<u>Topic Brief 2: Scaling up & Growth in Start-ups</u>

Anthropic : *Arch nemesis of Open AI*

- Vishnu Kariyattu (April 17, 2024)

Introduction

The tech giants and the venture capitalists have been throwing more than $29 billion is the last year itself at AI based start-ups following the success of Open AI's ChatGPT. **[1]** In 2021, a group of seven safety conscious researchers defected from Open AI to start Anthropic. While Open AI is going for scale and reachability by building models that are easy for developers to work with and build new and smart things, Anthropic, with its Constitutional AI feature, is focused on the safety as their primary motto. With Claude versions of Anthropic surpassing GPT-4 in full document mastery and the widespread acceptance of AI technologies, thanks to Open AI, has the cautious and steady approach given Anthropic the edge and with the latest versions of Anthropic in market, is Anthropic winning the race? If they are to scale successfully, what are they to keep in mind? **[2]**

Background of the company

Anthropic PBC is an artificial intelligence (AI) startup company based out of US that is researching artificial intelligence as a public-benefit company to deploy safe, and reliable models for the public based on their research of safety properties of AI. Anthropic was founded by Daniela Amodei and Dario Amodei after their directional differences, specifically with OpenAI' s ventures with Microsoft in 2019. The public benefit registration is a cautious strategy which will require the company's directors to align the company's priorities with those of the public and this will prevent it from deviating from its original mission. The main stakeholders of the firms are Dario Amodei (Co-Founder & Chief Executive Officer), Daniela Amodei (Co-Founder & President), Jason Clinton (Chief Information Security Officer), Jared Kaplan (Co-Founder & Chief Science Officer), Ben Mann (Co-Founder & Member of Technical Staff), Jack Clark (Co-Founder & Head of Policy).**[3] [Exhibit 1]**

Anthropic' s model has two stages. The Supervised Learning (SL) Stage and the Reinforcement Learning (RL) stage and in these stages, it generates responses to harmfulness prompts, then critiques and revises the response and then fine tunes with supervised learning on the final revised response. After this, the Anthropic AI generates multiple datasets of preferences for harmfulness and decides which is better according to the constitutional preferences. Then it trains the preferred model and proceeds to fine tune the original SL model with RL model using the new preference model that was accepted and trained in the previous steps. **[4] [Exhibit 2]**

(Constitutional Model: The model is given a list of general principles (a "constitution") to follow when crafting its responses instead of human feedback and this is also the thought process behind physical models which are better than statistical models.)

Revenue Growth and Funding

As of March 2024, Anthropic has achieved an annual recurring revenue (ARR) of $316 million and has grown 1,341% annually from the $22 million ARR in March 2023. This means Anthropic achieved a 61x revenue multiple from its last valuation of $18.4 billion.**[Exhibit 3]** Three months prior to March 2023, Anthropic had reported a $100 million ARR and expected the figure to touch $500 million ARR by the end of 2024. However, Reuter' s report suggests that Anthropic has projected the ARR to $850 million and then there are people close to Anthropic who believe that this figure could touch an astounding ARR value of $1 billion. **[5]**

Due to the advanced model and above said continued revenue growth, Anthropic, has attracted the attention of several investors. Google & Salesforce were two of the early investors who invested $450 million in May 2023 in Anthropic. This was followed by a $100 million investment by Korean telco SK Telecom in August 2023. Since Anthropic is already using Amazon's cloud services, this attracted an investment of $1.25 billion from Amazon in September 2023 followed by a commitment to invest an additional $2.75 billion by first quarter of 2024, which Amazon fulfilled in March 2024. This is apart from the additional funds of $1.5 billion from Google in October 2023 and $750 million from Menlo Ventures in February 2024.**[6]** The price per share during this duration grew from $2.6 in May 2021 to $30 in Jan 2024.**[Exhibit 4]** The A.I. start-up has raised $7.3 billion, and the deals have some comparative advantage where Anthropic will purchase chips and use services from its investors. This also means that some of the invested amount goes back to the investors. **[7][8]**

Considering the above statics on revenue growth, continued interest from investors, and a workforce of 240 employees, it is safe to conclude that the company is in Delegation stage (*Phase 3 in Greiner' s "The five Phases of Growth" framework*). This is where the culture and values must be firmly set as there is increased number of specialist hires, increased accountability for those hires and reduced burden on the leadership team on tasks they are not good at. This can also lead to a crisis of control where senior leaders may feel that they're losing control over the decentralized structure. **[9]**

Product Development & Product Offerings & Value for Anthropic

"A rock doesn't hallucinate, but it isn't very useful."— Jared Kaplan, co-founder and chief science officer of Anthropic.

