



EXPLORE  
NEW  
HORIZONS

# TRANSFORMATION WORLD 2023

**Transforming Business: Generative AI as a new operating system - Software becomes Business.**

Peter Gentsch – 15-06-2023





# **Generative AI: Development and Building Blocks**



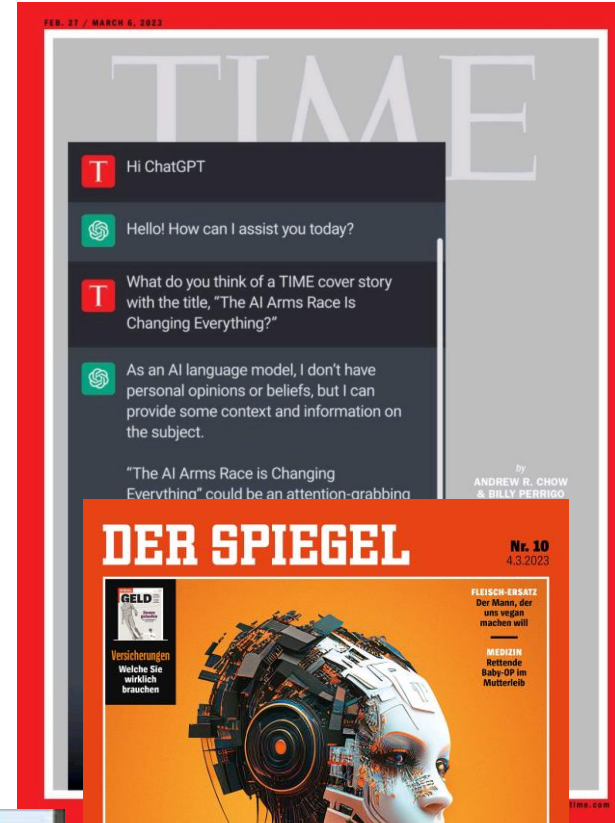
# Generative AI is as Mega-Trend in all the media

**Sam Altman** @sama  
ChatGPT launched on wednesday. today it crossed 1 million users!  
7:35 AM · Dec 5, 2022



**Linus Beličunas** @linas.beliunas  
Time it took to reach **1 million users**:

- Netflix - 3.5 years
- Airbnb - 2.5 years
- Facebook - 10 months
- Spotify - 5 months
- Instagram - 2.5 months
- iPhone - 74 days
- ChatGPT - 5 days

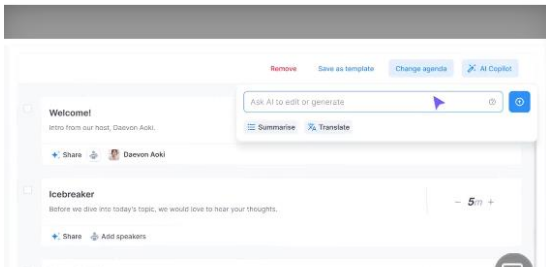


**Generative AI**  
833015 Follower:innen  
47 Min · 🌐

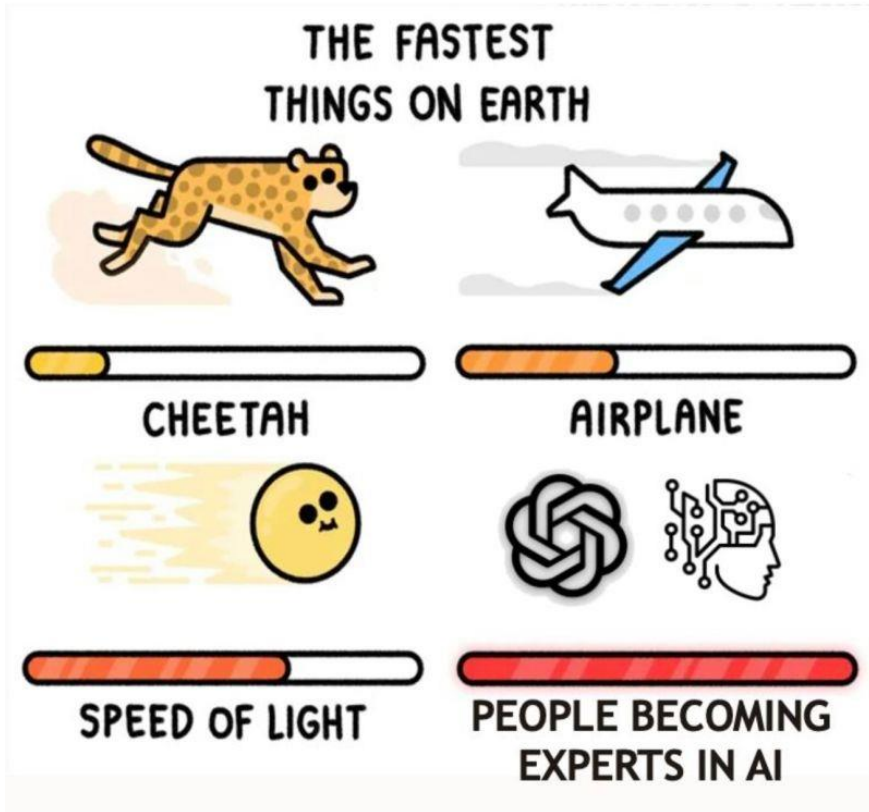
Sessions AI Copilot just took off 📈 📊

Introducing the world's most advanced AI webinar assistant. Create an ... mehr anzeigen

