AI Opportunities and Challenges in Public Administration and Food Safety

Luís Paulo Reis

lpreis@fe.up.pt

Associate Professor at FEUP - Faculty of Engineering of the University of Porto
Director of LIACC – Artificial Intelligence and Computer Science Lab., UPorto
Member of Coord. Commission of LASI – Intelligent Systems Associate Laboratory
President of the GA of APPIA – Portuguese Association for Artificial Intelligence

Agenda

- Artificial Intelligence (AI) and Machine Learning (ML)
- Generative AI, Chat GPT and AI Tools
- AI Revolution: Adapt or Resist
- AI Challenges and Oportunities
- AI in Public Administration and Food Safety
- Robotics, Deep RL and the Future
- Conclusions

AI and the Discovery of the New World

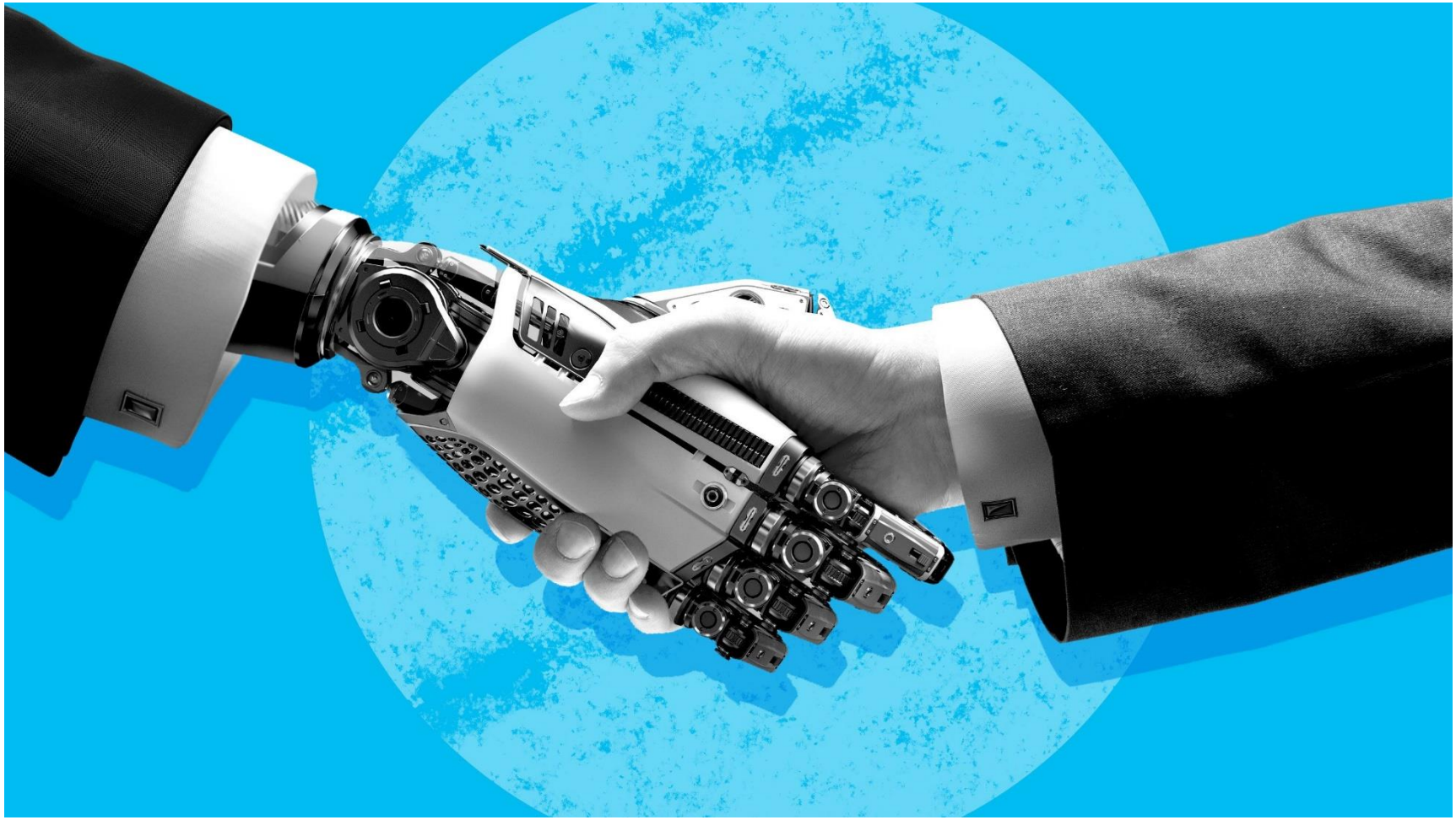
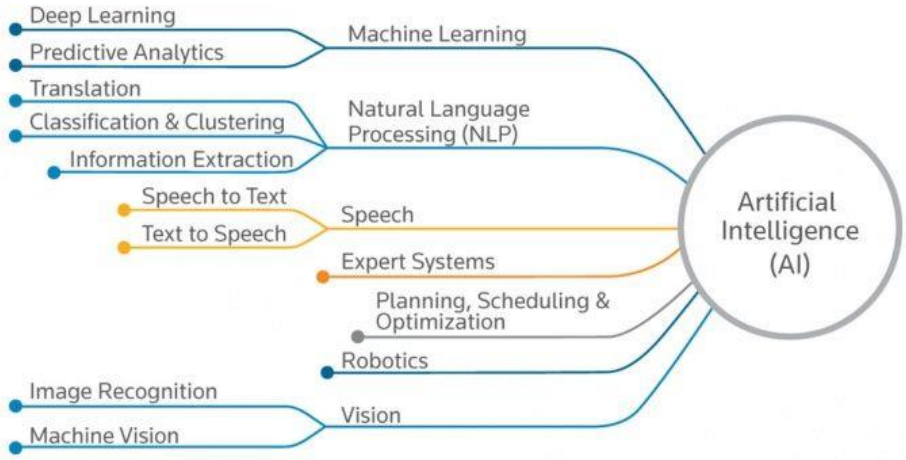


Image: Shutterstock / CNN - ttps://edition.cnn.com/2019/02/17/investing/artificial-intelligence-investors-machine-learning/

Artificial Intelligence (AI)

- **Intelligence**
  - "Capacity to **solve new problems** through the use of knowledge"

- **Artificial Intelligence**
  - "Science concerned with building **intelligent machines**, that is, machines that perform tasks that when performed by humans require intelligence"



Weak and Strong AI

- **Weak Artificial Intelligence**

  weak AI, also known as narrow AI is artificial intelligence that is focused on one **single narrow task**

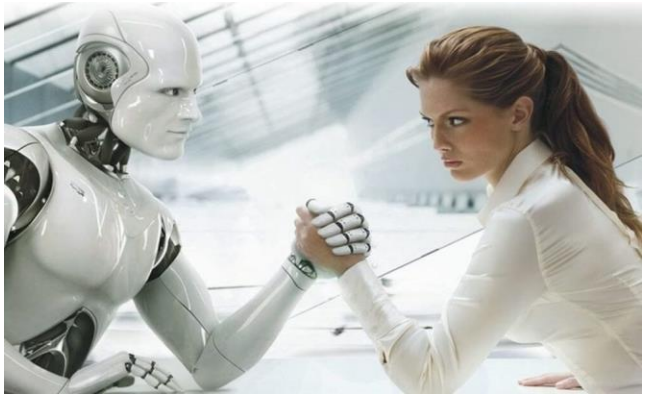


Image: https://livingsmartheart.com/what-is-strong-ai/

- **Strong Artificial Intelligence**

  Strong AI or Artificial General Intelligence (AGI) is the intelligence of a machine that could successfully perform **any intellectual task** that a human being can! Science fiction ?

AI Timeline

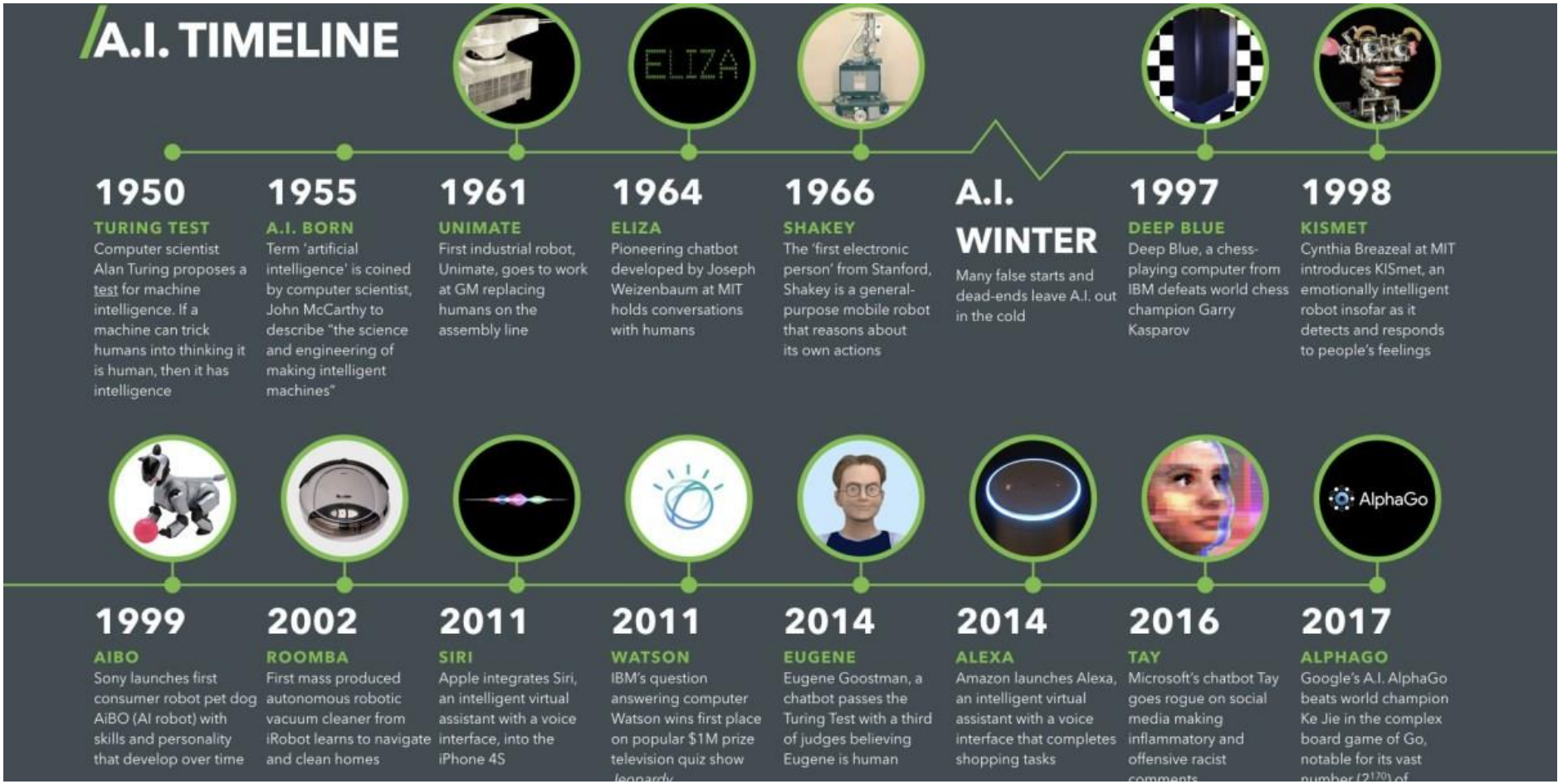
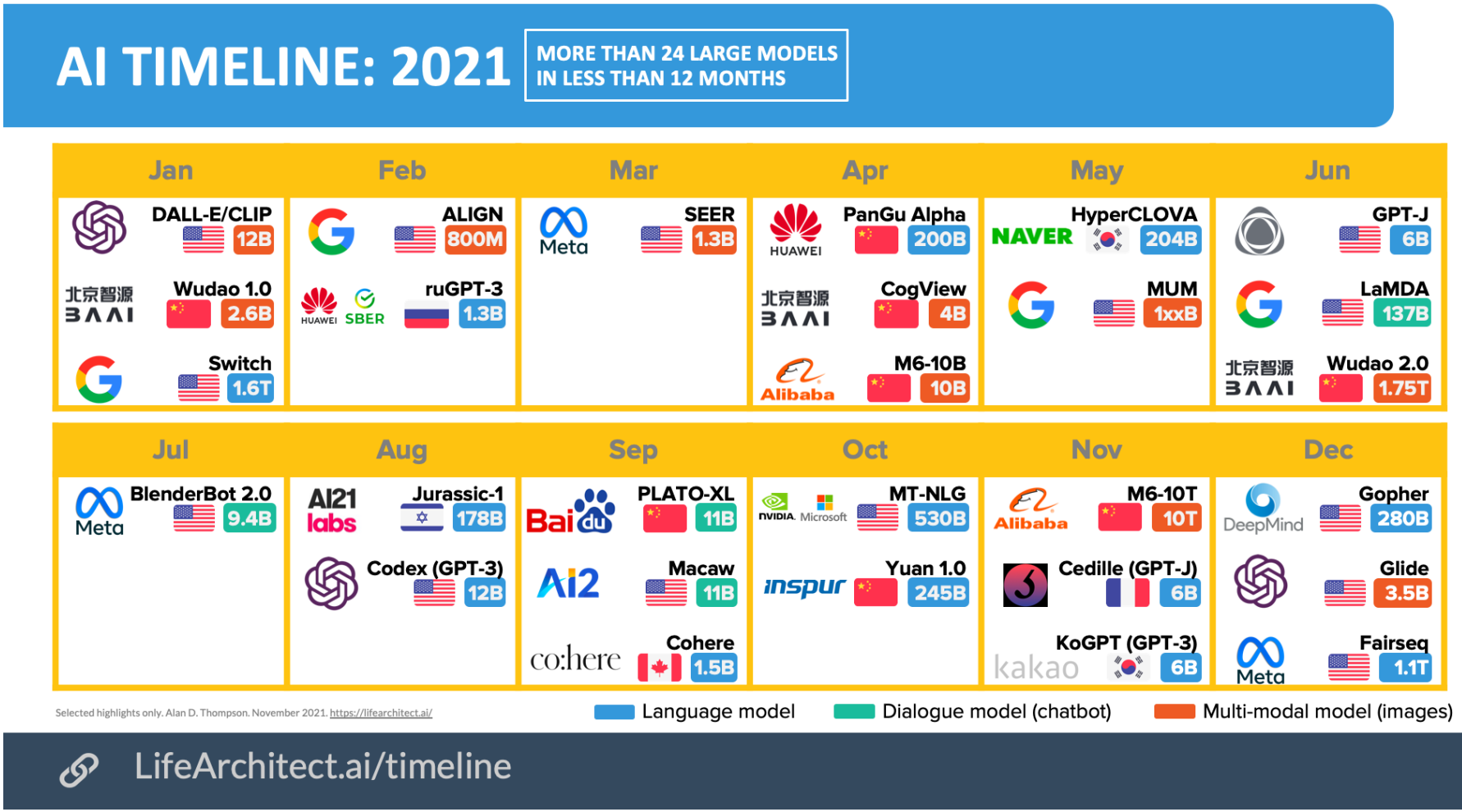


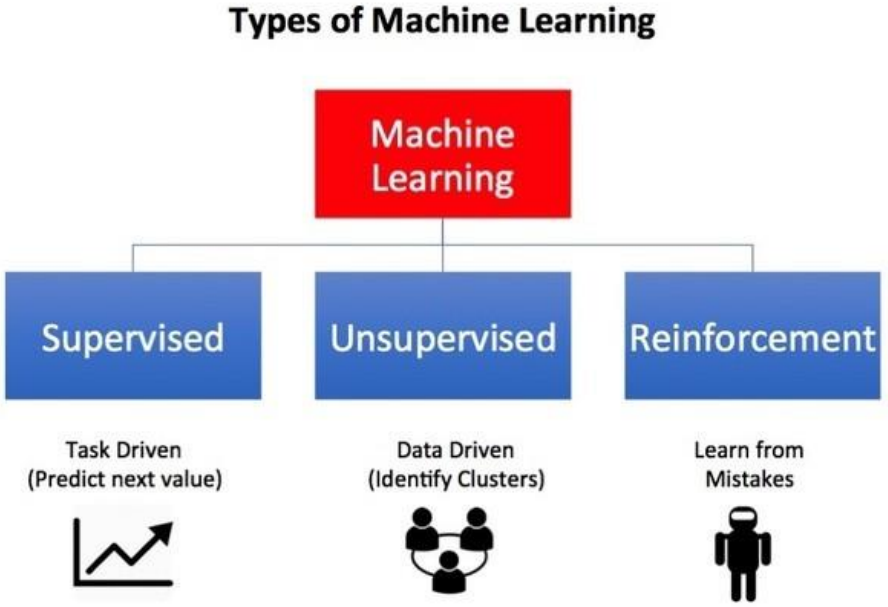
Image: [Paul Marsden, 2017] - https://digitalwellbeing.org/artificial-intelligence-timeline-infographic-from-eliza-to-tay-and-beyond/

AI Timeline


AI TIMELINE: 2021
MORE THAN 24 LARGE MODELS IN LESS THAN 12 MONTHS
Jan
DALL-E/CLIP 12B
Wudao 1.0 2.6B
Switch 1.6T
Feb
ALIGN 800M
ruGPT-3 1.3B
Mar
SEER 1.3B
Apr
PanGu Alpha 200B
CogView 4B
M6-10B 10B
May
HyperCLOVA 204B
MUM 1xxB
Jun
GPT-J 6B
LaMDA 137B
Wudao 2.0 1.75T
Jul
BlenderBot 2.0 9.4B
Aug
Jurassic-1 178B
Codex (GPT-3) 12B
Sep
PLATO-XL 11B
Macaw 11B
Cohere 1.5B
Oct
MT-NLG 530B
Yuan 1.0 245B
Nov
M6-10T 10T
Cedille (GPT-J) 6B
KoGPT (GPT-3) 6B
Dec
Gopher 280B
Glide 3.5B
Fairseq 1.1T
Selected highlights only. Alan D. Thompson. November 2021. https://lifearchitect.ai/
Language model
Dialogue model (chatbot)
Multi-modal model (images)
LifeArchitect.ai/timeline

Machine Learning

Machine learning is a field of artificial intelligence that gives **computer systems** the **ability to "learn"** (e.g., progressively **improve performance** on a specific task) **from data/results of their actions**, without being explicitly programmed



