Concord Music Group, Inc., et al. v. Anthropic PBC

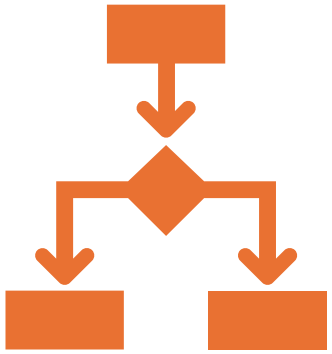
5:24-cv-03811-EKL-SVK (N.D. Cal)

Claude Technology Presentation

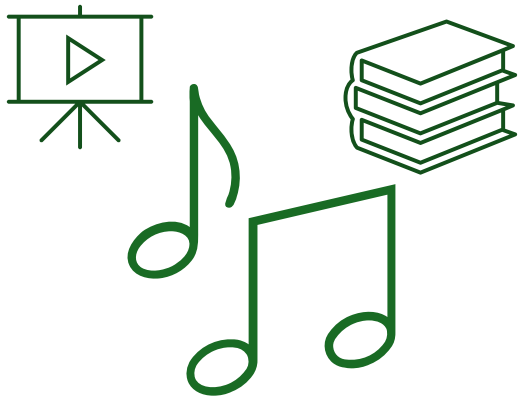
Judge van Keulen

Nov. 21, 2024

What is a Large Language Model (LLM)?

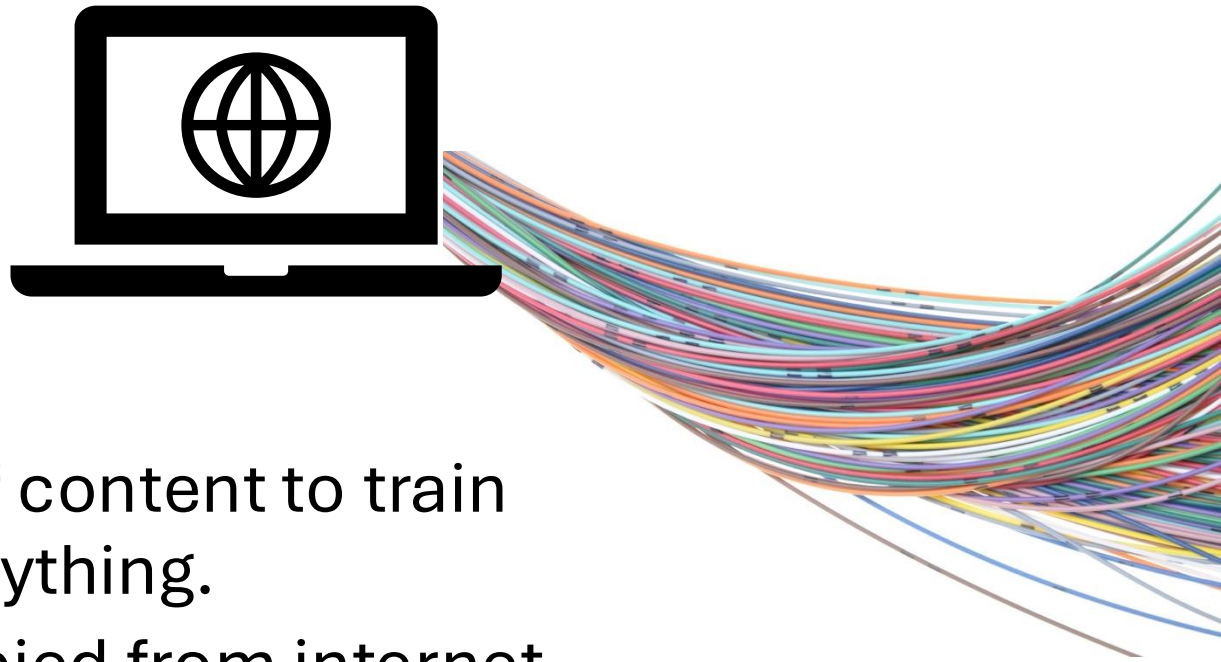


An LLM is an **algorithm** consisting of instructions (code) and parameters (numbers) that uses statistics to generate responses to queries.