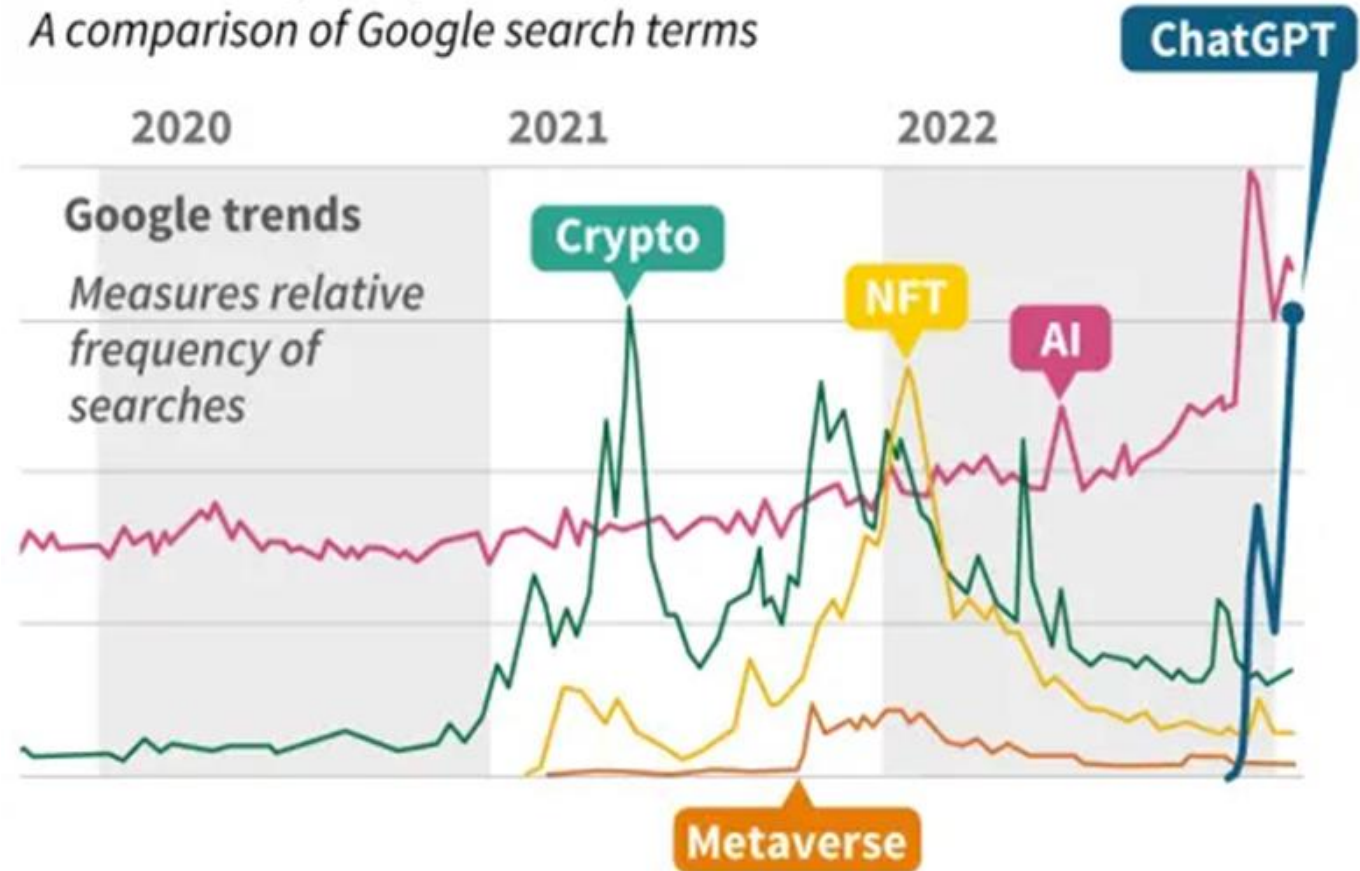
Übersetzung anzeigen



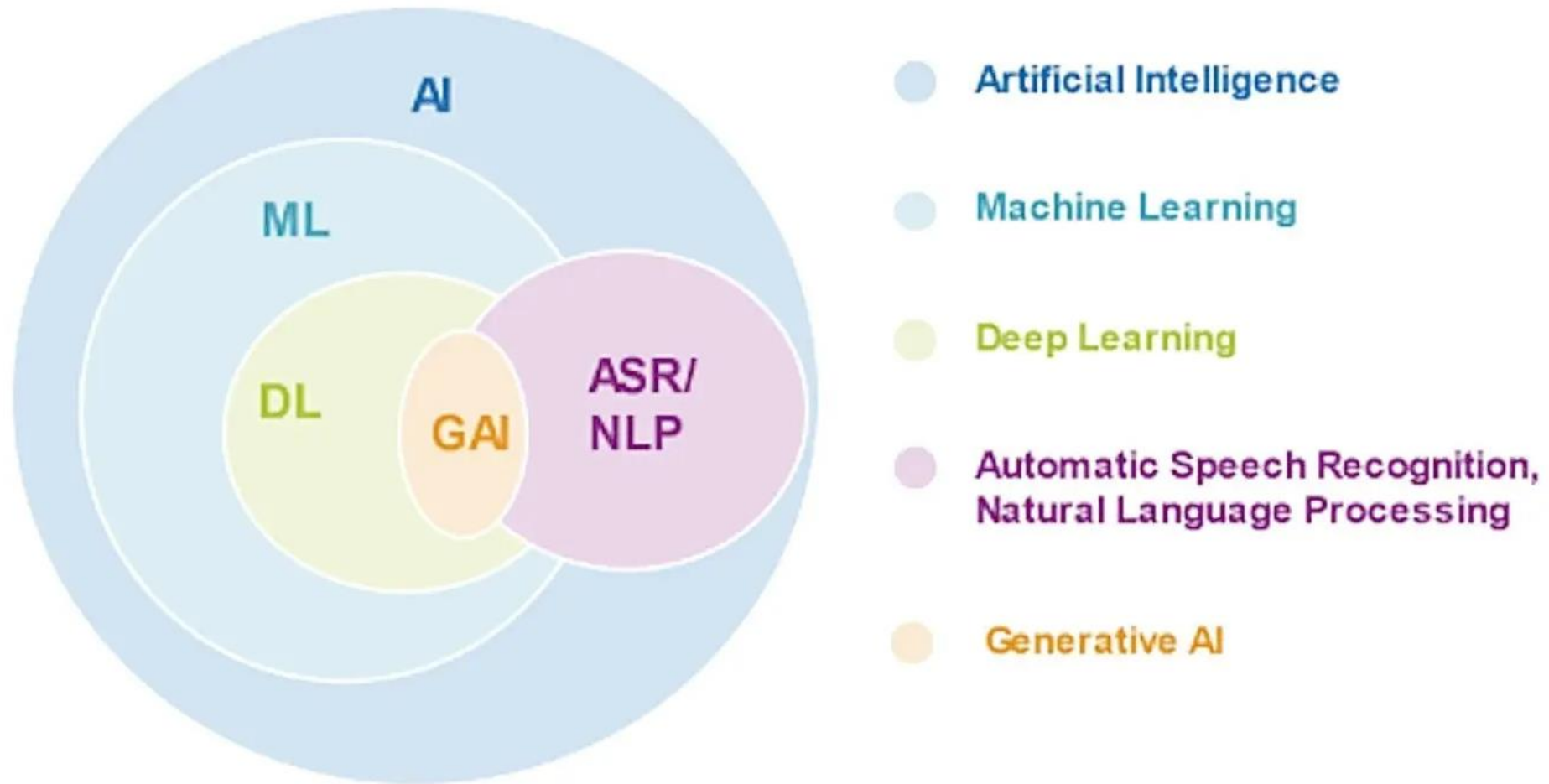
# ChatGPT & Co – the hype is on!



A comparison of Google search terms



# ChatGPT & Co as Generative AI is just one part of AI





# Google set up 2017 foundation for Transformer Models

Google Research

Philosophy

Research Areas

Publications

People

Tools & Downloads

BLOG >

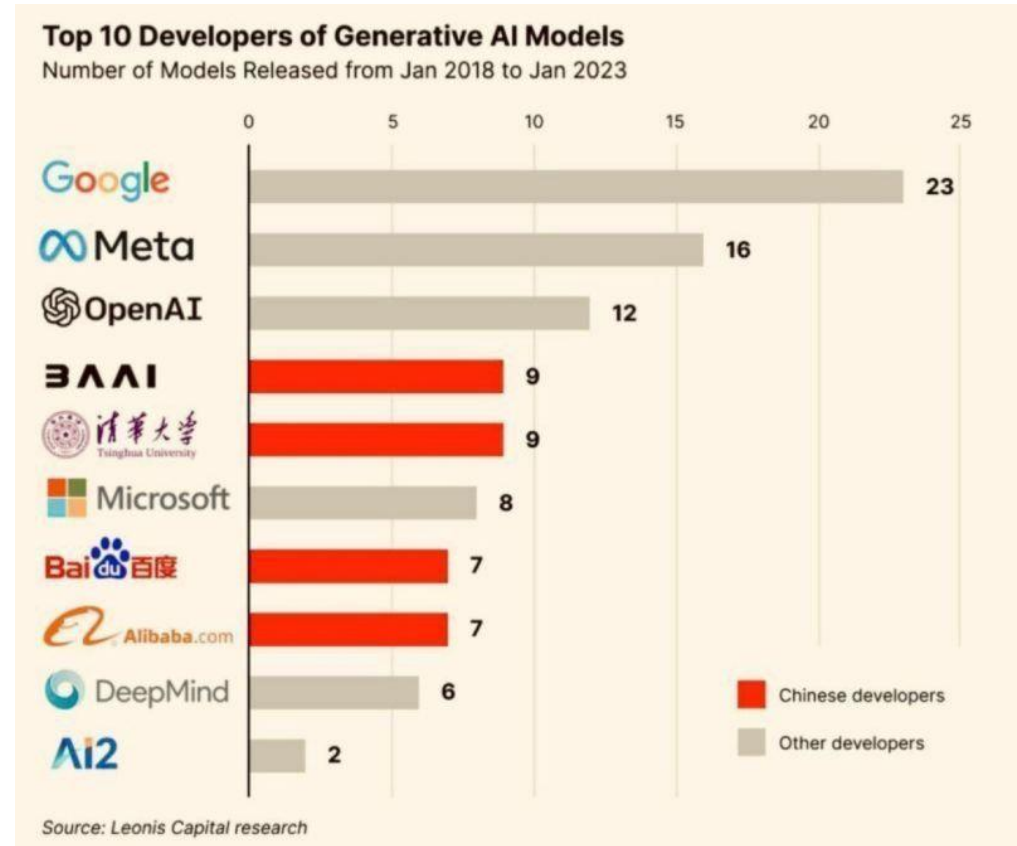
## Transformer: A Novel Neural Network Architecture for Language Understanding

THURSDAY, AUGUST 31, 2017

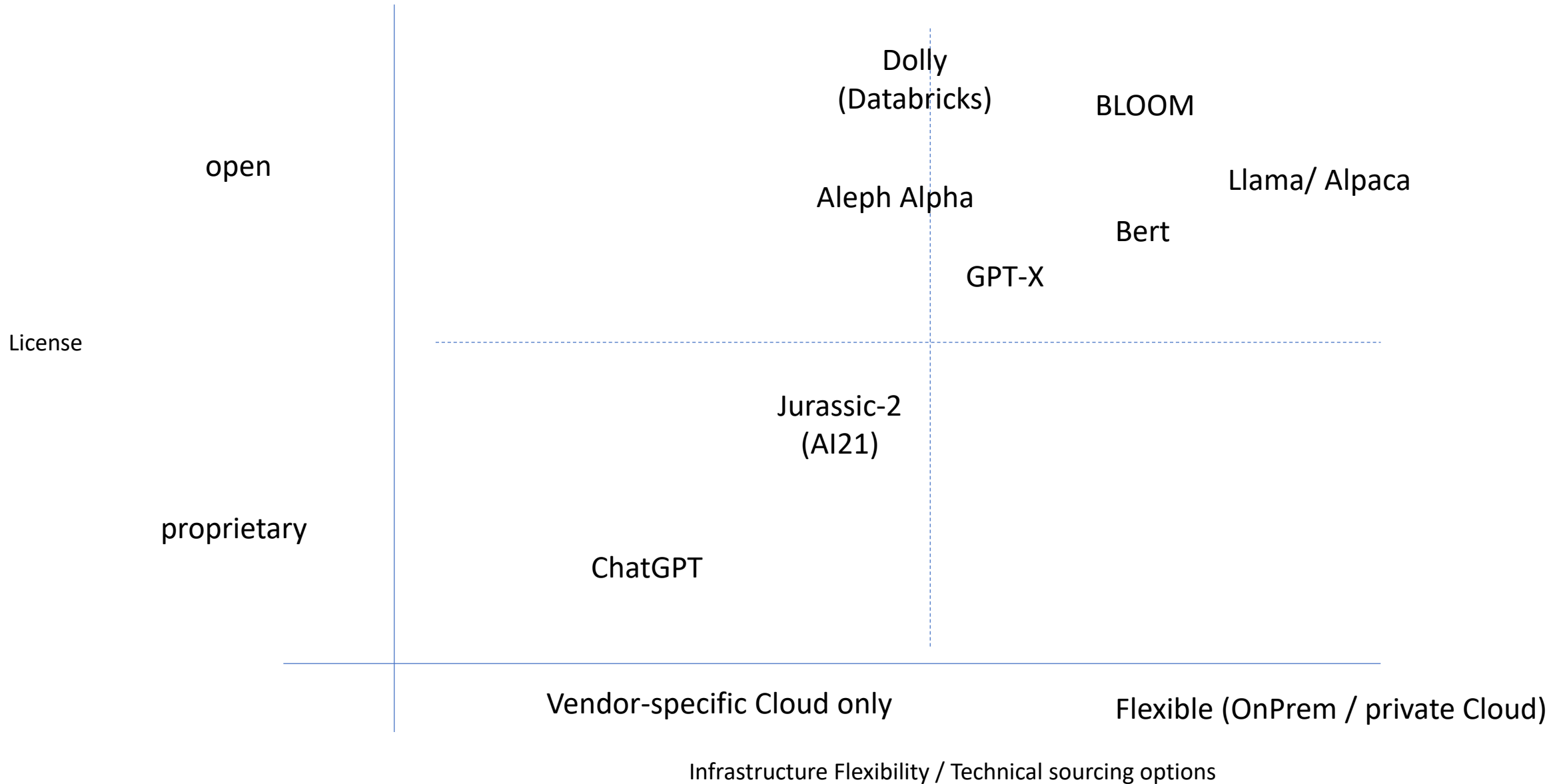
Posted by Jakob Uszkoreit, Software Engineer, Natural Language Understanding

