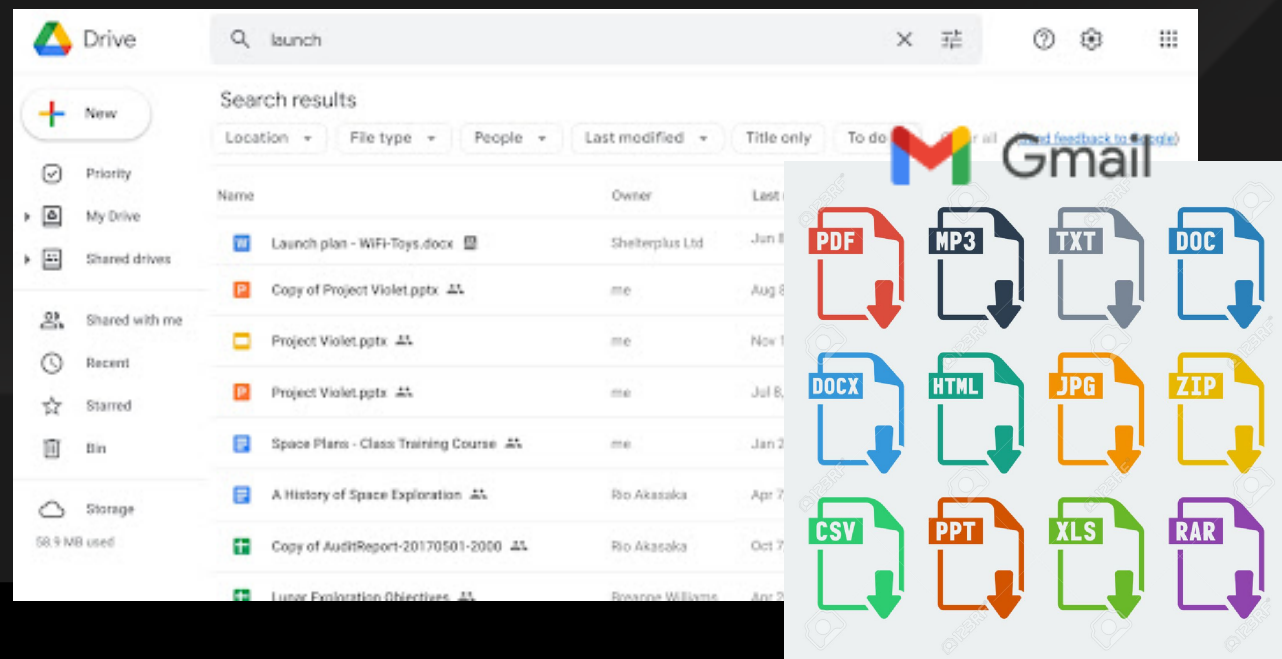# GENIS.AI

Let AI do the office works for your by turning your data into documents or website.
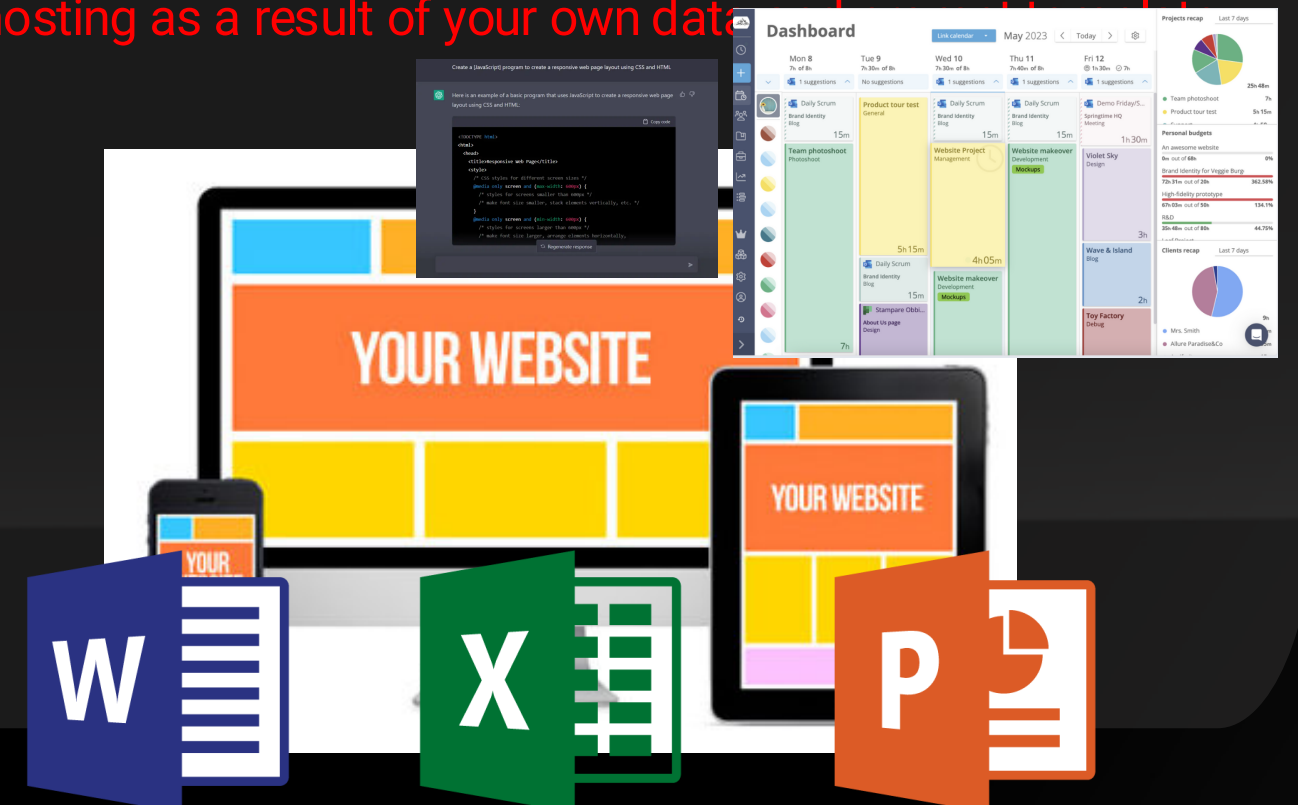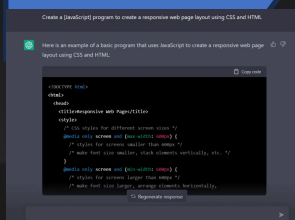
Problems

