

BUDAPEST **DATA** + **ML** FORUM

Kódolás + AI = Profit?

Arató Bence
Ügyvezető, BI Consulting





A Generatív AI robbanása

● AI
Search term

+ Compare

Worldwide ▼

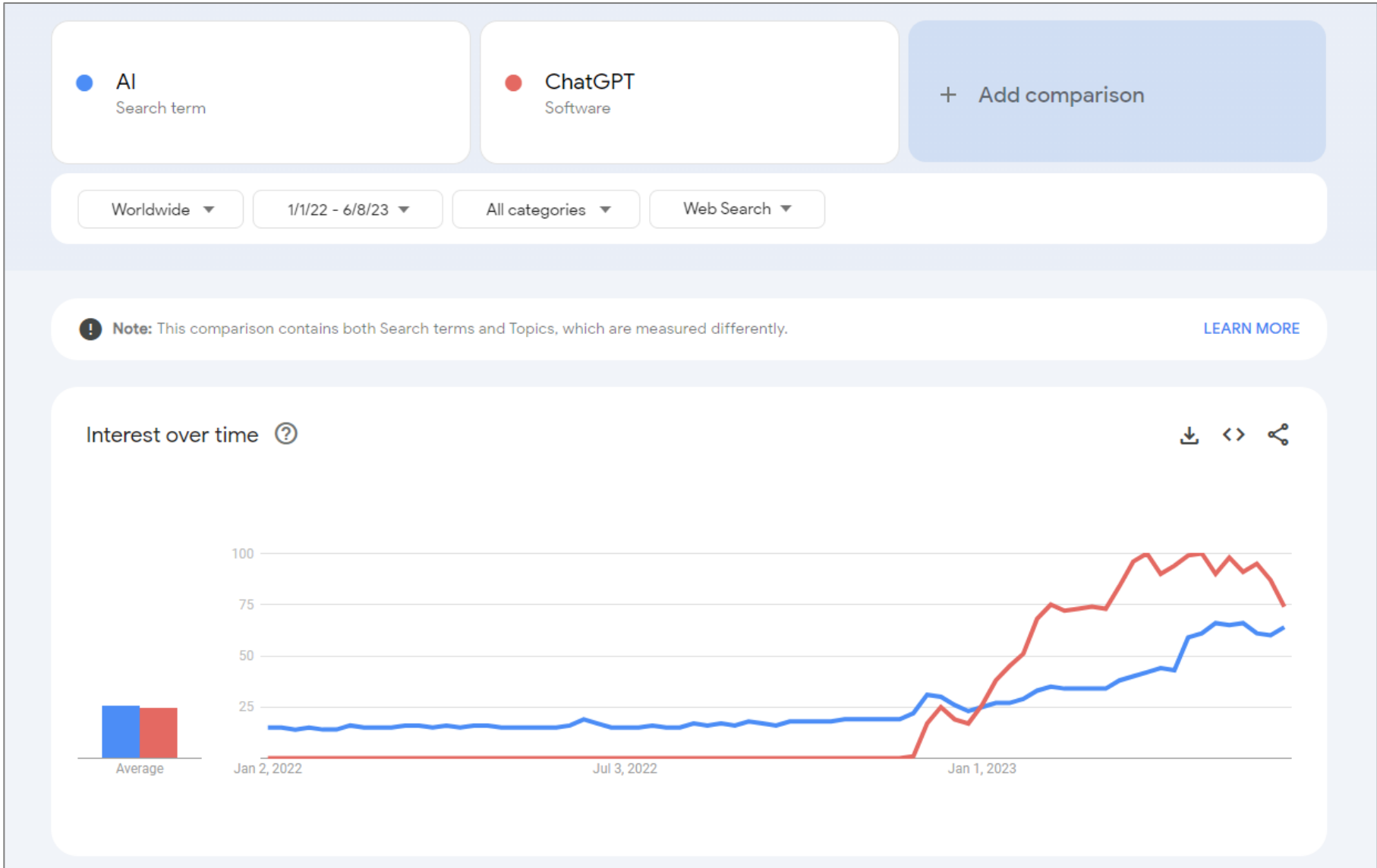
1/1/22 - 6/8/23 ▼

All categories ▼

Web Search ▼


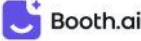
































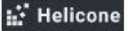

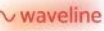















Interest over time ⓘ





Y Combinator W23

Y Combinator W23 Generative AI Landscape

	Example Use Case	YC W23 Startups
Business Function	Accounting Finance	ALPHAWATCH  Hadrius
	Marketing	 Booth.ai  Speedy
	Sales	 lightski  Tennr  Coldreach  Perspectiva  FABIUS  VECTOR
	Customer Success	 OpenSight  OfOne  Buff  inkeep  Parabolic  Yuma.ai  Deep BlueDot
	Knowledge Management	 Hazel  Credal.ai  type  Cardinal  Layup
	Media	 Iliad  decoherence  Texel  DREAM3D
Engineering Function	Data Analytics	 Defog.ai  Lume  turntable  Outerbase
	ML Ops Platform	 vellum  GRADIENT  Baseplate  Flower
	Infrastructure	 PoplarML  Helicone  stack  waveline  Chart  CHIMA  metal  BerriAI  ANARCHY  b.pyg
	Developer Tools	 extend  Wild Moose  Meru  CodeComplete  Lasso  FOUNDATION  Second  Rubbrband  vocode

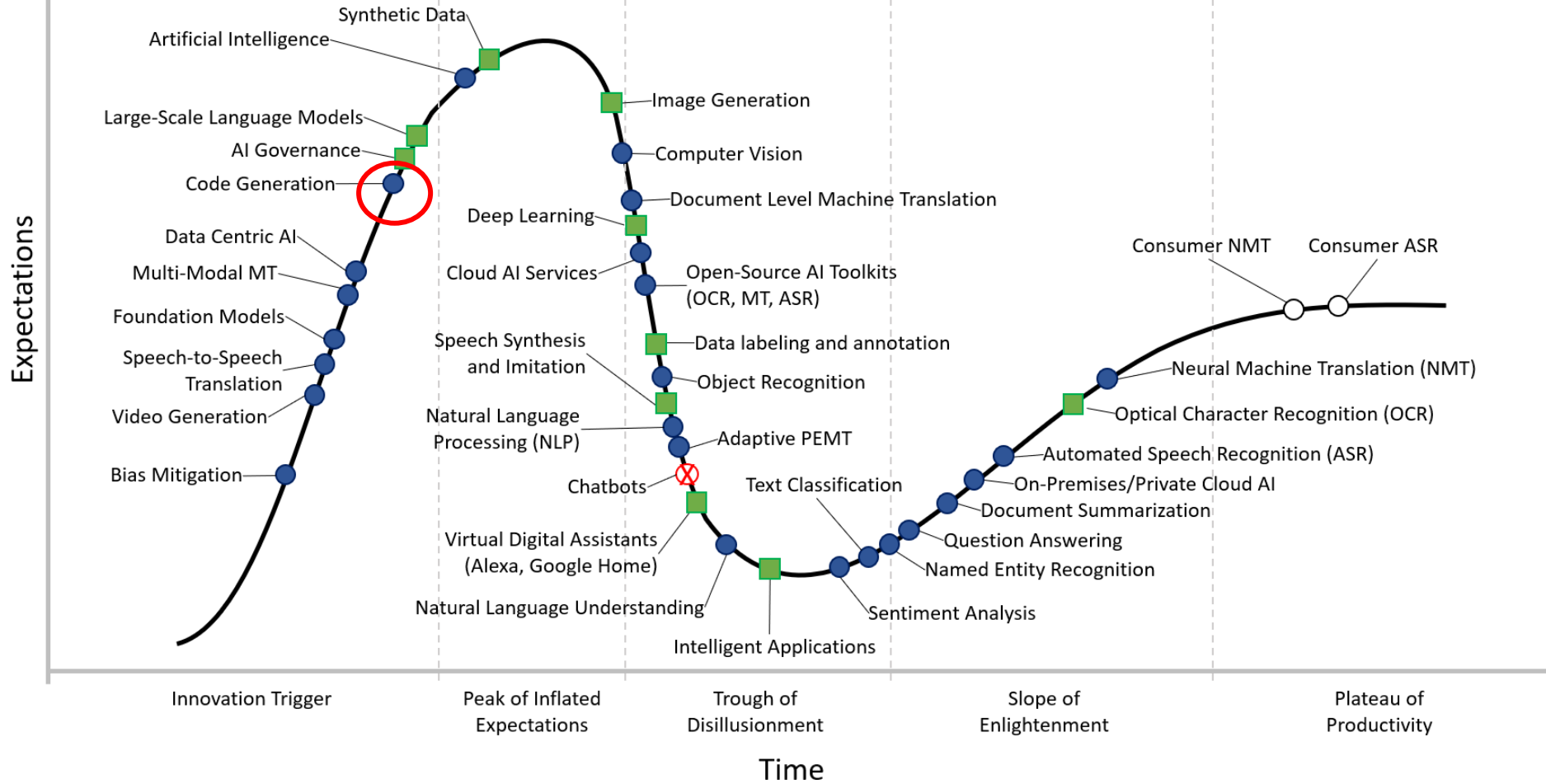
🔥 Created by your friends at Truewind (YC W23)

