



Presenter

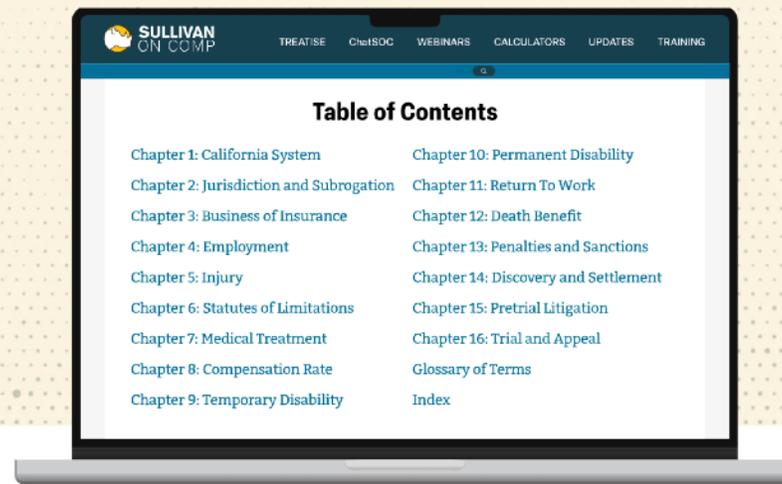
Michael Bell

From Treatise to Tech:
The Development of ChatSOC,
an AI Research Assistant



SULLIVAN ON COMP

The most comprehensive and up-to-date digital resource on California workers' compensation law for attorneys, adjusters, and other industry professionals



**How much time do you waste combing through different resources?
Or waiting for your colleagues to get back to you with answers?**

Save countless hours by having all the information on California workers' compensation law right at your fingertips.

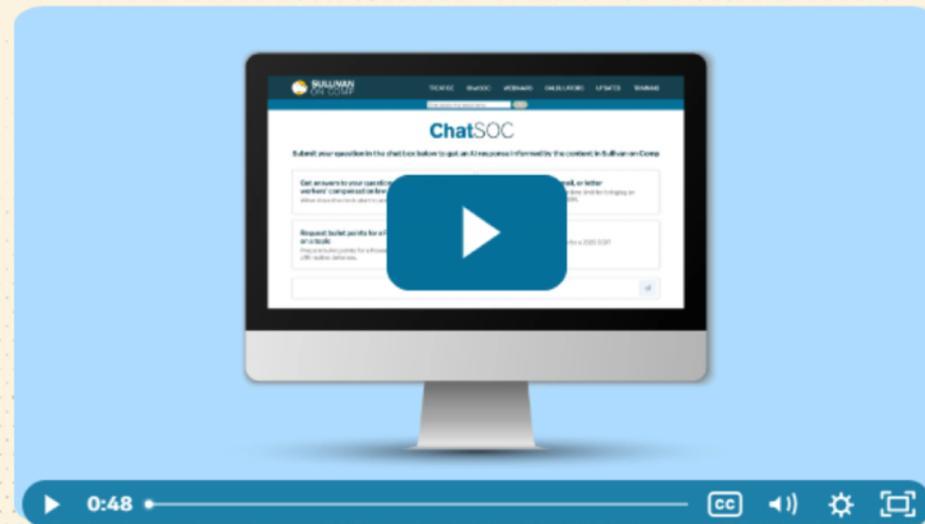


It's like Google, but for California workers' comp.

—Adam Dombchik, Co-Managing Partner, Gordon, Edelman, Krepack, Grant, Felton & Goldstein

ChatSOC

Your personal research and communication assistant that helps you work faster



WHAT IS ChatSOC

ChatSOC is a chatbot that has access to all the content of Sullivan on Comp. It gives you an answer to legal or procedural questions about California workers' compensation in seconds by connecting the dots from different chapters of the treatise.

Rather than spending time going through different treatise chapters and piecing the information together yourself, you can simply ask ChatSOC a question and get an answer quickly, like you would from a human.

A business meeting in a modern, glass-walled conference room. A woman in a black blazer stands and speaks to two men seated at a wooden table. The man on the left is writing in a notebook, and the man on the right is looking at her. A laptop is open on the table. The text "How much time do you waste" is overlaid in white.

How much time do you waste

2022





ChatGPT: Optimizing Language Models for Dialogue

We've trained a model called ChatGPT which interacts in a conversational way. The dialogue format makes it possible for ChatGPT to answer followup questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests. ChatGPT is a sibling model to [InstructGPT](#), which is trained to follow an instruction in a prompt and provide a detailed response.

November 30, 2022

13 minute read



Thank you to all who joined us in Miami for the 2022 inaugural TLTF Summit!



Get Updates on TLTF Summit 2023



Startup Advice from World Class Experts

VERIFIED EXPERT

Kate Kendall

Founder and CEO of CloudPeeps. Created The Fetch.

Fundraising

Marketplaces

Social Media Marketing

Email Marketing

Community Building

Find Your Expert

Learn More





Build and train a stand-alone GPT-3 instance

Posted last year

 Worldwide

Build and train a stand-alone GPT-3 instance to answer user questions based on our legal treatise. Also, create an API interface that can be used to submit queries and provide responses.

 Less than 30 hrs/week
Hourly

 1-3 months
Duration

 Intermediate
Experience Level

 \$40.00-\$75.00
Hourly

Project Type: Ongoing project

Skills and Expertise

Machine Learning Languages

Python

Other

Amazon Web Services

Artificial Intelligence

What did we have?



Our Content

- 16 chapters
- 636 Sections
- 4,273 sub-sections
- Over 2 million words



CPWC-IND
Indemnity Workers'
Compensation Adjuster

[Learn more >](#)

Quiz Data

- 500 True/False & Multiple-Choice Questions
- Derived from a 120-hour 24-module adjuster training course

First Steps...

Training self-hosted models

Training self-hosted models

GPT 3.5

QUIZ PAPER

Name _____

Roll Number _____

1. The process of...
2. The...
3. The...
4. The...
5. The...
6. The...
7. The...
8. The...
9. The...
10. The...

Q1. (A) (B) (C) (D)
Q2. (A) (B) (C) (D)
Q3. (A) (B) (C) (D)
Q4. (A) (B) (C) (D)
Q5. (A) (B) (C) (D)
Q6. (A) (B) (C) (D)
Q7. (A) (B) (C) (D)
Q8. (A) (B) (C) (D)
Q9. (A) (B) (C) (D)
Q10. (A) (B) (C) (D)
Q11. (A) (B) (C) (D)
Q12. (A) (B) (C) (D)
Q13. (A) (B) (C) (D)
Q14. (A) (B) (C) (D)
Q15. (A) (B) (C) (D)
Q16. (A) (B) (C) (D)
Q17. (A) (B) (C) (D)
Q18. (A) (B) (C) (D)
Q19. (A) (B) (C) (D)
Q20. (A) (B) (C) (D)
Q21. (A) (B) (C) (D)
Q22. (A) (B) (C) (D)
Q23. (A) (B) (C) (D)
Q24. (A) (B) (C) (D)
Q25. (A) (B) (C) (D)
Q26. (A) (B) (C) (D)
Q27. (A) (B) (C) (D)
Q28. (A) (B) (C) (D)
Q29. (A) (B) (C) (D)
Q30. (A) (B) (C) (D)
Q31. (A) (B) (C) (D)
Q32. (A) (B) (C) (D)
Q33. (A) (B) (C) (D)
Q34. (A) (B) (C) (D)
Q35. (A) (B) (C) (D)
Q36. (A) (B) (C) (D)
Q37. (A) (B) (C) (D)
Q38. (A) (B) (C) (D)
Q39. (A) (B) (C) (D)
Q40. (A) (B) (C) (D)
Q41. (A) (B) (C) (D)
Q42. (A) (B) (C) (D)
Q43. (A) (B) (C) (D)
Q44. (A) (B) (C) (D)
Q45. (A) (B) (C) (D)
Q46. (A) (B) (C) (D)
Q47. (A) (B) (C) (D)
Q48. (A) (B) (C) (D)
Q49. (A) (B) (C) (D)
Q50. (A) (B) (C) (D)
Q51. (A) (B) (C) (D)
Q52. (A) (B) (C) (D)
Q53. (A) (B) (C) (D)
Q54. (A) (B) (C) (D)
Q55. (A) (B) (C) (D)
Q56. (A) (B) (C) (D)
Q57. (A) (B) (C) (D)
Q58. (A) (B) (C) (D)
Q59. (A) (B) (C) (D)
Q60. (A) (B) (C) (D)
Q61. (A) (B) (C) (D)
Q62. (A) (B) (C) (D)
Q63. (A) (B) (C) (D)
Q64. (A) (B) (C) (D)
Q65. (A) (B) (C) (D)
Q66. (A) (B) (C) (D)
Q67. (A) (B) (C) (D)
Q68. (A) (B) (C) (D)
Q69. (A) (B) (C) (D)
Q70. (A) (B) (C) (D)
Q71. (A) (B) (C) (D)
Q72. (A) (B) (C) (D)
Q73. (A) (B) (C) (D)
Q74. (A) (B) (C) (D)
Q75. (A) (B) (C) (D)
Q76. (A) (B) (C) (D)
Q77. (A) (B) (C) (D)
Q78. (A) (B) (C) (D)
Q79. (A) (B) (C) (D)
Q80. (A) (B) (C) (D)
Q81. (A) (B) (C) (D)
Q82. (A) (B) (C) (D)
Q83. (A) (B) (C) (D)
Q84. (A) (B) (C) (D)
Q85. (A) (B) (C) (D)
Q86. (A) (B) (C) (D)
Q87. (A) (B) (C) (D)
Q88. (A) (B) (C) (D)
Q89. (A) (B) (C) (D)
Q90. (A) (B) (C) (D)
Q91. (A) (B) (C) (D)
Q92. (A) (B) (C) (D)
Q93. (A) (B) (C) (D)
Q94. (A) (B) (C) (D)
Q95. (A) (B) (C) (D)
Q96. (A) (B) (C) (D)
Q97. (A) (B) (C) (D)
Q98. (A) (B) (C) (D)
Q99. (A) (B) (C) (D)
Q100. (A) (B) (C) (D)

