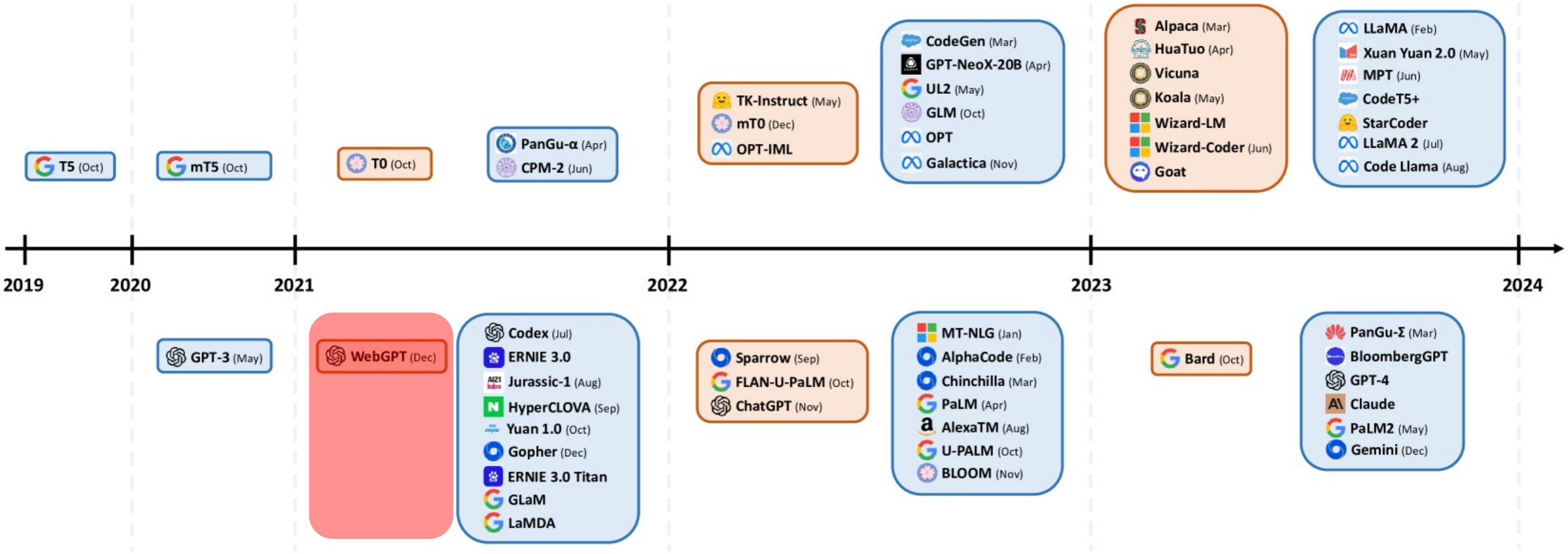
CONTEXT:

Waarom heeft Generatieve AI zo'n enorme impact op onderwijs & onderzoek?



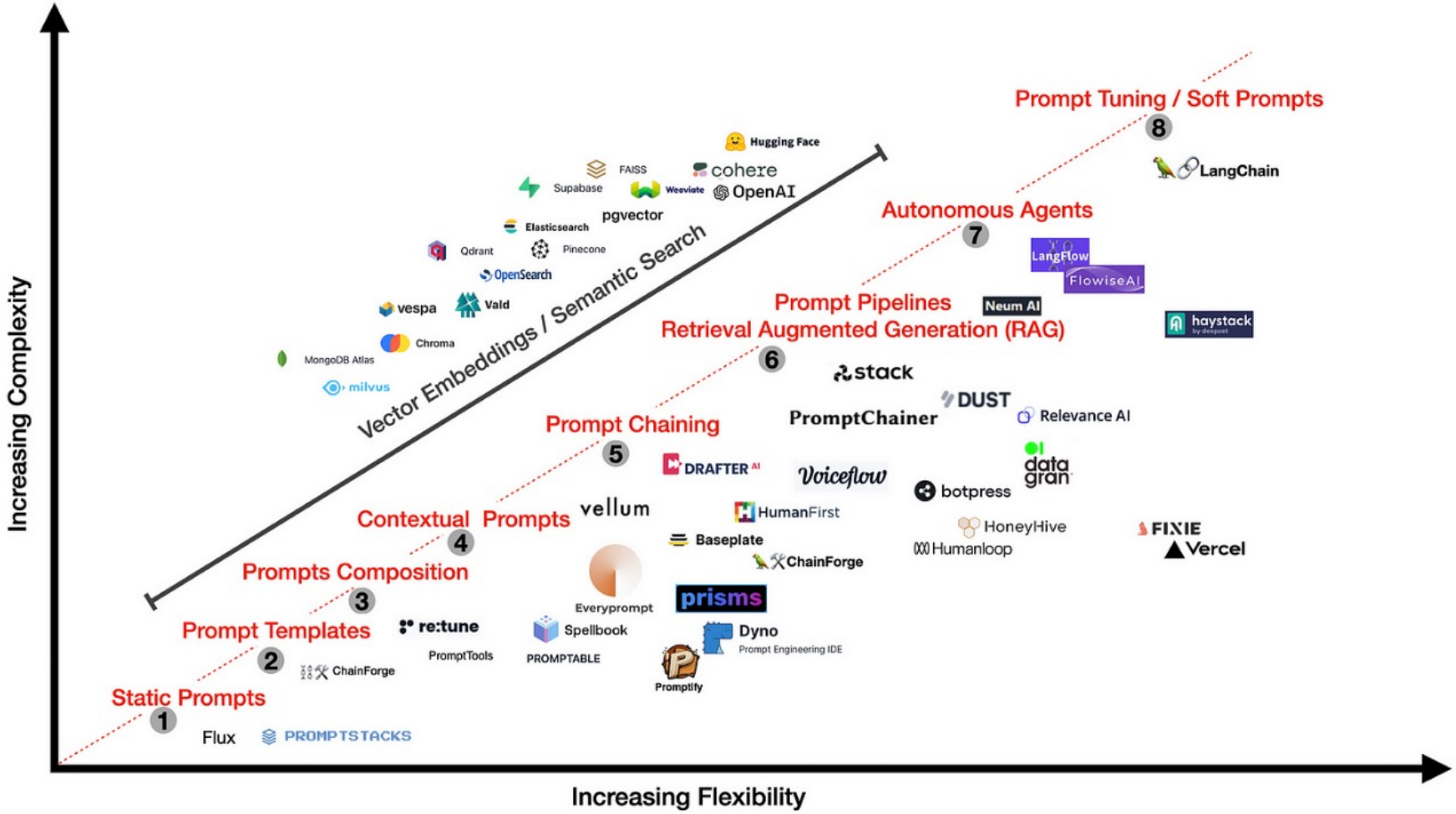
Ontstaansgeschiedenis + evolutie van grote taal modellen {LLM}

HOGESCHOOL ROTTERDAM



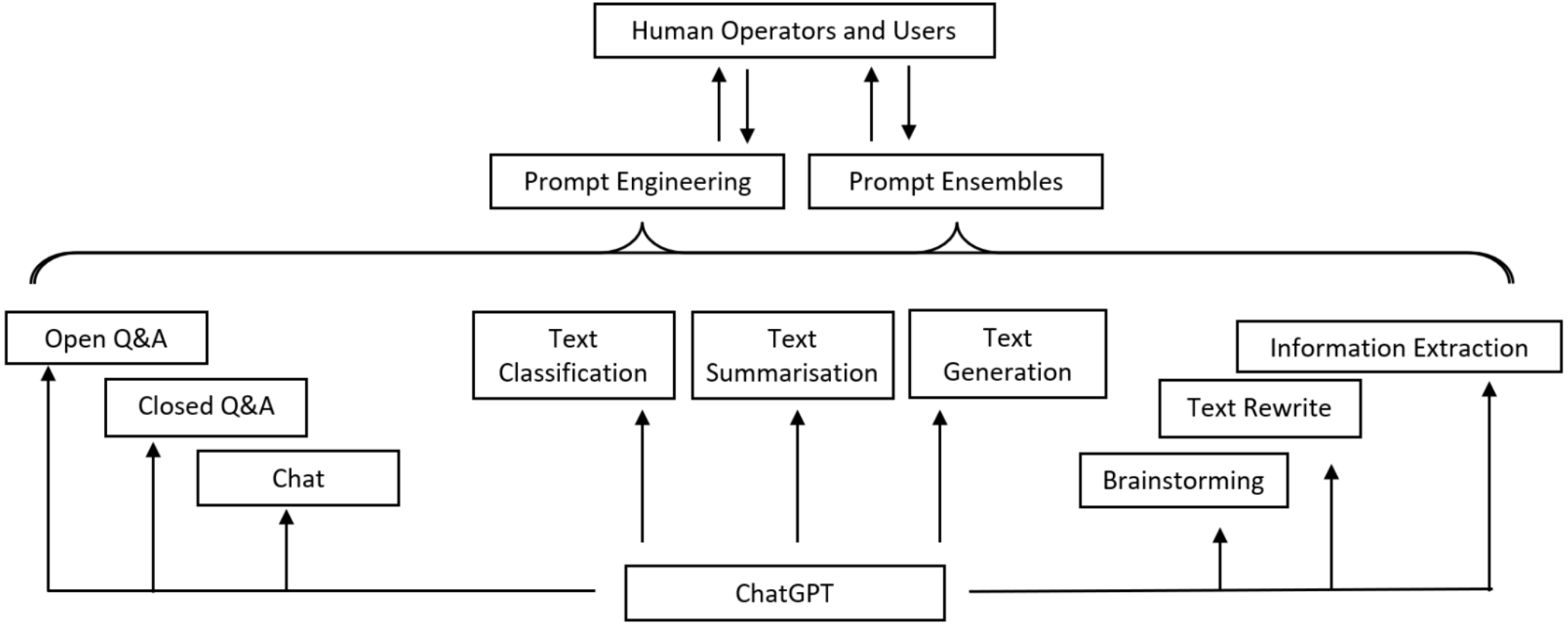
https://arxiv.org/html/2307.06435v9

LLM implementaties



https://blogs.novita.ai/exploring-architectural-structures-and-functional-capacities-of-llms/

ChatBot Use-Cases



Conferences > 2023 IEEE International Confe...

