

INDIAN CASE CHALLENGE 2024

PRE - CASE

Unravelling of OpenAI

The saga kicked off on the 17th of November when the Board of Directors abruptly terminated CEO and co-founder Sam Altman during a video call that Altman assumed would be an uneventful check-in. This bombshell and tragic news became public rapidly, with Microsoft discovering it around the same time as the entire world.

OpenAI has emerged as Silicon Valley's hottest and most valued AI startup of recent years, thanks to the viral user growth of its ChatGPT chatbot. Propelled by ChatGPT's runaway popularity, OpenAI now boasts over 100 million weekly active users interacting with its flagship conversational agent.

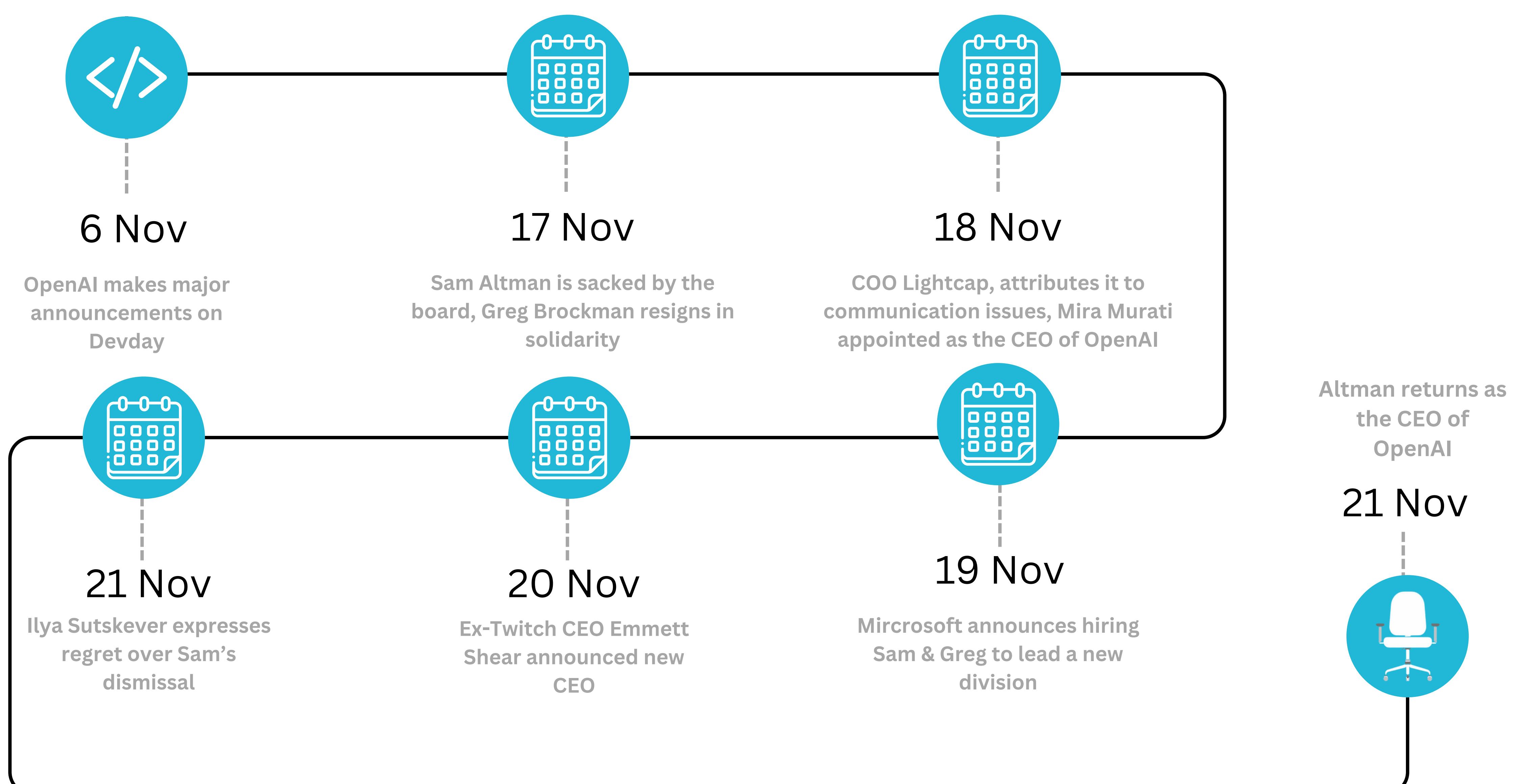
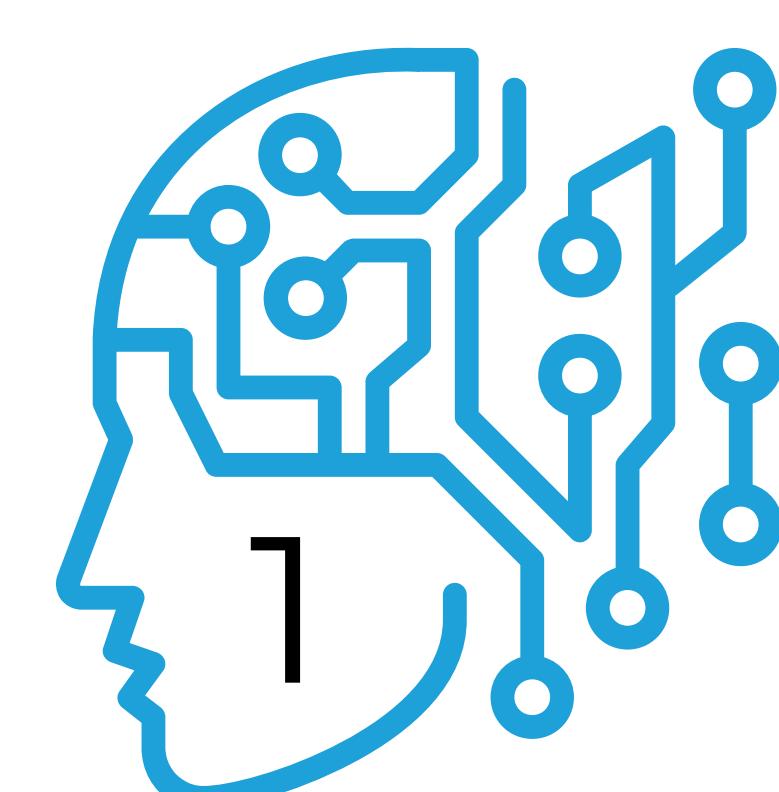
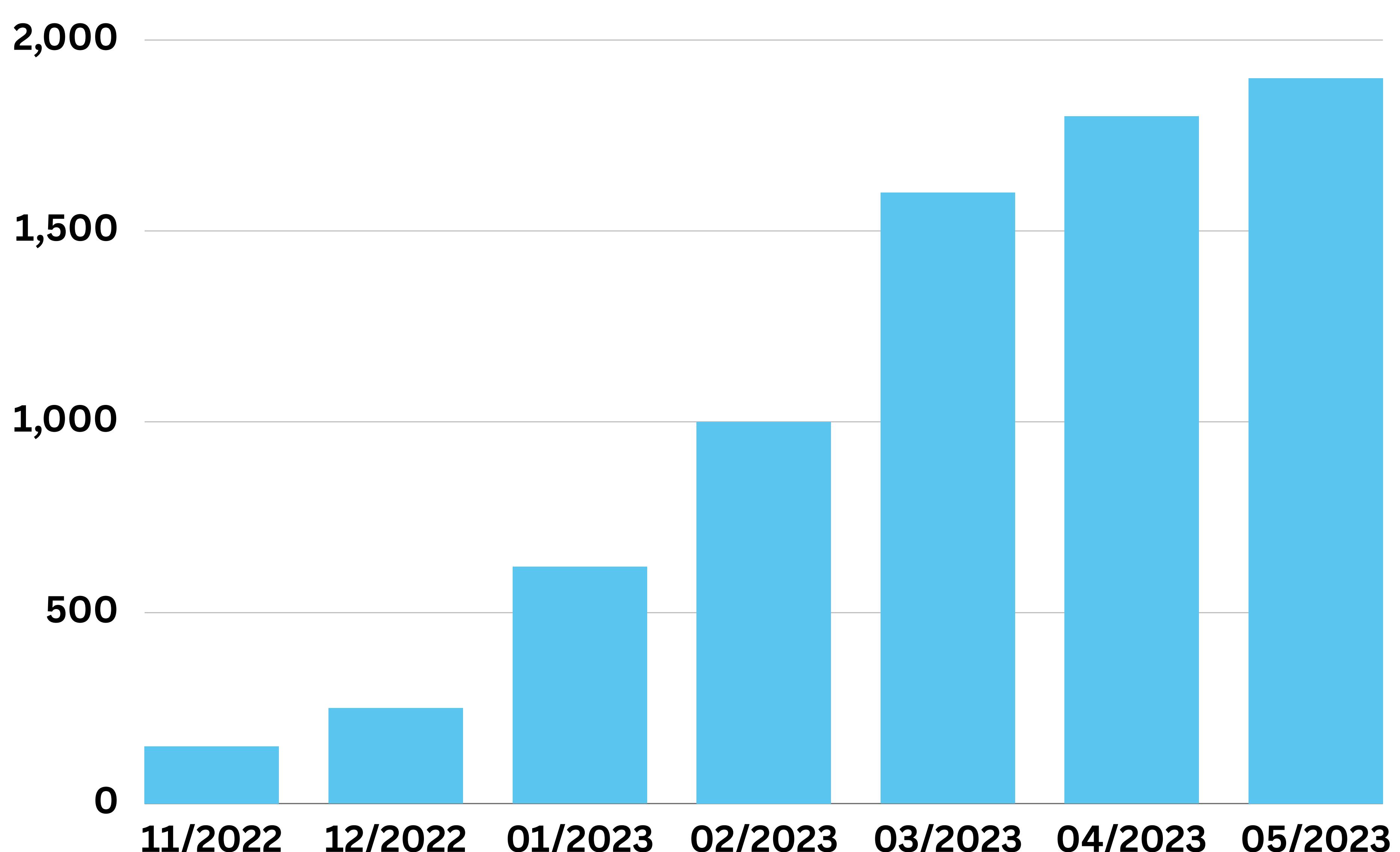


