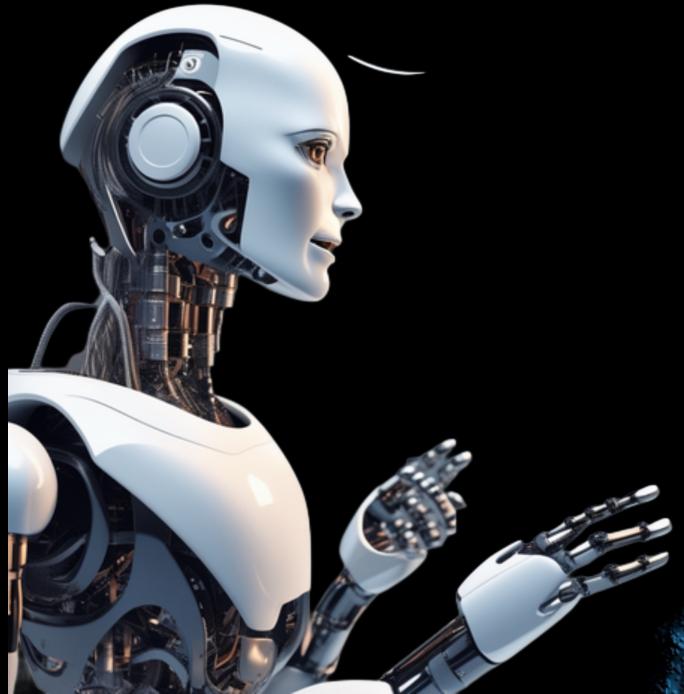


Current Job Market in Data Science!



Shovon Sengupta
~ Principal Data Scientist @Fidelity Investments

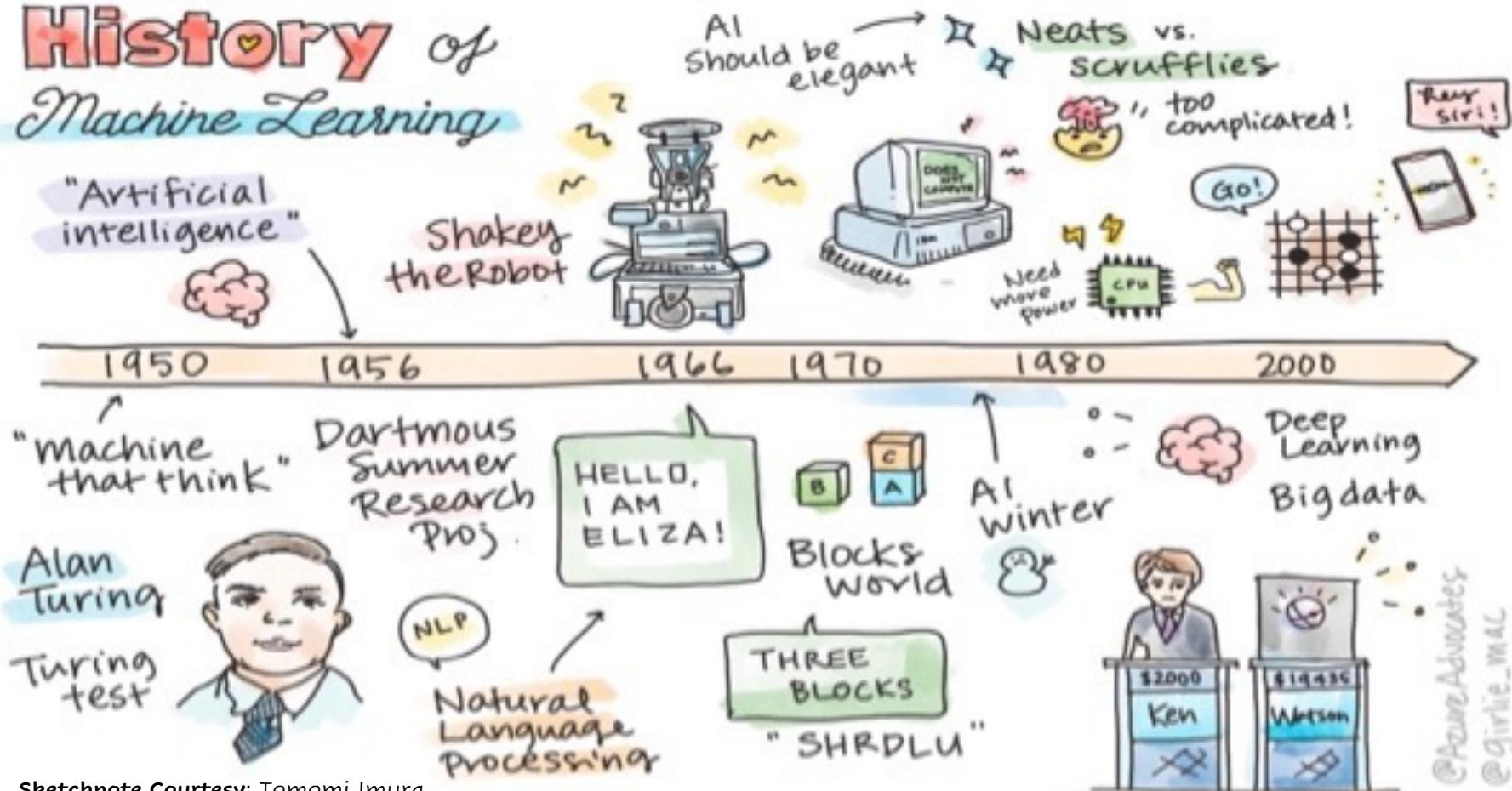
Who am I!