ChatGPT and Generative AI Guidelines for Addressing Academic Integrity and Augmenting Pre-Existing Chatbots

Publisher: IEEE Cite This PDF

Daswin De Silva ; Nishan Mills ; Mona El-Ayoubi ; Milos Manic ; Damminda Alahakoon All Authors

635 Full Text Views

Definiëring
karakteristieke kenmerken
Gen-AI *geeft inzicht*

AI-taxonomie

HOGESCHOOL ROTTERDAM

https://www.researchgate.net/publication/359424818_Designing_Neural_Networks_Through_Sensory_Ecology_Biology_to_the_rescue_of_AI_Produced_by_Living-Lab_AiRA_Hub_voor_Data_Responsible_AI_Hogeschool_Rotterdam_Lunch-Lezing_Creating-010_FEB_2022

Representations are localized and knowledgeable to humans

STRONG AI Human-level Intelligence

Good-Old-Fashioned Artificial Intelligence

GOFAI symbolic → Reasoning & Logic → Acting Humanly / Thinking Humanly

AI Capability Levels

Top-down approach

Bottom-up approach

Rule driven

Data driven

WEAK AI (narrow) single purpose Intelligence

CI sub-symbolic → Machine Learning & Self-organizing → Connect-ionist models / Embodied (embedded) Artificial lifeEm / Distributed Parallel Repre-sentation

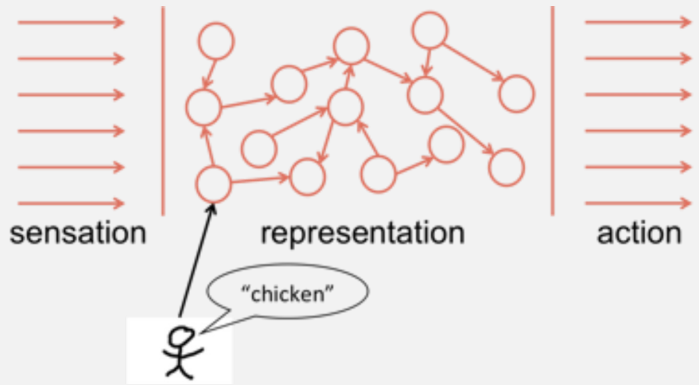
Modern Computational Intelligence

Representations are distributed not readable for humans

SYMBOLIC

Aoccdrnig to rscheearch at Cmabrigde Uinervtisy, it deosn't mttaer in waht oredr the ltteers in a wrod are, the olny iprmoetnt tihng is taht the frist and lsat ltteer be at the rghit pclae. The rset can be a toatl mses and you can sitll raed it wouthit porbelm. Tihs is bcuseae the huamn mnid deos not raed ervey lteter by istlef, but the wrod as a wlohe.

SUBSYMBOLIC



{Bottum-UP: Machinaal Leren [ML]}

HOGESCHOOL ROTTERDAM

What is machine learning?

> A field of study that gives computers the ability to learn without being explicitly programmed.
>
> —Attributed to Arthur Samuel

Given that deep learning is a subset of machine learning, what is machine learning? Most generally, it is what its name implies. Machine learning is a subfield of computer science wherein *machines learn* to perform tasks for which they were *not explicitly programmed*. In short, machines observe a pattern and attempt to imitate it in some way that can be either direct or indirect.

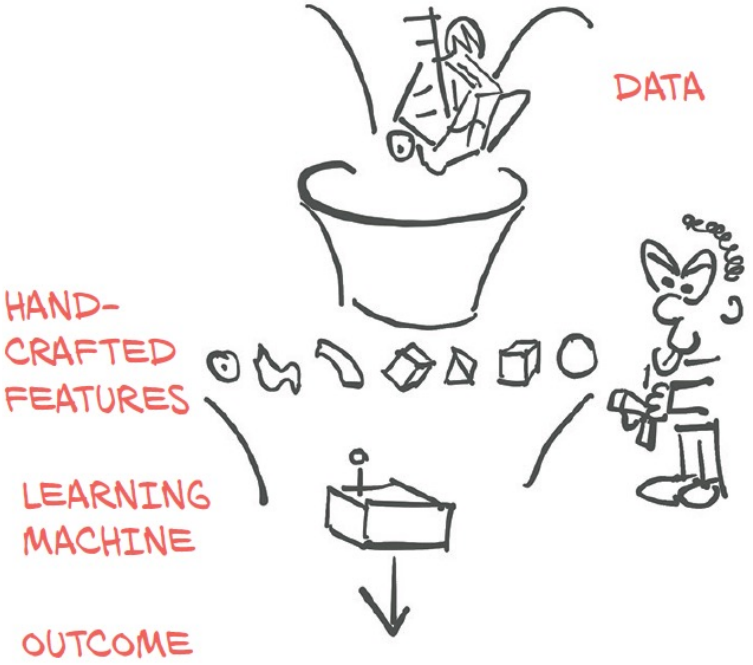
Machine learning ~= Monkey see, monkey do

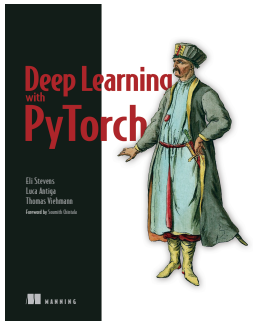
{Paradigm-shift ➔ NO Human input nodig }

HOGESCHOOL ROTTERDAM

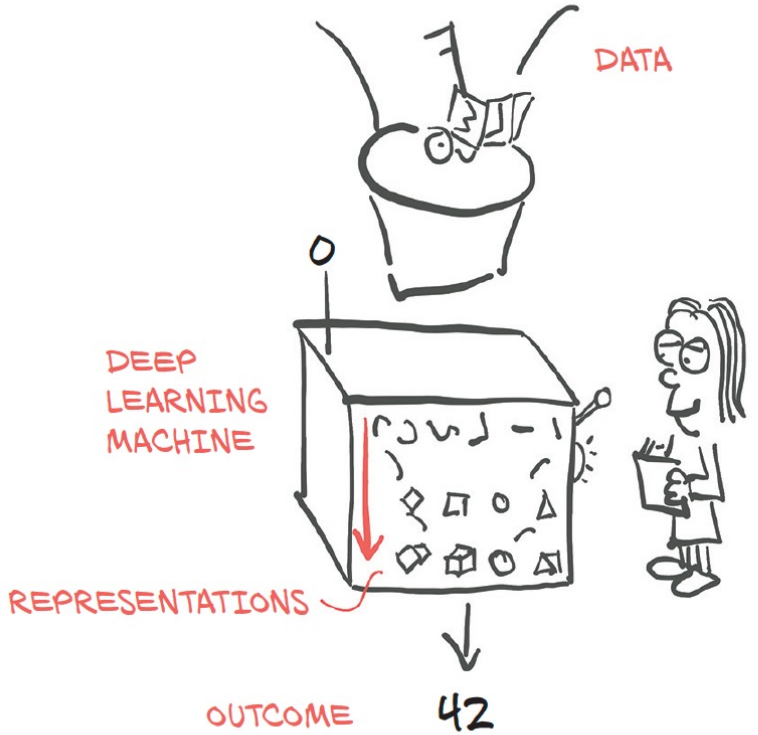
More data, parameters & computing power | Less human-in-the-loop

Machine Learning Paradigm {ML}





Deep Learning Paradigm {DL}

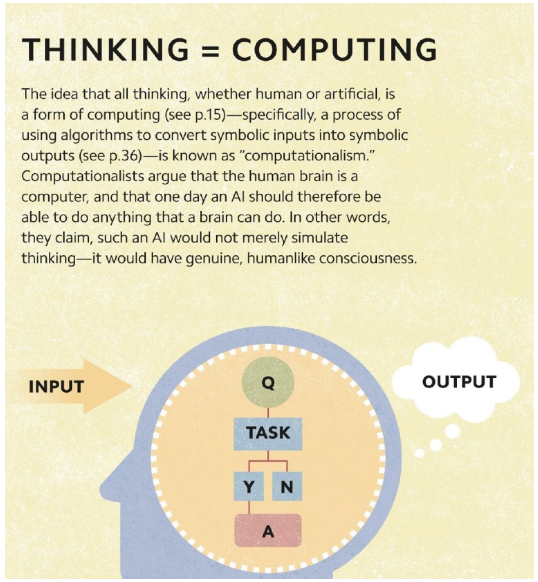


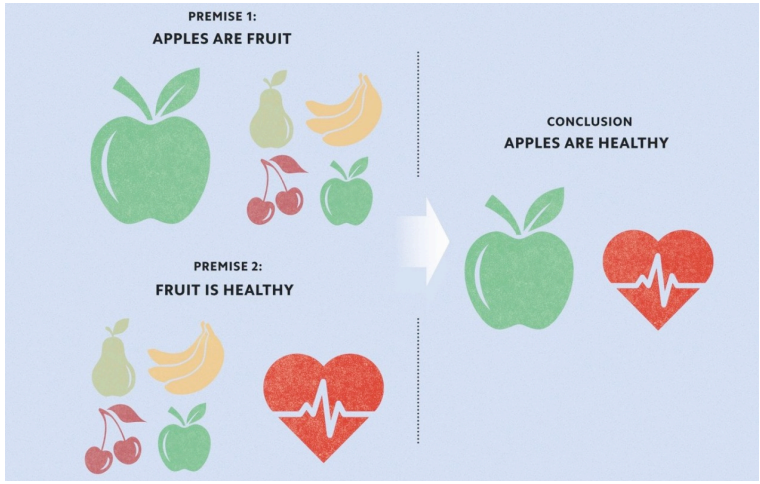
SYMBOLIC AI

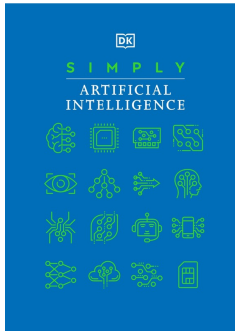
TOP-DOWN / open-loop

Intelligentie (denken) is
een vorm van Logica “Rekenkracht”
(computationalism)

Hersenen zijn een
“following the rules” Computer



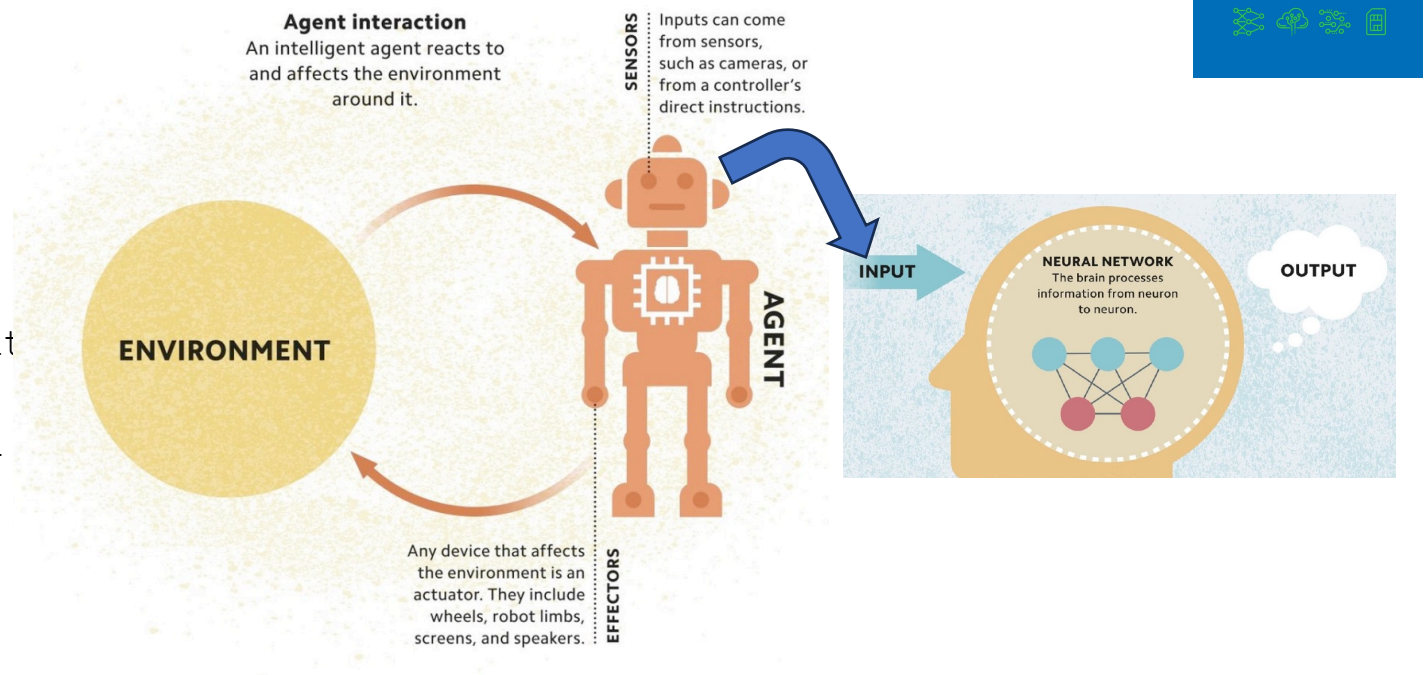




SUBSYMBOLIC AI

BOTTOM-UP / closed-loop

Intelligentie is adaptatief leervermogen
(trial & error) dat wordt mogelijk gemaakt
door netwerken bestaande uit simpele
rekeneenheden ***(connectionism)*** aangestuurd
door een algorithme