This is a work in progress. Reach out to us if you want to be added to the next iteration

Programozás

Omniscien Technologies Hype Cycle for AI Technologies in Business

January 2023



Plateau will be reached:

- Less than 2 years
- 2 to 5 years
- 5 to 10 years
- ▲ More than 10 years
- ⊗ Obsolete before plateau

Copyright © 2023 Omniscien Technologies

Kódasszisztensek

AR

Báttya, hogy is kell pandas Excel fájlt menteni?



Ahogy azt mondtad, hogy "báttya", gondolom, hogy magyarul beszélünk. Itt van egy példa arra, hogyan lehet a Pandas-t használni egy DataFrame mentésére Excel fájlba:



Először is, telepítened kell a Pandas csomagot, ha még nem tetted meg. Használhatod a következő parancsot a telepítéshez:

```
python
```

Copy code

```
pip install pandas
```

Miután telepítetted a Pandas-t, importáld be a következő módon:

```
python
```

Copy code

```
import pandas as pd
```

Ezután, feltéve hogy már van egy DataFrame objektumod (például a `df`), használhatod a `to_excel()` metódust a DataFrame mentésére Excel fájlba. Itt van egy példa arra, hogyan csináld:

```
python
```

Copy code

```
df.to_excel('kimeneti_fajl.xlsx', index=False)
```

Your AI pair programmer

GitHub Copilot uses the OpenAI Codex to suggest code and entire functions in real-time, right from your editor.

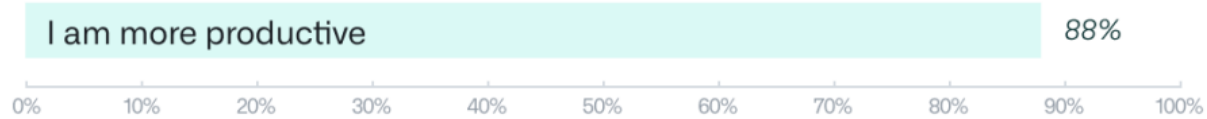
[Start my free trial >](#)

[Compare plans](#)

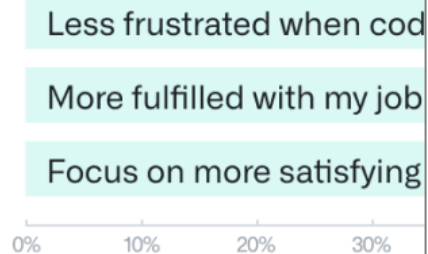
Github Copilot

When using GitHub Copilot...

Perceived Productivity



Satisfaction and Well-being



Efficiency and Flow*





Amazon CodeWhisperer

Build applications faster and more securely with your AI coding companion

Use CodeWhisperer for free

Start today with the Individual Tier:

- Free for individual use
- Unlimited code suggestions
- Reference tracking
- 50 security scans (per user, per month)

[Learn more »](#)

Get more done faster

CodeWhisperer is trained on billions of lines of code and can generate code suggestions ranging from snippets to full functions in real time based on your comments and existing code. Bypass time-consuming coding tasks and accelerate building with unfamiliar APIs.



main.js

```
1 // Parse a CSV string of songs and return the list (position, originalChartDate, artist, title).
2 // Ignore lines starting with #.
3 function parseCSV(csvString) {
4   const songs = [];
5   const lines = csvString.split('\n');
6   for (let i = 0; i < lines.length; i++) {
7     const line = lines[i];
8     if (line.startsWith('#')) continue;
```

Amazon CodeGuru

Automate code reviews and optimize application performance with ML-powered recommendations

Get started with Amazon CodeGuru

Detect security vulnerabilities and automate code reviews

Amazon CodeGuru is a developer tool that provides intelligent recommendations to improve code quality and identify an application's most expensive lines of code. Integrate CodeGuru into your existing software development workflow to automate code reviews during application development and continuously monitor application's performance in production and provide recommendations and visual clues on how to improve code quality, application performance, and reduce overall cost.

CodeGuru Reviewer uses machine learning and automated reasoning to identify critical issues, security vulnerabilities, and hard-to-find bugs during application development and provides recommendations to improve code quality. [Learn more »](#)

StarCoder: A State-of-the-Art LLM for Code

Published May 4, 2023

[Update on GitHub](#)



[lvwerra](#)

[Leandro von Werra](#)