Exhibit 1 - Timeline of events

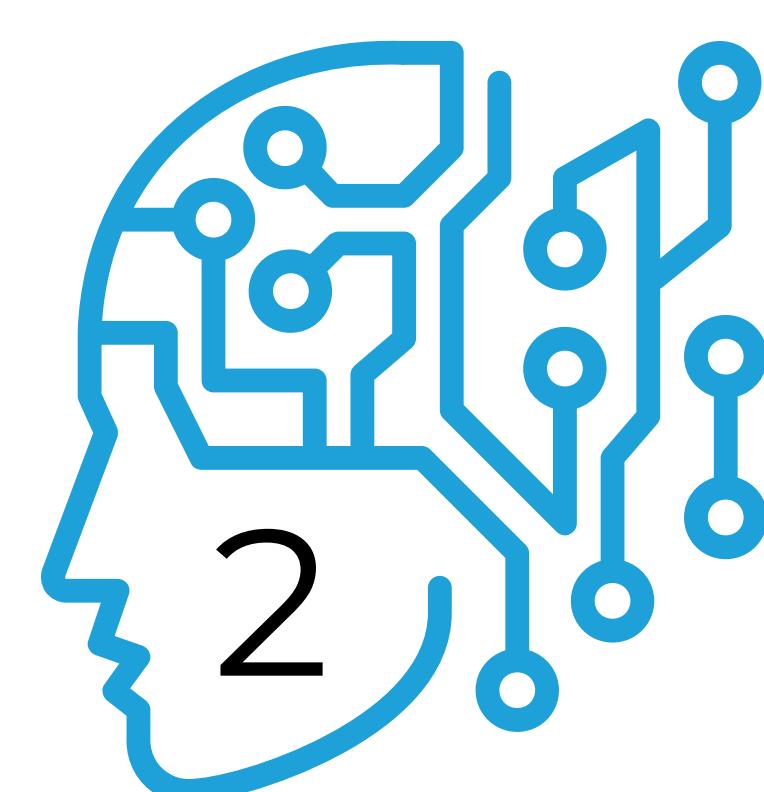


Altman's dismissal came as a massive shock across the technology industry and internally at OpenAI. In an act of solidarity, President Greg Brockman resigned on the same day that Altman was let go. Initial speculative reports suggested that a strained relationship and communication breakdown between Altman and Chief Operating Officer Brad Lightcap may have been an underlying factor behind Altman's removal.

Things seemed stable for a short time, but on November 20th, unexpected events unfolded, shaking up the situation. OpenAI revealed they had selected former Twitch CEO Emmett Shear as the company's next leader. Over 500 staffers signed a petition demanding Altman's return, or they would resign en masse. Even Chief Scientist Ilya Sutskever voiced doubts over the board's decision-making regarding Altman.



Exhibit(2) - Monthly visits to OpenAI

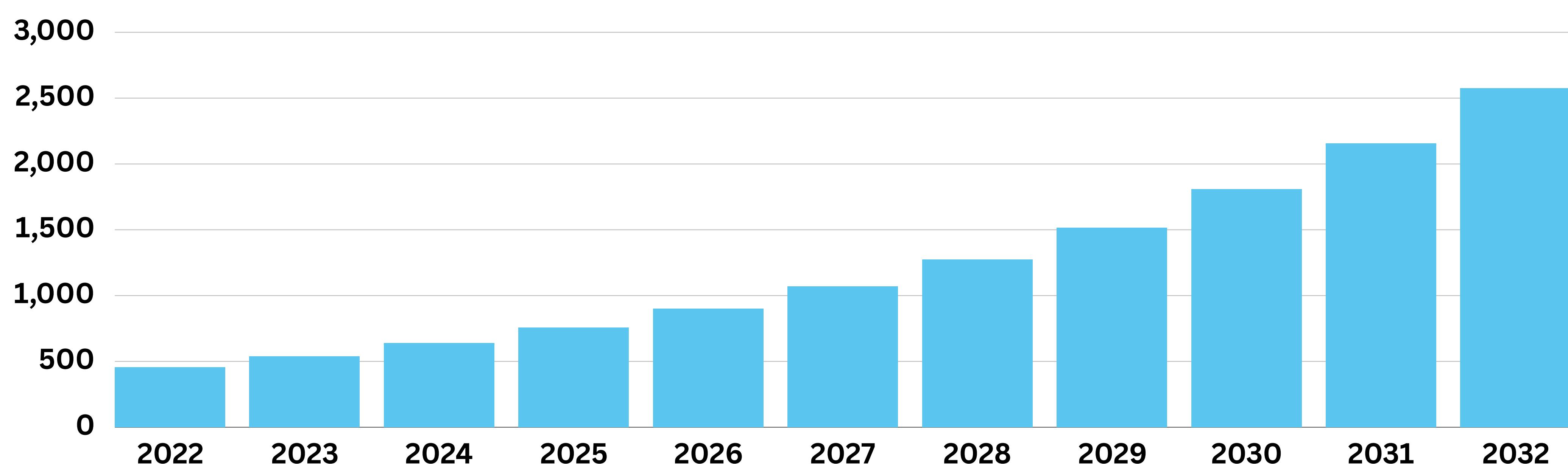


By November 21st, after internal discontent, OpenAI declared they had agreed that Altman would be reinstated as Chief Executive Officer under a partially new and reconfigured board. This dramatic reversal was completed over a week of unprecedented turmoil and volatility surrounding the company. In the end, the ordeal reaffirmed Altman as the leader.