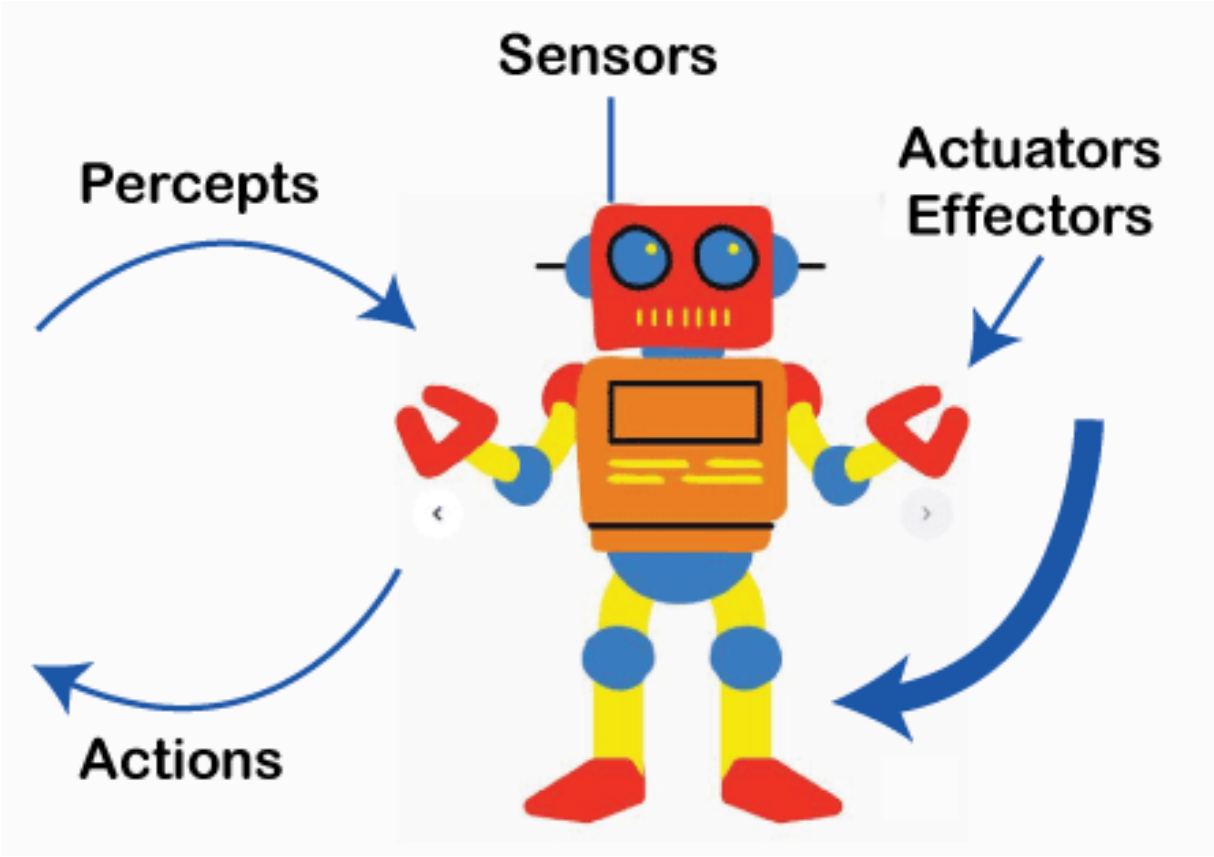


{MULTIMODALE AGENT}

HOGESCHOOL ROTTERDAM

```
AGENT representeert input / output model
```

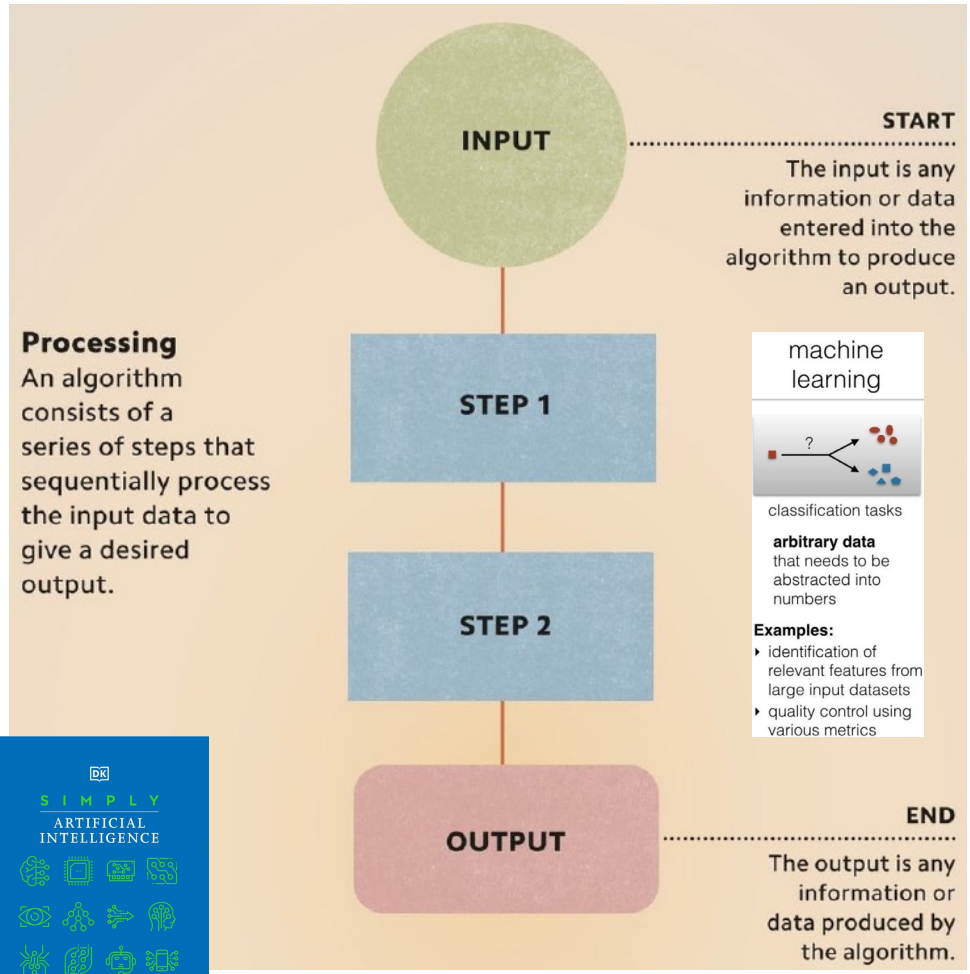
EXTERNAL INPUT

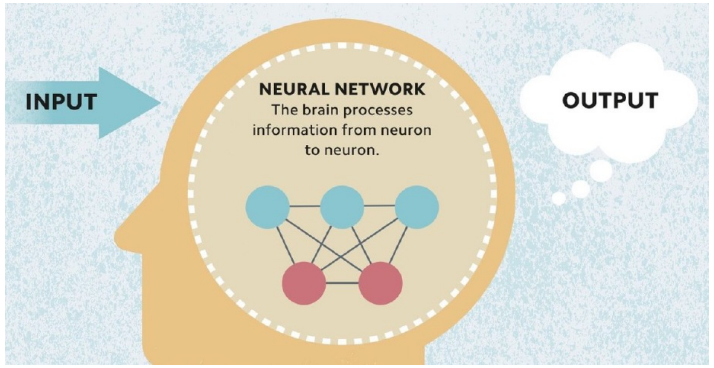


{ALGORITME}

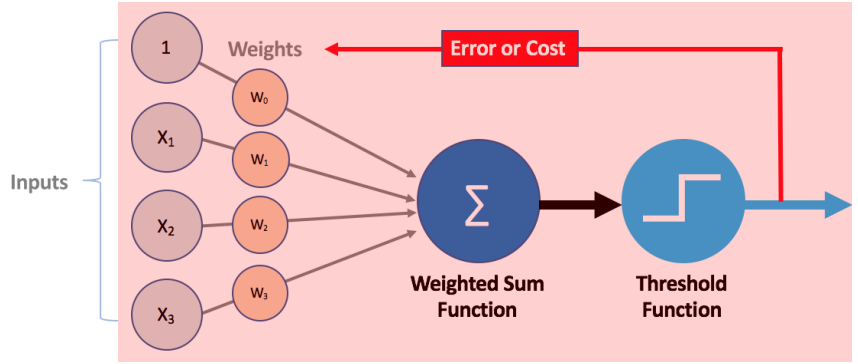
{AI}

https://learn.microsoft.com/nl-nl/dotnet/machine-learning/deep-learning-overview



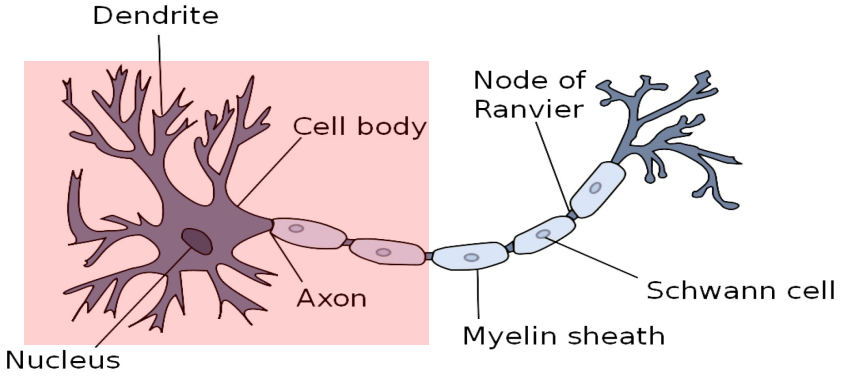


The weights are updated through a process of minimising a cost function



Names for Artificial Neurons

{unit}
{cell}
{node}
{perceptron}



{Top-down}

Top-down Encoding Capacity increases by adding hidden layers

HOGESCHOOL ROTTERDAM

What are the limits of deep learning?