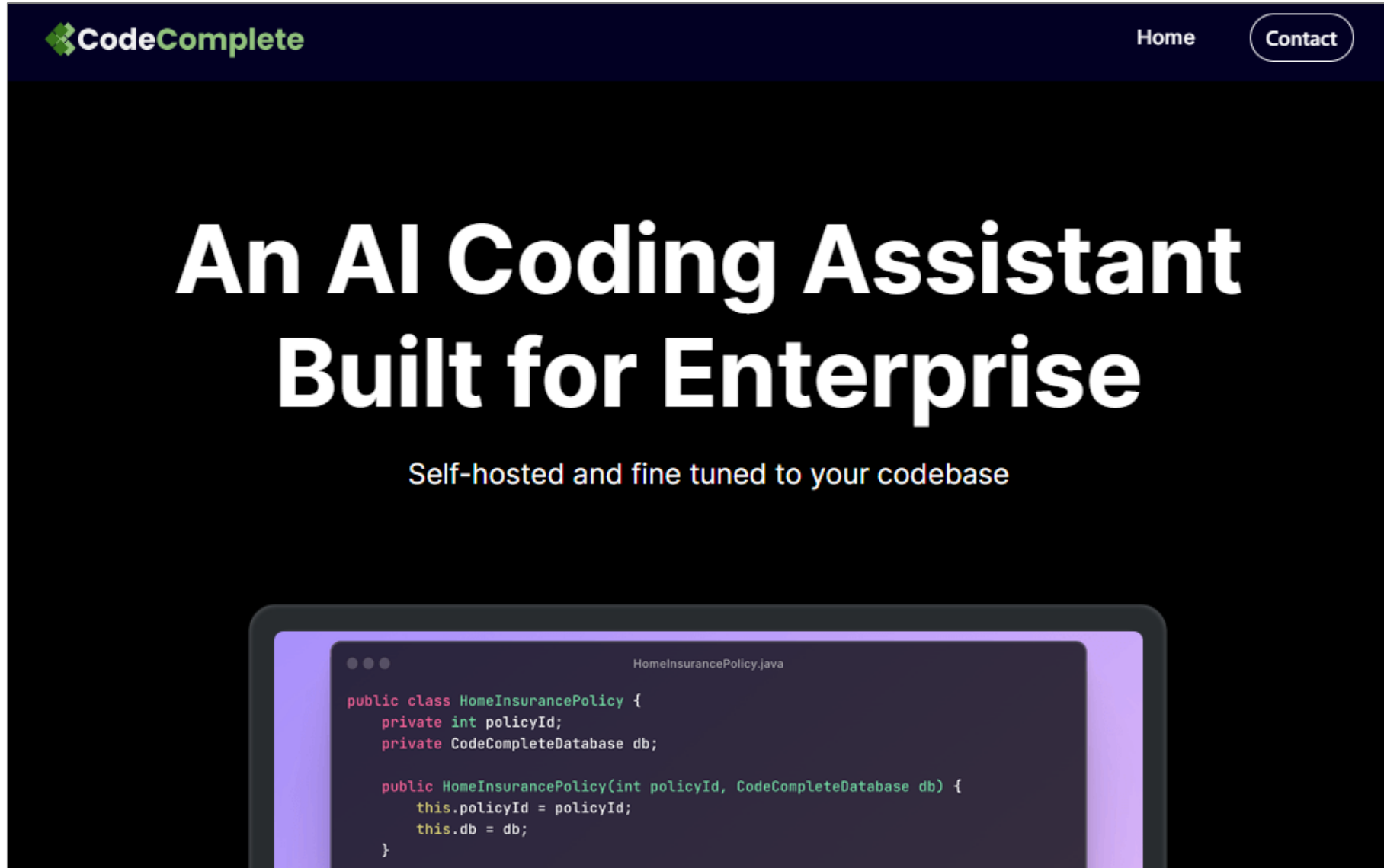
[loubnabnl](#)

[Loubna Ben Allal](#)

Introducing StarCoder

StarCoder and StarCoderBase are Large Language Models for Code (Code LLMs) trained on permissively licensed data from GitHub, including from 80+ programming languages, Git commits, GitHub issues, and Jupyter notebooks. Similar to LLaMA, we trained a ~15B parameter model for 1 trillion tokens. We fine-tuned StarCoderBase model for 35B Python tokens, resulting in a new model that we call StarCoder.

We found that StarCoderBase outperforms existing open Code LLMs on popular programming benchmarks and matches or surpasses closed models such as `code-cushman-001` from OpenAI (the original Codex model that powered early versions of GitHub Copilot). With a context length of over 8,000 tokens, the StarCoder models can process more input than any other open LLM, enabling a wide range of interesting applications. For



CodeComplete

Home Contact

An AI Coding Assistant Built for Enterprise

Self-hosted and fine tuned to your codebase

```
HomeInsurancePolicy.java

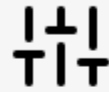
public class HomeInsurancePolicy {
    private int policyId;
    private CodeCompleteDatabase db;

    public HomeInsurancePolicy(int policyId, CodeCompleteDatabase db) {
        this.policyId = policyId;
        this.db = db;
    }
}
```



Self Hosted

In your cloud or on-premises
to protect your security and
IP



Fine Tuned

To your codebase and
coding style to better
integrate into your developer
workflow



Permissively Licensed

Trained only on
permissively-licensed repos
to avoid legal risks

Currently in private beta

Dokumentálás

Mintlify Writer

rating 4.9/5 (79)

chat 123 online

Stars

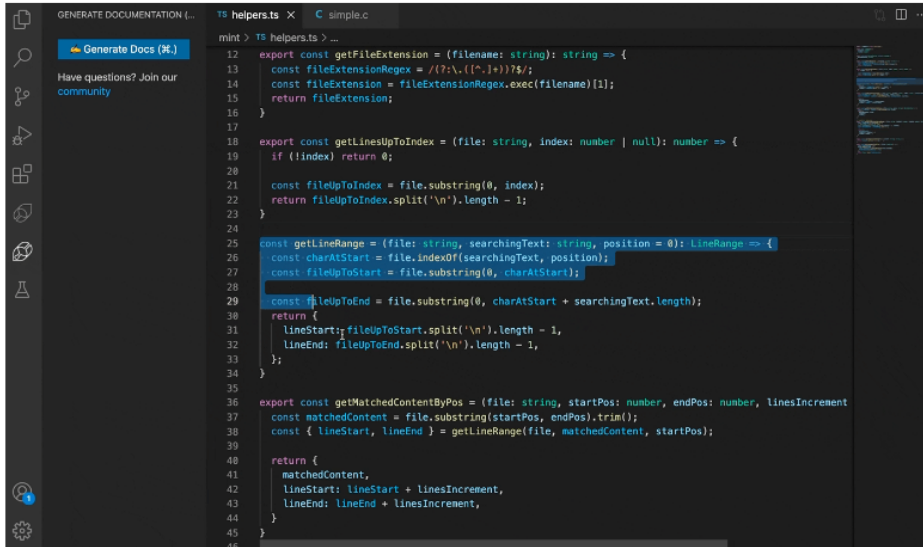
2.1k

Follow @mintlify

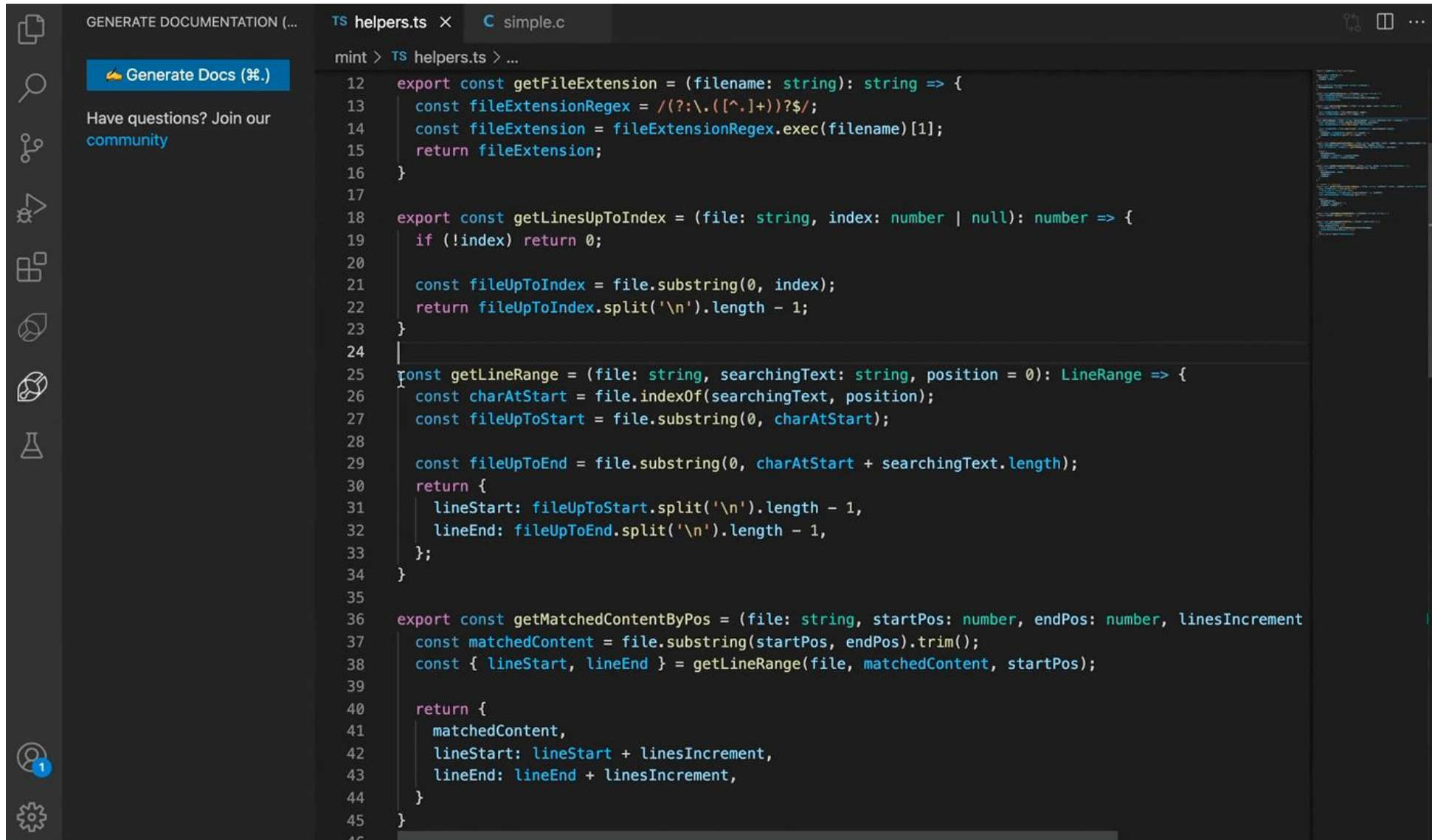
Writing documentation sucks. Let Mintlify take care of it. Just highlight code and see the magic.