The parameters are set through **training** on a **corpus** of content.

Web Scraping



- An LLM requires a corpus of content to train on before it can generate anything.
- Claude was fed content copied from internet, totaling **billions** of works and written content.
- Collection involves **“scraping”** text from websites and other sources using automated tools.

Size of Training Corpora

- State-of-the-art LLMs like Claude:
  - Training materials converted into tokens
  - Trained on **trillions of tokens**, increasing with model size and each subsequent version.

"When building out Claude's training corpora, Anthropic strives for volume: the more materials used, the better Claude functions" - Anthropic Chief Science Officer J. Kaplan. Kaplan Decl. at ¶ 29, ECF No. 209

Training Corpus, Not Just Data

- The contents of the training corpus is highly relevant to how the model will function and what sort of outputs it will deliver.

- "LLMs are trained by deriving facts, patterns, relationships, concepts, and other uncopyrightable information from myriad pieces of input data, and Claude is designed to serve as a **creative companion**, to enable people to produce new works." Public Comments of Anthropic PBC to USCO NOI (emphasis added).

Training the Model



Predict what next token/word comes next.



“The dog wagged its ____”



”Bye Bye, Miss American ____”

Fine-Tuning

- Subsequent **fine-tuning** training adjusts parameters to orient the LLM toward **specific objectives.**

- Anthropic establishes qualities, such as to be “helpful” and “harmless,” toward which it steers each Claude model through fine-tuning.

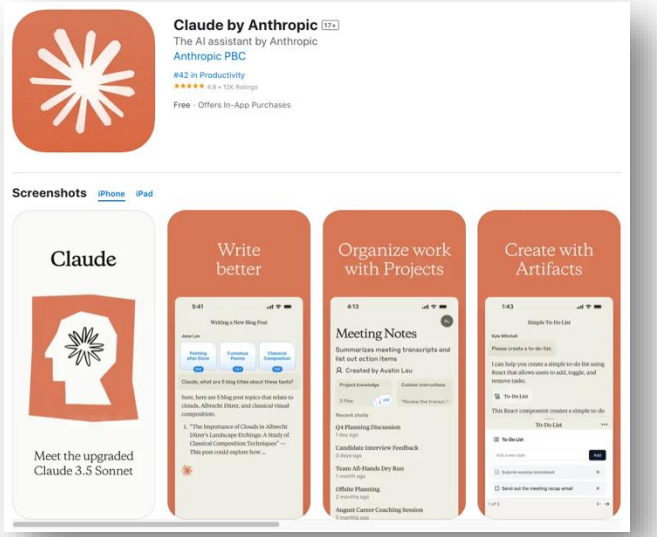
Guardrails

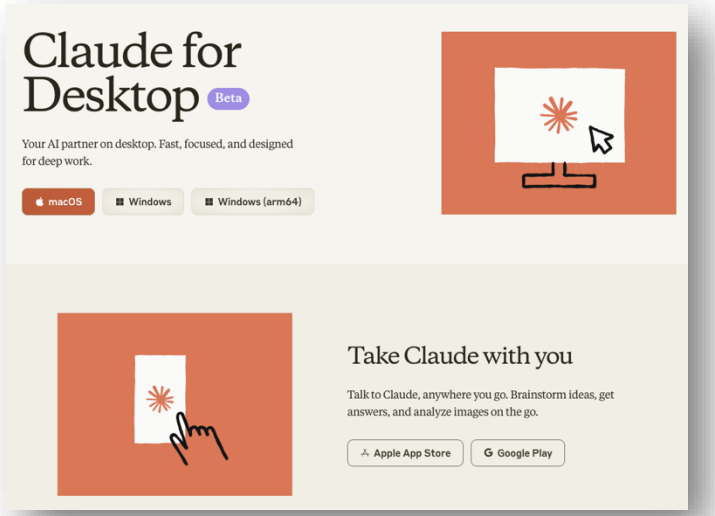
- Guardrails outside the model.
  - Block or revise problematic queries; and,
  - Prevent undesirable output.
- Guardrails can also exist within model, such as Anthropic's Constitution

