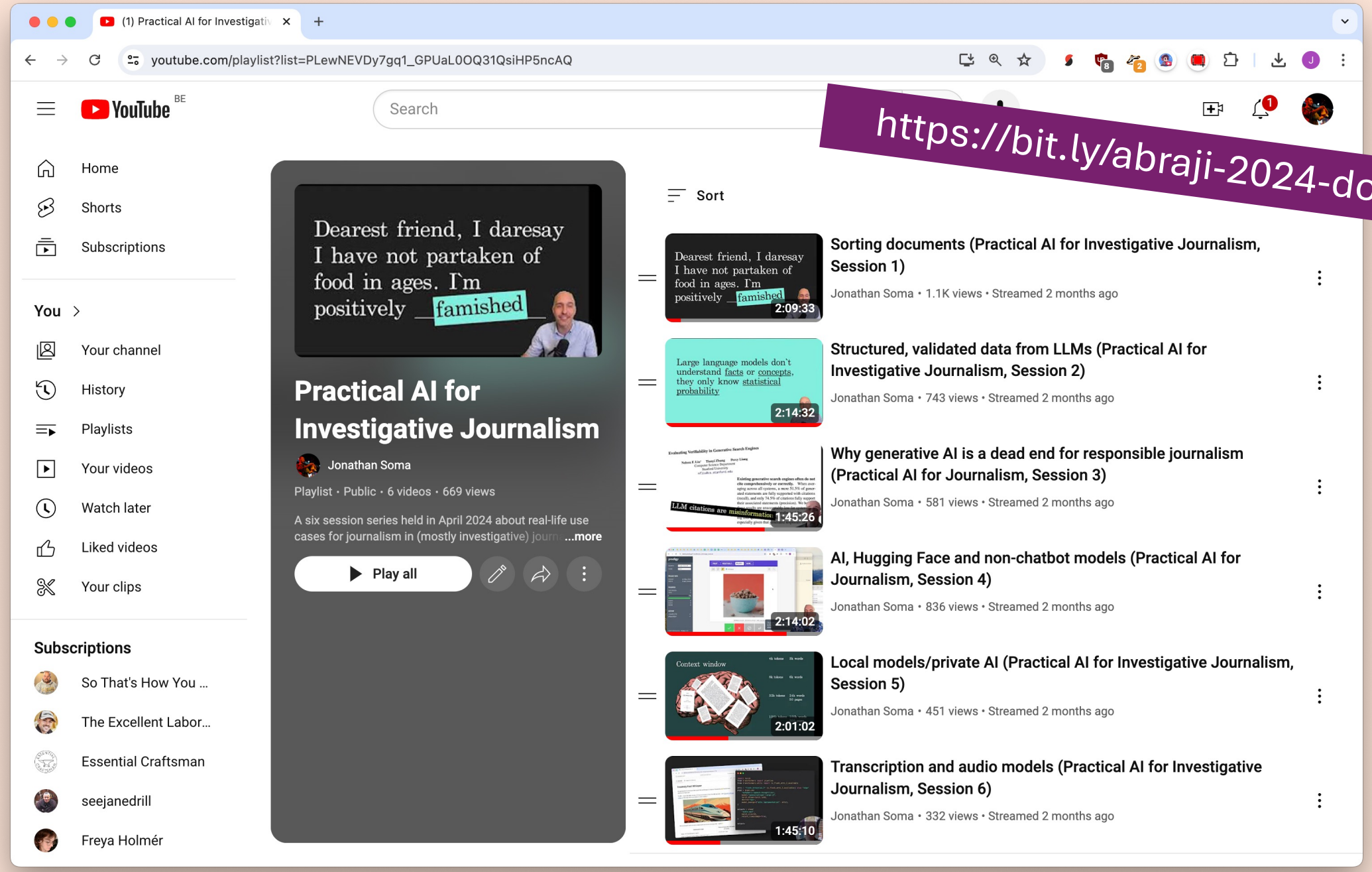
How to investigate documents using AI

https://bit.ly/abraji-2024-docs

Jonathan Soma
Knight Chair in Data Journalism
Columbia Graduate School of Journalism
js4571@columbia.edu • @dangerscarf
https://jonathansoma.com

Hi, I'm Soma!

- Professor at **Columbia's University's Journalism School**, with two data journalism programs
  - Data Journalism MS, a 12-month degree
  - Lede Program, 10-week summer intensive
- Teaching journalists machine learning and artificial intelligence for ten years, long before ChatGPT!
- You can just **ignore me and watch YouTube videos**


(1) Practical AI for Investigati
youtube.com/playlist?list=PLewNEVDy7gq1_GPUaL0OQ31QsiHP5ncAQ
YouTube BE
Search
https://bit.ly/abraji-2024-docs
Home
Shorts
Subscriptions
You
Your channel
History
Playlists
Your videos
Watch later
Liked videos
Your clips
Subscriptions
So That's How You ...
The Excellent Labor...
Essential Craftsman
seejanedrill
Freya Holmér
Dearest friend, I daresay I have not partaken of food in ages. I'm positively famished
Practical AI for Investigative Journalism
Jonathan Soma
Playlist · Public · 6 videos · 669 views
A six session series held in April 2024 about real-life use cases for journalism in (mostly investigative) journ ...more
Play all
Sort
Sorting documents (Practical AI for Investigative Journalism, Session 1)
Jonathan Soma • 1.1K views • Streamed 2 months ago
2:09:33
Structured, validated data from LLMs (Practical AI for Investigative Journalism, Session 2)
Large language models don't understand facts or concepts, they only know statistical probability
Jonathan Soma • 743 views • Streamed 2 months ago
2:14:32
Why generative AI is a dead end for responsible journalism (Practical AI for Journalism, Session 3)
LLM citations are misinformation
Jonathan Soma • 581 views • Streamed 2 months ago
1:45:26
AI, Hugging Face and non-chatbot models (Practical AI for Journalism, Session 4)
Jonathan Soma • 836 views • Streamed 2 months ago
2:14:02
Local models/private AI (Practical AI for Investigative Journalism, Session 5)
Context window
Jonathan Soma • 451 views • Streamed 2 months ago
2:01:02
Transcription and audio models (Practical AI for Investigative Journalism, Session 6)
Jonathan Soma • 332 views • Streamed 2 months ago
1:45:10

I actually don't trust A.I.

A.I. is too stupid to be trusted

We should be very surprised every time A.I. gives an accurate answer to a question

Text inviting
ChatGPT can make mistakes. Check important info.
Message ChatGPT
ChatGPT can make mistakes. Check important info.

You

Write a sonnet about my fa

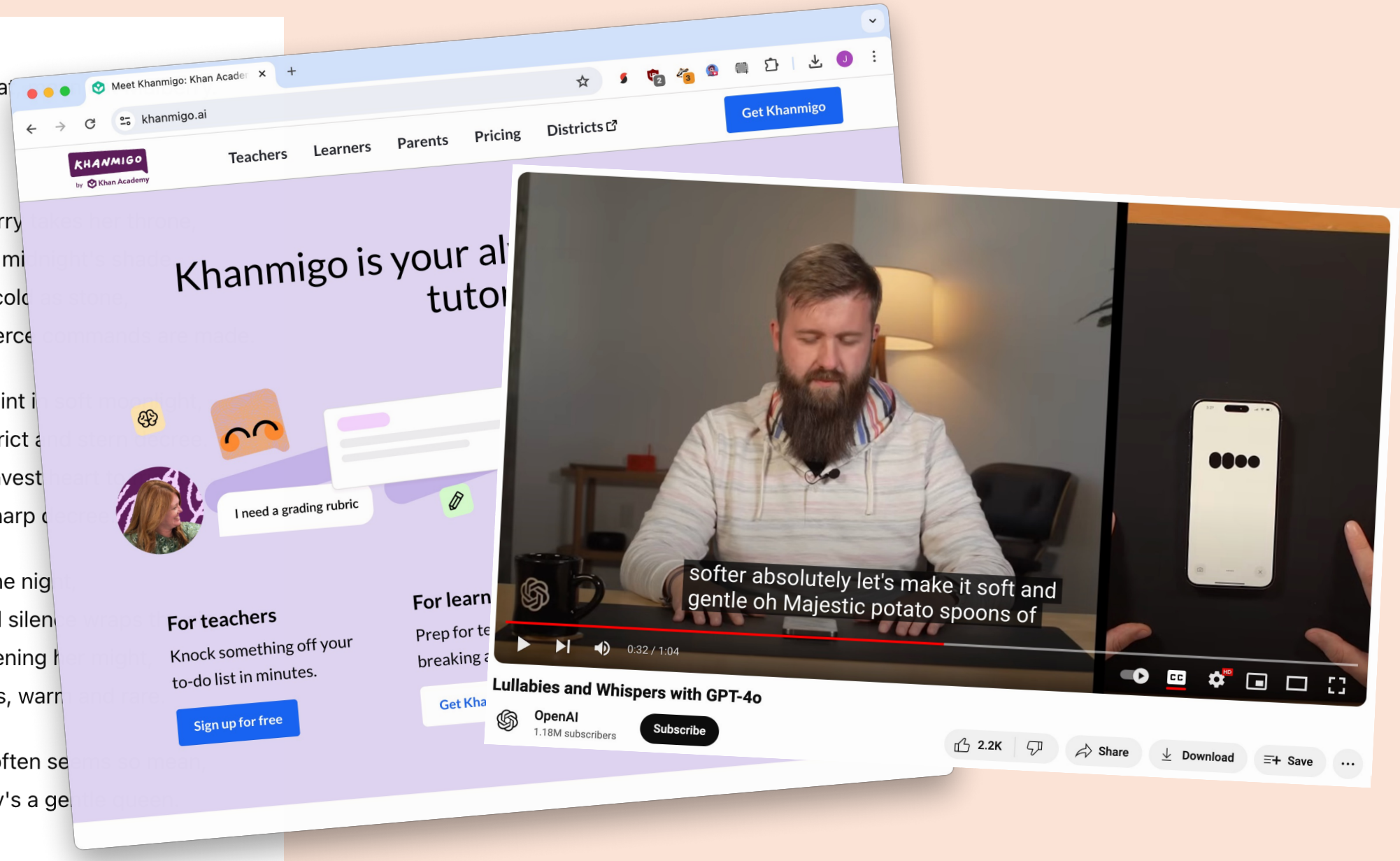
ChatGPT

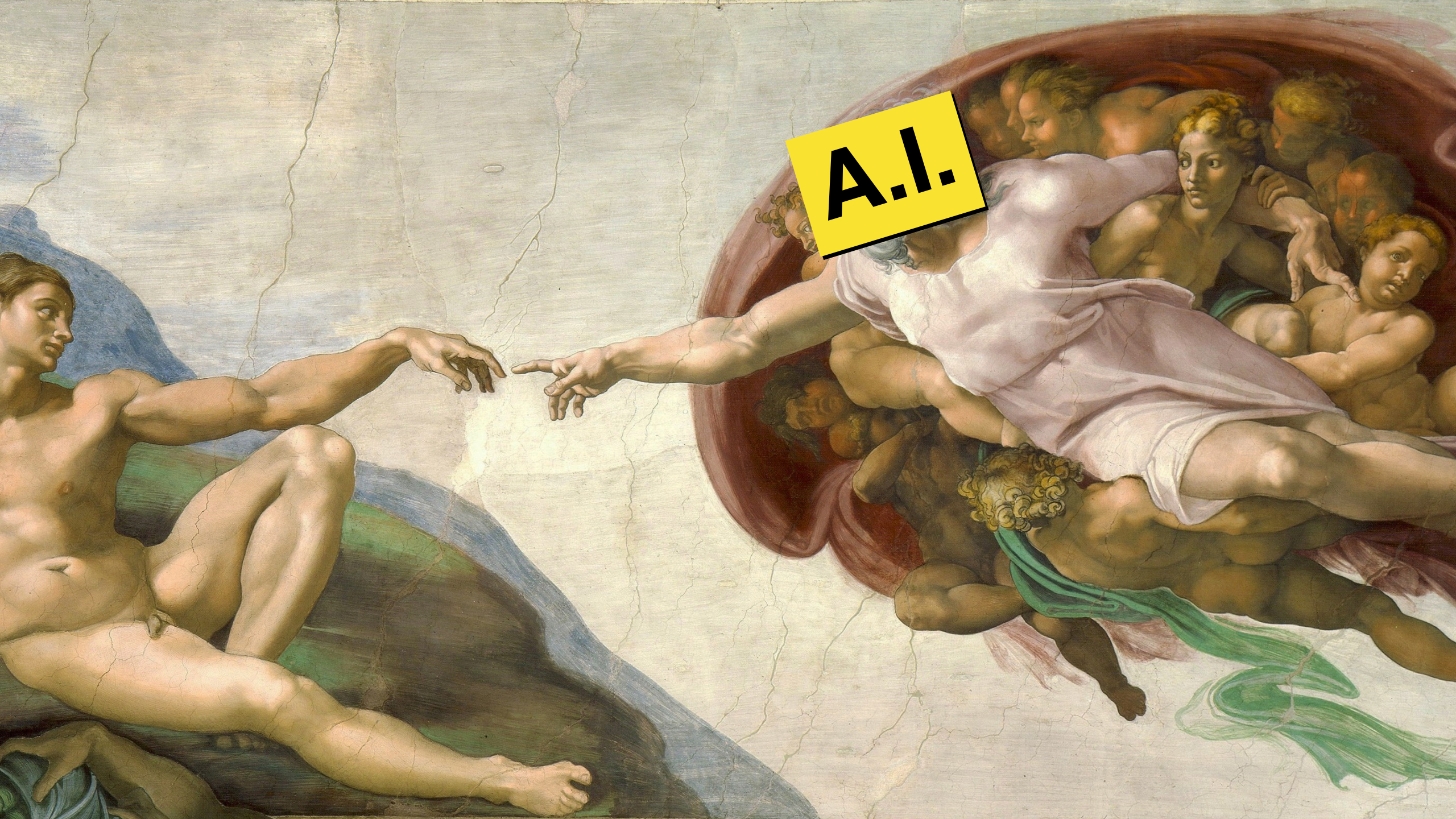
Beneath the moon, Mulberry
A cat of girth, with fur like mi
Her eyes, twin emeralds, cold
Within their depths, her fierce

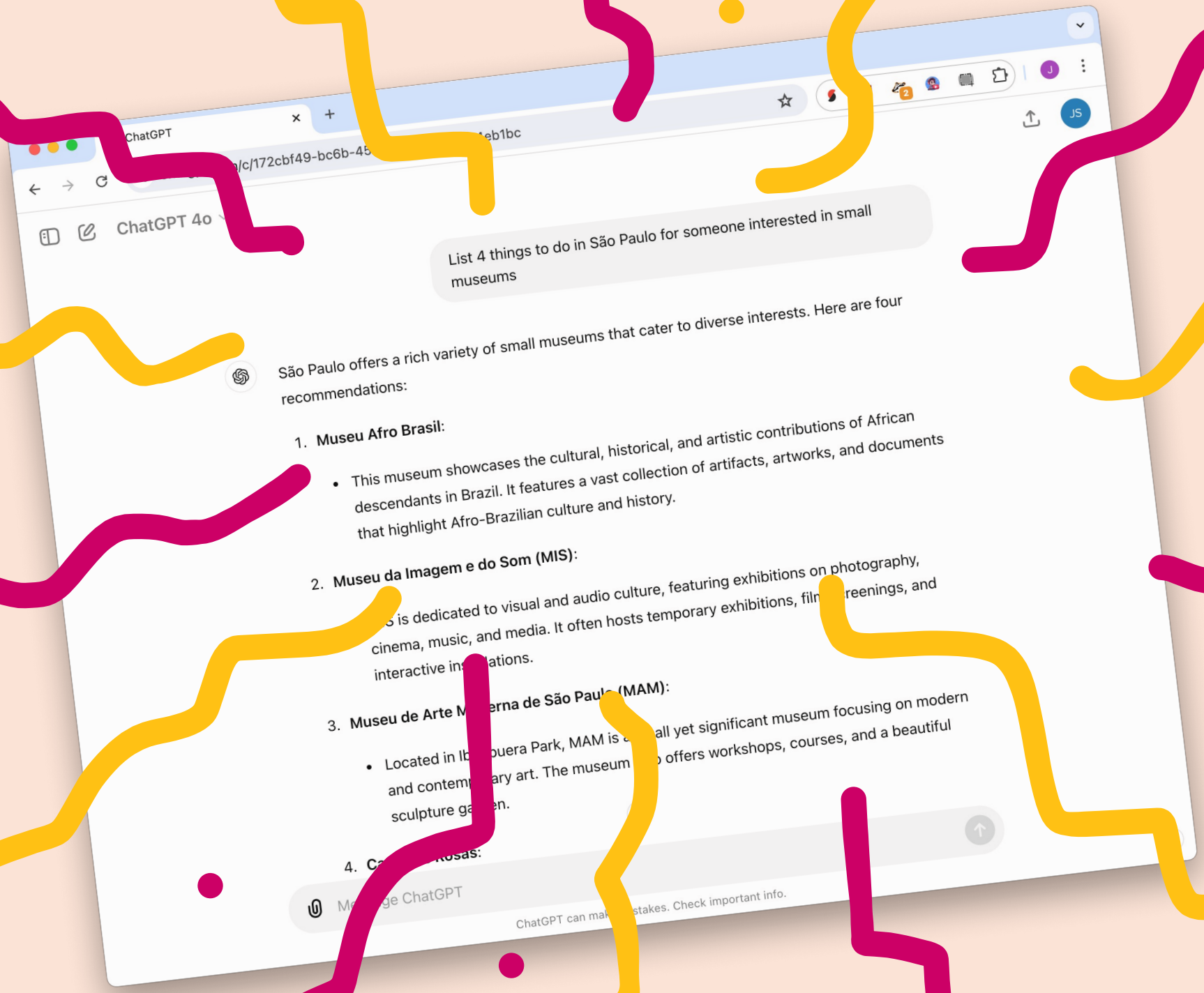
Her claws, like daggers, glint i
A queen who rules with strict a
Her growl can turn the bravest
Her hiss a symphony of sharp

Yet in the quiet hours of the nig
When shadows dance and silen
She curls beside me, softening h
Revealing tender moments, warm

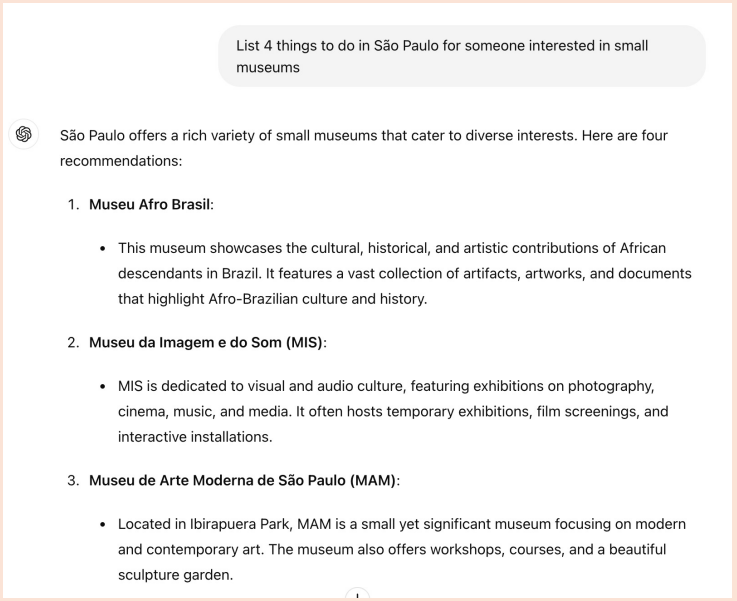
For though she's fat and often se
In dreams, sweet Mulberry's a ge



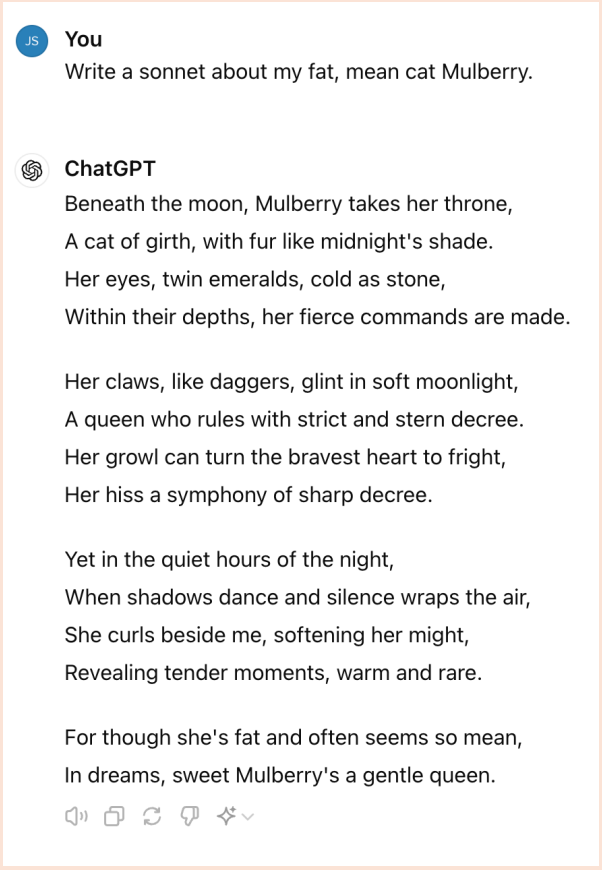

A.I.


ChatGPT
ChatGPT 4o
List 4 things to do in São Paulo for someone interested in small museums
São Paulo offers a rich variety of small museums that cater to diverse interests. Here are four recommendations:
1. Museu Afro Brasil:
This museum showcases the cultural, historical, and artistic contributions of African descendants in Brazil. It features a vast collection of artifacts, artworks, and documents that highlight Afro-Brazilian culture and history.
2. Museu da Imagem e do Som (MIS):
is dedicated to visual and audio culture, featuring exhibitions on photography, cinema, music, and media. It often hosts temporary exhibitions, film reenings, and interactive in lations.
3. Museu de Arte M erna de São Paul (MAM):
Located in Ib uera Park, MAM is a all yet significant museum focusing on modern and contem ary art. The museum offers workshops, courses, and a beautiful sculpture ga en.
4. Ca Rosas:
Message ChatGPT
ChatGPT can mak stakes. Check important info.

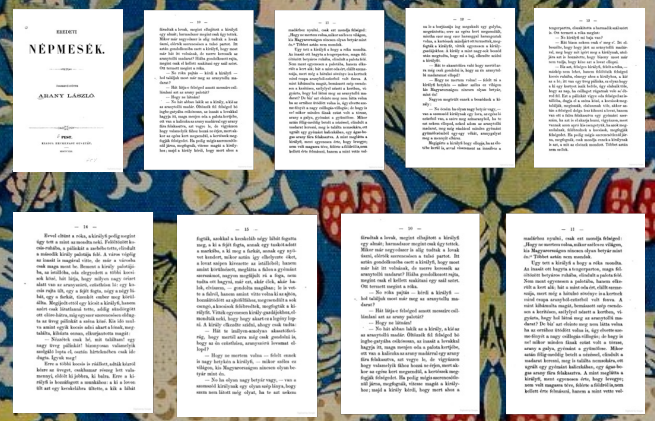
A.I. is *magic*

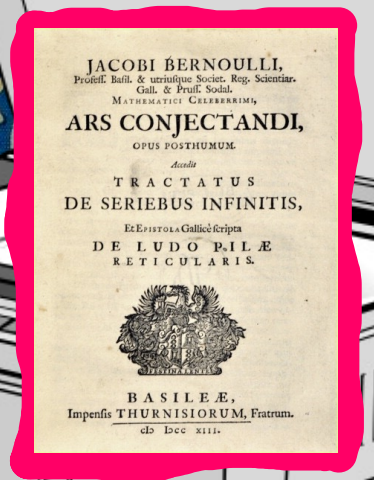

List 4 things to do in São Paulo for someone interested in small museums
São Paulo offers a rich variety of small museums that cater to diverse interests. Here are four recommendations:
1. Museu Afro Brasil:
• This museum showcases the cultural, historical, and artistic contributions of African descendants in Brazil. It features a vast collection of artifacts, artworks, and documents that highlight Afro-Brazilian culture and history.
2. Museu da Imagem e do Som (MIS):
• MIS is dedicated to visual and audio culture, featuring exhibitions on photography, cinema, music, and media. It often hosts temporary exhibitions, film screenings, and interactive installations.
3. Museu de Arte Moderna de São Paulo (MAM):
• Located in Ibirapuera Park, MAM is a small yet significant museum focusing on modern and contemporary art. The museum also offers workshops, courses, and a beautiful sculpture garden.

A.I. is ~~magic~~

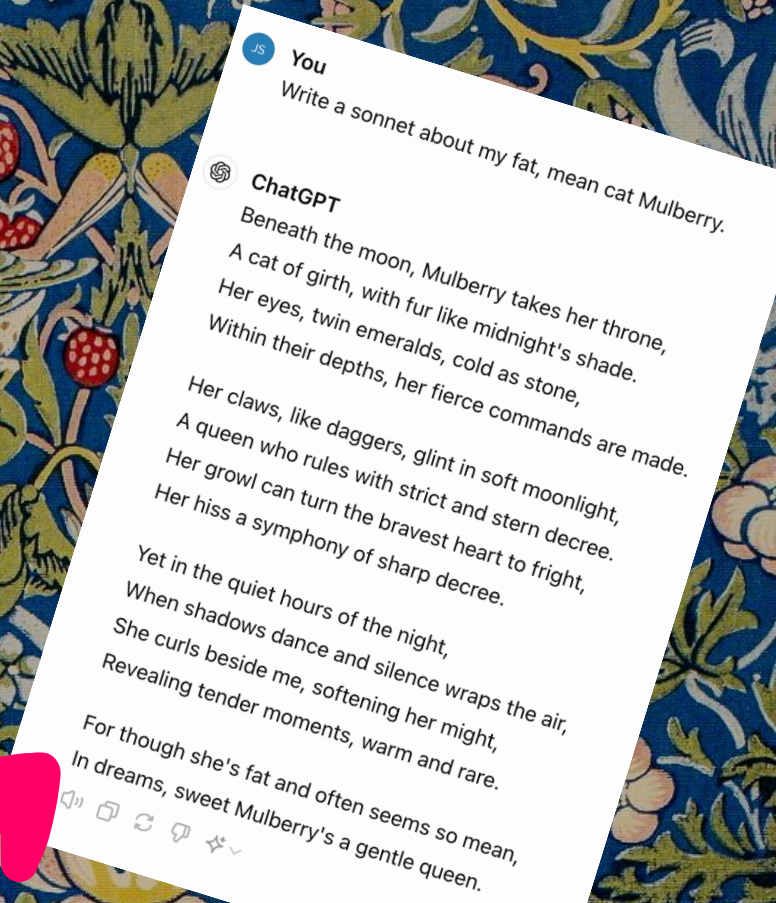

JS
You
Write a sonnet about my fat, mean cat Mulberry.
ChatGPT
Beneath the moon, Mulberry takes her throne,
A cat of girth, with fur like midnight's shade.
Her eyes, twin emeralds, cold as stone,
Within their depths, her fierce commands are made.
Her claws, like daggers, glint in soft moonlight,
A queen who rules with strict and stern decree.
Her growl can turn the bravest heart to fright,
Her hiss a symphony of sharp decree.
Yet in the quiet hours of the night,
When shadows dance and silence wraps the air,
She curls beside me, softening her might,
Revealing tender moments, warm and rare.
For though she's fat and often seems so mean,
In dreams, sweet Mulberry's a gentle queen.