The big-name customers of Anthropic include Notion, DuckDuckGo, Quora and more, who use Anthropic' s products for uses from summarizing documents to chatting (Poe chatbot can hold conversation better than ChatGPT). Anthropic offers Claude Pro, which is available in three versions based on the complexity and nature of the problem it can handle and the cost one has to pay, which in the increasing order of complexity, are Haiku (Fast lightweight actions), Sonnet (Speed & performance for efficient, high-throughput tasks) and Opus (Most intelligent, can handle complex analysis, longer tasks with multiple steps, and higher-order math and coding tasks). It also offers the Legacy models Claude Instant, Claude 2.0 and Claude 2.1. The fact that Claude 2.1 can handle 200,000 words gives it the advantage of analyzing texts in its entirety compared to ChatGPT which has limited capacity of 32,000 words and hence can only analyze text in segments and must rely on chunking and retrieval (RAG) pipelines, which are crude and difficult to optimize. The detailed comparison between different models of Claude, ChatGPT and Gemini are given in the appendix in **Exhibit 5**. On the API side, Anthropic provides the users the option to Build or Scale. In the Build option, the user can create a proof-of-concept and launch their own generative AI solution and in the Scale option, the user can scale their own generative AI solution with custom rate limits and hands-on support from the Anthropic team.

Due to the above-mentioned advantages and increased need for privacy and safety, Anthropic' s products are being widely accepted. Considering Anthropic's product development cycle, Anthropic identified that safety and privacy are issues in AI world, for which, they created the constitutional models. When customers identified the same issue of safety and privacy, they wanted to adopt Anthropic' s AI solutions. Anthropic gave customers the option to integrate with existing AI or to build one from scratch. The more the number of clients onboarded the better the model Anthropic can build, which results in better solutions for each sector/ industry, and which drives more product development and more sales.

The pricing strategy is such that Anthropic can accommodate the financial constraints and technological needs of varied user groups, from individuals to enterprises, without compromising the quality or performance. As such, the value for Anthropic lies in scaling and can be identified by the **Exhibit 6**. It is also important to understand that the value for Anthropic increases because of reduction of costs and being a SAAS business, the marginal cost eventually becomes zero and the weighted increase in revenue can be considered as increase in profits.

Success factors of Anthropic can be understood by analyzing:

1. Strategy: The strategy of Anthropic is building systems that are reliable, interpretable, and steerable through conducting research, applying it to Anthropic' s products, feeding those insights back into

Anthropic' s research, and regularly sharing the results with the world. Anthropic also collaborates with civil society, government, academia, nonprofits, and industry to promote safety industry-wide.

2. Leadership and People: The leadership and the team are highly skilled and to attract great talent, Anthropic pays way above the industry norms for its hires. The team comprises of researchers, engineers, policy experts and operational leaders from a variety of disciplines. The research team, the policy team, the product team and the operations team, work hand in hand and the team has had previous careers at NASA, startups, and the armed forces and have rich diverse experiences.

3. Capabilities: From 200,000 words processing to being multimodal in nature, the capabilities surpass those of the other current market leaders. The products can be integrated into existing AIs or AIs can be built from scratch. The products are secure, accessible through AWS, and are trustworthy (10x resistant to jailbreaks & misuse) and reliable (Very low hallucination rates & accurate on long documents). These features help in creating operational efficiency, use in legal applications and code generation.