- LLMs do not have any knowledge of recent or up to date nor private data and hence are unable to generate content relevance to specific needs.
- Training or even finetuning for customization is highly difficult.
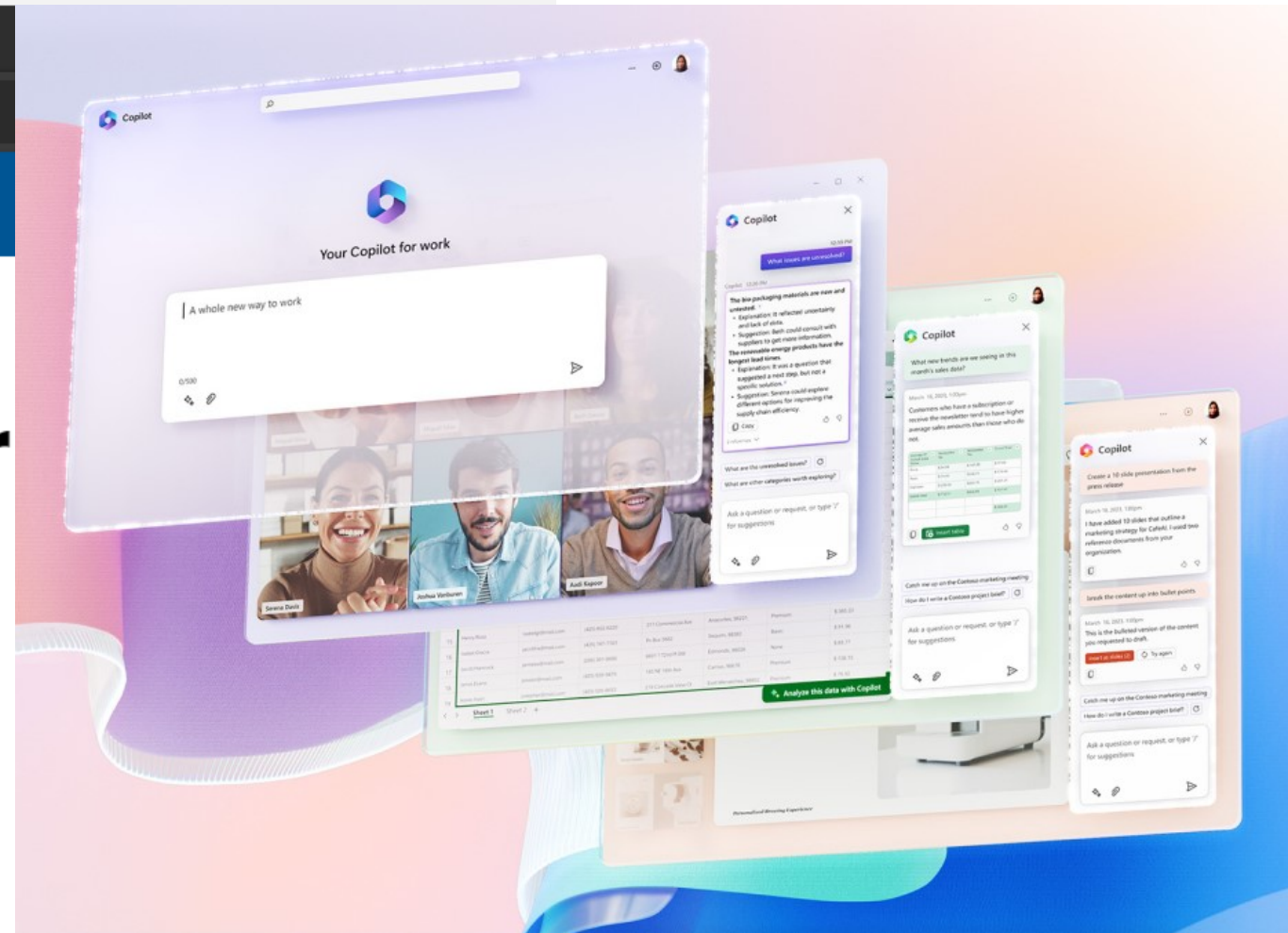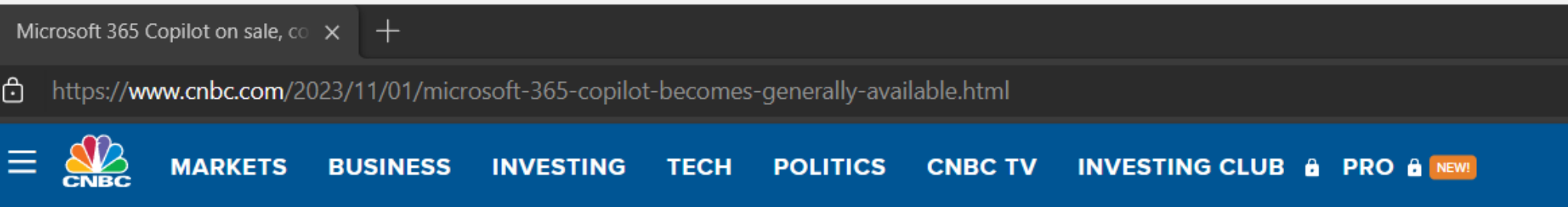- Current RAG solutions are either incomplete or costly and require manual integrations.