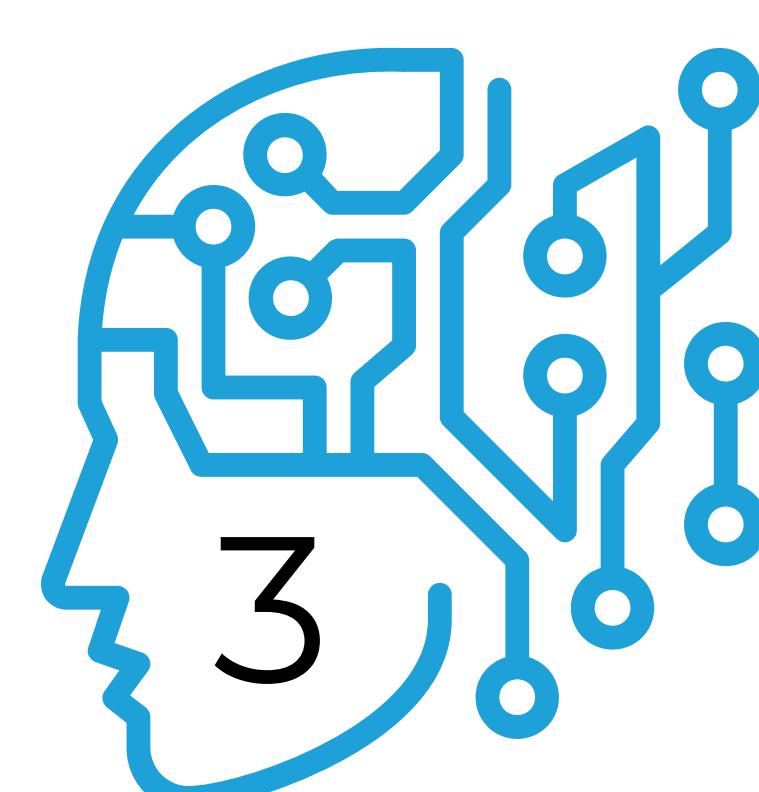
Altman's removal coincided closely with internal debates concerning the safe use of AI within the company. A report by The Information indicated internal disagreements and conflict among employees regarding the safety measures being undertaken in the development of AI. Despite the turmoil, OpenAI launched an assemblage of new features on 6th November and unveiled their latest innovations, garnering global coverage in an event called DevDay.

DevDay

"92% of Fortune 500 companies are using our solutions, some two million developers are building on top of the technologies via our API, and ChatGPT has 100 million weekly users, growing only by word of mouth", said the CEO Sam Altman, in high spirits. OpenAI's DevDay 2023 marked a significant milestone in AI; it was explicitly crafted for developers. The introduction of the Assistants API represented a pivotal tool for developers crafting sophisticated and interactive AI-driven applications.

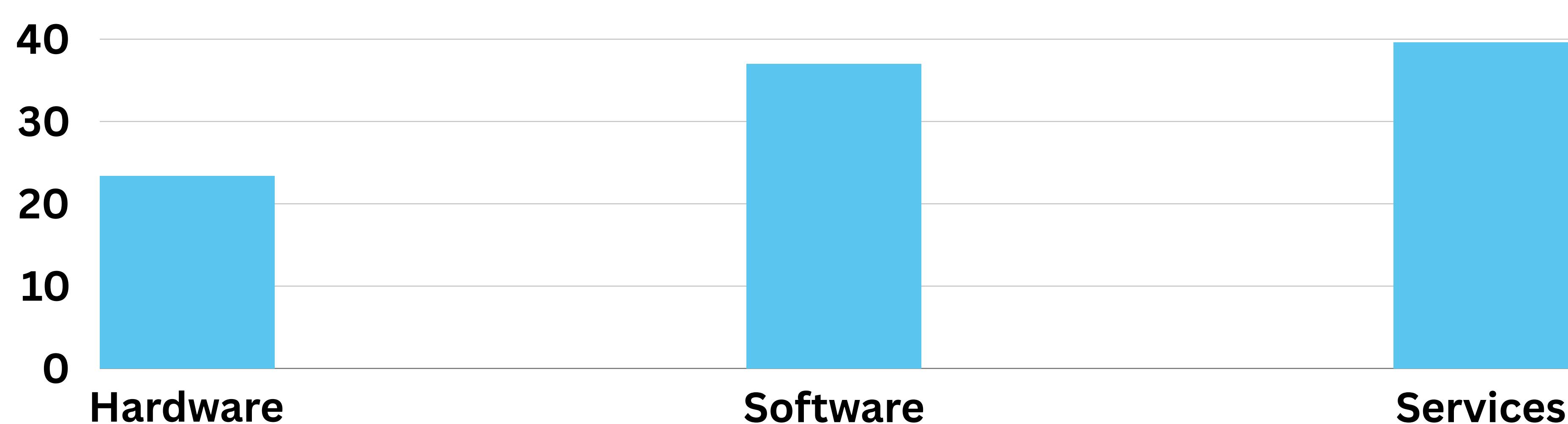


Exhibit(3) - Market size of AI industry (USD Billions)



The newly announced tool would significantly reduce development time and costs, allowing a much smoother API integration. The firm announced a more economical and efficient GPT4 Turbo and, most significantly, GPTs. With the rollout of GPTs, users could now train their custom-trained chat assistants like a GPT to plan and update their calendars in real-time. They would be available as plug-ins in the ChatGPT window and can be published for everyone to use; many of them could be built with little or no code.

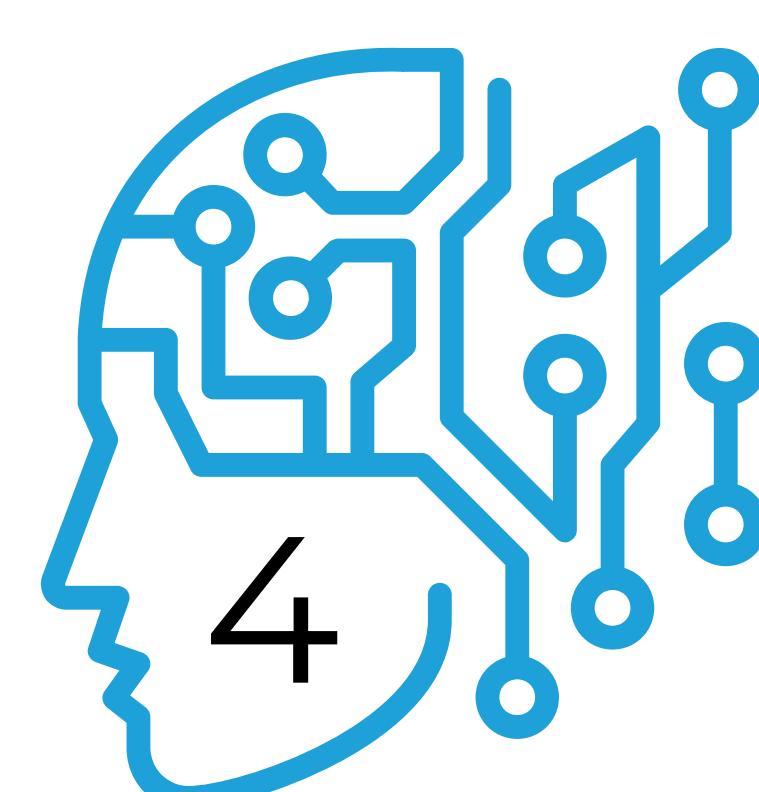
GPTs open up opportunities for innovators to build them with ease and share them with the entire user base. To facilitate this, OpenAI introduced GPT store and GPT builder, making the LLM more accessible.



Exhibit(4) - Artificial Intelligence Market Size By Solution, 2022 (%)

GPT Store

The GPT Store is an innovative, collaborative AI marketplace featuring custom GPT models. It aims to democratise access to AI models, enabling developers to utilise pre-trained models and contribute and share their models within a collaborative ecosystem. The GPT Store and GPT Builder work hand in hand to make customised AI models accessible. These AI models can be customised to fit user needs. For instance, an educator could use the GPT builder to create a model to teach complex scientific concepts. This process of building models has been made significantly straightforward, thanks to the GPT builder.



The GPT builder takes instructions in a conversational tone, enabling the user to train the model effectively. The builder thus simplifies the process of prompt engineering the model by taking the user's simple instructions, additional knowledge and functionalities. Once the user is satisfied with their model, they can publish it on the GPT store for the world to use.

Industry leaders have speculated that the GPT Store will only be accessible to paid users initially, though that may change. Subscribers will build the models through prompting and be able to push them onto the platform.