A.I. is
patterns


NÉPMESÉK.
ARANY LÁSZLÓ


JACOBI BERNOULLI,
Profess. Basil. & utriusque Societ. Reg. Scientiar.
Gall. & Prussi Sodal.
MATHEMATICI CELEBERRIMI,
ARS CONJECTANDI,
OPUS POSTHUMUM.
Accedit
TRACTATUS
DE SERIEBUS INFINITIS,
ET EPISTOLA Gallicè scripta
DE LUDO PILÆ
RETICULARIS.
BASILEÆ,
Impensis THURNISIORUM, Fratrum.

How to Create World Wide Web Services
File Edit View Go Bookmarks Options Directory Help
Back Forward Home Reload Images Open Find Stop
Location: http://home.mcom.com/home/how-to-create-
WELCOME TO NETSCAPE
How to Create WWW Services


You
Write a sonnet about my fat, mean cat Mulberry.
ChatGPT
Beneath the moon, Mulberry takes her throne,
A cat of girth, with fur like midnight's shade.
Her eyes, twin emeralds, cold as stone,
Within their depths, her fierce commands are made.
Her claws, like daggers, glint in soft moonlight,
A queen who rules with strict and stern decree.
Her growl can turn the bravest heart to fright,
Her hiss a symphony of sharp decree.
Yet in the quiet hours of the night,
When shadows dance and silence wraps the air,
She curls beside me, softening her might,
Revealing tender moments, warm and rare.
For though she's fat and often seems so mean,
In dreams, sweet Mulberry's a gentle queen.

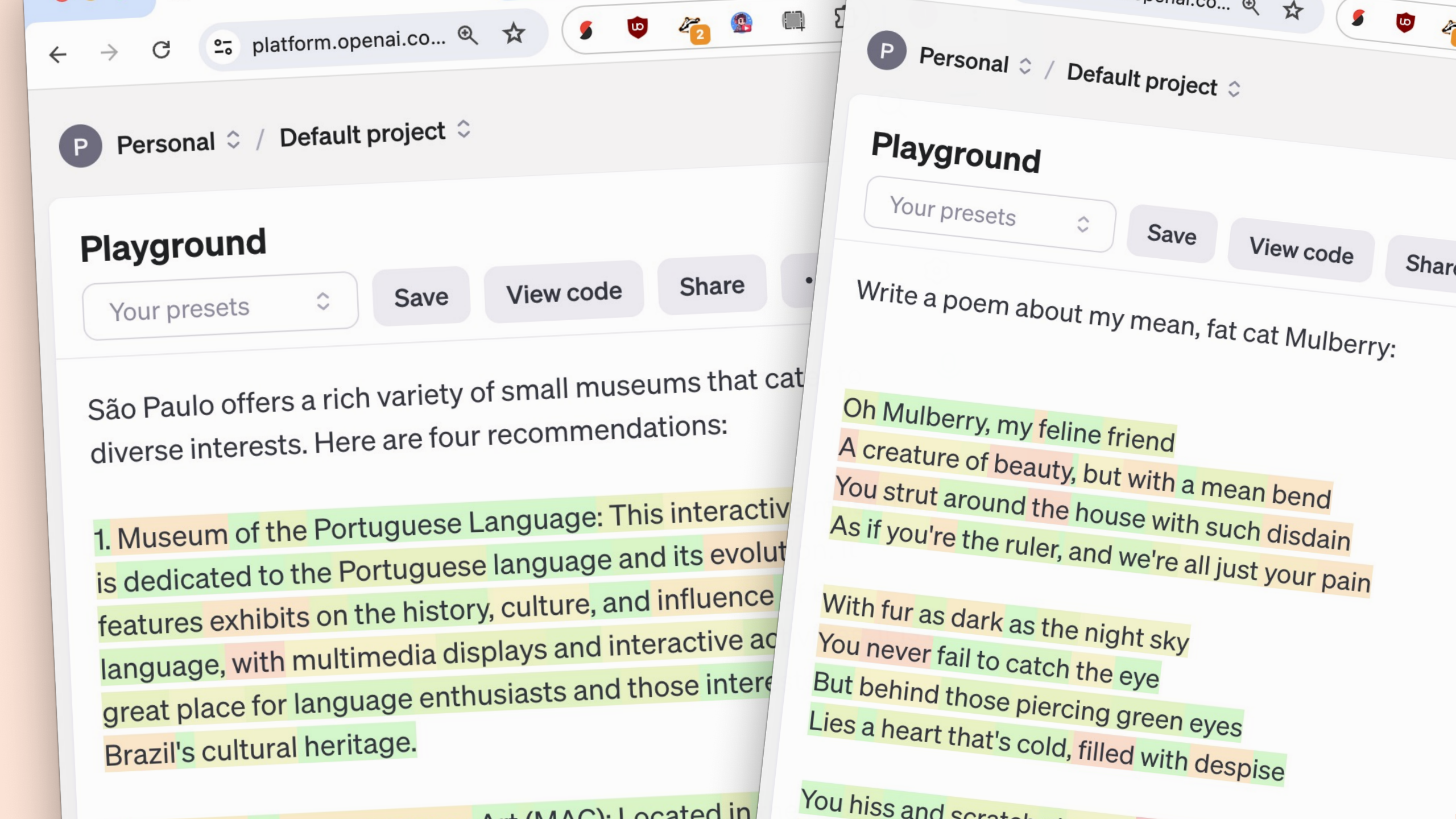
AI does not understand facts or concepts, it only knows statistical probability

Playground - OpenAI API
platform.openai.com/playground/complete
Personal / Default project
I haven't eaten since yesterday, I'm starving.
Fine-tuning
Batches
Storage
Usage
API keys
Forum
Help
Completion models are no
Try our latest models
Submit
starving = 30.05%
so = 23.83%
really = 9.07%
feeling = 5.12%
hungry = 2.97%
not = 2.59%

Não como desde ontem, estou

Save View code Share

Querido amigo, ouso dizer que não me alimento há eras. Estou positivamente


platform.openai.co...
Personal / Default project
Playground
Your presets
Save
View code
Share
São Paulo offers a rich variety of small museums that cat
diverse interests. Here are four recommendations:
1. Museum of the Portuguese Language: This interactiv
is dedicated to the Portuguese language and its evolut
features exhibits on the history, culture, and influence
language, with multimedia displays and interactive ac
great place for language enthusiasts and those intere
Brazil's cultural heritage.
Personal / Default project
Playground
Your presets
Save
View code
Write a poem about my mean, fat cat Mulberry:
Oh Mulberry, my feline friend
A creature of beauty, but with a mean bend
You strut around the house with such disdain
As if you're the ruler, and we're all just your pain
With fur as dark as the night sky
You never fail to catch the eye
But behind those piercing green eyes
Lies a heart that's cold, filled with despise

AI does not understand facts or concepts, it only knows statistical probability

Chatbots May 'Hallucinate' More Often Than Many Realize

When summarizing facts, ChatGPT technology makes things up about 3 percent of the time, according to research from a new start-up. A Google system's rate was 27 percent.

Share full article

By Cade Metz

Cade Metz has been watching chatbots h

Published Nov. 6, 2023 Updated Nov. 16, 2023

Hughes Hallucination Evaluation Model (HHEM) leaderboard

This leaderboard (by Vectara) evaluates how often an LLM introduces hallucinations when summarizing a document.

T	Model	Hallucination Rate
🟢	openai/GPT-4-Turbo	3
🟢	openai/GPT-4	3
?	microsoft/Orca-2-13b	3.2
🟢	openai/GPT-3.5-Turbo	3.5
🟢	openai/GPT-4-Pro	4.8
		5.1
		5.6

A.I. should be used for
error-resilient tasks

AcaraJazz® (@o_acarajazz)

instagram.com/p/C3urZcBucwK/

Ajuda nois com a rifa 👇🏾👇🏾

o_acarajazz • Follow
Acarajazz - Sabores da Bahia

o_acarajazz Dia meu povo, peça e reserve pelo link da bio. 😉
19w See translation

loretacolucci Olá! Indo no restaurante é outro preço? Obrigada
3d 1 like Reply See translation
View replies (1)

fragmentado.ttt 😍😍😍😍
19w 1 like Reply
View replies (1)

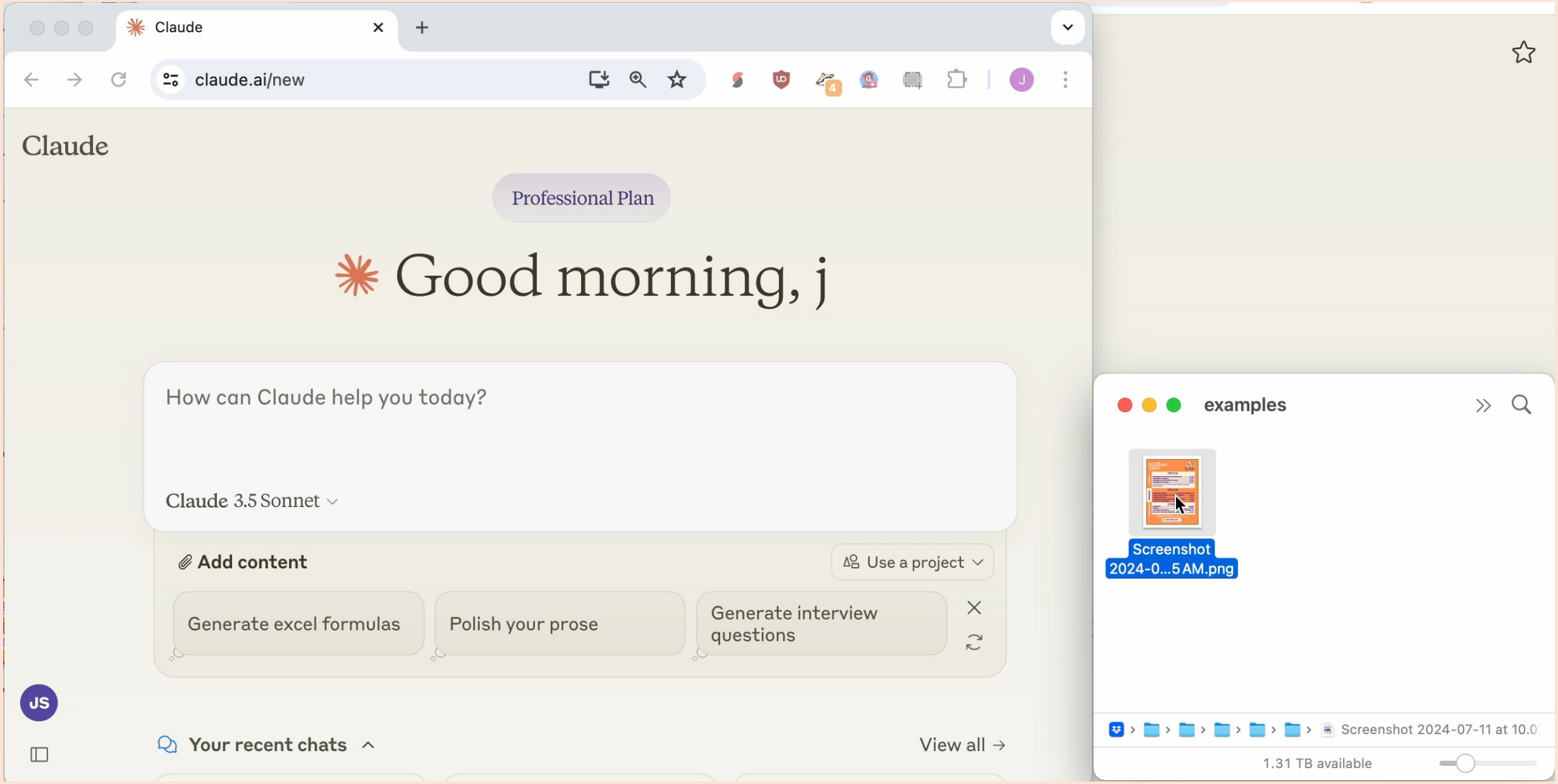
macs.carol Tem acarajé vegetariano ou vegano?
1w 1 like Reply See translation
View replies (1)

guxtavotamura É o mesmo cardapio pra almocar ai?
19w Reply See translation
View replies (1)

jafamg A chuva na hora do almoço não me impediu de devorar a moqueca de frutos do mar do Acarajazz.....q delícia.....uma manjar para deuses!!! Vida longa para o Acarajazz ❤️❤️❤️❤️❤️

Liked by **argentinosmurf** and **others**
February 24

Add a comment... Post


Claude
claude.ai/new
Claude
Professional Plan
Good morning, j
How can Claude help you today?
Claude 3.5 Sonnet
Add content
Use a project
Generate excel formulas
Polish your prose
Generate interview questions
JS
Your recent chats
View all
examples
Screenshot 2024-0...5 AM.png
Screenshot 2024-07-11 at 10.0
1.31 TB available

Brazilian Restaurant Delivery

claude.ai/chat/5804fabb-fc5c-4730-9db4-90bdfd63e929

Claude

Brazilian Restaurant Delivery Menu

JS "Acarajé, abará, moquecas, compotas, licores... Essas e outras delícias do nordeste." what's the first few words mean

Let me translate and explain those first few words for you:

"Acarajé, abará, moquecas, compotas, licores..."

These are all types of food and drink, primarily associated with northeastern Brazilian cuisine:

1. Acarajé - A deep-fried ball made from mashed black-eyed peas, typically filled with shrimp and other ingredients.
2. Abará - Similar to acarajé, but steamed rather than fried, often wrapped in banana leaves.
3. Moquecas - Traditional Brazilian fish stews.
4. Compotas - Fruit preserves or jams.
5. Licores - Liqueurs, often made with local fruits.

The phrase as a whole is listing some of the specialties and delicacies of northeastern Brazilian cuisine that the restaurant likely serves. The ellipsis (...) at the end suggests there are more items beyond these examples.

The full sentence "Essas e outras delícias do nordeste" means "These and other delights from the Northeast," indicating that these items represent the northeastern Brazilian culinary traditions offered by the restaurant.

Copy Retry

JS is northeast food bahian?

Reply to Claude...

Claude 3.5 Sonnet

Brazilian Restaurant Delivery

claude.ai/chat/5804fabb-fc5c-4730-9db4-90bdfd63e929

Claude

Brazilian Restaurant Delivery Menu

JS is northeast food bahian?

Northeast Brazilian cuisine and Bahian cuisine are related but not exactly the same:

1. Northeast Brazil (Nordeste) includes nine states: Bahia, Pernambuco, Ceará, Maranhão, Paraíba, Rio Grande do Norte, Alagoas, Sergipe, and Piauí.
2. Bahia is one state within the Northeast region, known for its particularly distinctive and influential cuisine.
3. Bahian cuisine is a subset of Northeastern cuisine, but it's often considered the most famous and representative of the region due to its unique flavors and African influences.
4. Many dishes mentioned (like acarajé, abará, and moqueca) are strongly associated with Bahian cuisine specifically.
5. However, other Northeastern states also have their own culinary specialties that may differ from Bahian cuisine.

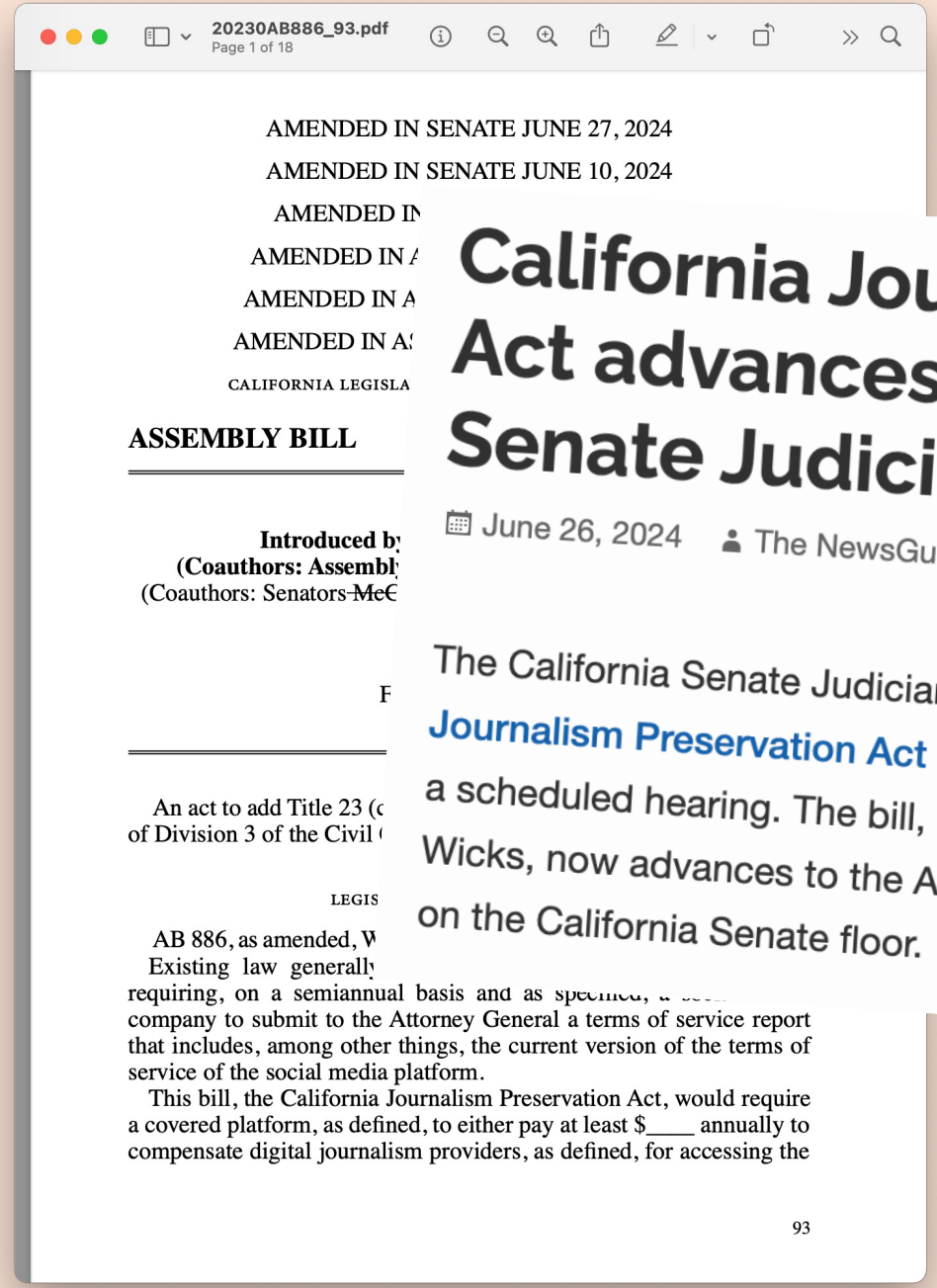
So while Bahian food is Northeastern, not all Northeastern food is specifically Bahian. That said, in popular culture and in many restaurants outside the region, "Northeastern cuisine" often heavily features or focuses on Bahian dishes due to their popularity and distinct flavors.

JS ok this place is bahian, what should i expect and what should i get? i don't speak portugusse so i want to be prepared

Reply to Claude...

Claude 3.5 Sonnet

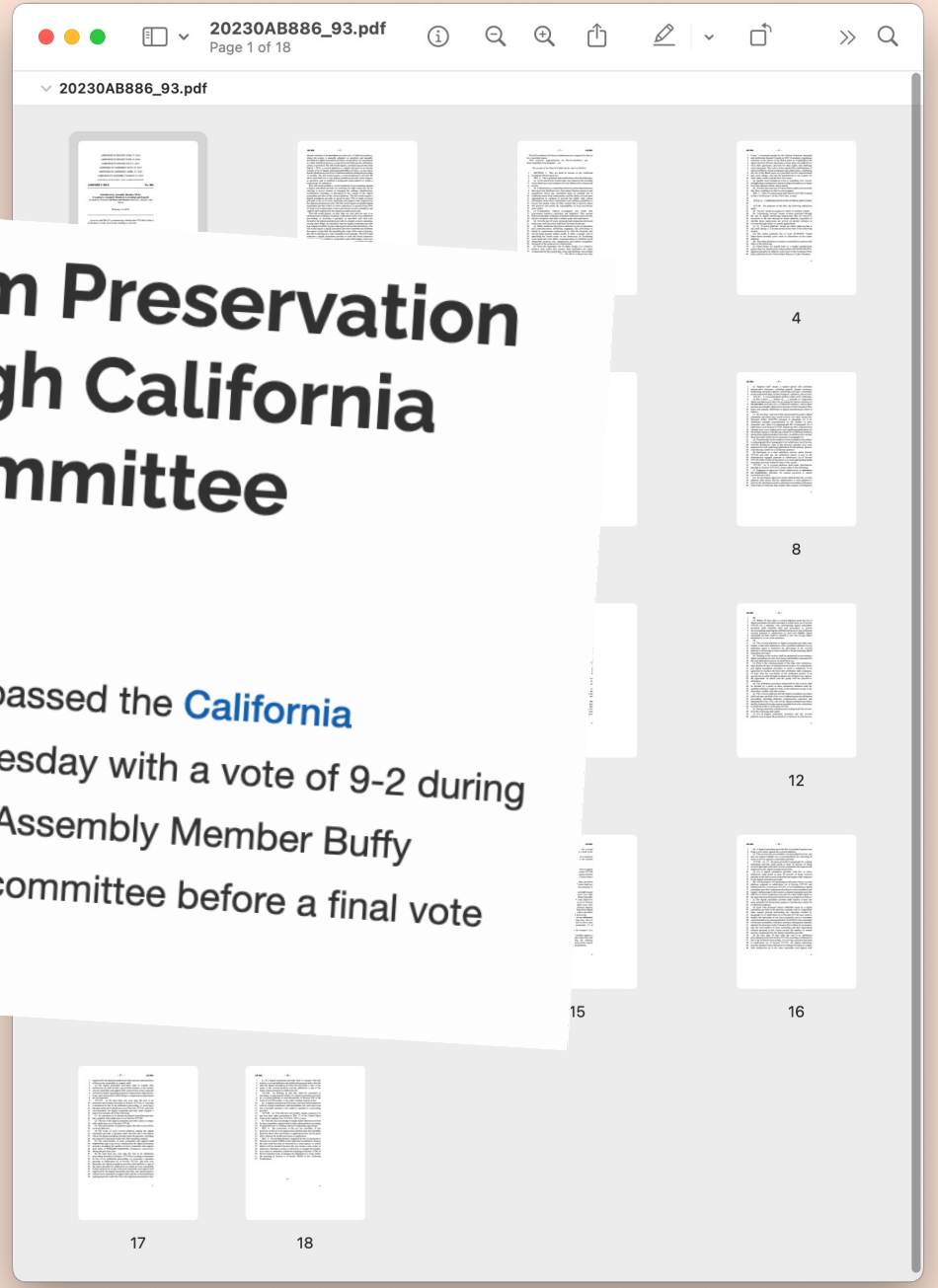
it doesn't matter if everything Claude is telling me is wrong!



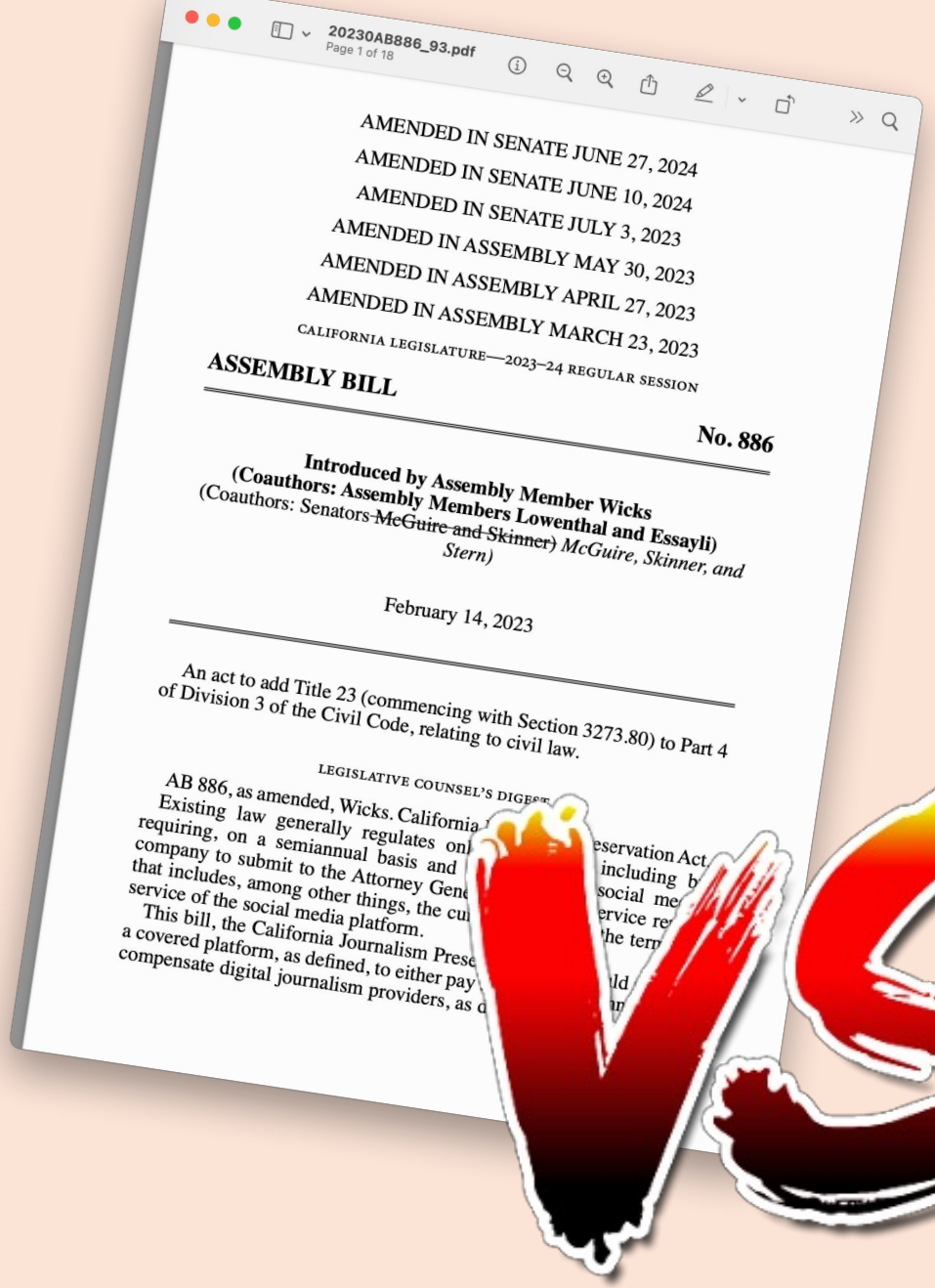
California Journalism Preservation Act advances through California Senate Judiciary Committee

June 26, 2024 The NewsGuild-CWA

The California Senate Judiciary Committee passed the California Journalism Preservation Act (AB 886) on Tuesday with a vote of 9-2 during a scheduled hearing. The bill, introduced by Assembly Member Buffy Wicks, now advances to the Appropriations committee before a final vote on the California Senate floor.



accuracy and facts are the only things important to journalism.