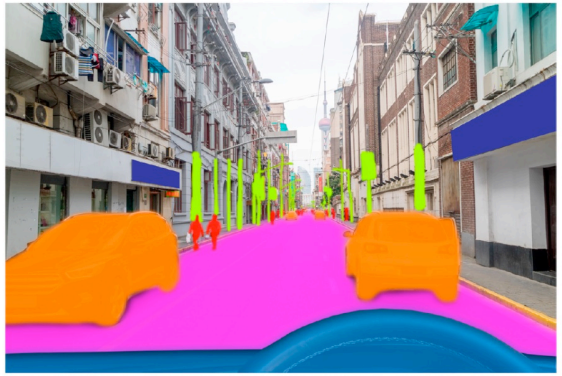
The much-ballyhooed artificial intelligence approach boasts impressive feats but still falls short of human brainpower. Researchers are determined to figure out what's missing.

M. Mitchell Waldrop, *Science Writer*

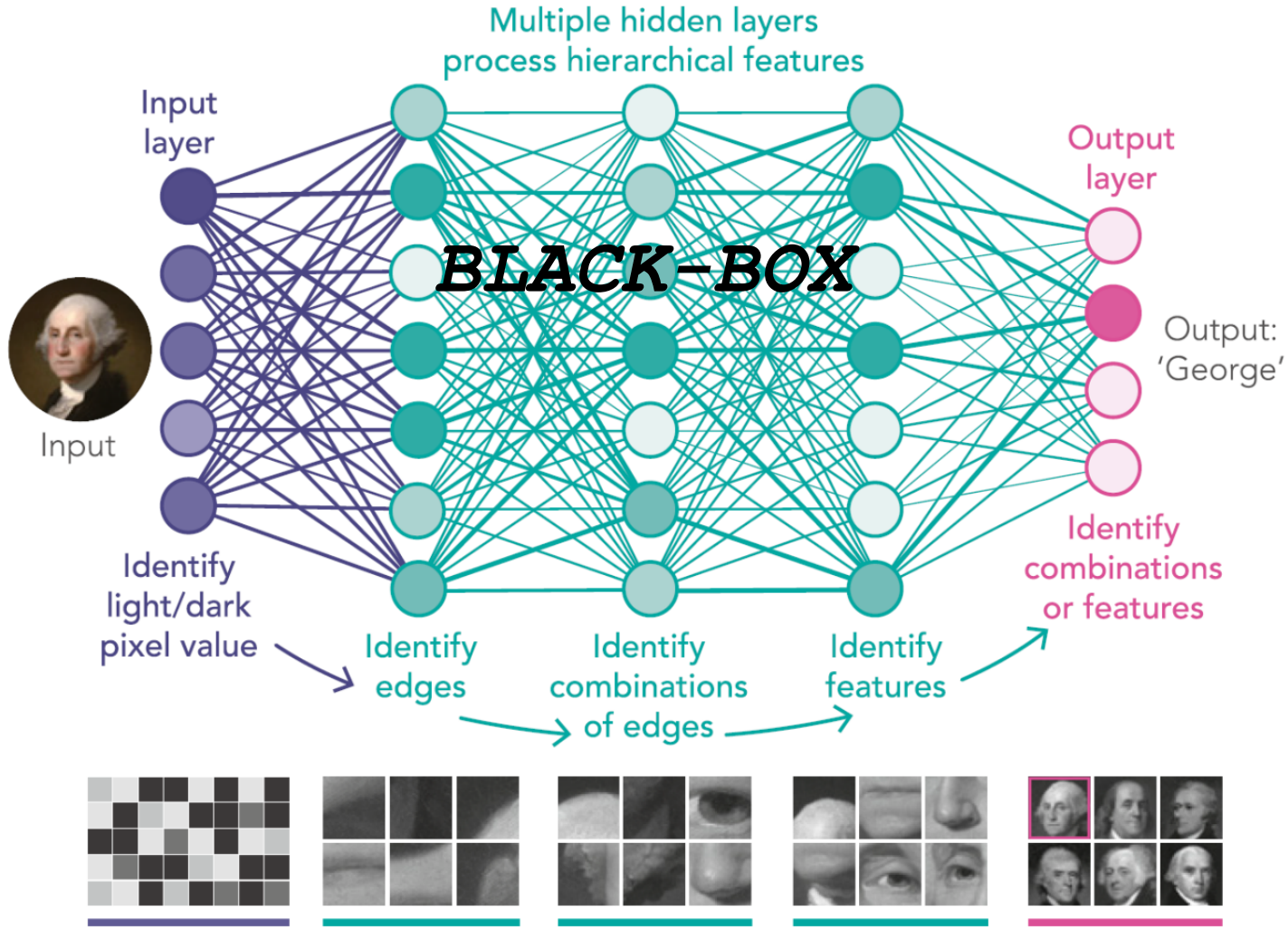
There's no mistaking the image: It's a banana—a big, ripe, bright-yellow banana. Yet the artificial intelligence (AI) identifies it as a toaster, even though it was trained with the same powerful and oft-publicized deep-learning techniques that have produced a white-hot revolution in driverless cars, speech understanding, and a multitude of other AI applications. That means the AI was shown several thousand photos of bananas, slugs, snails, and similar-looking objects, like so many flash cards, and then drilled on the answers until it had the classification down cold. And yet this advanced system was quite easily confused—all it took was a little day-glow sticker, digitally pasted in one corner of the image.

This example of what deep-learning researchers call an "adversarial attack," discovered by the Google Brain team in Mountain View, CA (1), highlights just how far AI still has to go before it remotely approaches human capabilities. "I initially thought that adversarial examples were just an annoyance," says Geoffrey Hinton, a computer scientist at the University of Toronto and one of the pioneers of deep learning. "But I now think they're probably quite profound. They tell us that we're doing something wrong."

That's a widely shared sentiment among AI practitioners, any of whom can easily rattle off a long list of deep learning's drawbacks. In addition to its vulnerability



Apparent shortcomings in deep-learning approaches have raised concerns among researchers and the general public as technologies such as driverless cars, which use deep-learning techniques to navigate, get involved in well-publicized mishaps. Image credit: Shutterstock.com/MONOPOLY919.



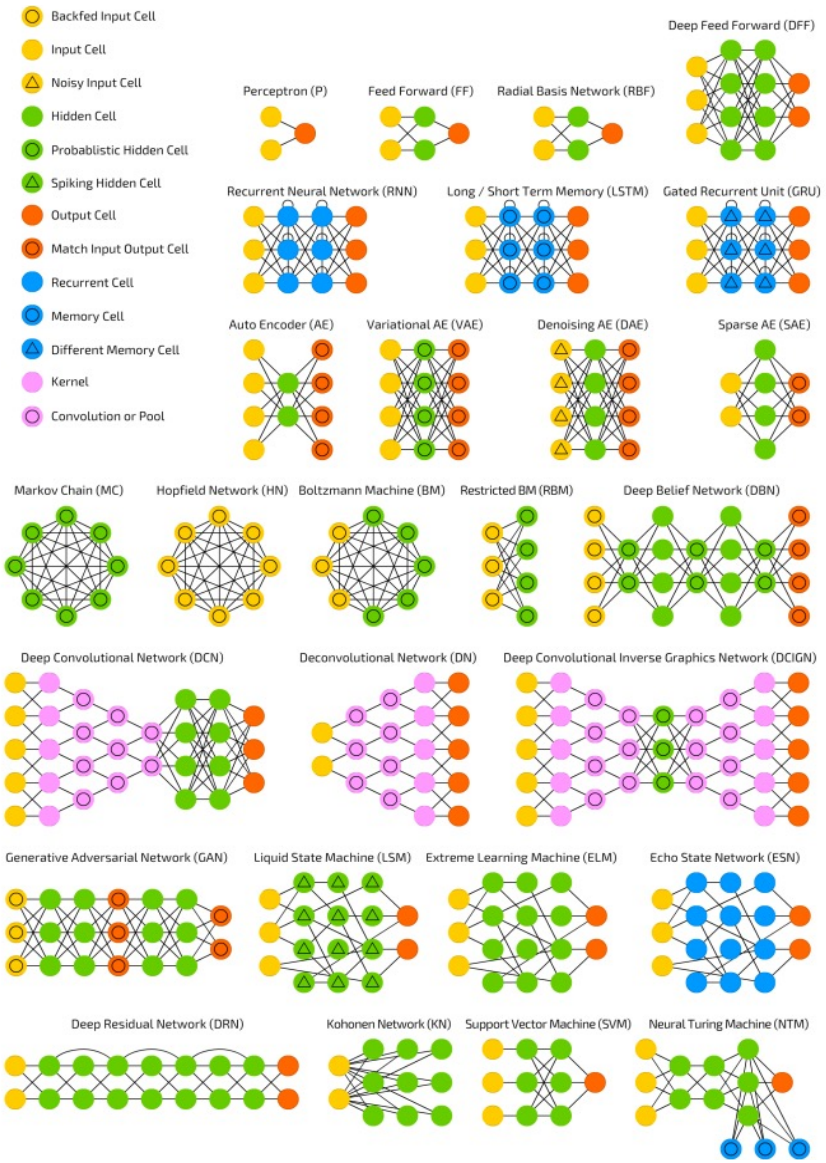
{Topology}

Topology of a neural network refers to the way artificial neurons are connected to form a network.

Form follows function!

The topology of a network determines the degree of perplexity of the tasks it can

https://pub.towardsai.net/main-types-of-neural-networks-and-its-applications-tutorial-734480d7ec8e



https://www.mdpi.com/2504-3900/47/1/9

HOGESCHOOL ROTTERDAM
CREATING

{Big-data}

Big-data is needed to avoid hand-crafted feature extraction

HOGESCHOOL ROTTERDAM

CREATING

A Unified Approach to Interpreting Model Predictions

Scott M. Lundberg
Paul G. Allen School of Computer Science
University of Washington
Seattle, WA 98105
slund1@cs.washington.edu

Su-In Lee
Paul G. Allen School of Computer Science
Department of Genome Sciences
University of Washington
Seattle, WA 98105
suinlee@cs.washington.edu

Abstract

Understanding why a model makes a certain prediction can be as crucial as the prediction's accuracy in many applications. However, the highest accuracy for large modern datasets is often achieved by complex models that even experts struggle to interpret, such as ensemble or deep learning models, creating a tension between *accuracy* and *interpretability*. In response, various methods have recently been proposed to help users interpret the predictions of complex models, but it is often unclear how these methods are related and when one method is preferable over another. To address this problem, we present a unified framework for interpreting predictions, SHAP (SHapley Additive exPlanations). SHAP assigns each feature an importance value for a particular prediction. Its novel components include: (1) the identification of a new class of additive feature importance measures, and (2) theoretical results showing there is a unique solution in this class with a set of desirable properties. The new class unifies six existing methods, notable because several recent methods in the class lack the proposed desirable properties. Based on insights from this unification, we present new methods that show improved computational performance and/or better consistency with human intuition than previous approaches.

1 Introduction

The ability to correctly interpret a prediction model's output is extremely important. It engenders appropriate user trust, provides insight into how a model may be improved, and supports understanding of the process being modeled. In some applications, simple models (e.g., linear models) are often preferred for their ease of interpretation, even if they may be less accurate than complex ones. However, the growing availability of big data has increased the benefits of using complex models, so bringing to the forefront the trade-off between accuracy and interpretability of a model's output. A wide variety of different methods have been recently proposed to address this issue [5, 8, 9, 3, 4, 1]. But an understanding of how these methods relate and when one method is preferable to another is still lacking.