How to generate documentation

Step 1 Highlight code or place cursor on the line you want to document **Step 2** Click on the `Write Docs` button (or hit `⌘ + .`)



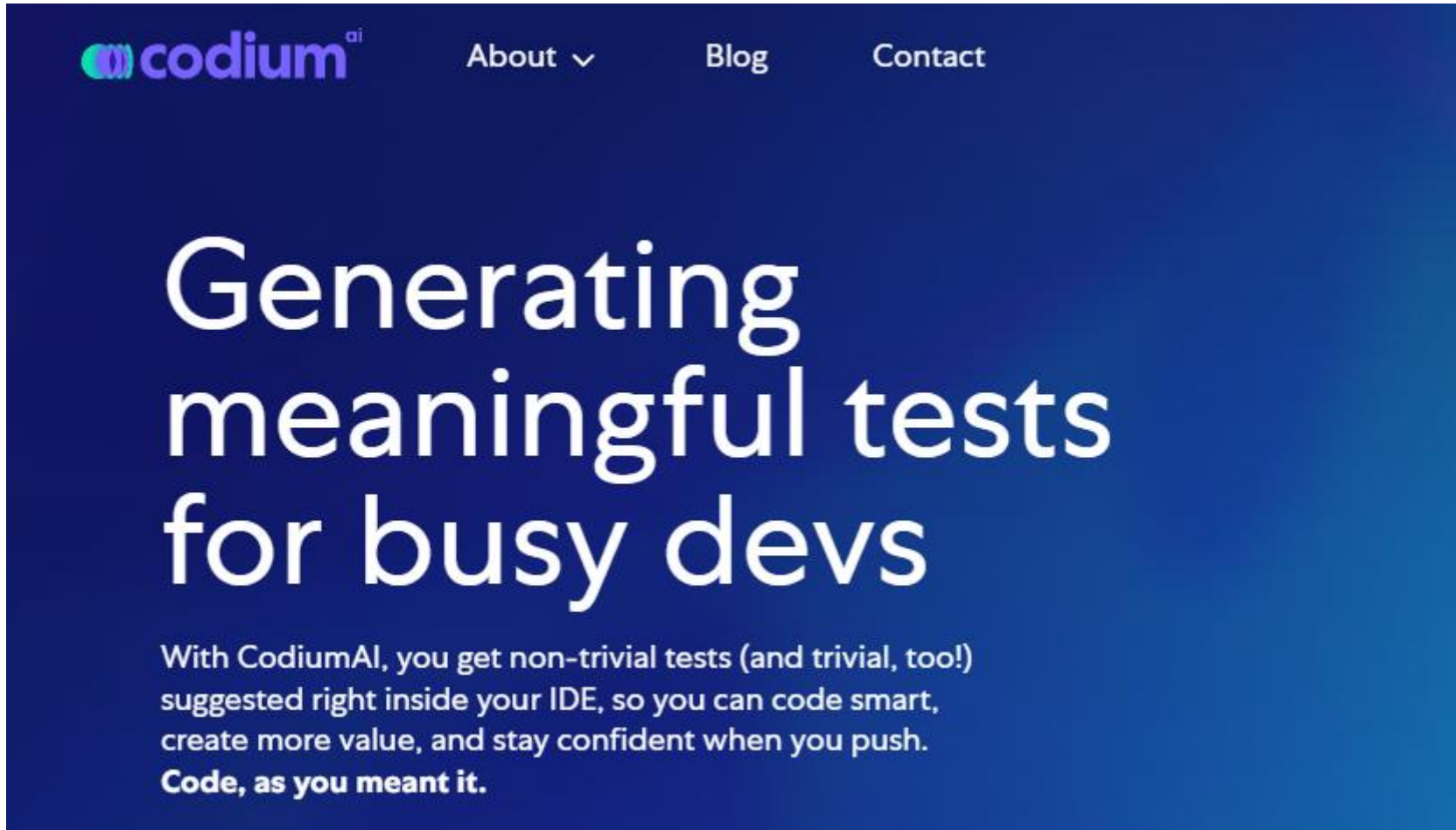
```
mint > TS helpers.ts > ...
12 export const getFileExtension = (filename: string): string => {
13   const fileExtensionRegex = /^(?:(?!.*\.)?)/?$/;
14   const fileExtension = fileExtensionRegex.exec(filename)[1];
15   return fileExtension;
16 }
17
18 export const getLinesUpToIndex = (file: string, index: number | null): number => {
19   if (!index) return 0;
20
21   const fileUpToIndex = file.substring(0, index);
22   return fileUpToIndex.split('\n').length - 1;
23 }
24
25 const getLineRange = (file: string, searchingText: string, position = 0): LineRange => {
26   const charAtStart = file.indexOf(searchingText, position);
27   const fileUpToStart = file.substring(0, charAtStart);
28
29   const fileUpToEnd = file.substring(0, charAtStart + searchingText.length);
30   return {
31     lineStart: fileUpToStart.split('\n').length - 1,
32     lineEnd: fileUpToEnd.split('\n').length - 1,
33   };
34 }
35
36 export const getMatchedContentByPos = (file: string, startPos: number, endPos: number, linesIncrement
37   const matchedContent = file.substring(startPos, endPos).trim();
38   const { lineStart, lineEnd } = getLineRange(file, matchedContent, startPos);
39
40   return {
41     matchedContent,
42     lineStart: lineStart + linesIncrement,
43     lineEnd: lineEnd + linesIncrement,
44   }
45 }
46 }
```