- ❖ Principal Data Scientist (Director) at AI CoE!
- ❖ 14+ years of experience in Applied Data Science across **Fintech, Pharmaceuticals, Retail industries**
- ❖ Couple of US – patents. And hopefully few – more in line!
- ❖ Presenter at Various International conferences on **Trustworthy ML, Financial forecasting, Graph Machine Learning, RL and its applications** in the realm of Finance!
- ❖ Published & presented at various prestigious conferences like **KDD'23, EMNLP'22, Ai42023, ISF'22/23, LREC'22**
- ❖ Authored a few (not so famous!) online courses for **Packt** and Ec-council/**redcode**
- ❖ A repeat offender – Fidelity awarded 22 times in last 7.5 years!

01

The world of Machine Learning!

History of Machine Learning



A Brief History of Machine Learning

What is Machine Learning?

"Field of study that gives computer the ability to learn without being explicitly programmed" – Arthur Samuel (1959)

AI can be broadly perceived as "**the study of machines that exhibits traits associated with a human mind such as perception, learning, reasoning, planning and problem solving**" – Kush R. Varshney ["Trustworthy Machine Learning"]

Waves of AI!

Classical AI (waves 1 &2)

Expert systems, rule-based systems, general problem solvers, heuristic search, knowledge representation, planning, **scheduling, reasoning about uncertainty**, causal reasoning, machine vision, natural language understanding, Pattern Recognition, **Machine Learning**

Predictive AI (waves 2+ and 3)

Empirical/Data-Driven methods - predictive analytics, decision trees, Bayesian Models, and many classification and regression methods in **data mining/data science**
Optimization-driven approaches – Linear & Integer programming, intensive search, advanced satisfiability search and optimization, **Deep Learning**

Generative AI (wave 3)

The generation of “content” (text, images, voice, video, etc.) from a textual or multimodal prompt
Analyzes and learns from data to create (“generate”) something new - data, images, sounds, or other types of information.
Recent buzz around generative AI driven by the simplicity of new user interfaces for creating high-quality text, graphics, and videos within seconds (e.g. ChatGPT)

Relevance of AI in Fintech

Scale is a must – human processing is not scalable or feasible

- AI algorithms for “understanding” – of context and customer
 - Leverage human judgement in delivery to build the right training data sets and KB
 - Need to make sure AI algos are subject to Responsible AI criteria (often overlooked)
- Complexity of products requires “reasoning”: Still a big challenge in AI



New capabilities tech/network require automation

- Novel risk and credit scoring opportunities
- Integration into micro-services
- Leverage networks (e.g. social, economic, commerce networks) and other viral services

Why is AI needed in Fintech?