Neural networks, in particular [recurrent neural networks](#) (RNNs), are now at the core of the leading approaches to language understanding tasks such as [language modeling](#), [machine translation](#) and [question answering](#). In “[Attention Is All You Need](#)”, we introduce the Transformer, a novel neural network architecture based on a self-attention mechanism that we believe to be particularly well suited for language understanding.

In our paper, we show that the Transformer outperforms both recurrent and convolutional models on academic English to German and English to French translation benchmarks. On top of higher translation quality, the Transformer requires less computation to train and is a much better fit for modern machine learning hardware, speeding up training by up to an order of magnitude.

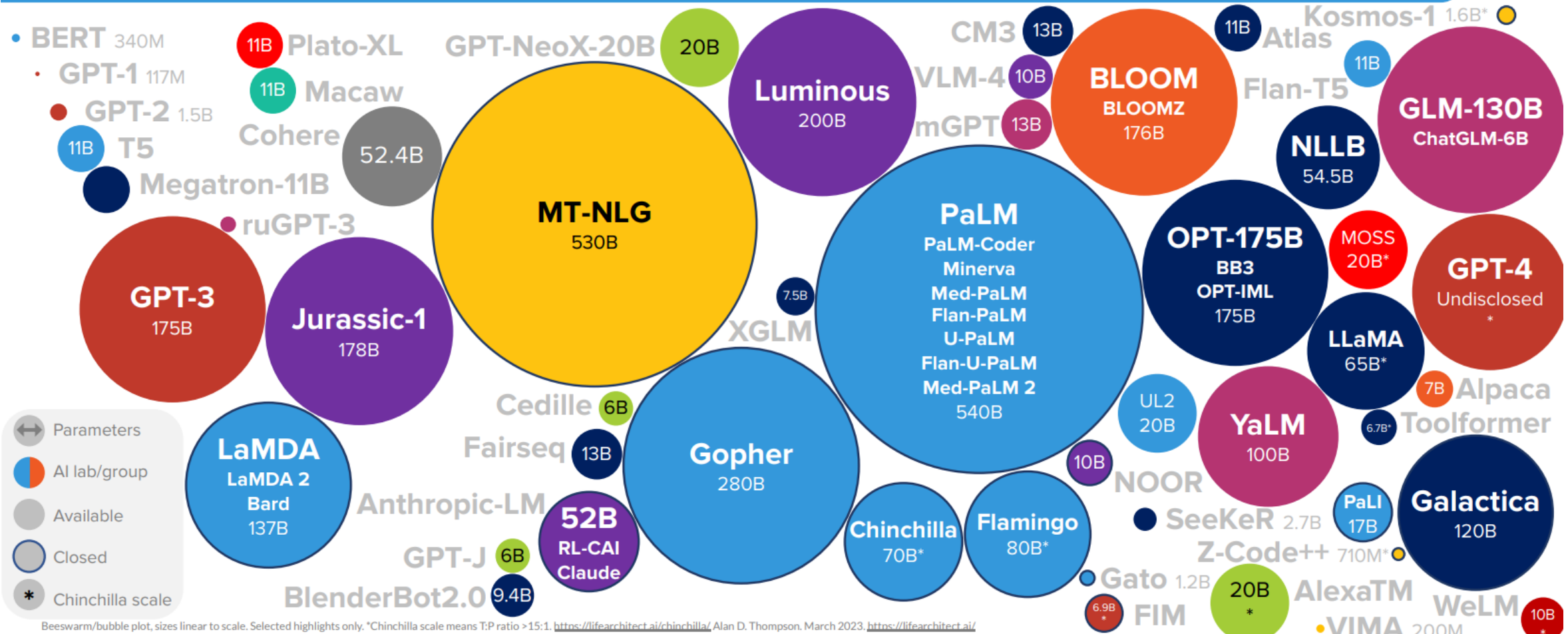


# LLM Landscape: License / Infrastructure Flexibility



# Global Large Language Model (LLM) Landscape

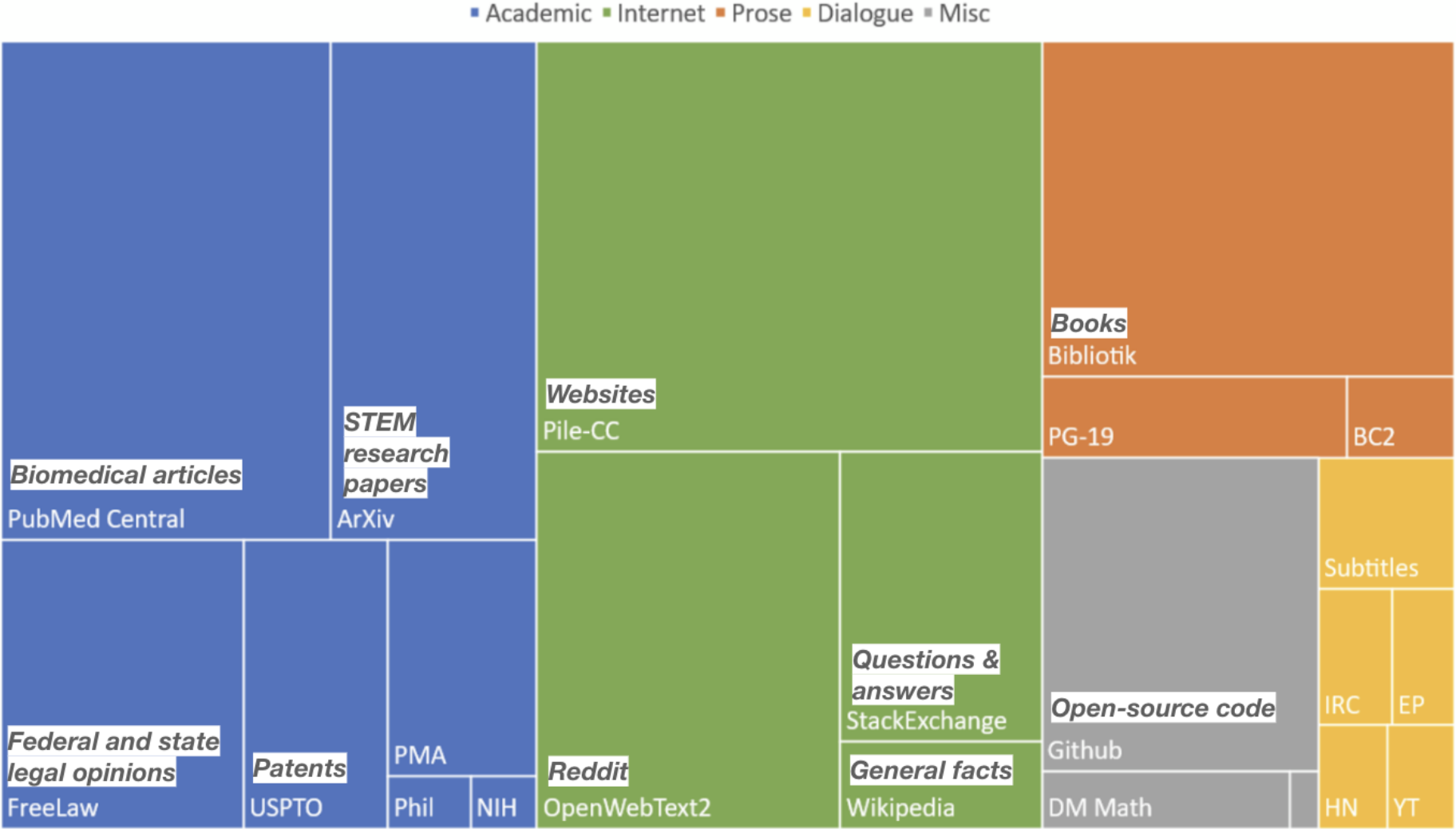
## LANGUAGE MODEL SIZES TO MAR/2023



Beeswarm/bubble plot, sizes linear to scale. Selected highlights only. \*Chinchilla scale means T:P ratio > 15:1. <https://liferchitect.ai/chinchilla/> Alan D. Thompson, March 2023. <https://liferchitect.ai/>