Close

65

CPE

25

5

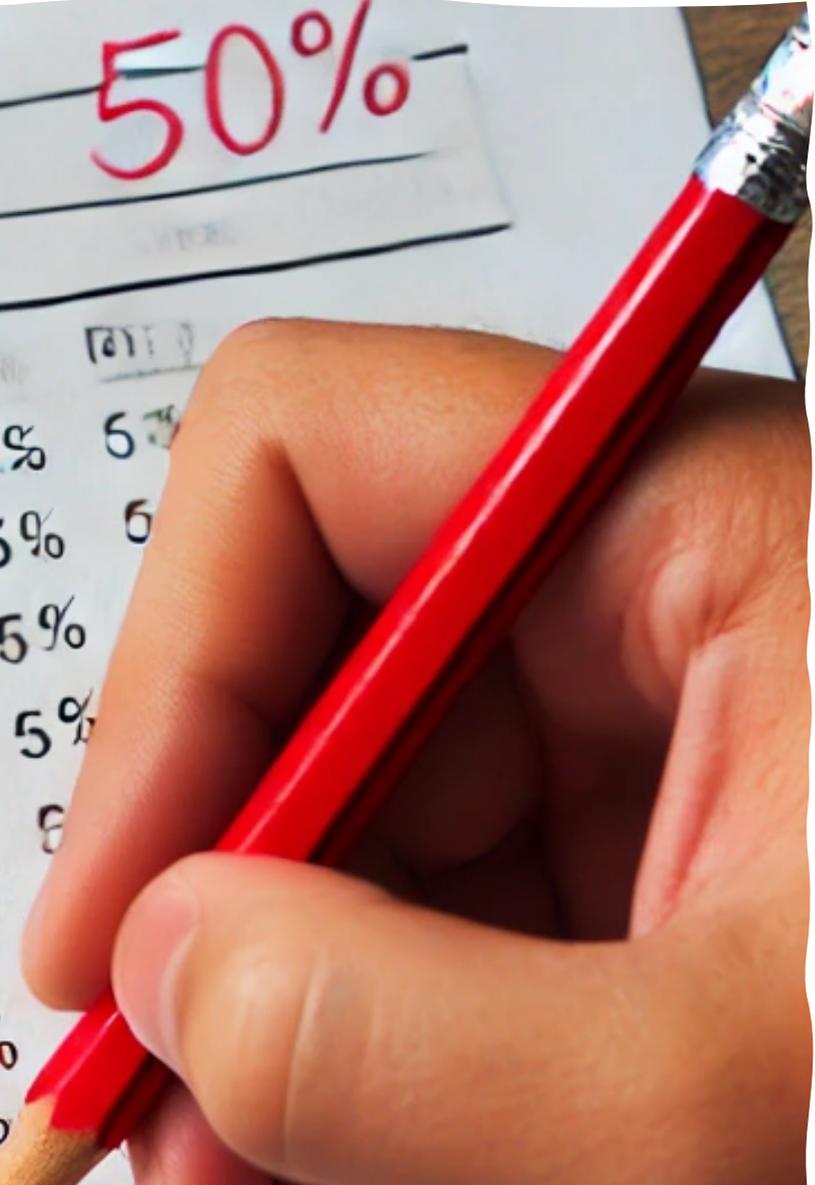
QUIZ PAPER

50%

QUIZ

1	Boonvateletoosentvoot	5%	5%	5%
1	Coontaan noretessan)	5%	5%	5%
A	Goedofert theertenee)	5%	5%	5%
2	Goederonabvarlantise)	5%	5%	5%
2	Googha beedersomee)	5%	5%	5%
2	Pleersto bon suraten)	5%	5%	5%
A	Bohce lal potestona)	5%	5%	5%
	...	5%	5%	5%

F



Training self-hosted models

~~GPT 3.5~~

Training self-hosted models

~~GPT 3.5~~

LLaMA-1 7B

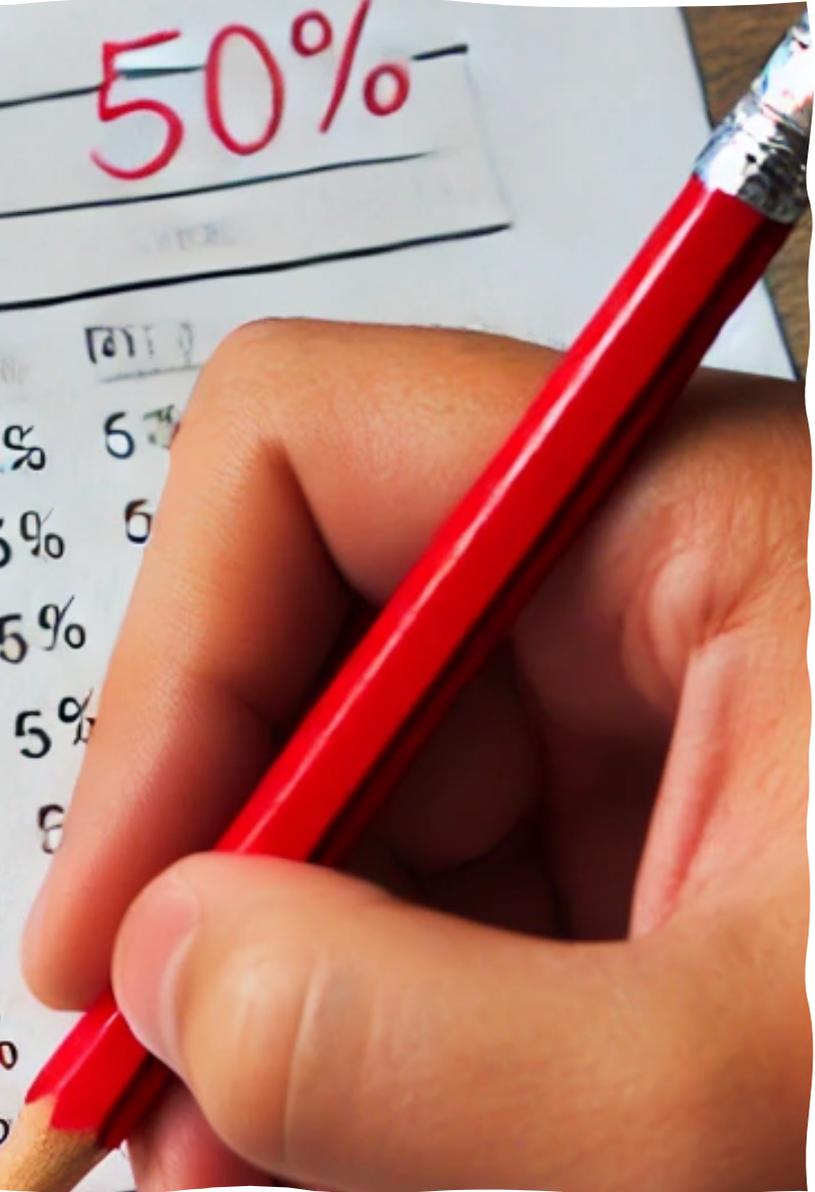
QUIZ PAPER

50%

QUIZ

1	Boonvateletoosentvoot	5%	5%	5%
1	Coontaan noretessan)	5%	5%	5%
A	Goedofert theertenee)	5%	5%	5%
2	Goederobstvarlantise)	5%	5%	5%
2	Googha beedersomee)	5%	5%	5%
2	Pleersto bon suraten)	5%	5%	5%
A	Bohce lal potestona)	5%	5%	5%
	...	5%	5%	5%