Much of finance digitization has focused on workflows, but little advance on customer experience and interactions



“Front Office”

Need to inject some of the human intelligence/understanding in customer interactions to achieve the right *customer experience*

- Understanding customer context
- Understanding intent, challenges, and issues
- Reduce costs of customer service operations (chatbots, problem understanding, problem resolution)
- Next-best-action, X-sell, up-sell, reactivation, etc... at scale



“Back Office”

Still expensive and complex and needs intelligent automation

- Compliance
- Financial Crime
- “Operations” – e.g. forms into actions



“Investment Banking”

- Market data operations
- Modelling micro and macro trends based on document injection - news, reports, analysts
- Market research and report generation and updates

Applications of AI in Financial Services



Credit and Risk

- ✓ Underwriting & pricing model optimization
- ✓ Dynamic Credit limit
- ✓ Risk Modeling and Scenario analysis



AI & Trading

- ✓ Data Analysis and trading support
- ✓ Portfolio optimization
- ✓ Modeling and Scenario Analysis



Customer Service and Personalization

- ✓ Dynamic Customer Profiling & segmentation
- ✓ Personalized content, product/service recommendations
- ✓ Intelligent Call routing
- ✓ Agent/chatbot support



Process Automation

- ✓ Data Extraction
- ✓ Document Automation
- ✓ Digitization of Service
- ✓ Document and agreement review
- ✓ Document Q/A



Fraud detection & prevention

- ✓ Credit card fraud
- ✓ Insurance claim fraud
- ✓ Money Laundering



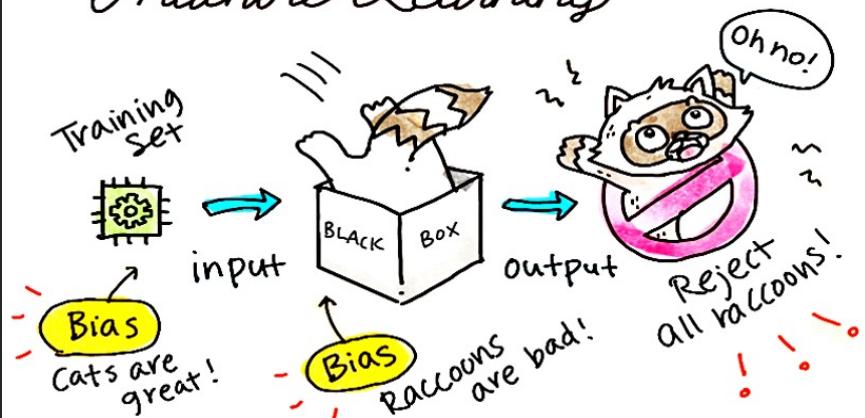
Trustworthy AI

- ✓ AI Ethics & governance
- ✓ Technical AI audit
- ✓ Responsible AI training

02

Fairness in Machine Learning!

Fairness in Machine Learning



Complex sociotechnical challenges

Purely social
Purely technical

@Azure Advocates
@girlie_mac

Sketchnote Courtesy: Tomomi Imura

Fairness-related harms

Unfairness = negative impacts for group of people such as those defined in terms of

- race • age
- gender • disability status

Harms:

- ★ Allocation
- ★ Quality of service
- ★ Stereotyping
- ★ Denigration
- ★ Over- / under-representation

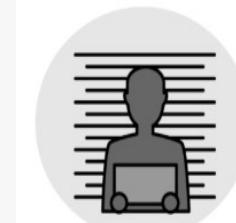


Assessment & mitigation

False negatives
False positives

- ❖ Identify the harm (+ benefits)
- ❖ Identify the affected groups
- ❖ Define fairness metrics

	False-Pos.	False-Neg.	Counts
men	0.35	0.27	6239
women	0.29	0.35	3124

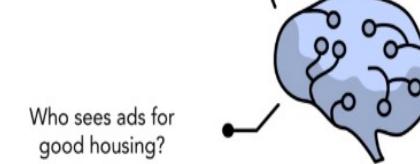


Who is likely to commit another crime?