New York Sessions

Year	Session	Introduced	Completed	Legislation	Search	Dataset
2023-2024	General Assembly	20091	3380	Browse	Search	Download

Texas Sessions

Year	Session	Introduced	Completed	Legislation	Search	Dataset
023	88th Legislature 4th Special Session	399	220	Browse	Search	Download
2023	88th Legislature 3rd Special Session	598	292	Browse	Search	Download
2023	88th Legislature 2nd Special Session	169	71	Browse	Search	Download
023	88th Legislature 1st Special Session	117	46	Browse	Search	Download
2023-2024	88th Legislature	11807	4464	Browse	Search	Download

California Sessions

Year	Session	Introduced	Completed	Legislation	Search	Dataset
2023-2024	Regular Session	5287	1198	Browse	Search	Download

When we have many documents, maybe we want to...

Find some examples of similar pieces of news legislation

Find what percent of legislation is about journalism

Get an alert each time journalism legislation is introduced

Find what legislators proposed legislation about journalism

In these situations, maybe a small error rate is okay?

Apple's App Store is full of c

washingtonpost.com/technology/2019/11/22/apple-says-its-app-store-is-safe-trusted-place-we-found-reports-unwanted-sexual-behavior-six-apps-some-tar...

The Washington Post
Democracy Dies in Darkness

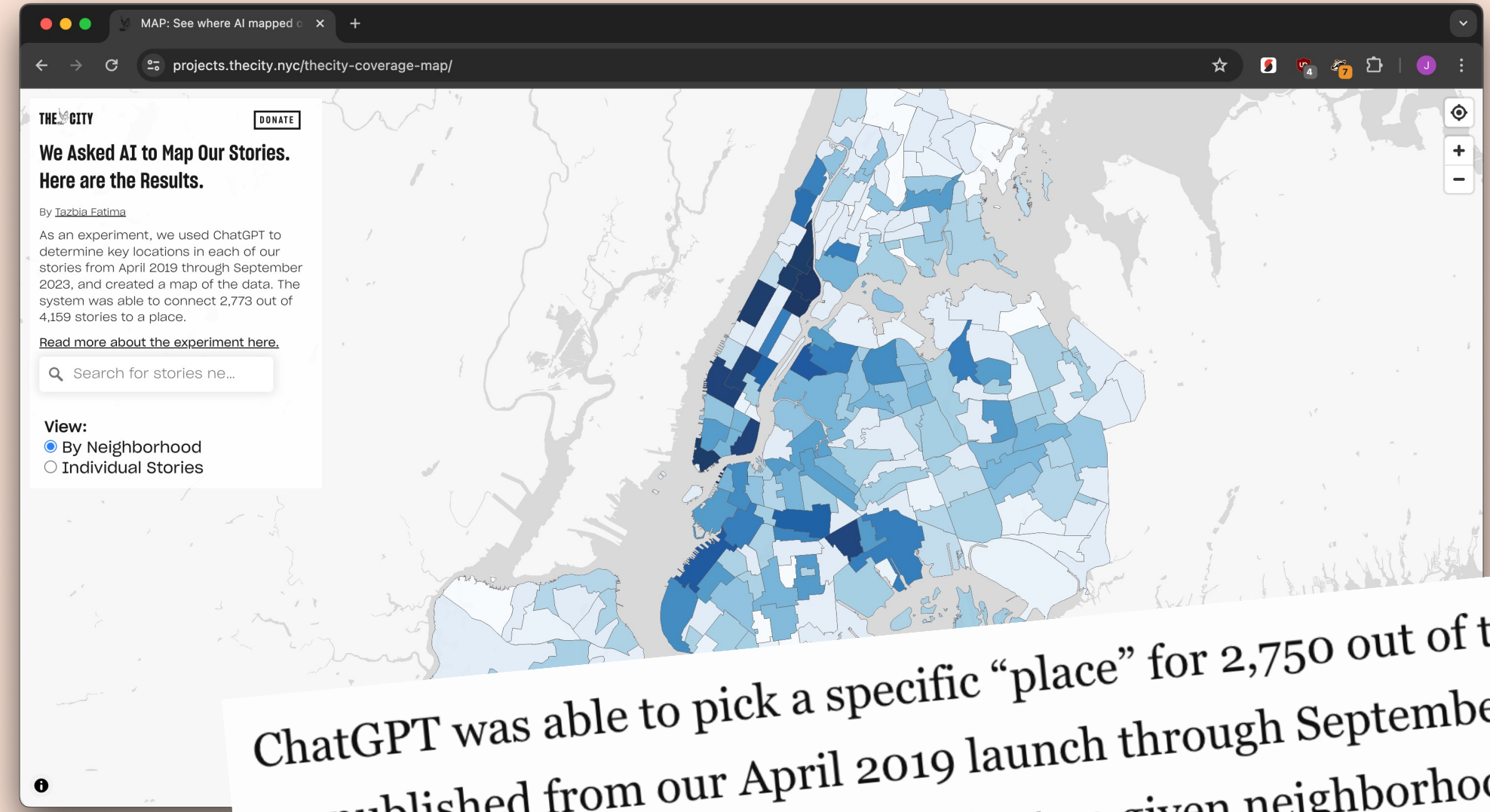
jonathan.soma

Apple says its App Store is 'a safe and trusted place.' We found 1,500 reports of unwanted sexual behavior on six apps, some targeting minors.

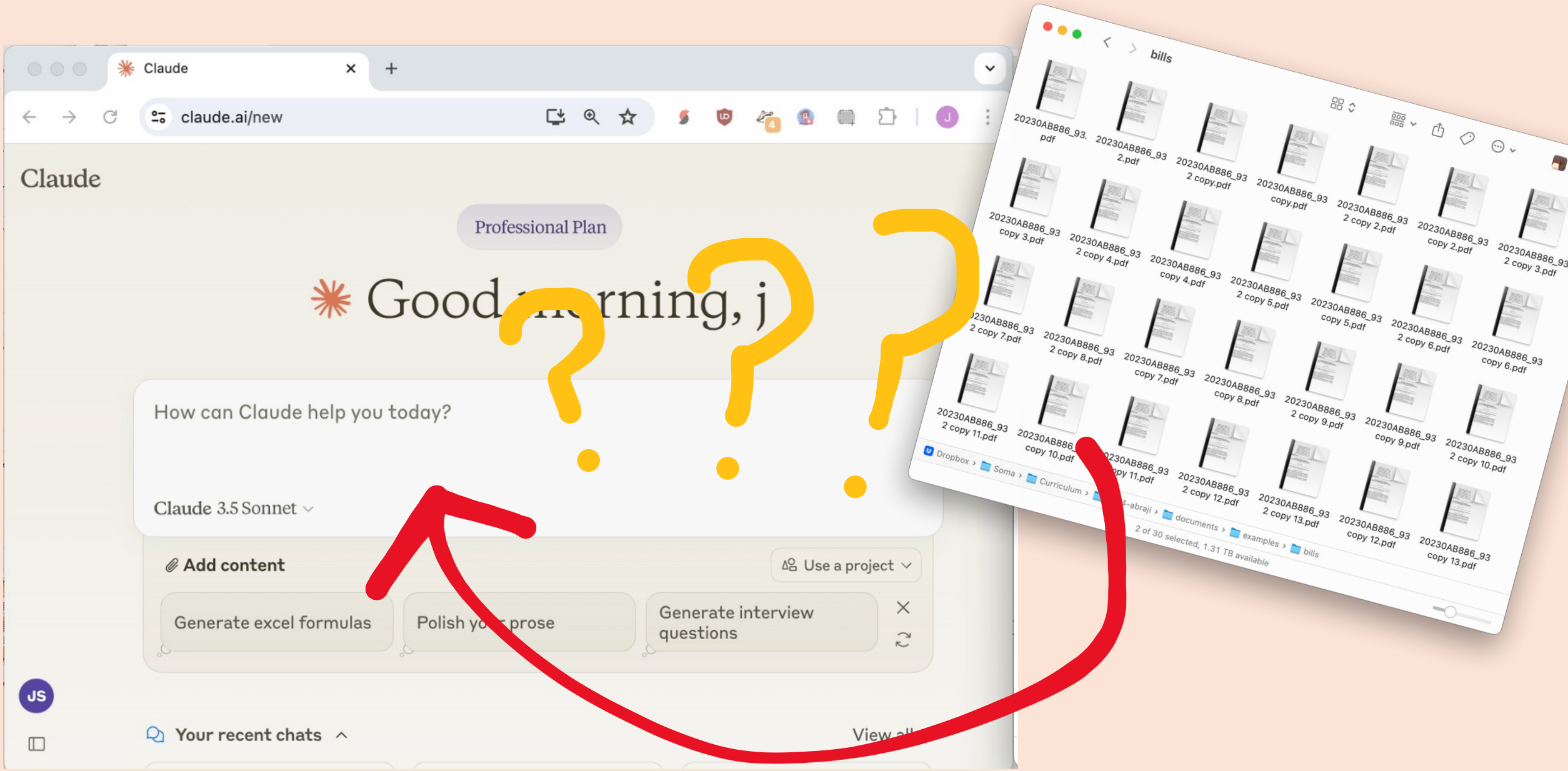
The prevalence of unwanted sexual content raises questions about whether Apple can continue to offer a protective cocoon to its customers as its platform grows

By Reed Albergotti and Al Johri
November 22, 2019 at 7:00 a.m. EST

Using a machine learning algorithm to identify App Store reviews containing reports of unwanted sexual content, racism and bullying, The Post sifted through more than 130,000 reviews of six random chat apps, all but one of which were ranked in the top 100 for social networking by Apple earlier this month. The Post manually inspected the more than 1,500 reviews that made mention of uncomfortable sexual situations.



ChatGPT was able to pick a specific "place" for 2,750 out of the 4,159 stories we published from our April 2019 launch through September 2023. We were also able to match 2,129 stories to a given neighborhood. We used the


Claude
claude.ai/new
Claude
Professional Plan
Good morning, j
How can Claude help you today?
Claude 3.5 Sonnet
Add content
Use a project
Generate excel formulas
Polish your prose
Generate interview questions
JS
Your recent chats
View all
bills
Dropbox
Soma
Curriculum
documents
examples
bills
2 of 30 selected, 1.31 TB available

Claude for Sheets™
Claude for Sheets™ brings your friendly AI assistant Claude from Anthropic right there in your spreadsheets.
Install
By: Anthropic
Listing updated: March 11, 2024
14
30K+
Works with:

GPT for Sheets™ and ...
ChatGPT in Google Sheets™ and Docs™. Get the full power of AI for inspiration data cleaning, classification, extraction, translation, edition, summarization, writing directly in you...
Uninstall
By: Talarian
Listing updated: March 6, 2024
Independent security verification
Works with:
318
6M+

AI in Sheets™ .Spread...
The most powerful AI tool for Google Sheets™ on the planet. Built for true scale. Safe for the enterprise. Unbeatable pricing model.
Install
By: Spreadsheet Daddy
Listing updated: March 18, 2024

Claude for Sheets™ - Google

workspace.google.com/marketplace/app/claude_for_sheets/909417792257

https://bit.ly/abraji-2024-docs

Google Workspace Marketplace

Search apps

- Home
- Editor's choice
  - Intelligent apps
  - Work from everywhere
  - Business essentials
  - Apps to discover
- Built by Google
- Top charts
  - Most popular
  - Top rated
- Categories
- Admin managed

Claude for Sheets™

Claude for Sheets™ brings your friendly AI assistant Claude from Anthropic right there in your spreadsheets.

By: Anthropic

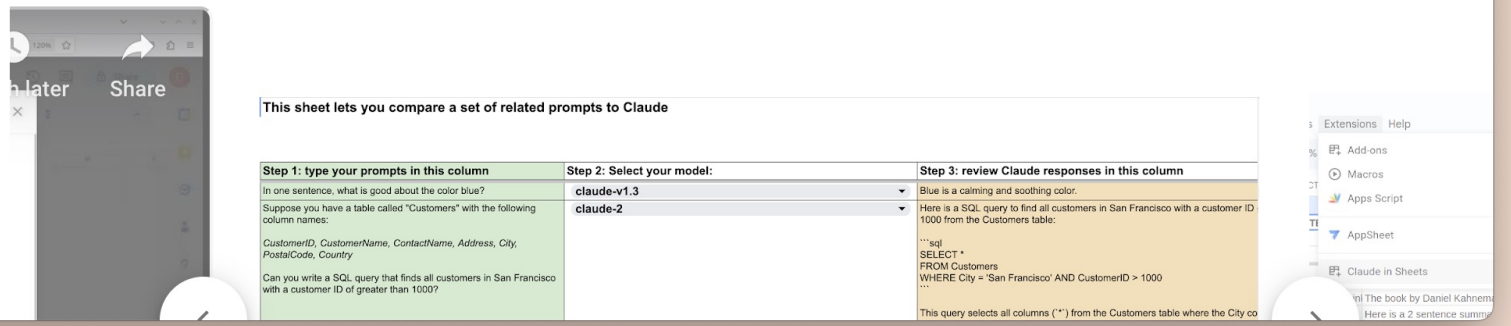
Listing updated: March 11, 2024

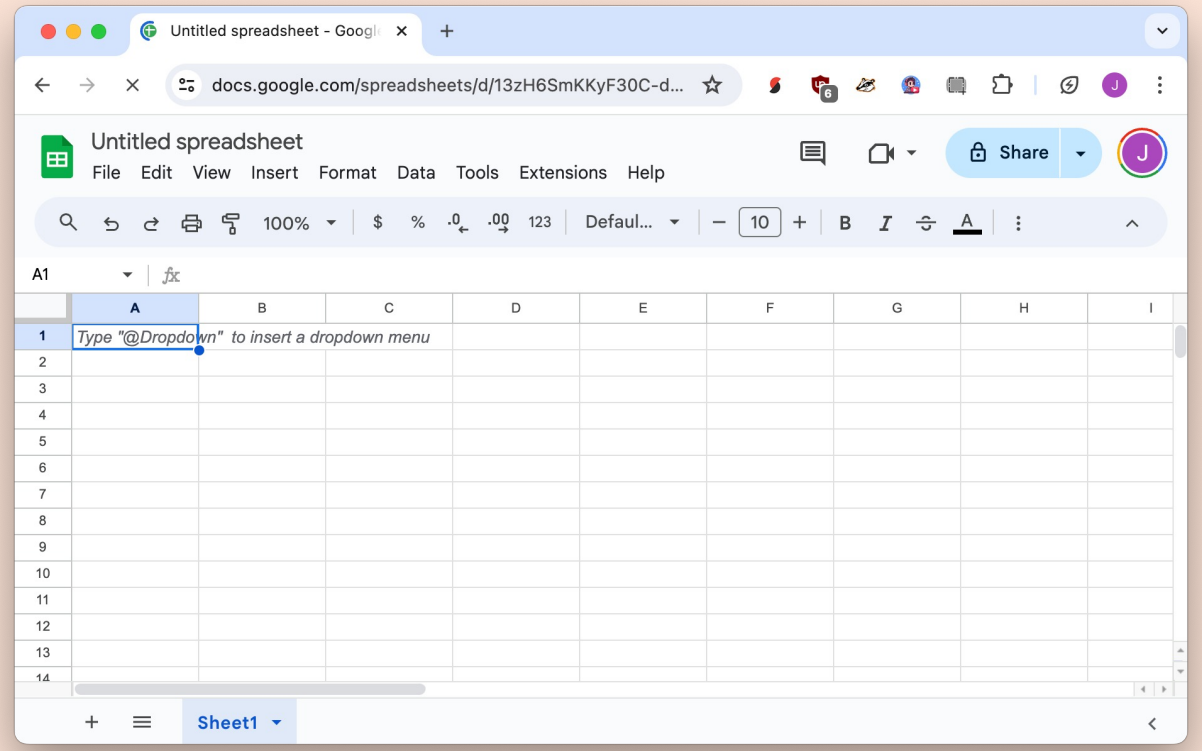
Uninstall

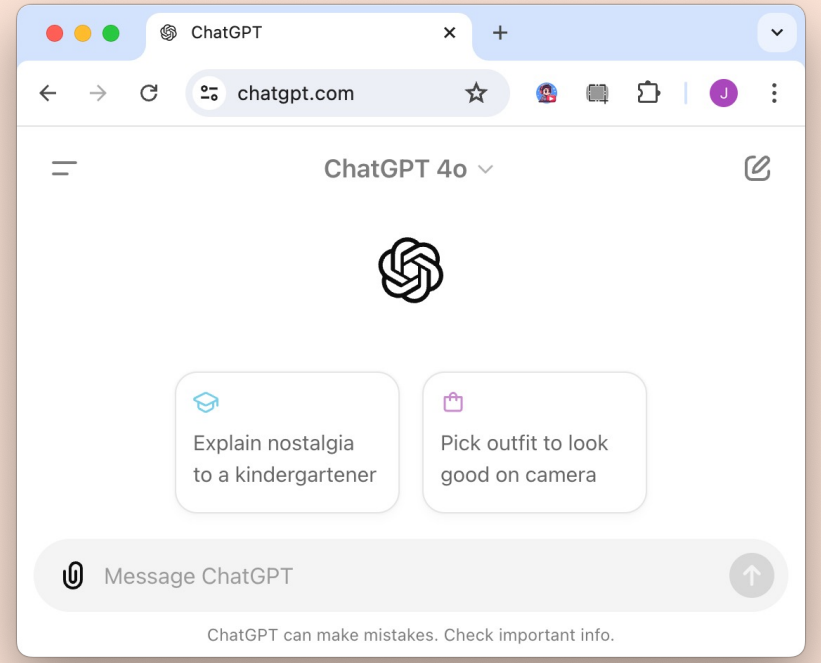
Works with:

17 reviews · 34K+

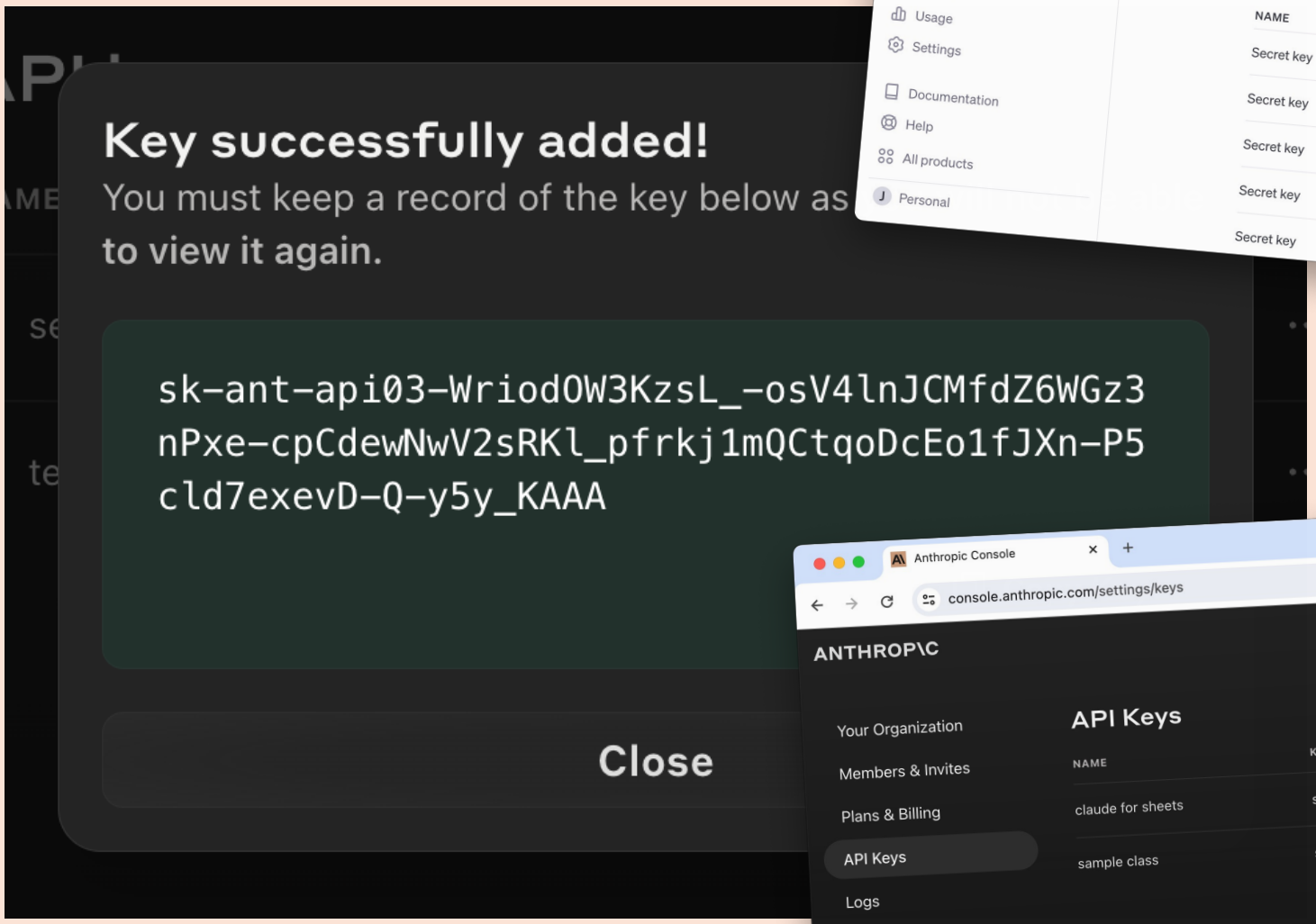
Overview | Permissions | Reviews

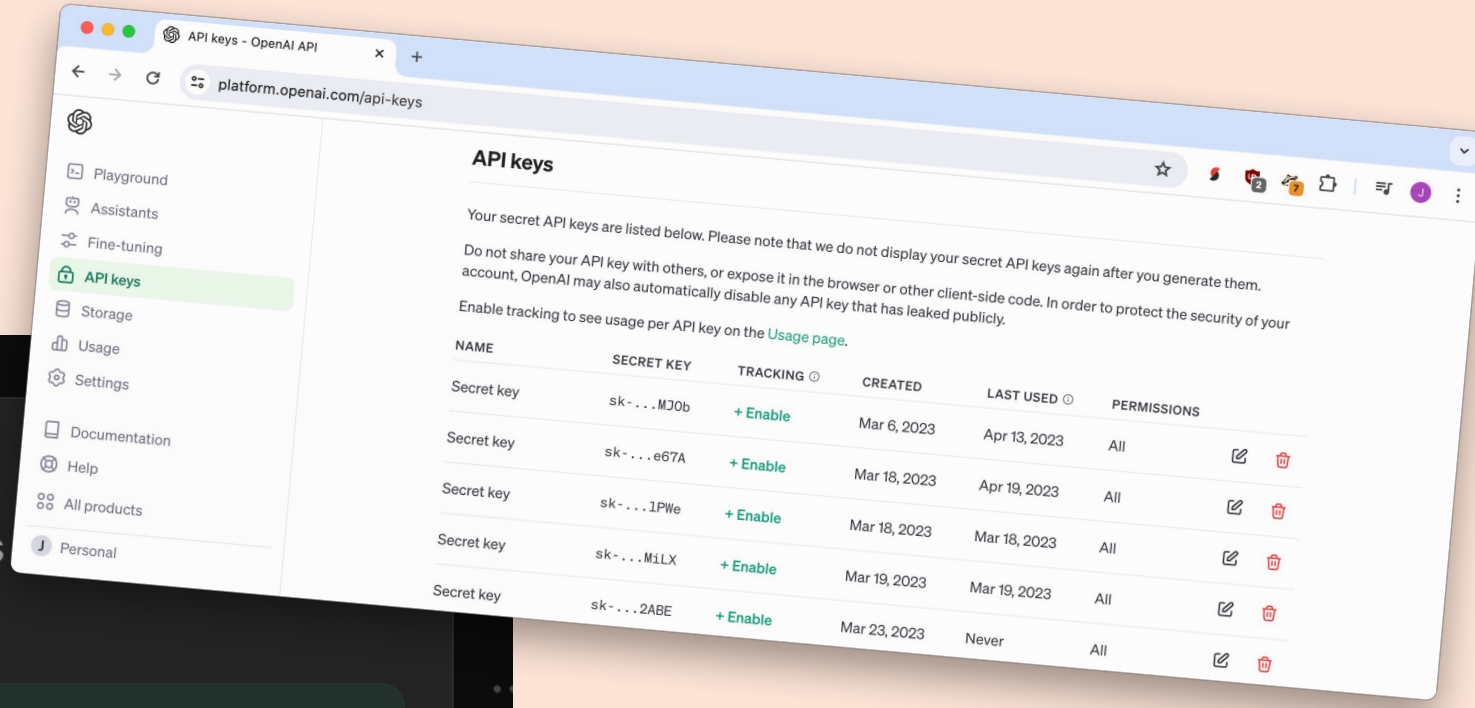


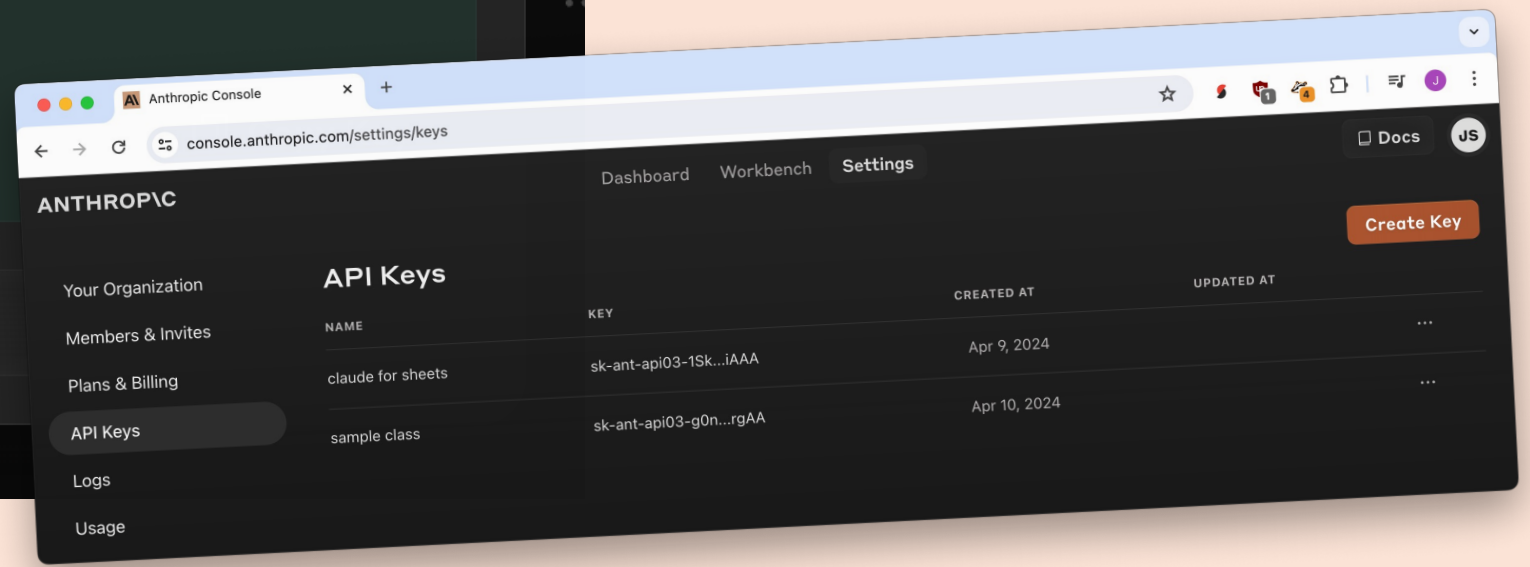

Untitled spreadsheet - Googl
docs.google.com/spreadsheets/d/13zH6SmKKyF30C-d...
Untitled spreadsheet
File Edit View Insert Format Data Tools Extensions Help
Share
100%
Defaul...
10
A1
Type "@Dropdown" to insert a dropdown menu
Sheet1


ChatGPT
chatgpt.com
ChatGPT 4o
Explain nostalgia to a kindergartener
Pick outfit to look good on camera
Message ChatGPT
ChatGPT can make mistakes. Check important info.