Supervised Learning

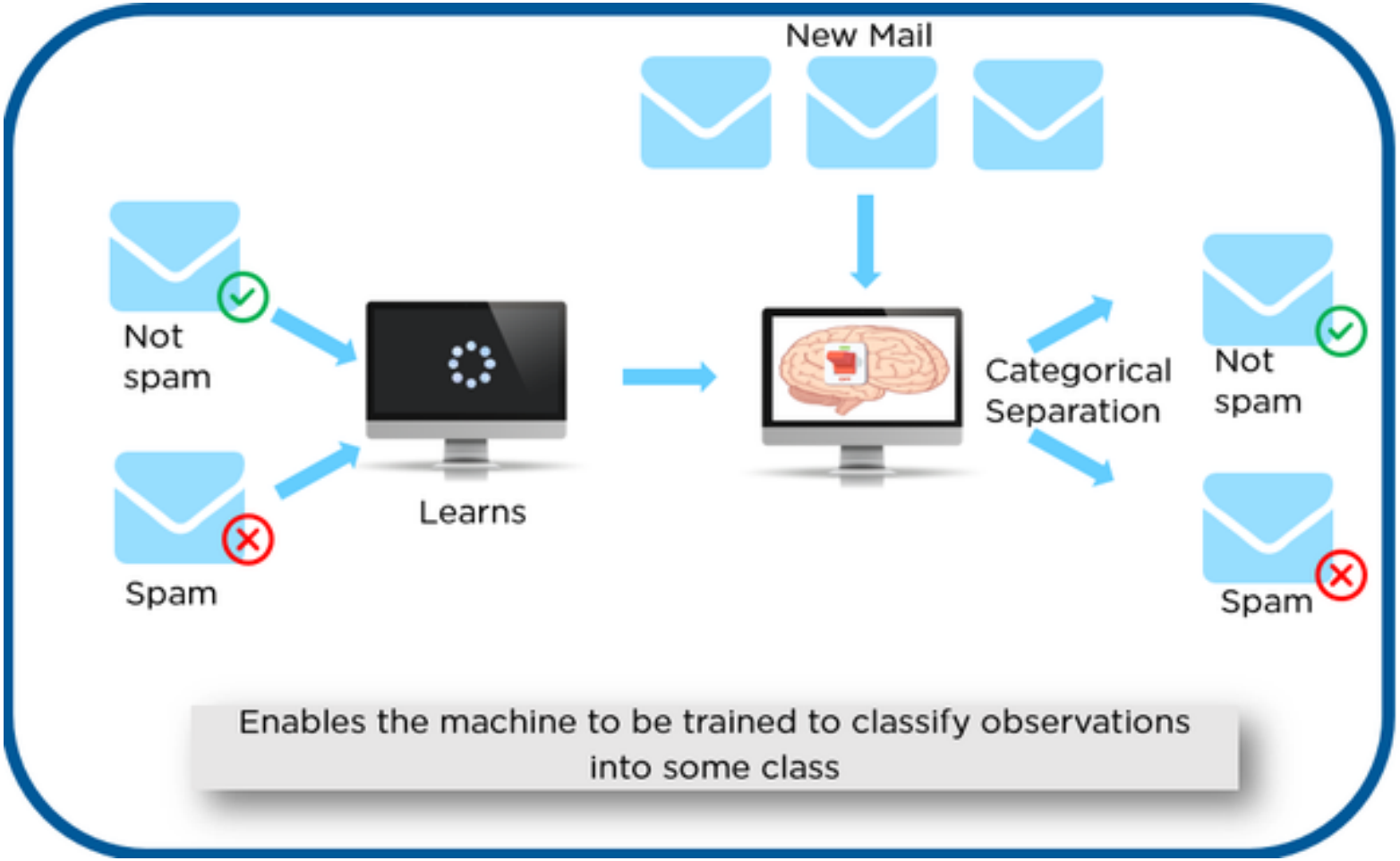
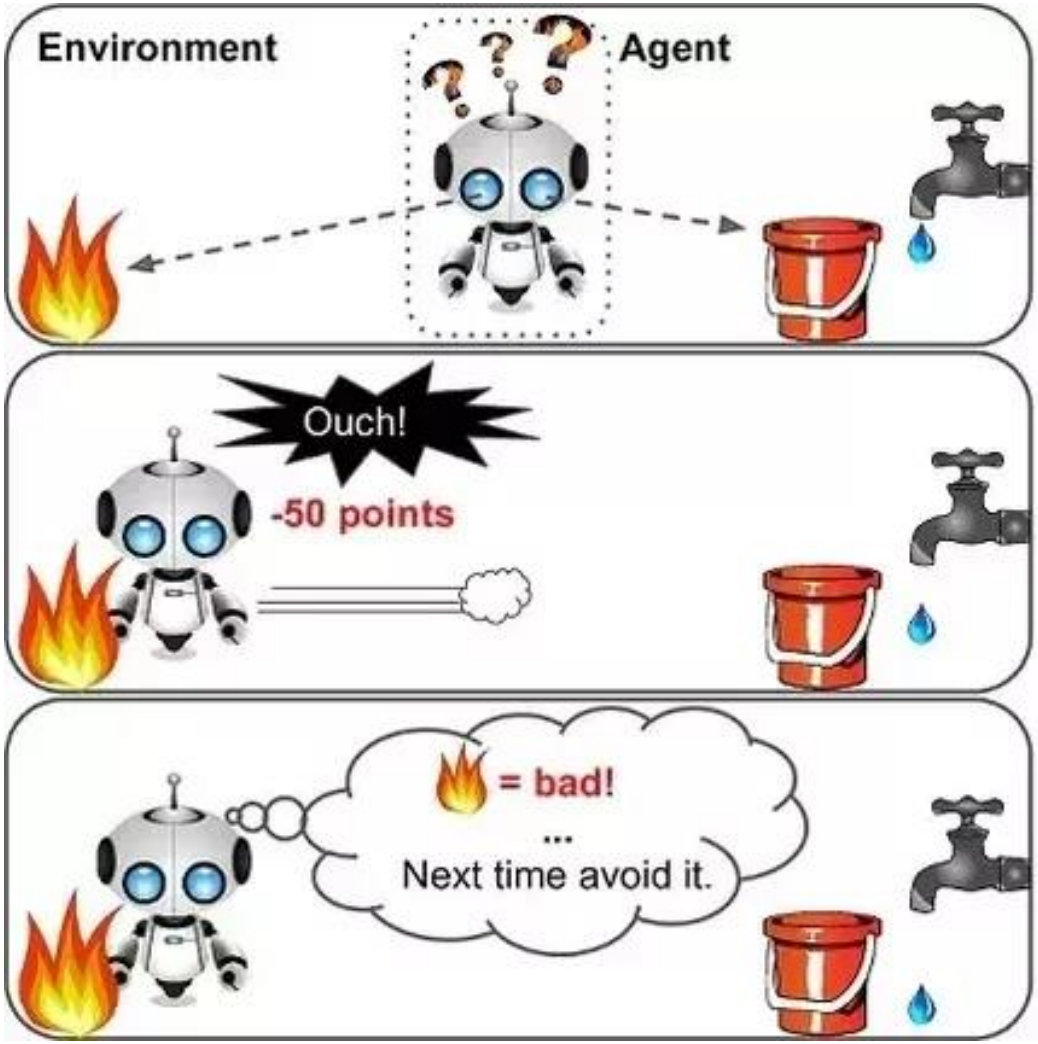


Image: https://towardsdatascience.com/what-are-the-types-of-machine-learning-e2b9e5d1756f

Reinforcement Learning



1. Observe
2. Select action using policy
3. Action!
4. Get reward or penalty
5. Update policy (learning step)
6. Iterate until an optimal policy is found

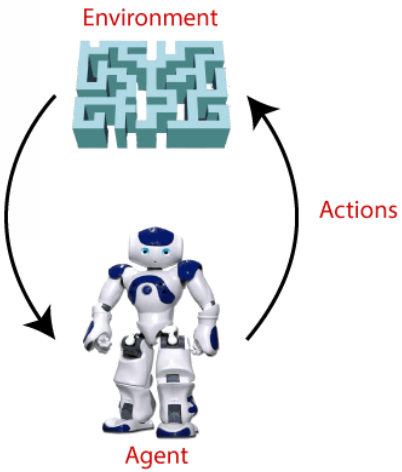
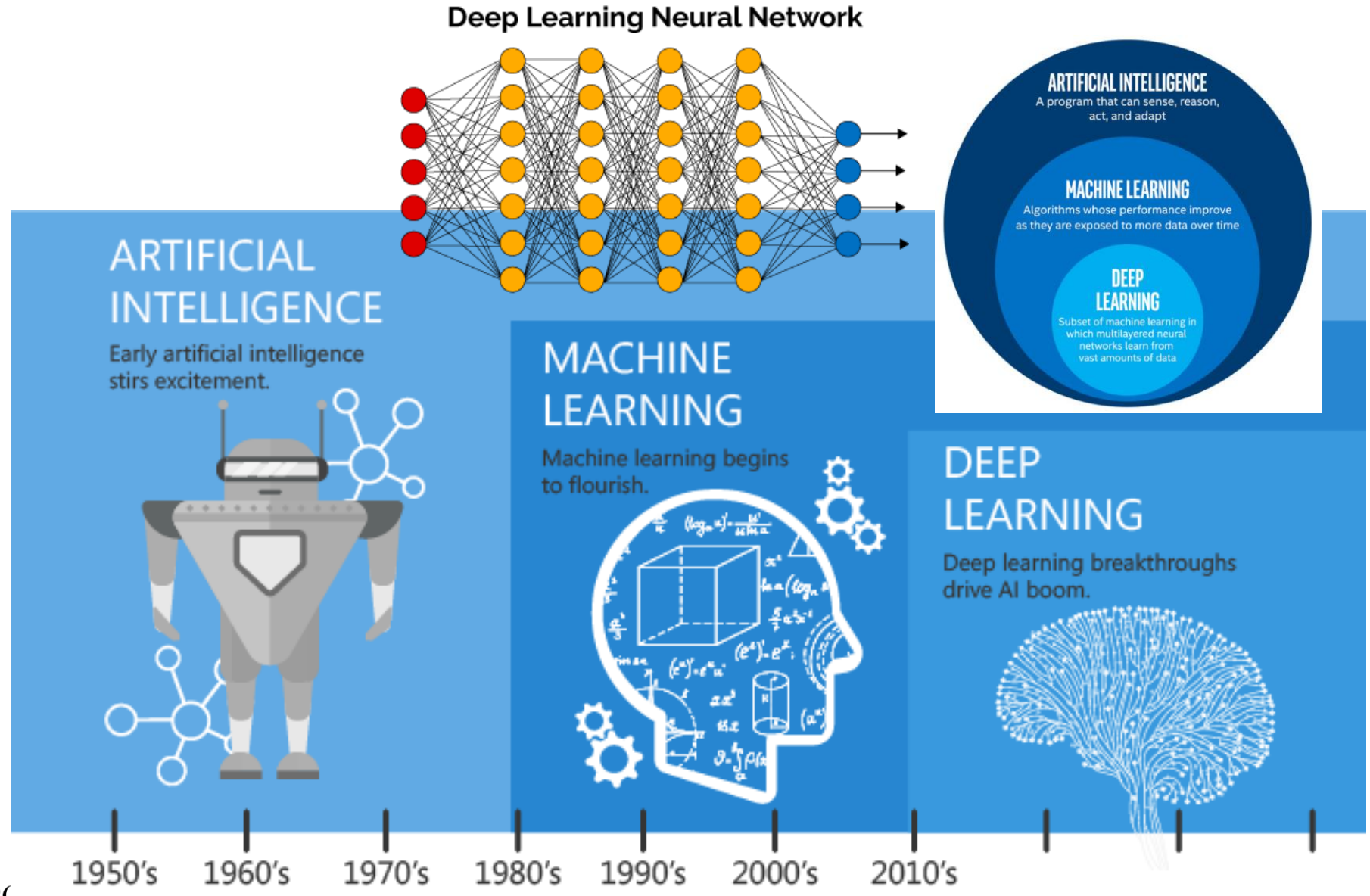


Image: https://becominghuman.ai/an-introduction-to-machine-learning-33a1b5d3a560

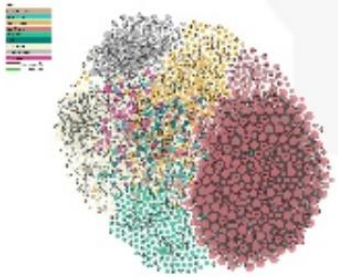
Machine Learning - History



Artificial Intelligence Today

Hardware

Data



Software

ML frameworks

PyTorch

TensorFlow

scikit learn

torch

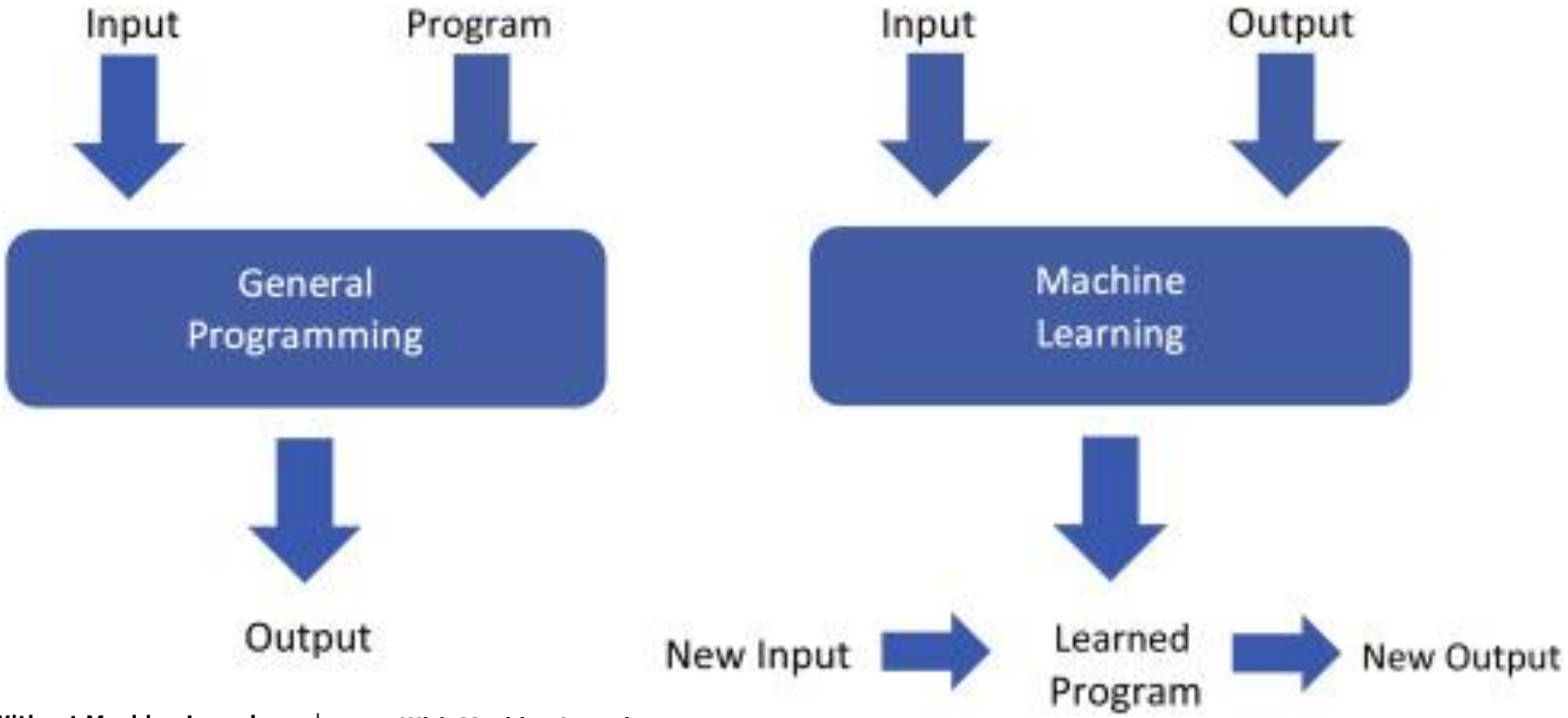
theano

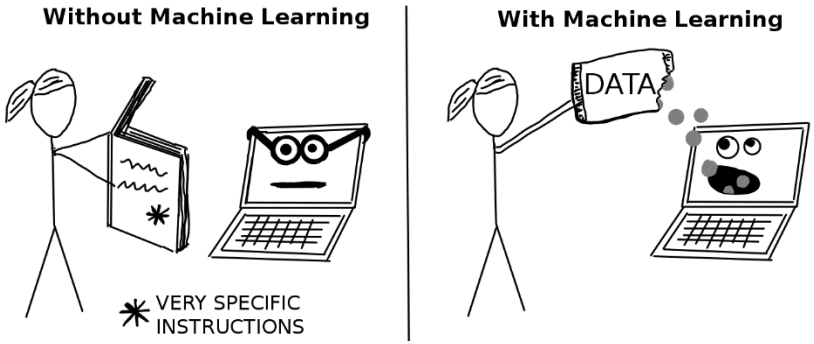
Keras

APACHE Spark ML

HUGGING FACE

Programming vs Machine Learning





The End of Programming!

Nvidia CEO predicts the death of coding — Jensen Huang says AI will do the work, so kids don't need to learn

News By Benedict Collins published February 26, 2024

Jensen Huang believes coding languages are a thing of the past

(Image credit: Nvidia)

Nvidia CEO Jensen Huang has once again announced the death of coding, but this time in front of a potentially far more influential audience.

StarCoder 2 is a code-generating AI that runs on most GPUs

Kyle Wiggers @kyle_l_wiggers / 2:00 PM UTC • February 28, 2024 Comment

Image Credits: Tippapatt / Getty Images

Developers are adopting AI-powered code generators — services like GitHub Copilot and Amazon CodeWhisperer, along with open access models such as Meta's Code Llama — at an astonishing rate. But the tools are far from ideal. Many aren't free. Others are, but only under licenses that preclude them from being used in common commercial contexts.

Perceiving the demand for alternatives, AI startup Hugging Face several years ago teamed up with ServiceNow, the workflow automation platform, to create StarCoder, an open source code generator with a less restrictive license than some of the others out there. The original came online early last year, and work has been underway on a follow-up, StarCoder 2, ever since.

StarCoder 2 isn't a single code-generating model, but rather a family. Released today, it comes in three variants, the first two of which can run on most modern consumer GPUs:

- A 3-billion-parameter (3B) model trained by ServiceNow
- A 7-billion-parameter (7B) model trained by Hugging Face
- A 15-billion-parameter (15B) model trained by Nvidia, the newest supporter of the StarCoder project

The End of Software Engineering!

Is Devin AI the End of Software Engineers? Here's What Experts Say

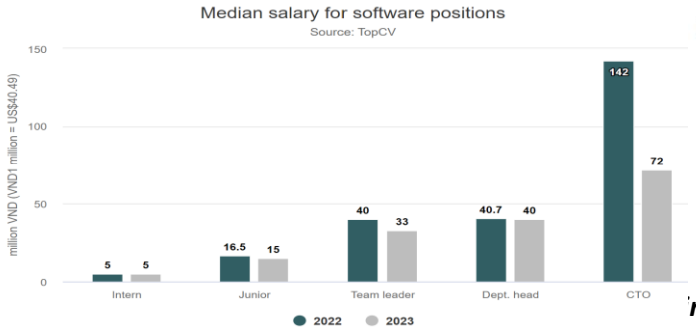
Eddie Wrenn
Editor | Fact Checked by Neil C. Hughes
Last updated: 19 March, 2024

Disclosure | Why Trust Us

Devin AI, 'the world's first fully autonomous software engineer,' likely marks the next disruption that artificial intelligence (AI) leaves on the world: Who needs a skilled coder anymore?

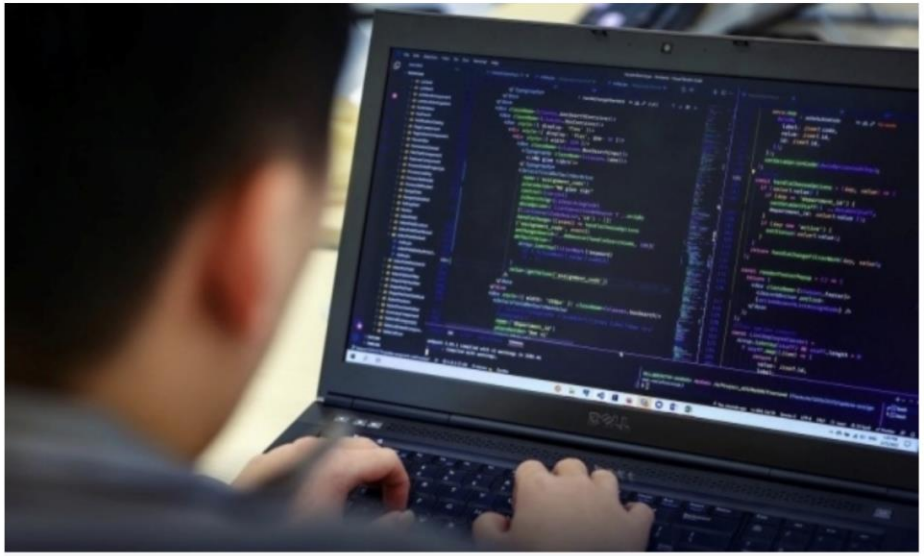
That's the question we pose to various experts across the AI and software engineering fields, as we are not convinced that human coding is going anywhere soon.