# Foundation-Models are trained on 100s of Gigabytes of text data to build an understanding of the world



# Foundation Models: Very simple mechanics, but very powerful! But not really intelligent....

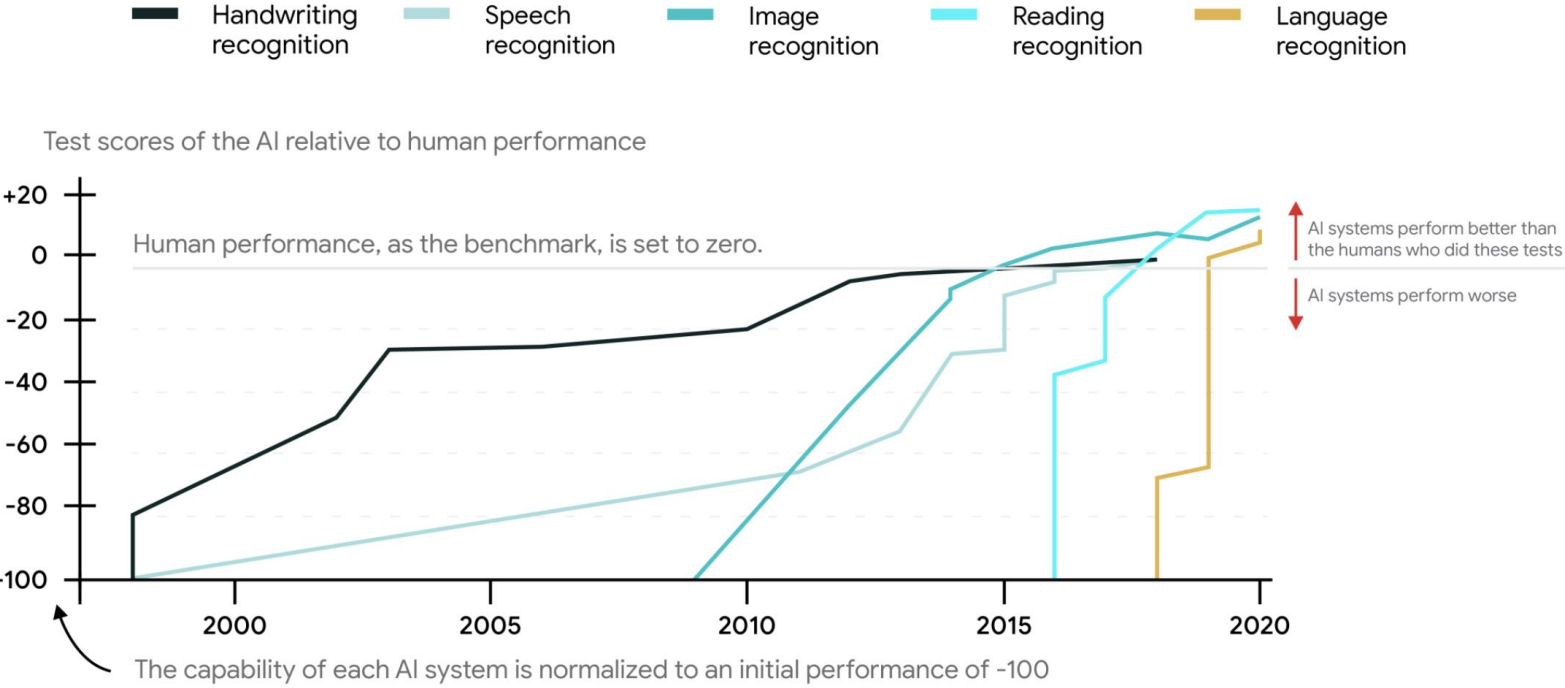
## PRETRAINING

- Training of a model with simple tasks that do not require human annotation of training data. e.g. by filling gaps or by completing sentences or texts

\_\_\_\_\_ released a China version of Model 3. -- Tesla  
Tesla \_\_\_\_\_ a China version of Model 3. -- released  
Tesla released \_\_ China version of Model 3. -- a  
Tesla released a \_\_\_\_\_ version of Model 3. -- China  
Tesla released a China \_\_\_\_\_ of Model 3. -- version  
Tesla released a China version \_\_ Model 3. -- of  
Tesla released a China version of \_\_\_\_\_. -- Model 3

training data	labels
sentence with gap 0-1	word 1
sentence with gap 1-2	word 2
sentence with gap 2-3	word 3
sentence with gap 3-4	word 4
sentence with gap 4-5	word 5
sentence with gap 5-6	word 6
sentence with gap 6-7	word 6

# The development pace of model capabilities has massively increased in the last years



Data source: Kiela et al. (2021) - Dynabench: Rethinking Benchmarking in NLP  
 OurWorldinData.org - Research and data to make progress against the world's largest platforms.

Licensed under CC-BY by the author Max Roser

## What, like it's hard?

Academic and professional exam scores (and percentiles) of Open AI's latest models

Exam	GPT-4	GPT-3.5
Uniform Bar Exam	298/400 (~90th)	213/400 (~10th)
LSAT	163/180 (~88th)	149/180 (~40th)
SAT Reading & Writing	710/800 (~93rd)	670/800 (~87th)
SAT Math	700/800 (~89th)	590/800 (~70th)
GRE Verbal	169/170 (~99th)	154/170 (~63rd)
GRE Writing	4/6 (~54th)	4/6 (~54th)
AP Biology	5/5 (85th-100th)	4/5 (62nd-85th)
AP Calculus BC	4/5 (43rd-59th)	1/5 (0th-7th)
AP Chemistry	4/5 (71st-88th)	2/5 (22nd-46th)
AP English Language and Composition	2/5 (14th-44th)	2/5 (14th-44th)
AP English Literature and Composition	2/5 (8th-22nd)	2/5 (8th-22nd)
AP Macroeconomics	5/5 (84th-100th)	2/5 (33rd-48th)
Introductory Sommelier	92%	80%
Advanced Sommelier	77%	46%
Leetcode (easy)	31/41	12/41
Leetcode (hard)	3/45	0/45

Table: J.D. Capelouto/Semafor • Source: OpenAI







# Foundation AI as continuum of applications: Best Practices

# Foundation AI as continuum of applications

No Quality Management/  
Adjustment

Output is not directly  
used as final result

Inspiration  
Innovation  
Brainstorming  
Research

Minor Quality Management/  
Adjustment

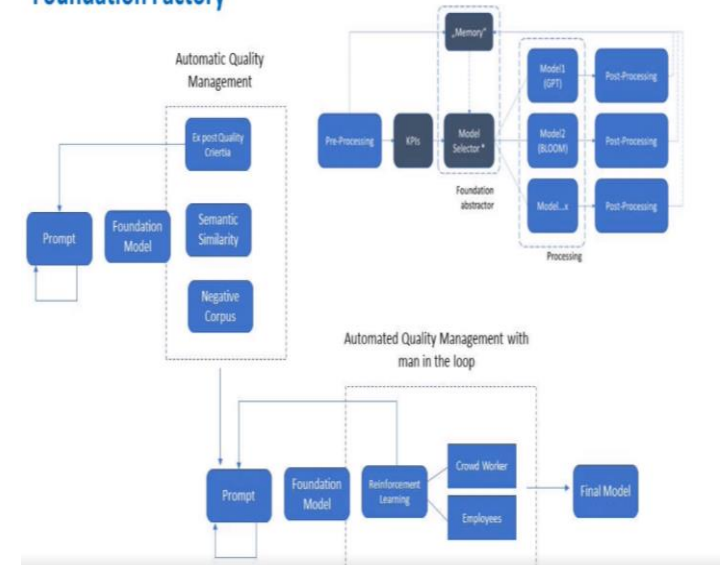
In order for the output to be  
used as the final result, no or  
only minor adjustments have to  
be made (usually ex post)

Imitation  
Summarization  
Paraphrasing  
Translation

Blog Post  
Images  
SEO-Optimization  
E-Mail responses  
...