Here, we present a novel unified approach to interpreting model predictions.[1] Our approach leads to three potentially surprising results that bring clarity to the growing space of methods:

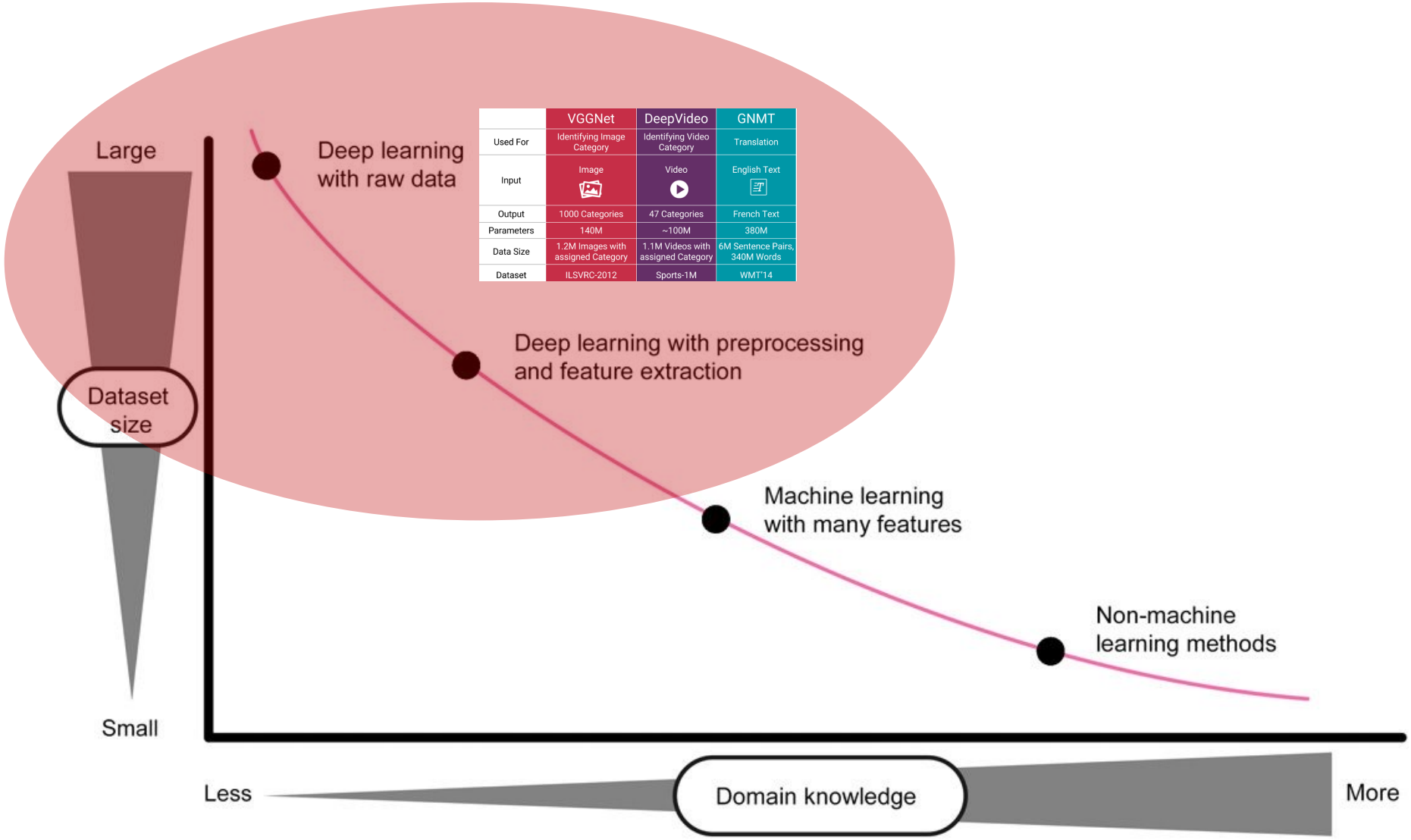
1. We introduce the perspective of viewing *any* explanation of a model's prediction as a model itself, which we term the *explanation model*. This lets us define the class of *additive feature attribution methods* (Section 2), which unifies six current methods.

[1] https://github.com/slundberg/shap

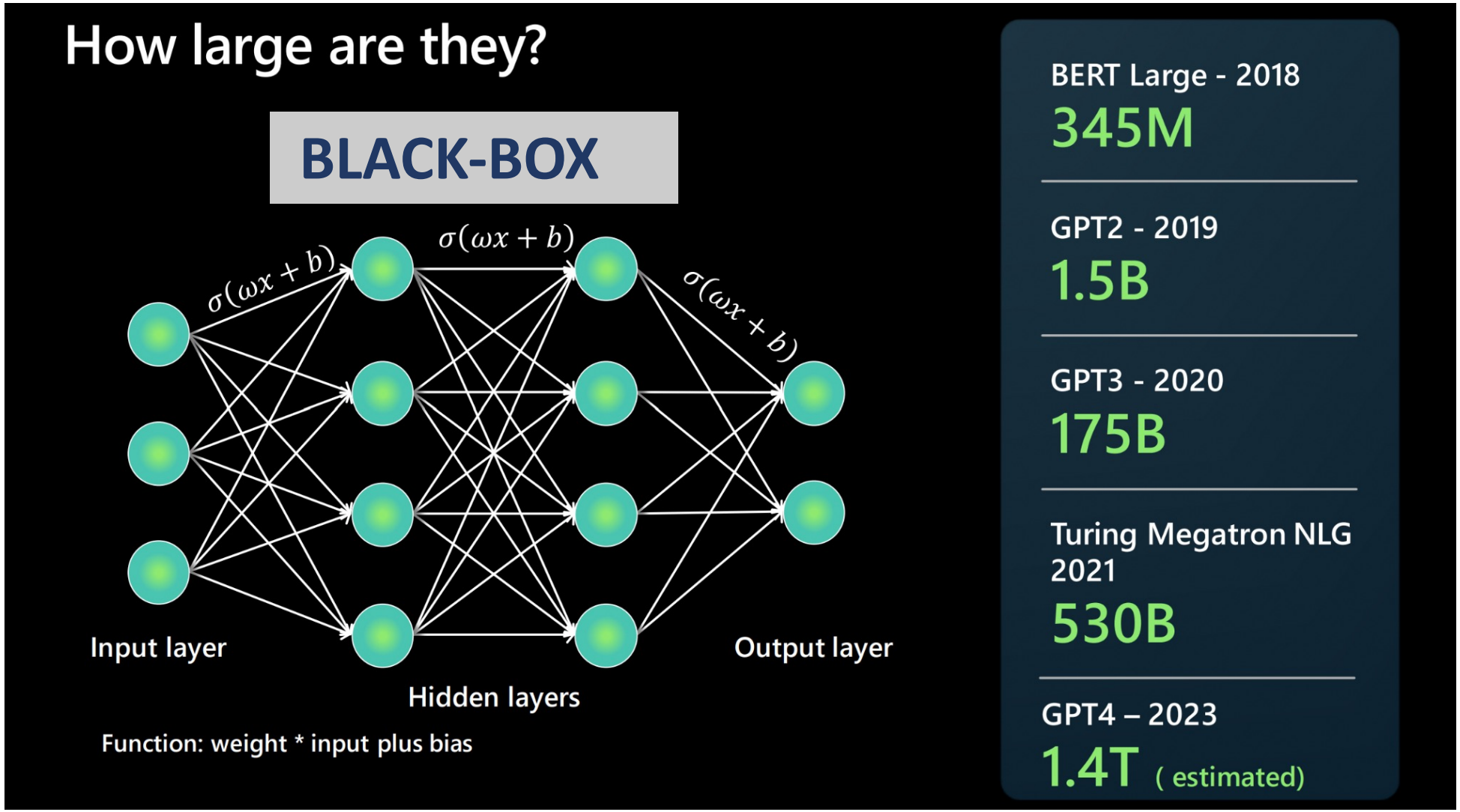
31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.

https://proceedings.neurips.cc/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf

	VGGNet	DeepVideo	GNMT
Used For	Identifying Image Category	Identifying Video Category	Translation
Input	Image	Video	English Text
Output	1000 Categories	47 Categories	French Text
Parameters	140M	~100M	380M
Data Size	1.2M Images with assigned Category	1.1M Videos with assigned Category	6M Sentence Pairs, 340M Words
Dataset	ILSVRC-2012	Sports-1M	WMT'14



{Diep Neuraal Netwerk [NN] onstaat “vanzelf”}



Het gebruik van in onderwijs en onderzoek grote taal modellen is problematisch

Generatieve AI

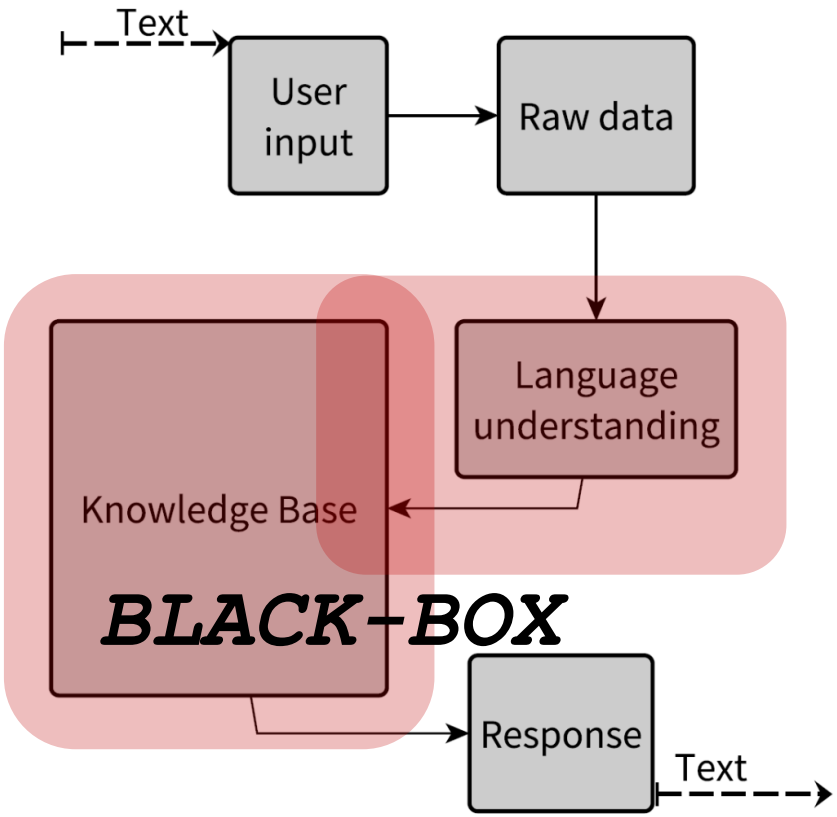
Agents gebaseerd op Neurale netwerk modellen die machinaal hebben geleerd op basis van bestaande multimodale content (trainen van het model)
nieuwe inhoud te creëren, zoals tekst, afbeeldingen, muziek en code.