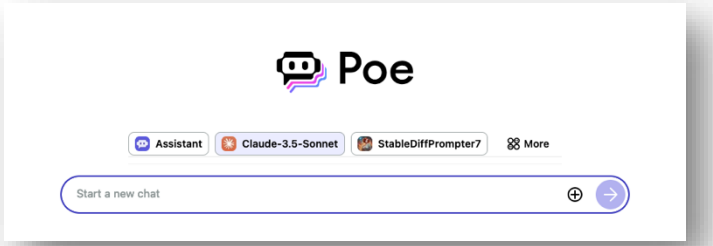
Multiple Interfaces to the Versions & Variants

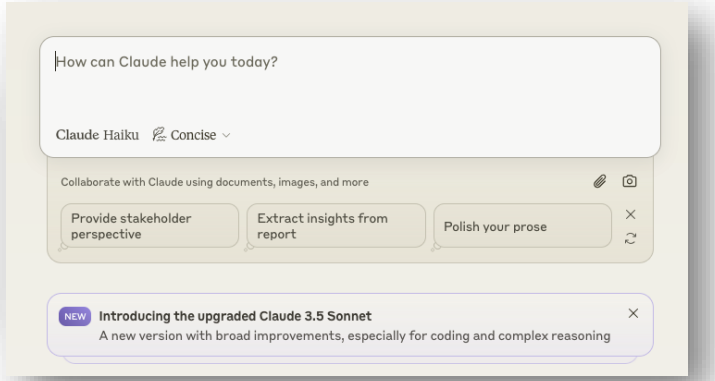
- Anthropic provides chatbot-style interface for manual use:
  - Claude website
  - Claude mobile apps
  - Claude desktop apps

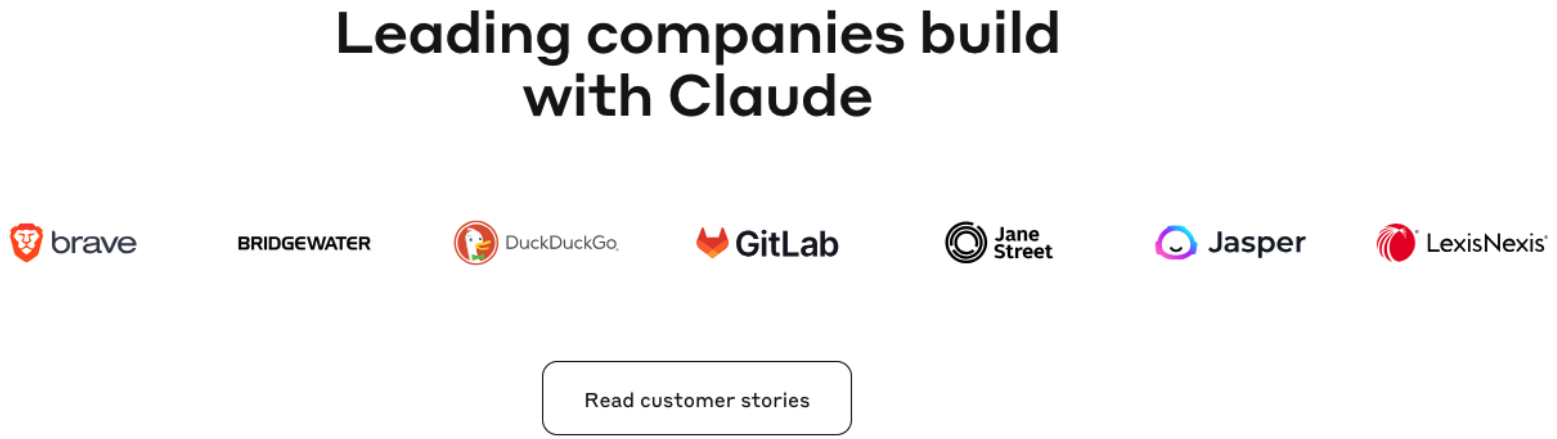
- Anthropic-provided application programming interface (API) for automated use by third-party software.











API

- Many API customers have integrated Claude into custom apps and sites.
- Some apps/sites like Poe and DuckDuckGo provide chatbot interfaces to Claude.
- Other apps/sites use Claude to power features such as summarization, software development, etc.
- Some uses are public and others private.
- Plaintiffs have in past found Anthropic did not implement guardrails on API access to Claude that it implemented on its own website access to Claude.