Advanced Quality  
Management/ Adjustment  
(Fine Tuning/ Embeddings)

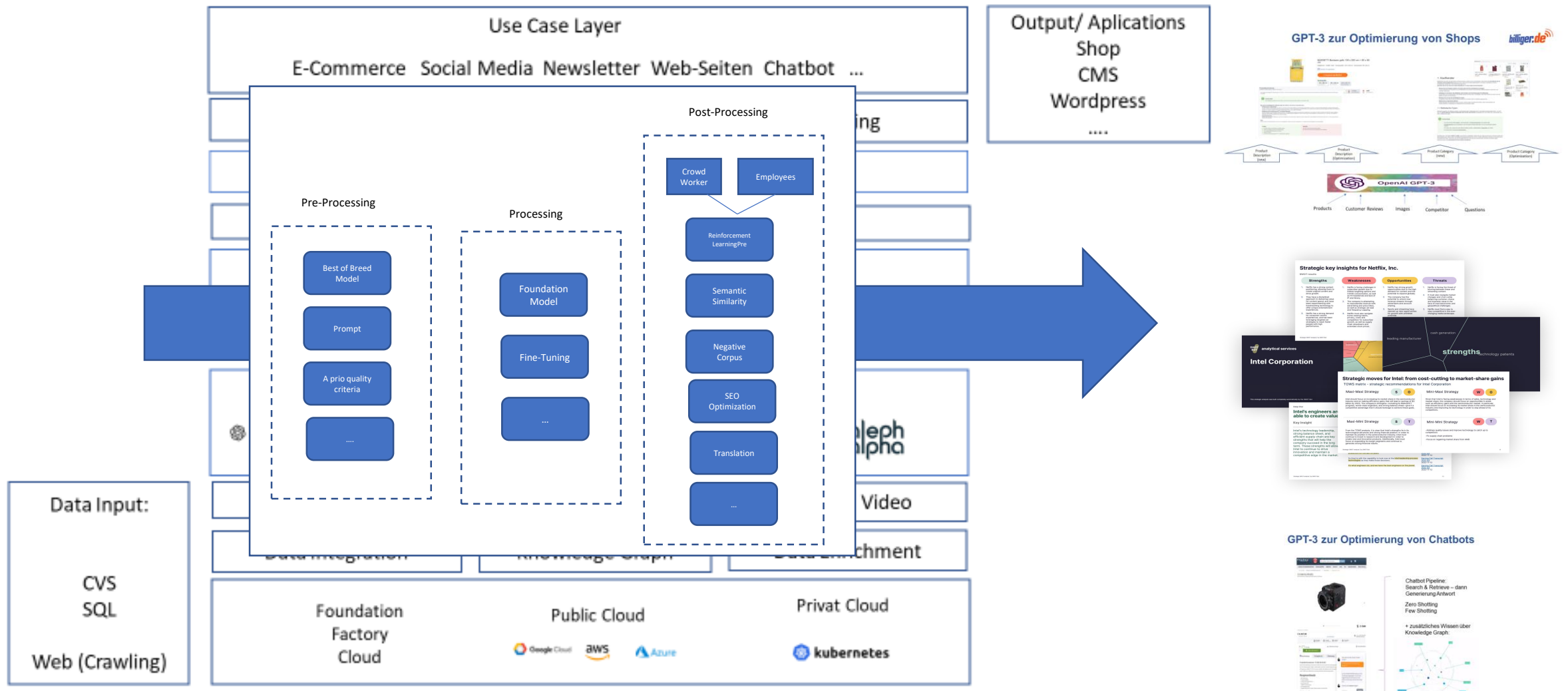
Foundation Factory



„Human in the loop“

Error Tolerance  
Halluzination Tolerance

# Foundation Factory: Building professional processes and products







# GPT-4 + Knowledge Graph: New quality of Customer communication and interaction



Kamera- & Produktionstechnik Kamerazubehör Objektive Post & IT Licht Ton Speichermedien Demo & Used

Sie sind hier: Kamera- & Produktionstechnik / Camcorder / Digital Cine Cam

### Z-CAM E2-F8 (EF)

8K Cine-Kamera, Vollformat CMOS Sensor, EF-Mount

Artikelnummer: 12286262

€ 6.357,98  
Brutto € 7.566,00

Bestpreis Garantie 100 Tage Zahlergüdel Produkt Schutz Vollkasko Schutz

Lieferzeit: 3-7 Werktage

In den Warenkorb

Beschreibung Zubehör (5) Beratung

#### Produktinformationen "Z-CAM E2-F8 (EF)"

Mit der 4K-CineKamera E2 brachte Z-CAM schon eine super Preis-Leistungs-Kam einer 4K Auflösung bei 120fps, 10-Bit-Farben und einem auf alle Situationen an nicht genug und legt nicht nur einer, sondern drei Kameras nach. Dazu geht einem Vollformat CMOS Sensor atemberaubende Aufnahmen garantiert.

#### Hauptmerkmale

- 8K Auflösung
- 10-bit Farbtiefe
- Vollformat CMOS Sensor
- Dual Native ISO
- 28AW Unterstützung
- EF Objektivmount
- Gigabit Ethernet für Daten, Steuerung & Live-Streaming

#### Lieferumfang

Was willst du über dieses Produkt wissen?

Ist die Kamera für Livestreaming geeignet?

Ja, die Z-CAM E2-F8 (EF) Kamera ist für Livestreaming geeignet, da sie einen Gigabit Ethernet Anschluss für die Daten, Steuerung und Live-Streaming hat.

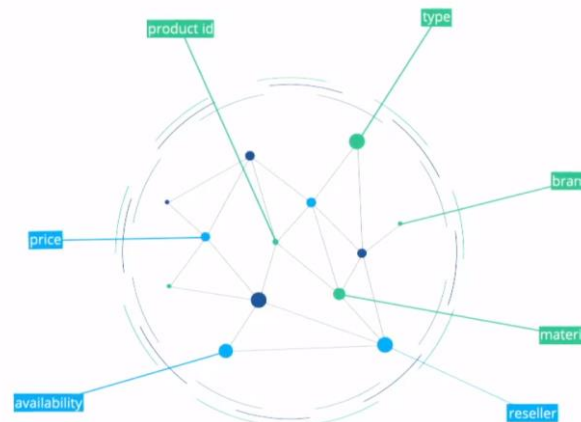
Hast du noch weitere Fragen?

Tippe hier... Senden

Chatbot Pipeline:  
Search & Retrieve – After that  
generating answers

Zero Shotting  
Few Shotting

+ additional Knowledge via  
Knowledge Graph:



Fine-Tuning

+

Domain Knowledge

# Category-Manager as Co-Pilot – interactive evaluation of Store-Performance and Action Plan

Category Manager



Dynamically generated Analyses and Insights



Indication of anomalies



Customized action



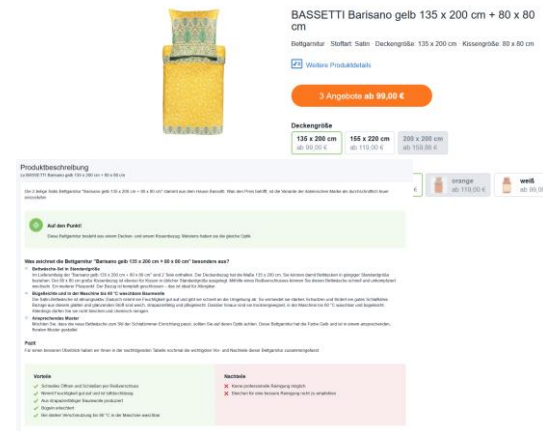


# GPT-4 for optimization of E-Commerce



## Tasks:

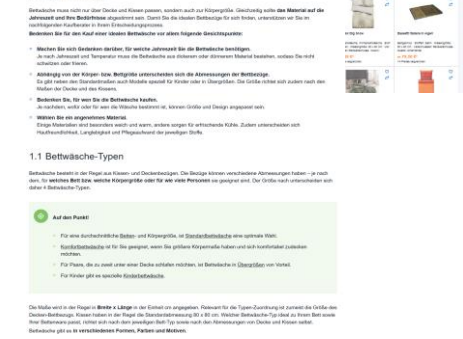
- Generation of category texts
- Product text generation



## Business Benefits:

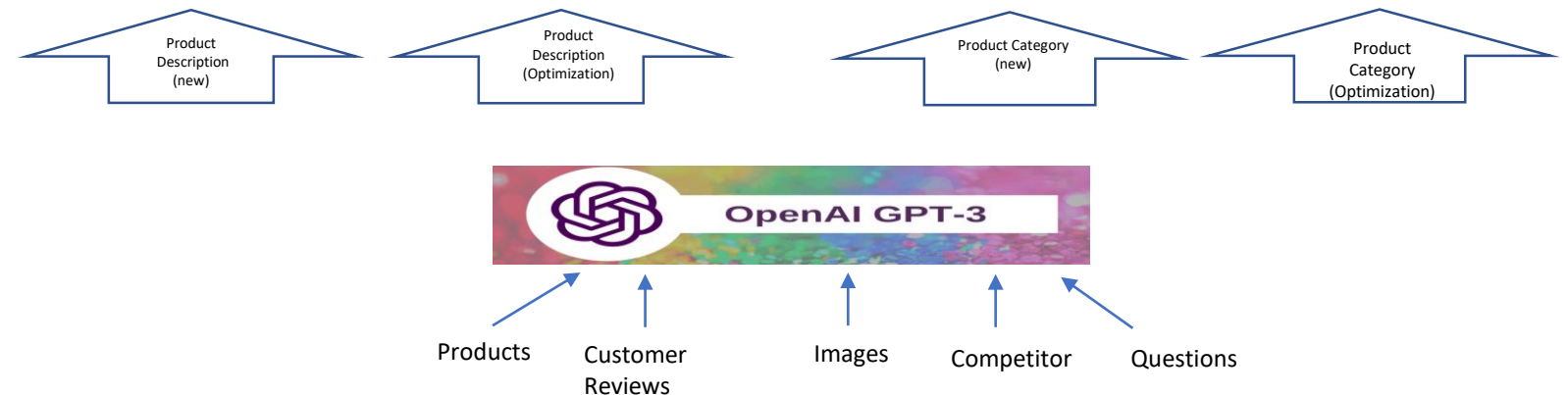
- Efficiency / cost benefits
- 83% savings
- Increase Conversion Rate by 8%

### 1. Kaufberater



## Quality:

- Quality assured
- Domain optimized
- SEO optimized
- Customer-centric
- Personalization



# Generative AI for product visuals





# New Business Models: Generative AI will create new business opportunities and models

## Generative AI as lever for IoT-Solutions

### 1 Code generation for IoT

Large language models can be used to create, complete or combine IoT code, sourced from code snippets or natural language descriptions.