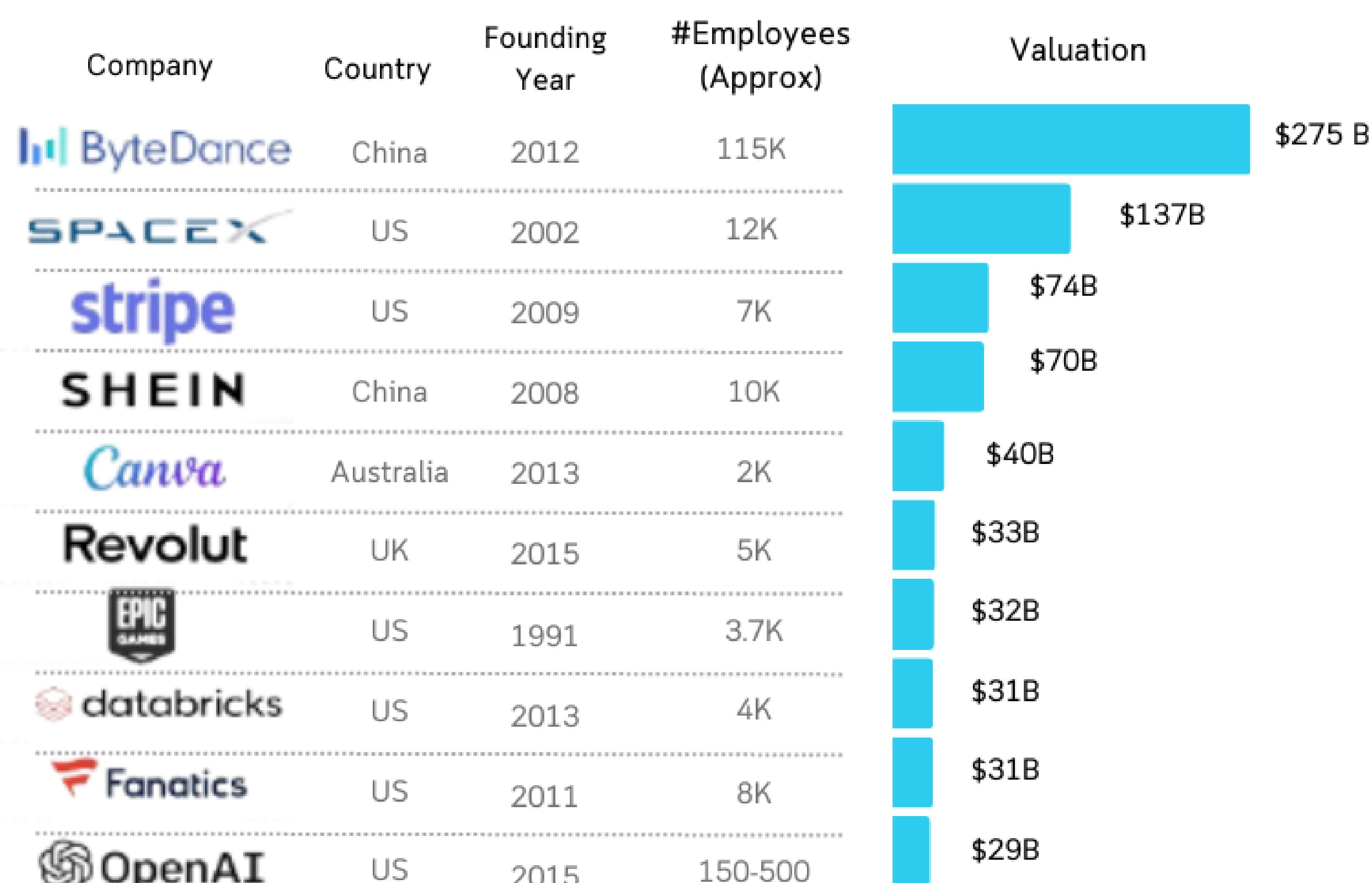
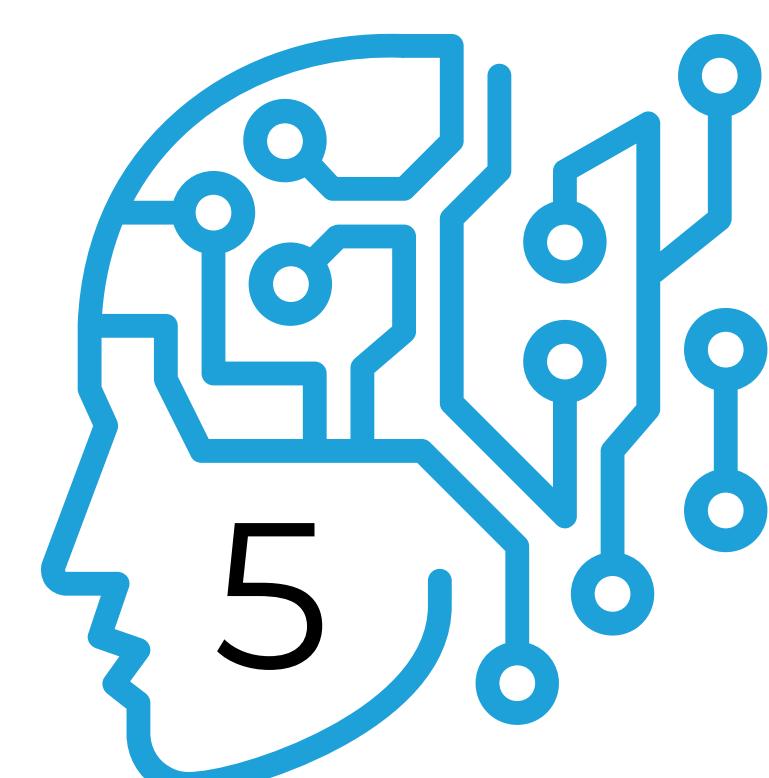
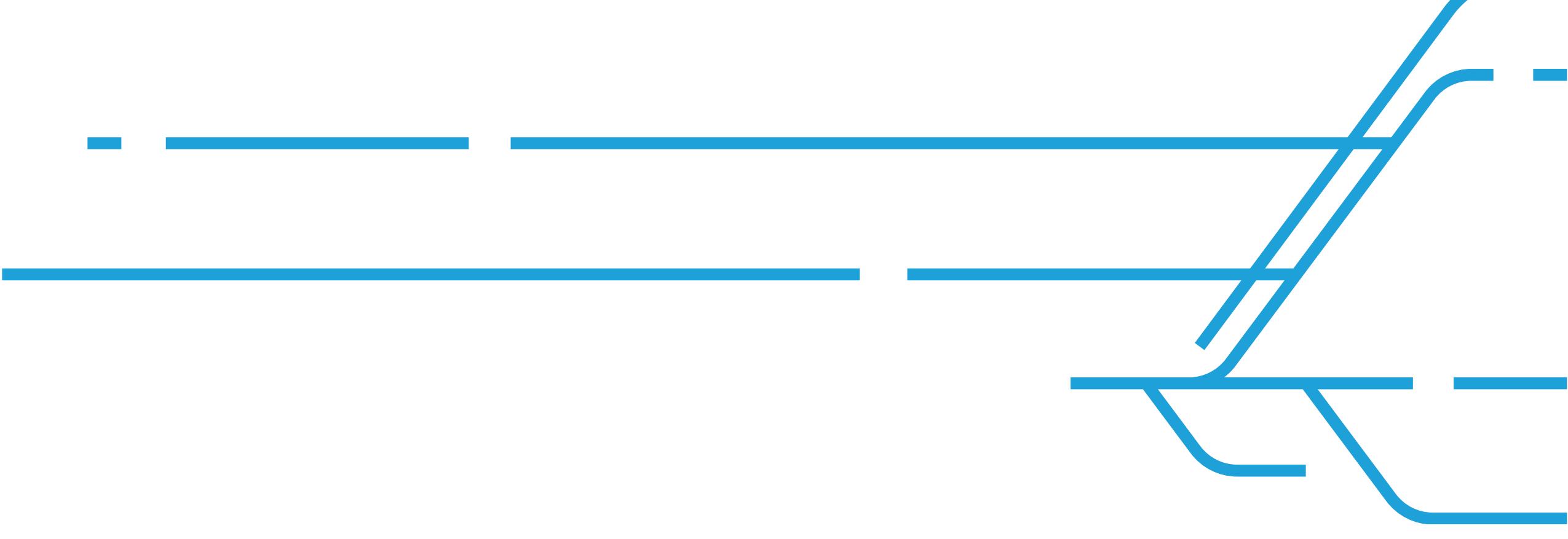


Exhibit (5) - OpenAI is now world's 10th largest Unicorn

Revenue Sharing

The GPT Store's revenue-sharing aspect is intriguing and critical. While the initial focus is on user engagement, there's room for alternative approaches. Each approach will present its unique set of circumstances to Open AI.