4. Governance & Culture: Anthropic believes in the mission of cultivating the ethical AI Framework as its mission. The values of Anthropic are 'Here for the mission', 'Unusually high trust', 'One big team' and 'Do the simple thing that works'. These tangible values are put into the interview loop through cultural interviews to understand the fit of the candidate with the company. Since Anthropic is in delegation stage, top notch talents are hired and are highly paid because moving forward, these talents will be second and third line of command and would decide the future of Anthropic. **[10]**

Competition & Risks

There is increased competition from Open AI' s ChatGPT, AI21 Labs' s Jurassic, Google's Gemini created by merging Google's DeepMind and Google Brain AI divisions, Cohere (mirror's Anthropic), Character.ai' s specialized Chatbots, and Meta' s Pytorch, Cicero, Segment Anything and RecD.

The risks of an upcoming technology are even more. The quality of the data that is used for training matters a lot and if it is garbage in, then it will be garbage out. Once the data is used to train these models, criticality is in achieving the balance between a well-trained model that hallucinates optimally and is not overly cautious. An overly cautious system becomes very, very worried about making mistakes and will say 'I don't know the context' for everything. The origin of model training data becomes an issue because Anthropic or any AI for that matter cannot exploit content without permission and if a content is accidently taken in by the AI in training, there is no specific way now to remove it. Once the data is incorporated, it is incorporated. Even when everything goes right, the output must be observed cautiously, and the source of the data has to be verified. Another problem that bugs Anthropic' s AI tools is limited availability, capacity, and cost of hardware like AI chips used in training. **[11]**

The challenge managing existing AI' s action and introduction of new AI on the regulatory side is that these systems are growing faster than humanity' s availability as researchers to evaluate them. **[12]**

Conclusion

With increased turbulence at Open AI, and Anthropic' s Claude beating Open AI's ChatGPT on many frontiers, many businesses, startups and enterprises alike, are going to look to Anthropic as a viable option. Navigating the murky waters is going to be a challenge for Anthropic especially because of the limitations mentioned in risks (above). To sustain the growth rate and maintain its status as a unicorn, Anthropic will have to provide particular attention to its organizational structure and culture. It adds to the advantage that Anthropic is good at filtering the talent that enters its system, but it will have to sustain the same quality when the organization grows big. One bad apple is enough to spoil the bunch. As Claude's pricing structures **[Exhibit 7]** and capabilities evolve, Anthropic will stay a big player but it must continue innovating to keep itself at the forefront of AI technology and being vital for users.

As more people want a piece of the pie, more AI startups will start getting funds, and the barrier to entry will go down. Businesses will most likely choose AI partners who are more aligned with them in terms of innovation, ethics, and technology. Anthropic will have to continue leveraging technological advancements, expand enterprise applications and cultivating the ethical AI framework.

From an industry point of view, the new AI things powered by the likes of Open AI need to keep happening and at the same time, players like Anthropic should make things safe and fair. And as things progress, it will become vital for companies in the AI space to strike a balance between both the worlds.

References:

1. Anthropic, Company Profile, Available at https://www.forbes.com/companies/anthropic/?sh=70bc02d2b94e (Accessed: 16 April 2024).

2. Sunil Rajaraman, OpenAI Faces Competitive Pressure From Anthropic (*Forbes*, 21 November 2023) <https://www.forbes.com/sites/sunilrajaraman/2023/11/21/openai-faces-competitive-pressure-from-anthropic/?sh=4f7be4955352> accessed 16 April 2024.

3. Jonathan Gillham, Anthropic Company Profile: 75+ Statistics and Trends in Q4 2023 (*Originality.ai*, April 8,2023) <https://originality.ai/blog/anthropic-ai-statistics > accessed 16 April 2024.

4. Michael Spencer, OpenAI vs. Anthropic is the Key 2023 Startup Battle to Watch in Generative A.I. (*AI Supremacy*, May 24,2023) <https://aisupremacy.substack.com/p/openai-vs-anthropic-is-the-key-2023> accessed April 16, 2024.

5. Reuters, Anthropic forecasts more than $850 mln in annualized revenue rate by 2024-end - report, Available at https://www.reuters.com/technology/anthropic-forecasts-more-than-850-mln-annualized-revenue-rate-by-2024-end-report-2023-12-26/ (Accessed: April 16, 2024).