Example



GitHub Copilot

```
def max_val_list(x):
    max_val = 0
    for i in range(len(x)):
        if x[i] > max_val:
            max_val = x[i]
    return max_val
```

GitHub Copilot

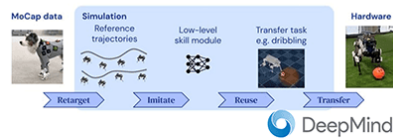
### 2 Robot control

Generative AI can be used to generate controls and commands of IoT connected robots by capturing motion data from animals or humans.

Example



Deepmind Robot Control Framework



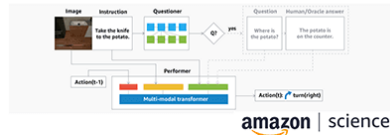
### 3 Social IoT devices

Generative AI can be used to change the interaction with devices. Devices, for example, could provide the user with answers to complex question.

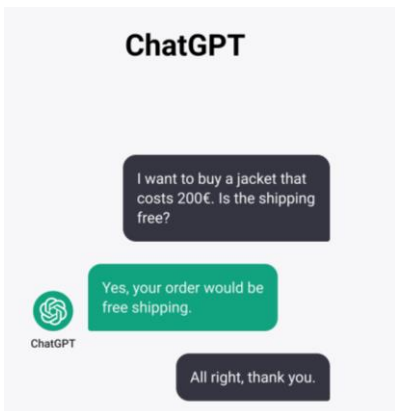
Example



Amazon DialFRED



## Generative AI as lever for teh Metaverse



Example Smart Fridge, delivering proactively services

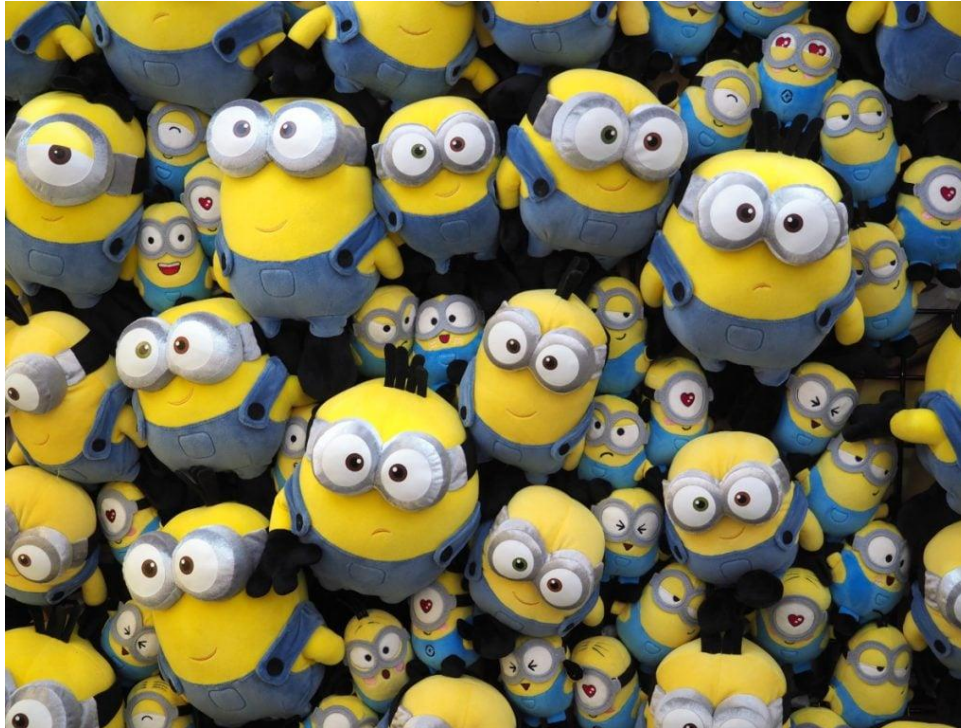




**Paradigm Shift: LLM as the new  
Business Operating System**



# Auto-GPT tries to increase the autonomy of tasks and processes....



Linas Beliūnas   
@linas.beliunas

**AutoGPT** is the biggest technology breakthrough ever and **the next BIG thing in AI**.

They're basically **AI agents** that run **autonomously** and complete tasks for you. Today they can already:

- **Build an app/website** from scratch
- **Order your coffee** at Starbucks
- Run your **social media accounts**
- Find and **negotiate a lease**
- Perform **financial analysis**
- Do **market research**

And this is just the tip of the iceberg.

AutoGPT will change every industry and every job forever.

...but a bit overhyped and underdelivered. Execution should not be part of LLM. But the underlying idea of LLM as agents is promising.

# Foundation System: Primary & Secondary System

## Roles & Rights Concept

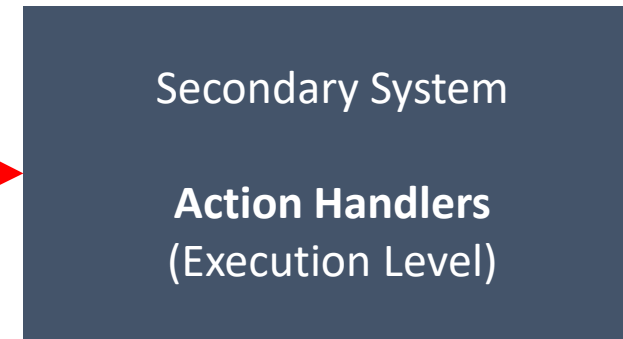
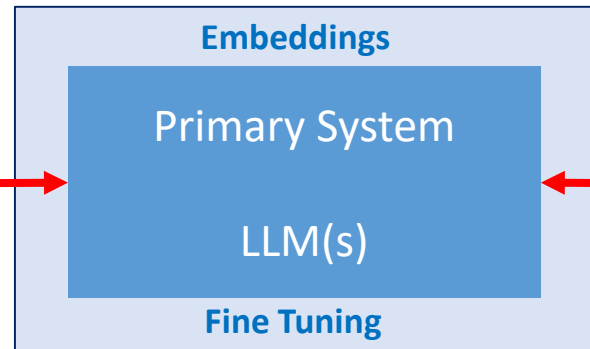
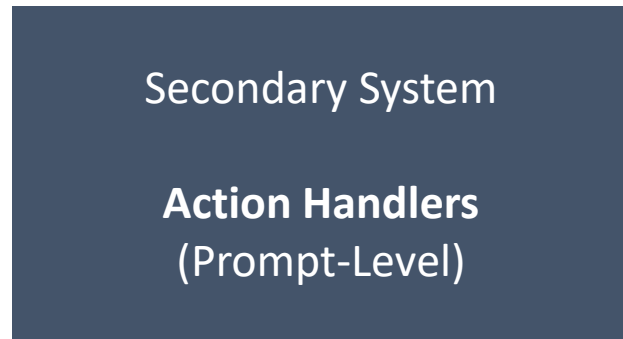
Layer III: Custom LLM Policies

Layer II: Domain-Specific LLM Policies

Layer I: Corporate LLM Policies

Secondary System  
Policy Engine

Safety Layer



Data Layer

(Cloud) Infrastructure Layer



# Large European Retailer – Customer Service Center

you act as a member of the customer services department of a large European retailer. you are responsible for handling customer complaints.

here are the general company policies:

- if spare parts are required with less than 20 EUR total value, you can send them straight away
- for spare parts larger than 20 EUR write an approval email to approval@email.com and wait for approval email back

here are information about information system that you can access:

Below is a list of all the relevant Postgres databases with their description and public schema. \n flow\_id - 72 \tflow\_name - query ordering system \ttable\_name - order\_history \tdescription - This flow is used to query the order history for a customer. \ttable\_schema - orderID:serial,customerID:serial,product:VARCHAR(512),purchase\_date:TIMESTAMP,price:NUMERIC(10,2) \tinput\_parameters\_keys - customerID \t\nflow\_id - 72 \tflow\_name - spare parts \tdescription - This flow is used to query all available spare parts for a product. \ttable\_schema - spare\_part\_id:serial,product\_name:VARCHAR(512),product\_part:VARCHAR(512),product\_spare\_part:VARCHAR(512),price: NUMERIC(10,2) \tinput\_parameters\_keys - product\_name \t\nflow\_id - 72 \tflow\_name - describe logistic system \ttable\_name - logistic\_system \tdescription - This flow is used to populate the logistic system table with shipment of the spare parts. \ttable\_schema - customerID:serial,claim\_id:serial,part\_id:serial,customer\_name:VARCHAR(512),customer\_address:VARCHAR(512),customer\_city:VARCHAR(512),shipping\_date:TIMESTAMP \tinput\_parameters\_keys - customerID \t\nflow\_id - 72 \tflow\_name - Query customer master \ttable\_name - customer\_data \tdescription - This flow is used to query customer master data for a customer. \ttable\_schema - customerID:serial,customer\_firstname:VARCHAR(512),customer\_lastname:VARCHAR(512),customer\_address:VARCHAR(512),customer\_city:VARCHAR(512) \tinput\_parameters\_keys - customerID

Customer Message:

Dear Product Team,

The foot of one of the armchairs in my ""Toscana"" set is broken. I would like to have it repaired.