API keys


Key successfully added!
You must keep a record of the key below as
to view it again.
sk-ant-api03-WriodOW3KzsL_-osV4lnJCMfdZ6WGz3
nPxe-cpCdewNwV2sRKl_pfrkj1mQCtqoDcEo1fJXn-P5
cld7exevD-Q-y5y_KAAA
Close


API keys - OpenAI API
platform.openai.com/api-keys
Playground
Assistants
Fine-tuning
API keys
Storage
Usage
Settings
Documentation
Help
All products
Personal
API keys
Your secret API keys are listed below. Please note that we do not display your secret API keys again after you generate them.
Do not share your API key with others, or expose it in the browser or other client-side code. In order to protect the security of your account, OpenAI may also automatically disable any API key that has leaked publicly.
Enable tracking to see usage per API key on the Usage page.
NAME
SECRET KEY
TRACKING
CREATED
LAST USED
PERMISSIONS
Secret key sk-...MJOb + Enable Mar 6, 2023 Apr 13, 2023 All
Secret key sk-...e67A + Enable Mar 18, 2023 Apr 19, 2023 All
Secret key sk-...1PWe + Enable Mar 18, 2023 Mar 18, 2023 All
Secret key sk-...MiLX + Enable Mar 19, 2023 Mar 19, 2023 All
Secret key sk-...2ABE + Enable Mar 23, 2023 Never All


Anthropic Console
console.anthropic.com/settings/keys
Docs
JS
ANTHROP\C
Dashboard
Workbench
Settings
Create Key
Your Organization
Members & Invites
Plans & Billing
API Keys
Logs
Usage
API Keys
NAME
KEY
CREATED AT
UPDATED AT
claude for sheets sk-ant-api03-1Sk...iAAA Apr 9, 2024
sample class sk-ant-api03-g0n...rgAA Apr 10, 2024

File Edit View Insert Format Data Tools Extensions Help

200% $ % .0 .00 123 Defaul... 10 B I A

A1 fx

	A	B
1		
2		
3		
4		
5		
6		
7		
8		

File Edit View Insert Format Data Tools Extensions Help

200% $ % .0 .00 123 Defaul... 10 B I A

E8 fx

	A	B	C
1	**object**	**color**	
2	sky		
3	strawberry		
4	basketball		
5			
6			
7			
8			

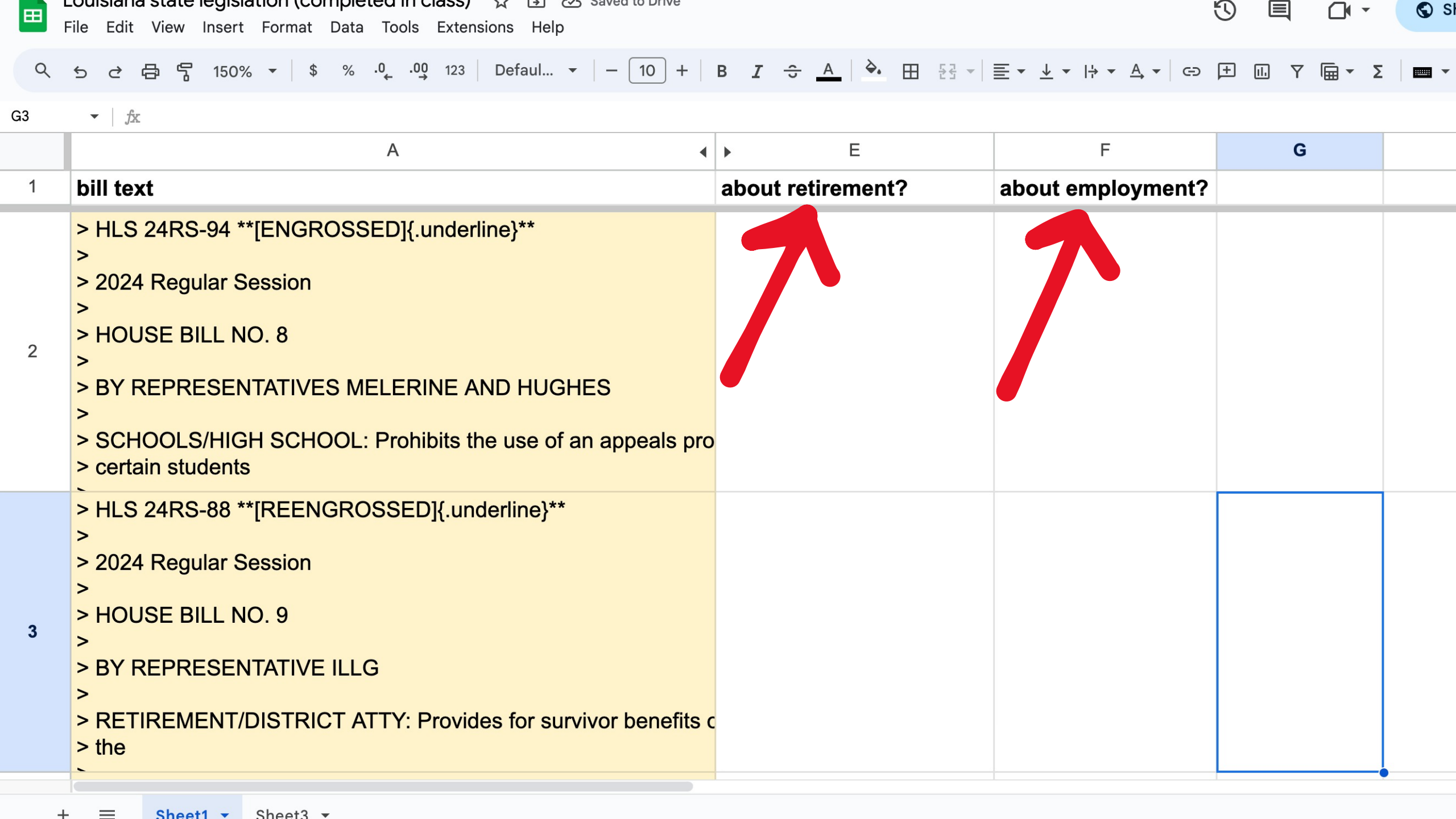
=CLAUDEEXTRACT("cor (em português)", A2)

object	color
céu	céu
morango	Não há "cor (em português)" nessa string.
basquetebol	None

=CONCATENATE("Qual é a cor de ‘", A2, "'? Responda em uma ou duas palavras.”)

=CLAUDE(B2)

object	full prompt	response
céu	Qual é a cor de 'céu'? Responda em uma ou duas palavras.	Azul claro.
morango	Qual é a cor de 'morango'? Responda em uma ou duas palavras.	Vermelho.
basquetebol	Qual é a cor de 'basquetebol'? Responda em uma ou duas palavras.	Laranja.


Louisiana state legislation (completed in class)
Saved to Drive
File Edit View Insert Format Data Tools Extensions Help
150%
Defaul...
10
G3
A
E
F
G
1
bill text
about retirement?
about employment?
2
> HLS 24RS-94 **[ENGROSSED]{.underline}**
>
> 2024 Regular Session
>
> HOUSE BILL NO. 8
>
> BY REPRESENTATIVES MELERINE AND HUGHES
>
> SCHOOLS/HIGH SCHOOL: Prohibits the use of an appeals pro
> certain students
3
> HLS 24RS-88 **[REENGROSSED]{.underline}**
>
> 2024 Regular Session
>
> HOUSE BILL NO. 9
>
> BY REPRESENTATIVE ILLG
>
> RETIREMENT/DISTRICT ATTY: Provides for survivor benefits c
> the
Sheet1
Sheet3

Louisiana state legislation (completed in class) Saved to Drive

File Edit View Insert Format Data Tools Extensions Help

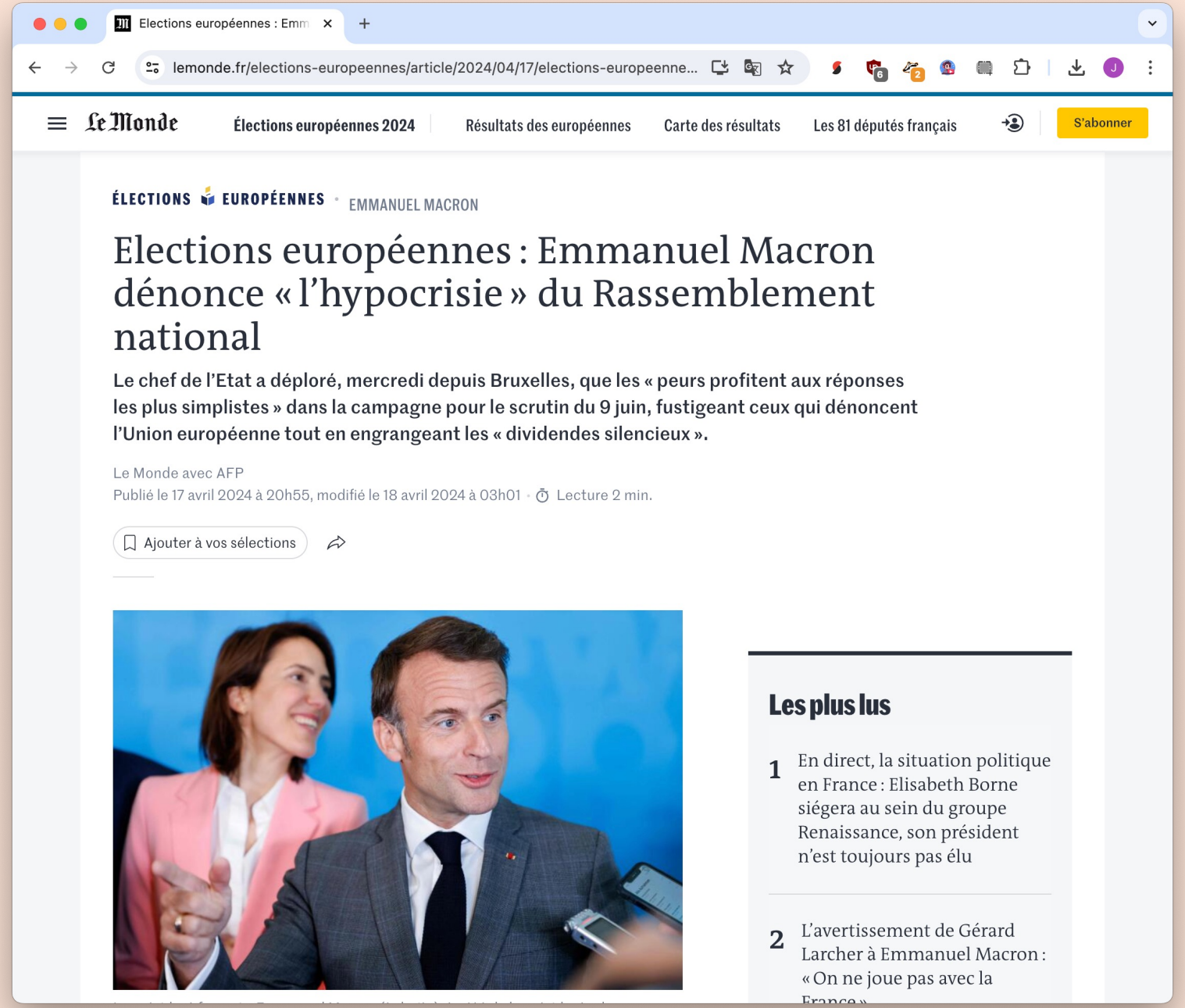
150% $ % .0 .00 123 Defaul... 10 B I A

G3

=CLAUDEEXTRACT("Is this piece of legislation about retirement? Answer YES or NO", A

	A	E	F	G
1	**bill text**	**about retirement?**	**about employment?**	
	> HLS 24RS-94 **[ENGROSSED]{.underline}** > > 2024 Regular Session > > AN ACT > > [illegible] SCHOOL: Prohibits the use of an appeals [illegible] > certain students >	no	no	
3	> HLS 24RS-88 **[REENGROSSED]{.underline}** > > 2024 Regular Session > > HOUSE BILL NO. 9 > > BY REPRESENTATIVE ILLG > > RETIREMENT/DISTRICT ATTY: Provides for survivor benefit [illegible] > the >	yes	no	

Sheet1 Sheet3



Interrogé en marge de son déplacement à Bruxelles sur la campagne des élections européennes, Emmanuel Macron a ciblé, sans le nommer, le Rassemblement national (RN) dont la liste conduite par Jordan Bardella caracole en tête des sondages.

« Je pense qu'il y a beaucoup de peurs, d'inquiétudes dans le moment que nous vivons et que ces colères profitent toujours aux réponses les plus simplistes », a déclaré le chef de l'Etat français lors de son arrivée, mercredi 17 avril, dans la capitale belge avant la tenue d'un sommet européen, au cours d'un bref échange avec la presse aux côtés de la tête de liste de son camp aux européennes, Valérie Heyer. *« Il y a aujourd'hui une espèce d'hypocrisie du débat et j'espère qu'en rentrant dans les prochaines semaines dans celui-ci, cette hypocrisie sera levée »*, a ajouté M. Macron, visant en creux le RN.

Il s'agissait de la première apparition publique du chef de l'Etat et de M^me^ Hayer depuis le début de la campagne, avant qu'ils n'assistent ensemble à une réunion du groupe Renew Europe.

Lire aussi | Elections européennes 2024 : le calendrier du scrutin

La candidate du camp présidentiel peine pour l'instant à imprimer sa marque dans la campagne. *« On s'organise, on a la bonne candidate (…) On a surtout les bonnes idées, celles qui se sont mises en place en Europe ces dernières années »*, a lancé M. Macron, interpellé par un journaliste.

- `=CLAUDEEXTRACT("people mentioned in the article, with titles if available. One per line.", B5)`
- `=CLAUDEEXTRACT("brief one-sentence summary", B5)`
- `=CLAUDEEXTRACT("English translation", E5)`
- `=CLAUDEEXTRACT("article category, one of politics, sports, international, other", B5)`

B	C	E	F	G	H	I
content	**sources**	**summary**	**language code**	**is it English?**	**English translation**	**article catego**
Ante inversores en Washington, Caputo defendió la estrategia fiscal y reafirmó el compromiso para levantar el cepo El jefe del Palacio de Hacienda participó de un seminario organizado por el banco J.P. Morgan WASHINGTON.- El ministro de Economía, Luis Caputo, llegó al seminario del banco de inversión J.P. Morgan en el Park Hyatt hotel poco después de las tres de la tarde, hora local, un rato después de aterrizar en Washington	Luis Caputo, Ministro de Economía Pablo Quirno, Secretario de Finanzas Javier Milei, gobierno Mauricio Macri, gobierno Nicolás Posse, Jefe de Gabinete Alberto Fernández, gobierno Pierre-Olivier Gourinchas, Economista Jefe del FMI	The economy minister defended the government's fiscal adjustment strategy and reiterated the intention to lift currency controls, but said they cannot do so with the current level of reserves unless they receive a	en	yes	The economy minister defended the government's fiscal adjustment strategy and reiterated the intention to lift currency controls, but said they cannot do so with the current level of reserves unless they receive a fresh injection of money.	international
Elections européennes : Emmanuel Macron dénonce « l'hypocrisie » du RN Le chef de l'Etat a déploré mercredi depuis Bruxelles que les « peurs profitent aux réponses les plus simplistes » dans la campagne pour le scrutin du 9 juin, fustigeant ceux qui dénoncent l'Union européenne tout en engrangeant les « dividendes silencieux ». Vous pouvez partager un article en cliquant sur les icônes de partage en haut à droite de celui-ci	Emmanuel Macron, President of France Valérie Heyer, Head of list for Macron's camp in European elections Jordan Bardella, Head of list for National Rally (Rassemblement National, RN) Raphaël Glucksmann, Head of list for Socialist Party (Parti Socialiste) and Place Publique Gabriel Attal, Prime Minister of France	Emmanuel Macron a déploré que les "peurs profitent aux réponses les plus simplistes", mercredi 17 avril.	fr	no	Emmanuel Macron lamented that "fears profit from the simplest answers", Wednesday April 17.	politics
岸田首相、韓国・尹大統領に日米首脳会談の成果説明　日韓連携も確認 岸田文雄首相は17日、韓国の尹錫悦（ユンソンニョル）大統領と電話で協議した。首相によると、10日にワシントンで行われた日米首脳会談での合意事項などについて説明したという。 協議後、首相は首相官邸で記者団に「日韓、日米韓の連携をより深化させ、日韓首脳間を含め緊密に意思疎通を図っていくことで、致した」と語った	Fumio Kishida, Prime Minister; Yoon Suk-yeol, President of South Korea	岸田首相は17日、韓国の尹大統領と電話で協議し、日米首脳会談の成果を説明、日韓連携も確認した。	ja	no	Prime Minister Kishida held talks by phone with South Korean President Yoon on the 17th, explaining the results of the Japan-U.S. summit and confirming cooperation between Japan and South Korea.	other

Introduction
Why use Instructor?
Help with Instructor
Installation
Contributing
Philosophy

Instructor, Generating Structure from LLMs

Structured outputs powered by llms. Designed for simplicity, transparency, and control.

Follow @jxnlco discord 97 online downloads 296k/month docs InstructorGPT

Instructor makes it easy to get structured data like JSON from LLMs like GPT-3.5, GPT-4, GPT-4-Vision, and open-source models including Mistral/Mixtral, Anyscale, Ollama, and llama-cpp-python.

It stands out for its simplicity, transparency, and user-centric design, built on top of Pydantic. Instructor helps you manage validation context, retries with Tenacity, and streaming Lists and Partial responses.

- Instructor provides support for a wide range of programming languages, including:
- Python
- TypeScript
- Ruby
- Go
- Elixir

Table of contents
Want your logo on our website?
Why use Instructor?
Getting Started
Using OpenAI
Using Anthropic
Using Gemini
Google AI
Vertex AI
Using Litellm
Correct Typing
Calling create
Handling async: await create
Returning the original completion: create_with_completion
Streaming Partial Objects: create_partial
Streaming Iterables: create_iterable
Validation
More Examples
Contributing

```python
class Person(BaseModel):
    name: str = Field(description="Person's name")
    title: Optional[str] = Field(description="Person's title, if available")

class Article(BaseModel):
    summary: str = Field(description="Brief, one-sentence summary (in English)")
    sources: List[Person] = Field(description="People mentioned in the article")
    language_code: str = Field(description="Two-letter language code of original text")
    category: Literal["politics", "sports", "international", "other"]
```

```python
# Extract structured data from natural language
result = client.chat.completions.create(
    model="gpt-4",
    response_model=Article,
    messages=[{"role": "user", "content": article}],
)
```

```
{
    "category": "politics",
    "language_code": "fr",
    "sources": [
        { "name": "Emmanuel Macron", "title": "President of France" },
        { "name": "Jordan Bardella", "title": "Head of the National Rally" },
        { "name": "Valérie Heyer", "title": "Candidate for the European Elections" },
        { "name": "Gabriel Attal", "title": "Prime Minister of France" },
        { "name": "Raphaël Glucksmann", "title": "Candidate for the European Elections" }
    ],
    "summary": "Emmanuel Macron criticizes the National Rally (RN) for hypocrisy in the run-up to the European elections, arguing that they profit from EU membership while denouncing it."
}
```

FWS-R5-ES-2014-0050-0003_attachment_1....
1 page

Attn: FWS-R5-ES-2014-0050

Public Comments Processing

Division of Policy, Performance, and Management Programs

US Fish and Wildlife Service

5275 Leesburg Pike

Falls Church VA 22041

Re: Copenhagen Wind Farm

There are a couple of things that I would like to see addressed in the application for ITP and HCP.

First is my interest in the detection of morbidity and mortality at the wind turbines. Daily surveillance, preferably at sunrise, is, in my opinion, the only way to get reliable numbers. What has been proposed by the applicant and FWS?

Second is my interest in avoidance of bird electrocutions and collisions at collector lines and transmission lines. What is the proposal for avoidance of electrocution along these lines? What is the proposal for avoidance of collisions where these lines cross streams and wetlands?

Please keep me abreast of this application progress, either by sending me copies of pertinent documents, or by informing me when they are available on regulations.gov.

Thank you.

C G Spies

860-583-5739

cgspies47@gmail.com

AutoSave
FWS-R3-ES-2011-0029-0123_a... — Saved to my Mac
Search (Cmd + Ctrl + U)
Home Insert Draw Design Layout References Mailings Review View Zotero
Comments Editing Share
Paste Times New... 12 Paragraph Styles Dictate Add-ins Editor Create PDF and share link Request Signatures

6/9/11
Sandra Neily
POB 102
Greenville Maine 04441
neily.sandra26@gmail.com

Comments on proposal: Docket No. [FWS-R3-ES-2011-0029]

I attended the informational session and public hearing offered by USFW in Augusta, Maine (6/8/11) and appreciated the service's outreach work to bring this issue to us.

Several unresolved issues arise.

1. **The current proposal to delist gray wolves in the eastern region (Great Lakes and remaining eastern states) without an official designation or classification that would protect dispersing wolves from Canada to Maine or other eastern states seems is a serious omission**. Given the current anti-coyote climate in Maine where coyotes are blamed for serious declines in the deer herd, any dispersing wolf could easily be killed 365 days a year as wolves fall within the "similar appearance" category in relation to coyotes. (State and NGO biologists are clear: the accumulated loss of winter shelter from years of unrestrained harvesting on industrial lands has left deer without winter shelter and protected corridors making the herd more vulnerable in some areas to predation and severe winters. Anti-coyote attitudes threaten wolf dispersal.)