F



Training self-hosted models

~~GPT 3.5~~

~~LLaMA-1 7B~~

Training self-hosted models

~~GPT 3.5~~

~~LLaMA-1 7B~~

LLaMA-1 13B

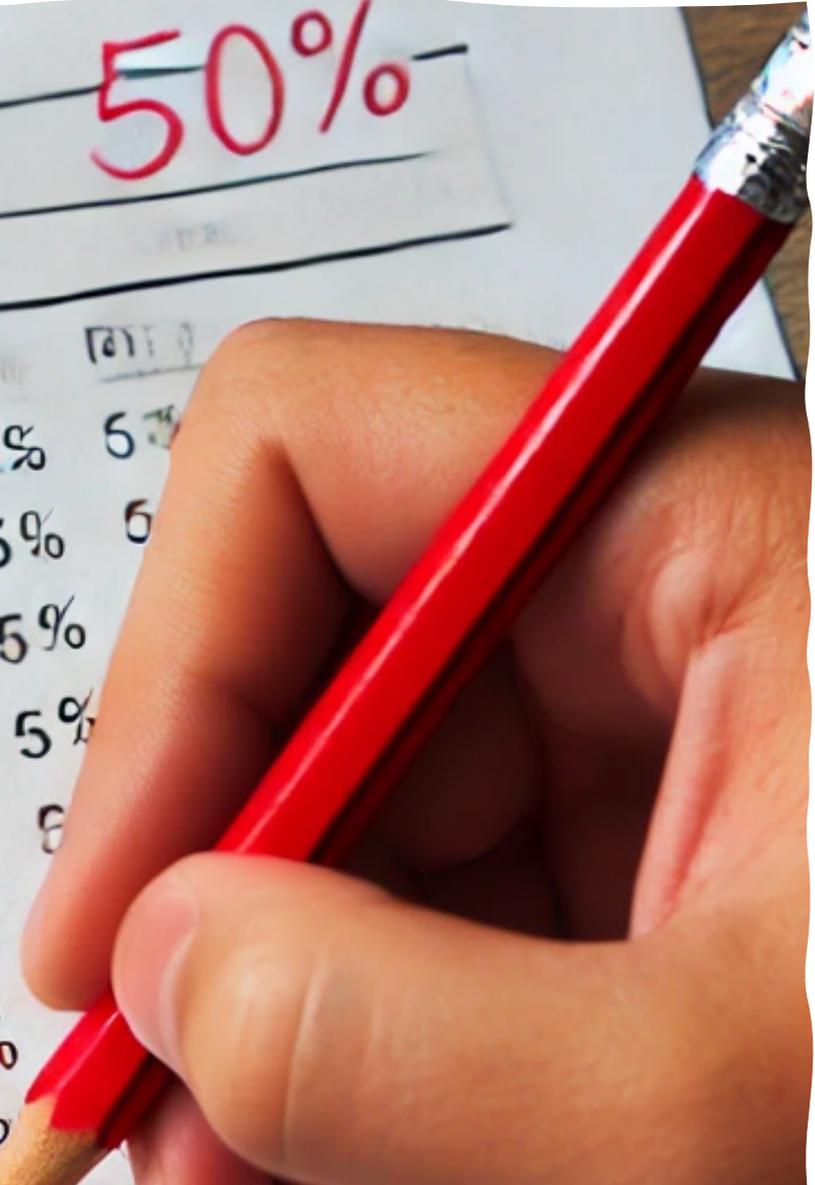
QUIZ PAPER

50%

QUIZ

1	Boonvateletoosentvoot	5%	5%	5%
1	Coontaan noretessan)	5%	5%	5%
A	Goedofert theertenee)	5%	5%	5%
2	Goederobstvarlantise)	5%	5%	5%
2	Googha beedersomee)	5%	5%	5%
2	Pleersto bon suraten)	5%	5%	5%
A	Bohce lal potestona)	5%	5%	5%
	...	5%	5%	5%

F

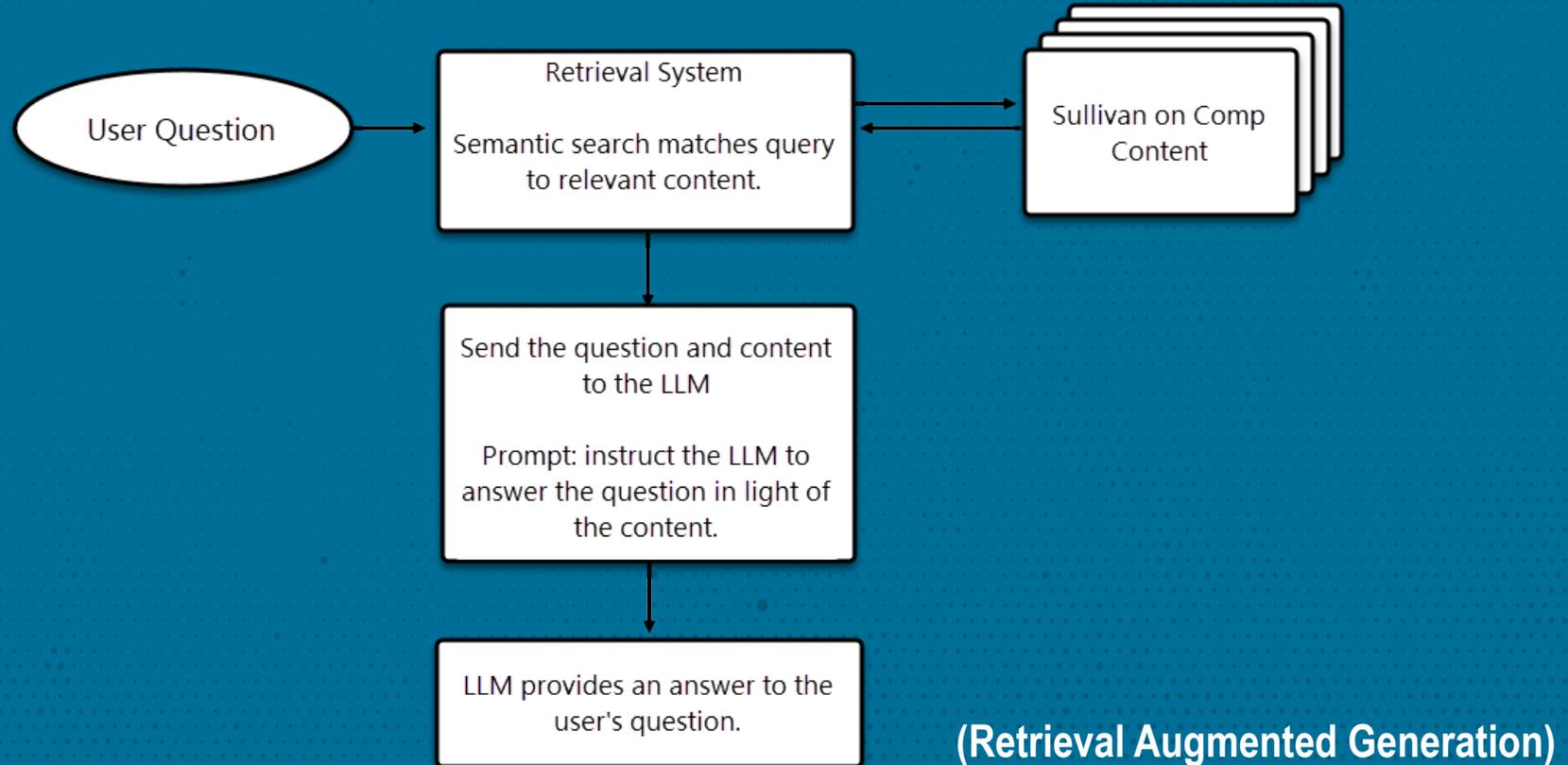




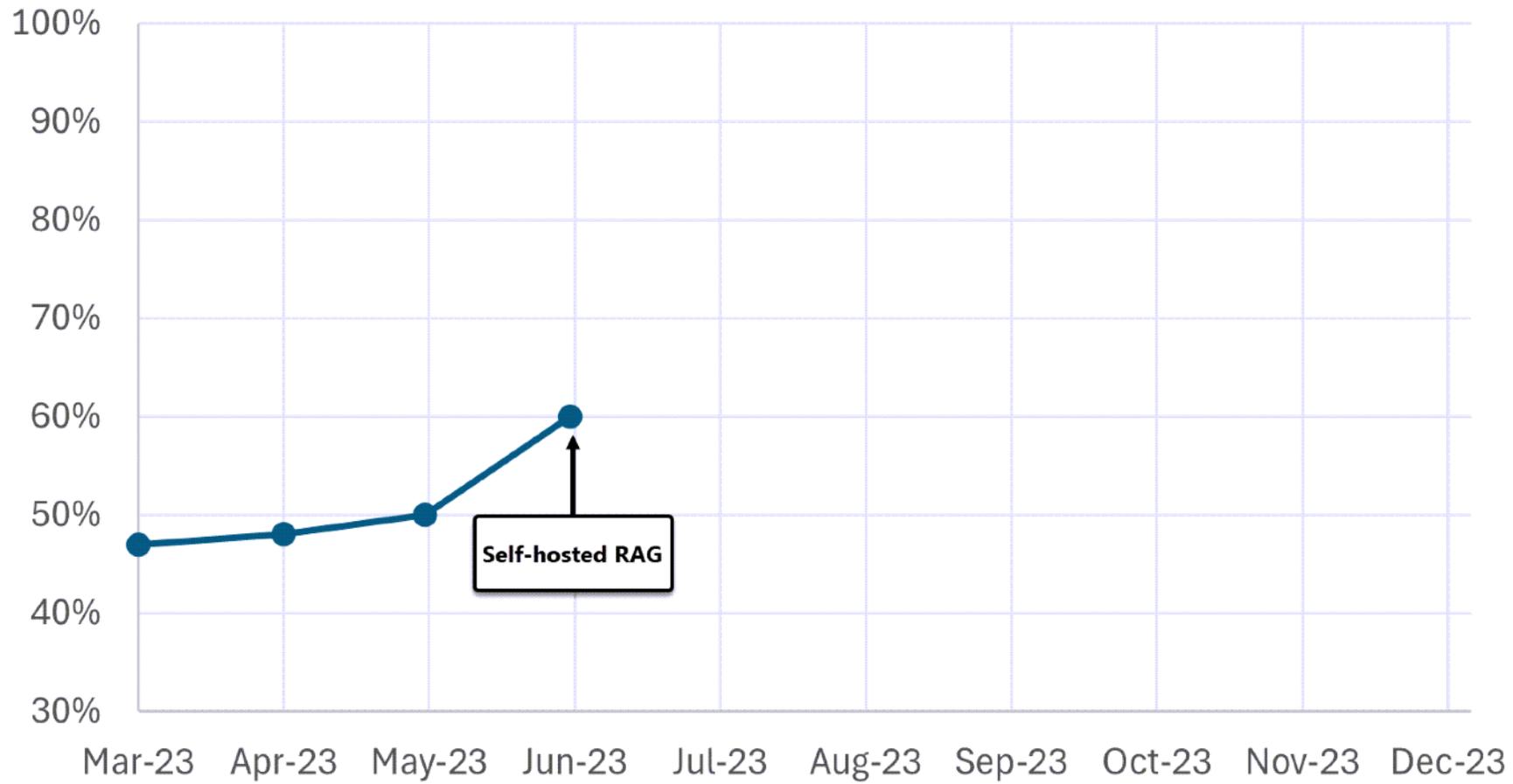
Next steps...

Self-hosted RAG

(LLaMA Vicuna)



Quiz Score



March 2023



[Research](#) [Products](#) [Safety](#) [Company](#)



Enterprise privacy at OpenAI

Trust and privacy are at the core of our mission at OpenAI. We're committed to privacy and security for ChatGPT Team, ChatGPT Enterprise, and our API Platform.