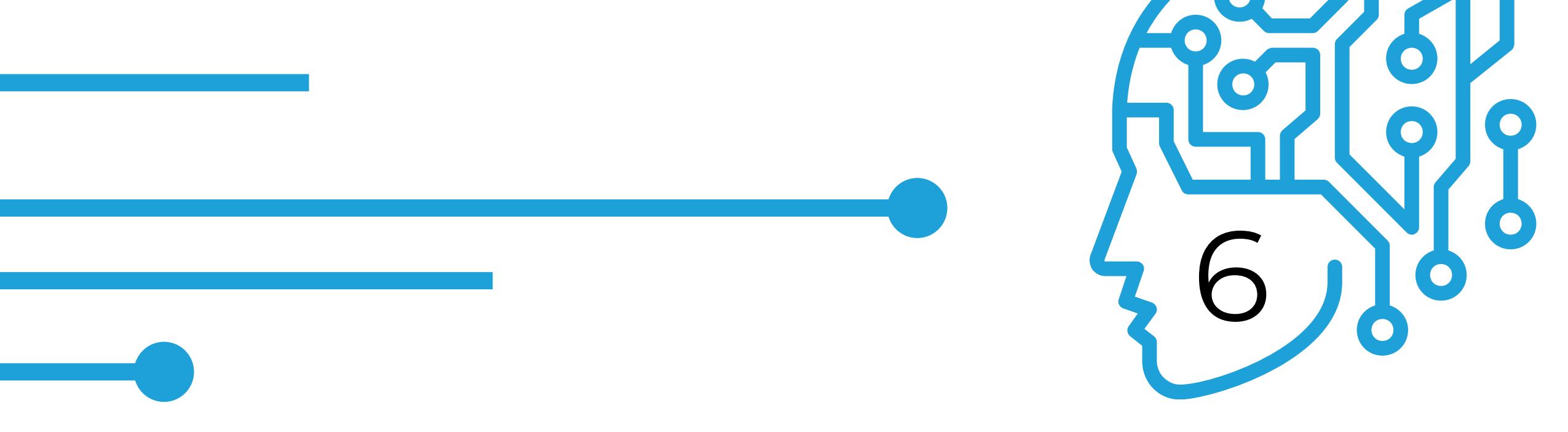
While keeping it accessible to premium users can incentivise people to buy their subscription, there's always a tradeoff as OpenAI loses its share of new subscribers who would've tuned in to try the all-new and customised models. Furthermore, this new revenue-sharing model would require a compromise between developer incentives and company profits.

Determining the revenue-sharing model that aligns with OpenAI's commitment to advancing AI development is crucial. Should access to the powerful GPT Builder be restricted to paid users, or should some level of functionality be available to everyone? OpenAI has aimed to incentivise developers by promising monetisation. However, the choice to restrict access exclusively to premium users warrants justification. While opening models to the general public might increase traffic to OpenAI, the incentive for premium subscriptions may outweigh the advantages of broader accessibility.

Potential Pitfalls

The GPT Store could also face several other problems, oversaturation being a major concern. Due to ease of creation and the limited revenue per user, the store could quickly be flooded with custom GPTs. This could cause several issues:

1. An influx of low-quality GPTs significantly hinders user experience and technology acceptance.
2. Increased competition among similar GPTs, causing financial strain for developers and taking away existing apps and SaaS solutions' user base and revenue.
3. Users might need help discovering relevant models, leading to fatigue and frustration navigating the vast array of available GPTs.



AI-Specific Chips

Another possible concern with the GPT Store is the high influx of users, which could strain the computational resources of OpenAI, which are stretched thin to begin with. Computational resources, specifically AI chips, have become important in AI development. These chips offer notable performance and efficiency gains over conventional CPUs and GPUs because they are specifically engineered to meet the demands of AI workloads.

Companies like Microsoft and OpenAI are creating specialised AI chips and are driving this revolution. Microsoft has released the Azure Maia 100 and Cobalt 100 chips, designed primarily for general-purpose computing and AI training. In the meantime, OpenAI, led by Sam Altman, is researching and developing the TIGRIS chip, which will challenge Nvidia's hegemony in the AI hardware market.

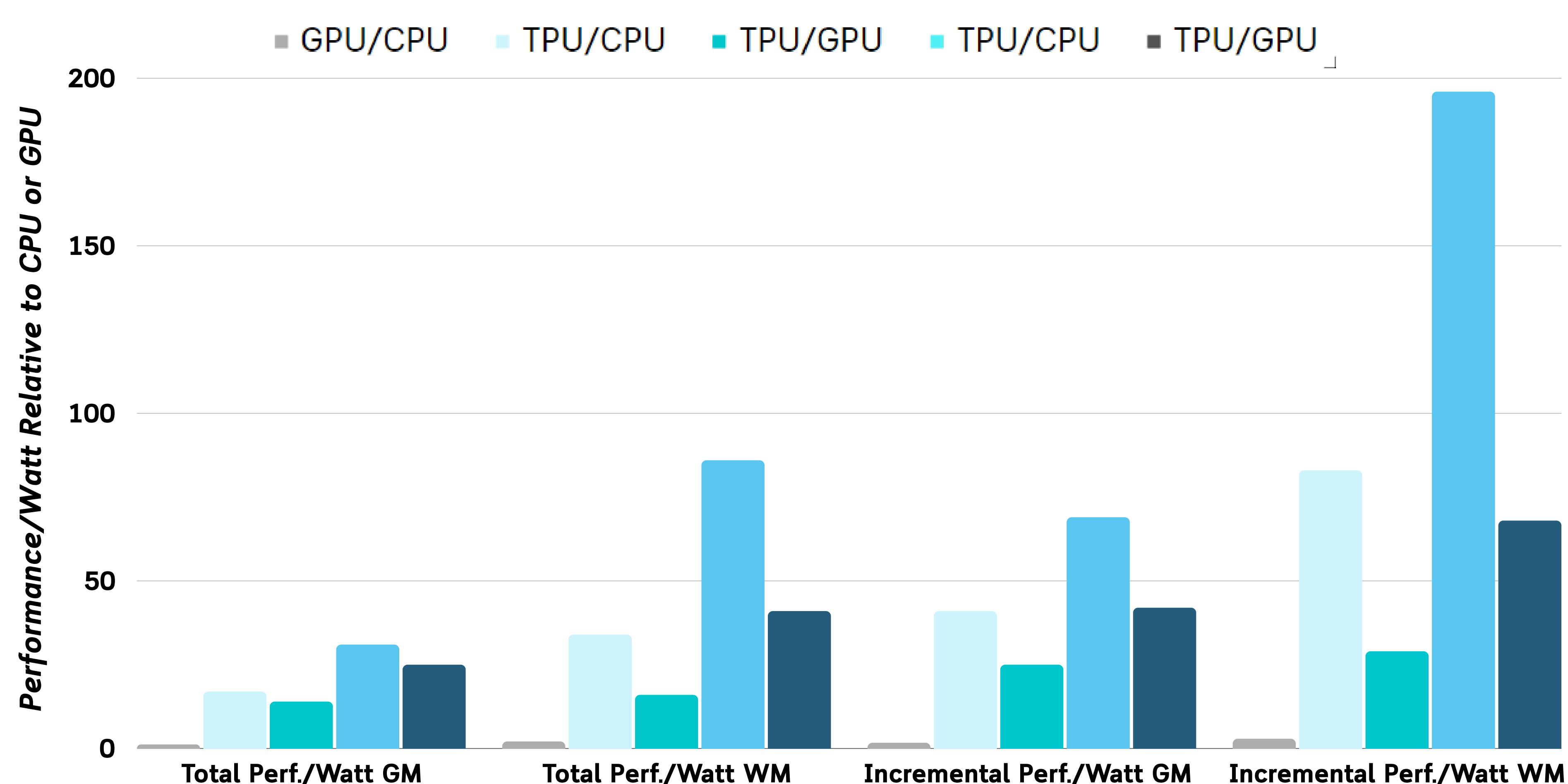
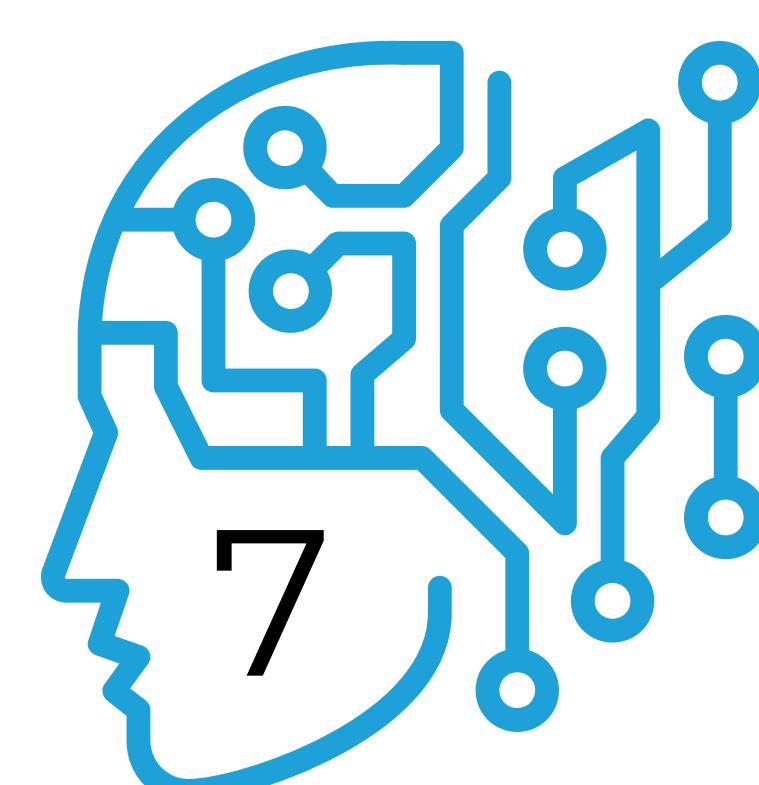