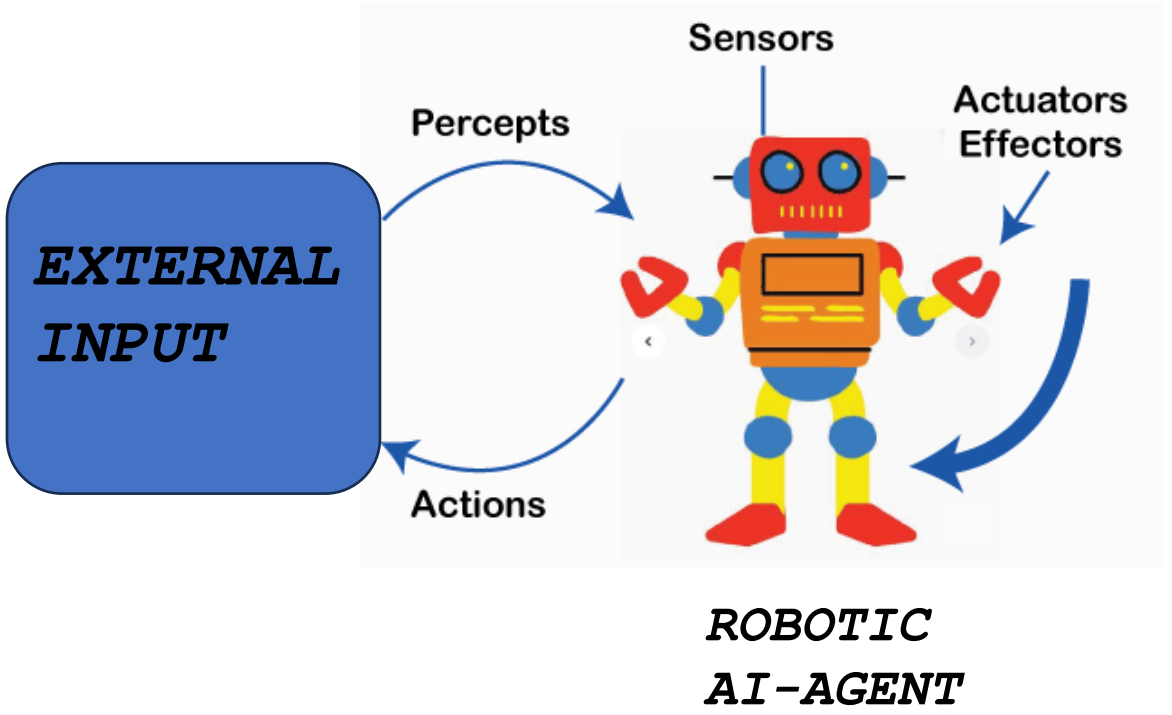
Generatieve AI zijn geen informatiedatabases of deterministische informatiezoeksystemen, omdat het voorspellingssystemen zijn.

Gen-AI maakt dus geen onderscheid tussen goed/fout of waar/niet-waar maar produceert een uitkomst die met grote waarschijnlijkheid kan worden gerelateerd aan de geven input (prompt).

https://ai.google/static/documents/google-about-generative-ai.pdf

ChatGPT is een Conversationele tekst-in/tekst-uit ChatBot



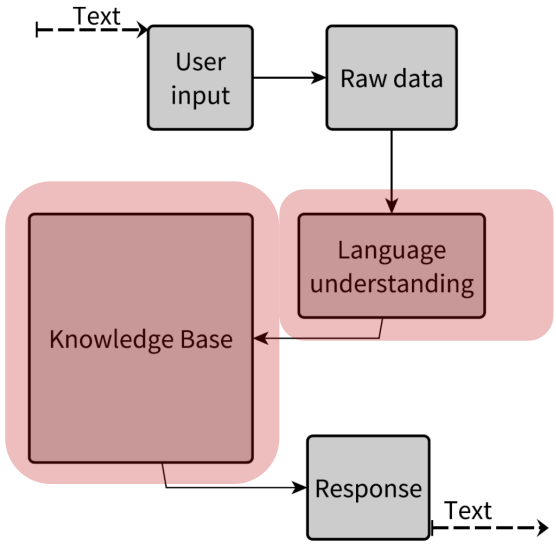


ROBOTIC AI-AGENT

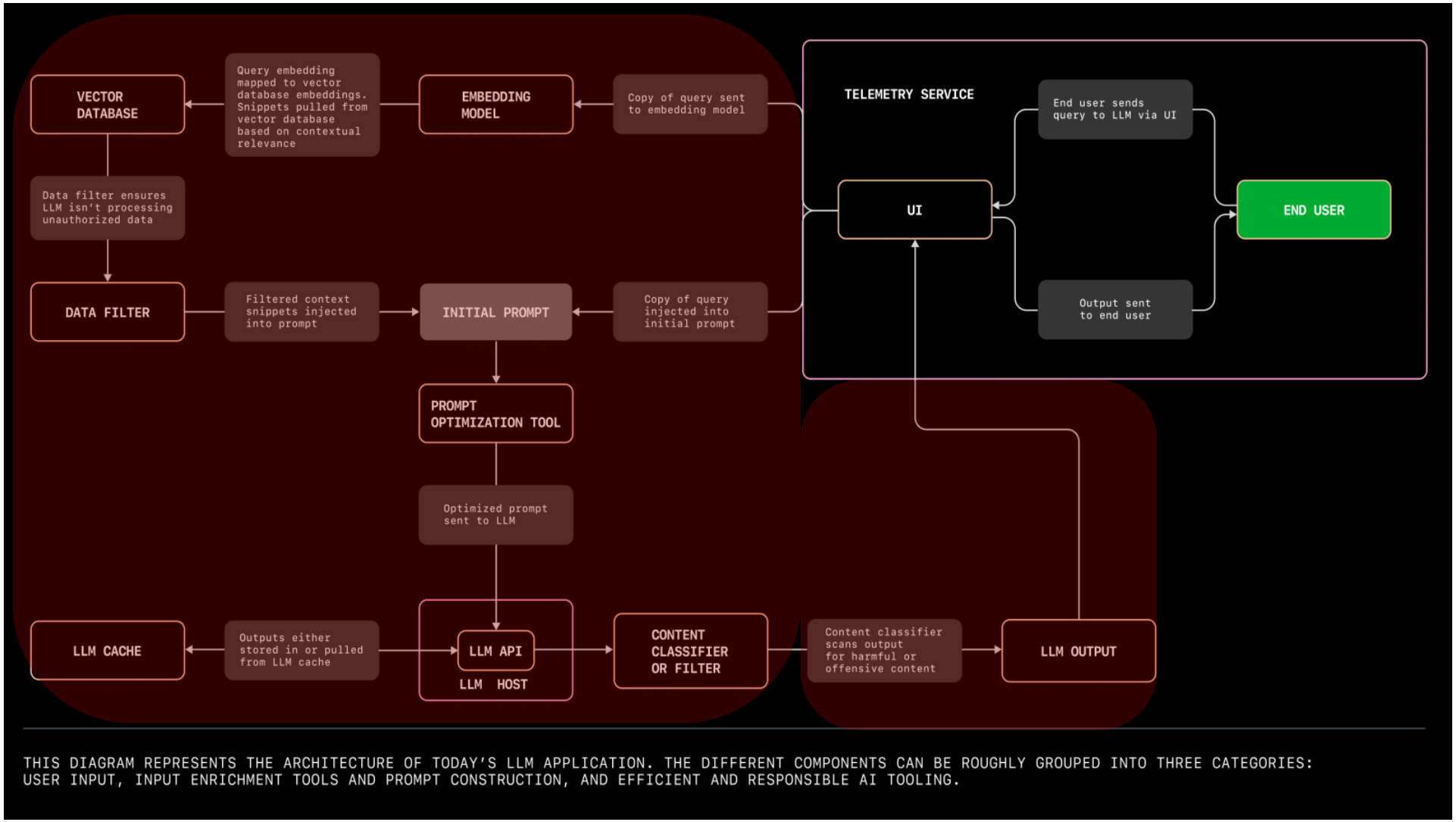
Sánchez-Díaz, X., Ayala-Bastidas, G., Fonseca-Ortiz, P., & Garrido, L. (2018). A knowledge-based methodology for building a conversational chatbot as an intelligent tutor. https://doi.org/10.1007/978-3-030-04497-8_14

ChatGPT == 99% BLACKBOX + 1% user-interface

HOGESCHOOL ROTTERDAM



BLACK-BOX



https://blogs.novita.ai/exploring-architectural-structures-and-functional-capacities-of-llms/

{Skewed}

HOGESCHOOL ROTTERDAM

Big-Data is Inherently Skewed

When artificial intelligence learns from humans, it's bad

James Crowder*

Learned bias

ML algorithms are not biased in and of themselves; they *learn* to be biased. This *algorithmic bias* (Danks and London 2017) has received a great deal of attention. It occurs when the learning algorithm is trained on biased datasets and subsequently "accurately" learns the patterns of bias inherent in the data (see, e.g., Caliskan et al. 2017). In some cases, the learned representations within ML algorithms can even exaggerate these biases (Zhao et al. 2017). Algorithmic bias has two sources: incomplete datasets and datasets that represent biased social phenomena.

Incomplete datasets are those that are not representative of the entire range of potential examples. Consequently, an algorithm trained on an incomplete dataset will perform poorly when given an example that falls outside the scope of the available data. For example, a facial recognition algorithm that is not trained on a wide variety of skin colours might not function accurately for faces of all skin tones. Indeed, one AI researcher with dark skin discovered that an otherwise functional facial recognition algorithm failed to recognize her face unless she put on a white mask[45] . Many similar examples have been reported.[46]

When critical decisions are made based on input from an algorithm trained on a database that is not representative of the entire user population, the results can impact health and well-being. Most clinical trials have highly selective criteria that exclude women (especially pregnant women), the elderly, and those with conditions beyond those being studied. Thus, participants tend to be white males.[47] In some cases, the findings of these studies do not generalize well across the broader population, with outcomes potentially compromised for individuals not represented in the research. As a result, AI algorithms trained on this data are

* bit.ly/2HFetvY

Photo by Marius Masalar

{Disparities}

Big Data causes racial & gender disparities

The Race Gap in Speech Recognition Technology

Researchers from Stanford University found that state-of-the-art ASR systems exhibited significant racial and gender disparity—they misunderstand Black speakers twice as often as white speakers. In the paper, titled "Racial Disparities in Automated Speech Recognition," authors ran thousands of audio snippets of white and Black speakers, transcribed from interviews conducted with 42 white speakers and 73 Black speakers, through leading speech-to-text services by Amazon, Apple, Google, IBM, and Microsoft.

The results suggest that, on average, systems made 19 errors every hundred words for white speakers and 35 errors for Black speakers—nearly twice as many. Moreover, the systems performed particularly poorly for Black men, with more than 40 errors for every hundred words (Figure 2.5.3). The breakdown by ASR systems shows that gaps are similar across companies (Figure 2.5.4). This research emphasizes the importance of addressing the bias of AI technologies and ensuring equity as they become mature and deployed.