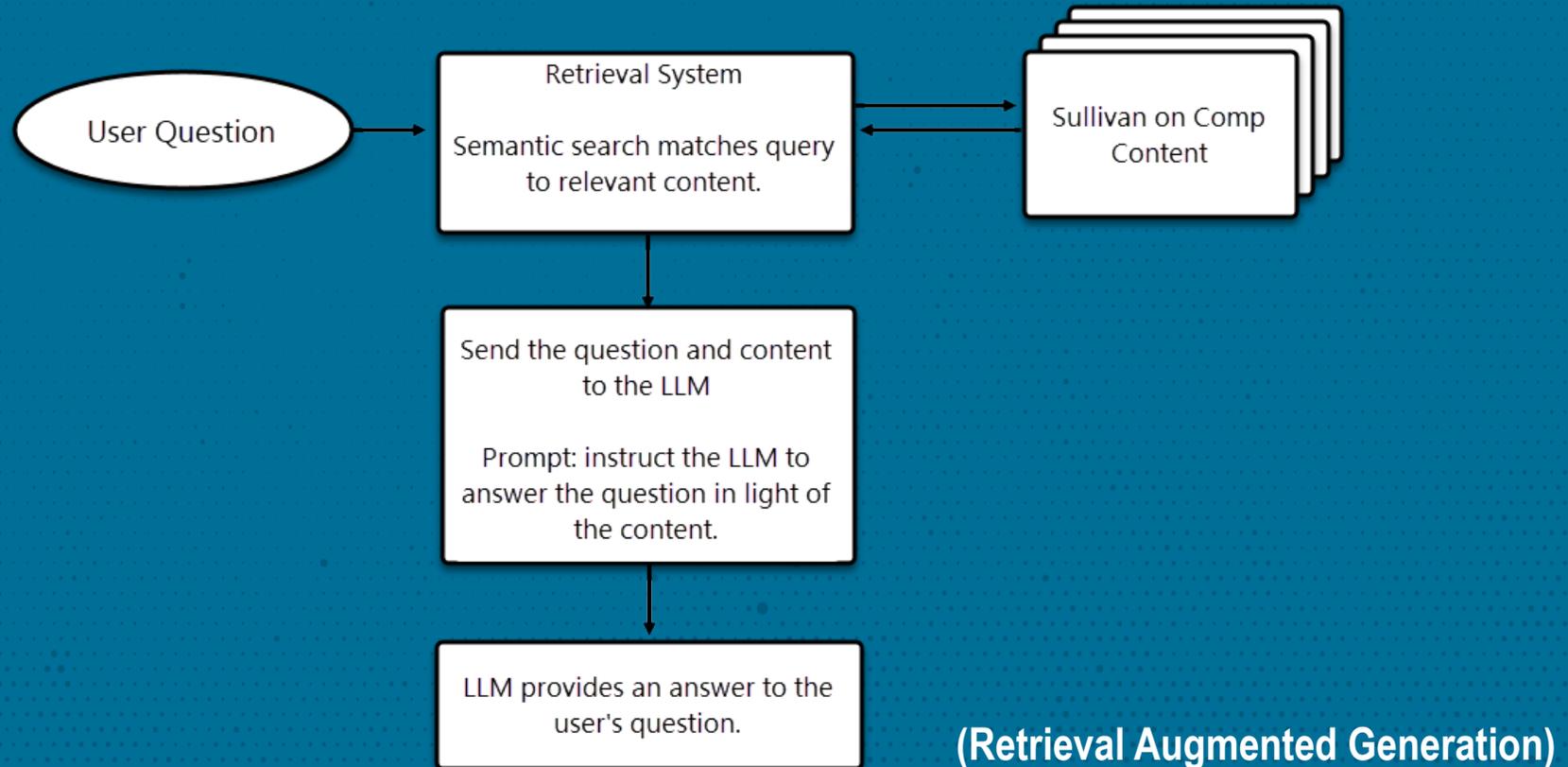
Our commitments

Ownership: You own and control your data

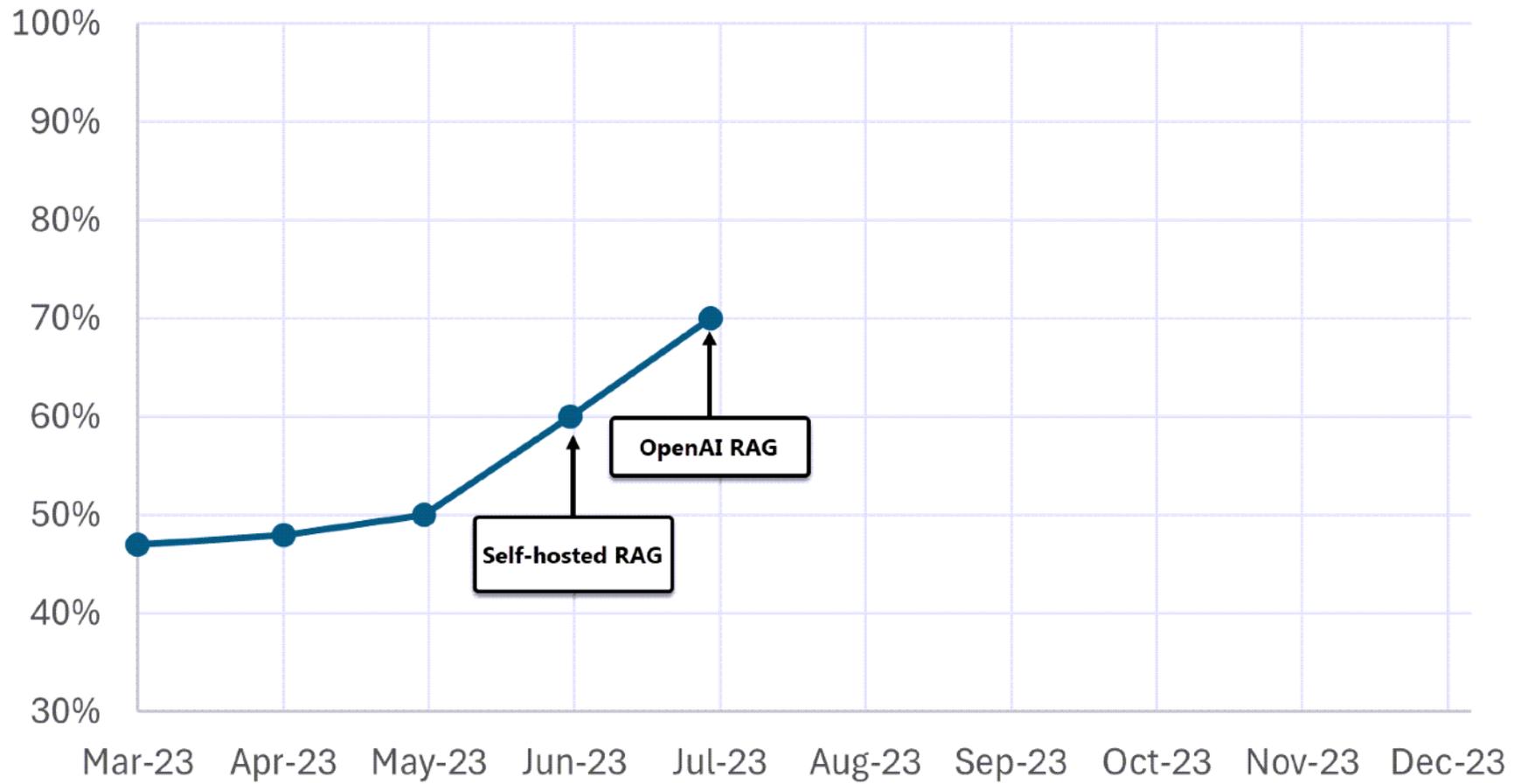
- ✓ We do not train on your business data (data from ChatGPT Team, ChatGPT Enterprise, or our API Platform)
- ✓ You own your inputs and outputs (where allowed by law)
- ✓ You control how long your data is retained (ChatGPT Enterprise)



OpenAI API RAG



Quiz Score



Subjective user feedback

Subjective user feedback

Our content experts

Subjective user feedback

Our content experts



Subjective user feedback

Our content experts



Attorneys at our firm

Subjective user feedback

Our content experts



Attorneys at our firm



Subjective user feedback

Our content experts



Attorneys at our firm



Beta users (adjusters)