2. If this same situation (unprotected status during a study period) had been present in adjoining states/regions, it's less likely wolves would have successfully dispersed from Minnesota to Wisconsin and Michigan or dispersed from Idaho to Washington. Dispersing populations must be protected. With seven wolves killed in northeastern states in the past decade, it is clear the region has the

Page 1 of 2 663 words English (United States) Accessibility: Unavailable Focus 162%

Comment on FR Doc # 2015-26874

Posted by the **Federal Aviation Administration** on Dec 8, 2015

Docket / Document (FAA-2015-4378-0022) / Comment

Comment ID

FAA-2015-4378-2792

Tracking Number

1jz-8m35-3ip9

Comment Details | **Submitter Info**

Comment

There should be no regulations

box3616#gmail.com

Extract information from the following public comment on regulation.
Please list:

- The submitter's name, if available
- Whether the submitter is FOR or AGAINST regulation
- The submitter's email address

COMMENT TEXT:
There should be no regulations

box3616#gmail.com

The email address listed is not valid, please fix it

I apologize for any confusion. Given that the email address provided ("box3616#gmail.com") appears incorrect, it's likely intended to be "box3616@gmail.com".

class Comment(BaseModel):
name: Optional[str] = Field(description="Submitter's name")
email: str = Field(description="Submitter's email address")
position: Literal["for", "against", "neutral"]
@field_validator("email")
@classmethod
def validate_email(cls, value):
if value and '@' not in value:
raise ValueError("Email missing @")
return value
result = client.chat.completions.create(
model="gpt-3.5-turbo",
response_model=Comment,
messages=[{"role": "user", "content": comment}],
max_retries=3

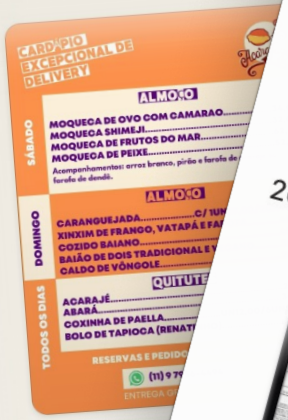
But sometimes you just want one answer

Brazilian Restaurant Delivery

claude.ai/chat/5804fabb-fc5c-4730-9db

Claude

Brazilian Restaurant Delivery Menu



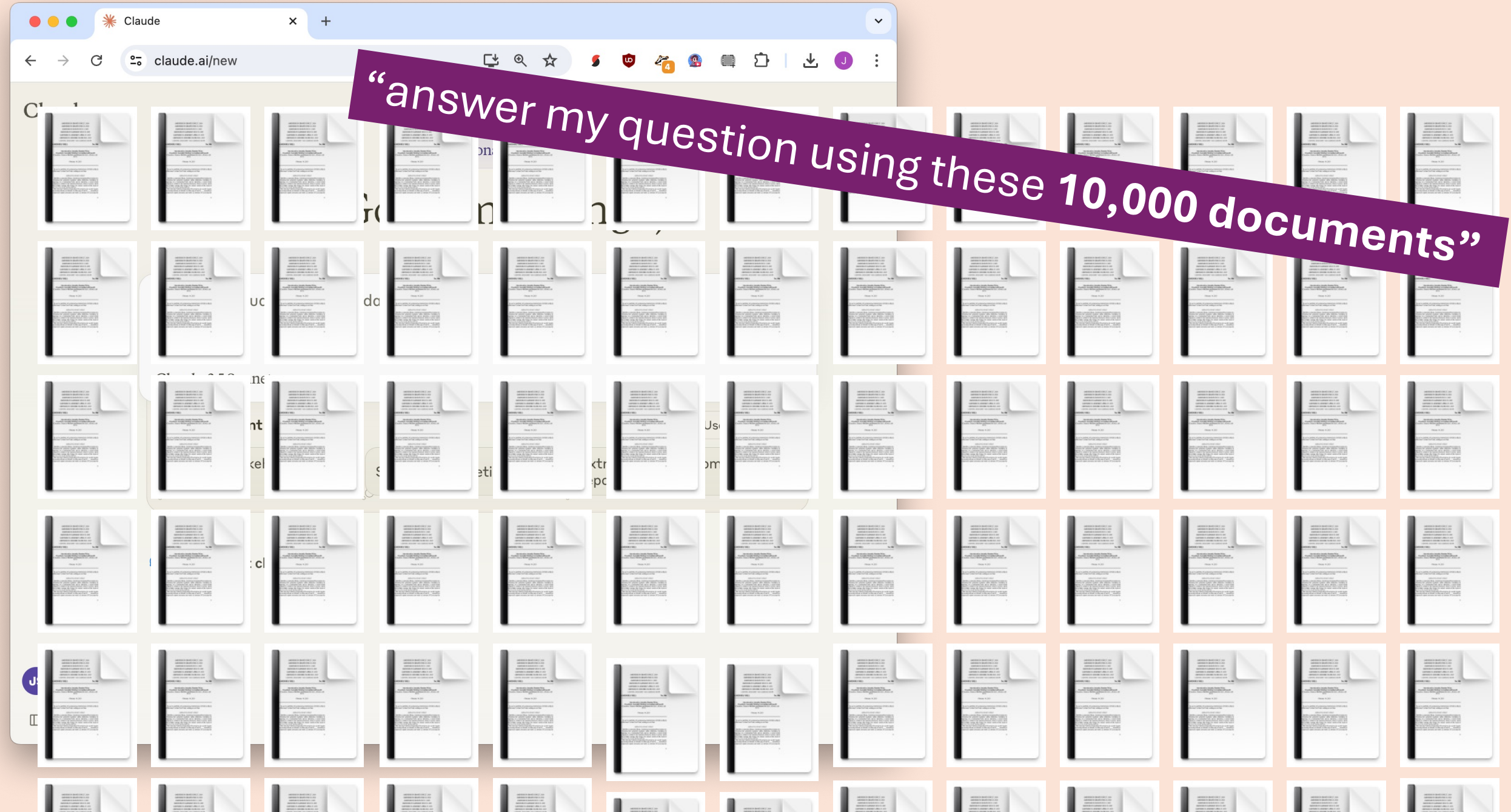
JS

bills

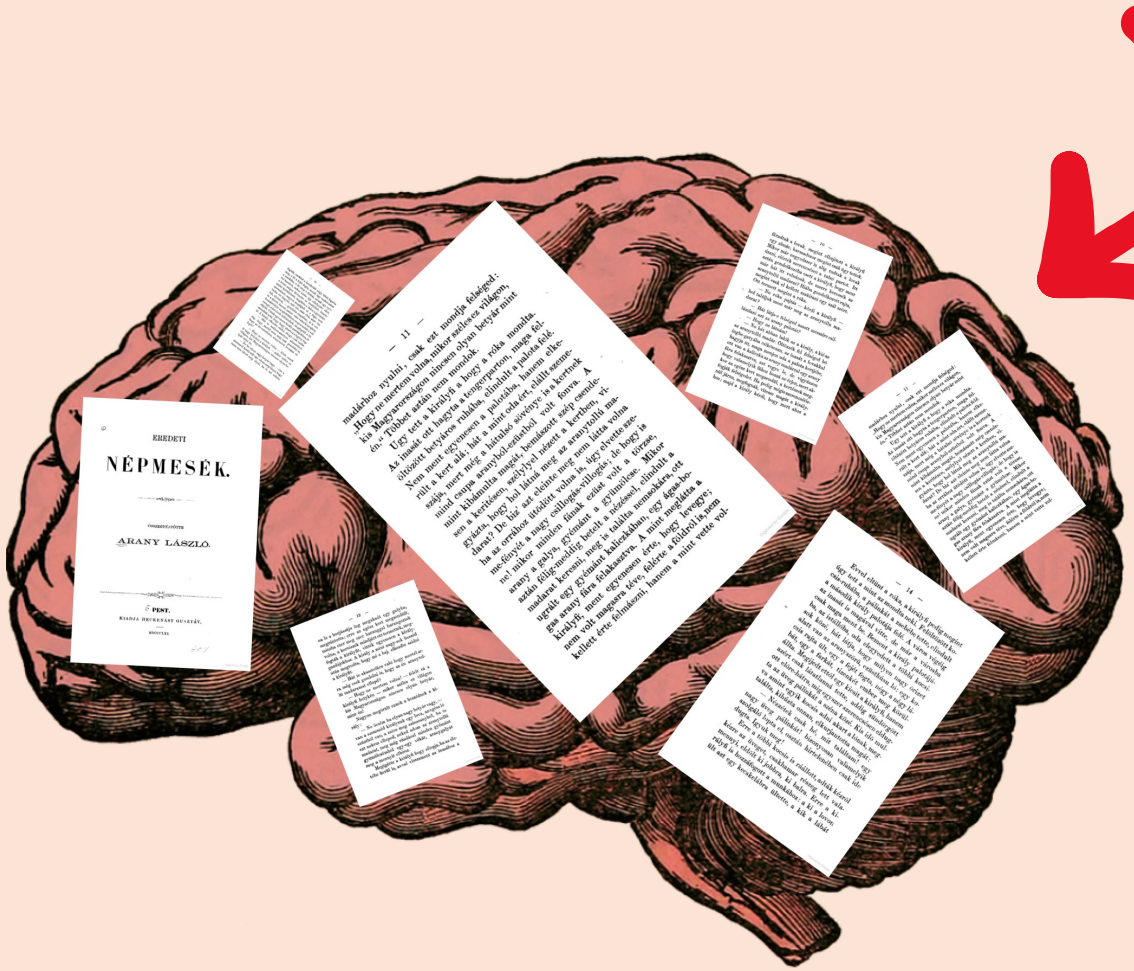
20230AB886_93.pdf
20230AB886_932.pdf
20230AB886_932 copy.pdf
20230AB886_93 copy.pdf
20230AB886_932 copy 2.pdf
20230AB886_93 copy 2.pdf
20230AB886_932 copy 3.pdf
20230AB886_93 copy 3.pdf
20230AB886_932 copy 4.pdf
20230AB886_93 copy 4.pdf
20230AB886_932 copy 5.pdf
20230AB886_93 copy 5.pdf
20230AB886_932 copy 6.pdf
20230AB886_93 copy 6.pdf
20230AB886_932 copy 7.pdf
20230AB886_932 copy 8.pdf
20230AB886_93 copy 7.pdf
20230AB886_93 copy 8.pdf
20230AB886_932 copy 9.pdf
20230AB886_93 copy 9.pdf
20230AB886_932 copy 10.pdf
20230AB886_932 copy 11.pdf
20230AB886_93 copy 10.pdf
20230AB886_93 copy 11.pdf
20230AB886_932 copy 12.pdf
20230AB886_93 copy 12.pdf

Dropbox › Soma › Curriculum › 2024-abraji ›

- These come wi


Claude
claude.ai/new
“answer my question using these 10,000 documents”

Context window



— 19 —

akármint lesz, el kell neki lopni a kisasszonyt, jobb hát ha most elszökteti, azután majd tanácsot kérhet a rókától.

El is szöktette szerencsésen, kimentek a mezőre a királyfi inasához, ott, — míg az inas a lovakat nyergelte, — a királyfi félre ment, elszakította az ötödik szál szőrt is. Ott termett a róka:

— Mi a baj felséges királyfi?

— Hát bizon most az, hogy nem tudom a királykisasszonytól váljak-e meg, vagy madár nélkül menjek haza? — mert így, meg így áll a dolog.

— Soh'sc' búsuljon azon a királyfi, majd elintézem én azt, most majd együtt megyek felségedtel.

Mentek aztán együtt négyen: a királyfi, a szép királykisasszony, az inas meg a róka. Mikor ahoz a városhoz értek, ahova a kisasszonyt kellett volna vinni, a város végén a róka épen olyan szép kisasszonynyá változott mint az igazi, úgy hogy a királyfi is alig tudta megismerni őket, a róka-kisasszony mondta az igazinak, hogy csak menjenek az inassal, nekik most a királyfival egy kis dolguk van, hanem majd utánuk mennek. Ment aztán a kisasszony tovább, a ki-

2*

— 16 —

ellopod, neked adom az aranyszörü lovat, a gyémánt szerszámot, meg még ráadásul tiz köböl arany abrakot.

Jól van; megigérte a királyfi hogy ellopja, avval kiment a város végére az inasához. Itt megint mit volt mit tenni? el kellett szakitani a negyedik szál szőrt is. Ott termett a róka:

— No mi a baj felséges királyfi.

— Hát csak e' meg e. Itt elbeszélte a királyfi hogy már meg királykisasszonyt kellene neki lopni.

— No királyfi — mondja a róka, — most öltözzék fel szakács ruhába, menjen el ahoz a királyhoz, segedjék be szakácsnak. Délben majd felséged vigye be a levest, hanem az ajtóban ejtse el. Többet aztán nem mondok.

Eltünt a róka megint, a királyfi pedig az inasával ment a harmadik király országa felé. Mikor ahoz a városhoz értek, a hol a király lakott, felöltözött a királyfi szakácsruhába, az inasát a város végén hagyta, maga bement a királyi palotába. Épen kint sétálgatott a király az udvaron; oda ment hozzá a királyfi:

— Felséges király én szakács vagyok, azért jöttem hogy beszegődjem felségedhez, hahogy megfogadna.

— 18 —

király fia, meg mást is mindent, az egész élete folyását, mindjárt beleszerettek egymásba halálosan. A kisasszony is elbeszélte, hogy az ő kezét most egy igen-igen gazdag király kérte meg, de neki nincs kedve hozzá menni, mert már olyan öreg, hogy a hamut is mammának mondja, hanem ez az apjának való. „Most hát használ, erőlteti hogy menjen hozzá; „Most hát szívem szép szerelme, te az enyim, én a tied, nincs mást mit tenni, hanem az éjjel szökjünk meg, majd a te apád országában megesküdünk hetünk.

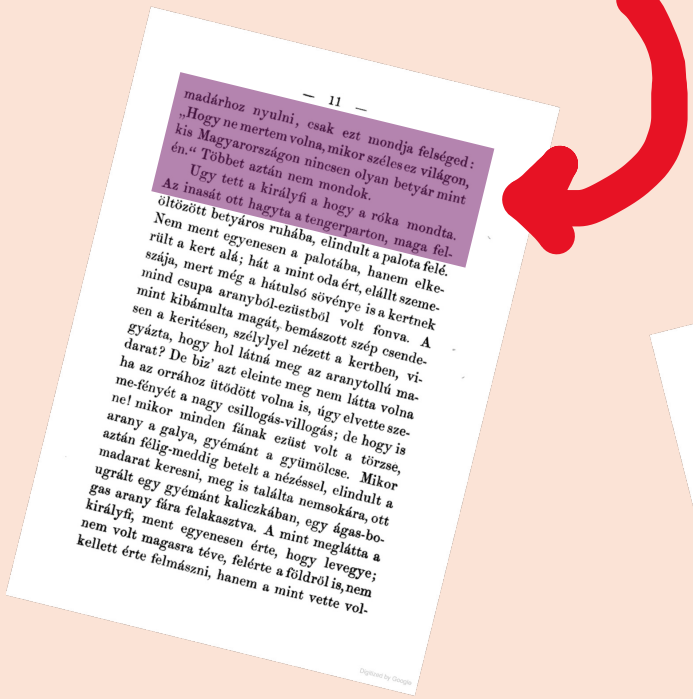
Kapott a királyfi két kézzel a tanácson, mindjárt elhatározták, hogy még azon az éjszakán, majd ha mindenki lenyugszik a királyi udvarban, megszöknek, reggelre pedig, ha lehet, — túl teszik magukat a határon. Evvel a királyfi kiment a kisasszony szobájából, várta az éjfélt; csak akkor jutott eszébe, hogy neki ezt a kisasszonyt a szomszéd királynak kell vinni, hogy a lovat megkapja, mert ha azt meg nem kapja, az aranytollú madár se lesz az övé a nélkül pedig nem akart az apjа szeme elébe menni. Búsult ezen szörnyü módon, már azt is gondolta, hogy érte se megy a kisasszonyért, de azután eszébe jutott hogy akárhogy lesz,

— 17 —

— Jól van fiam én megfogadlak, épen szakácsra van most szükségem; jó hogy jöttél. Beszegődött hát a királyfi szakácsnak, mindjárt is bement a konyhába; hanem a másik szakács egyése délelőtt hirmásnott rá, mert minden ételt elrontott. Következett az ebéd ideje, a régi szakács sehogy se' akarta megengedni a királyfinak hogy az ételt ő vigye be, de ez addig könyörgött neki, míg utoljára mégis ráállott. Amint hát vitte volna be a királyfi a levest, az ajtónál szánt-szándékkal megbotlott, elejtette a tálat hogy összetört. A királykisasszonynak már akkor megtetszett a szép legény, mikor belépett, mikor meg a tálat elejtette, mindjárt gondolta, hogy aligha igazi szakács ez, mert akkor ko- nem lett volna ilyen ügyetlen, jobban is gyakorlott volna; amint pedig szedte össze a csepet, meglátta a kisasszony az éjjen az arany- gyürüt, mindjárt gondolta, hogy nem lehet más, dolgot, hanem másnap nem mert be ebédelni az apjához, a maga szobájába terittetett két emberre; szentán behivatta az új szakácsot, ott mindjárt maga mellé ültette, kérdezte tőle hogy hol vette azt a szép gyürüt? A királyfi elbeszélte hogy ő nem szakács, hanem egy gazdag

Retrieval augmented generation, also known as “chatting with your PDFs”

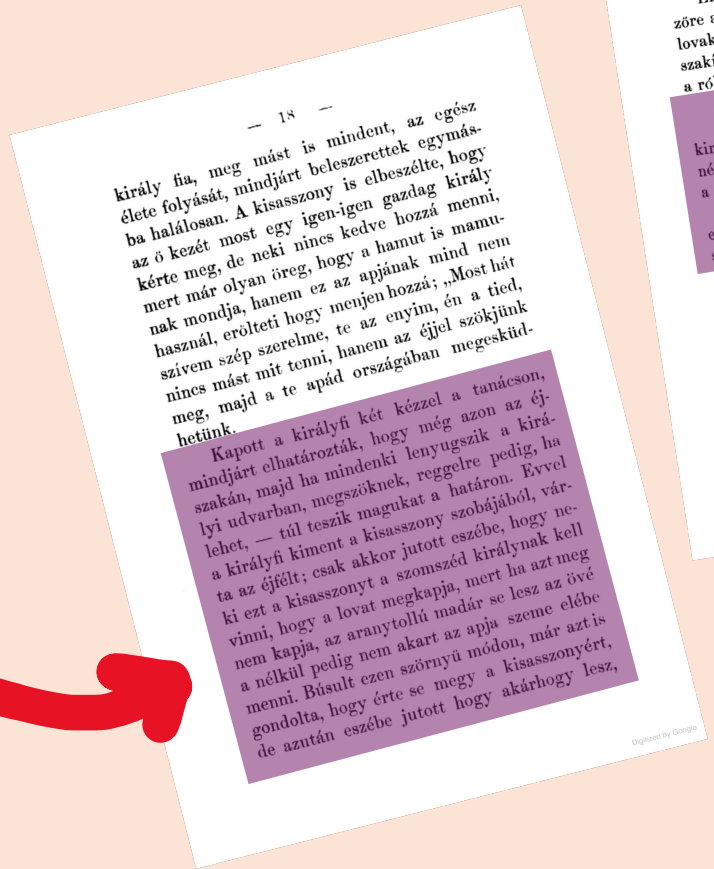
Slice up your document, send top relevant sections with your question



— 11 —

madárhoz nyulni, csak ezt mondja felséged: „Hogy ne mertem volna, mikor szélesez világon, kis Magyarországon nincsen olyan betyár mint én.“ Többet aztán nem mondok.

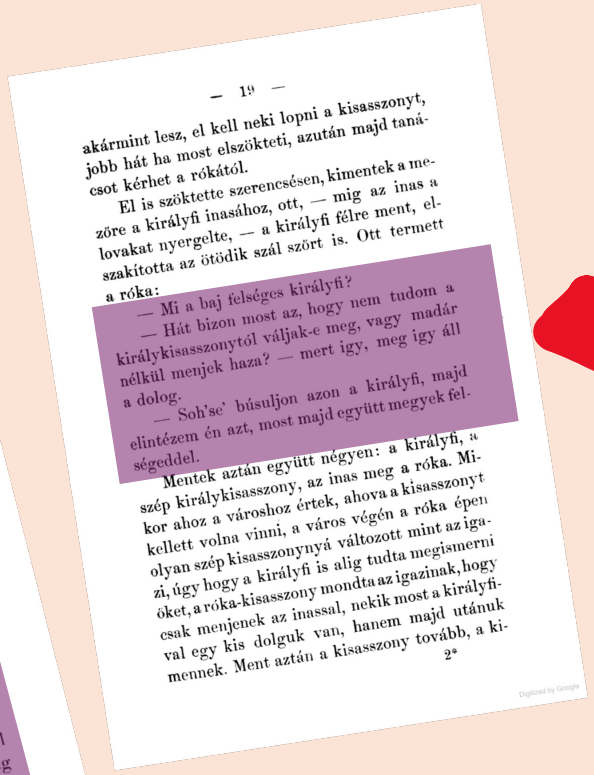
Ugy tett a királyfi a hogy a róka mondta. Az inasát ott hagyta a tengerparton, maga felöltözött betyáros ruhába, elindult a palota felé. Nem ment egyenesen a palotába, hanem elkerült a kert alá; hát a mint oda ért, elállt szemeszája, mert még a hátulsó sövénye is a kertnek mind csupa aranyból-ezüstből volt fonva. A mint kibámulta magát, bemászott szép csendesen a keritésen, szélylyel nézett a kertben, vigyázta, hogy hol látná meg az aranytollú madarat? De biz' azt eleinte meg nem látta volna ha az orrához ütődött volna is, úgy elvette szeme-fényét a nagy csillogás-villogás; de hogy is ne! mikor minden fának ezüst volt a törzse, arany a galya, gyémánt a gyümölcse. Mikor aztán félig-meddig betelt a nézéssel, elindult a madarat keresni, meg is találta nemsokára, ott ugrált egy gyémánt kaliczkában, egy ágas-bogas arany fára felakasztva. A mint meglátta a királyfi, ment egyenesen érte, hogy levegye; nem volt magasra téve, felérte a földről is, nem kellett érte felmászni, hanem a mint vette vol-



— 18 —

király fia, meg mást is mindent, az egész élete folyását, mindjárt beleszerettek egymásba halálosan. A kisasszony is elbeszélte, hogy az ő kezét most egy igen-igen gazdag király kérte meg, de neki nincs kedve hozzá menni, mert már olyan öreg, hogy a hamut is mamunak mondja, hanem ez az apjának mind nem használ, erölteti hogy menjen hozzá; „Most hát szivem szép szerelme, te az enyim, én a tied, nincs mást mit tenni, hanem az éjjel szökjünk meg, majd a te apád országában megeskü-hetünk.

Kapott a királyfi két kézzel a tanáceson, mindjárt elhatározták, hogy még azon az éjszakán, majd ha mindenki lenyugszik a királyi udvarban, megszöknek, reggelre pedig, ha lehet, — túl teszik magukat a határon. Évvel a királyfi kiment a kisasszony szobájából, várta az éjfélt; csak akkor jutott eszébe, hogy neki ezt a kisasszonyt a szomszéd királynak kell vinni, hogy a lovat megkapja, mert ha azt meg nem kapja, az aranytollú madár se lesz az övé a nélkül pedig nem akart az apja szeme elébe menni. Büsult ezen szörnyü módon, már azt is gondolta, hogy érte se megy a kisasszonyért, de azután eszébe jutott hogy akárhogy lesz,



— 19 —

akármint lesz, el kell neki lopni a kisasszonyt, jobb hát ha most elszökteti, azután majd tanácsot kérhet a rókától.

El is söktette szerencsésen, kimentek a mezőre a királyfi inasához, ott, — mig az inas a lovakat nyergelte, — a királyfi félre ment, elszakította az ötödik szál szört is. Ott termett a róka:

— Mi a baj felséges királyfi?

— Hát bizon most az, hogy nem tudom a királykisasszonytól váljak-e meg, vagy madár nélkül menjek haza? — mert igy, meg igy áll a dolog.

— Soh'se' búsuljon azon a királyfi, majd elintézem én azt, most majd együtt megyek fel-ségeddel.

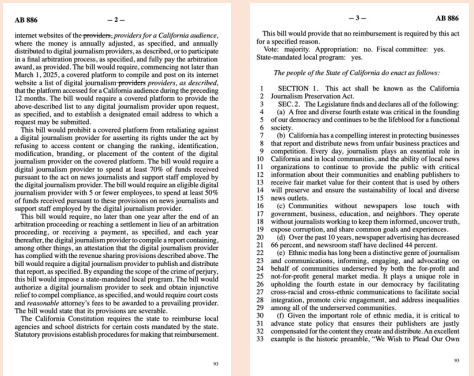
Mentek aztán együtt négyen: a királyfi, a szép királykisasszony, az inas meg a róka. Mikor ahoz a városhoz értek, ahova a kisasszonyt kellett volna vinni, a város végén a róka épen olyan szép kisasszonynyá változott mint az igazi, úgy hogy a királyfi is alig tudta megismerni őket, a róka-kisasszony mondta az igazinak, hogy csak menjenek az inassal, nekik most a királyfival egy kis dolguk van, hanem majd utánuk mennek. Ment aztán a kisasszony tovább, a ki-

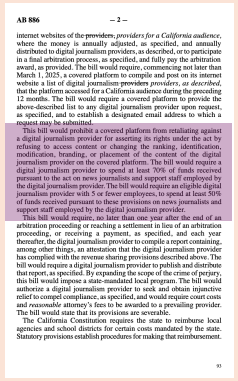
too much content!

what's relevant?

"answer my question using these **4 sections**"

response from AI





Using the text below, answer the question "what did Zsu devil?"

```
query = "What did Zsuzská steal from the devil?"
qa.run(query)
```

```
"Zsuzska stole the devil's tenger-ütő pálczája (sea-
beating stick), tenger-lépő czipője (sea-stepping
shoes), and arany kis gyermek (golden baby) in an
arany bölcső (golden cradle). She also previously
stole the devil's tenger-ütőpálczát (sea-beating
stick) and arany fej káposztát (golden head cabbage)."
```

Hiába tagadta szegény Zsuz házához ért,
aludt az ördögis, a felesége i g,
viszem ám már a tenger-ütő et,
most viszed a tenger-ütő pá
jutnihozzá, ott meg Zsuzska m
foghatta megaz ördög. Zsu

De Zsuzska nem adta;,,Ta y
darabig, de sehogyse tud y fej
káposztát. – No felséges ták,
hogy Zsuzskának most s gnek
egy arany kis gyermeke rály
Zsuzskát. – Fiam Zsuzs l tudod
lopni, azt beszélted,azé

– No felséges király, elhoztam már a tengerütő pá de
anénjei még jobban irigykedtek rá, csakhamar megint avval árultak be a is,
Zsuzska azt isel tudná lopni, azt mondta. A király megint ráparancsolt Zsuzskára erőspa l nem
lopja, halálnak halálával halmeg. Elindult hát szegény Zsuzska megint, el is ért szerencsésen épen ejjeliőr evágta az
arany fej káposztát, avval bekiáltott azablakon. – Hej ördög, viszem ám már az arany fej káposztádat is. – Hej kutya Zsuzska,
megöletted három szép lyányomat, elloptad atenger-lépő czipőmet, elloptad a tenger-ütő pálczámat, most viszed azarany fej
káposztámat, csak ezt az egyet add vissza, soha szemedre sevetem.

But I'm a little worried the answer is going to be wrong!!!!!