The image shows a code editor interface with a dark theme. On the left, there is a sidebar with a 'Generate Docs (%)' button and a link to 'Have questions? Join our community'. The main editor area shows a file named 'helpers.ts' with the following TypeScript code:

```
mint > TS helpers.ts > ...
12 export const getFileExtension = (filename: string): string => {
13   const fileExtensionRegex = /^(?:\.([\^.]*)?)?$/;
14   const fileExtension = fileExtensionRegex.exec(filename)[1];
15   return fileExtension;
16 }
17
18 export const getLinesUpToIndex = (file: string, index: number | null): number => {
19   if (!index) return 0;
20
21   const fileUpToIndex = file.substring(0, index);
22   return fileUpToIndex.split('\n').length - 1;
23 }
24
25 const getLineRange = (file: string, searchingText: string, position = 0): LineRange => {
26   const charAtStart = file.indexOf(searchingText, position);
27   const fileUpToStart = file.substring(0, charAtStart);
28
29   const fileUpToEnd = file.substring(0, charAtStart + searchingText.length);
30   return {
31     lineStart: fileUpToStart.split('\n').length - 1,
32     lineEnd: fileUpToEnd.split('\n').length - 1,
33   };
34 }
35
36 export const getMatchedContentByPos = (file: string, startPos: number, endPos: number, linesIncrement
37   const matchedContent = file.substring(startPos, endPos).trim();
38   const { lineStart, lineEnd } = getLineRange(file, matchedContent, startPos);
39
40   return {
41     matchedContent,
42     lineStart: lineStart + linesIncrement,
43     lineEnd: lineEnd + linesIncrement,
44   }
45 }
46
```

Tesztelés

The image shows a screenshot of the Codium AI website. The header is dark blue with the Codium AI logo on the left and navigation links for 'About', 'Blog', and 'Contact' on the right. The main content area features a large white headline and a sub-headline on a dark blue background.

 **codium**^{ai} About ▾ Blog Contact

Generating meaningful tests for busy devs

With CodiumAI, you get non-trivial tests (and trivial, too!) suggested right inside your IDE, so you can code smart, create more value, and stay confident when you push.
Code, as you meant it.

Generate test

```
def gaussian_elimination(  
    coefficients: NDArray[float64], vector: NDArray[float64]) -> NDArray[float64]:  
  
    if (check_valid_input(coefficients, vector) == False):  
        return np.array([], dtype=float)  
  
    rows, columns = np.shape(coefficients)  
  
    # augmented matrix  
    augmented_mat = create_augmented_matrix(coefficients, vector)  
  
    # scale the matrix leaving it triangular  
    augmented_mat = scale_matrix(augmented_mat, rows)  
  
    x = retroactive_resolution_full_augmented_mat(augmented_mat, columns)  
    return x
```

Generate test

```
def gaussian_elimination(
    coefficients: NDArray[float64], vector: NDArray[float64]) -> NDArray[float64]:
    if (check_valid_input(coefficients, vector) == False):
        return np.array([], dtype=float)

    rows, columns = np.shape(coefficients)

    # augmented matrix
    augmented_mat = create_augmented_matrix(coefficients, vector)

    # scale the matrix leaving it triangular
    augmented_mat = scale_matrix(augmented_mat, rows)

    x = retroactive_resolution_full_augmented_mat(augmented_mat, vector)
    return x
```

```
import unittest
from math import nan, isnan
import numpy as np
from ..gaussian_elimination import gaussian_elimination
```

"""

Code Analysis:

- The function 'gaussian_elimination' solves a system of linear equations.
- It takes two inputs: 'coefficients' and 'vector', which represent the matrix and the right-hand side of the equations, respectively.
- The function first checks if the input is valid by calling 'check_valid_input()'. If the input is invalid, it returns an empty array.
- If the input is valid, the function creates an augmented matrix by combining the coefficients and the vector.
- The function then scales the augmented matrix to make it triangular using 'scale_matrix()'. This step involves dividing each row by its leading coefficient.
- Finally, the function solves the system of equations using 'retroactive_resolution_full_augmented_mat()'. This step involves back-substitution to find the values of the variables.
- The output of the function is an array containing the solution vector.

"""

"""

Test Plan:

- test_valid_input_with_unique_solution(): tests the function with a valid input and a unique solution.
- test_valid_input_with_multiple_solutions(): tests the function with a valid input and multiple solutions.
- test_valid_input_with_no_solution(): tests the function with a valid input and no solution.
- test_empty_input_array(): tests the function with empty input arrays.

Make this with PyTest

Generate test

```
def gaussian_elimination(
    coefficients: NDArray[float64], vector: NDArray[float64]) -> NDArray[float64]:
    """
    Solves a system of linear equations using Gaussian elimination.
    """
    if (check_valid_input(coefficients, vector) == False):
        return np.array([], dtype=float)

    rows, columns = np.shape(coefficients)

    # augmented matrix
    augmented_mat = create_augmented_matrix(coefficients, vector)

    # scale the matrix leaving it triangular
    augmented_mat = scale_matrix(augmented_mat, rows)

    x = retroactive_resolution_full_augmented_mat(augmented_mat,
    return x
```

```
import pytest
from math import nan, isnan
import numpy as np
from ..gaussian_elimination import gaussian_elimination
```

```
"""
```

Code Analysis:

- The function 'gaussian_elimination' solves a system of linear equations.
- It takes two inputs: 'coefficients' and 'vector', which represent the matrix and the right-hand side of the equations, respectively.
- The function first checks if the input is valid by calling 'check_valid_input'.
- If the input is valid, the function creates an augmented matrix by combining the coefficients and the vector.
- The function then scales the augmented matrix to make it triangular using 'scale_matrix'.
- Finally, the function solves the system of equations using 'retroactive_resolution_full_augmented_mat'.
- The output of the function is an array containing the solution vector.

```
"""
```

```
"""
```

Test Plan:

- test_valid_input_with_unique_solution(): tests the function with a valid input that has a unique solution.
- test_valid_input_with_multiple_solutions(): tests the function with a valid input that has multiple solutions.
- test_valid_input_with_no_solution(): tests the function with a valid input that has no solution.
- test_empty_input_arrays(): tests the function with empty input arrays.

Give me more

Generate test

```
def gaussian_elimination(  
    coefficients: NDArray[float64], vector: NDArray[float64]) ->  
  
    if (check_valid_input(coefficients, vector) == False):  
        return np.array([], dtype=float)  
  
    rows, columns = np.shape(coefficients)  
  
    # augmented matrix  
    augmented_mat = create_augmented_matrix(coefficients, vector)  
  
    # scale the matrix leaving it triangular  
    augmented_mat = scale_matrix(augmented_mat, rows)  
  
    x = retroactive_resolution_full_augmented_mat(augmented_mat,  
    return x
```

```
import pytest  
from math import nan, isnan  
import numpy as np  
from ..gaussian_elimination import gaussian_elimination  
  
"""  
Test Plan:  
- test_matrix_already_triangular(): tests the function with a  
- test_singular_matrix(): tests the function with a singular  
- test_ill_conditioned_matrix(): tests the function with an ill  
  
Additional instructions:  
- Assert with isnan in case of no solution  
"""  
  
class TestGaussianElimination:  
  
    def test_matrix_already_triangular(self):  
        # Test with a matrix that is already triangular  
        coefficients = np.array([[1, 0, 0], [2, 3, 0], [4, 5,  
        expected_output = np.array([[1, 0, 0], [2, 3, 0], [4, 5, 6]])
```

Try negative values & funny pirate-like comments

Kódmegértés

Buildt

**Buildt is an AI tool to help
developers quickly search
and understand large
codebases**

Engineers at companies like Stripe and Airbnb have to work with million-line codebases; our **LLM-powered** tool makes this simple

Current code search tools rely only on string matches and won't allow you to find what you want by describing the function vaguely

By mapping your codebase with our fine tuned LLM that we trained with proprietary synthetic data sets, our AI has a better than human understanding of any codebase.