When Cognition, the team behind Devin AI, unveiled videos of its early-access master coder last week — writing code from prompts, bug-fixing on the fly, even handling paid-for Upwork tasks for a cheeky ROI — there have been breathless exclamations that this is the end of coding as we know it.



Salaries decline across board for software engineers

By **Khuong Nha** March 17, 2024 | 08:05 pm PT

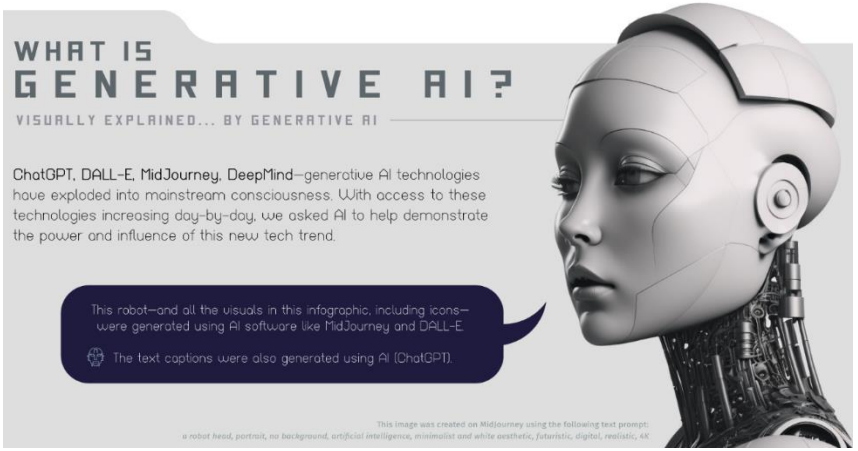


Program engineer in Vietnam. Photo by Funix

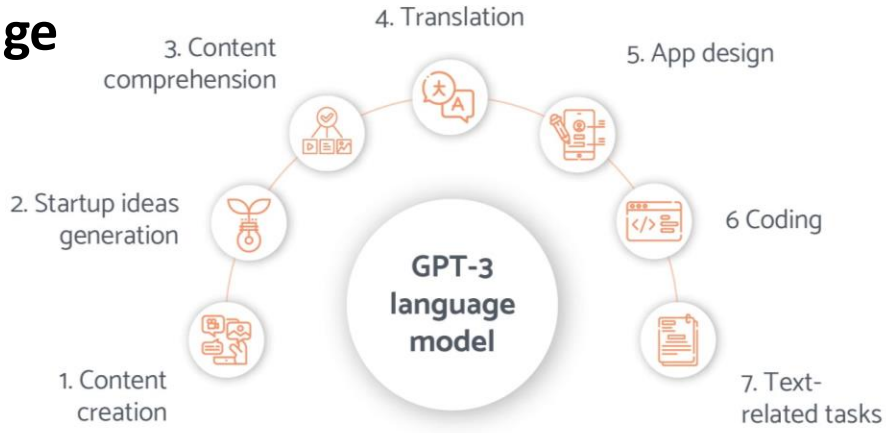
The median salaries for software engineers plummeted in 2023, according to human resources technology platform TopCV.

TopCV's latest annual report shows that salaries for software jobs, while still high, were much lower than in 2022.

Generative AI

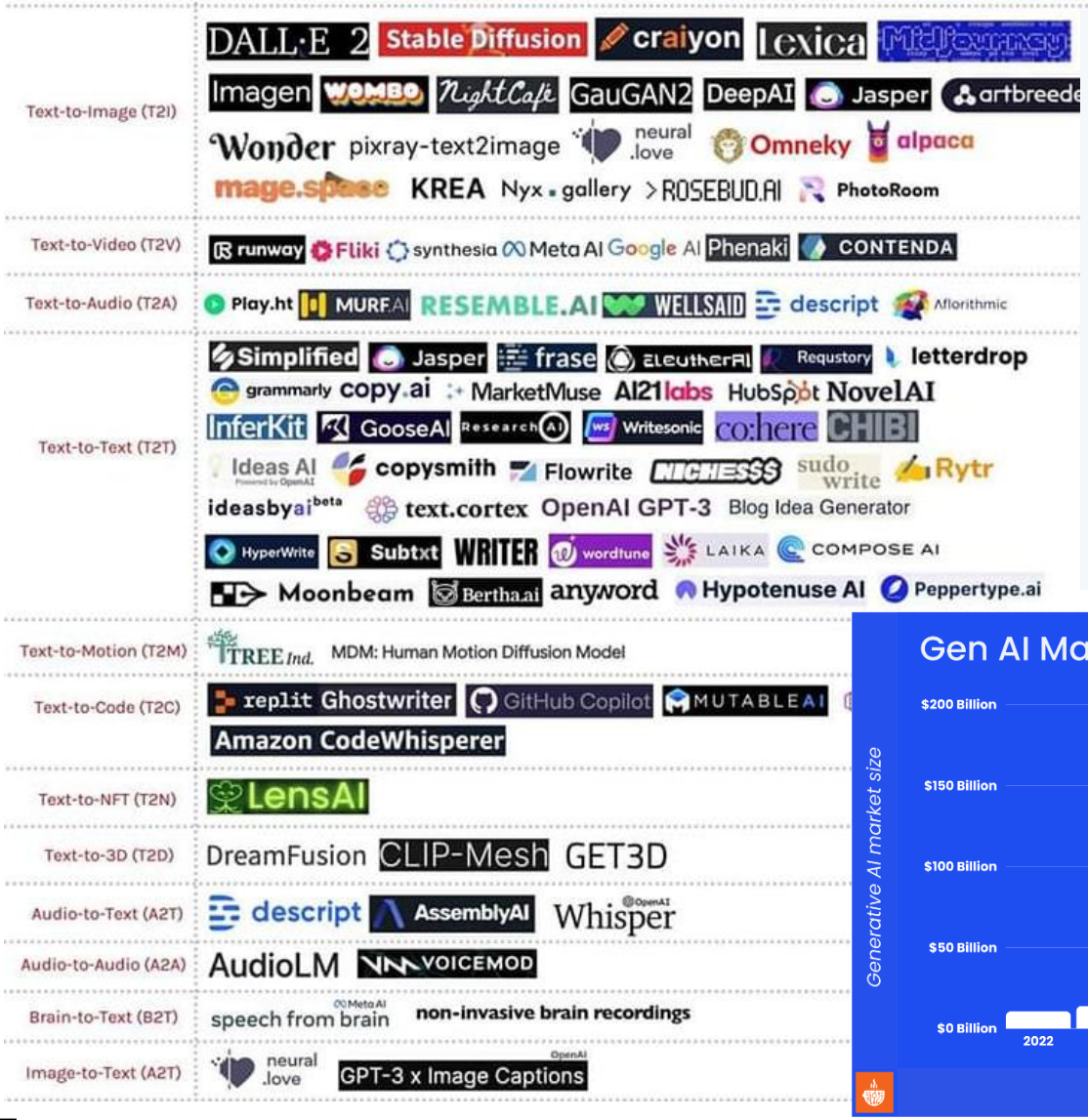


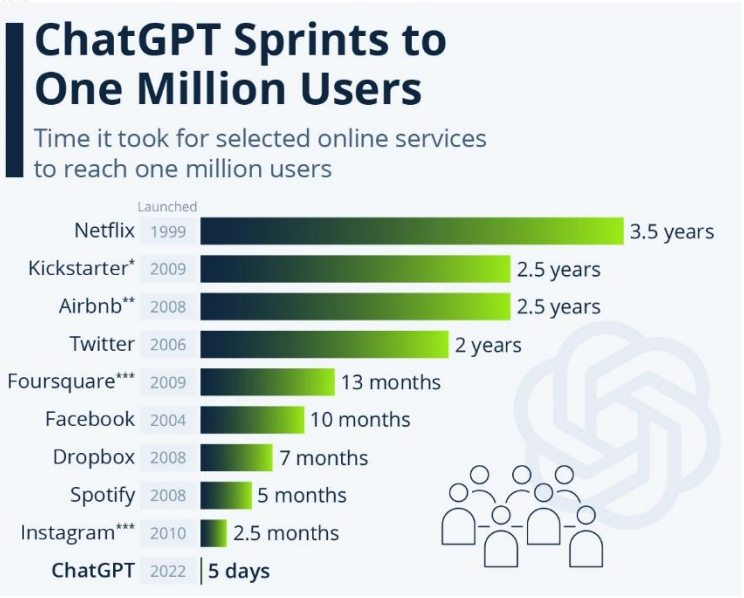
- **Generative AI (GenAI)** can **create a wide variety of data, such as images, videos, audio, text, and 3D models**
- **Learning patterns from existing data**, and then **using this knowledge to generate new** and unique **outputs in response to prompts**
- GenAI produce **highly realistic and complex content** that mimics human creativity

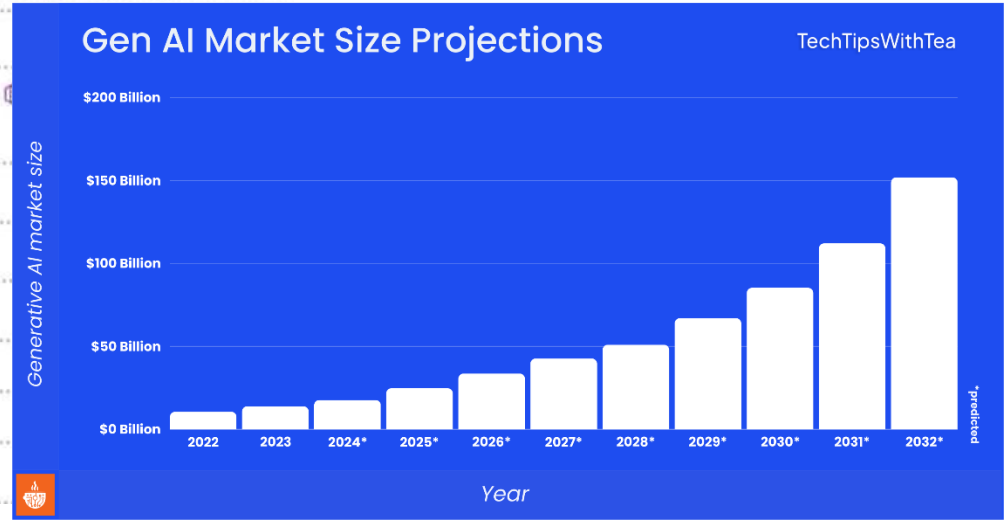


- **Large language models (LLM)** consist of a **very large neural network** with billions of weights, trained on very large quantities of text

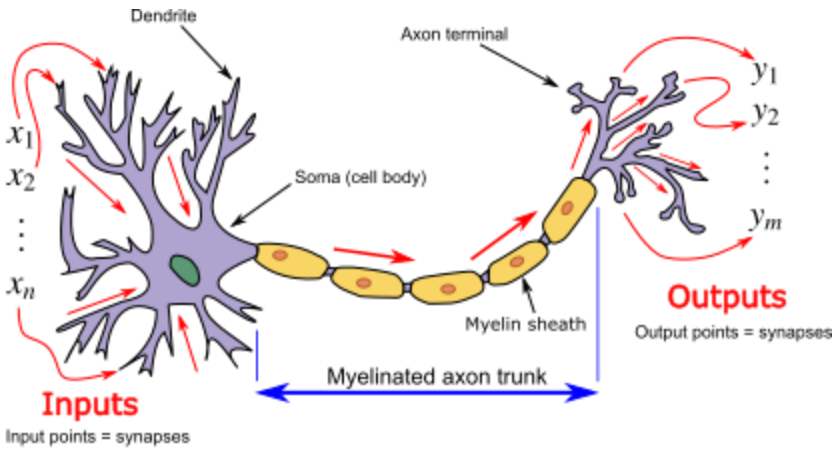
Generative AI

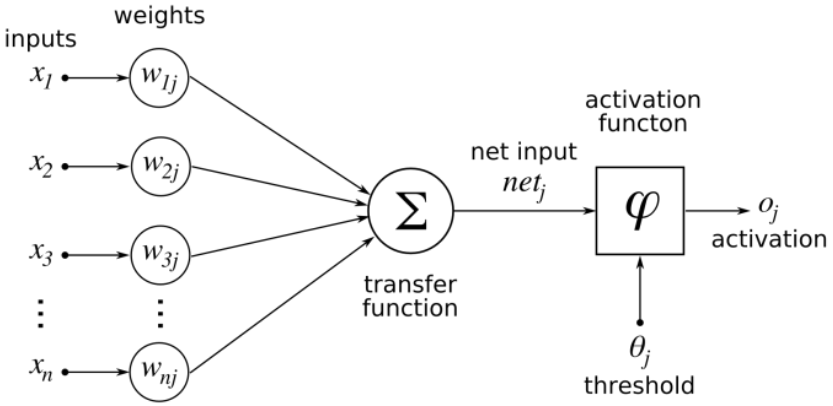


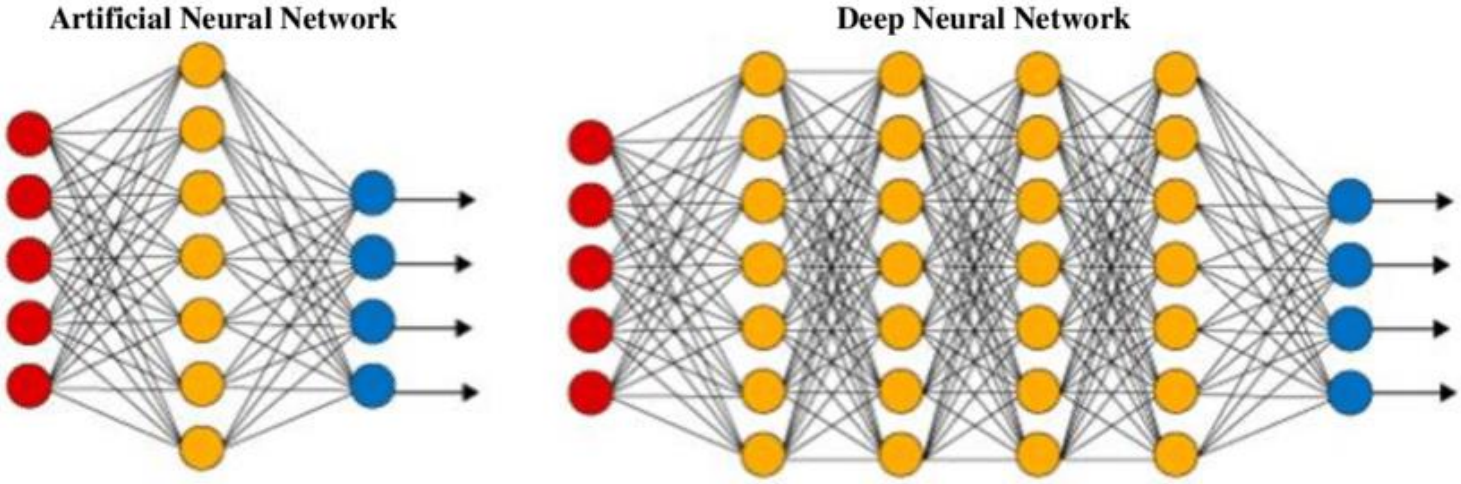




Artificial Neural Networks







GenAI Competition

Google Search Engine

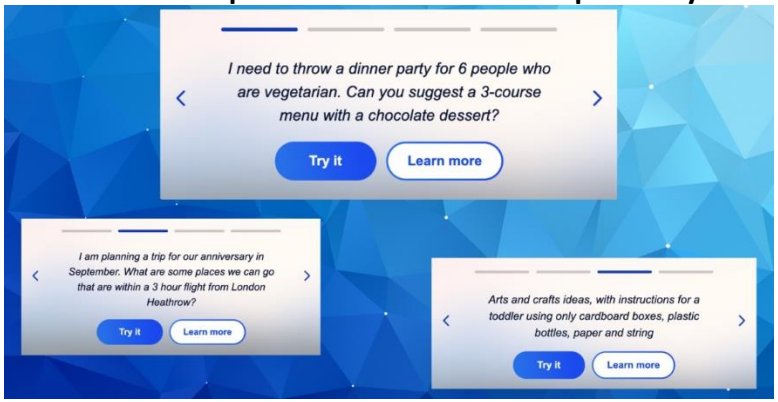
- **Hummingbird Algorithm**
  - **PageRank algorithm** that covers a specific way of **giving pages credit based on the links from other pages** pointing at them
  - Panda, Penguin and Payday to fight **spam**
  - Pigeon designed to improve **local results**
  - Top Heavy designed to demote **ad-heavy** pages,
  - Mobile Friendly designed to reward m**obile-friendly** pages
  - Pirate designed to fight **copyright infringement**

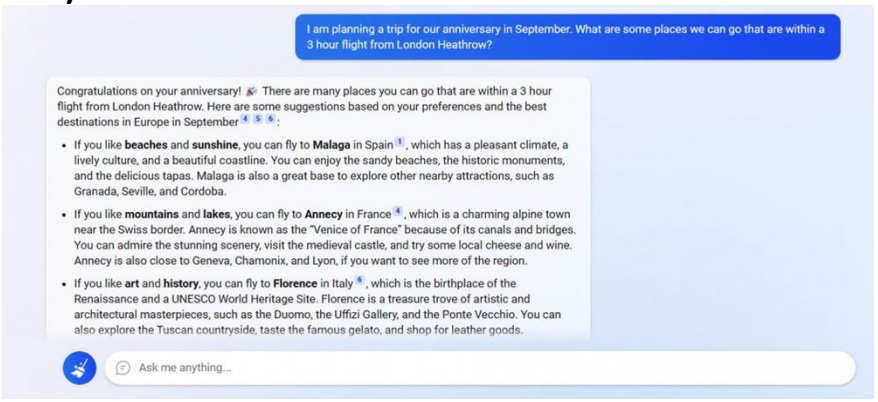
- **200 major ranking signals (up to 10,000 variations or sub-signals)**
- **Google RankBrain Algorithm**
  - **Machine-Learning** System
  - **3rd most important signal**
  - Interpret searches people submit to **find pages that might not have the exact words** that were searched for

Bing Search Engine

Reinventing search with AI with four technical breakthroughs:

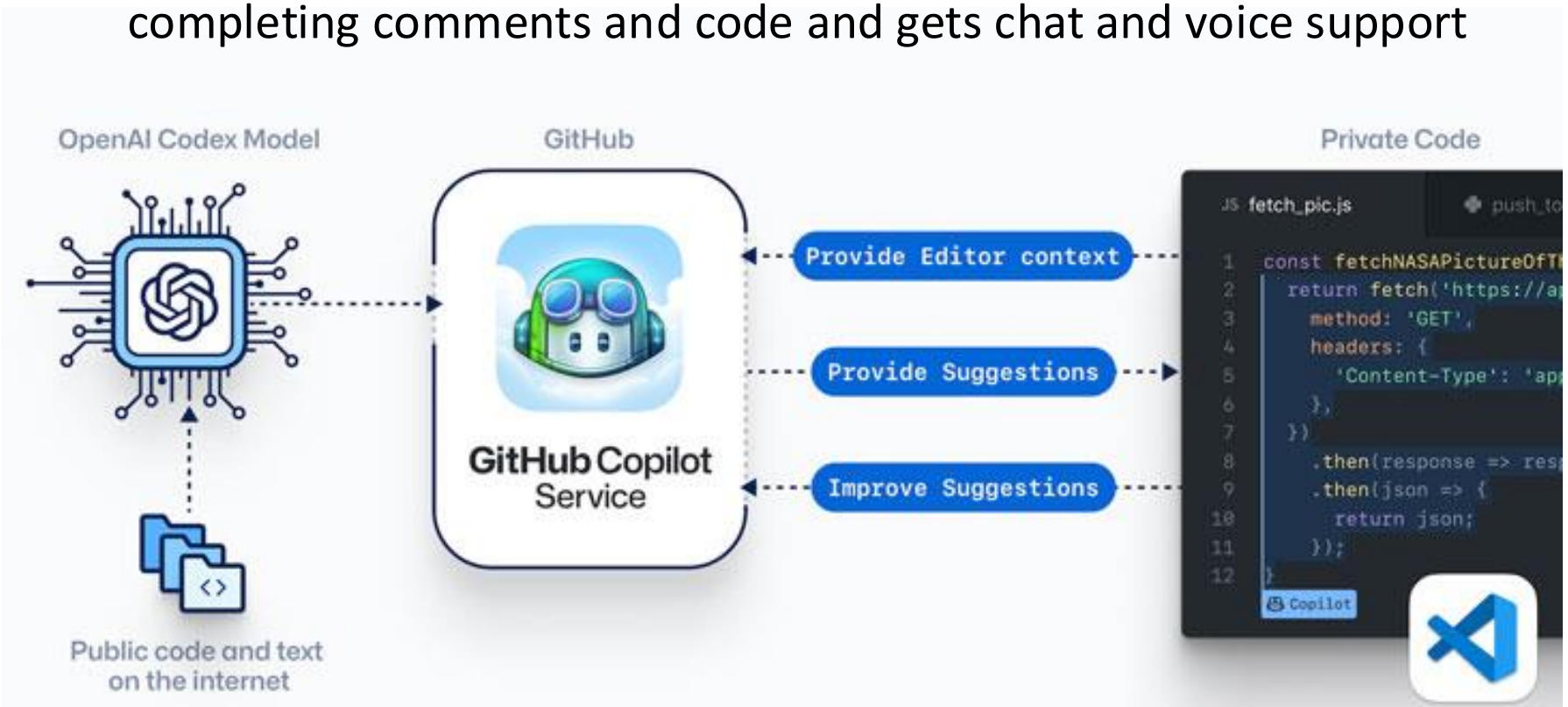
- **Next-generation OpenAI model more powerful than ChatGPT** and customized specifically for search. It takes key learnings and advancements from ChatGPT and GPT-3.5/4.0 and it is even faster, more accurate and more capable
- **Microsoft Prometheus model.** Proprietary way of working with OpenAI model that allows to best leverage its power
- **Applying AI to core search algorithm**. AI model to core Bing search ranking engine, which led to the largest jump in relevance in two decades. Search queries are more accurate and more relevant
- **New user experience.** Reimagining interaction with search, browser and chat into a unified experience in a completely new way to interact with the web.