6. **ApeVue,** FTX's Anthropic Shares Liquidation is Approved (March 5, 2024), Available at https://www.apevue.com/insights/ftxs-anthropic-shares-liquidation-is-approved (Accessed: April 16, 2024).

7. Erin Griffith and Cade Metz, Inside the Funding Frenzy at Anthropic, One of A.I.'s Hottest Start Ups (*The New York Times*, February 20,2024) <https://www.nytimes.com/2024/02/20/technology/anthropic-funding-ai.html#:~:text=Last%20May%2C%20Anthropic%2C%20one%20of,million%20from%20two%20Asian%20telecoms> accessed April 16, 2024

8. **Amazon,** Amazon and Anthropic deepen their shared commitment to advancing generative AI (March 27, 2024), Available at https://www.aboutamazon.com/news/company-news/amazon-anthropic-ai-investment (Accessed: April 16, 2024).

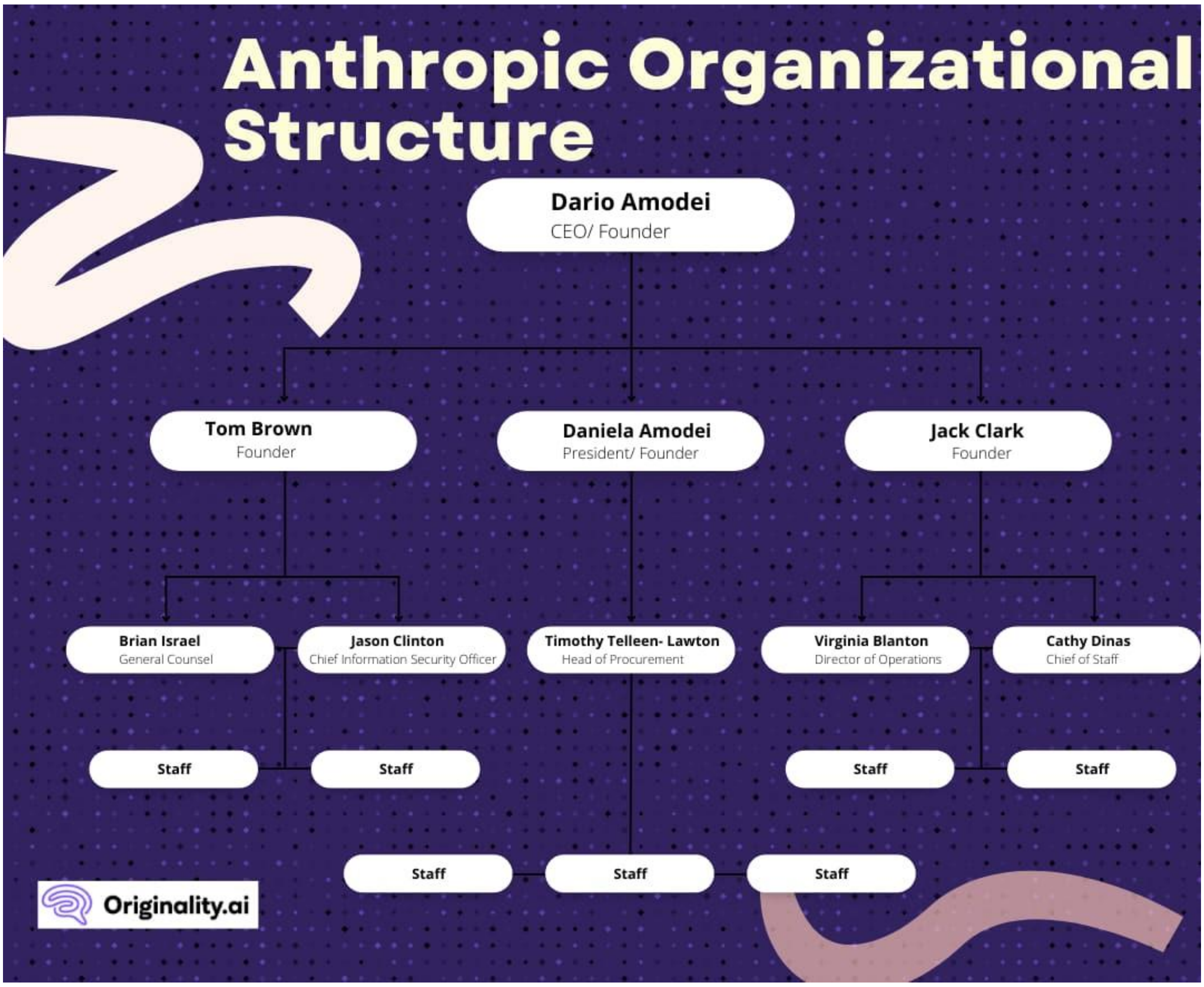
9. Larry E. Greiner, Evolution and Revolution as Organizations Grow (*Harvard Business Review*, May-June, 1998) <https://hbr.org/1998/05/evolution-and-revolution-as-organizations-grow> accessed April 16, 2024