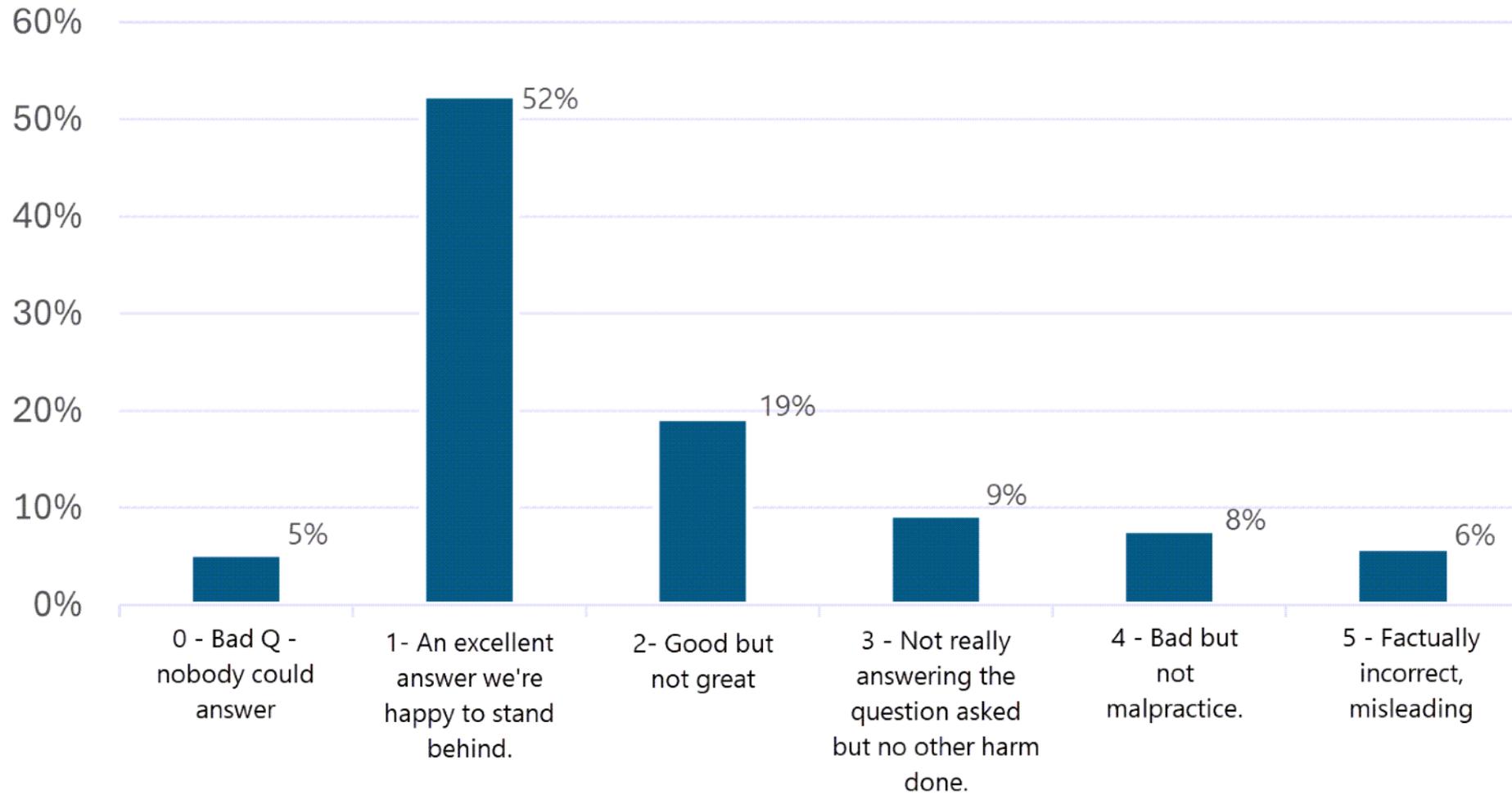
Subjective user feedback

Our content experts 😊

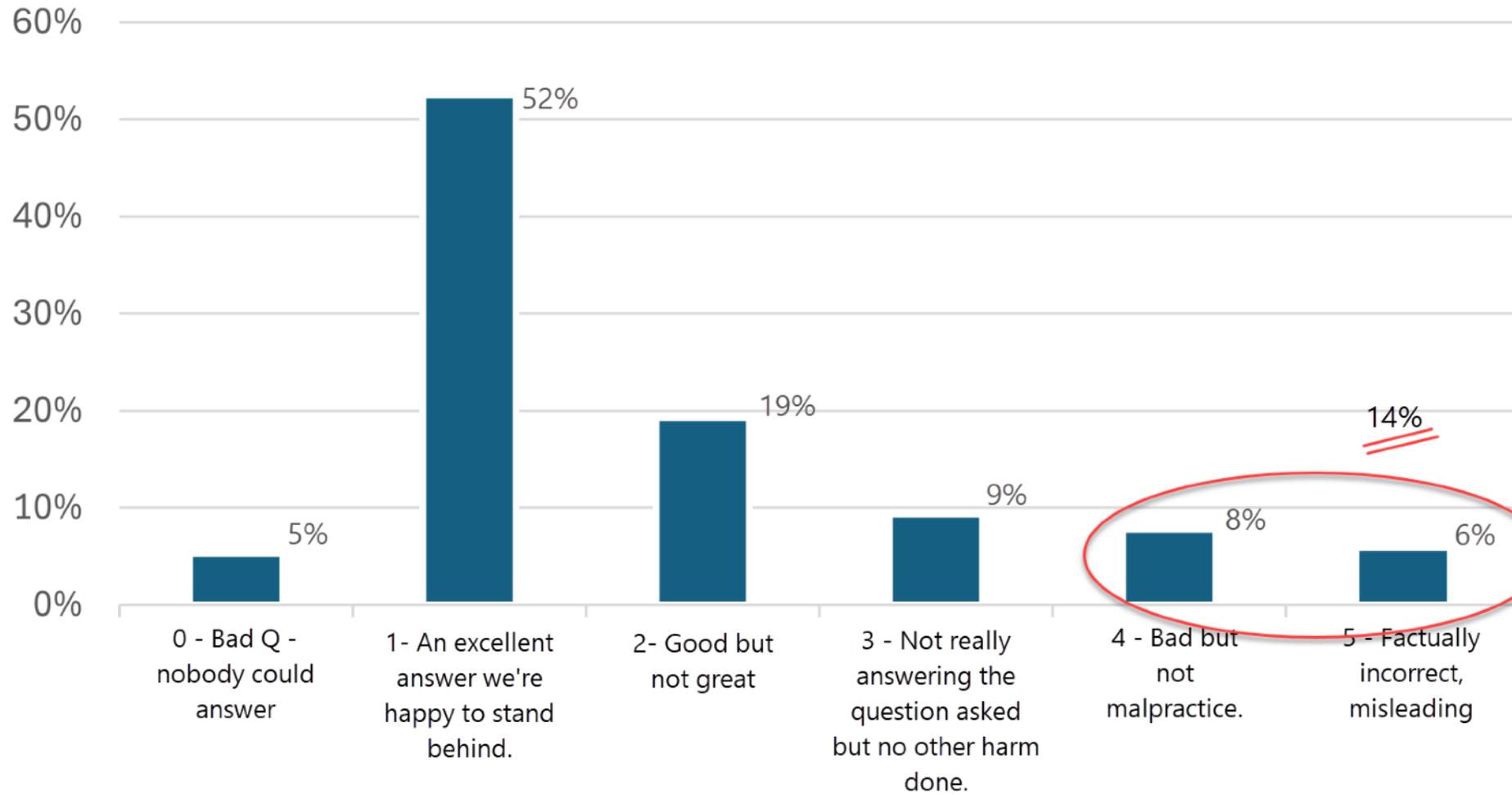
Attorneys at our firm 😐

Beta users (adjusters) 😞

Expert Review of 322 User Qs & AI Answers



Expert Review of 322 User Qs & AI Answers



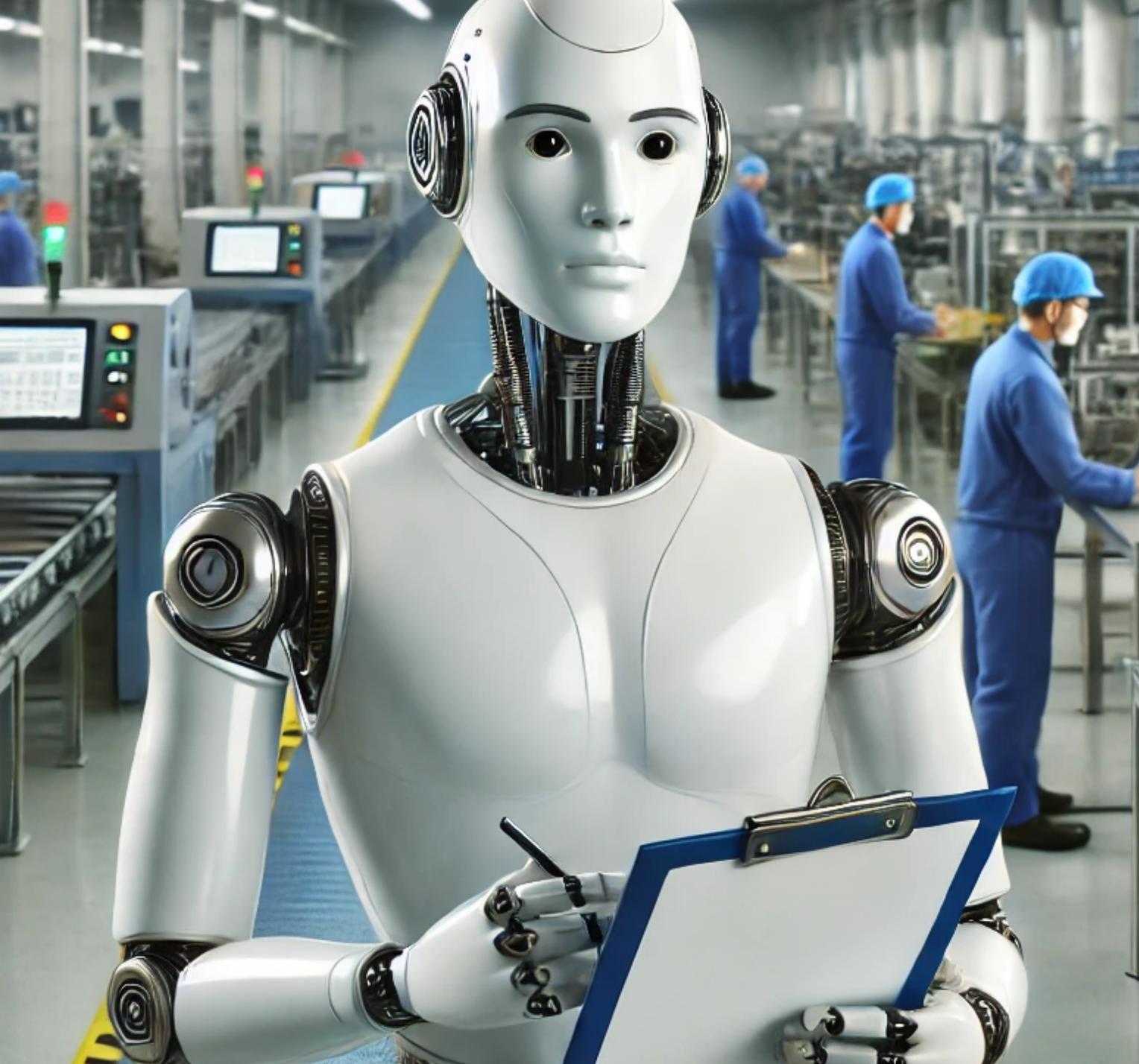


Sample questions from experts

- What are proper grounds for apportionment of permanent disability?
- Utilization review usually decides whether proposed medical care is reasonable and necessary. Are there any times where the WCJ may decide this instead?