TESTINGS on LEADING SPEECH-to-TEXT SERVICES: WORD ERROR RATE by RACE and GENDER, 2019

Source: Koenecke et al., 2020 | Chart: 2021 AI Index Report

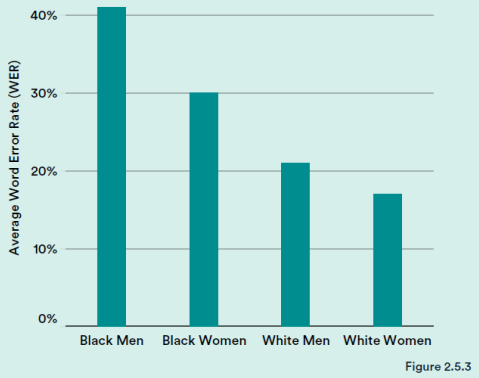


Figure 2.5.3

TESTINGS on LEADING SPEECH-to-TEXT SERVICES: WORD ERROR RATE by SERVICE and RACE, 2019

Source: Koenecke et al., 2020 | Chart: 2021 AI Index Report

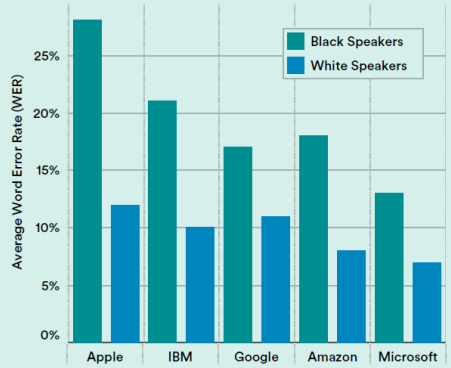


Figure 2.5.4

{Augmentation}

HOGESCHOOL ROTTERDAM

Big Data that is *not* augmented causes Overfitting

SURVEY PAPER Open Access

A survey on Image Data Augmentation for Deep Learning

Connor Shorten* and Taghi M. Khoshgoftaar

*Correspondence:
cshorten2015@fau.edu
Department of Computer and Electrical Engineering and Computer Science, Florida Atlantic University, Boca Raton, USA

Abstract

Deep convolutional neural networks have performed remarkably well on many Computer Vision tasks. However, these networks are heavily reliant on big data to avoid overfitting. Overfitting refers to the phenomenon when a network learns a function with very high variance such as to perfectly model the training data. Unfortunately, many application domains do not have access to big data, such as medical image analysis. This survey focuses on Data Augmentation, a data-space solution to the problem of limited data. Data Augmentation encompasses a suite of techniques that enhance the size and quality of training datasets such that better Deep Learning models can be built using them. The image augmentation algorithms discussed in this survey include geometric transformations, color space augmentations, kernel filters, mixing images, random erasing, feature space augmentation, adversarial training, generative adversarial networks, neural style transfer, and meta-learning. The application of augmentation methods based on GANs are heavily covered in this survey. In addition to augmentation techniques, this paper will briefly discuss other characteristics of Data Augmentation such as test-time augmentation, resolution impact, final dataset size, and curriculum learning. This survey will present existing methods for Data Augmentation, promising developments, and meta-level decisions for implementing Data Augmentation. Readers will understand how Data Augmentation can improve the performance of their models and expand limited datasets to take advantage of the capabilities of big data.

Keywords: Data Augmentation, Big data, Image data, Deep Learning, GANs

Introduction

Deep Learning models have made incredible progress in discriminative tasks. This has been fueled by the advancement of deep network architectures, powerful computation, and access to big data. Deep neural networks have been successfully applied to Computer Vision tasks such as image classification, object detection, and image segmentation thanks to the development of convolutional neural networks (CNNs). These neural networks utilize parameterized, sparsely connected kernels which preserve the spatial characteristics of images. Convolutional layers sequentially downsample the spatial resolution of images while expanding the depth of their feature maps. This series of convolutional transformations can create much lower-dimensional and more useful representations of images than what could possibly be hand-crafted. The success of CNNs has spiked interest and optimism in applying Deep Learning to Computer Vision tasks.

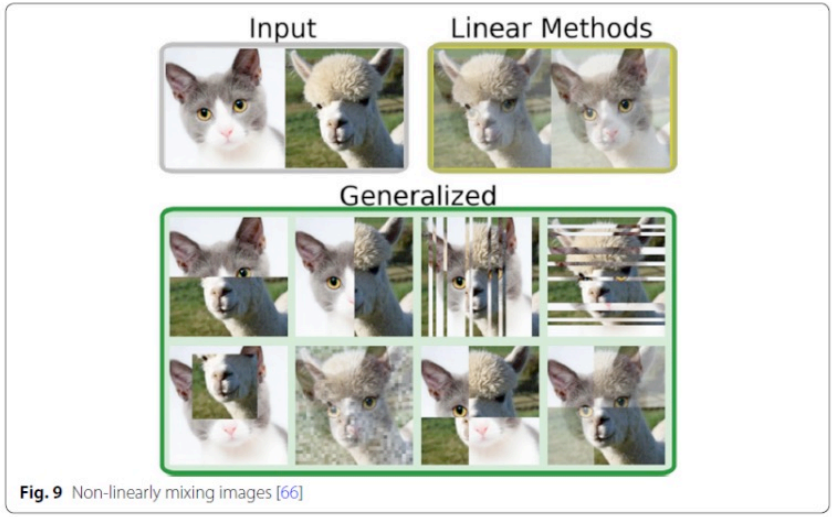
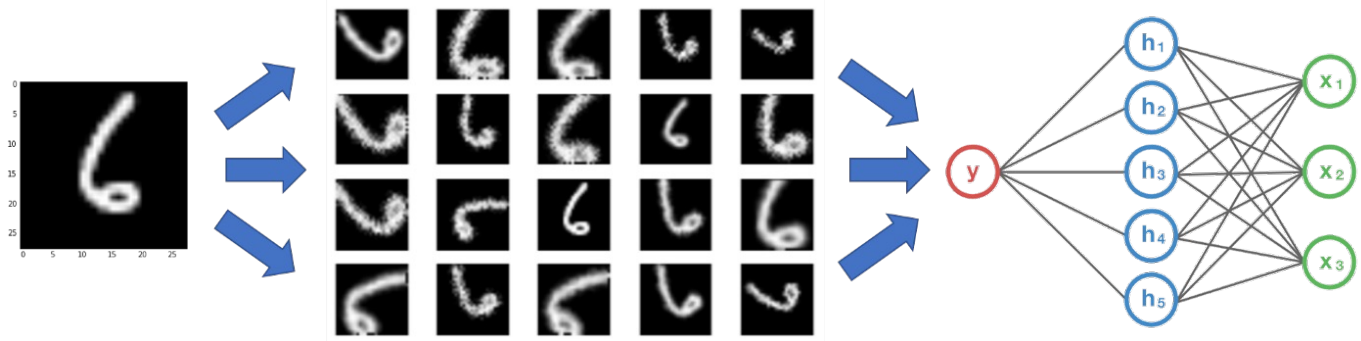
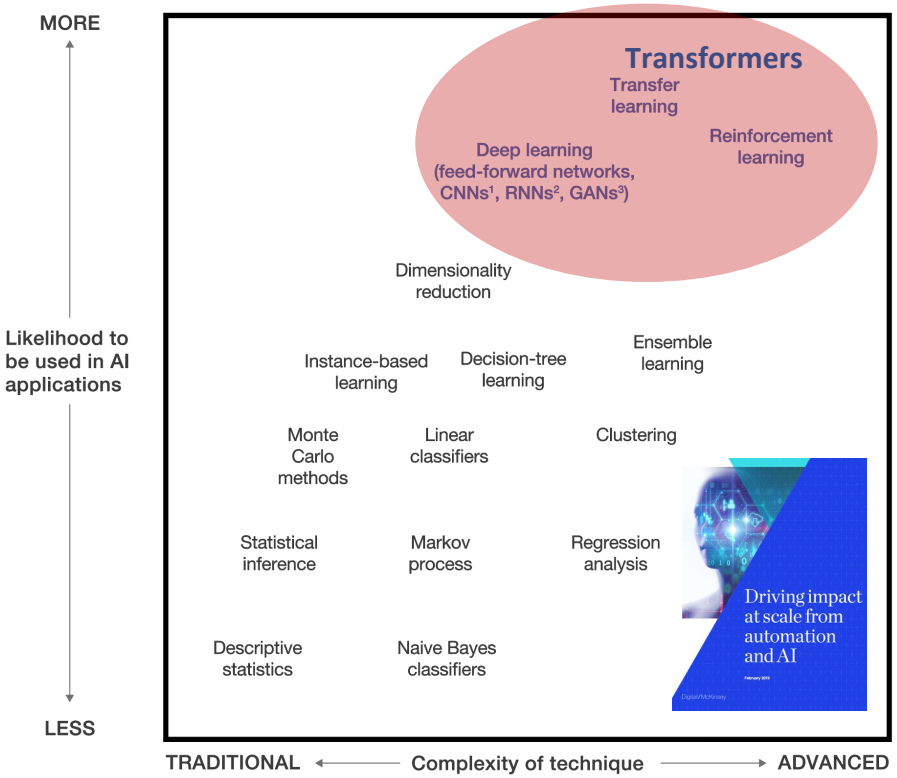


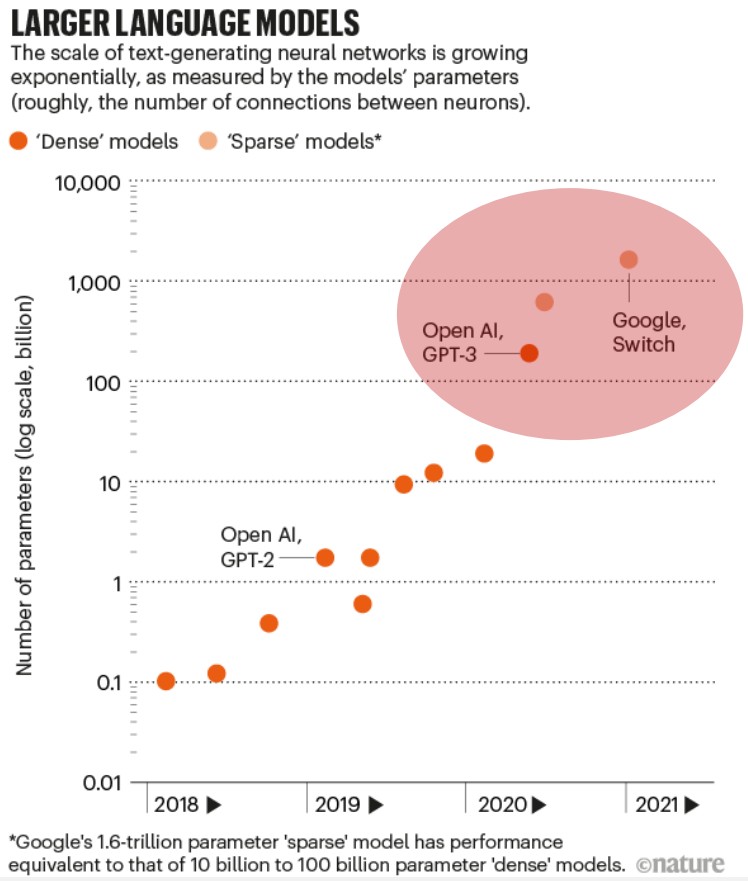
Fig. 9 Non-linearly mixing images [66]



{large scale}

Only very large scale {DNNs} are useful [can compete with human performance]