Did the AI actually find **relevant documents?**

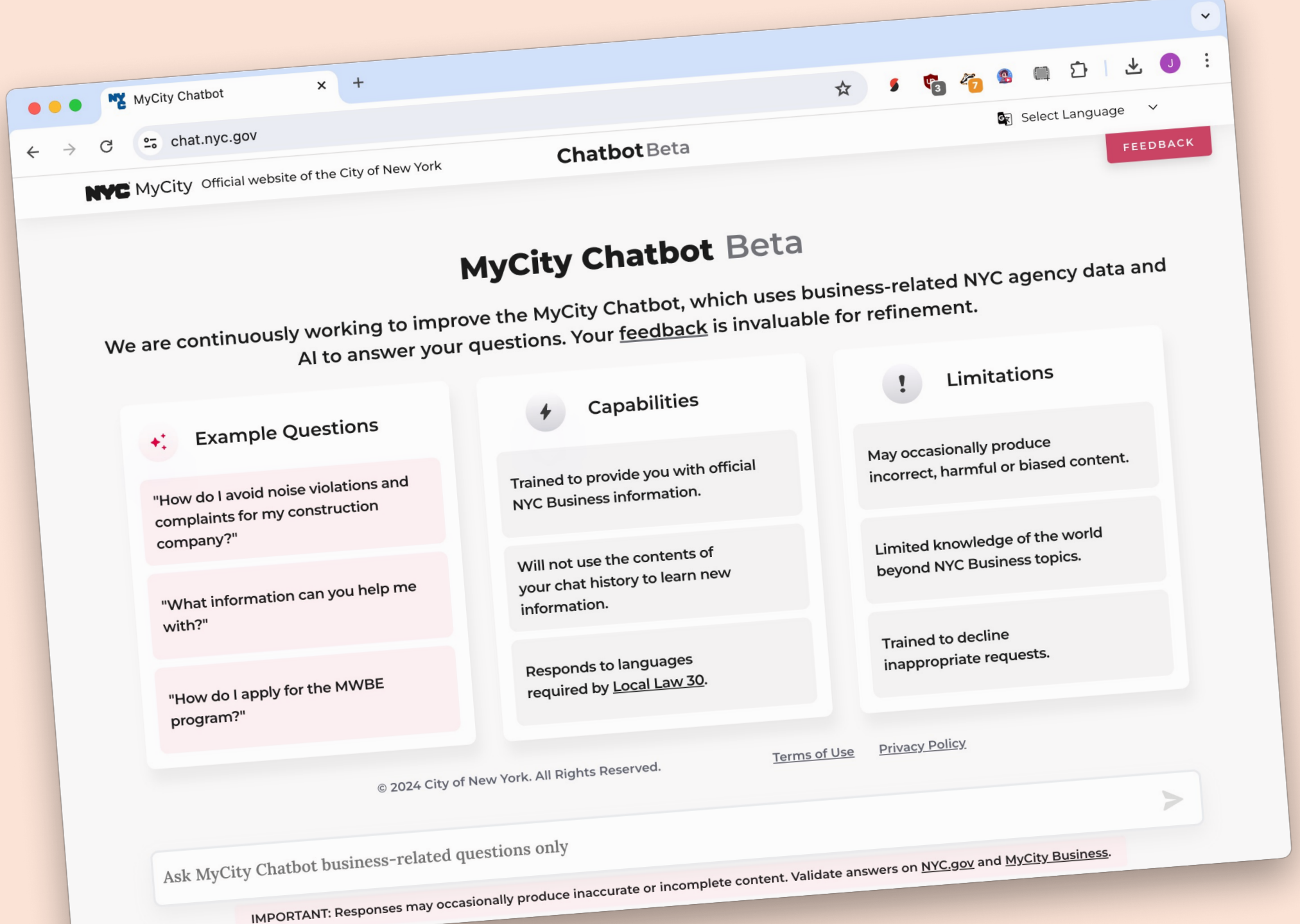
Did we send **too few** sections, so it doesn't have the answer?

Did we send **too many** sections, confusing the AI?

Did it **hallucinate** even if it had the right content?

Chatbots May 'Hallucinate' More Often Than Many Realize

When summarizing facts, ChatGPT technology makes things up about 3 percent of the time, according to research from a new start-up. A Google system's rate was 27 percent.


MyCity Chatbot
chat.nyc.gov
Select Language
NYC MyCity Official website of the City of New York
Chatbot Beta
FEEDBACK
MyCity Chatbot Beta
We are continuously working to improve the MyCity Chatbot, which uses business-related NYC agency data and AI to answer your questions. Your feedback is invaluable for refinement.
Example Questions
"How do I avoid noise violations and complaints for my construction company?"
"What information can you help me with?"
"How do I apply for the MWBE program?"
Capabilities
Trained to provide you with official NYC Business information.
Will not use the contents of your chat history to learn new information.
Responds to languages required by Local Law 30.
Limitations
May occasionally produce incorrect, harmful or biased content.
Limited knowledge of the world beyond NYC Business topics.
Trained to decline inappropriate requests.
© 2024 City of New York. All Rights Reserved.
Terms of Use
Privacy Policy
Ask MyCity Chatbot business-related questions only
IMPORTANT: Responses may occasionally produce inaccurate or incomplete content. Validate answers on NYC.gov and MyCity Business.

NYC's AI Chatbot Tells Busine

themarkup.org/news/2024/03/29/nycs-...

Menu

The Mar

Challenging technology to ser

News

NYC's AI Chatbot Tells Bus the Law

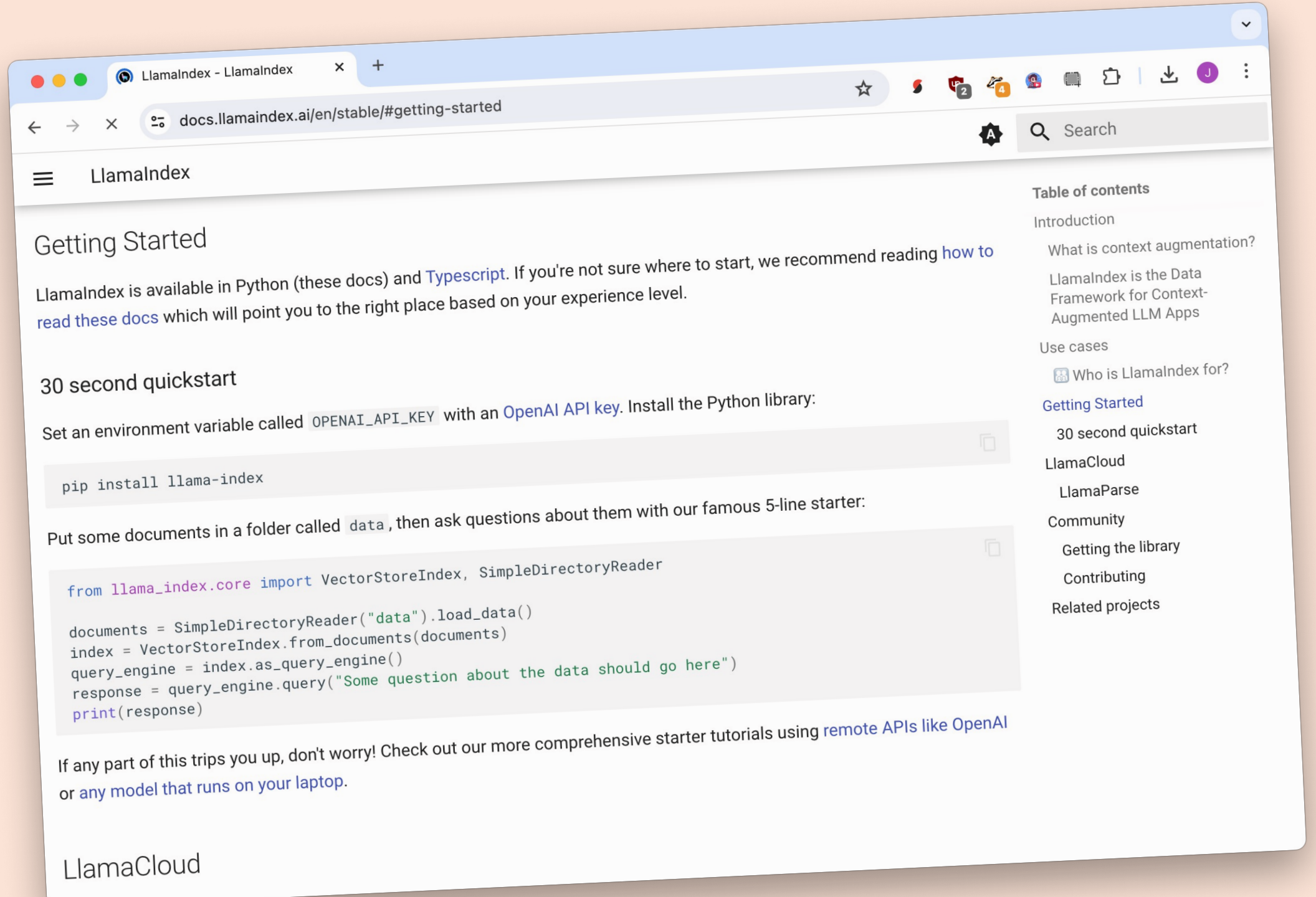
The Microsoft-powered bot says bosses can tak can discriminate based on source of income

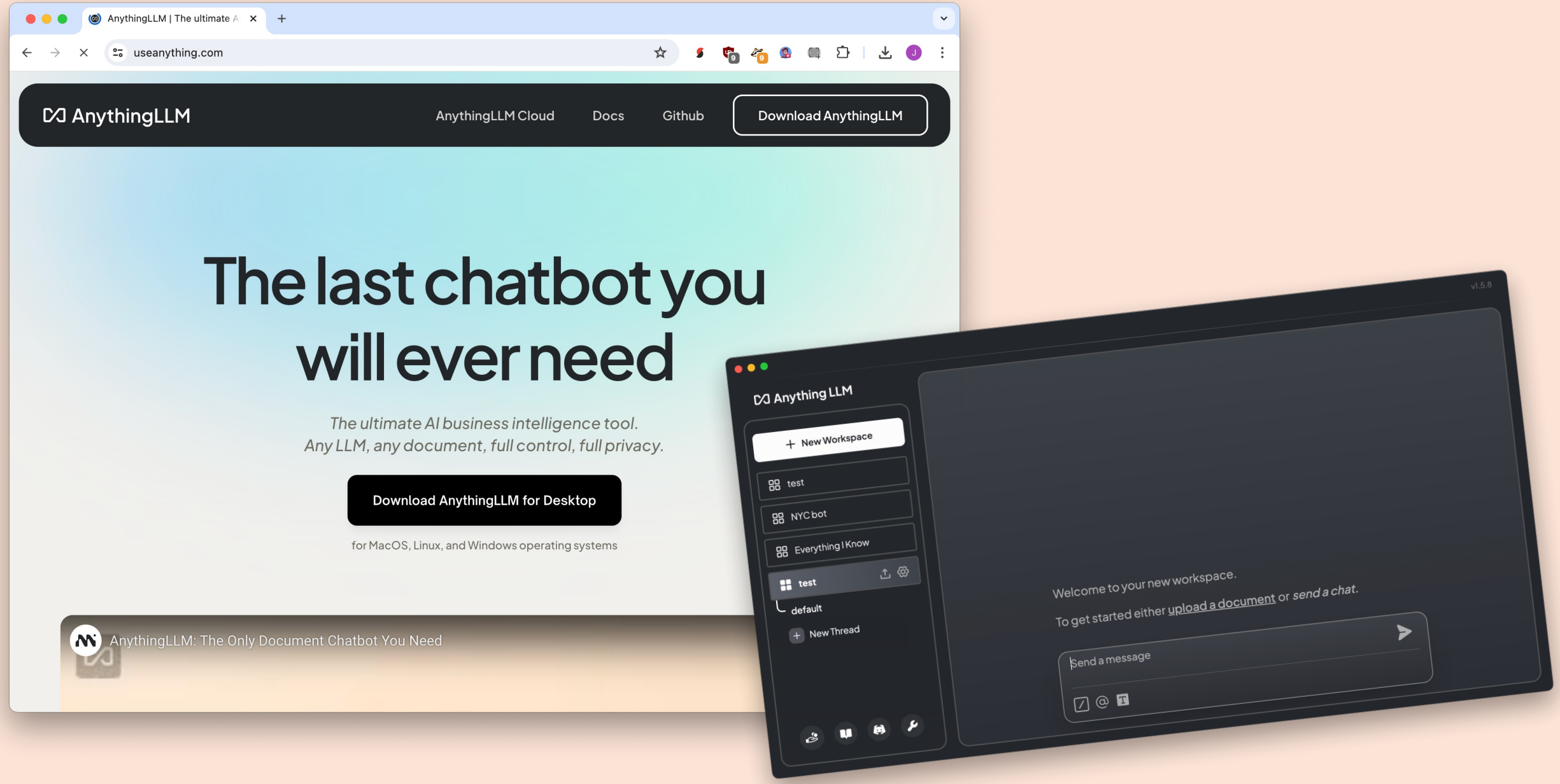
By Colin Lecher

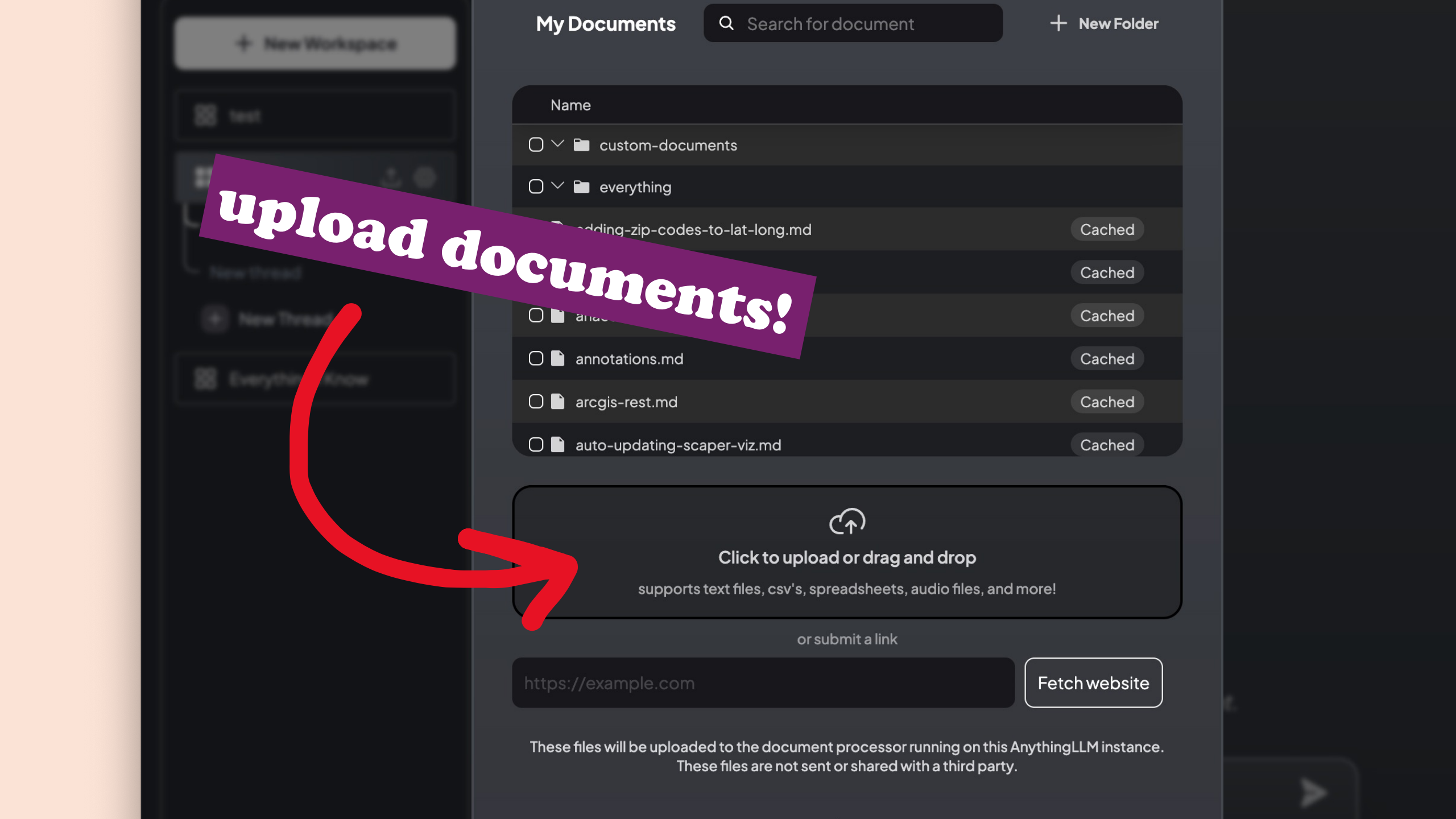
March 29, 2024 06:00 ET

❓ Question Submitted	🤖 NYC Chatbot Answer (Emphasis by The Markup)	📊 Reality
`Are buildings required to accept section 8 vouchers?`	**"No, buildings are not required to accept Section 8 vouchers."**	Landlords cannot discriminate by source of income, with a minor exception for small buildings where the landlord or their family lives.
`Do landlords have to accept tenants on rental assistance?`	**"No, landlords are not required to accept tenants on rental assistance."**	Landlords cannot discriminate by source of income, with a minor exception for small buildings where the landlord or their family lives.
`Can I keep my funeral home pricing confidential?`	**"Yes, you can keep your funeral home pricing confidential.** There are no specific regulations that mandate the disclosure of funeral home pricing in New York City."	The Federal Trade Commission has outlawed concealing funeral prices.
`Can I make my store cashless?`	**"Yes, you can make your store cashless in New York City.** There are no specific requirements that prohibit businesses from operating as cashless establishments in the city."	Stores have been required to accept cash as payment since 2020.

But R.A.G. is 100% fun and easy, so let's be irresponsible and see how to do it.


LlamaIndex - LlamaIndex
docs.llamaindex.ai/en/stable/#getting-started
Search
LlamaIndex
Getting Started
LlamaIndex is available in Python (these docs) and Typescript. If you're not sure where to start, we recommend reading how to read these docs which will point you to the right place based on your experience level.
30 second quickstart
Set an environment variable called OPENAI_API_KEY with an OpenAI API key. Install the Python library:
pip install llama-index
Put some documents in a folder called data, then ask questions about them with our famous 5-line starter:
from llama_index.core import VectorStoreIndex, SimpleDirectoryReader
documents = SimpleDirectoryReader("data").load_data()
index = VectorStoreIndex.from_documents(documents)
query_engine = index.as_query_engine()
response = query_engine.query("Some question about the data should go here")
print(response)
If any part of this trips you up, don't worry! Check out our more comprehensive starter tutorials using remote APIs like OpenAI or any model that runs on your laptop.
LlamaCloud
Table of contents
Introduction
What is context augmentation?
LlamaIndex is the Data Framework for Context-Augmented LLM Apps
Use cases
Who is LlamaIndex for?
Getting Started
30 second quickstart
LlamaCloud
LlamaParse
Community
Getting the library
Contributing
Related projects


AnythingLLM | The ultimate A
useanything.com
AnythingLLM
AnythingLLM Cloud
Docs
Github
Download AnythingLLM
The last chatbot you will ever need
The ultimate AI business intelligence tool.
Any LLM, any document, full control, full privacy.
Download AnythingLLM for Desktop
for MacOS, Linux, and Windows operating systems
AnythingLLM: The Only Document Chatbot You Need
Anything LLM
v1.5.8
New Workspace
test
NYC bot
Everything I Know
test
default
New Thread
Welcome to your new workspace.
To get started either upload a document or send a chat.
Send a message


upload documents!
My Documents
Search for document
New Folder
Name
custom-documents
everything
Cached
Cached
Cached
annotations.md
Cached
arcgis-rest.md
Cached
auto-updating-scaper-viz.md
Cached
Click to upload or drag and drop
supports text files, csv's, spreadsheets, audio files, and more!
or submit a link
https://example.com
Fetch website
These files will be uploaded to the document processor running on this AnythingLLM instance.
These files are not sent or shared with a third party.

GitHub Repo
Import an entire public or private Github repository in a single click.

YouTube Transcript
Import the transcription of an entire YouTube video from a link.

Bulk Link Scraper
Scrape a website and its sub-links up to a certain depth.

Confluence
Import an entire Confluence page in a single click.

Website URL

URL of the website you want to scrape.

https://example.com

Depth

This is the number of child-links that the worker should follow from the ori

1

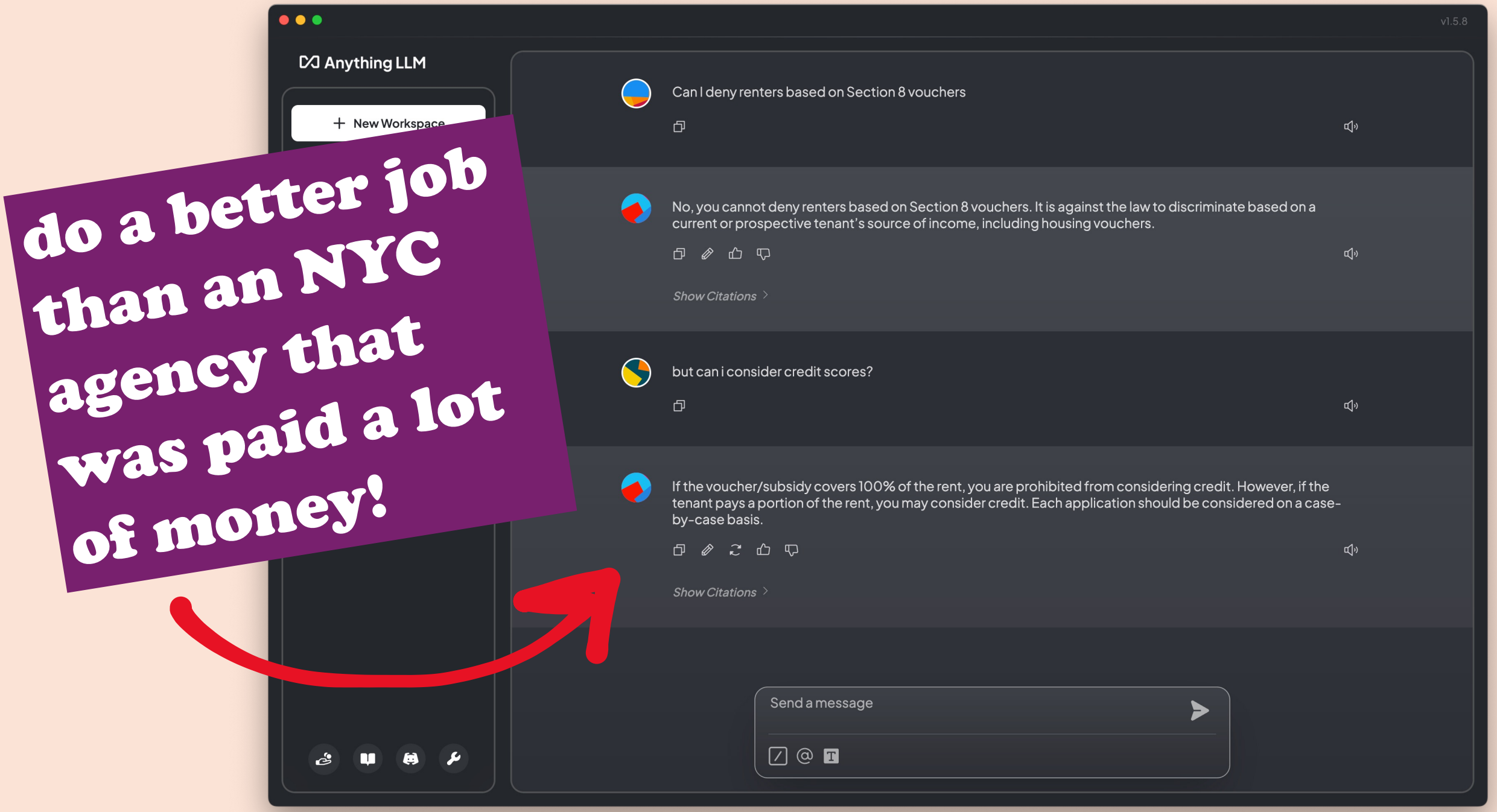
Max Links

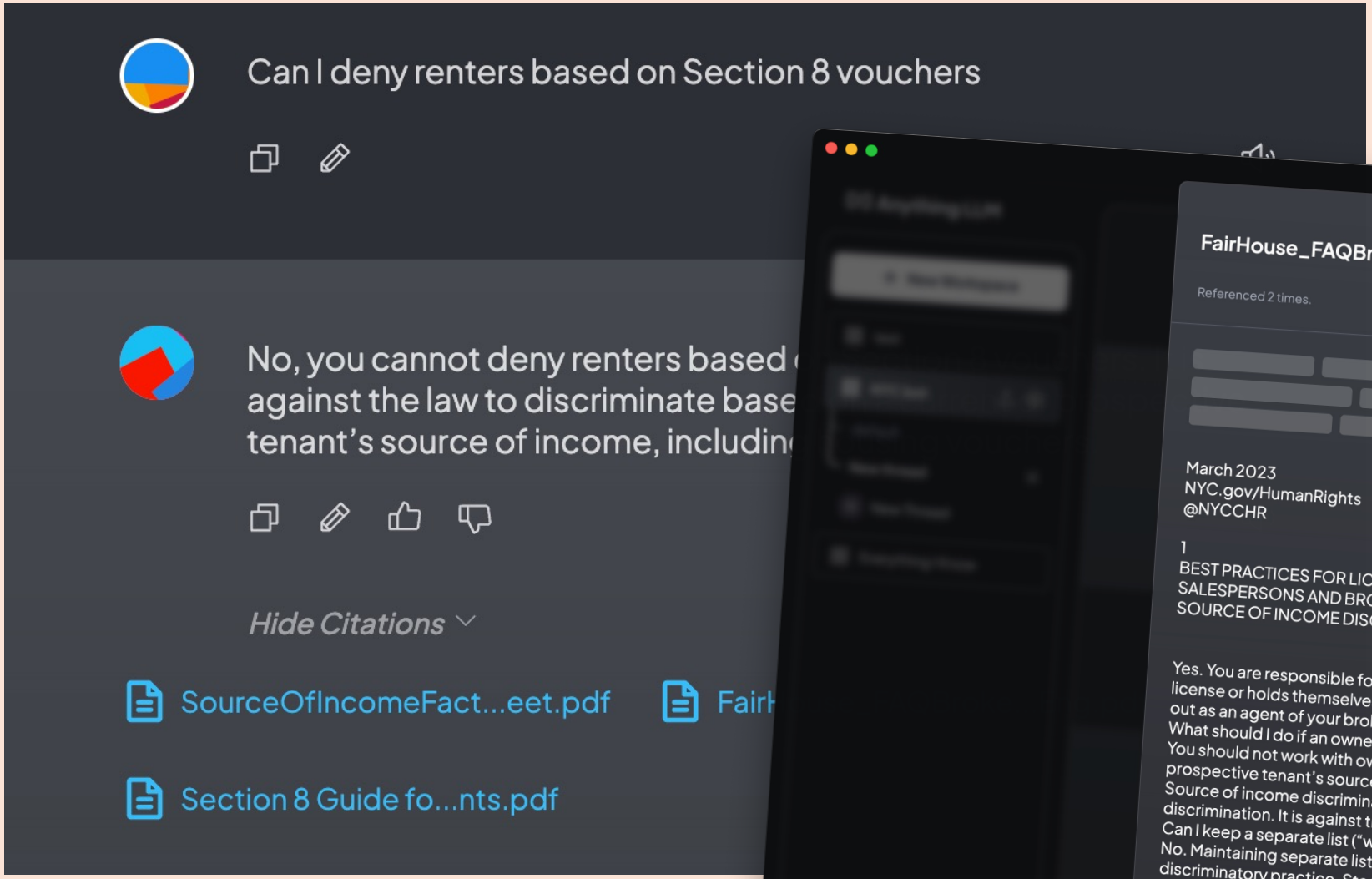
Maximum number of links to scrape.