So with Buildt you can ask questions about your own project as if an expert was next to you!

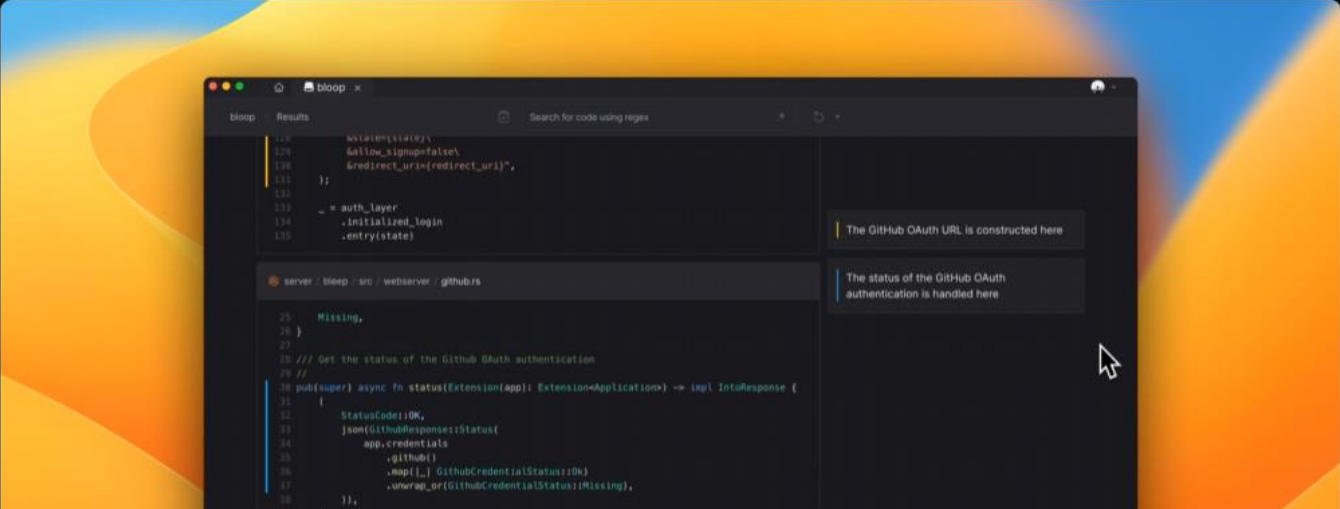
Bloop

bloop. Docs Pricing Blog GitHub · 5990 ★ Login [Get started](#)

Find Code. Fast.

Understand codebases with GPT-4
and semantic code search

[Try bloop for free](#) [Talk to sales](#)



The screenshot displays the Bloop application interface. At the top, there is a navigation bar with the Bloop logo, links for Docs, Pricing, and Blog, a GitHub repository link (5990 stars), and buttons for Login and Get started. The main content area features the headline "Find Code. Fast." followed by the subtext "Understand codebases with GPT-4 and semantic code search". Below this are two buttons: "Try bloop for free" and "Talk to sales". The central part of the image is a large screenshot of the Bloop application in use. It shows a dark-themed code editor with a search bar at the top that says "Search for code using regex". The search results are displayed in a list on the left, with the selected result showing code snippets. On the right side of the code editor, there are two annotations: "The GitHub OAuth URL is constructed here" and "The status of the GitHub OAuth authentication is handled here". The code editor also shows a file path "server / bloop / src / webserver / github.rs" and a snippet of Rust code.

<https://bloop.ai/>

What's our current approach to GitHub auth?



Our current approach to GitHub auth is to use the `octocrab: :auth` module to poll the GitHub API to get an authorization token, which is then stored in the d app's credentials. The relevant code can be found below:

```
let auth = loop {
  clock.tick().await;

  if Instant::now().duration_since(start) > Duration
    error! ("Github authorization timed out!");
    return;
}

match codes-poll_once(&github, &client_id).await {
  Ok(Either::Left (auth)) => break auth,
  Ok(Either::Right (cont)) => match cont {
    octocrab: tauth::Continue::SlowDown => {
      / We were request to slow down. We add five sec
      // duration.
```

Human Explanation

bloop understands your codebase. When responding to natural language searches, complex concepts are summarised and the intention behind code is explained.



Igor Susmelj

Cofounder, Lightly

Understanding code is a major component of every developer's day. bloop's summarization helps close the development loop by speeding up code review, planning and many other tasks.

Kódmigráció

Automatically

Put code migrations and dependency upgrades on autopilot

Get Early Access →

Schedule a Demo

JS to TS Migration

Class to Functional Components

JS to Arrow Functions

AngularJS to Angular



```
export function heroDetail Directive() {  
  return {  
    restrict: "E",  
    scope: {},  
    bindToController: {  
      hero: "=",  
      deleted: "&",  
    },  
    template: `  
      <h2>{{ctrl.hero.name}} details!</h2>  
      <div><label>id: </label>{{ctrl.hero.id}}</div>  
      <button type="button" ng-click="$ctrl.onDelete()">  
        Delete  
      </button>
```


R&D

BLOG ›

Large sequence models for software development activities

WEDNESDAY, MAY 31, 2023

Posted by Petros Maniatis and Daniel Tarlow, Research Scientists, Google

Software isn't created in one dramatic step. It improves bit by bit, one little step at a time — editing, running unit tests, fixing build errors, addressing code reviews, editing some more, appeasing [linters](#), and fixing more errors — until finally it becomes good enough to merge into a code repository. Software engineering isn't an isolated process, but a dialogue among human developers, code reviewers, bug reporters, software architects and tools, such as compilers, unit tests, linters and static analyzers.

The novelty of DIDACT is that it uses *the process of software development as the source of training data for the model, rather than just the polished end state of that process, the finished code.*

finished code. By exposing the model to the contexts that developers see as they work, paired with the actions they take in response, the model learns about the dynamics of software development and is more aligned with how developers spend their time. We leverage instrumentation of Google's software development to scale up the quantity and diversity of developer-activity data beyond previous works. Results are extremely promising along two dimensions: usefulness to professional software developers, and as a potential basis for imbuing ML models with general software development skills.

DIDACT

Debugging & Repair Tasks

Build Repair

Build Error Prediction

TFix

Code Review Tasks

Comment Prediction

Tip Prediction

Comment Repair

Pre-Submit Cleanup

Readability Renaming

Code Review
Autocompletion

Code Editing

Edit Prediction

Span Denoising

Variable Renaming

History-Augmented
Code Completion

Prompting & Chaining

Python ökoszisztéma

PandasAI

PandasAI

Release **v0.3.0**  ci **passing**  cd **failing** docs **passing**  PandasAI downloads/month **28k** License **MIT**  Open in Colab

Pandas AI is a Python library that adds generative artificial intelligence capabilities to Pandas, the popular data analysis and manipulation tool. It is designed to be used in conjunction with Pandas, and is not a replacement for it.