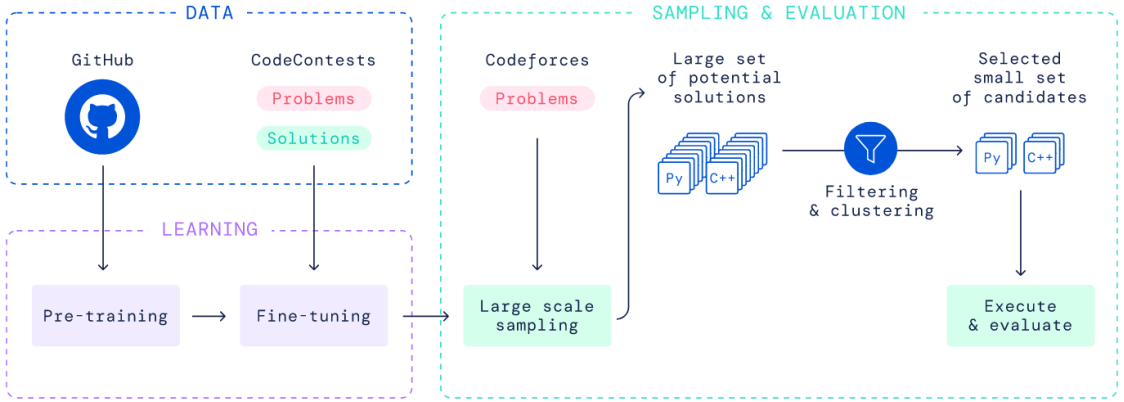
GitHub Copilot

- **June 2021: GitHub Releases Copilot**
- **March 2023: GitHub uses OpenAI's GPT-4** to go way beyond auto-completing comments and code and gets chat and voice support



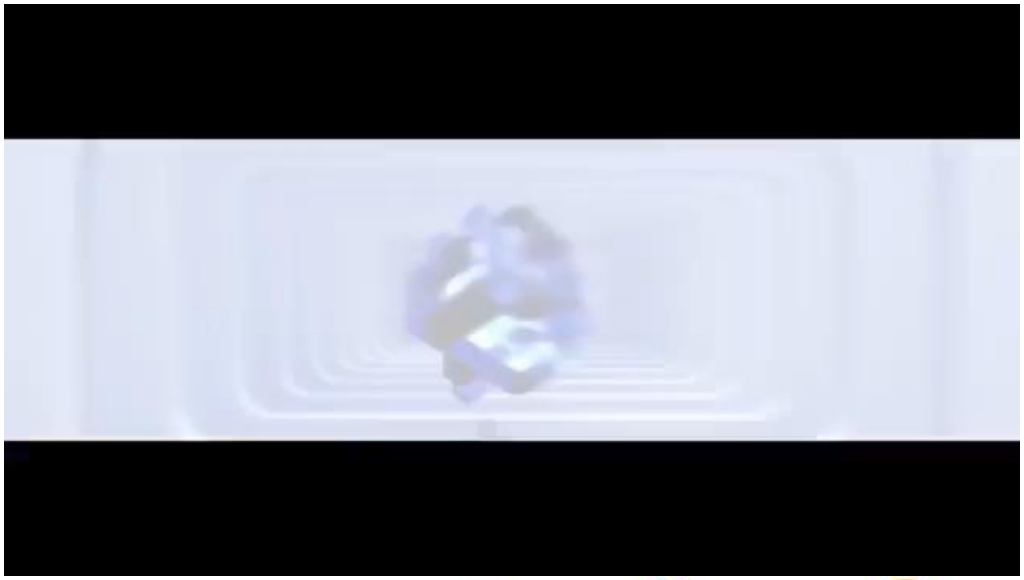
Alpha Code - DeepMind

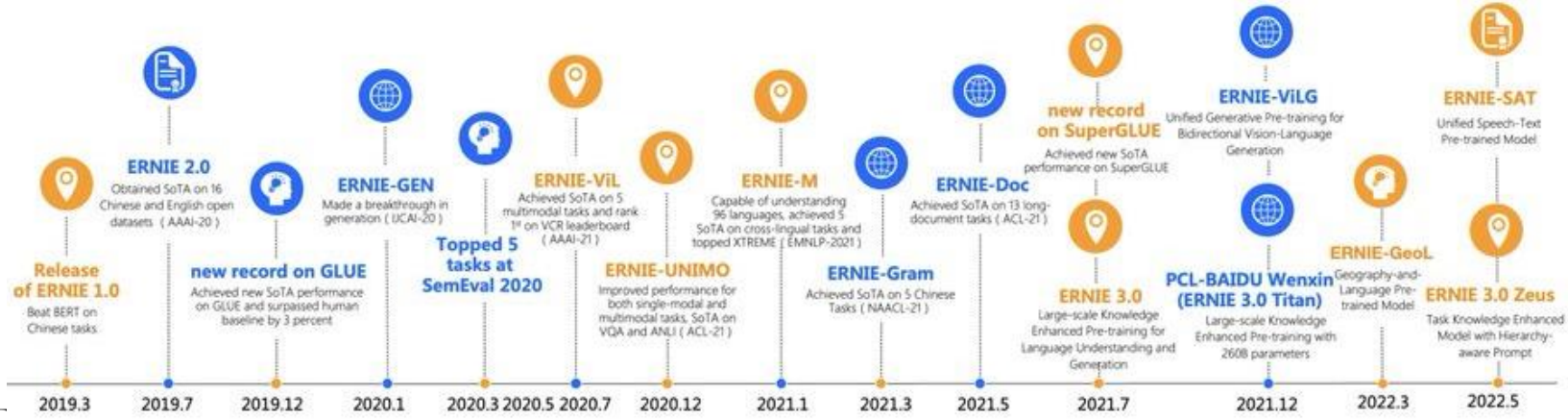
- Solving novel problems and setting a new milestone in competitive programming
- Achieved an estimated rank within the **top 54% of participants in programming competitions**
- Solving new problems that require a combination of **critical thinking, logic algorithms, coding, and natural language understanding**



Ernie Bot - Baidu

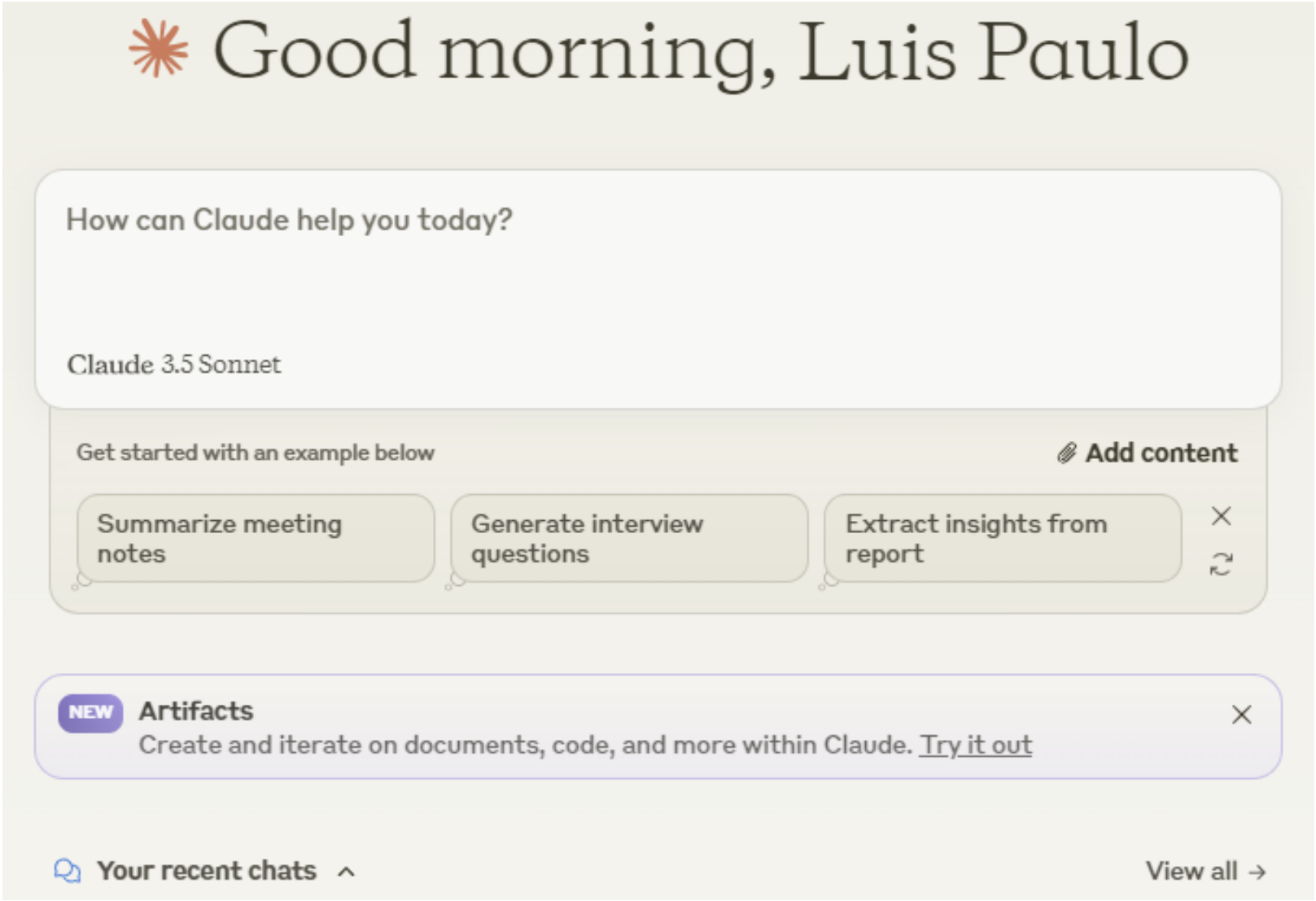
- **ERNIE - Enhanced Representation through Knowledge Integration**
- AI chatbot product of **Baidu**, started 2019
- Based on LLM "**Ernie 3.0-Titan**“
- Released on **March 17, 2023**





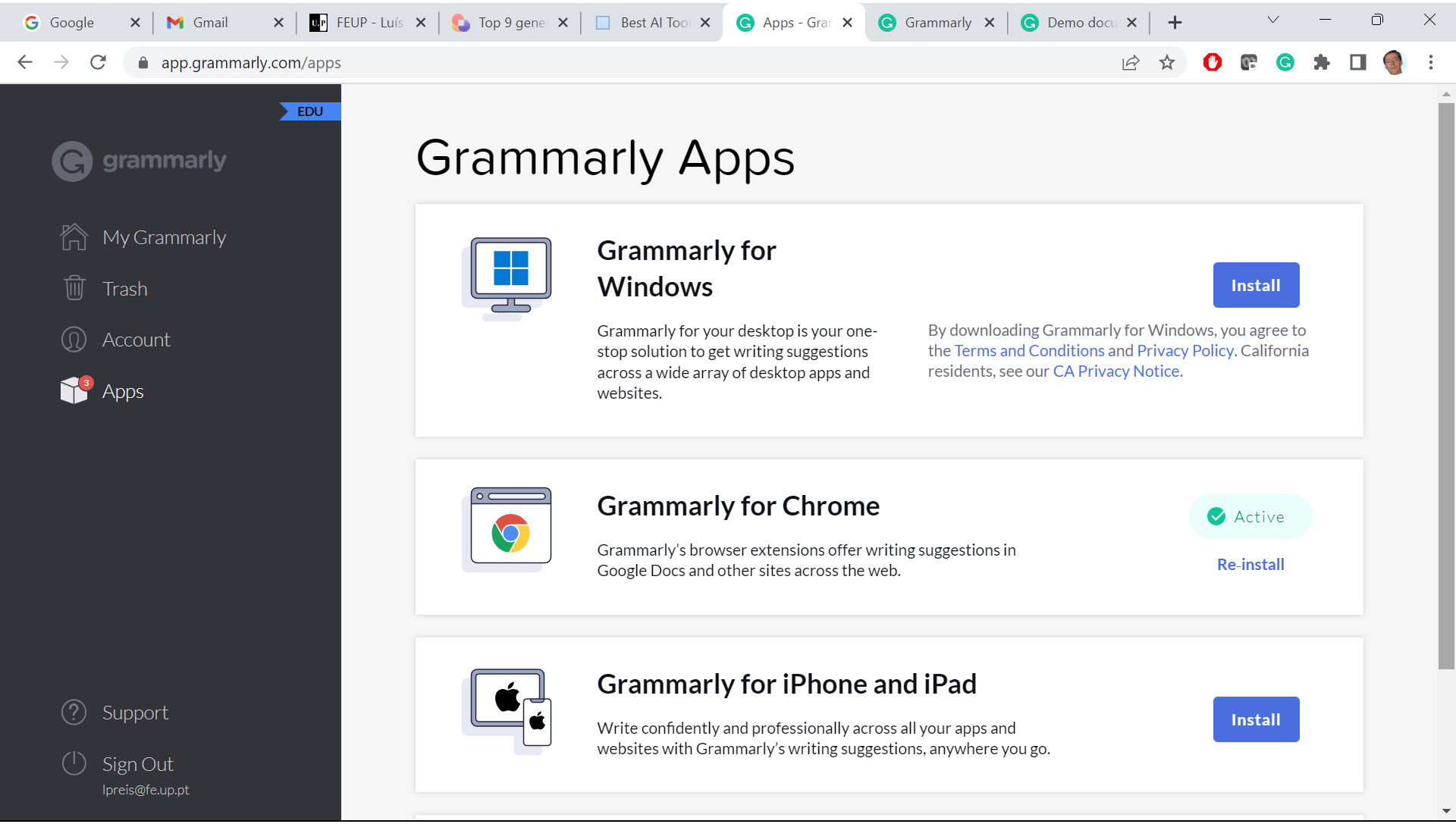
Gemini

Claude 3.5


Good morning, Luis Paulo
How can Claude help you today?
Claude 3.5 Sonnet
Get started with an example below
Add content
Summarize meeting notes
Generate interview questions
Extract insights from report
NEW
Artifacts
Create and iterate on documents, code, and more within Claude. Try it out
Your recent chats
View all

Microsoft Copilot

Grammarly


Google
Gmail
FEUP - Luís
Top 9 gene
Best AI Tool
Apps - Gra
Grammarly
Demo docu
app.grammarly.com/apps
EDU
grammarly
My Grammarly
Trash
Account
Apps
Support
Sign Out
lpreis@fe.up.pt
Grammarly Apps
Grammarly for Windows
Grammarly for your desktop is your one-stop solution to get writing suggestions across a wide array of desktop apps and websites.
Install
By downloading Grammarly for Windows, you agree to the Terms and Conditions and Privacy Policy. California residents, see our CA Privacy Notice.
Grammarly for Chrome
Grammarly's browser extensions offer writing suggestions in Google Docs and other sites across the web.
Active
Re-install
Grammarly for iPhone and iPad
Write confidently and professionally across all your apps and websites with Grammarly's writing suggestions, anywhere you go.
Install

DeepL

DeepL Translator | DeepL Pro | Why DeepL? | API | Plans and pricing | Apps FREE | Start free trial | Login

Translate text
31 languages

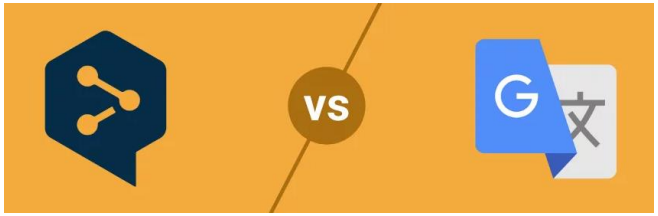
Translate files
.pdf, .docx, .pptx

Pick up where you left off with your **30-day free trial** of DeepL Pro.

Detect language | English (UK) | Glossary

Type to translate.

Drag and drop to translate PDF, Word (.docx), and PowerPoint (.pptx) files with our document translator.



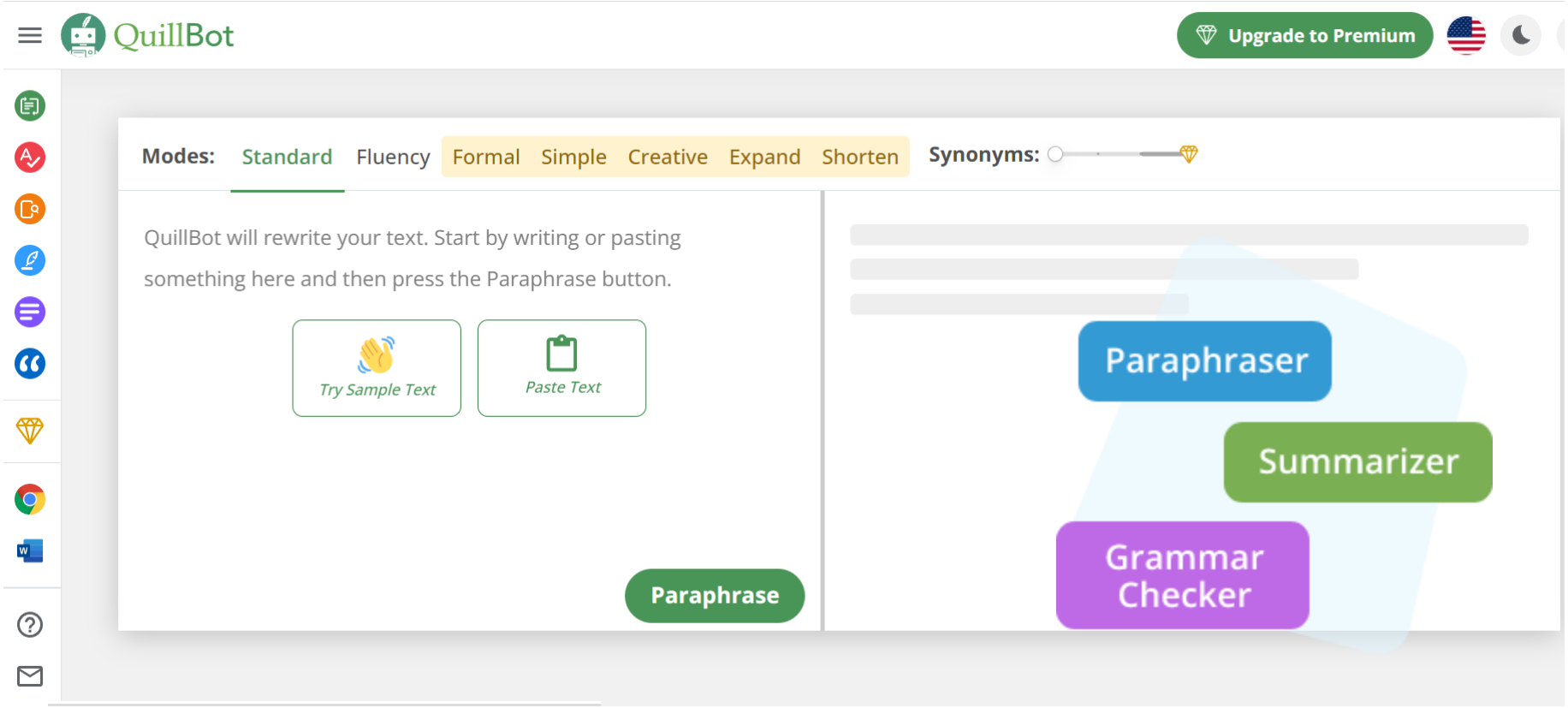
DeepL Translator

Tech giants Google, Microsoft and Facebook are all applying the lessons of machine learning to translation, but a small company called DeepL has outdone them all and raised the bar for the field.
Its translation tool is just as quick as the outsized competition, but more accurate and nuanced than any we've tried.