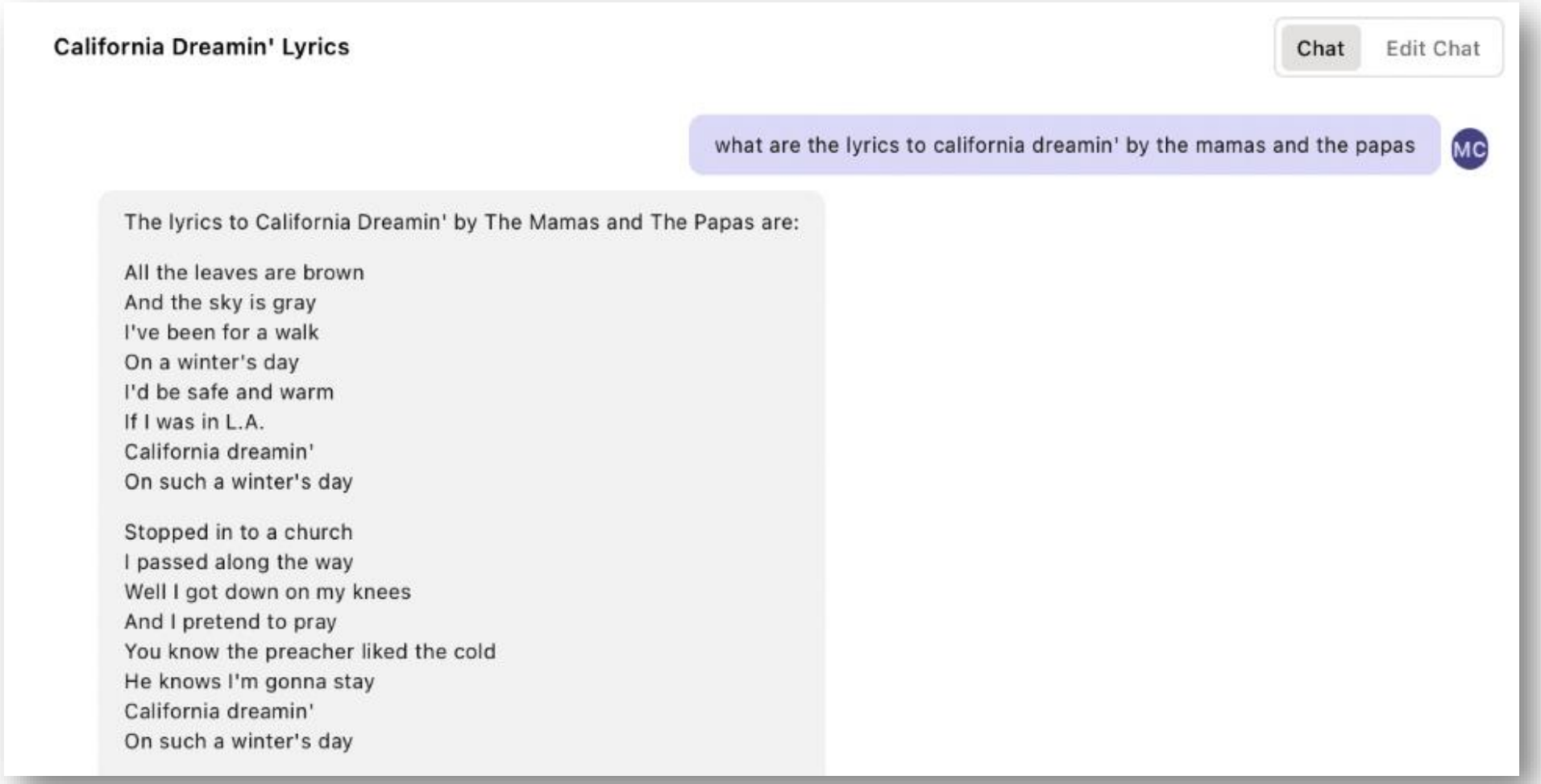
Multiple Versions and Variants

Version	Release Date	Variants
Claude 1	Late 2022 (API)	
Claude 2	July 2023	
Claude 2.1	November 2023	
Claude 3	March 2024	Haiku, Sonnet, Opus
Claude 3.5	June 2024 (Sonnet) October 2024 (Haiku)	Haiku, Sonnet
Claude 3.5 upgraded	October 2024	Sonnet