Solutions

- Freemium personal and sharable cloud storage and vector database that keep track of changes of the original files.
- Freemium office document, slide, or spreadsheet from your prompt and combined with other data.
- Freemium webapp hosting with content generation and hosting as a result of your own data

RAG+LLM

Microsoft 365 Copilot on sale, co

https://www.cnbc.com/2023/11/01/microsoft-365-copilot-becomes-generally-available.html

TECH

Microsoft starts selling AI tool for Office, which could generate $10 billion a year by 2026

PUBLISHED WED, NOV 1 2023•7:58 AM EDT | UPDATED WED, NOV 1 2023•1:44 PM EDT

Jordan Novet
@JORDANNOVET

SHARE f 𝕏 in ✉

KEY POINTS

- On Wednesday, Microsoft started selling the Microsoft 365 Copilot artificial intelligence add-on for its Office app subscriptions targeting businesses.

- Piper Sandler analysts estimate the tool could generate more than $10 billion in annualized revenue by 2026.

- That doesn't include uptake of additional tools such as Microsoft's Purview for managing data.

In this article

MSFT -0.15 (-0.04%) 🌙 ⊞

Open Source
Coding Generation and
Execution Environment

GitHub - KillianLucas/open-inter...

https://github.com/KillianLucas/open-interpreter/

Product
Solutions
Open Source
Pricing

Search or jump to...

Sign in   Sign up

KillianLucas / open-interpreter   Public

Code   Issues 119   Pull requests 25   Actions   Wiki   Security   Insights

Notifications   Fork 2.8k   Star 31.5k

main   7 branches   18 tags

Go to file   Code

KillianLucas Gotta go --fast   2317f27 16 hours ago   838 commits

.github   Check for potential duplicates for Github issues   3 days ago
docs   Update MACOS.md: fix command   2 days ago
interpreter   Gotta go --fast   16 hours ago
tests   Reversed local procedures (embedding model is too big, RAG is out of ...   17 hours ago
.git-blame-ignore-revs   chore: ignore second reformatting commit   4 days ago
.gitignore   Removed unecessary dependencies, moved yaspin to [safe], removed...   yesterday
.pre-commit-config.yaml   chore: add black, isort, and pre-commit   4 days ago
CONTRIBUTING.md   Merge pull request #699 from InterwebAlchemy/feature/automated-for...   3 days ago
LICENSE   Generate.   2 months ago
README.md   Addressed feedback on the PR   yesterday
README_IN.md   Translated Open Interpreter README into Hindi   2 months ago
README_JA.md   docs: move CONTRIBUTING to common path   2 months ago

About

OpenAI's Code Interpreter in your terminal, running locally

openinterpreter.com/

python   interpreter   gpt-4   chatgpt
llama2   codellama

Readme
MIT license
Security policy
Activity
31.5k stars
250 watching
2.8k forks

Report repository

Releases
18 tags

Blocks display code and outputs in different languages.

__init

Define

elf.lang

elf.outp

elf.code

self.acti

self.live

Terminal

~ $ interpreter -y

> Make a virtual environment here

Sure, I can help with

Create Local
Environment

Terminal

In Boulder\n\n- Aug 17, 2023\n - At Ty's house\n\n- Aug 18, 2023\n - Meet with Jon from 8:00 AM to 9:00 AM\n - Tentative: Dinner with Alex from 7:00 PM to 8:00 PM\n\n- Aug 19, 2023\n - See the boys", visible:true}
tell newMessage
make new to recipient at end of to recipients with properties
{address:"killian@water.ai"}
end tell
send newMessage
end tell

Email Sending

The email with your calendar events for the week has been sent to killian@water.ai. Please check your inbox. Let me know if there's anything else I can assist you with.