Machine learning creates bias



Who should be eligible for same-day delivery?



Who sees ads for good housing?



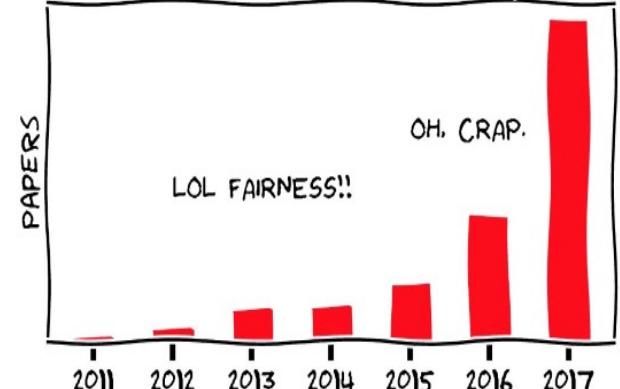
when it's tasked with answering questions like...



Who hears about career opportunities in STEM?

Examples of how bias in ML models can affect our daily life!

BRIEF HISTORY OF FAIRNESS IN ML



publications on fairness from 2011 to 2017

Gen AI and its relevance!

Artificial Intelligence

Programs with the ability to simulate human intelligence

Machine Learning

Programs with the ability to learn without being explicitly programmed

Generative Models

Programs with the ability to learn how to generate new data that is similar to a given set of training data

Power of Generative AI stems from



Self-Supervised Learning

Learning from unlabeled data without explicit supervision or labeling.



Deep Learning

Neural Networks that enable automatic feature extraction by modeling complex distributions from large amounts of data.



Transformer Architecture

Neural Network that leverage **self-attention** to learn context by tracking relationships in the sequence of data.

Gen AI : Examples!

Generative AI for Text

- ✓ ChatGPT (OpenAI)
- ✓ BARD (Google)
- ✓ LLaMa (Meta)

Generative AI for Images

- ✓ DALL-E (Open AI)
- ✓ Stable Diffusion (Stability AI)
- ✓ Midjourney (idem)

Generative AI for short videos

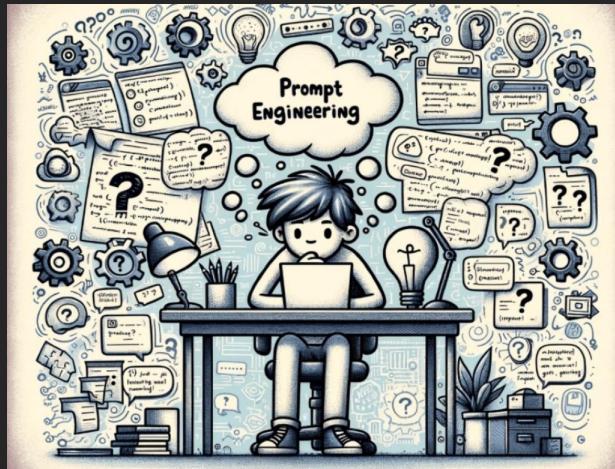
- ✓ Runway
- ✓ Midjourney (with a parameter)