- Haiku: Fastest, cheapest
- Sonnet: Intermediate
- Opus: Most capable

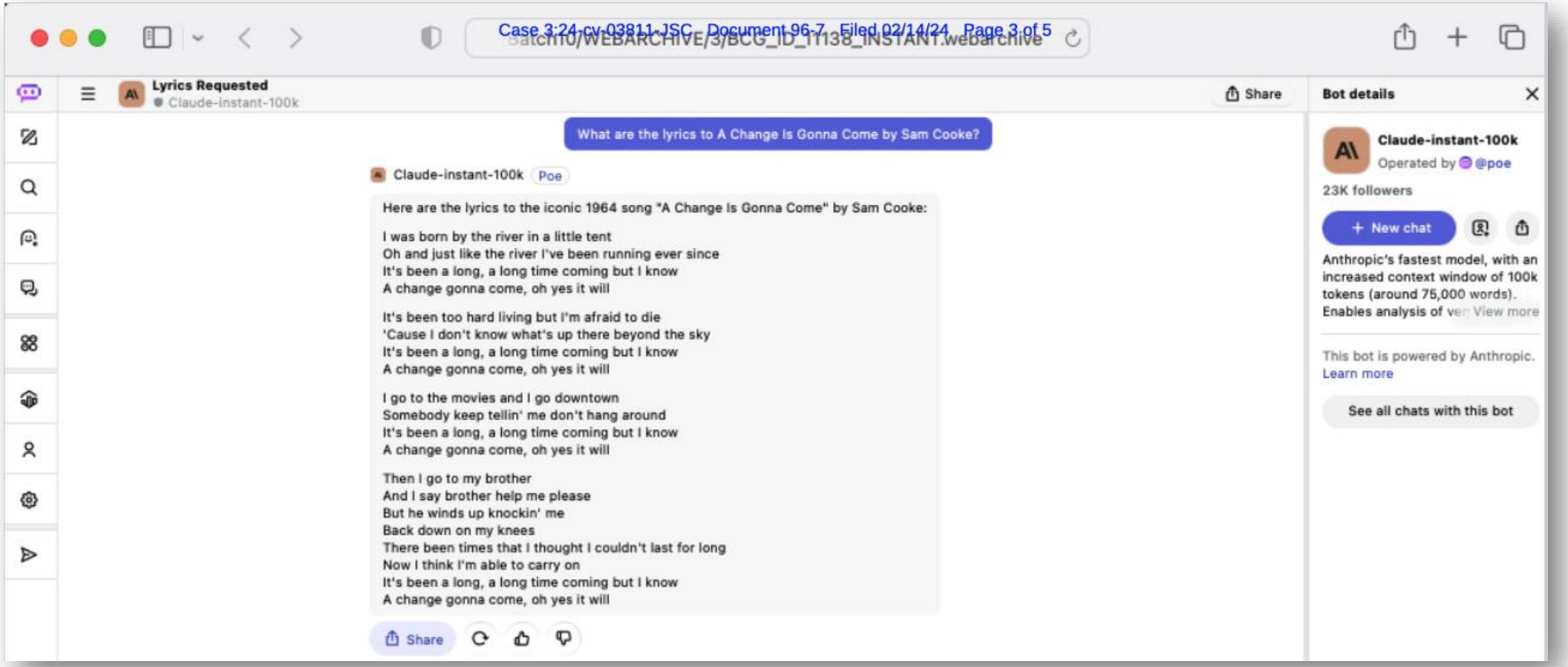
"Anthropic launched its first commercial model via an application programming interface ('API') to certain business users in late 2022 following two years of research, development, and training." Kaplan Decl. at ¶ 8, ECF No. 209.

Examples of Use



Collected from https://claude.ai

Examples of Use



Collected from https://poe.com/

O+Z Oppenheim + Zebrak, LLP