TechCrunch, USA

TechCrunch

QuillBot



QuillBot

Vocabulary enhancement feature, Thesaurus powered by AI, Integrations (Word, GoogleDocs), Grammar check, Summarizing tool

Dall-E 3

- **DALL·E 3 is an AI system that can create realistic images and art from a description in natural language**

An astronaut riding a horse in photorealistic style

A beach similar to "praia da baia" in Espinho with surfers, surfing in the sea and big waves

DALL·E 2 | Midjourney | Stable Diffusion | Canva AI-ASSISTED DESIGN GENERATOR

Explore GPTs

GPTs

Discover and create custom versions of ChatGPT that combine instructions, extra knowledge, and any combination of skills.

Search public GPTs

1 **Canva**
Effortlessly design anything: presentations, logos, social media posts and more.
By canva.com

2 **Logo Creator**
Use me to generate professional logo designs and app icons!
By Chase Lean

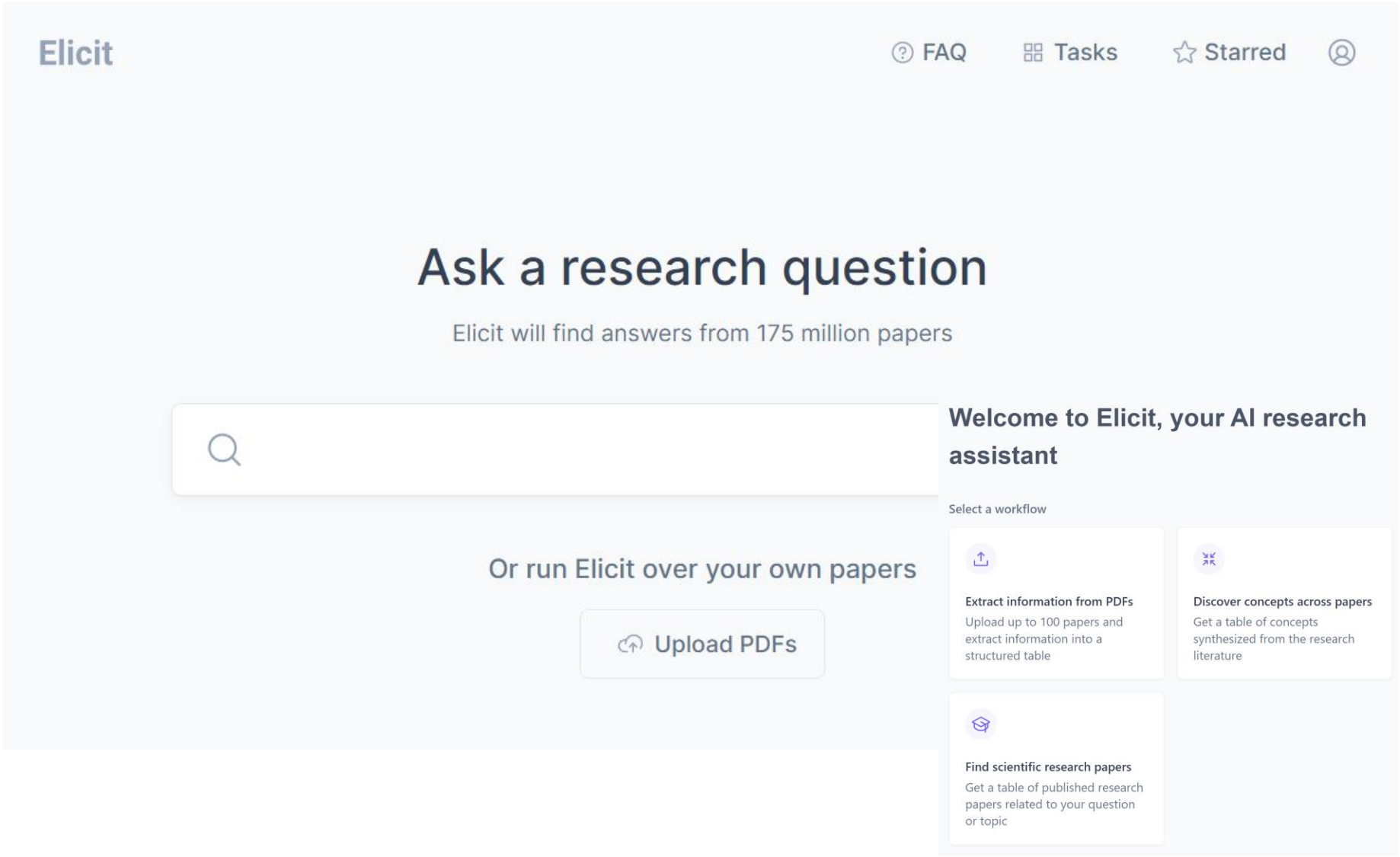
3 **Consensus**
Your AI Research Assistant. Search 200M academic papers from Consensus, get science-based answers, and draft content...
By Consensus

4 **image generator**
A GPT specialized in generating and refining images with a mix of professional and friendly tone.image generator
By NAIF J ALOTAIBI

5 **Video GPT by VEED**
AI Video Maker. Generate videos for social media - YouTube, Instagram, TikTok and more! Free text to video & speech tool wit...
By veed.io

6 **Grimoire**
Code Wizard 🧙 Learn to Prompt-gram! Create a website (or anything) with a sentence. 20+ Hotkeys for coding flows....
By gptavern.mindgoblinstudios.com

Elicit


Elicit
FAQ
Tasks
Starred
Ask a research question
Elicit will find answers from 175 million papers
Or run Elicit over your own papers
Upload PDFs
Welcome to Elicit, your AI research assistant
Select a workflow
Extract information from PDFs
Upload up to 100 papers and extract information into a structured table
Discover concepts across papers
Get a table of concepts synthesized from the research literature
Find scientific research papers
Get a table of published research papers related to your question or topic

Consensus

You

write me a state of the art about AI ethics

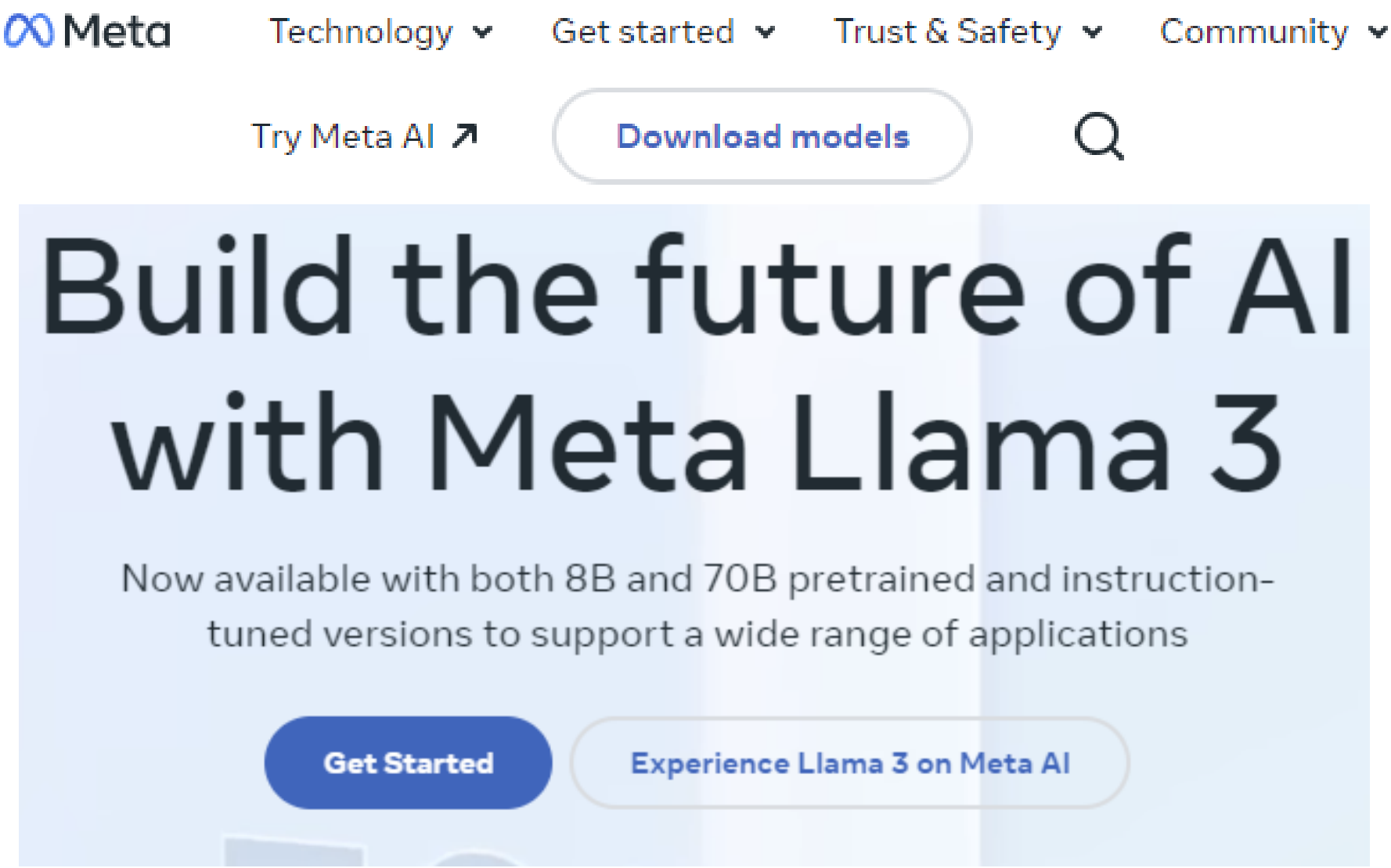
Consensus

Talked to chat.consensus.app

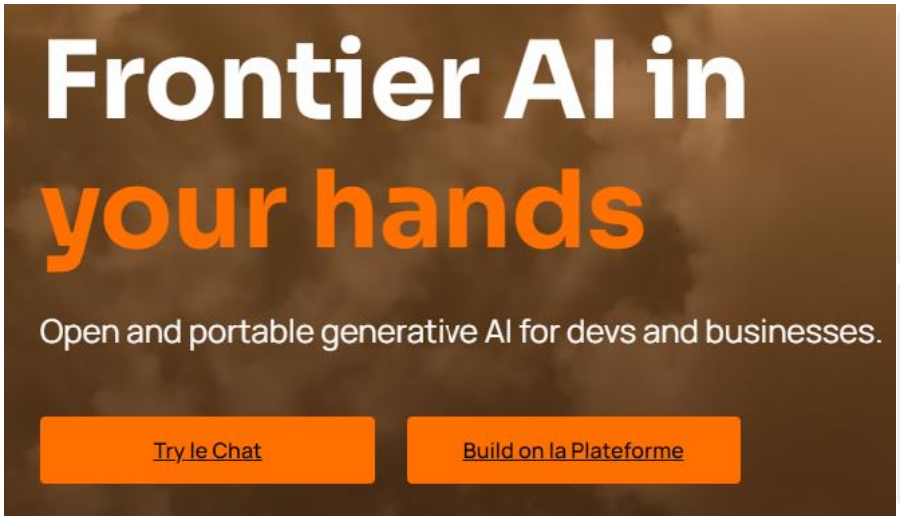
The state of the art in AI ethics encompasses a wide range of topics, from algorithmic injustice and discrimination to labor impacts, privacy concerns, and the societal implications of AI. Here's a summary of key findings from recent research:

1. **Algorithmic Injustice and Discrimination**: AI technologies often reflect and perpetuate existing societal biases, leading to concerns about algorithmic injustice and discrimination. This issue is especially pertinent in the context of societal and labor impacts, as well as in the development of ethical AI frameworks (Gupta et al., 2021).
2. **Ethical AI Development**: There is a growing emphasis on the need to develop AI systems ethically. This includes addressing issues of privacy, security, and the risk of misinformation. The role of AI in societal contexts, such as labor and social media, necessitates careful ethical consideration (Vakkuri et al., 2020)
3. **Challenges in Implementing AI Ethics**: Despite the proliferation of ethical guidelines for AI, there's a notable gap between these guidelines and their implementation in practice. This gap is evident in industries and companies dealing with AI, indicating a need for better integration of ethical principles in AI development (Vakkuri et al., 2022).
4. **AI and Labor Impact**: The impact of AI on labor and employment is a significant ethical concern. AI's potential to automate jobs raises questions about job displacement and the need for new skills and training for the workforce (Gupta et al., 2021).
5. **Privacy and Security**: With the increasing use of AI in various domains, issues of data privacy and security are paramount. Ensuring the protection of personal data and preventing misuse of AI technologies are key ethical challenges (Gupta et al., 2021).
6. **Ethical Implications in Specific AI Applications**: Specialized areas like AI-generated art and machine translation also present unique ethical challenges. For instance, AI-generated art raises questions about the representation of artistic styles and cultural nuances (Srinivasan & Uchino, 2021).

Llama 3 - Meta


Meta
Technology
Get started
Trust & Safety
Community
Try Meta AI
Download models
Build the future of AI with Meta Llama 3
Now available with both 8B and 70B pretrained and instruction-tuned versions to support a wide range of applications
Get Started
Experience Llama 3 on Meta AI

Mistral



Openness

We lead the market of open source generative technologies to bring trust and transparency in the field and foster decentralised technology development

Portability

Our technology is available through serverless APIs, public cloud services (on Azure AI, on Amazon Bedrock), and for VPC/on-premise deployment. Our independence from cloud providers is a guarantee of our customers' independence.

Value and speed

Our flagship model, Mistral Large, has independently validated top-tier reasoning in multiple languages. All our models bring unmatched value and latency at their price points.

Customisation

Our models can be fine-tuned and modified at will for your business to create differentiated AI applications.

Build on Open Source

Under the Apache 2.0 license, our 3 open source models Mistral 7B, Mixtral 8x7B, Mixtral 8x22B are usable and customisable for a variety of use cases. They can be downloaded or used on demand via our platform.

- Download them for deployment in your own environment
- Use them on La Plateforme at market-leading availability, speed, and quality control

Gemma 2 – Gemini Google

Gemma 2 is now available to researchers and developers

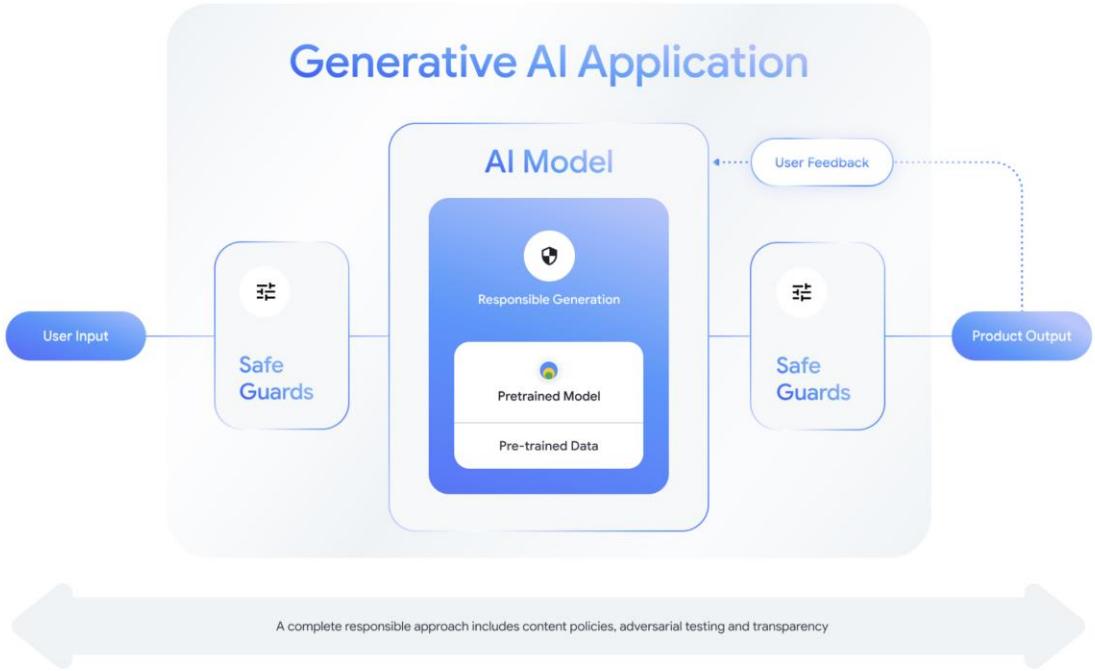
Gemma

google/gemma

Gemma 2 offers best-in-class performance, runs at incredible speed across different hardware and easily integrates with other AI tools.

Jun 27, 2024 · 4 min read

Gemma is a family of lightweight, open models built from the research and technology that Google used to create the Gemini models.



Responsible Generative AI Toolkit

Send feedback

This toolkit provides resources to apply best practices for responsible use of open models such as the Gemma models, including:

- Guidance on setting safety policies, safety tuning, safety classifiers and model evaluation.
- The Learning Interpretability Tool (LIT) for investigating Gemma's behavior and addressing potential issues.
- A methodology for building robust safety classifiers with minimal examples.

This version of the toolkit focuses on English text-to-text models only. You can provide feedback to make this toolkit more helpful through the feedback mechanism link at the bottom of the page.

When building with Gemma, you should take a holistic approach to responsibility and consider all the possible challenges at the application and model levels. This toolkit covers risk and mitigation techniques to address safety, privacy, fairness, and accountability.