10. **Anthropic,** Company, Available at https://www.anthropic.com/company (Accessed: April 16, 2024).

11. Belle Lin, Google and Anthropic Are Selling Generative AI to Businesses, Even as They Address Its Shortcomings (*The Wall Street Journal*, February 13, 2024) <https://www.wsj.com/articles/google-and-anthropic-are-selling-generative-ai-to-businesses-even-as-they-address-its-shortcomings-ff90d83d > accessed April 16, 2024

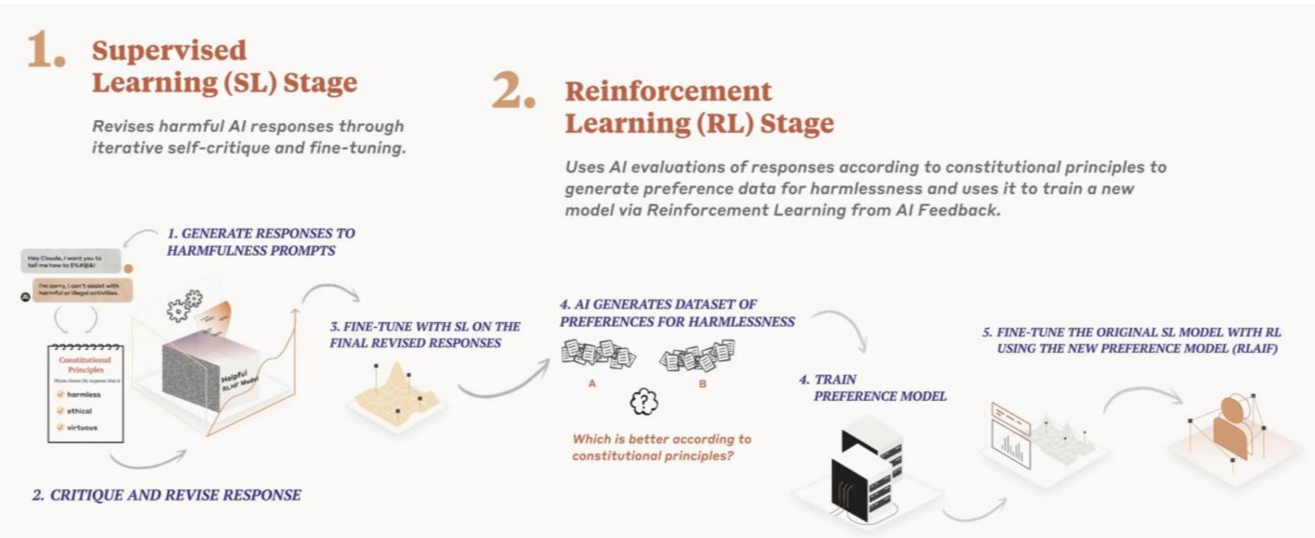
12. Ryan Tracy, Risks Posed by AI Technology Spur Calls for New Regulatory Agency (*The Wall Street Journal*, May 18, 2023) <https://www.wsj.com/articles/risks-posed-by-ai-technology-spur-calls-for-new-regulatory-agency-8cc2dd76 > accessed April 16, 2024

Exhibit 1: Anthropic Organizational Structure:



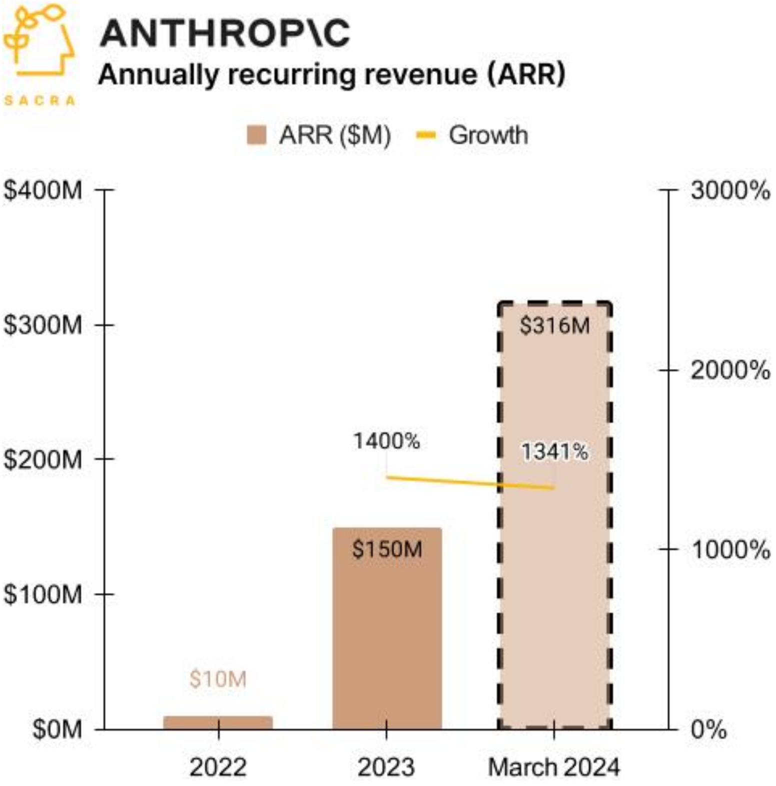
Source: https://originality.ai/blog/anthropic-ai-statistics

Exhibit 2: Anthropic' s model:

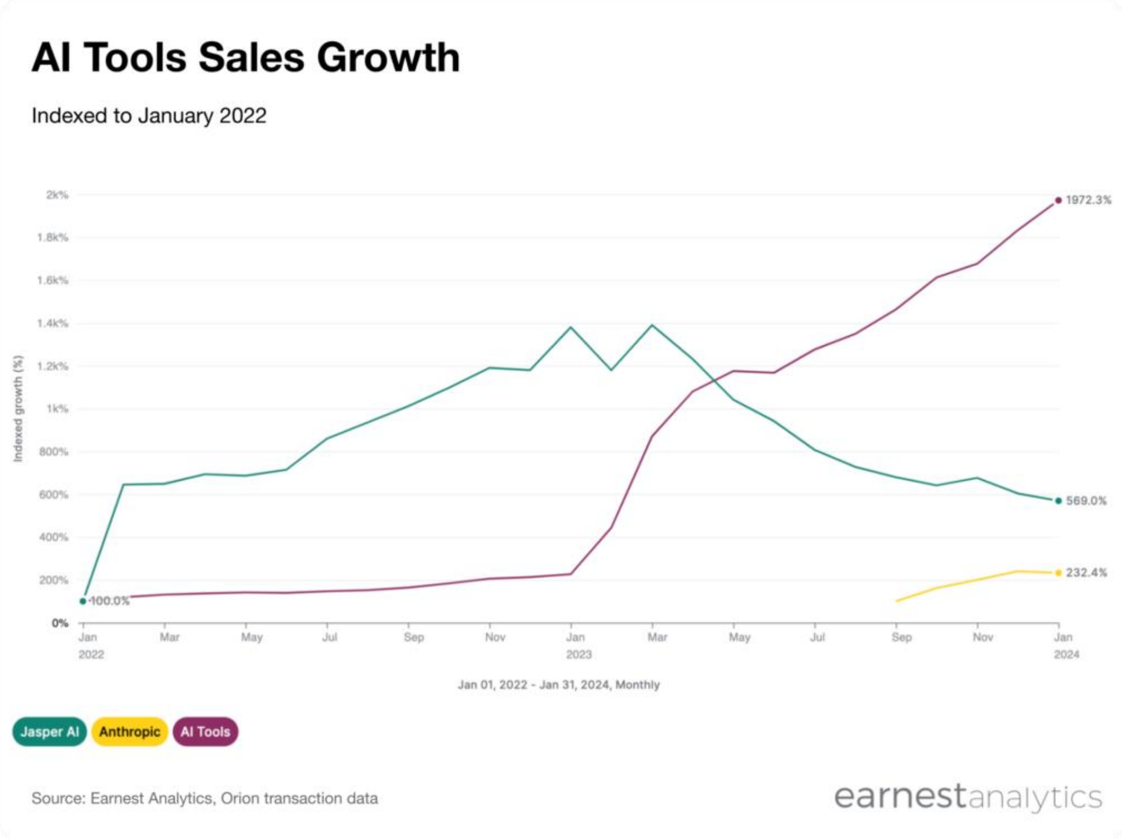


Source: https://aisupremacy.substack.com/p/openai-vs-anthropic-is-the-key-2023

Exhibit 3: Anthropic' s Annual Recurring Revenue & Sales Growth

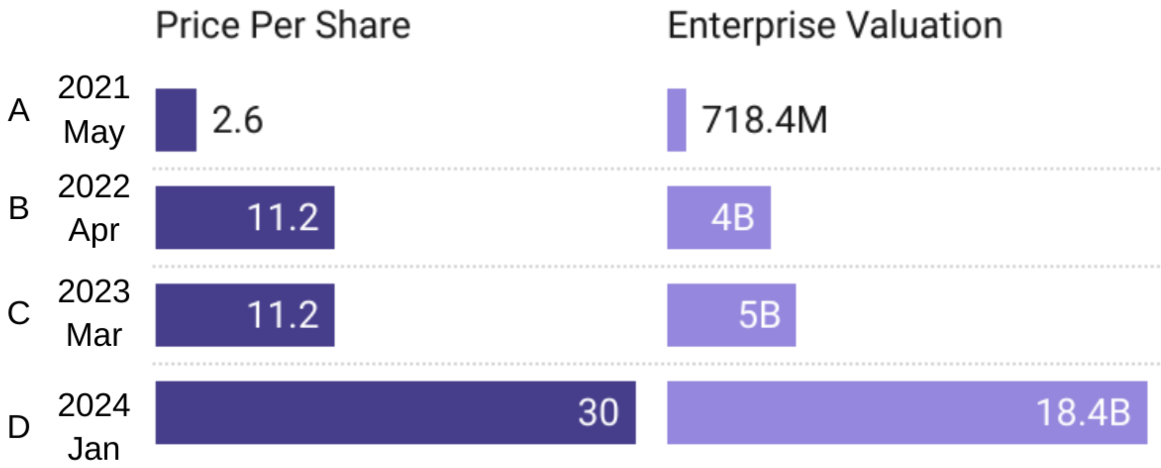


Source: https://sacra.com/c/anthropic/



Source: https://sacra.com/c/anthropic/

Exhibit 4: Funding Round in USD



Source: https://www.apevue.com/insights/ftxs-anthropic-shares-liquidation-is-approved

Exhibit 5: Comparison of the Claude 3 models to those of our peers on multiple benchmarks of capability