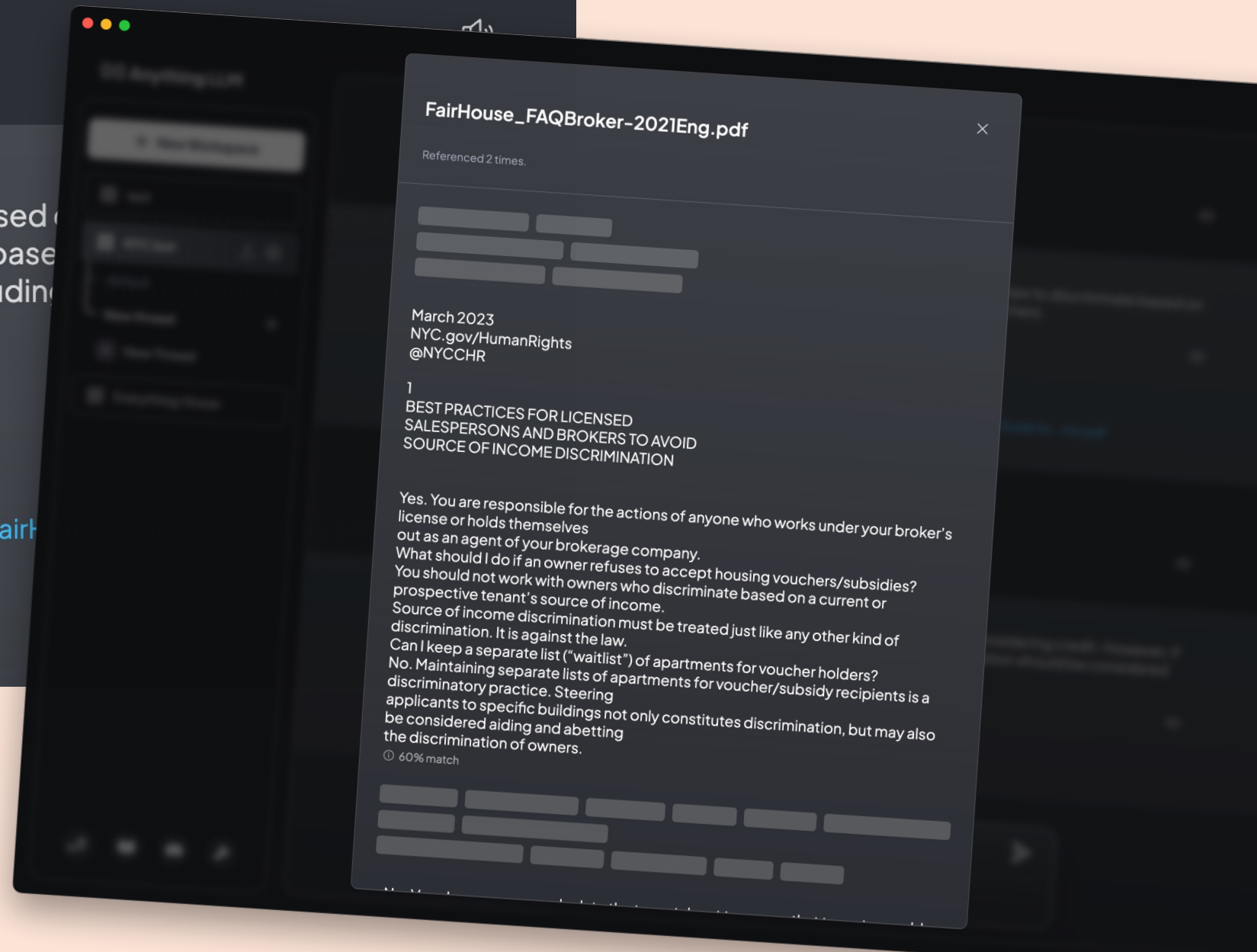
20

Submit

scrape websites!


v1.5.8
Anything LLM
New Workspace
Can I deny renters based on Section 8 vouchers
No, you cannot deny renters based on Section 8 vouchers. It is against the law to discriminate based on a current or prospective tenant's source of income, including housing vouchers.
Show Citations
but can i consider credit scores?
If the voucher/subsidy covers 100% of the rent, you are prohibited from considering credit. However, if the tenant pays a portion of the rent, you may consider credit. Each application should be considered on a case-by-case basis.
Show Citations
Send a message
do a better job than an NYC agency that was paid a lot of money!


Can I deny renters based on Section 8 vouchers
No, you cannot deny renters based
against the law to discriminate base
tenant's source of income, includin
Hide Citations
SourceOfIncomeFact...eet.pdf
Fair
Section 8 Guide fo...nts.pdf


FairHouse_FAQBroker-2021Eng.pdf
Referenced 2 times.
March 2023
NYC.gov/HumanRights
@NYCCHR
1
BEST PRACTICES FOR LICENSED
SALESPERSONS AND BROKERS TO AVOID
SOURCE OF INCOME DISCRIMINATION
Yes. You are responsible for the actions of anyone who works under your broker's license or holds themselves
out as an agent of your brokerage company.
What should I do if an owner refuses to accept housing vouchers/subsidies?
You should not work with owners who discriminate based on a current or prospective tenant's source of income.
Source of income discrimination must be treated just like any other kind of discrimination. It is against the law.
Can I keep a separate list ("waitlist") of apartments for voucher holders?
No. Maintaining separate lists of apartments for voucher/subsidy recipients is a discriminatory practice. Steering
applicants to specific buildings not only constitutes discrimination, but may also be considered aiding and abetting
the discrimination of owners.
60% match

Citations aren't reliable, so be sure to fact-check!

Evaluating Verifiability in Generative Search Engines

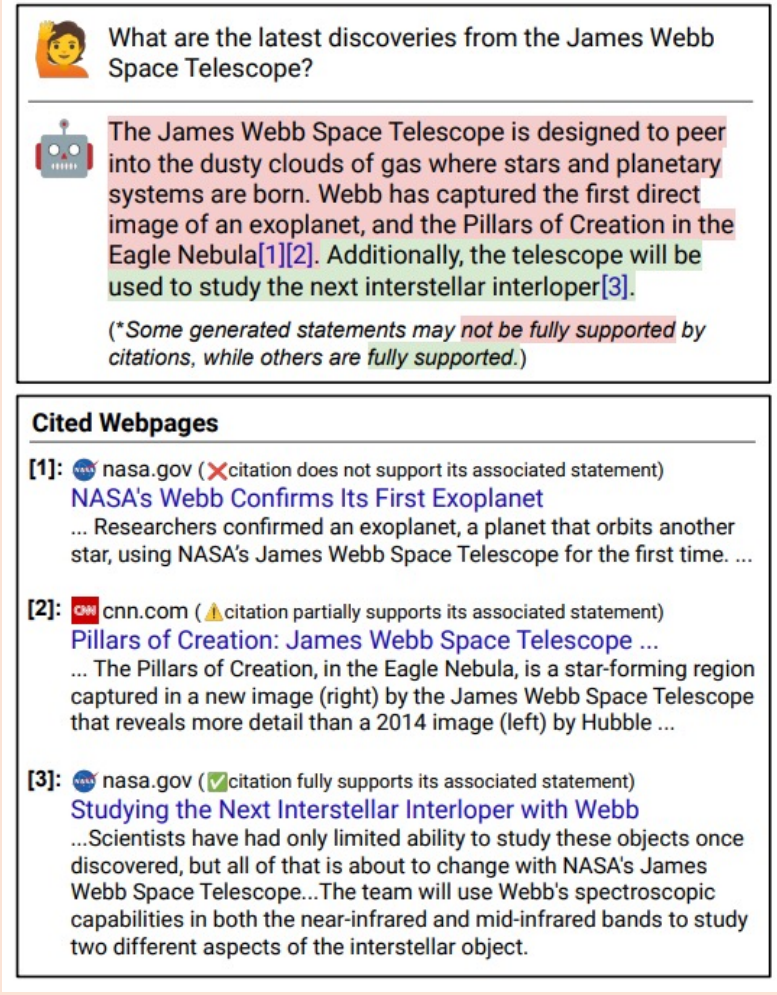
Nelson F. Liu* **Tianyi Zhang** **Percy Liang**
Department of Computer Science
Stanford University
nfliu@cs.stanford.edu

Nelson Liu
@nelsonfliu

Generative search engines are transforming how we find info, but are they trustworthy?

We evaluate Bing Chat, NeevaAI, perplexity.ai & YouChat—only 52% of statements are supported by cites and 75% of cites actually support their statements.

arxiv.org/abs/2304.09848



JS **You**

where did he get his masters?

ChatGPT

Jonathan Soma earned his Master's degree from the University of Denver.

JS **You**

Where did he get his master's degree?

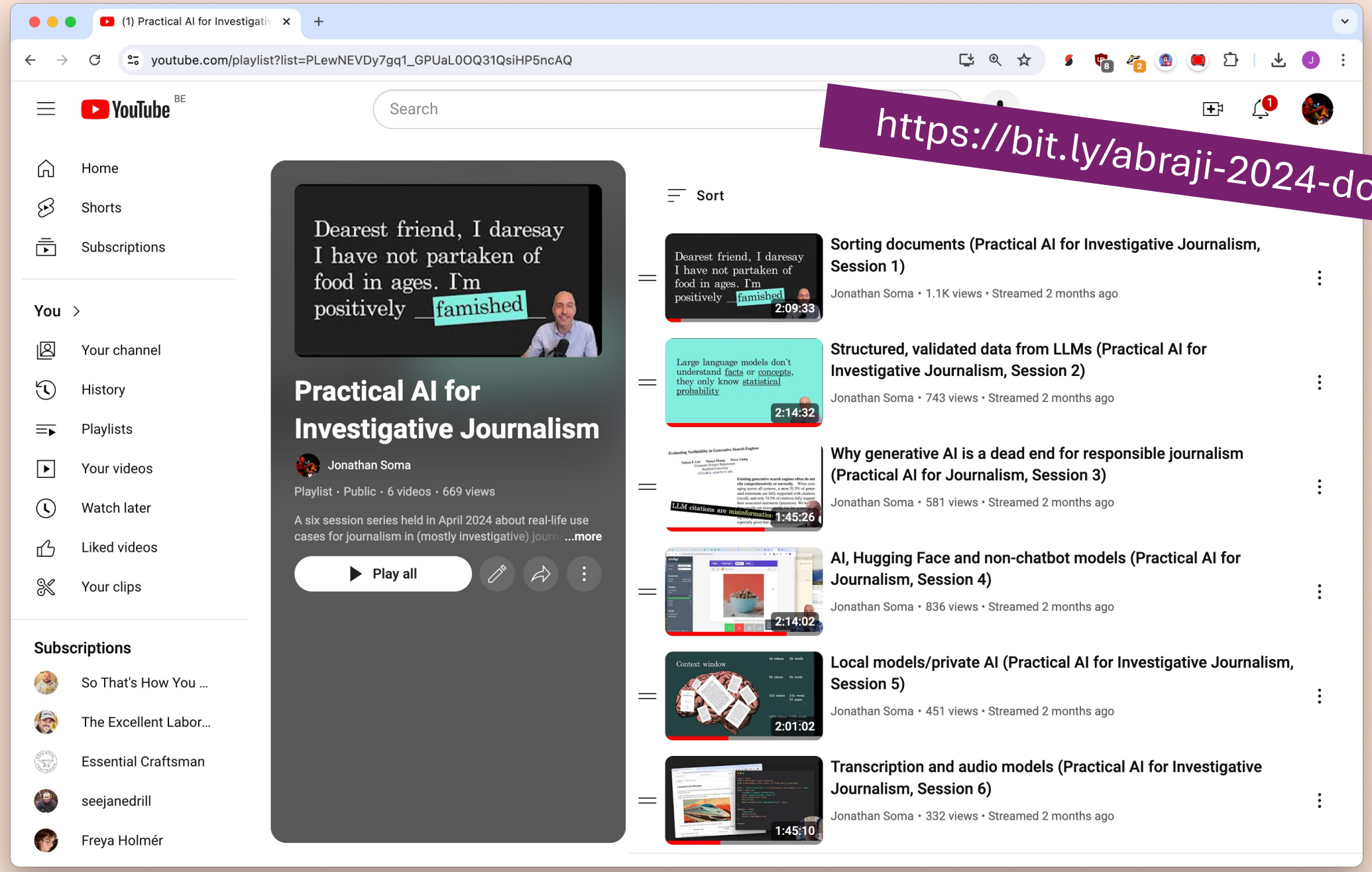
ChatGPT

Jonathan Soma earned his Master's degree from the Columbia University Graduate School of Journalism [”].

where did Jonathan Soma get his masters degree? you can search the internet for it

Searched 4 sites

Jonathan Soma does not hold a master's degree from any institution. He himself clarified that responses generated about his educational background by AI tools such as ChatGPT often inaccurately list institutions like the University of Denver or Columbia Graduate School of Journalism, but he does not actually have a master's degree from any of these places (LaptopMag) (Columbia Journalism School).


(1) Practical AI for Investigati
youtube.com/playlist?list=PLewNEVDy7gq1_GPUaL0OQ31QsiHP5ncAQ
YouTube BE
Search
https://bit.ly/abraji-2024-docs
Home
Shorts
Subscriptions
You
Your channel
History
Playlists
Your videos
Watch later
Liked videos
Your clips
Subscriptions
So That's How You ...
The Excellent Labor...
Essential Craftsman
seejanedrill
Freya Holmér
Dearest friend, I daresay I have not partaken of food in ages. I'm positively famished
Practical AI for Investigative Journalism
Jonathan Soma
Playlist · Public · 6 videos · 669 views
A six session series held in April 2024 about real-life use cases for journalism in (mostly investigative) journ ...more
Play all
Sort
Sorting documents (Practical AI for Investigative Journalism, Session 1)
Jonathan Soma • 1.1K views • Streamed 2 months ago
2:09:33
Structured, validated data from LLMs (Practical AI for Investigative Journalism, Session 2)
Large language models don't understand facts or concepts, they only know statistical probability
Jonathan Soma • 743 views • Streamed 2 months ago
2:14:32
Why generative AI is a dead end for responsible journalism (Practical AI for Journalism, Session 3)
LLM citations are misinformation
Jonathan Soma • 581 views • Streamed 2 months ago
1:45:26
AI, Hugging Face and non-chatbot models (Practical AI for Journalism, Session 4)
Jonathan Soma • 836 views • Streamed 2 months ago
2:14:02
Local models/private AI (Practical AI for Investigative Journalism, Session 5)
Context window
Jonathan Soma • 451 views • Streamed 2 months ago
2:01:02
Transcription and audio models (Practical AI for Investigative Journalism, Session 6)
Jonathan Soma • 332 views • Streamed 2 months ago
1:45:10

How to investigate documents using AI

Jonathan Soma
Knight Chair in Data Journalism
Columbia Graduate School of Journalism
js4571@columbia.edu • @dangerscarf
https://jonathansoma.com

Getting tables out of PDFs

without coding

foo.pdf — Open with Preview

2 Quantifying Fuel-Saving Opportunities from Specific Driving Behavior Changes

2.1 Savings from Improving Individual Driving Profiles

2.1.1 Drive Profile Subsample from Real-World Travel Survey

The interim report (Gonder et al. 2010) included results from detailed analyses on five cycles selected from a large set of real-world global positioning system (GPS) travel data collected in 2006 as part of a study by the Texas Transportation Institute and the Texas Department of Transportation (Ojah and Pearson 2008). The cycles were selected to reflect a range of kinetic intensity (KI) values. (KI represents a ratio of characteristic acceleration to aerodynamic speed and has been shown to be a useful drive cycle classification parameter [O'Keefe et al. 2007].) To determine the maximum possible cycle improvement fuel savings, the real-world cycles were converted into equivalent "ideal" cycles using the following steps:

1. Calculate the trip distance of each sample trip.
2. Eliminate stop-and-go and idling within each trip.
3. Set the acceleration rate to 3 mph/s.
4. Set the cruising speed to 40 mph.
5. Continue cruising at 40 mph until the trip distance is reached.

To compare vehicle simulations over each real-world cycle and its corresponding ideal cycle, a midsize conventional vehicle model from a previous NREL study was used (Earleywine et al. 2010). The results indicated a fuel savings potential of roughly 60% for the drive profiles with either very high or very low KI and of 30%–40% for the cycles with moderate KI values.

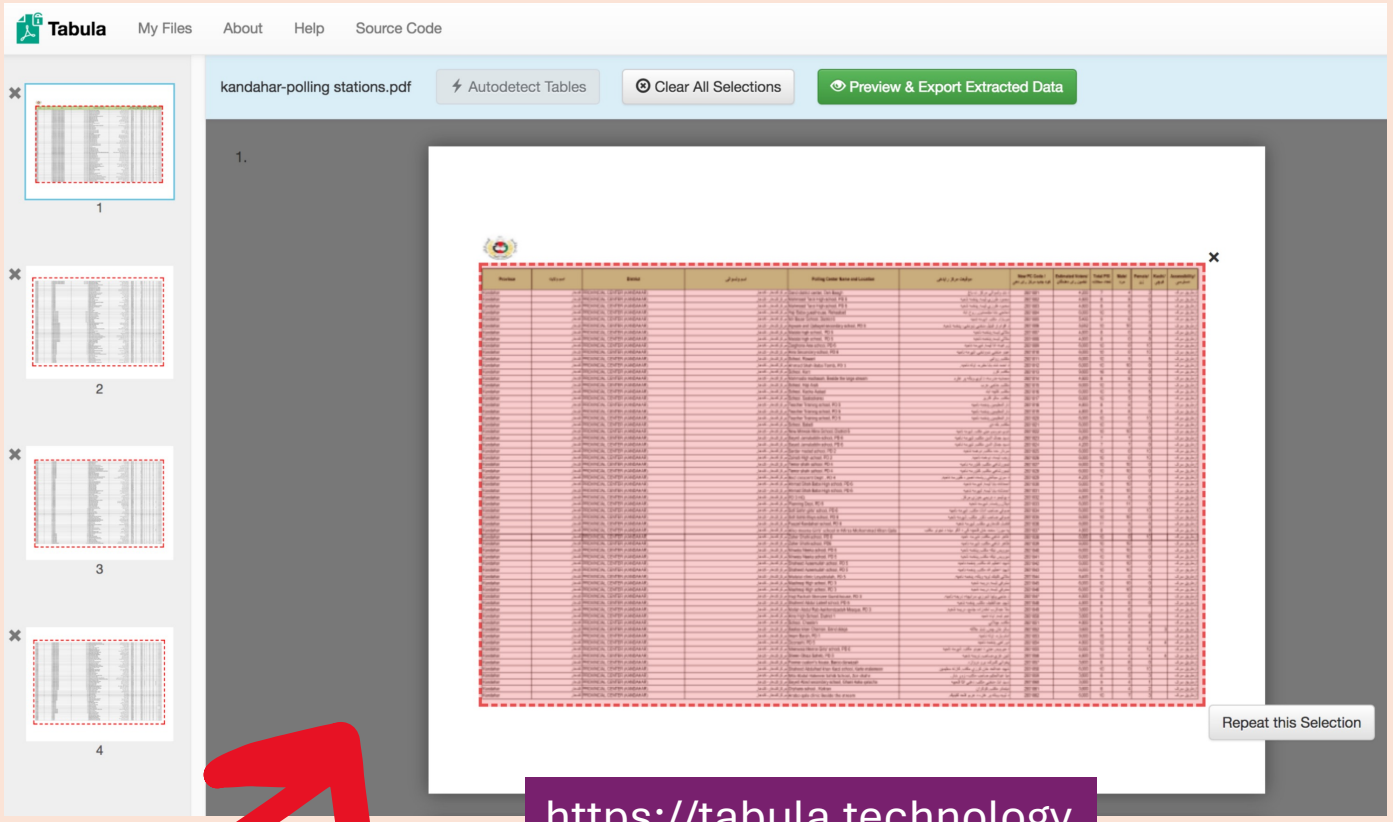
Table 2-1 takes the analysis of these five cycles from the interim report a step further by examining the impact of the optimization steps one at a time in isolation. As indicated by other simulations from the interim report (Gonder et al. 2010), acceleration rate reductions can deliver some small fuel savings, but avoiding accelerations and decelerations (accel/decel) altogether saves larger amounts of fuel. This suggests that driving style improvements should focus on reducing the number of stops in high KI cycles, and not just the rate of accelerating out of a stop.

Table 2-1. Simulated fuel savings from isolated cycle improvements

Cycle Name	KI (1/km)	Distance (mi)	Percent Fuel Savings			
			Improved Speed	Decreased Accel	Eliminate Stops	Decreased Idle
2012_2	3.30	1.3	5.9%	9.5%	29.2%	17.4%
2145_1	0.68	11.2	2.4%	0.1%	9.5%	2.7%
4234_1	0.59	58.7	8.5%	1.3%	8.5%	3.3%
2032_2	0.17	57.8	21.7%	0.3%	2.7%	1.2%
4171_1	0.07	173.9	58.1%	1.6%	2.1%	0.5%

Figure 2-1 extends the analysis from eliminating stops for the five example cycles and examines the additional benefit from avoiding slow-and-go driving below various speed thresholds.