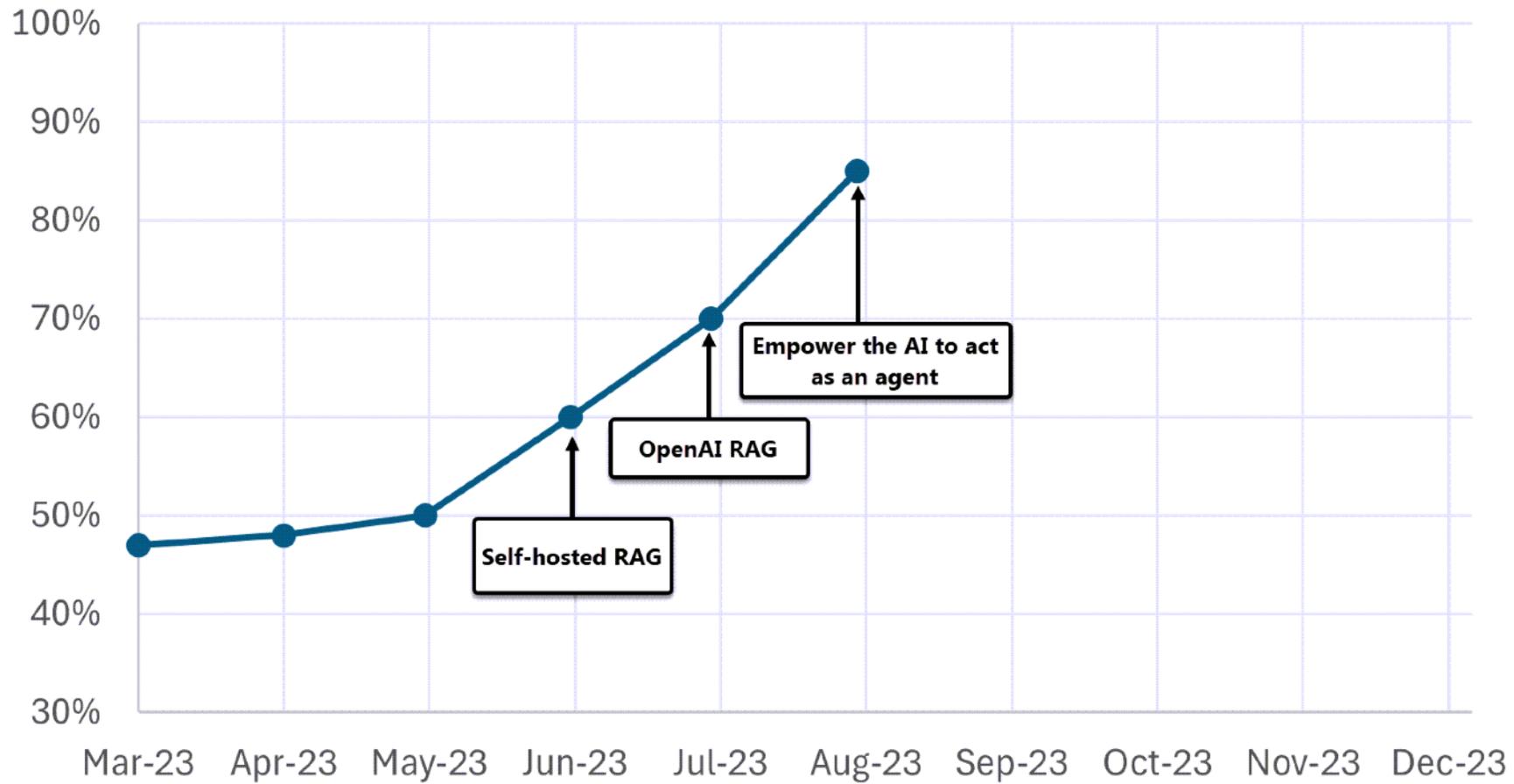
Sample questions from users

- Temporary partially disabled after layoff
- What if the check was mailed to the wrong address?
- What's the SOL for a cumulative trauma claim



**Let's have the
AI interpret the
question before
we try to match
the content!**

Quiz Score



**Time to build our next
benchmark test**

Time to build our next benchmark test

200 easy questions

Time to build our next benchmark test

200 easy questions

200 medium questions

Time to build our next benchmark test

200 easy questions

200 medium questions

200 hard/expert questions

Time to build our next benchmark test

200 easy questions

200 medium questions

200 hard/expert questions

600 total questions

Time to build our next benchmark test

200 easy questions

200 medium questions

200 hard/expert questions

600 total questions

Answered by our top content expert

**How did we use the new
benchmark test?**

How did we use the new benchmark test?

- Make a change to the model

How did we use the new benchmark test?

- Make a change to the model
- Have it answer the 600 user questions

How did we use the new benchmark test?

- Make a change to the model
- Have it answer the 600 user questions
- Have the LLM compare each answer given by the model to the expert's answer and assign a score based on similarity

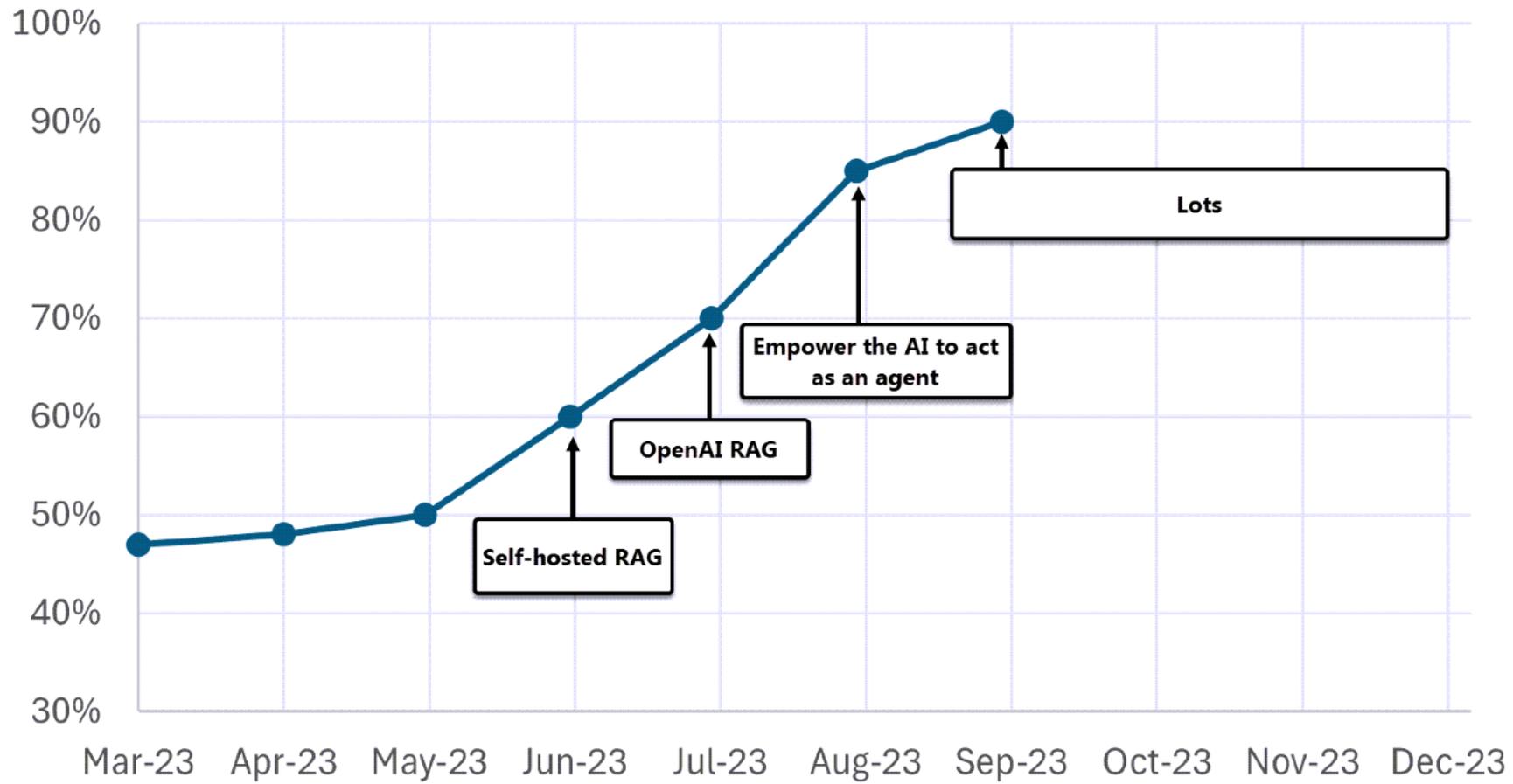
How did we use the new benchmark test?