Many greetings

CustomerID: 4711

use your process know how for handling this specific customer complain.

write a step-by-step explanation.

if you need information from any system, write the SQL query to extract the relevant data.

make sure all company policies are adhered to.

# Large European Retailer – Customer Service Center

Step 1: Identify the product

First, we need to identify the product that the customer is referring to. In this case, it's the "Toscana" armchair set.

Step 2: Query spare parts system

Next, we need to query the spare parts system to find the spare part required for repairing the broken foot of the armchair.

SQL Query:

sql

Copy code

```
SELECT spare_part_id, product_part, product_spare_part, price
FROM spare_parts
WHERE product_name = 'Toscana';
```

Step 3: Determine spare part cost

After obtaining the spare part information from the spare parts system, determine the cost of the spare part needed for the repair.

Step 4: Check company policy on spare part cost

If the cost of the spare part is less than 20 EUR, we can send it directly to the customer. If the cost is greater than 20 EUR, we need to send an approval request email to the customer.

Step 5: Send approval request email if required

If the cost of the spare part is greater than 20 EUR, send an approval request email to [approval@email.com](mailto:approval@email.com) with the necessary details, such as customer name, product name, and spare part details.

Action Handler  
(secondary system)

External  
System

Input for next prompt

here is the specific result from the SQL query from step 2:

SparePartID	Product Name	Part	Spare Part	Price
12	Roma Tisch	Tischplatte		150
23	Roma Tisch	Tischbein		35
34	Toscana Sessel	Sesselbezug		50
45	Toscana Sessel	Fuß		15
56	Toscana Sofa	Bezug		130

use this as an input to run step 3

# Large European Retailer – Customer Service Center

here is the specific result from the SQL query from step 2:

SparePartID	Product Name	Part	Spare Part	Price
12	Roma	Tisch	Tischplatte	150
23	Roma	Tisch	Tischbein	35
34	Toscana	Sessel	Sesselbezug	50
45	Toscana	Sessel	Fuß	15
56	Toscana	Sofa	Bezug	130

use this as an input to run step 3

Step 3: Determine spare part cost

From the provided SQL query results, we can see that the spare part needed for repairing the broken foot of the "Toscana" armchair is:

SparePartID: 45  
Product Name: Toscana  
Part: Sessel (Armchair)  
Spare Part: Fuß (Foot)  
Price: 15

The cost of the spare part (foot) required for the repair is 15 EUR.

run step 4

Step 4: Check company policy on spare part cost

Since the cost of the spare part (foot) is 15 EUR, which is less than the 20 EUR threshold, we can send it directly to the customer without requiring any additional approval.

Action Handler  
(secondary system)

External  
System(s)





# Large European Retailer – Customer Service Center

here is the  
SparePart  
12 Roma  
23 Roma  
34 Toscar  
45 Toscar  
56 Toscar

use this as

Step 3: De  
From the p  
broken foc

SparePart  
Product N  
Part: Sess  
Spare Part  
Price: 15

The cost o

run step 4

Step 4: Ch  
Since the c  
can send it directly to the customer without requiring any additional approval.

Subject: Your Toscana Armchair Repair - Claim #34593-234

Dear Peter Gentsch,

Thank you for contacting us regarding your Toscana armchair set. We apologize for any inconvenience caused by the broken foot on one of the armchairs.

We are pleased to inform you that we have processed your claim under the reference number 34593-234. As the cost of the required spare part (foot) is 15 EUR, we will be sending it to you without any additional approval needed.

Your spare part has been shipped with the following details:

Shipment ID: 464-824-789  
Spare Part: Toscana Armchair Foot  
Price: 15 EUR

Please allow a few business days for delivery. Once you receive the spare part, you can proceed with the repair. If you require any further assistance, please do not hesitate to contact us.


We appreciate your understanding and patience, and we hope that this resolution meets your satisfaction. Thank you for choosing our products, and we look forward to serving you in the future.

Best regards,

Customer Service Department  
Non-disclosed Company

Action Handler  
(secondary system)

External  
System(s)



# **Generative AI: ReThinking Business**

# Every industry is impacted by GenAI

## Verticals

Financial Services	Retail	Media/ Entertainment
Manufacturing	Utilities	Utilities


## Value Chain/ Functions

Management		
Marketing	Sales	Procurement
E-Commerce	After Sales/ Service	Legal
HR/ Training	Category Management	IT
Knowledge Management		
Admin/ Communication		
Market Research/ Insights		



# Every industry is impacted by GenAI: Example Retail

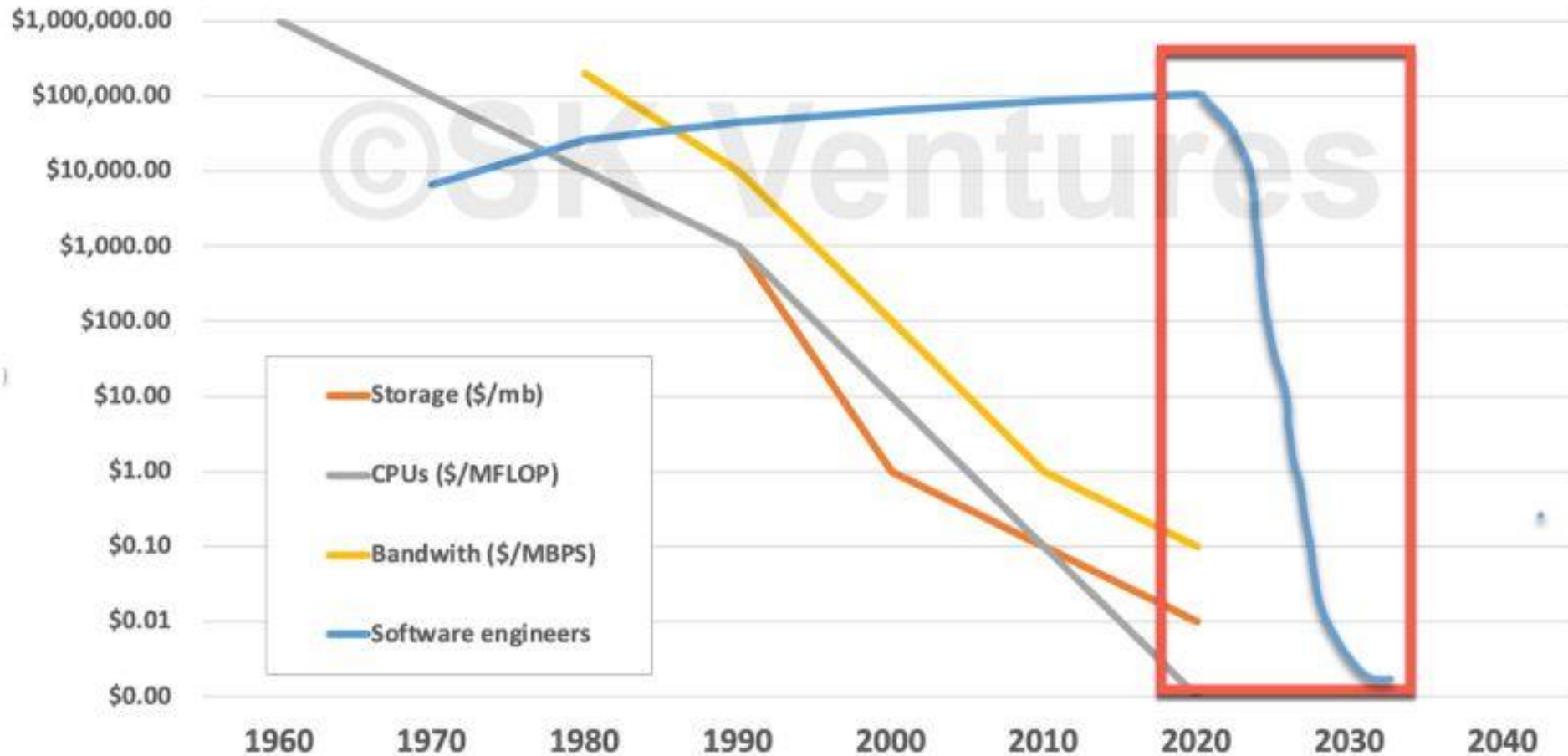
Strategic advice (Co-Pilot), SWOT-Bot, New Business Opportunities/ Models, ...		
Content Production, Hyper Personalization, Engagement (Bots), Automated Campaigning...	Sales Bot, Lead Generation and Classification, Automated Campaigning...	RFP Creation, Supplier negotiations, Tailored Vendor Communication, ....
Product description, Recommendation, Virtual Shopping Assistant, ...	24/7/365 Service (Bots), Personalization, Q&A, Automated Campaigning...	Contract Approval/ Generation, Drafting documents, ....
Personalized Training and Onboarding, Job Descriptions, ..	Category Recommendation, Pricing Optimization, ...	Code Generation, Code Translation, Code Approval,...
Knowledge Co-Pilot, Generation/ Enrichment of Knowledge Graphs, ...		
E-Mail Generation/ Sorting, Meeting summarization, Action Follow-up/ Scheduling		
Automated Market Research, Customer Insights, Store/ Shop Optimization		