Generative AI for Code

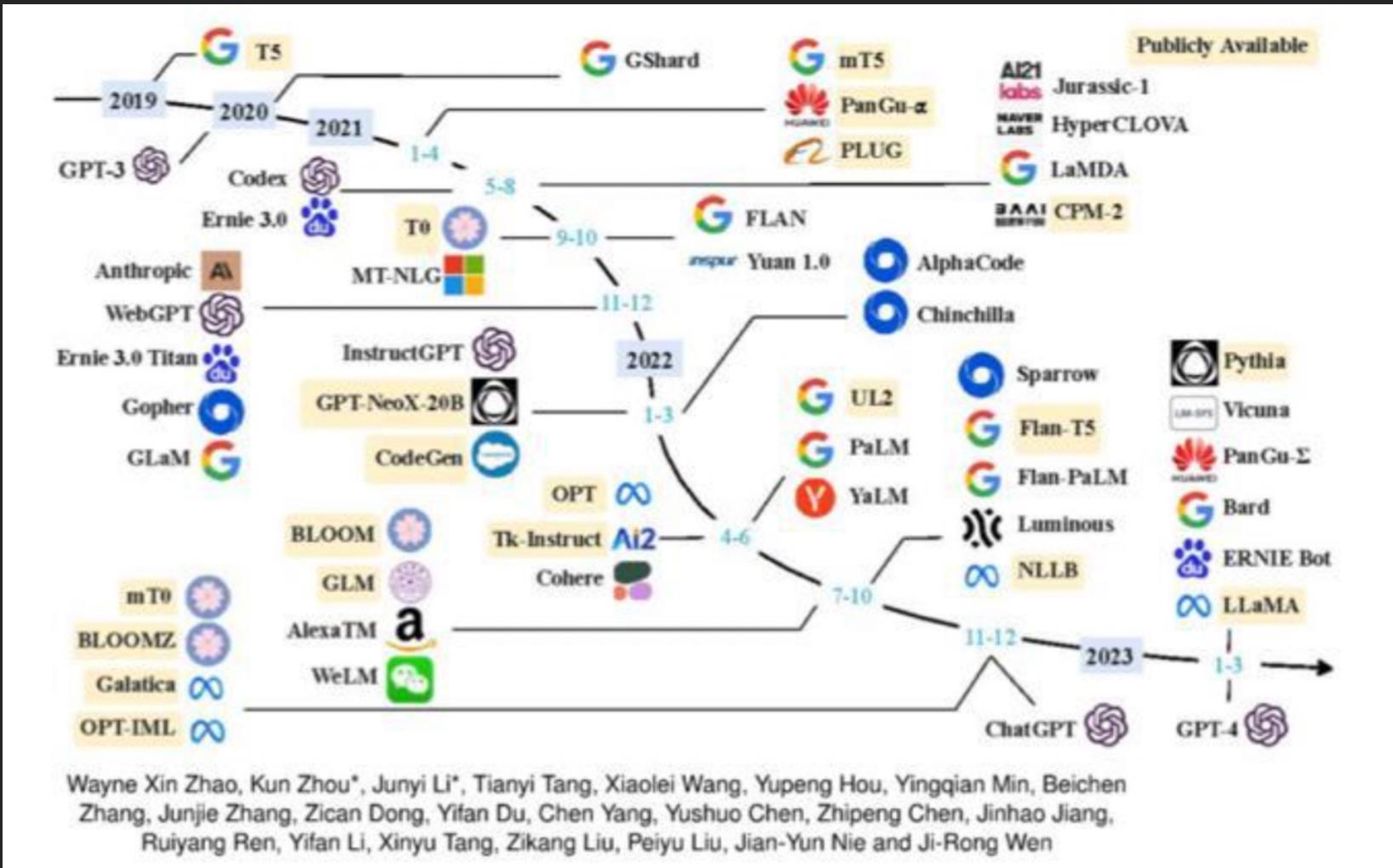
- ✓ Codex (Open AI)
- ✓ Microsoft CoPilot (used Codex for Github)

Generative AI for Voice & Music

- ✓ Aiva.ai
- ✓ Endel
- ✓ Lyricstudio.net



A survey of LLMs!



Where LLM Shines!

The New York Times

The Brilliance and Weirdness of ChatGPT

A new chatbot from OpenAI is inspiring awe, fear, stunts and attempts to circumvent its guardrails.