Exhibit (6) - Performance comparission of CPU/GPU



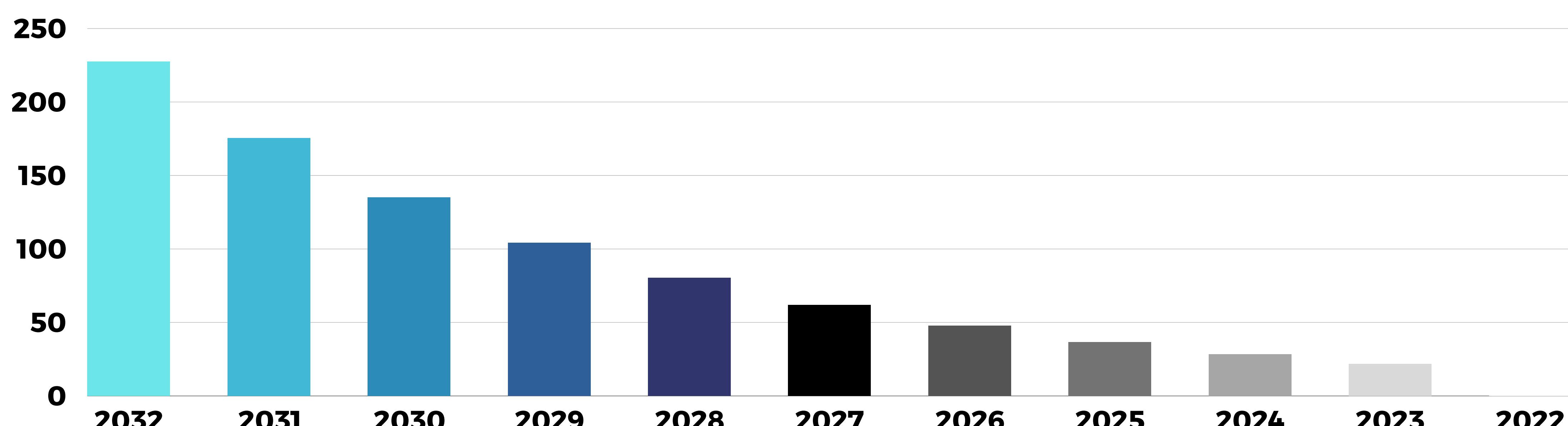


Exhibit (7) - Artificial Intelligence (AI) Chip Market Size (USD Billion)

A number of startups have also sprung up in recent times to capitalise on this relatively new industry. Few of them are given below.

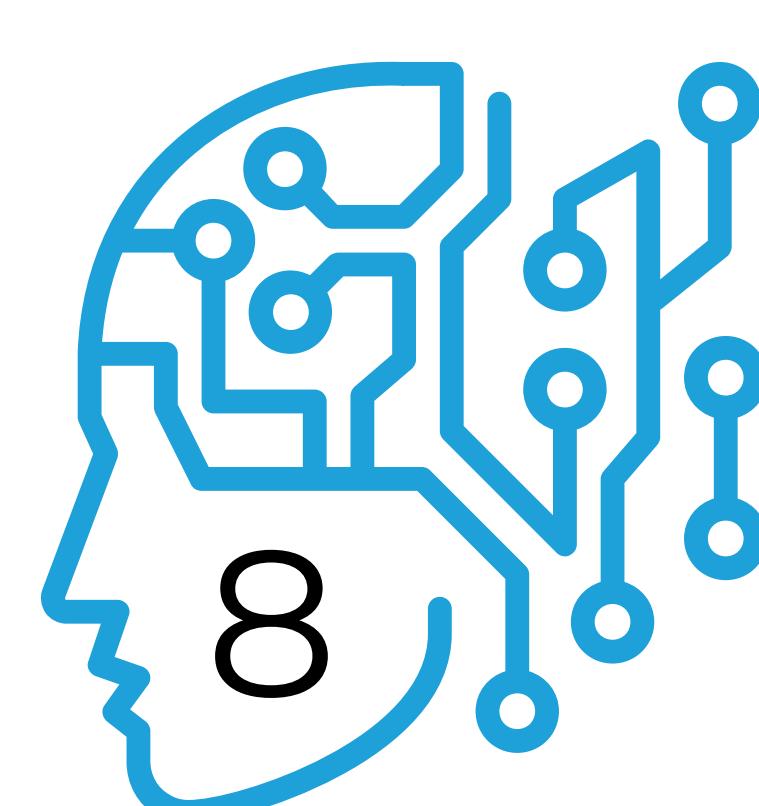
Cerebras: Founded in 2016 by Andrew Ng, a renowned AI expert, Cerebras is pushing the boundaries of AI hardware with its wafer-scale chip architecture. Unlike traditional chips with billions of transistors on a single silicon wafer, Cerebras uses hundreds of interconnected wafers to create a massive chip. This "one-chip computer" boasts 1.4 trillion transistors, overpowering even the most powerful GPUs. Cerebras chips can handle massive datasets. They provide a higher level of efficiency, opening new avenues to explore natural language processing, protein folding, and climate modelling.

Key Features:

- Wafer-scale architecture with 1.4 trillion transistors
- Massively parallel processing for large datasets and complex models
- Reduced power consumption compared to traditional AI hardware
- Focus on high-performance computing and scientific research

Problems:

- Chips are not readily available due to complex architecture and manufacturing for wafer-scale processors.
- Excels at raw processing power are not optimised for specific AI tasks like deep learning.



Graphcore: While Cerebras focuses on sheer scale, Graphcore takes a different approach with its "intelligent processing unit" (IPU) architecture. These particular chips are made to perform the calculations required by graph algorithms and deep learning. IPUs offer substantial performance gains due to their specialised cores and memory, which optimise data flow and communication, setting them apart from CPUs or GPUs. Pfizer, BMW, and Facebook already use Graphcore's IPUs for various AI applications, from financial modelling to drug discovery.