>

Terminal

~ $ interpreter -y
> Can you make me a simple Pomodoro app?

Sure, I can help you with that. Here is the plan:

1 I will create an HTML file with e
2 The CSS will be used to style the
  be used to control the timer func
3 The app will have input fields fo
  times, and buttons to start and s
4 The timer will be displayed and u

Let's start by creating the HTML

Pomodoro Timer

Work Time (minutes)
25

Break Time (minutes)
5

Start

Stop

25:00

Website Creation

Market Growth

Global Cloud Storage Volume in Numbers

According to recent research by Nasuni,
there is over 1 Exabyte of data stored in the cloud,
or 1024 Petabytes of data

https://www.globaldots.com/resources/blog/how-much-is-stored-in-the-cloud/

Technology & Telecommunications › IT Services

Size of the cloud storage market worldwide from 2022 to 2030
(in billion U.S. dollars)

Market size in billion U.S. dollars

Year	Value
2022	90.17
2023*	108.69
2024*	134.12
2025*	165.51
2026*	204.24
2027*	252.03
2028*	311
2029*	383.78
2030*	472.47
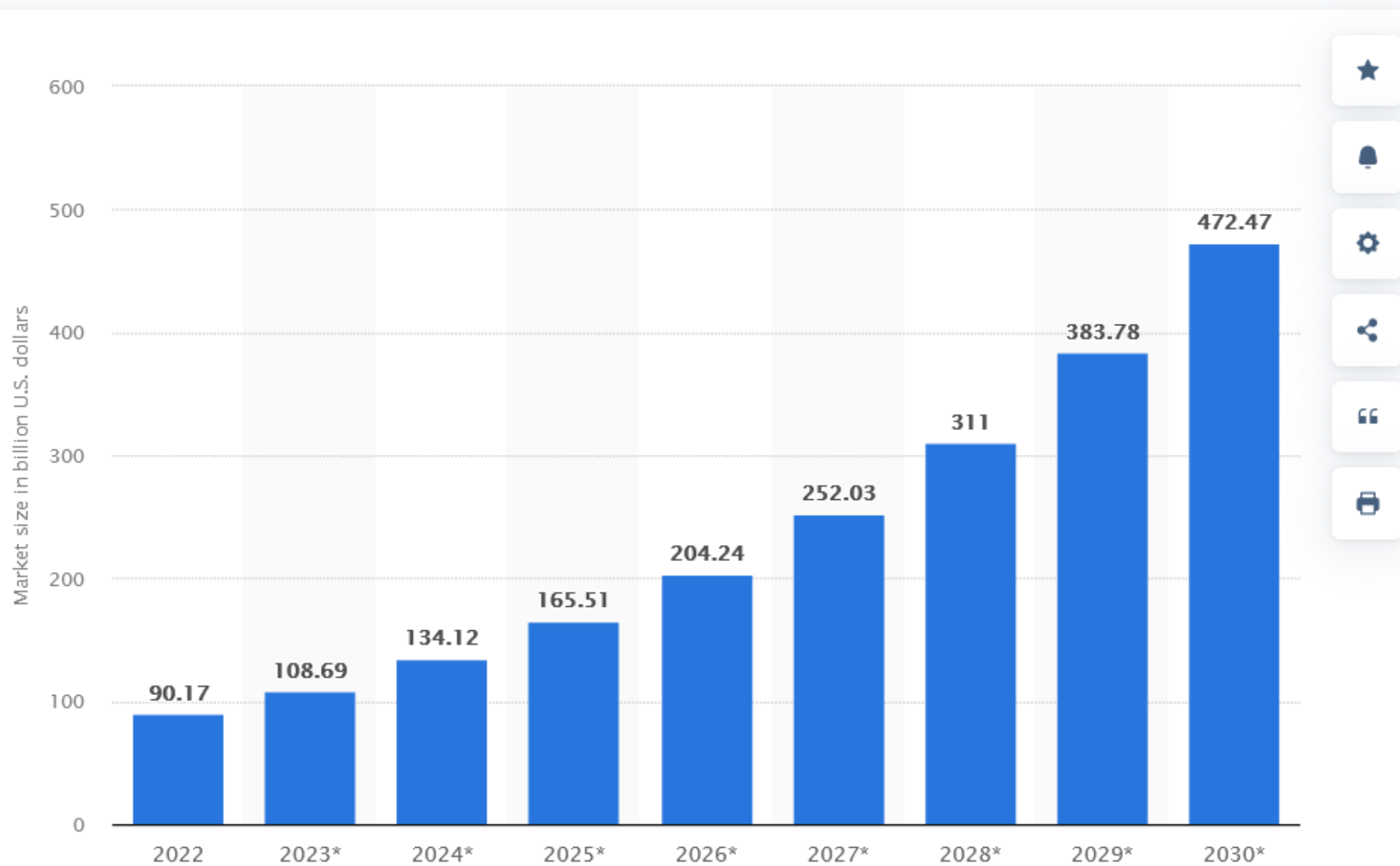
© Statista 2023

ⓘ Additional Information

Show source ⓘ

gen ai market value

News Images Graph Shopping Videos Books Maps Flights Finance

About 43,600,000 results (0.54 seconds)

Value in the Generative AI market is projected to reach US$45bn in 2023. Value is expected to show an annual growth rate (CAGR 2023-2030) of 24.40%, resulting in a market volume of US$207bn by 2030. In global comparison, most value will be generated in the United States (US$16bn in 2023).