	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

Source: https://www.anthropic.com/news/claude-3-family

Exhibit 6: Value – Price graph over time

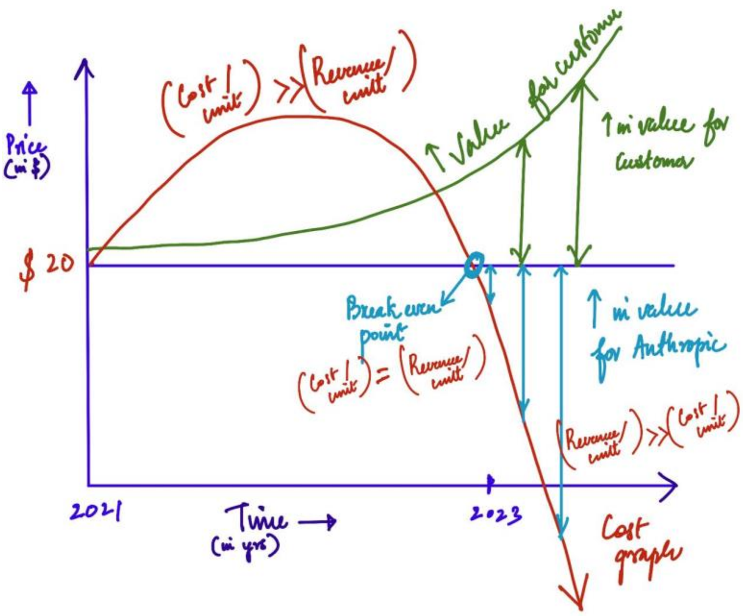
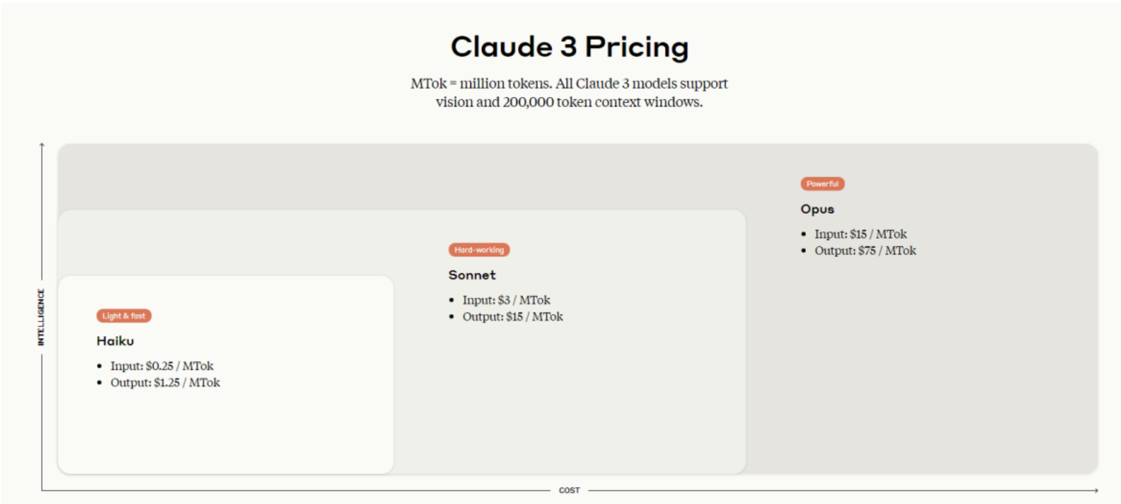
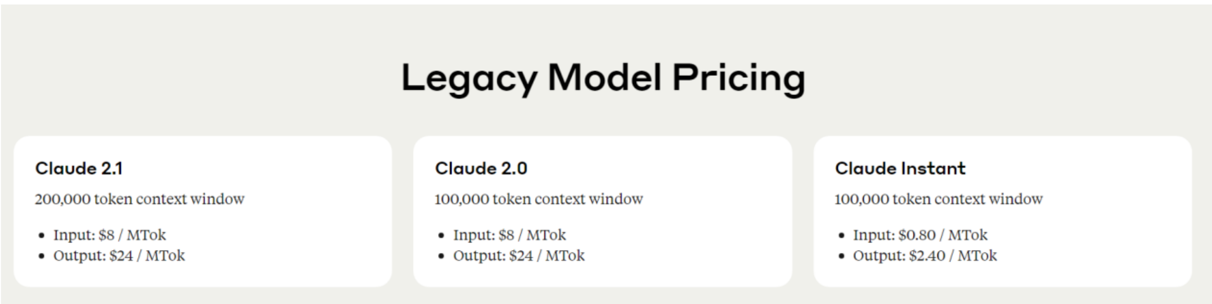


Exhibit 7: Pricing Models of Anthropic



Source: https://www.anthropic.com/api

Legacy Models: Claude 2.1 and Claude 2.0



Source: https://www.anthropic.com/api