<https://github.com/gventuri/pandas-ai>

```
import pandas as pd
from pandasai import PandasAI

# Sample DataFrame
df = pd.DataFrame({
    "country": ["United States", "United Kingdom", "Germany", "Canada", "Australia"],
    "gdp": [19294482071552, 2891615560000, 3568145000000, 1714611000000, 1536016070000],
    "happiness_index": [6.94, 7.16, 7.02, 7.23, 7.18]
})

# Instantiate a LLM
from pandasai.llm.openai import OpenAI
llm = OpenAI(api_token="YOUR_API_TOKEN")

pandas_ai = PandasAI(llm, conversational=False)
pandas_ai(df, prompt='Which are the 5 happiest countries?')
```

The above code will return the following:

```
6          Canada
7          Australia
1    United Kingdom
3           Germany
0    United States
Name: country, dtype: object
```

```
import pandas as pd
from pandasai import PandasAI
```

```
pandas_ai(df, prompt='What is the sum of the GDPs of the 2 unhappiest countries?')
```

The above code will return the following:

```
19012600725504
```

```
pandas_ai = PandasAI(llm, conversational=False)
pandas_ai(df, prompt='Which are the 5 happiest countries?')
```


Jupyter AI

Jupyter AI

Welcome to Jupyter AI, which brings generative AI to Jupyter. Jupyter AI provides a user-friendly and powerful way to explore generative AI models in notebooks and improve your productivity in JupyterLab and the Jupyter Notebook. More specifically, Jupyter AI offers:

- An `%%ai` magic that turns the Jupyter notebook into a reproducible generative AI playground. This works anywhere the IPython kernel runs (JupyterLab, Jupyter Notebook, Google Colab, VSCode, etc.).
- A native chat UI in JupyterLab that enables you to work with generative AI as a conversational assistant.
- Support for a wide range of generative model providers and models (AI21, Anthropic, Cohere, Hugging Face, OpenAI, SageMaker, etc.).

Documentation is available on [ReadTheDocs](#).

Then, you can use the `%%ai` magic command to specify a model and natural language prompt:

```
[15]: %%ai chatgpt
Please generate the Python code to solve the 2D Laplace equation in cartesian coordinates.
Solve the equation on the square domain x=(0,1) and y=(0,1) with vanishing boundary conditions.
Plot the solution using Matplotlib.
Please also provide an explanation.
```

Here's the Python code to solve the 2D Laplace equation in Cartesian coordinates:

```
import numpy as np
import matplotlib.pyplot as plt

# Set up grid
nx = 101
ny = 101
nt = 100

dx = 1. / (nx - 1)
dy = 1. / (ny - 1)

x = np.linspace(0, 1, nx)
y = np.linspace(0, 1, ny)

# Initialize solution
u = np.zeros((nx, ny))
```

Data Science notebooks

Data Science Notebooks

Notebook Landscape



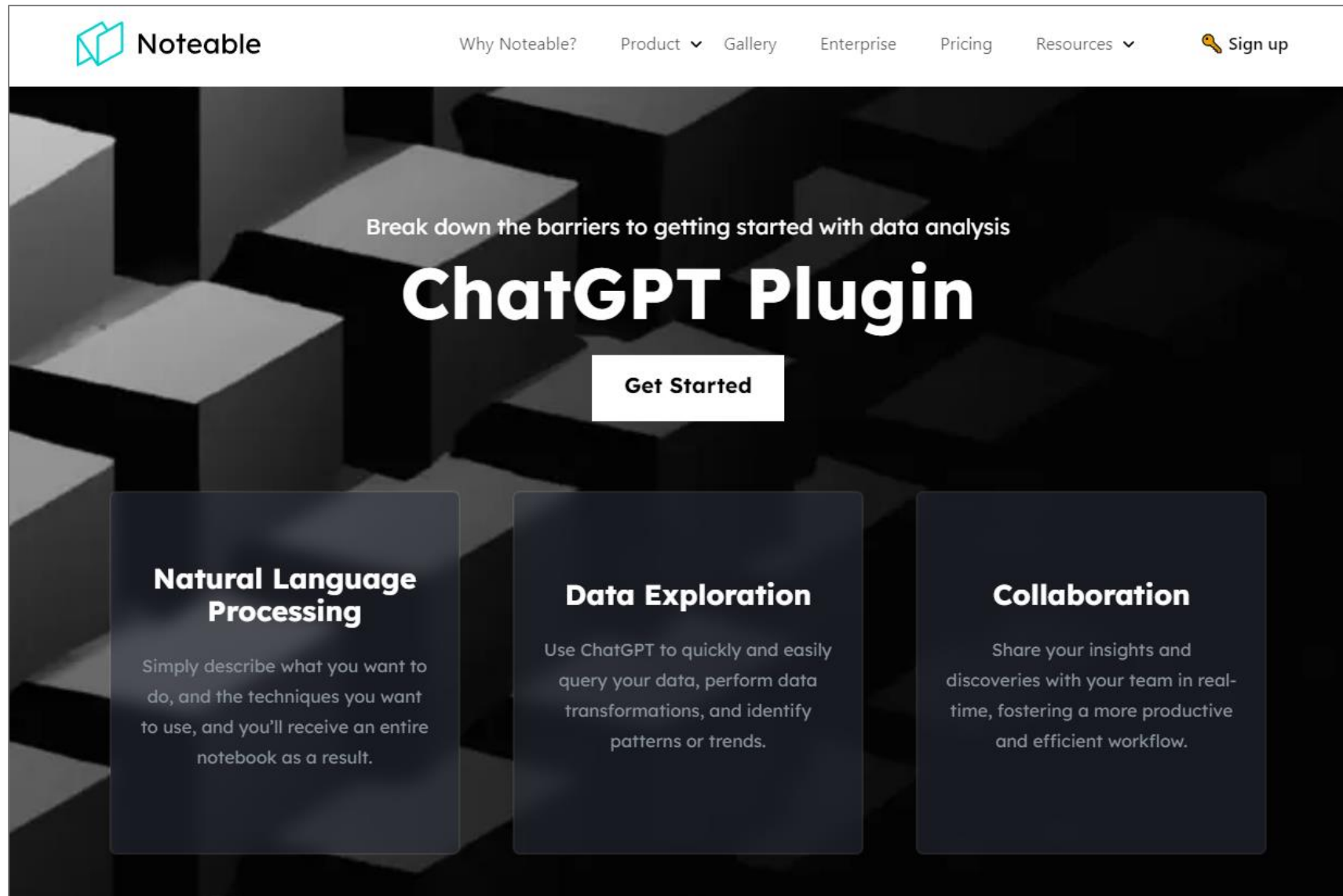
amazon web services ANACONDA CLOUDERA COCALC count databricks

Deepnote DOMINO colab HEX jovian.ml jupyter

kaggle Microsoft Azure MODE Nextjournal noteable nteract

Observable Polynote Saturn Cloud APACHE Spark notebooks Apache Zeppelin

@AstasiaMyers



The screenshot shows the Noteable website's landing page for the ChatGPT Plugin. The page features a dark background with a 3D geometric pattern of cubes. At the top, there is a navigation bar with the Noteable logo and several menu items: 'Why Noteable?', 'Product', 'Gallery', 'Enterprise', 'Pricing', 'Resources', and 'Sign up'. The main heading reads 'Break down the barriers to getting started with data analysis' followed by 'ChatGPT Plugin' in a large, bold font. Below the heading is a 'Get Started' button. Three feature cards are displayed: 'Natural Language Processing', 'Data Exploration', and 'Collaboration', each with a brief description of its capabilities.

Noteable Why Noteable? Product Gallery Enterprise Pricing Resources Sign up

Break down the barriers to getting started with data analysis

ChatGPT Plugin

Get Started

Natural Language Processing

Simply describe what you want to do, and the techniques you want to use, and you'll receive an entire notebook as a result.

Data Exploration

Use ChatGPT to quickly and easily query your data, perform data transformations, and identify patterns or trends.

Collaboration

Share your insights and discoveries with your team in real-time, fostering a more productive and efficient workflow.