Key Features:

- IPU architecture optimised for deep learning and graph algorithms
- High-bandwidth memory and specialised cores for efficient data flow
- Scalable architecture for large-scale AI deployments
- Focus on practical AI applications in various industries

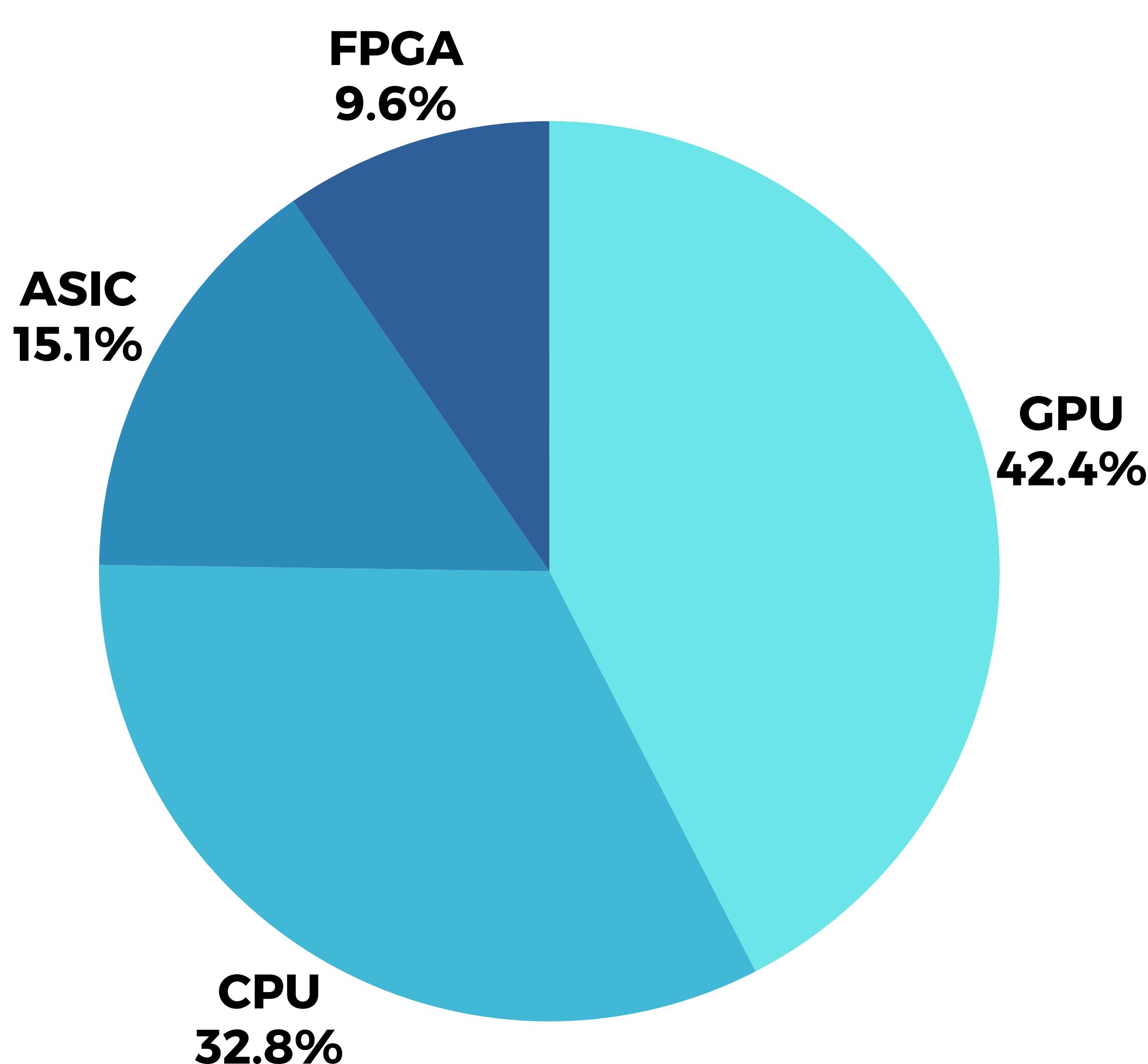
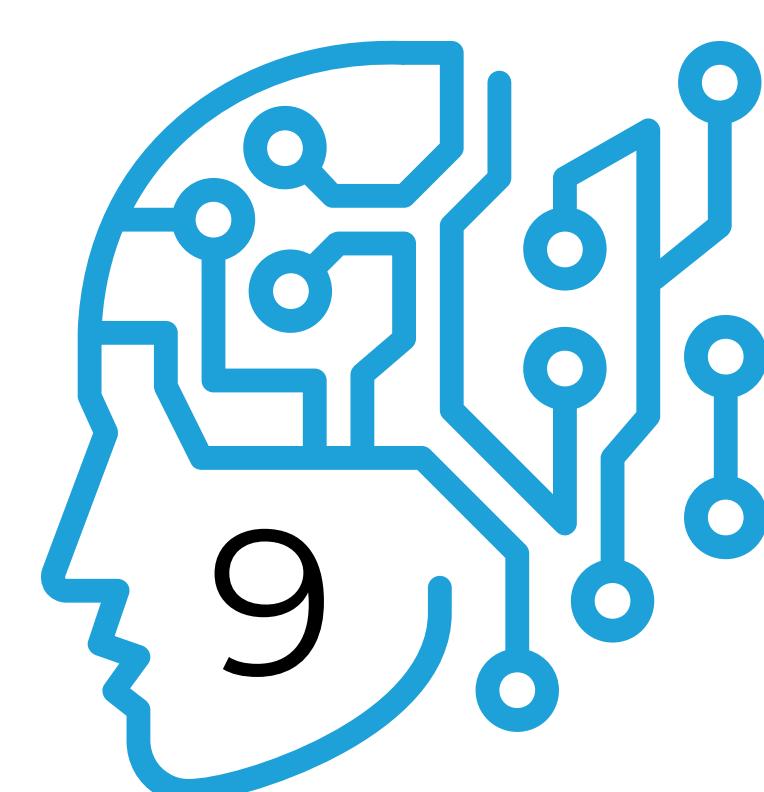


Exhibit (8) -Global AI chipset revenue by market share



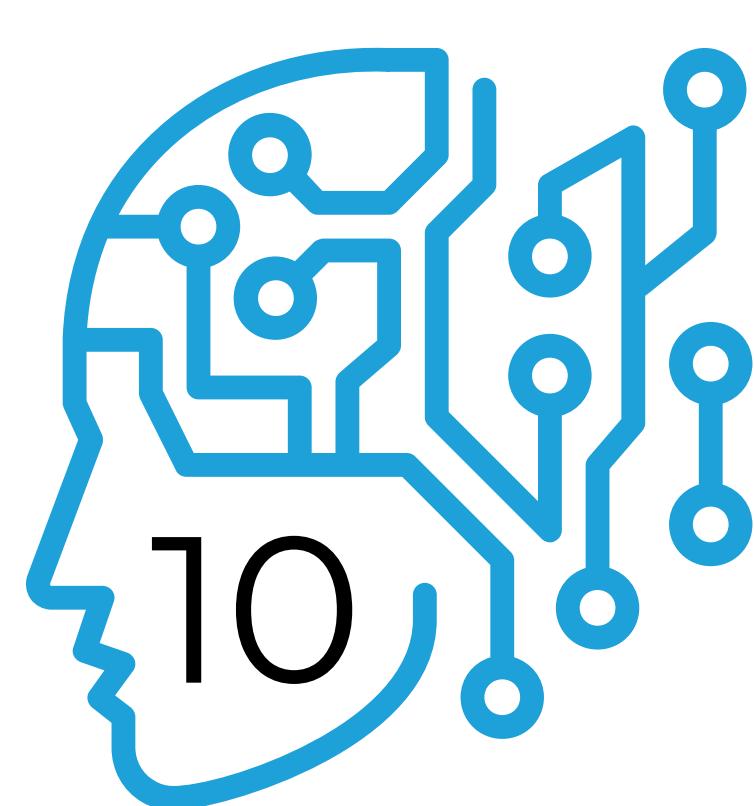
Problems:

- Excels at specific tasks but limited scalability for massively parallel tasks.
- It uses its domain-specific language (Poplar), which might bug developers.