Statista
https://www.statista.com › ... › Artificial Intelligence

Generative AI - Worldwide | Statista Market Forecast

LLM Models

ChatGPT: World renowned.
Claude: Rising star with Amazon backed
Bard: Google Cutting-Edge research

Llama: Prominent open source alternative
Falcon: Commercial license available
Mistral: High Performance
Zephyr: Outperforms Chat Llama

LLM Tools

ChatGPT Function Call and Open Interpreter

LlamaIndex: Optimized RAG with indexing

Lang Chain: Chain of prompt

AutoGPT: From a prompt to autonomous

ragas: RAG benchmarking tool

VectorDB

Qdrant: Top Performance
Wewiate: Rich Integration
Milvus: Widely used by majority tech
ChromaDB: Robust
Pinecone: Original
vectorDB: Efficient and Promising by jina.ai

RAG Tools (Competitors)

ChatGPT Plugin: Global Market Leader
Office365Copilot: Tightly integrated.
Bard: Google Eco-system Integrated
saydata.tech: BI from prompt on web
Alisa (Line OA): Local Market Leader
Locusive.com: Closely related with ours
Typingmind.com: Market leader
Quivr.app: require self-implementation
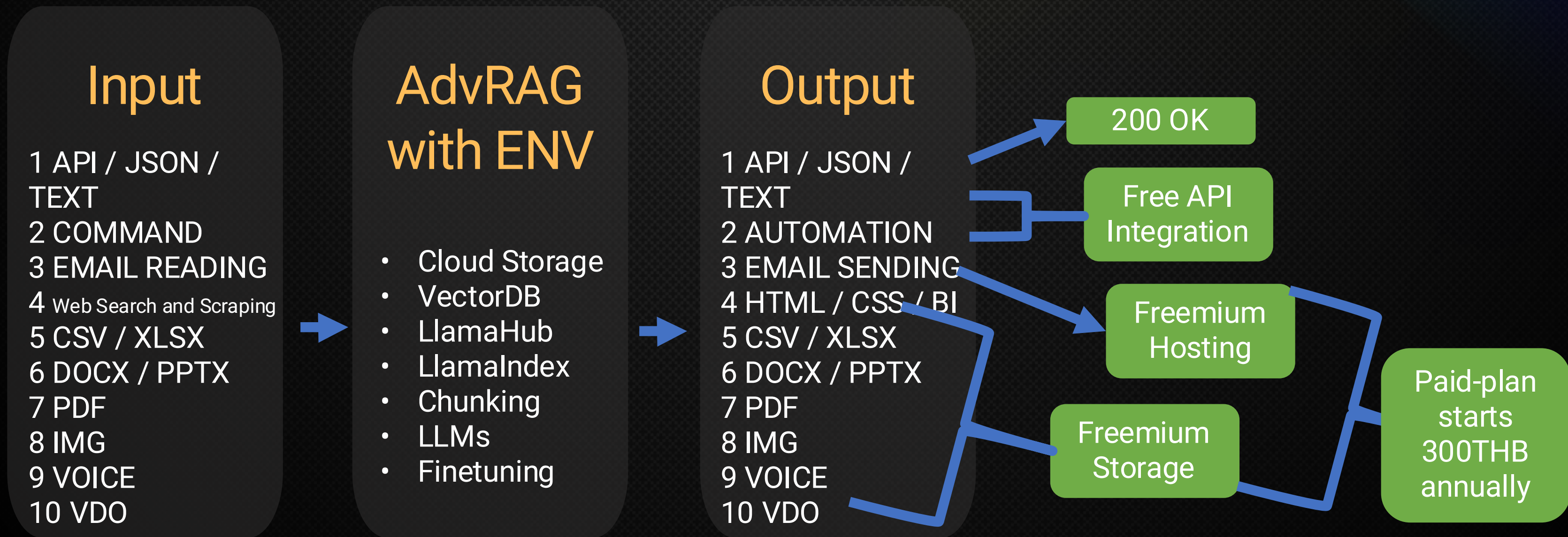ChatDoc, klu.ai, etc.

GENIS.AI

- Cutting-Edge innovation for SEO website/app creation from LLM and RAG with hosting and private cloud storage/vectorDB for long-term memory included. Codeserver for customize.

- Allow automation to be executed in private VM/Container environment as arms and legs of AI. 3rd party API for extension to physical.

- Total solution in cloud storage for both raw and generated content with private/public view with social interactions and LLM ready.

- Customizable chatbot for public access with RAG function. Guardrail against hallucination as RAG is based on factual information.