- Make a change to the model
- Have it answer the 600 user questions
- Have the LLM compare each answer given by the model to the expert's answer and assign a score based on similarity
- Derive a total score for the entire benchmark

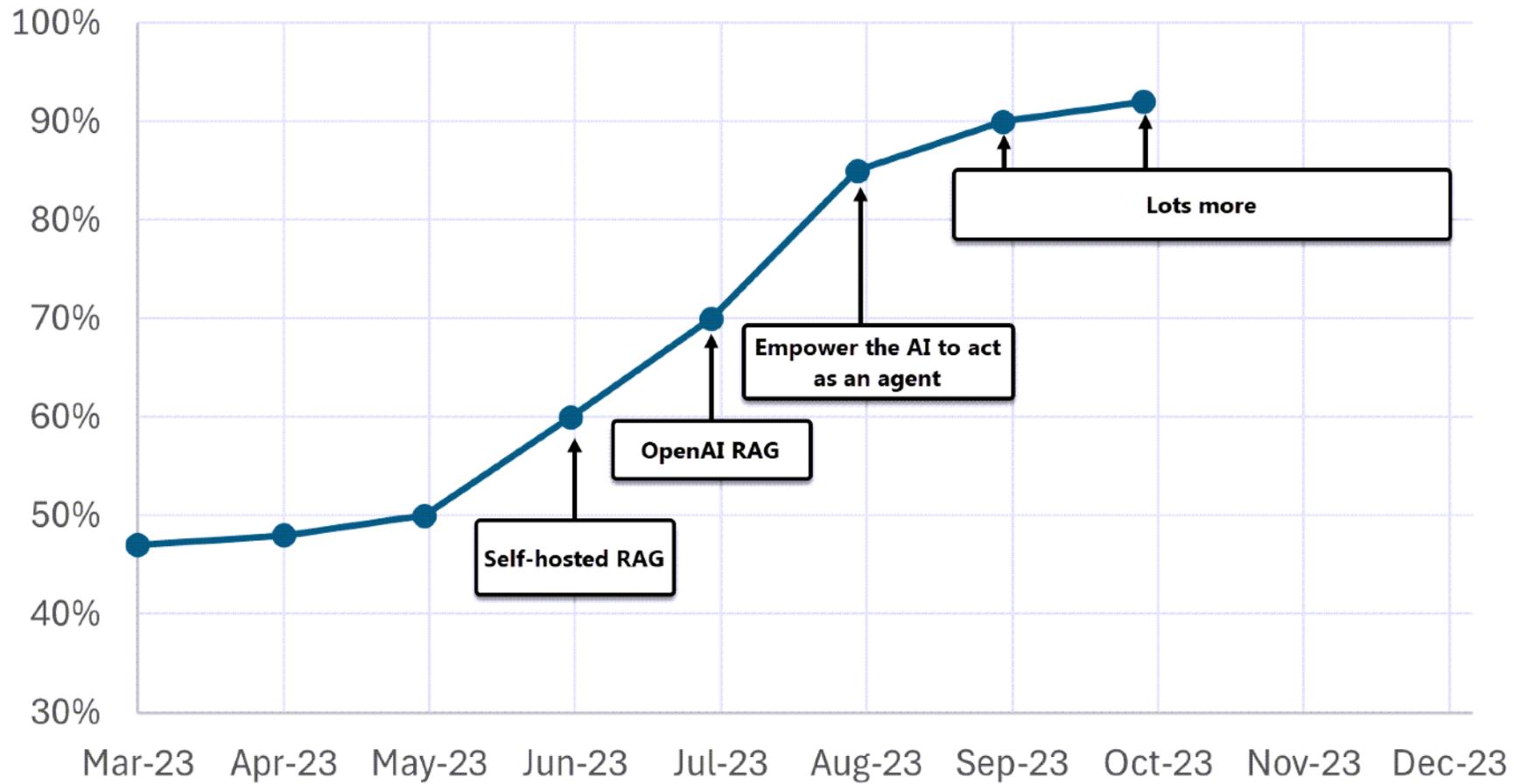
How did we use the new benchmark test?

- Make a change to the model
- Have it answer the 600 user questions
- Have the LLM compare each answer given by the model to the expert's answer and assign a score based on similarity
- Derive a total score for the entire benchmark
- Run the assessment for every iteration

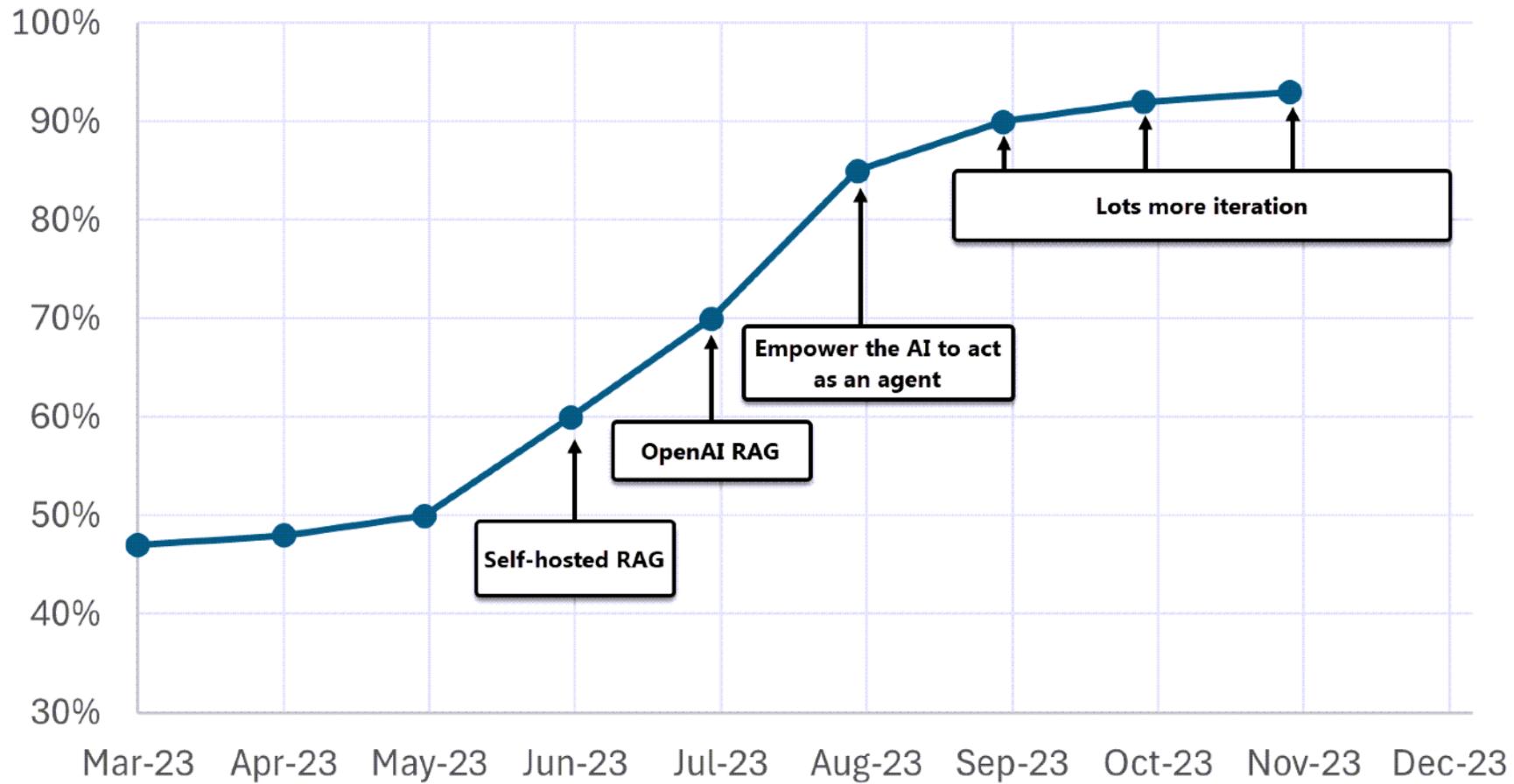
Quiz Score



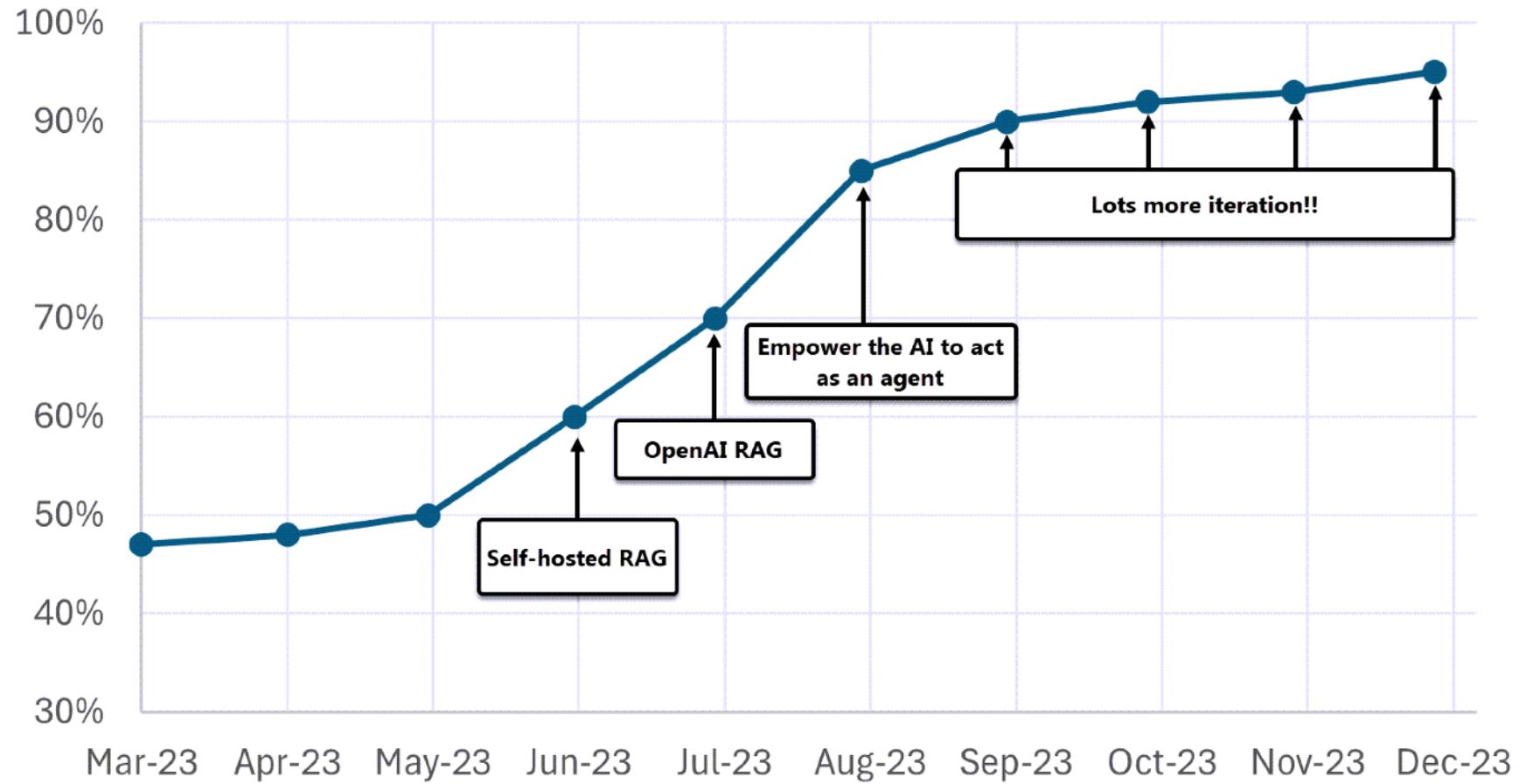
Quiz Score



Quiz Score



Quiz Score



Parameters for iteration

Parameters for iteration

- Language model selection

Parameters for iteration

- Language model selection
- Embeddings model optimization

Parameters for iteration

- Language model selection
- Embeddings model optimization
- Keyword search integration