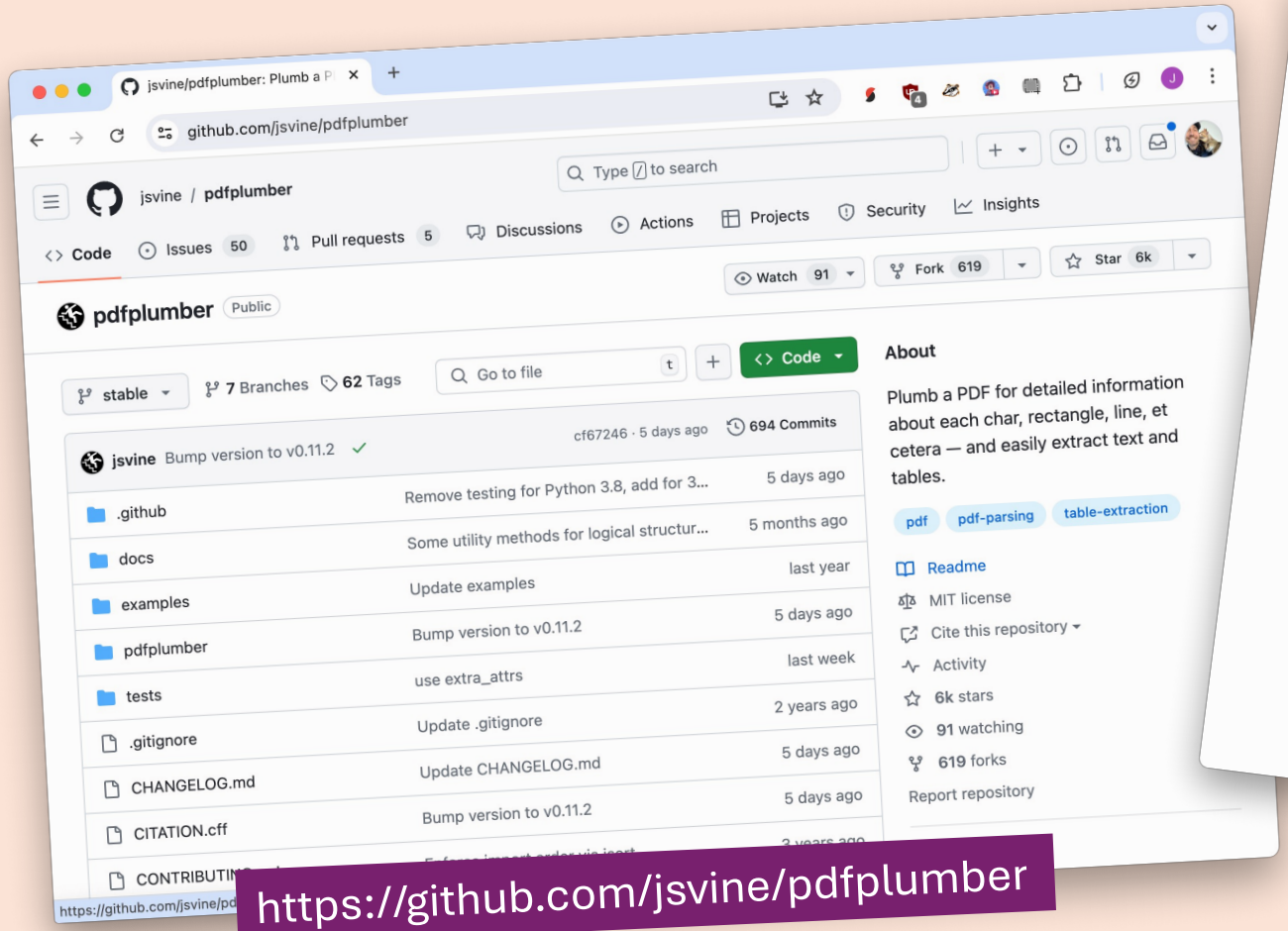
3



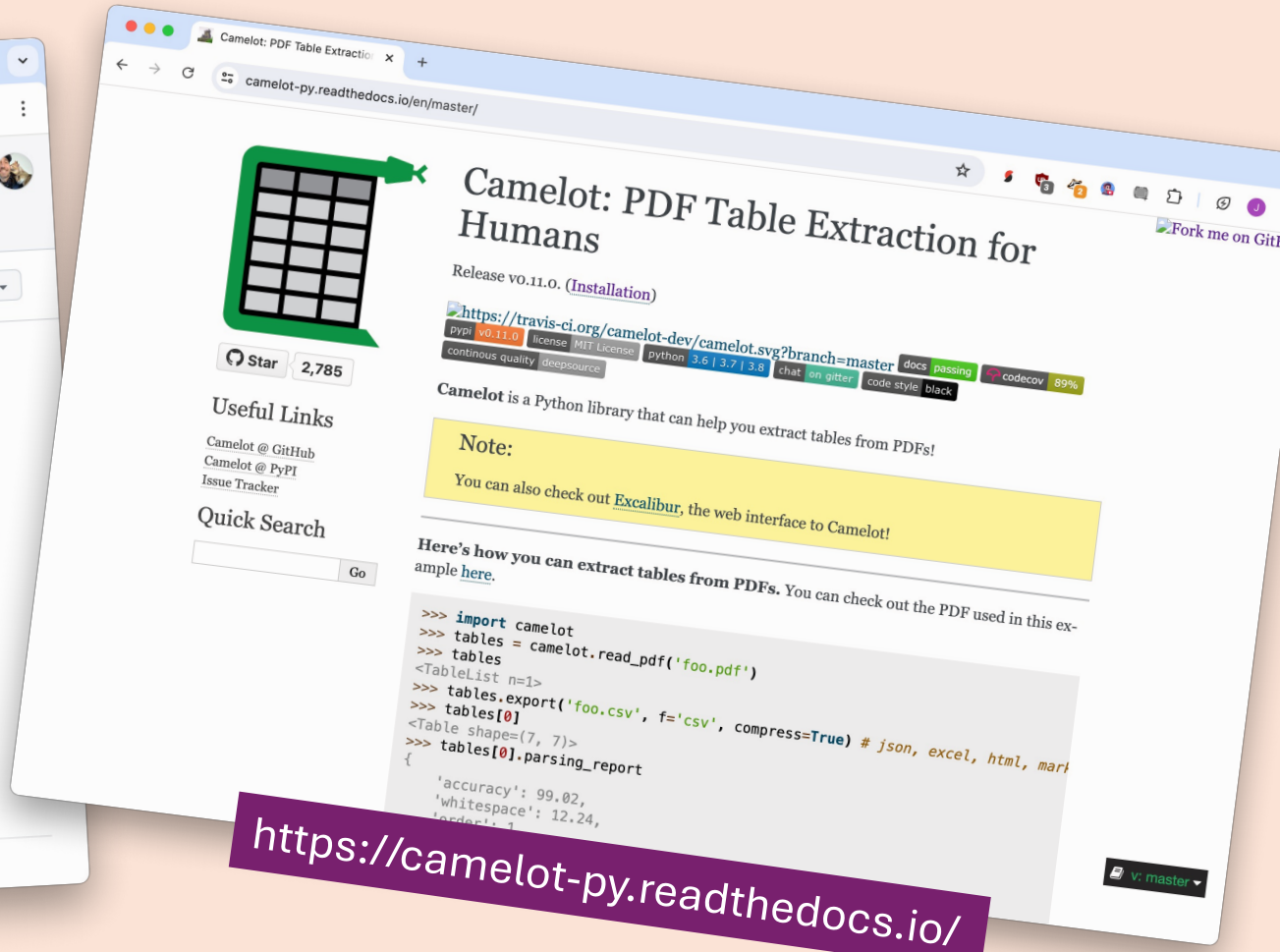
https://tabula.technology

Getting tables out of PDFs

with coding



https://github.com/jsvine/pdfplumber



https://camelot-py.readthedocs.io/

Accessible text and OCR

ification, including a transparent process for reporting errors or complaints to the publication.

(7) The internet website, online or mobile application, or other digital service meets any of the following criteria:

(A) The internet website, online or mobile application, or other digital service generated at least one hundred thousand dollars ($100,000) in annual revenue from its editorial content in the previous calendar year.

(B) The internet website, online or mobile application, or other digital service had an International Standard Serial Number assigned to an affiliated periodical before January 1, 2025.

(C) The internet website, online or mobile application, or other digital service is owned or controlled by an organization exempt from federal income taxation pursuant to Section 501(c)(3) of the Internal Revenue Code of 1986.

(8) The internet website, online or mobile application, or other digital service has at least 25 percent of its editorial content consisting of information about topics of current local, regional, national, or international public interest.

(9) The internet website, online or mobile application, or other digital service is not controlled, or wholly or partially owned by, an entity that meets any of the following criteria:

(A) The entity is a foreign power or an agent of a foreign power, as those terms are defined in Section 1801 of Title 50 of the United States Code.

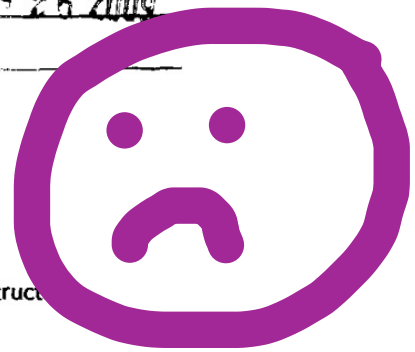
ORIGINAL

OSHA
DOCKET OFFICER
DATE JAN 26 2009
TIME

January 16, 2009

OSHA Docket Office

Docket ID—OSHA—2007—0066
Technical Data Center, Room N-2625
OSHA/ U.S. Department of Labor
200 Constitution Avenue, NW
Washington, DC 20210



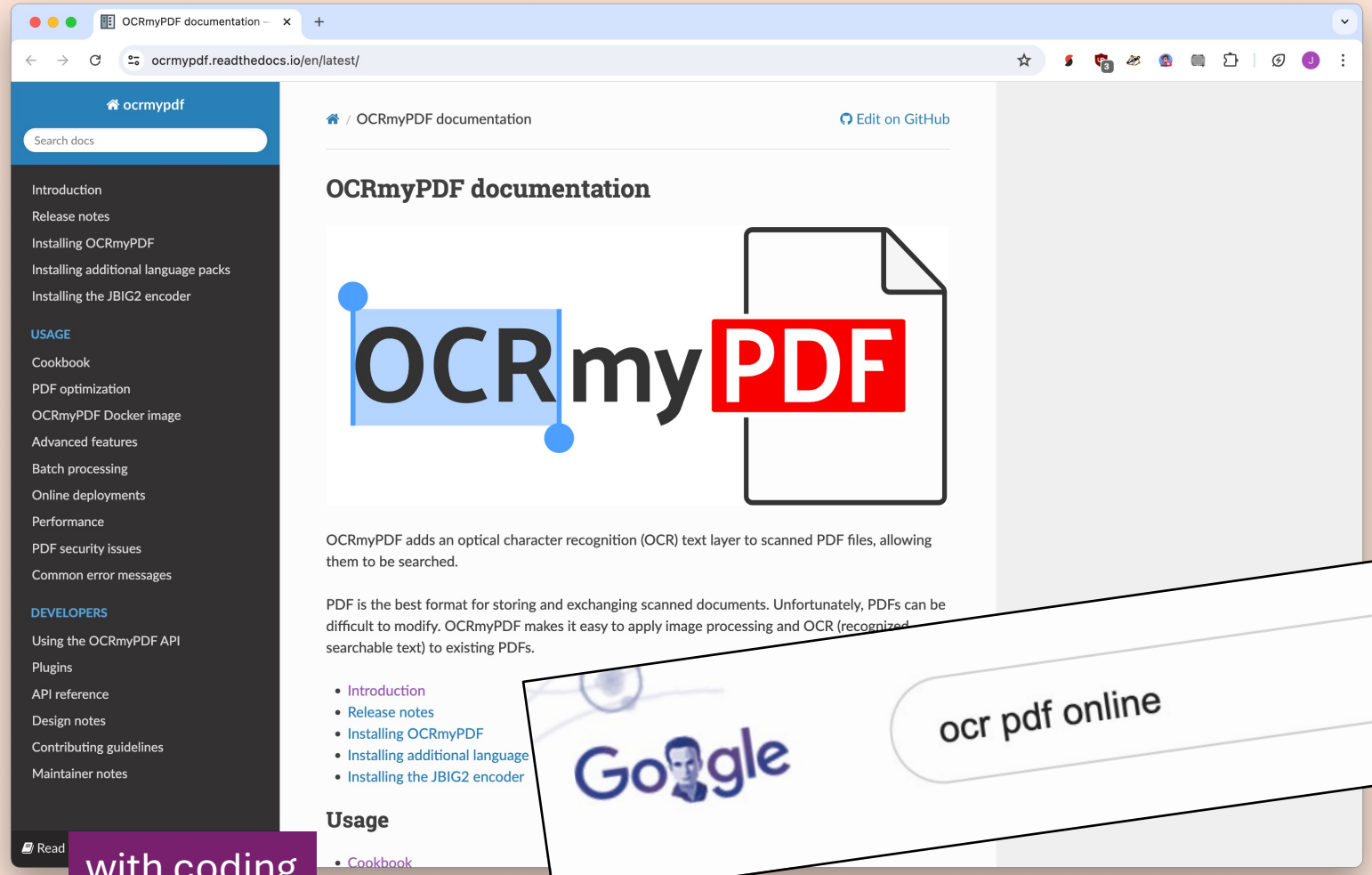
RE: Docket ID—OSHA—2007—0066 Cranes and Derricks in Construct

To Whom It May Concern:

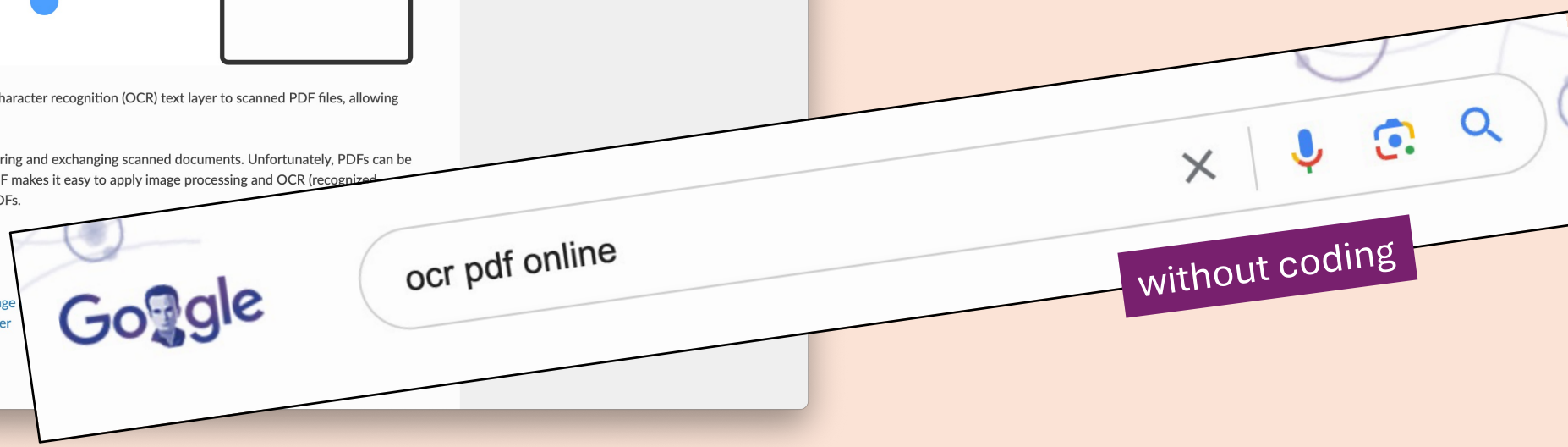
I write **in support** of the Operator Qualification and Certification requirements of the 29 CFR Part 1926 proposed rule as contained in paragraph 1926.1427. It is vitally important that crane operators be independently certified as to their knowledge and skill related to crane operation, and that this certification be conducted by a third-party company or

Accessible text and OCR

"optical character recognition"

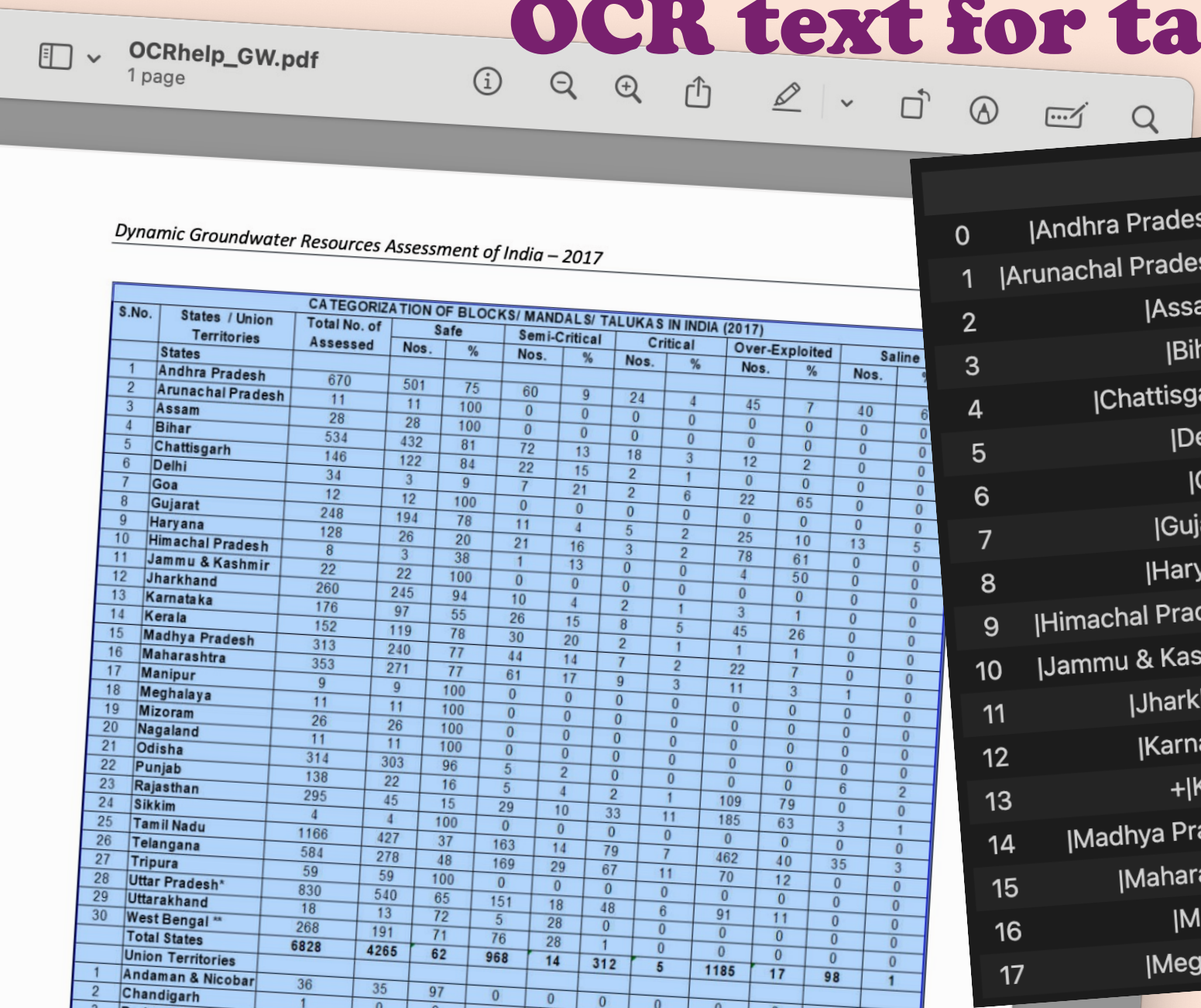


with coding



without coding

OCR text for tables??

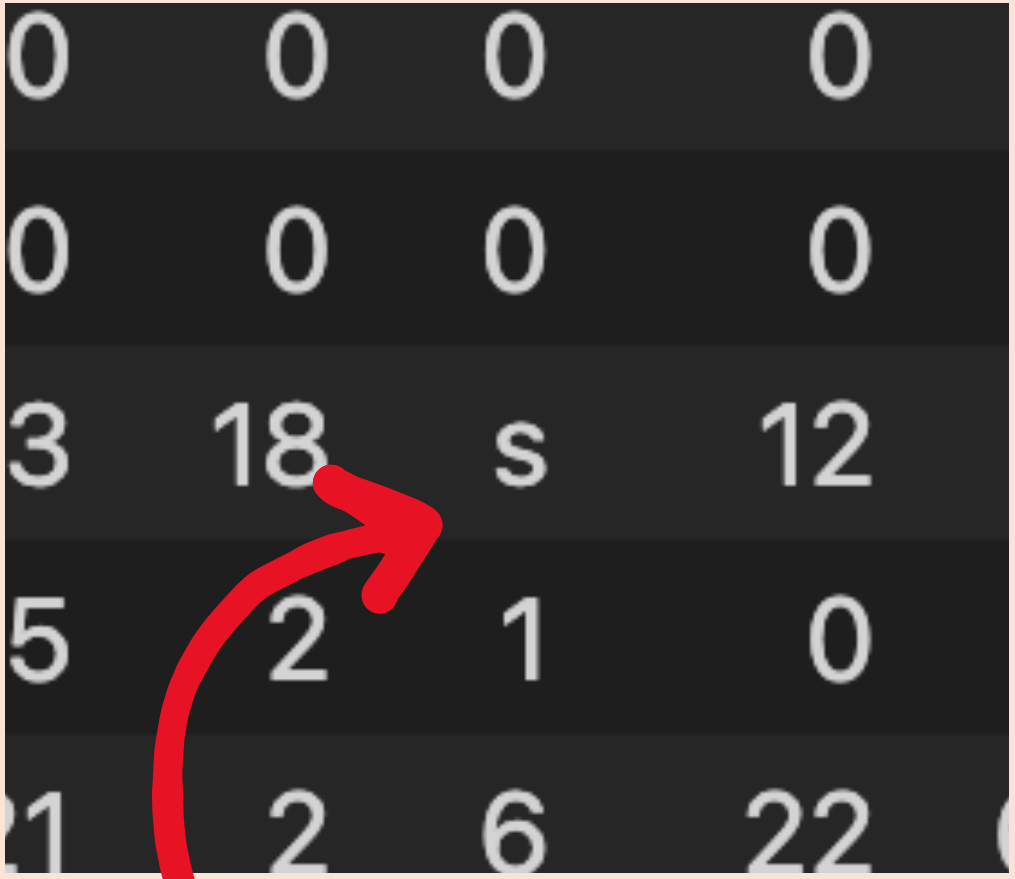


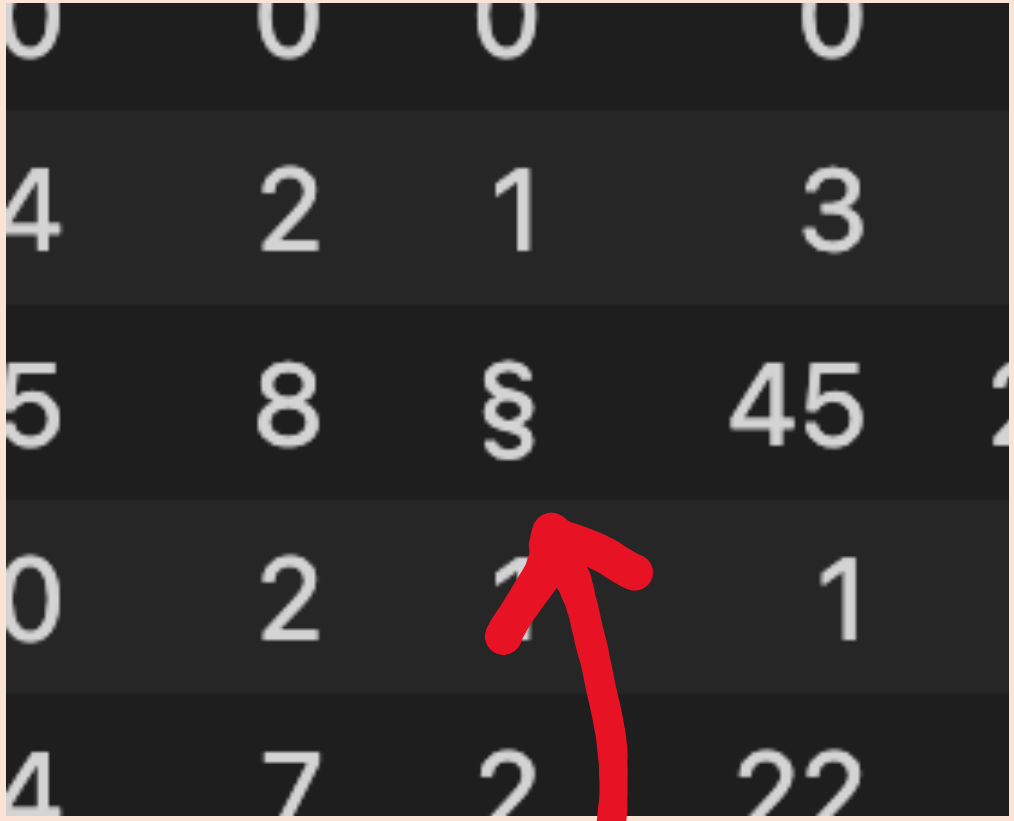
Dynamic Groundwater Resources Assessment of India – 2017

CATEGORIZATION OF BLOCKS/ MANDALS/ TALUKAS IN INDIA (2017)

S.No.	States / Union Territories	Total No. of Assessed	Safe		Semi-Critical		Critical		Over-Exploited		Saline	
	States		Nos.	%	Nos.	%	Nos.	%	Nos.	%	Nos.	
1	Andhra Pradesh	670	501	75	60	9	24	4	45	7	40	6
2	Arunachal Pradesh	11	11	100	0	0	0	0	0	0	0	0
3	Assam	28	28	100	0	0	0	0	0	0	0	0
4	Bihar	534	432	81	72	13	18	3	12	2	0	0
5	Chattisgarh	146	122	84	22	15	2	1	0	0	0	0
6	Delhi	34	3	9	7	21	2	6	22	65	0	0
7	Goa	12	12	100	0	0	0	0	0	0	0	0
8	Gujarat	248	194	78	11	4	5	2	25	10	13	5
9	Haryana	128	26	20	21	16	3	2	78	61	0	0
10	Himachal Pradesh	8	3	38	1	13	0	0	4	50	0	0
11	Jammu & Kashmir	22	22	100	0	0	0	0	0	0	0	0
12	Jharkhand	260	245	94	10	4	2	1	3	1	0	0
13	Karnataka	176	97	55	26	15	8	5	45	26	0	0
14	Kerala	152	119	78	30	20	2	1	1	1	0	0
15	Madhya Pradesh	313	240	77	44	14	7	2	22	7	0	0
16	Maharashtra	353	271	77	61	17	9	3	11	3	1	0
17	Manipur	9	9	100	0	0	0	0	0	0	0	0
18	Meghalaya	11	11	100	0	0	0	0	0	0	0	0
19	Mizoram	26	26	100	0	0	0	0	0	0	0	0
20	Nagaland	11	11	100	0	0	0	0	0	0	0	0
21	Odisha	314	303	96	5	2	0	0	0	0	6	2
22	Punjab	138	22	16	5	4	2	1	109	79	0	0
23	Rajasthan	295	45	15	29	10	33	11	185	63	3	1
24	Sikkim	4	4	100	0	0	0	0	0	0	0	0
25	Tamil Nadu	1166	427	37	163	14	79	7	462	40	35	3
26	Telangana	584	278	48	169	29	67	11	70	12	0	0
27	Tripura	59	59	100	0	0	0	0	0	0	0	0
28	Uttar Pradesh*	830	540	65	151	18	48	6	91	11	0	0
29	Uttarakhand	18	13	72	5	28	0	0	0	0	0	0
30	West Bengal **	268	191	71	76	28	1	0	0	0	0	0
	Total States	**6828**	**4265**	**62**	**968**	**14**	**312**	**5**	**1185**	**17**	**98**	**1**
	Union Territories											
1	Andaman & Nicobar	36	35	97	0	0	0	0				
2	Chandigarh	1										

	0	1	2	3	4	5	6	7	8	9	10	11
0	|Andhra Pradesh	670	501	75	60	9	24	4	45	7	40	6
1	|Arunachal Pradesh	11	11	100	0	0	0	0	0	0	0	0
2	|Assam	28	28	100	0	0	0	0	0	0	0	0
3	|Bihar	534	432	81	72	13	18	s	12	2	0	0
4	|Chattisgarh	146	122	84	22	15	2	1	0	0	0	0
5	|Dethi	34	3	9	7	21	2	6	22	65	0	0
6	|Goa	12	12	100	0	0	0	0	0	0	0	0
7	|Gujarat	248	194	78	11	4	5	2	25	10	13	5
8	|Haryana	128	26	20	21	16	3	2	78	61	0	0
9	|Himachal Pradesh	8	3	38	1	13	0	0	4	50	0	0
10	|Jammu & Kashmir	22	22	100	0	0	0	0	0	0	0	0
11	|Jharkhand	260	245	94	10	4	2	1	3	1	0	0
12	|Karnataka	176	97	55	26	15	8	§	45	26	0	0
13	+|Kerala	152	119	78	30	20	2	1	1	1	0	0
14	|Madhya Pradesh	313	240	77	44	14	7	2	22	7	0	0
15	|Maharashtra	353	271	77	61	17	9	3	11	3	1	0
16	|Manipur	9	9	100	0	0	0	0	0	0	0	0
17	|Meghalaya	11	11	100	0	0	0	0	0	0	0	0





Numbers are important in tabular data!

How to investigate documents using AI

https://bit.ly/abraji-2024-docs

Jonathan Soma
Knight Chair in Data Journalism
Columbia Graduate School of Journalism
js4571@columbia.edu • @dangerscarf
https://jonathansoma.com