- Option for one-click fine-tuning with your

Business model and the ability to promote business
Multi-Modal Combination
(for complex/creative input and output)

Input

1 API / JSON / TEXT
2 COMMAND
3 EMAIL READING
4 Web Search and Scraping
5 CSV / XLSX
6 DOCX / PPTX
7 PDF
8 IMG
9 VOICE
10 VDO

AdvRAG with ENV

- Cloud Storage
- VectorDB
- LlamaHub
- LlamaIndex
- Chunking
- LLMs
- Finetuning

Output

1 API / JSON / TEXT
2 AUTOMATION
3 EMAIL SENDING
4 HTML / CSS / BI
5 CSV / XLSX
6 DOCX / PPTX
7 PDF
8 IMG
9 VOICE
10 VDO

200 OK

Free API Integration

Freemium Hosting

Freemium Storage

Paid-plan starts 300THB annually

(Draft) Vision and Management Strategy
Feasible Technical Stacks

1 User Management
OAuth+
User Database
Cloud Storage

2 Front-End Chat UI

HTMX
TAILWIND
LINE API
FASTAPI Server

4 On-Prem Storage

MinIO
TrueNas
(Caching)

7 Container/VM Orchestration

Kubernetes /
Proxmox

Hosting and Storage

Query/Prompt

Data Management API

5 RAG/VectorDB

LlamaHub
LlamaIndex

Retrieved Context

3 Back-End

Python API

Similarity Search

6 Output

LLM API
OpenAI+Others
Hugging Face

Query/Prompt

Specificity of Clients' Pain Point

| Use-Cases/Solutions	Pain Points / Problems	Impact	User Personas
Free AI virtual assistant/organizer and hosting of personal raw/generated gallery with social interaction (Like, Share, Follow, Buy)	• RAG need to be integrated and uploaded data manually to LLM • Available solutions are inconvenience and not utilized.	• Broaden AI adoption with free virtual assistant for digital life. • Personal cloud storage with LLM+RAG organizer and LINE API.	Human and Business
Create a BI and office files with data analytic using NLP prompt	• Tedious tasks for human. • Available solutions are inconvenience and not utilized.	• AI application enable office workers to do their tedious routine tasks.	Employee
Data Analytic and BI from private data with web accessible and authentication.	• RAG need to be integrated and uploaded data manually to LLM • Available solutions are inconvenience and not utilized.	• Potential high impact for data-driven organization. • Reduce turn around time for BI production.	Business Enterprise
Launch customized static/web-app from LLM+RAG generated content with private data and web-based VSCode (code-server) for customization and automation.	• Manually created website/application and hosting requires DevOps/SWE and setup. • No available solution/application with LLM and RAG integration.	• Instantly deploy a sales page from prompt with customized product info / content creation. • Remove friction in development and marketplace reliance. • Analytic/ Promotion/ Interaction	Business Enterprise Business Entrepreneur
Customizable, personalized and specialized chatbot for customer interaction with public access or embed in the sales page for commerce and service support.	• Require time, personnel, and infrastructure for the maintenance and utilization of a chat-enabled application. • Customized solutions have not available as a publicly accessible web application.	• Develop a chatbot using RAG technology and deploy it as a tailored web application in a secure, private environment. • Enhance AI capabilities to coordinate and streamline more intricate tasks and beyond.	Business Enterprise Business Entrepreneur

The Team
Ability of the Management System to Execute

- Kind personality, communicate clearly

- Skilled AI & Data Scientist/Engineer and full-stack developers

- OSS💜Python and Tailwind/Flow-bite and DevOps/IT Infra

- Lead by heart, learning by doing, always a glass half full

- Passion in creating positive impact for societies with AI

Skunpoj Thanarojsophon, PhD
Full-Stack Dev /
Data Engineer

- Engineer Background

- Master in Big Data

- Winner of AI competition

- Winner of Gen-AI adoption plan competition for BOT

- Experienced Data Engineer and OSS contributor

Pitikorn Khlaisamniang
AI Engineer
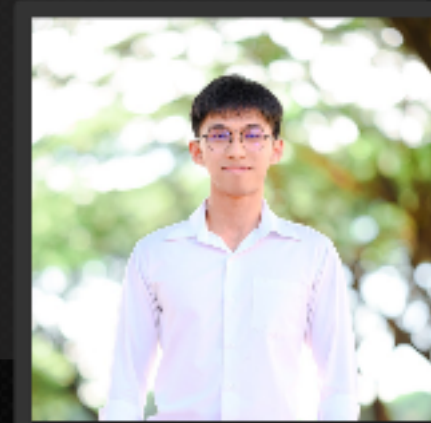
Next Steps

In Code-Server side

- Add localcopilot to code-server (https://github.com/danielgross/localpilot)