LLMs Comparison

	Claude 3 Opus	Claude 3 Sonnet	Claude 3 Haiku	GPT-4	GPT-3.5	Gemini 1.0 Ultra	Gemini 1.0 Pro
Undergraduate level knowledge *MMLU*	86.8% 5 shot	79.0% 5-shot	75.2% 5-shot	86.4% 5-shot	70.0% 5-shot	83.7% 5-shot	71.8% 5-shot
Graduate level reasoning *GPQA, Diamond*	50.4% 0-shot CoT	40.4% 0-shot CoT	33.3% 0-shot CoT	35.7% 0-shot CoT	28.1% 0-shot CoT	—	—
Grade school math *GSM8K*	95.0% 0-shot CoT	92.3% 0-shot CoT	88.9% 0-shot CoT	92.0% 5-shot CoT	57.1% 5-shot	94.4% Maj1@32	86.5% Maj1@32
Math problem-solving *MATH*	60.1% 0-shot CoT	43.1% 0-shot CoT	38.9% 0-shot CoT	52.9% 4-shot	34.1% 4-shot	53.2% 4-shot	32.6% 4-shot
Multilingual math *MGSM*	90.7% 0-shot	83.5% 0-shot	75.1% 0-shot	74.5% 8-shot	—	79.0% 8-shot	63.5% 8-shot
Code *HumanEval*	84.9% 0-shot	73.0% 0-shot	75.9% 0-shot	67.0% 0-shot	48.1% 0-shot	74.4% 0-shot	67.7% 0-shot
Reasoning over text *DROP, F1 score*	83.1 3-shot	78.9 3-shot	78.4 3-shot	80.9 3-shot	64.1 3-shot	82.4 Variable shots	74.1 Variable shots
Mixed evaluations *BIG-Bench-Hard*	86.8% 3-shot CoT	82.9% 3-shot CoT	73.7% 3-shot CoT	83.1% 3-shot CoT	66.6% 3-shot CoT	83.6% 3-shot CoT	75.0% 3-shot CoT
Knowledge Q&A *ARC-Challenge*	96.4% 25-shot	93.2% 25-shot	89.2% 25-shot	96.3% 25-shot	85.2% 25-shot	—	—
Common Knowledge *HellaSwag*	95.4% 10-shot	89.0% 10-shot	85.9% 10-shot	95.3% 10-shot	85.5% 10-shot	87.8% 10-shot	84.7% 10-shot

LLMs Comparison

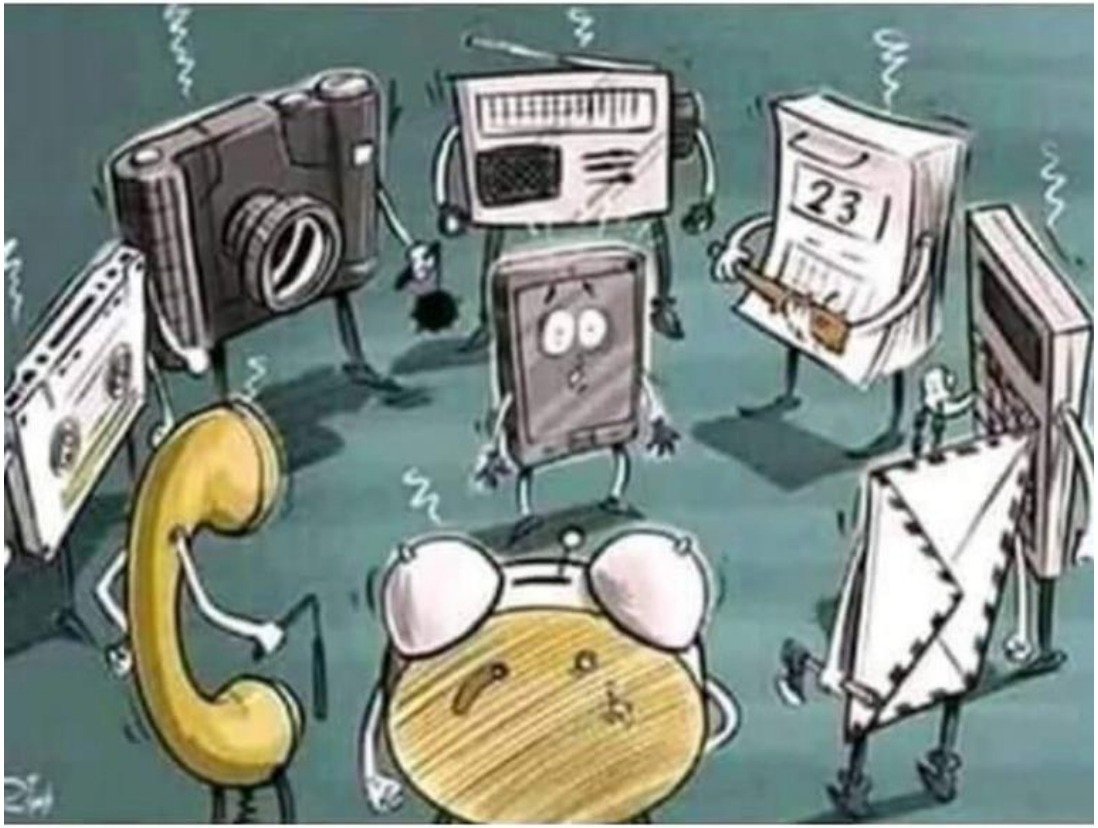
LMSYS Chatbot Arena Leaderboard

Rank* (UB)	Model	Arena Elo	95% CI	Votes	Organization	License	Knowledge Cutoff
1	GPT-4o-2024-05-13	1287	+3/-3	46179	OpenAI	Proprietary	2023/10
2	Claude 3.5 Sonnet	1272	+6/-8	7278	Anthropic	Proprietary	2024/4
2	Gemini-Advanced-0514	1267	+3/-4	36867	Google	Proprietary	Online
2	Gemini-1.5-Pro-API-0514	1263	+4/-4	39321	Google	Proprietary	2023/11
4	Gemini-1.5-Pro-API-0409-Preview	1257	+3/-3	55731	Google	Proprietary	2023/11
4	GPT-4-Turbo-2024-04-09	1257	+3/-3	68075	OpenAI	Proprietary	2023/12
7	GPT-4-1106-preview	1251	+2/-2	84482	OpenAI	Proprietary	2023/4
7	Claude 3 Opus	1248	+2/-2	136753	Anthropic	Proprietary	2023/8
7	GPT-4-0125-preview	1246	+3/-2	77746	OpenAI	Proprietary	2023/12
10	Yi-Large-preview	1240	+3/-3	43372	01 AI	Proprietary	Unknown
11	Gemini-1.5-Flash-API-0514	1229	+3/-4	36996	Google	Proprietary	2023/11
12	Yi-Large	1215	+6/-6	7930	01 AI	Proprietary	Unknown
12	Bard (Gemini Pro)	1208	+7/-5	11853	Google	Proprietary	Online
12	GLM-4-0520	1208	+8/-7	6258	Zhipu AI	Proprietary	Unknown
12	Llama-3-70b-Instruct	1207	+3/-2	139037	Meta	Llama 3 Community	2023/12
12	Nemotron-4-340B-Instruct	1205	+5/-5	13461	Nvidia	NVIDIA Open Model	2023/6
14	Claude 3 Sonnet	1201	+3/-2	105288	Anthropic	Proprietary	2023/8
15	Reka-Core-20240501	1200	+3/-3	53279	Reka AI	Proprietary	Unknown
19	Command R+	1189	+3/-2	71841	Cohere	CC-BY-NC-4.0	2024/3

Technology/AI: Adapt or Resist?!

- COVID-19 - Use of technology
- Educate New Generations
- Ethical Use
- Technical Use
- AI Policy
- Prompt Engineering

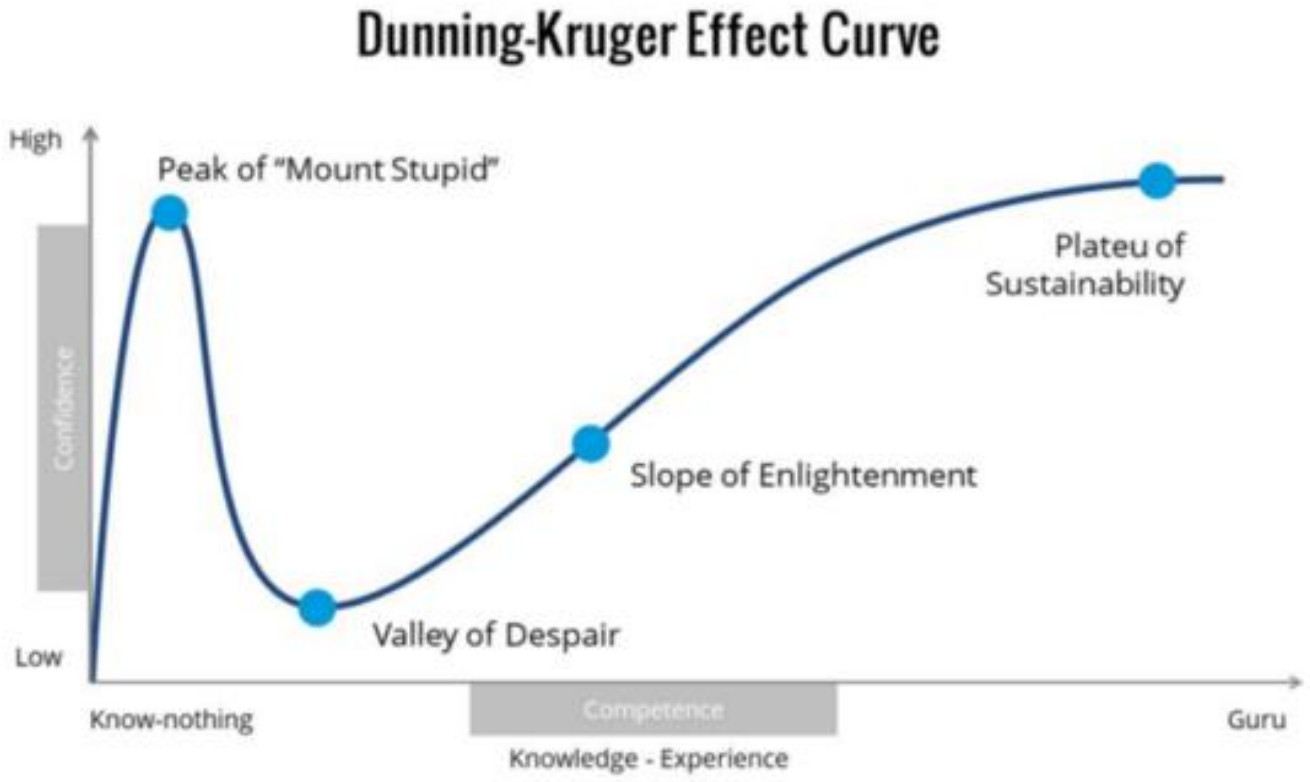
So! You are the guy…



That took all of our Jobs!

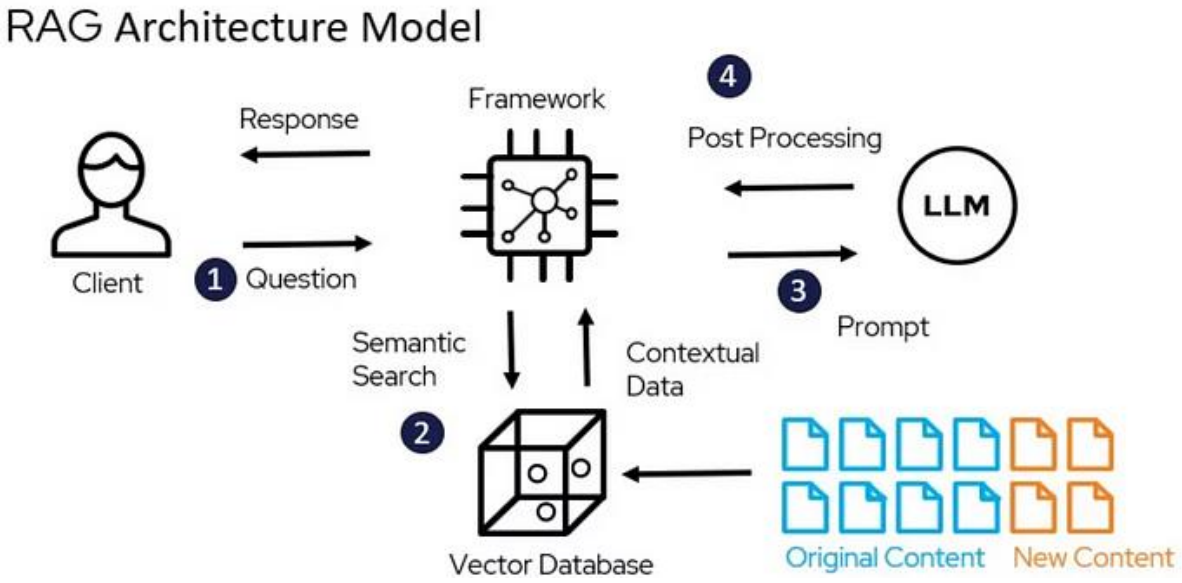
Chat GPT – Dunning-Kruger Effect

Dunning–Kruger effect: Cognitive bias whereby people with low ability, expertise, or experience regarding a certain type of task or area of knowledge tend to overestimate their ability or knowledge. Some researchers include the opposite effect for high performers: tendency to underestimate their skills.



RAG Architecture Model

- **Retrieval Augmented Generation (RAG) has emerged as a powerful technique for improving LLMs.**
- **Retrieving and conditioning external knowledge, RAG allows models to generate more accurate, relevant, and comprehensive text**



- **Advanced RAG enhances each module further with innovations like higher-order retrievers, cross-encoder rerankers, and evidence manipulation architectures**

LLMs Adaptation

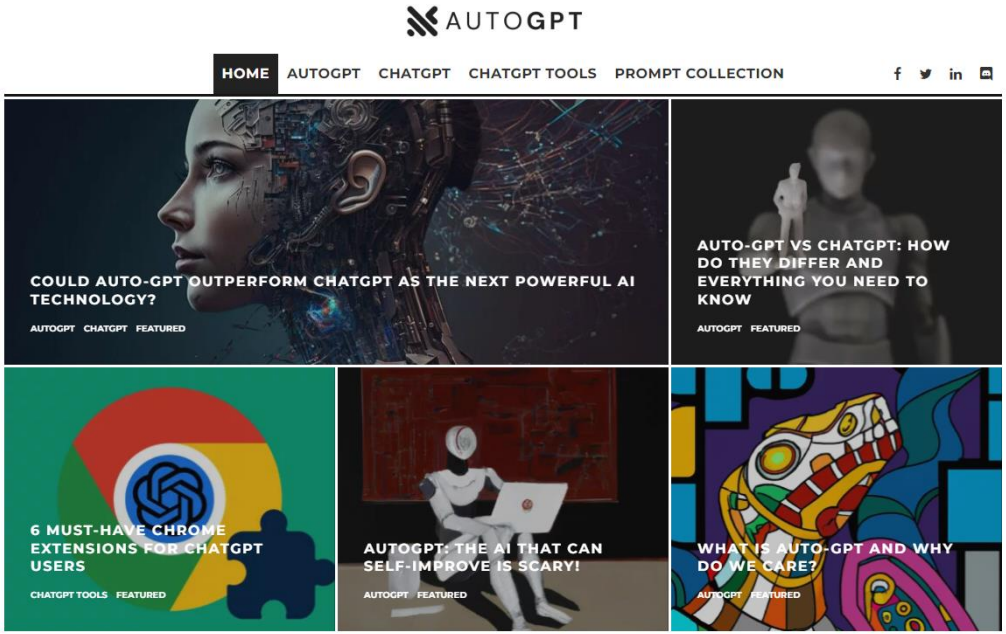
- **Fine-Tuning** - Further training it on a specific dataset relevant to the desired task
- **Prompt Engineering** - Crafting specific prompts that guide the model to generate desired outputs without modifying the model's weights
- **Low-Rank Adaptation (LoRA)** - Adding low-rank matrices to a pre-trained model to adapt it to new tasks with minimal computational overhead
- **Adapters** - Small bottleneck layers added between the layers of the pre-trained model (trained on new data without altering the original model's parameters significantly)
- **Knowledge Distillation** - A smaller model (student) is trained to replicate the behavior of a larger model (teacher). This makes the smaller model more efficient while retaining much of the performance of the larger model

LLMs Adaptation

- **Zero-Shot and Few-Shot Learning** - Using the LLM's ability to perform tasks with little to no task-specific training data by providing examples within the input prompt
- **Data Augmentation** - Expanding the training dataset with synthetic data generated by the model or by other means to improve model robustness and performance
- **Ensemble Methods** - Combining the outputs of multiple models to improve performance and robustness (averaging outputs or voting)
- **Task-Specific Pre-Training** - Pre-training the model on a large corpus of data related to the specific task before fine-tuning it
- **Reinforcement Learning** - Training the model using a reward signal to perform specific tasks better interacting with an environment or feedback from users
- **Continuous Learning** - Continuously improving the model by learning from user interactions and feedback over time

Auto-GPT

- **Auto-GPT is an "AI agent" that given a goal in natural language, will attempt to achieve it by breaking it into sub-tasks and using the internet and other tools in an automatic loop.**
- **Uses GPT-4 APIs to perform autonomous tasks**



Features

- Internet access for searches and information gathering
- Long-term and short-term memory management
- GPT-4 instances for text generation
- Access to popular websites and platforms
- File storage and summarization with GPT-3.5
- Extensibility with Plugins

AI & LLM Key Issues in 2024

- Small Language Models will drive enterprise AI adoption
- Mixture of Experts (MoE) + LoRA will enable small language models to outperform 10x larger LLMs
- Open-source will become the de facto way for LLMs use
- Large language models will adopt modular architectures
- Synthetic data revolution arrives to LLMs
- LLM hallucinations disappear as training techniques evolve
- Data (not LLMs) becomes the true competitive advantage
- LLM adoption grows as the Transformer architecture matures
- RAGs (Retrieval-Augmented Generation) used in most LLMs tasks
- Shift from closed to open-source democratizes AI (transparency)
- LLMs for machines, not just people
- LLMs for Robotics

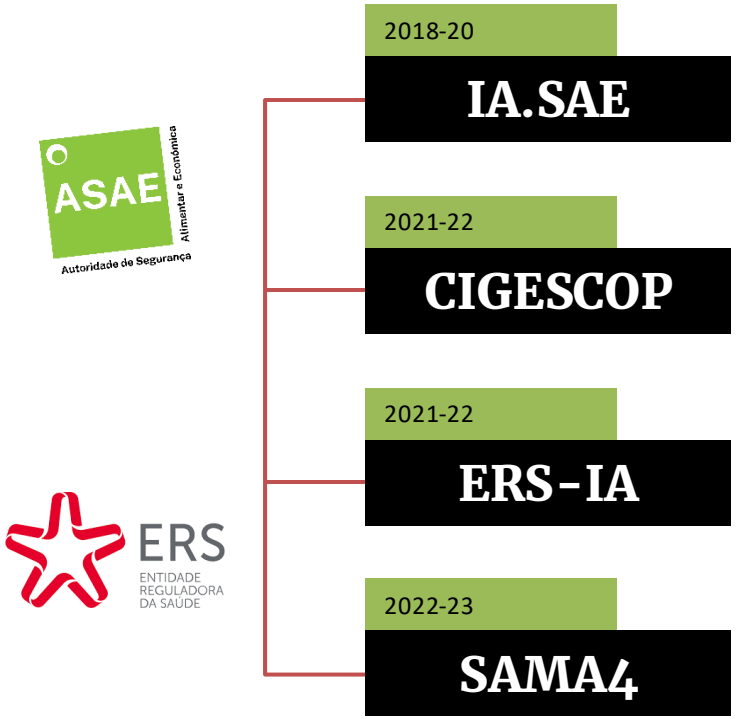
AI Opportunities

- **Data-Driven Decision Making:** Enhance decision-making by processing large datasets efficiently and extracting valuable insights. In food safety and regulatory science, AI can support risk assessments, automate data analysis, and provide predictive modeling for emerging risks.
- **Automation and Efficiency:** Automate repetitive tasks, such as data entry, analysis, and reporting, freeing up human resources to focus on complex decision-making. Reduce errors and increase speed of regulatory processes.
- **Predictive Analytics:** Predictive models can forecast potential food safety issues allowing for preventive measures. AI models can also identify patterns in large datasets.
- **Explainable AI (XAI):** Making AI decisions more transparent, is essential in food safety. Build trust with stakeholders by mitigating "black box" problem.
- **Collaborative Learning:** Federated learning and data sharing between organizations allow training models across diverse datasets without violating privacy.
- **AI for Regulatory Compliance:** AI can help EFSA to ensure compliance with safety standards by continuously monitoring data from various sources, identifying non-compliant products, and automating compliance checks.