Parameters for iteration

- Language model selection
- Embeddings model optimization
- Keyword search integration
- Prompt engineering

Parameters for iteration

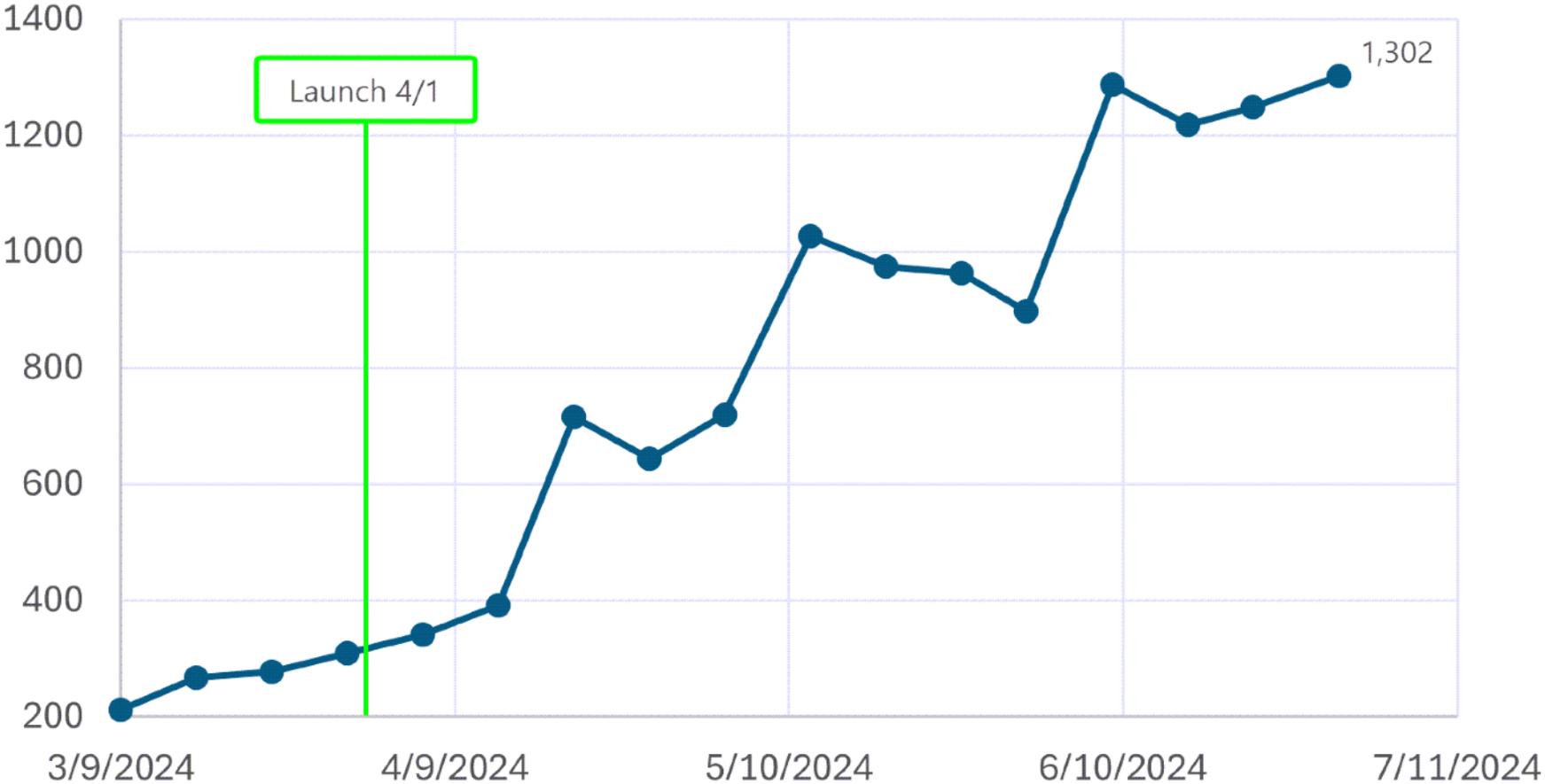
- Language model selection
- Embeddings model optimization
- Keyword search integration
- Prompt engineering
- Evaluation and feedback loops

Parameters for iteration

- Language model selection
- Embeddings model optimization
- Keyword search integration
- Prompt engineering
- Evaluation and feedback loops
- Knowledge base management

**How has it been received
by our users?**

Chats per week



What have we learned?

What have we learned?

- Find a strong developer

What have we learned?

- Find a strong developer
- You can accomplish a lot with a small team

What have we learned?

- Find a strong developer
- You can accomplish a lot with a small team
- Set a clear target

What have we learned?

- Find a strong developer
- You can accomplish a lot with a small team
- Set a clear target 

What have we learned?

- Find a strong developer
- You can accomplish a lot with a small team
- Set a clear target 
- Develop tests to measure your progress

What have we learned?

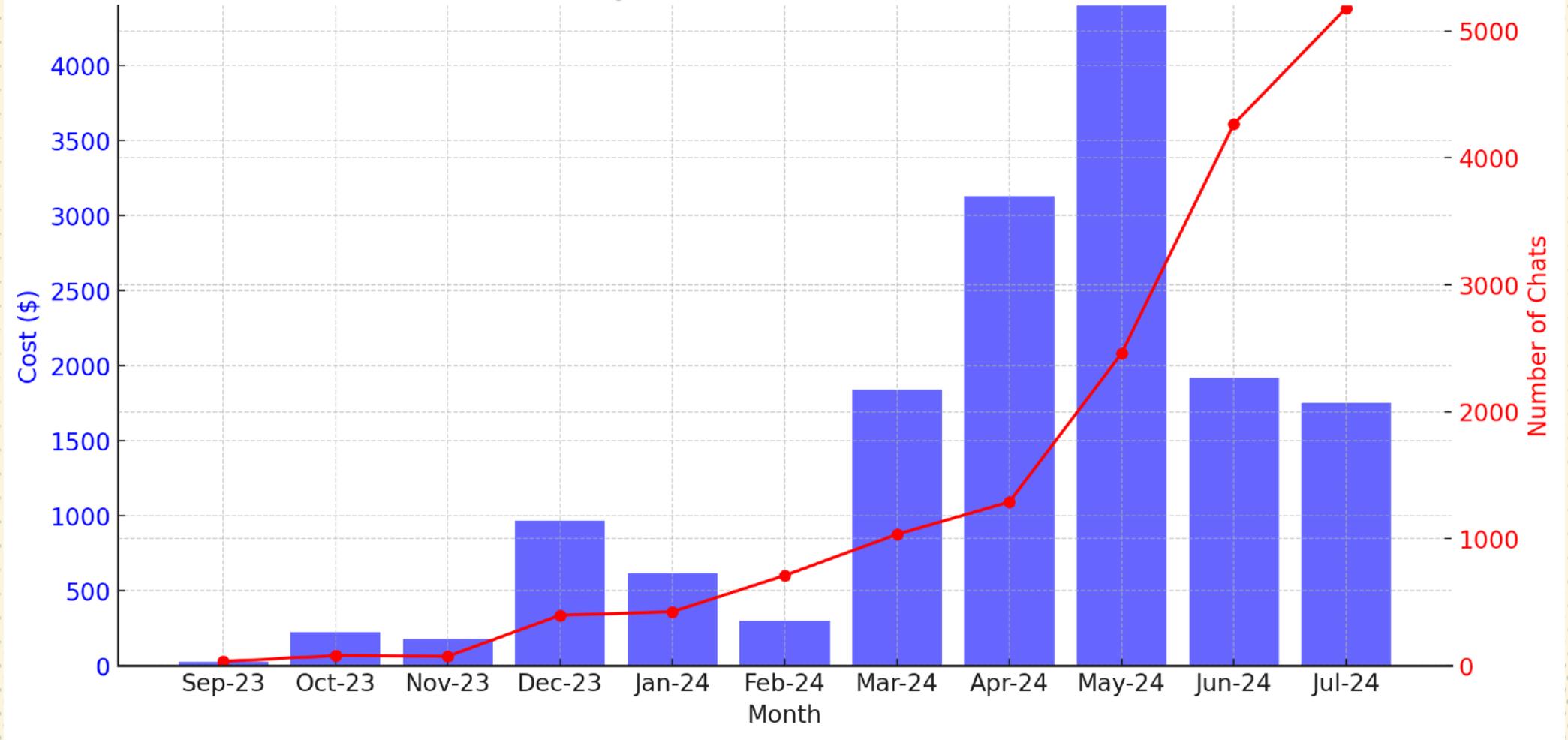
- Find a strong developer
- You can accomplish a lot with a small team
- Set a clear target 
- Develop tests to measure your progress
- Closed models are leading the way

What have we learned?

- Find a strong developer
- You can accomplish a lot with a small team
- Set a clear target
- Develop tests to measure your progress
- Closed models are leading the way
- Don't be daunted by compute costs



Monthly Cost and Number of Chats



***There are no experts,
only pioneers.***

Make it happen!



Presenter

Michael Bell

From Treatise to Tech:
The Development of ChatSOC,
an AI Research Assistant

Michaelbell@sullivanoncomp.com