- Add RepoCoder

  framework to tackle the repository-level code completion task, which is to continue writing the unfinished code based on a broader context of the repository.
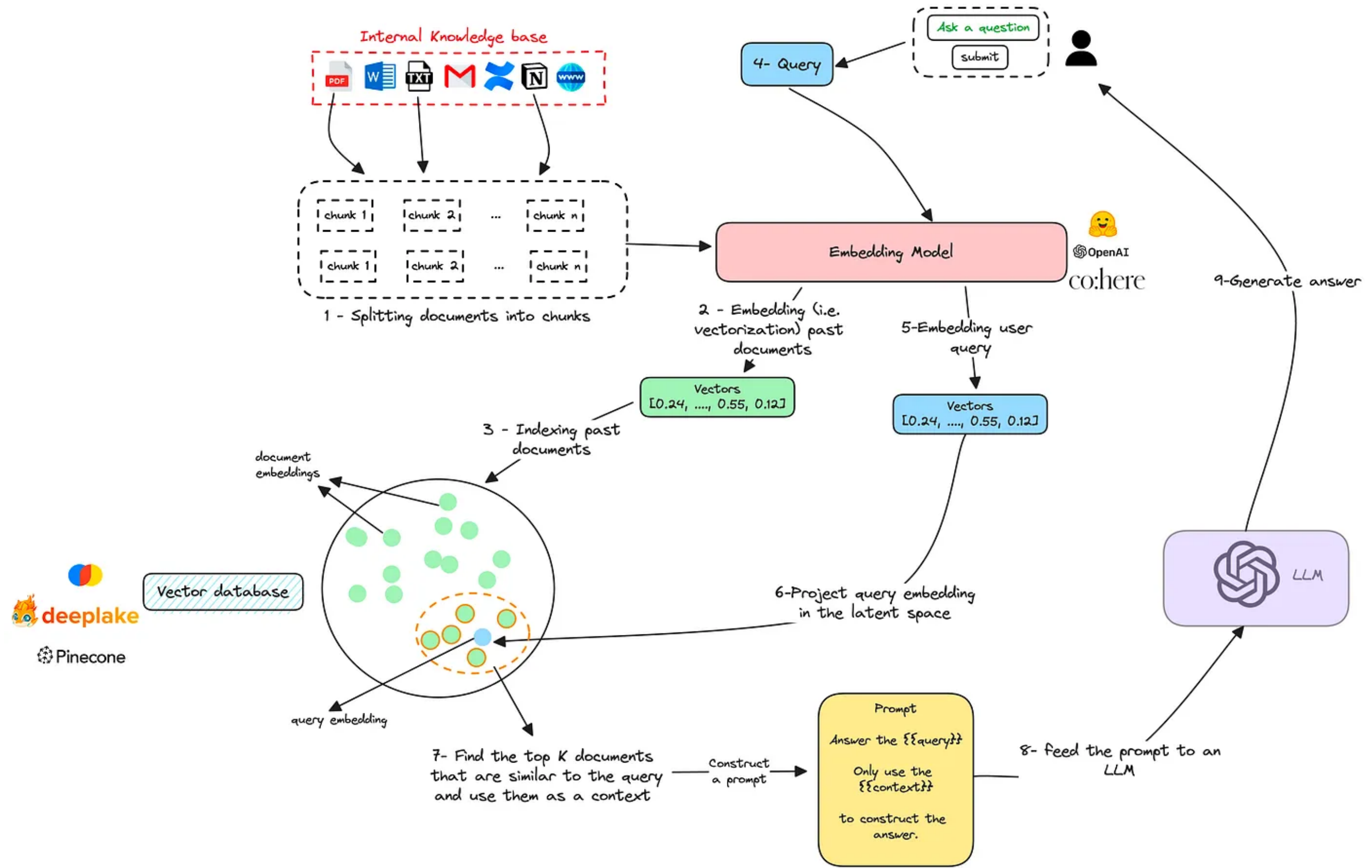
In Chat side

- sss

RAG References

Internal Knowledge base

Ask a question
submit

4- Query

Embedding Model
OpenAI
co:here

1 - Splitting documents into chunks

chunk 1   chunk 2   ...   chunk n
chunk 1   chunk 2   ...   chunk n

2 - Embedding (i.e. vectorization) past documents

5-Embedding user query

9-Generate answer

Vectors
[0.24, ..., 0.55, 0.12]

Vectors
[0.24, ..., 0.55, 0.12]

3 - Indexing past documents

document embeddings

deeplake
Pinecone

Vector database

6-Project query embedding in the latent space

LLM

query embedding

7- Find the top K documents that are similar to the query and use them as a context