AI Challenges

- **Data Readiness and Quality:** Lack of high-quality, standardized data. AI models need large datasets to train effectively, but with consistent, complete and unbiased data
- **Privacy and Ethical Concerns:** Concerns about data privacy, particularly when sharing sensitive information between organizations. GDPR must be carefully considered to ensure compliance
- **Explainability and Trust:** Lack of transparency in AI's decision-making process can erode trust. Explainable AI (XAI) needs further developments.
- **Regulatory and Legal Frameworks:** Regulatory bodies (and everybody…) still catching up with AI's rapid advancements. Ensuring that AI systems comply with international regulations is a challenge
- **Integration with Existing Systems:** Organizations lack infrastructure to fully integrate AI solutions into their workflows. Technology, staff training, and process adaptation
- **Bias in AI Models:** Biases in the data may lead to unfair or inaccurate results. Methods to detect and mitigate bias in AI systems in regulatory and food safety contexts

Successful AI in PA Projects



Artificial Intelligence in Food and Economic Security

IA.SAE

ASAE's Intelligent Operational Management and Control Center

POCI-05-5762-FSE-000215

Predictive Behavior in Health and Analysis of Complaints

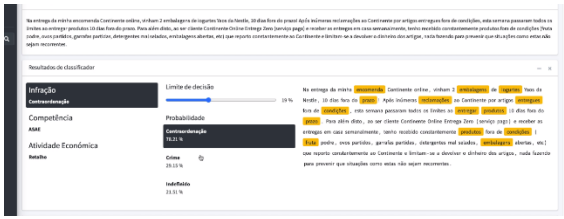
POCI-05-5762-FSE-000205

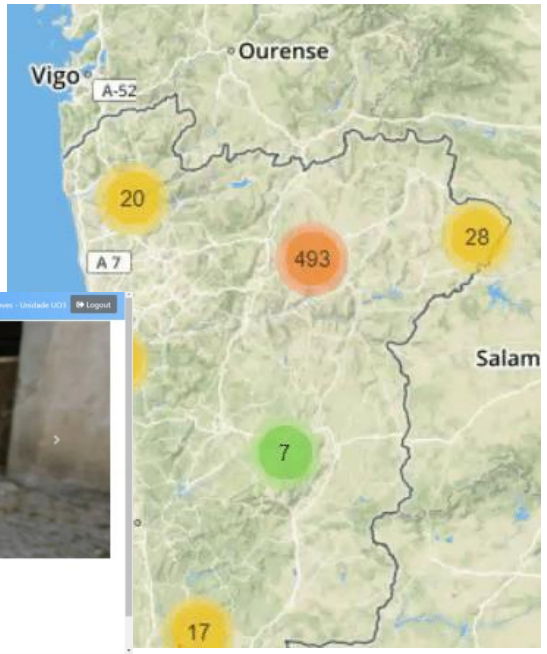
Supervision based on Information and Risk Analysis

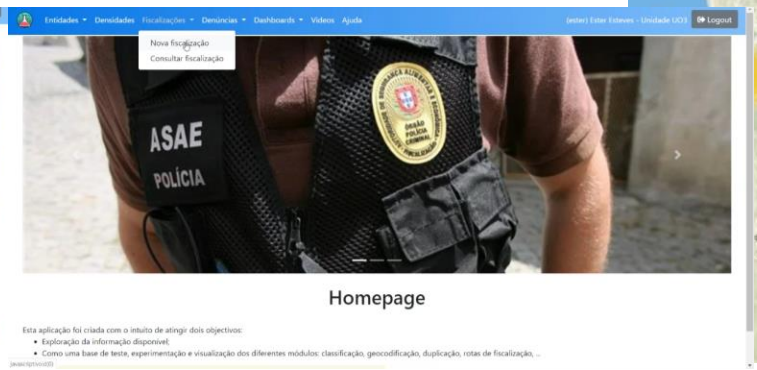
POCI-05-5762-FSE-000389

IA.SAE - Artificial Intelligence in Food and Economic Security

- Risk Analysis and Selection Models of Economic Agents to be Supervised, based on AI and ML applied to ASAE's databases
- Georeferencing of economic agents in Portugal
- Electronic complaint analysis, automatic classification and prioritization
- Intelligent selection of economic agents to be inspected and flexible routes generation
- Intelligent data and KPIs visualization

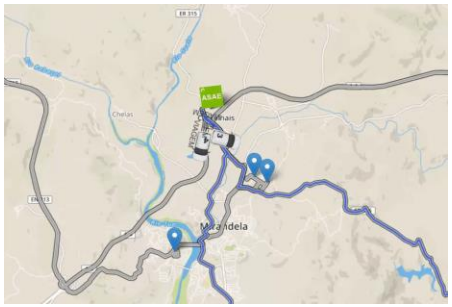


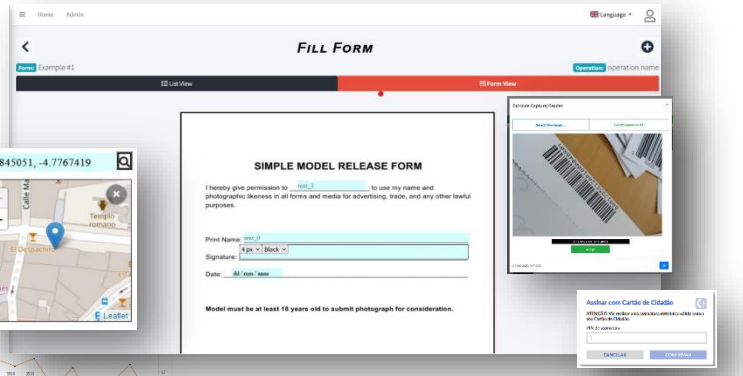


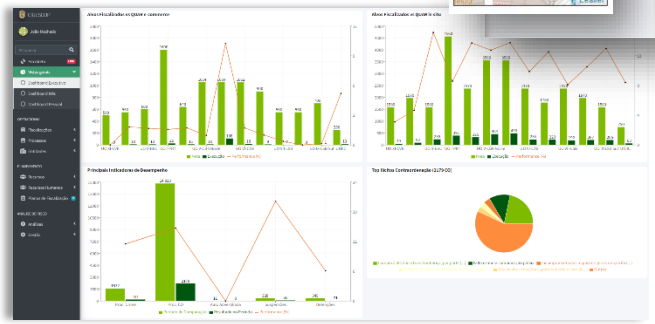


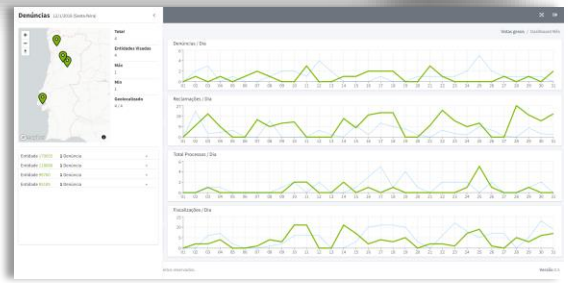
CIGESCOP – ASAE Intelligent Control Center

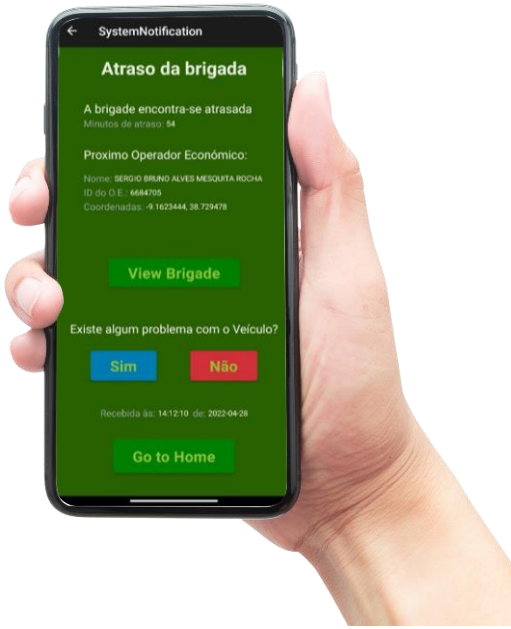
- Real-time Virtual Inspection Office, to optimize management and operational control
- Assisted disruption management
- Flexible global risk matrices
- Smart forms with filling assistant
- Costumizable dashboards that adapt to users' preferences





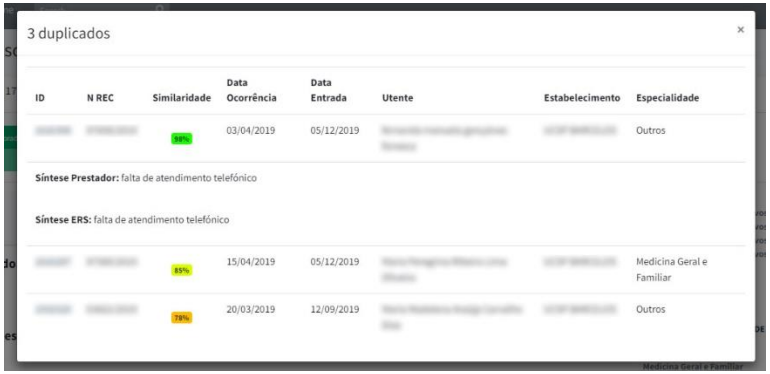


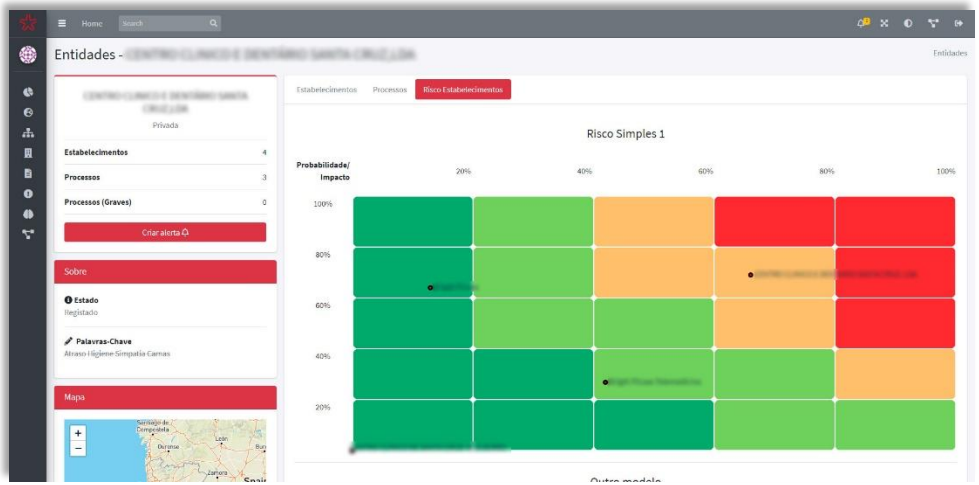




ERS AI – Health Regulatory Agency

- **Electronic complaints automated analysis and classification, as well as severity assessment**
- **Duplicate and related complaints detection**
- **Intelligent suggestions of the complaint's outcome**

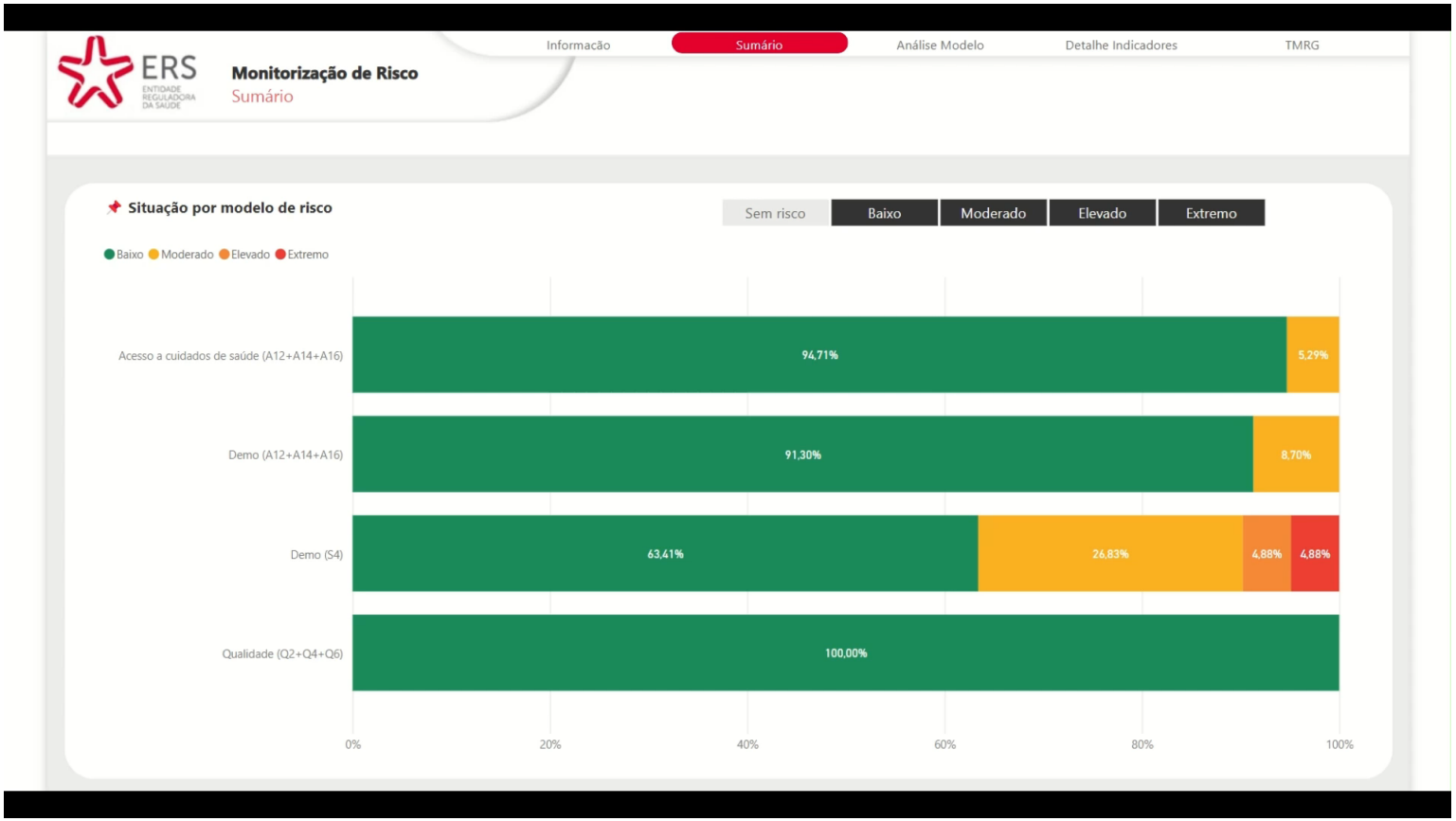




ERS
ENTIDADE REGULADORA DA SAÚDE

ERS AI 2 – Risk Based Supervision

- **Solution that automates the risk-based supervision model of health establishments**
- **Supervision based on risk analysis using a ML model**



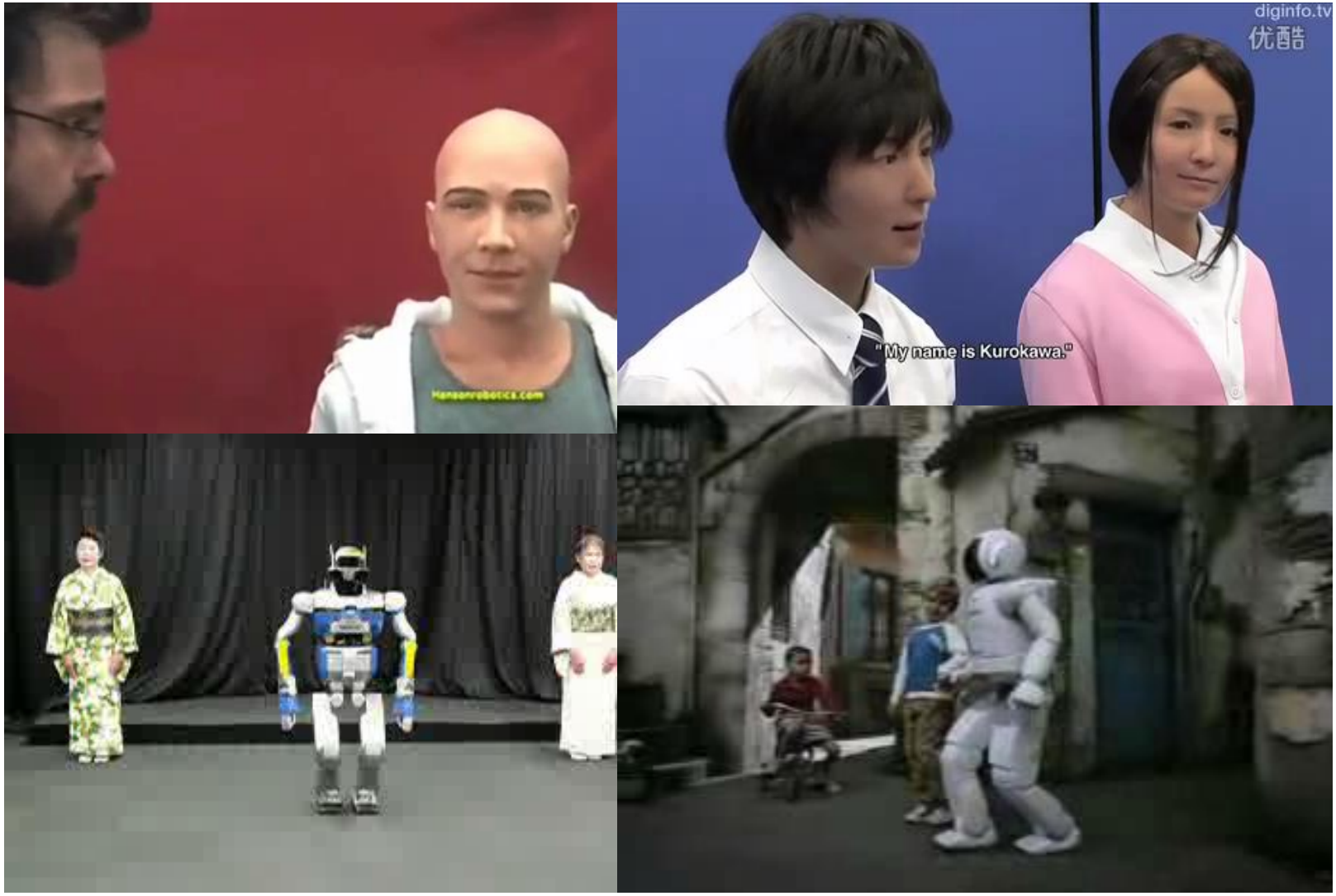
Strong AI – Star Wars and C3PO

Video: Lucasfilm Ltd. (1977). Star Wars: Episode IV - A New Hope. [Film]. 20th Century Fox.