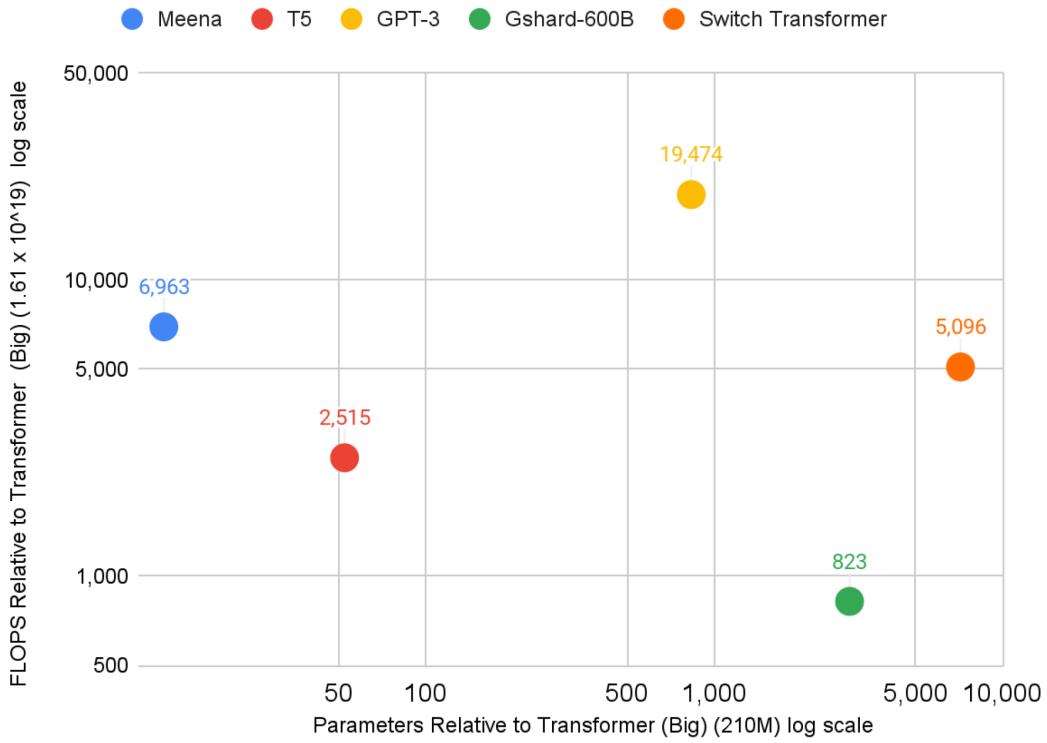


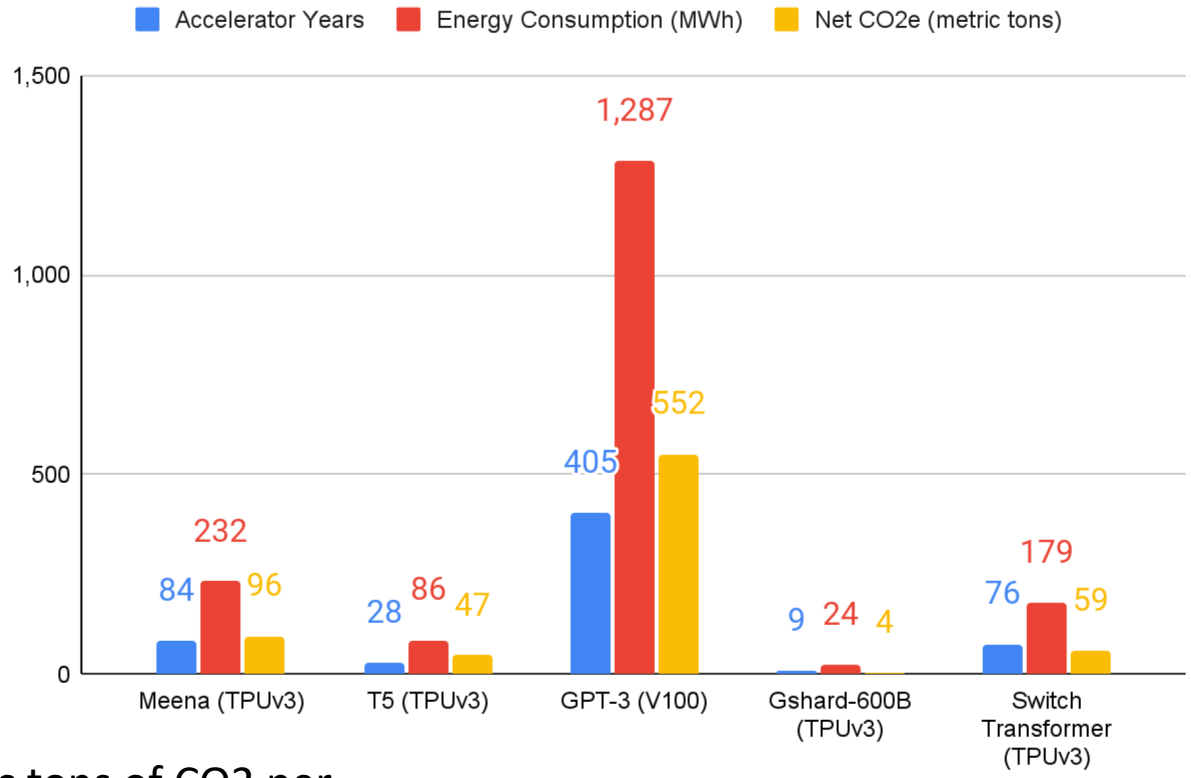


https://www.nature.com/articles/d41586-021-00530-0

{CO2 foot-print}

Training large scale transformer {DNNs} produce massive Carbon Emissions





As of 2007, the average U.S. household emits 20 metric tons of CO2 per year. In comparison to a world average of 4 tons.

Carbon Footprint_CSS09-05_e2021.pdf (umich.edu)

https://arxiv.org/ftp/arxiv/papers/2104/2104.10350.pdf

{computational unsustainability}

HOGESCHOOL ROTTERDAM

The scale of state-of-the-art {SOTA} –near human level– DNNs
– combined with a blind Brute-Force implementation + post-hoc analysis –
is becoming more and more
computationally unsustainable,
even to the point
that hypernetworks are employed
to help humans to make DDNs work.

2110.13100v1.pdf (arxiv.org)

https://paperswithcode.com/sota/

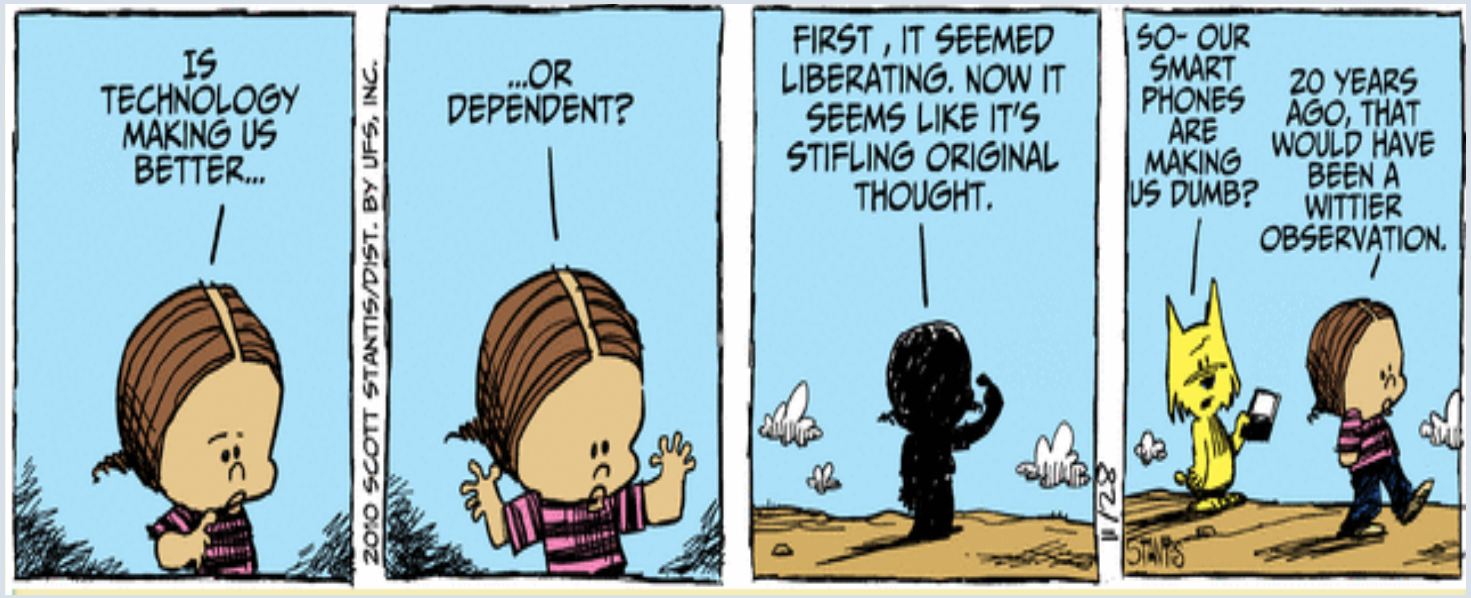
{biased towards coherence}

HOGESCHOOL ROTTERDAM

LLMs are pattern-followers,
not poets or truth-finders in any deep sense,
so "poetry," repetition, and odd keyword placement affect them by nudging which statistical patterns they continue, not by making them more truthful.

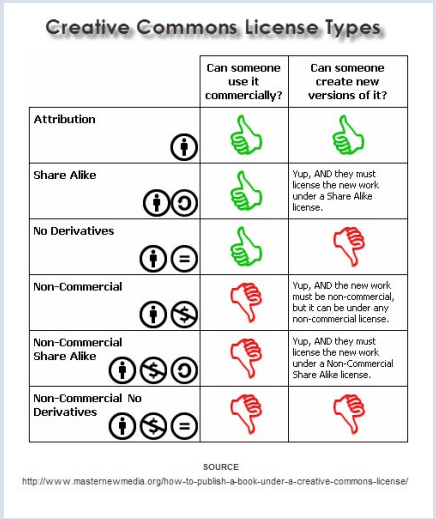
Their apparent sensitivity to style, rhyme, or scattered key phrases exposes how much they optimize for coherence and fluency over accuracy or epistemic care.

https://www.perplexity.ai/search/curious-facts-about-llms-poetr-Qk7SHS1ARKCPT3YLTMK0MQ#0



This seminar was developed by:
Programma AI & Ethiek
Lead-Tech: Rob van der Willigen

JANUARI 2026



Creative Commons License Types

	Can someone use it commercially?	Can someone create new versions of it?
Attribution	👍	👍
Share Alike	👍	Yup, AND they must license the new work under a Share Alike license.
No Derivatives	👍	👎
Non-Commercial	👎	Yup, AND the new work must be non-commercial, but it can be under any non-commercial license.
Non-Commercial Share Alike	👎	Yup, AND they must license the new work under a Non-Commercial Share Alike license.
Non-Commercial No Derivatives	👎	👎

SOURCE
http://www.masternewmedia.org/how-to-publish-a-book-under-a-creative-commons-license/

{Knowledge Dissemination & Curation}

HOGESCHOOL ROTTERDAM

Hight quality, insightful Dutch reviews on AI