Construct a prompt

Prompt

Answer the {{query}}

Only use the {{context}}

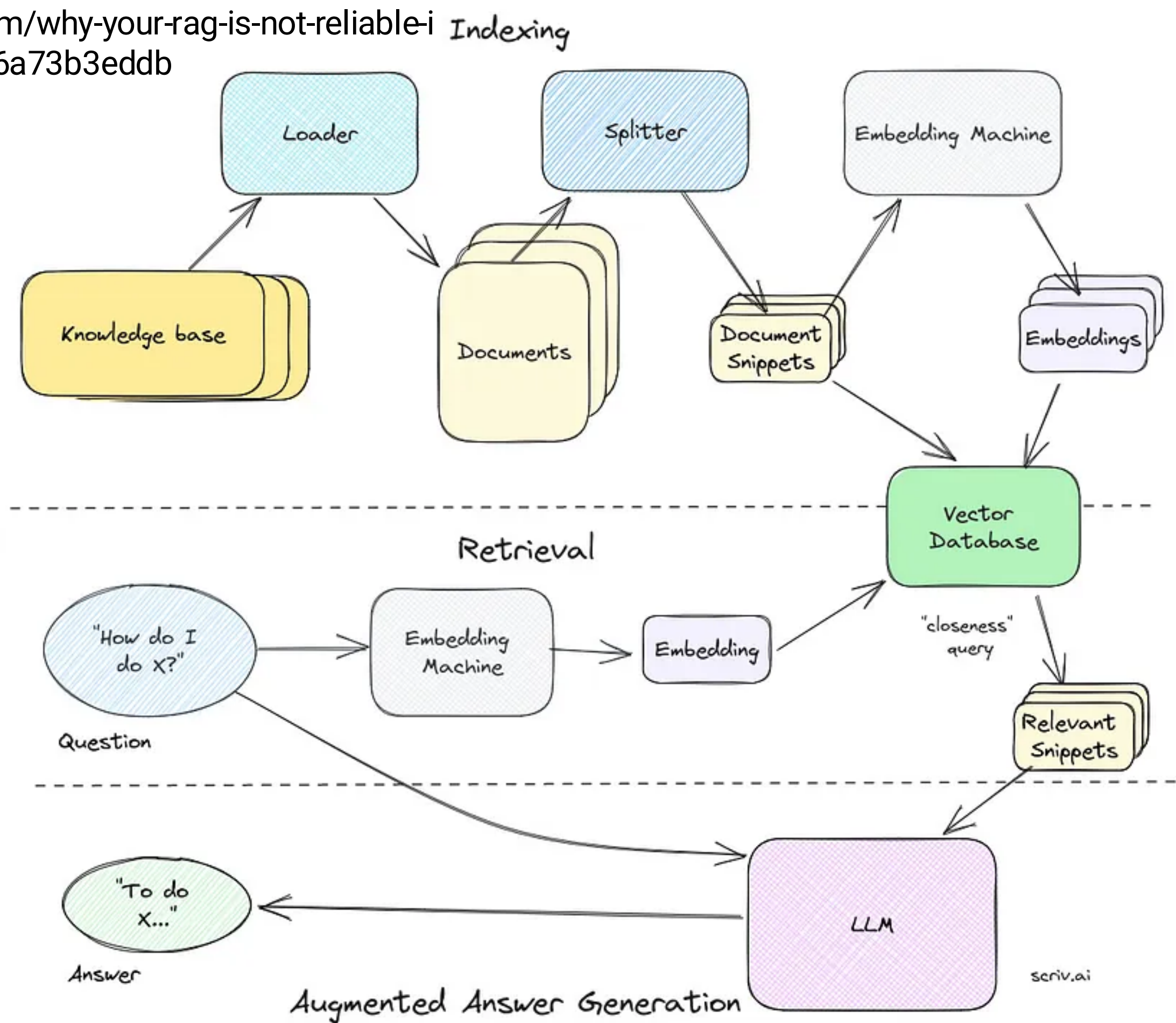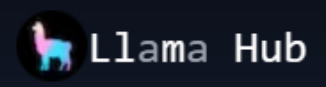to construct the answer.

8- feed the prompt to an LLM

- query is replaced with user input
- context is replaced with the similar documents

https://levelup.gitconnected.com/you-can-now-build-a-chatbo
t-to-talk-to-your-internal-knowledge-base-b6066cacf2d5

https://towardsdatascience.com/why-your-rag-is-not-reliable-in-a-production-environment-9e6a73b3eddb

Indexing

Loader

Splitter

Embedding Machine

Knowledge base

Documents

Document Snippets

Embeddings

Vector Database

Retrieval

"How do I do X?"

Question

Embedding Machine

Embedding

"closeness" query

Relevant Snippets

"To do X..."

Answer

LLM

Augmented Answer Generation

scriv.ai

Llama Hub

Add a tool or loader

Search for all...

Popular   Recently added

iamarunbrahma · 8 months ago · ⭐ 0   mommi84 · 8 months ago · ⭐ 0   ajhofmann · 8 months ago · ⭐ 0

file/simple_csv   Loader
vguillet · 8 months ago · ⭐ 0

file/unstructured   Loader
thejessezhang · 8 months ago · ⭐ 0

firebase_realtimedb   Loader
ajay · 8 months ago · ⭐ 0

firestore   Loader
rayzhudev · 8 months ago · ⭐ 0

github_repo   Loader
ahmetkca · 8 months ago · ⭐ 0

github_repo_collaborators   Loader
rwood-97 · 8 months ago · ⭐ 0

github_repo_issues   Loader
moncho · 8 months ago · ⭐ 0

gmail   Loader
bbornsztein · 8 months ago · ⭐ 0

gmail   Tool
ajhofmann · 8 months ago · ⭐ 0

google_calendar   Loader
ong · 8 months ago · ⭐ 0

google_calendar   Tool
ajhofmann · 8 months ago · ⭐ 0

google_docs   Loader
jerryjliu · 8 months ago · ⭐ 0

google_drive   Loader
ravi03071991 · 8 months ago · ⭐ 0

google_keep   Loader
pycui · 8 months ago · ⭐ 0

google_search   Tool
ajhofmann · 8 months ago · ⭐ 0