**Change has never been this fast  
- it will never be this slow again**

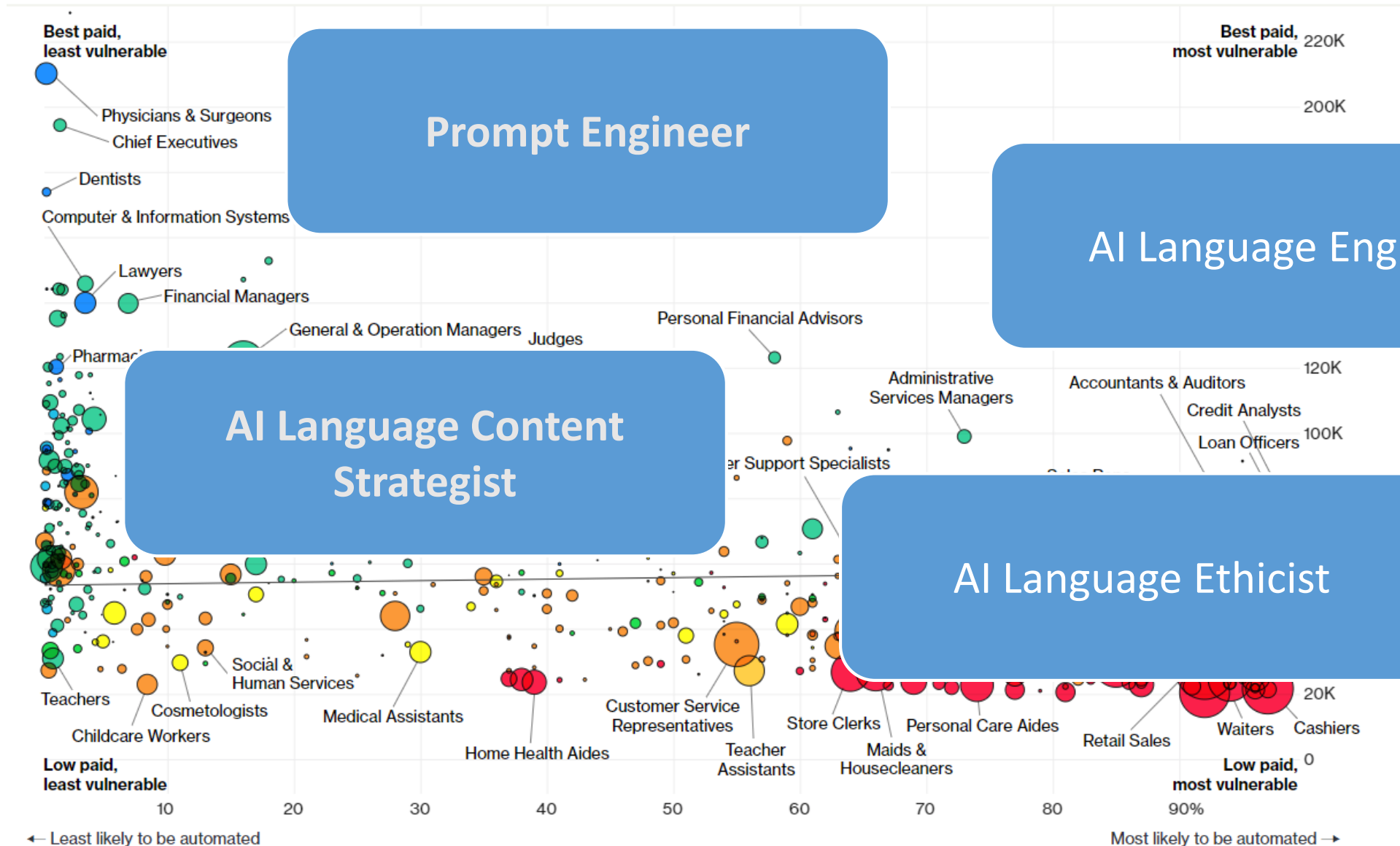
# Disruption of jobs: Example Software Engineer?

## The Next Collapsing Tech Cost: Software Itself

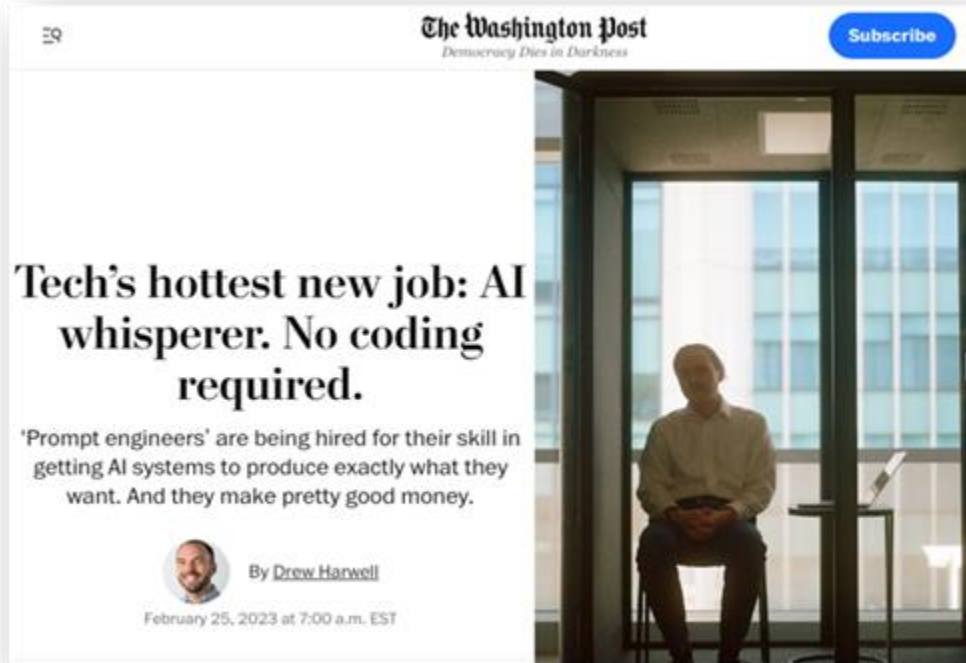




# Jobs are disrupted and created at the time



# PROMPT ENGINEER ALS "HOTTEST NEW JOB"



## SPIEGEL Job & Karriere

ChatGPT erschafft neuen Beruf

### 5+ »Ein Jahresgehalt von 300.000 Dollar ist natürlich verlockend«

Für diesen Job muss man keine Programmiersprache können: Prompt Writers bringen künstliche Intelligenzen dazu, optimale Ergebnisse zu liefern. Ein Literaturwissenschaftler erzählt, wie er sich so ein zweites Standbein aufgebaut hat.

Ein Interview von [Verena Töpfer](#)

24.03.2023, 17.54 Uhr

**Jeremiah O...** [+ Follow](#)



Career mission:  
I connect com...

**PREDICTION:** The next billion-dollar startup will have only 3 employees.







# Key Takeaways

# Key Takeaways

- To unleash the real value of LLMs they have to be incorporated into reliable and quality assured business infrastructure
- The biggest challenge of LLM is to ensure that the generated output is compliant with policy and quality standards of companies and industries
- The autonomy level of AI driven tasks will increase significantly
- It is not about the LLM itself: US - China - Europe - Germany. It will become more and more commodity. It is mainly about how to adapt and interact with LLM
- Therefore a paradigm shift is needed: To control quality and autonomous processes the LLM infrastructure of the future will consist of a primary system (LLM) and secondary system (policies)
- A Business Canvas for Generative AI can help to systematically identify use and business cases
- Job profiles, organizations, business models and industries will change dramatically
- Companies need to radically rethink the way they are doing business
- Development of AI has never been this fast – it will never be this slow again



**Teufel**

**G+J**

**Miele**

**arvato**  
BERTELSMANN

**solute**<sup>o</sup>

AI Foundation  
Circle 2023



**R T L**

grand  
centrix

**ProSiebenSat.1**  
Media SE

**TELTEC**  
moving-picture.experts

**DB**

<https://generative-ai.group>





Thank you  
very much!

peter.gentsch@intelligence-group.com

COMPANIES												
INVESTMENTS												
CUSTOMERS												
ACADEMICS					BOOKS							
AWARDS					TALKS							

PETERGENTSCH.COM