This rapidly rising segment of the AI specific chip industry has opened up a new facet. The success of many of these large AI firms will depend on how they can complement their state of the art software with specialised hardware which will boost the pace of development of AI.

Conclusion

OpenAI is undoubtedly at the forefront of AI development owing to the unprecedented adoption of ChatGPT and DALL-E, while also being in a continuous tussle of capital, regulation and innovation. Many firms, although not as prominent as OpenAI, are capitalising on the booming AI market fueled by heavy demands from individuals and enterprises alike. GPTs and GPT stores would mark a critical step in determining OpenAI's market position for building the most pragmatic solutions over a hyper-competitive landscape. In order to sustain the research the firm is best known for, developing AI-specific chips could be a viable option, but it comes not without its own drawbacks and challenges. Whether OpenAI navigates through these barriers successfully or not is unclear. However, what the future beholds for AI is indeed enticing.

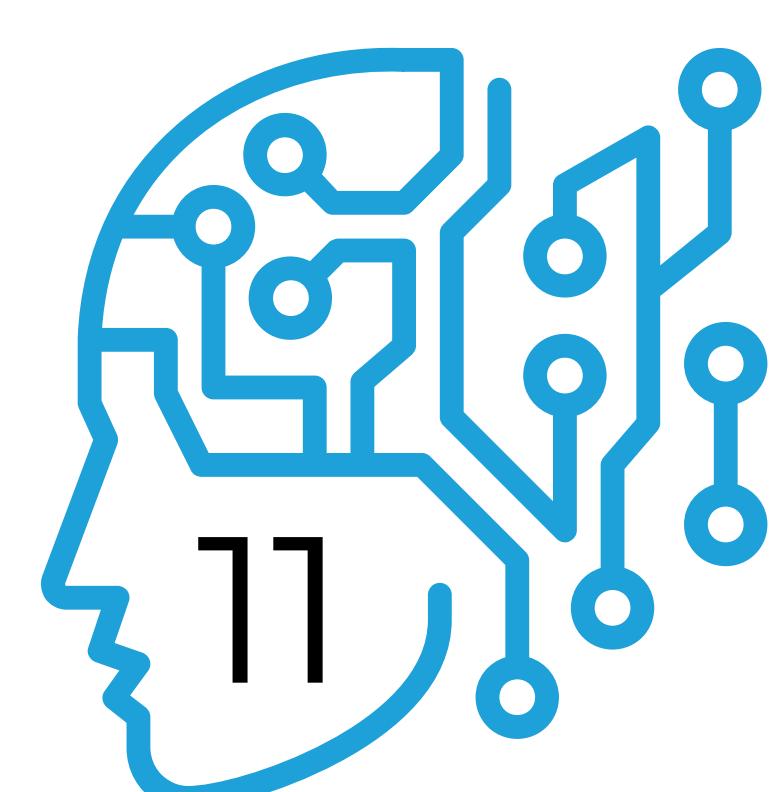


Questions

1) GPT Store with the GPT Builder has the potential to revolutionise chatbots, offering a level of customisation and freedom to its users never seen before in the industry. However, it also comes with a number of questions for OpenAI. The company can employ several different methods to monetise the GPT Store based on its long-term goals. Keeping various industry factors in mind, construct an appropriate revenue model for the GPT Store, exploring the proposed model's financial feasibility and a go-to-market strategy to stand out among its competitors.

Also, recommend strategies for OpenAI to address the potential for oversaturation in the GPT marketplace and ensure that the platform fosters the development of high-quality, specialised, and sustainable GPTs that meet the diverse needs of users.

2) OpenAI's recent venture into the semiconductor market opens up several opportunities for the company along with several challenges. Braving it out alone in an industry requiring huge capital investments could be a major challenge for the company. Keeping this in mind, suggest a strategy for OpenAI to onboard a firm or startup in the AI chip design industry and the feasibility of such an onboarding. Justify your suggestion considering various metrics such as economic and market synergies, goals, etc, and state the advantages of such a move for both sides.



Rating Criteria and Qualification Requirements

For the final round, the teams will be evaluated on the basis of the following criteria:

- Quality of Analysis
- Creativity
- Feasibility of the Solution
- The overall Presentation of the Case

Note

- Teams must submit their presentation by 11:59 PM, 3rd Jan, 2024.
- Submission should be in PPT format only(maximum 15 slides excluding the introduction, final slide, and Appendix) with an accompanying PDF copy of the same presentation.
- The solutions should be mailed to the iccs submissions2024@gmail.com with attachments as .pptx and .pdf.
- Relevant information can be added for the analysis from the web or other suitable resources.
- Any form of plagiarism will be heavily penalized.
- The inclusion of an executive summary slide is mandatory.

Copyright Notice

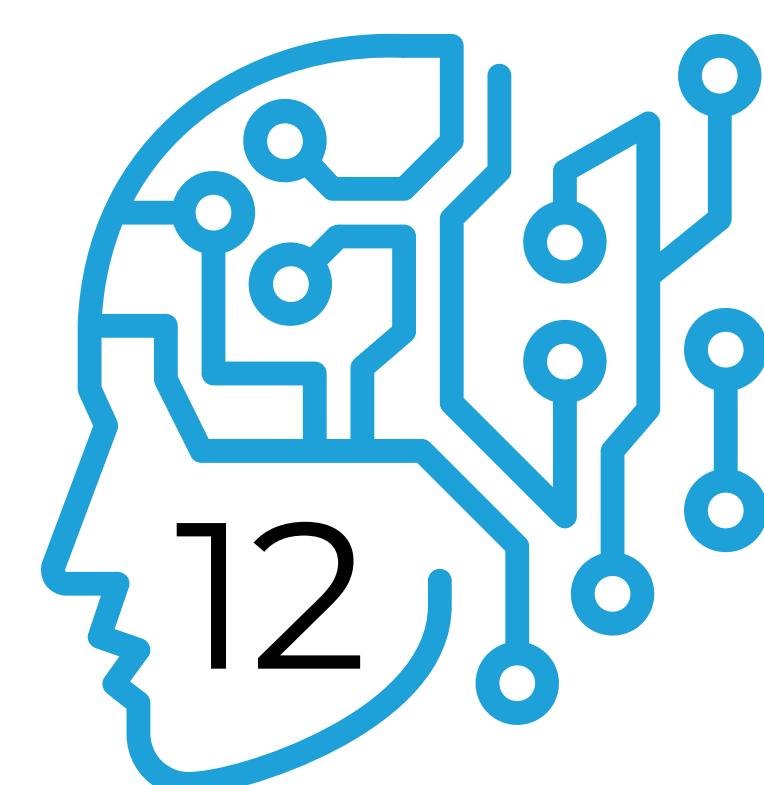
All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of Business Club, IIT Kharagpur.

This case is the intellectual property of Business Club, IIT Kharagpur.

For further information, visit our website at <https://icc.bclub.co.in/> or contact us at our FB page <https://www.facebook.com/bclubkgp>.

For any further queries, feel free to call or Whatsapp:

1. Rajshekhar Upadhyay - 9137454490
2. Arnav Bansal - 9522280723



References

- [Artificial Intelligence Market](#)
- [Microsoft-OpenAI partnership](#)
- [ChatGPT Operating Costs](#)
- [Cloud Infrastructure cost comparison](#)
- [The State of AI - Industry Report](#)
- [Developments in AI](#)
- [OpenAI: about the company](#)
- [OpenAI dependance on microsoft](#)
- [OpenAI products](#)
- [OpenAI Company Profile](#)
- [Sam Altman Dismissal Timeline](#)
- [Gemini vs Chatgpt](#)
- [Graphcore](#)
- [Devday Announcement blogs](#)
- [OpenAI current market position](#)
- [OpenAI board/management](#)
- [Global AI chipset Market](#)
- [OpenAI - Generative AI market opportunities](#)
- [AI Chip ventures](#)
- [Cerebras's chip development](#)
- [Understanding the semiconductor structure](#)
- [Graphcore IPU processors](#)
- [AI roadmap](#)
- [Google's gemini stacking against OpenAI](#)
- [Current state of the semiconductor industry](#)
- [Sam Altman's vision for future of AI](#)
- [Classification of AI](#)