Robotics

- Autonomous driving car (Google)
- Service, mars explor., medical robotics (Motorman, Miimo, Roomba, Oz, Asimo, Nao)
- Exosekeleton (exoAtlete)
- Ambient Assisted Living
- Drones & Delivery (PT ConnectRobotics)
- Military, Assistive, Eldery, …
- Education, entertainment, …

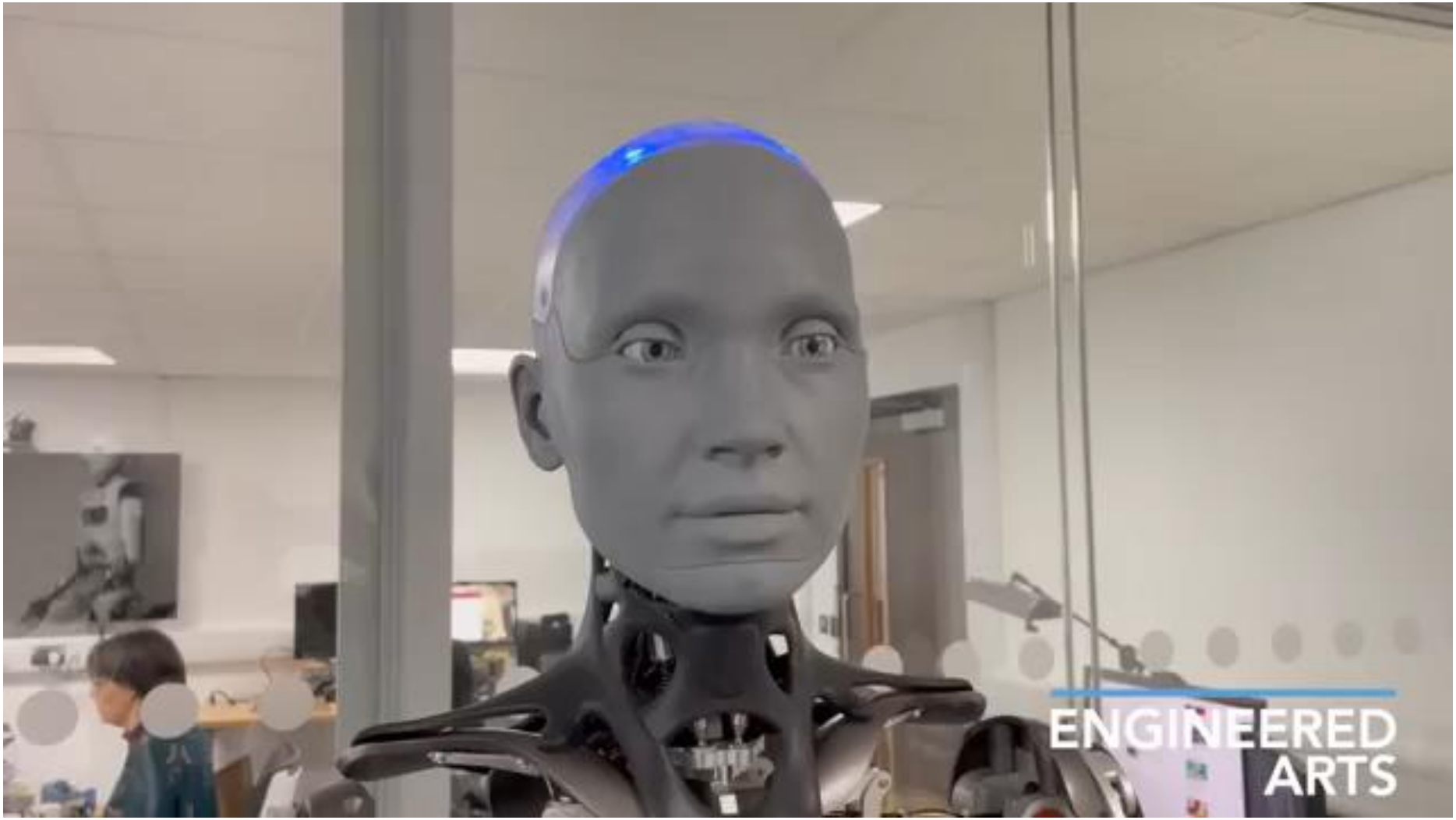
Humanoid Robotics



Robotics – Atlas

Video: Boston Dynamics: Atlas Gets a Grip - https://www.youtube.com/watch?v=-e1_QhJ1EhQ&ab_channel=BostonDynamics

Robotics – Ameca with GPT



Video: Engineered Arts: Ameca expressions with GPT3 /4 - https://www.youtube.com/watch?v=yUszJyS3d7A&ab_channel=EngineeredArts

Deep Reinforcement Learning (DRL)

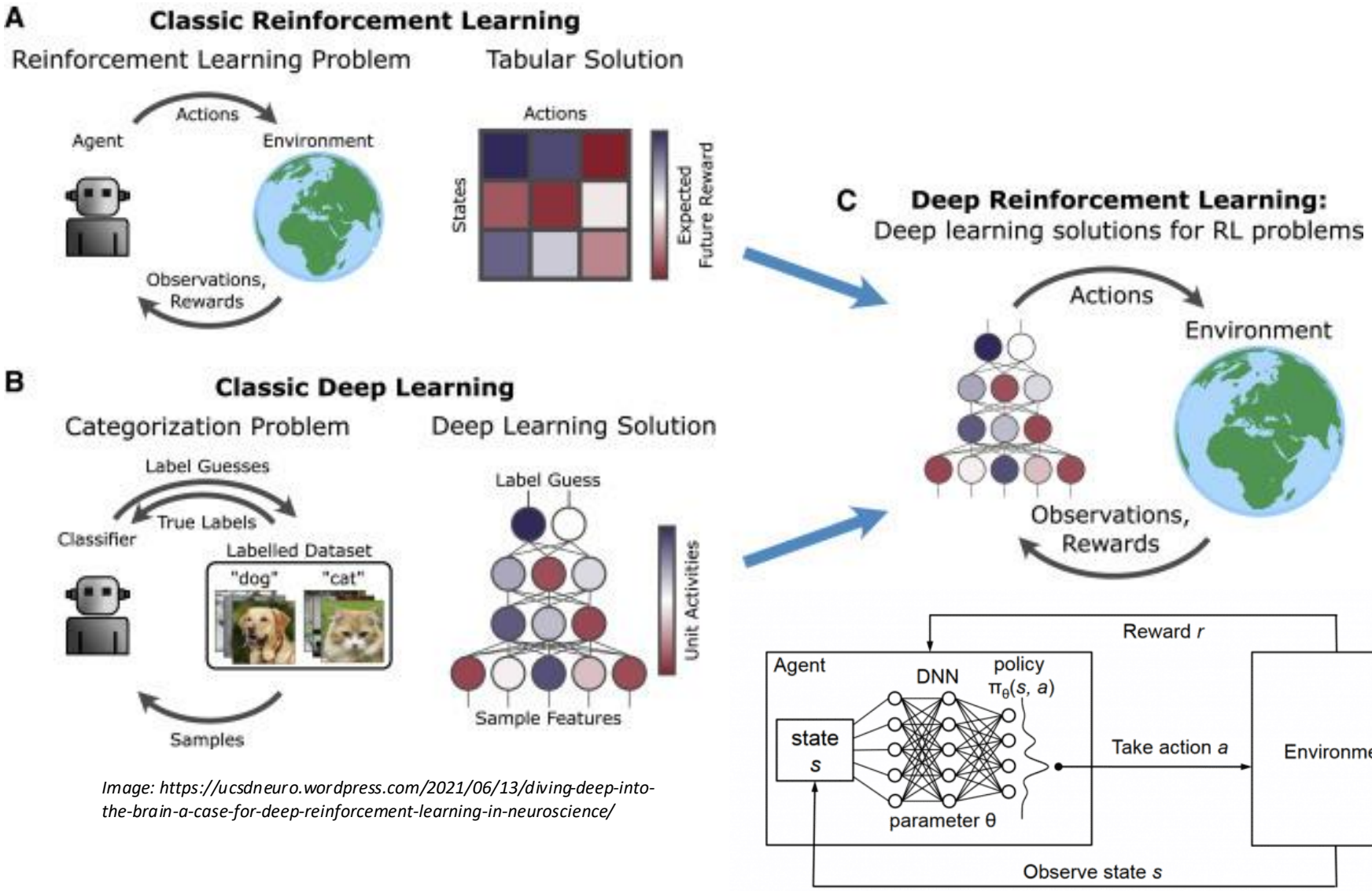
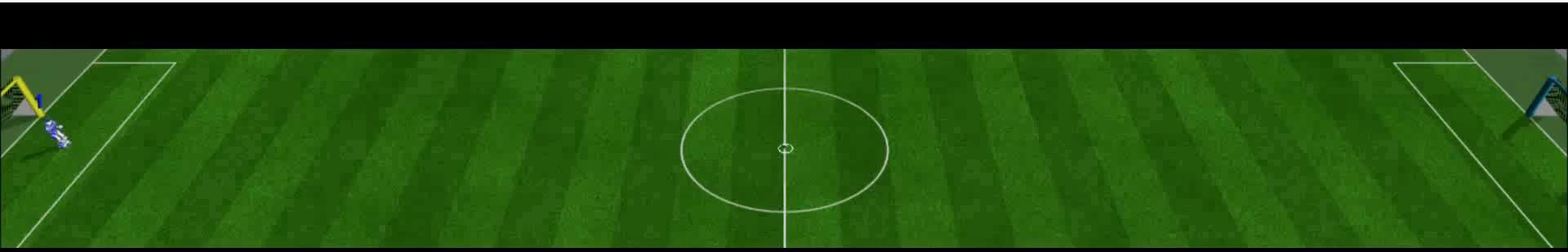
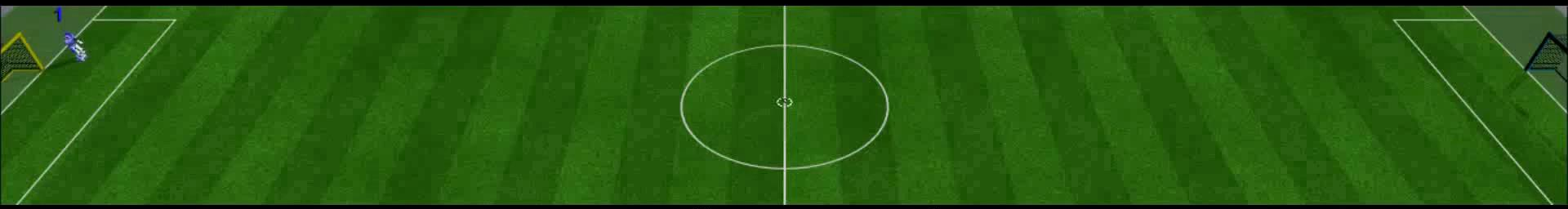


Image: https://ucsdneuro.wordpress.com/2021/06/13/diving-deep-into-the-brain-a-case-for-deep-reinforcement-learning-in-neuroscience/

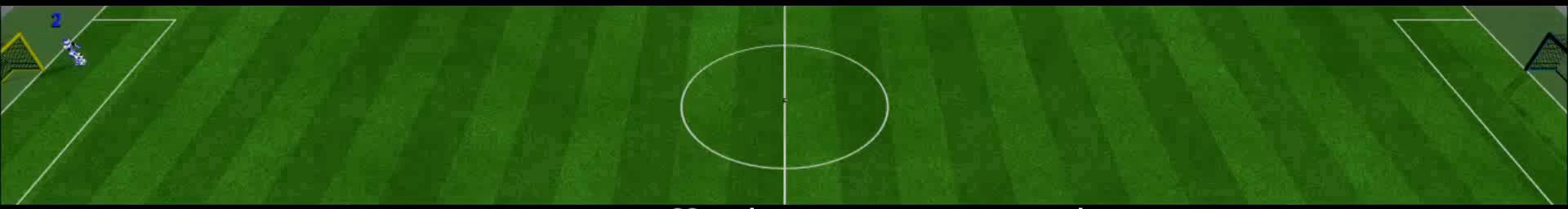
DRL for Learning to Sprint



(Our Approach) FCPortugal



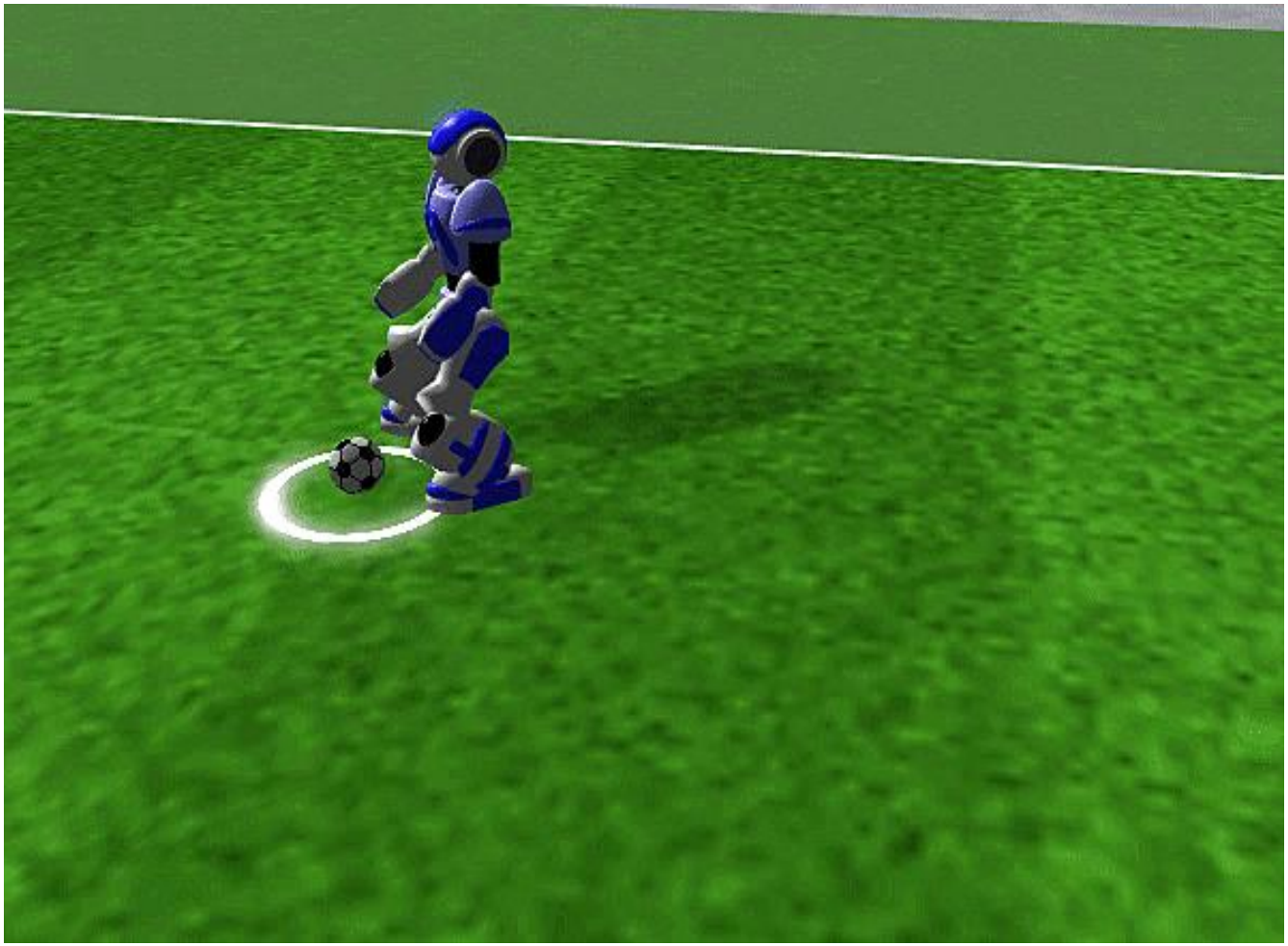
UT Austin Villa (3DSSL Champion)



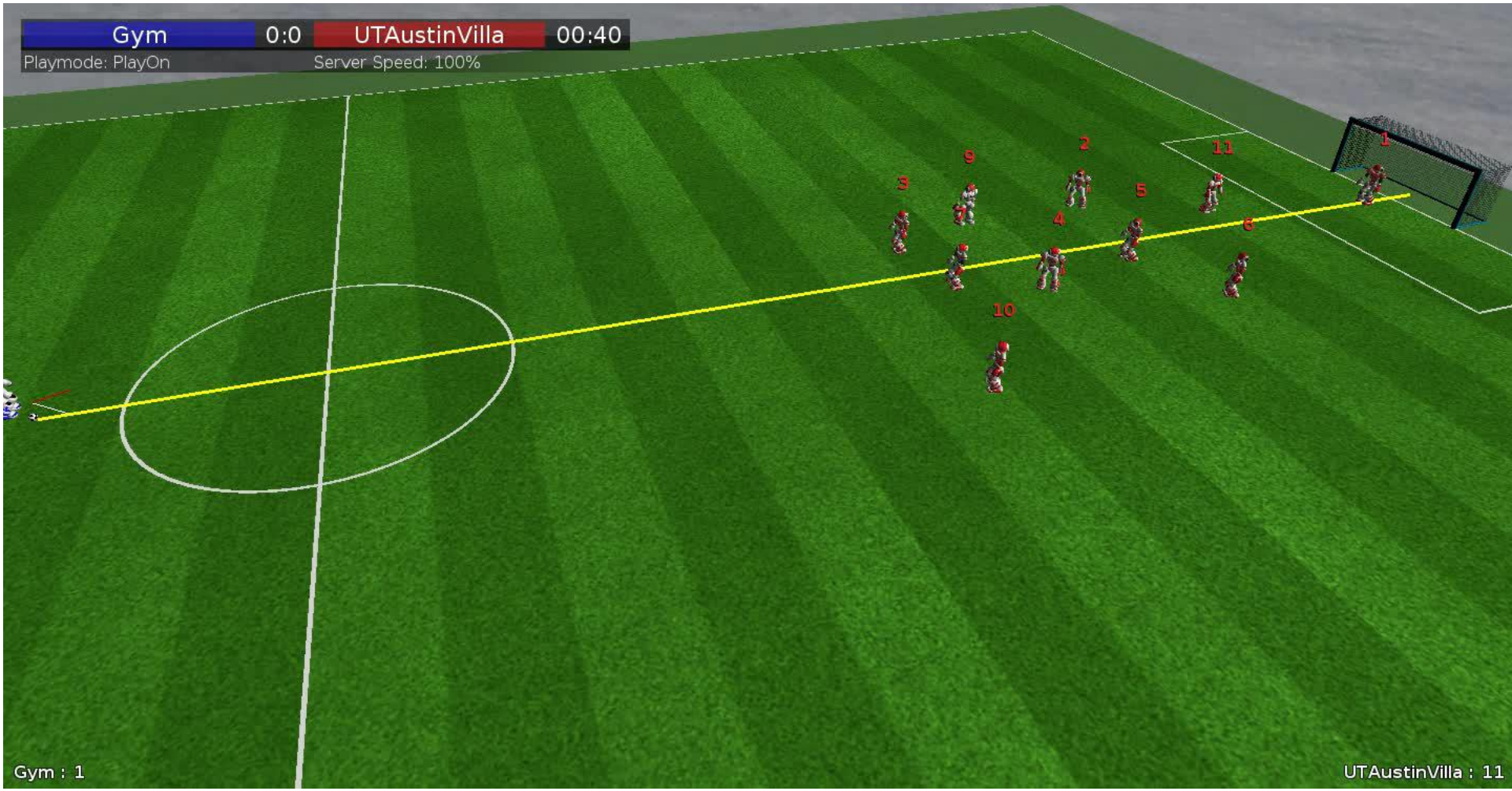
magmaOffenburg (3DSSL 2nd)

DRL for Learning to Sprint

DRL for Learning to Dribble



DRL for Learning to Dribble



https://www.youtube.com/channel/UCpMr8nYaN_EuBDRl0m9drGQ/videos

Conclusions

- **AI and the New Discovery of the New World**
- **AI is growing** steadily worldwide with emphasis for **Machine Learning, NLP** and Intelligent **Robotics**
- **AI** is probably the **hottest Research Topic** in the world!
- **Strong AI** is possible, and we are getting close to it!
- **Impossible to prevent** the use of **AI Tools in All Areas of Society and Work**
- **AI must be used with care**: **Explainability,** Transparency, Accuracy, Fairness, Integrity, Responsibility, **Privacy,** Trust **Bias&Discrimination**, Security, **Overreliance** and **Ethical**
- AI offers **huge potential** in the areas of **Public Administration** and **Food Safety**
- **New AI/GenAI and DRL powered and trained Robots**

AI Opportunities and Challenges in Public Administration and Food Safety

Luís Paulo Reis

lpreis@fe.up.pt

Associate Professor at FEUP - Faculty of Engineering of the University of Porto
Director of LIACC – Artificial Intelligence and Computer Science Lab., UPorto
Member of Coord. Commission of LASI – Intelligent Systems Associate Laboratory
President of the GA of APPIA – Portuguese Association for Artificial Intelligence