We made ChatGPT write a song for us

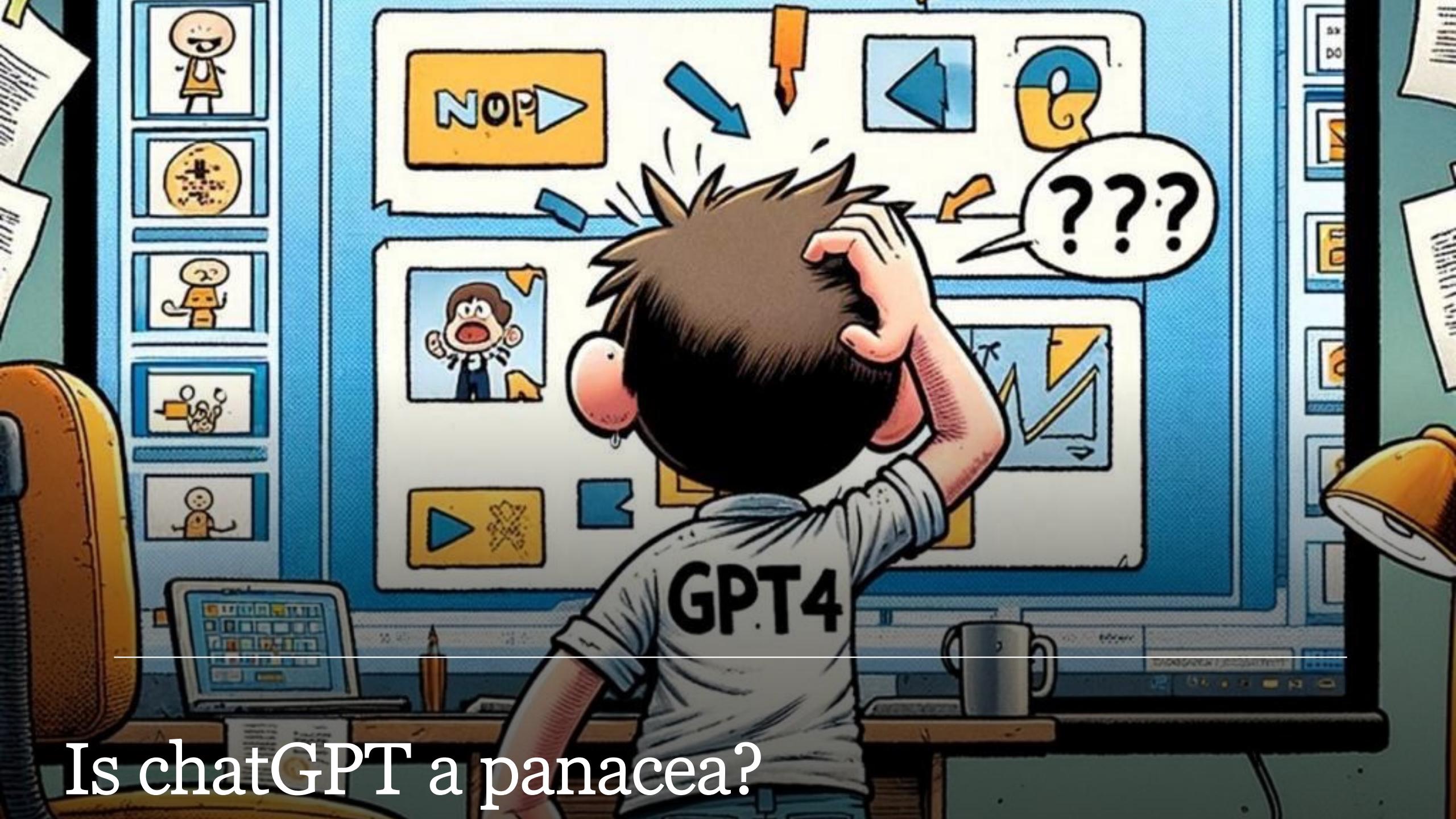
THE WALL STREET JOURNAL.

The Jobs Most Exposed to ChatGPT

New study finds that AI tools could more quickly handle at least half of the tasks that auditors, interpreters and writers do now

The
Economist

Is GPT-4 the dawn of true artificial intelligence?



Is chatGPT a panacea?



Thank You!