The image shows a screenshot of the Hex website's landing page for Hex Magic. The background is dark blue with a subtle grid pattern. At the top left is the 'HEX' logo. The navigation menu includes 'Product', 'Learn', 'Enterprise', 'Pricing', and 'About'. On the top right, there are 'Log in' and 'Get started' buttons. The central focus is a 'SQL cell' interface for 'Atriedes DB' with a 'Browse schema' button. The SQL query in the cell is:

```
select order_date,
sum(number_of_items) as total_items
from prod.dim_orders
group by 1
order by 1 desc
```

 A 'Watch demo (1 min)' button is overlaid on the SQL code. Below the SQL cell, the text reads 'INTRODUCING Hex Magic AI-powered tools for humans doing amazing things with data.' At the bottom, it says 'PUBLIC BETA'.

HEX

Product ▾ Learn ▾ Enterprise Pricing About ▾

Log in Get started

SQL cell

Atriedes DB ▾ Browse schema

EDIT Truncate order date to month and add a 3 month rolling average

▶ Watch demo (1 min)

```
select order_date,
sum(number_of_items) as total_items
from prod.dim_orders
group by 1
order by 1 desc
```

INTRODUCING

Hex Magic

AI-powered tools for humans doing amazing things with data.

PUBLIC BETA

Táblázatkezelő

Spreadsheets, meet AI Automation

Join thousands of analysts at the world's largest banks saving themselves from hours of repetitive work.

Install Mito

or see Pro plans →

```
[5]: import mitosheet  
      mitosheet.sheet()
```

	order_id	order_date	product_id	quantity_order	price_each	status	sales_person
	10	2021-01-19	prod_25f2d5b58b70	1	59.99	Shipped	Janessa Benton
	11	2021-01-19	prod_9c88bd37f2f6	2	15.99	Shipped	Brenda Webb
	12	2021-01-19	prod_3b5cac1a5a91	2	9.99	Shipped	Grace Patterson

Mito AI

Mito AI is one of the fastest ways to transform your data. This documentation explains how.

What is Mito AI?

Mito AI is ChatGPT for your pandas dataframes. Its the easiest way to apply simple edits, like adding filters or parsing strings, to your data.

Like ChatGPT, Mito AI is a chat interface for interacting with OpenAI. Unlike ChatGPT:

1. When you use Mito AI to transform your data, it automatically executes the code in the context of your analysis so you'll immediately see the effects on your data. This makes it easier to decide if the code generated by OpenAI was correct or whether you want to undo the edit and try again.
2. Mito AI has context about your data and your analysis. This additional information, which Mito automatically provides to OpenAI without you having to type it out, helps OpenAI generate code that is useful to you right away.

Edit Dataframes Columns Rows Graphs Format Code View Help Get Support Mito Pro

Undo Redo Clear Import Export Add Col Del Col Dtype Less More Number Pivot Graph AI Steps Fullscreen

member type | Member

	duration	start time	end time	start station	end station	bike number	member type
	int	str	str	str	str	str	str
0	1,012	2010-09-20 11:27...	2010-09-20 11:43...	M St & New Jersey...	4th & M St SW	W00742	Member
1	61	2010-09-20 11:41...	2010-09-20 11:42...	1st & N St SE	1st & N St SE	W00032	Member
2	2,690	2010-09-20 12:05...	2010-09-20 12:50...	5th & K St NW	19th St & Pennsylv...	W00993	Member
3	1,406	2010-09-20 12:06...	2010-09-20 12:29...	5th & K St NW	Park Rd & Holmea...	W00344	Member
4	1,413	2010-09-20 12:10...	2010-09-20 12:34...	19th St & Pennsylv...	15th & P St NW	W00883	Member
5	982	2010-09-20 12:14...	2010-09-20 12:30...	7th & T St NW	Massachusetts Av...	W00850	Member
6	930	2010-09-20 12:15...	2010-09-20 12:30...	7th & T St NW	14th & Harvard St ...	W00804	Member
7	1,659	2010-09-20 12:16...	2010-09-20 12:44...	10th & U St NW	M St & New Jersey...	W01084	Member
8	2,496	2010-09-20 12:18...	2010-09-20 13:00...	5th & K St NW	19th & East Capito...	W00812	Member
9	1,487	2010-09-20 12:19...	2010-09-20 12:44...	Minnesota Ave Me...	1st & M St NE	W00803	Member
10	1,007	2010-09-20 12:21...	2010-09-20 12:38...	4th & W St NE	Massachusetts Av...	W00256	Member
11	1,159	2010-09-20 12:26...	2010-09-20 12:45...	1st & M St NE	1st & M St NE	W01023	Member
12	88	2010-09-20 12:31...	2010-09-20 12:32...	Park Rd & Holmea...	Park Rd & Holmea...	W00472	Member
13	202	2010-09-20 12:33...	2010-09-20 12:36...	14th & Harvard St ...	16th & Harvard St ...	W00685	Member

+ bike_rides v (115597 rows, 7 cols)



Mélyvíz

Új programnyelvek



LMQL

A query language for programming (large) language models.

[Documentation](#) »

[Explore Examples](#) · [Playground IDE](#) · [Report Bug](#)

 chat **90 online**  package **0.0.6.3**

LMQL is a query language for large language models (LLMs). It facilitates LLM interaction by combining the benefits of natural language prompting with the expressiveness of Python. With only a few lines of LMQL code, users can express advanced, multi-part and tool-augmented LM queries, which then are optimized by the LMQL runtime to run efficiently as part of the LM decoding loop.

`argmax` specifies the use of `argmax` decoding for this query, but LMQL also supports `sample` and `beam` search. Decoding parameters like `sampling temperature` can also be specified.

The `from` clause specifies the identifier of a text generation model from the 🤗 Transformers model repository or an OpenAI model like `text-davinci-003`.

LMQL supports high-level constraints, where the language runtime automatically derives token-level prediction masks and validates the produced sequence eagerly, i.e. as soon as the provided validation condition is definitively violated, decoding stops or is redirected to a different branch.

```
LMQL
• argmax
  """A list of good dad jokes. A indicates
  - the punchline
  Q: How does a penguin build its house?
  A: Igloos it together.
  Q: Which knight invented King Arthur's
  - Round Table?
  A: Sir Cumference.
  Q: [JOKE]
  A: [PUNCHLINE] """
• from
  "openai/text-davinci-003"
• where
  len(JOKE) < 120 and
  STOPS_AT(JOKE, "?") and
  STOPS_AT(PUNCHLINE, "\n") and
  len(PUNCHLINE) > 1
```

MODEL OUTPUT

A list of good dad jokes. A indicates the punchline
Q: How does a penguin build its house?
A: Igloos it together.
Q: Which knight invented King Arthur's Round Table?
A: Sir Cumference.
Q: [JOKE] What did the fish say when it hit the wall?
A: [PUNCHLINE] Dam!

Highlighted text is model output.

Example of a simple LMQL program.



Mojo — a new programming language for all AI developers.

Mojo combines the usability of Python with the performance of C, unlocking unparalleled programmability of AI hardware and extensibility of AI models.

Write everything
in one language

Write Python or scale all the way down to the metal. Program the multitude of low-level AI hardware. No C++ or CUDA required.

Take a tour of Mojo

FEATURES INCLUDE:

PROGRESSIVE TYPES

Leverage types for better performance and error checking.

ZERO COST ABSTRACTIONS

Take control of storage by inline-allocating values into structures.

OWNERSHIP + BORROW CHECKER

Take advantage of memory safety without the rough edges.

PORTABLE PARAMETRIC ALGORITHMS

○○○ FILE_NAME 🔥

```
def reorder_and_process(owned x: HugeArray):  
    sort(x) # Update in place  
  
    give_away(x^) # Transfer ownership  
  
    print(x[0]) # Error: 'x' moved away!
```

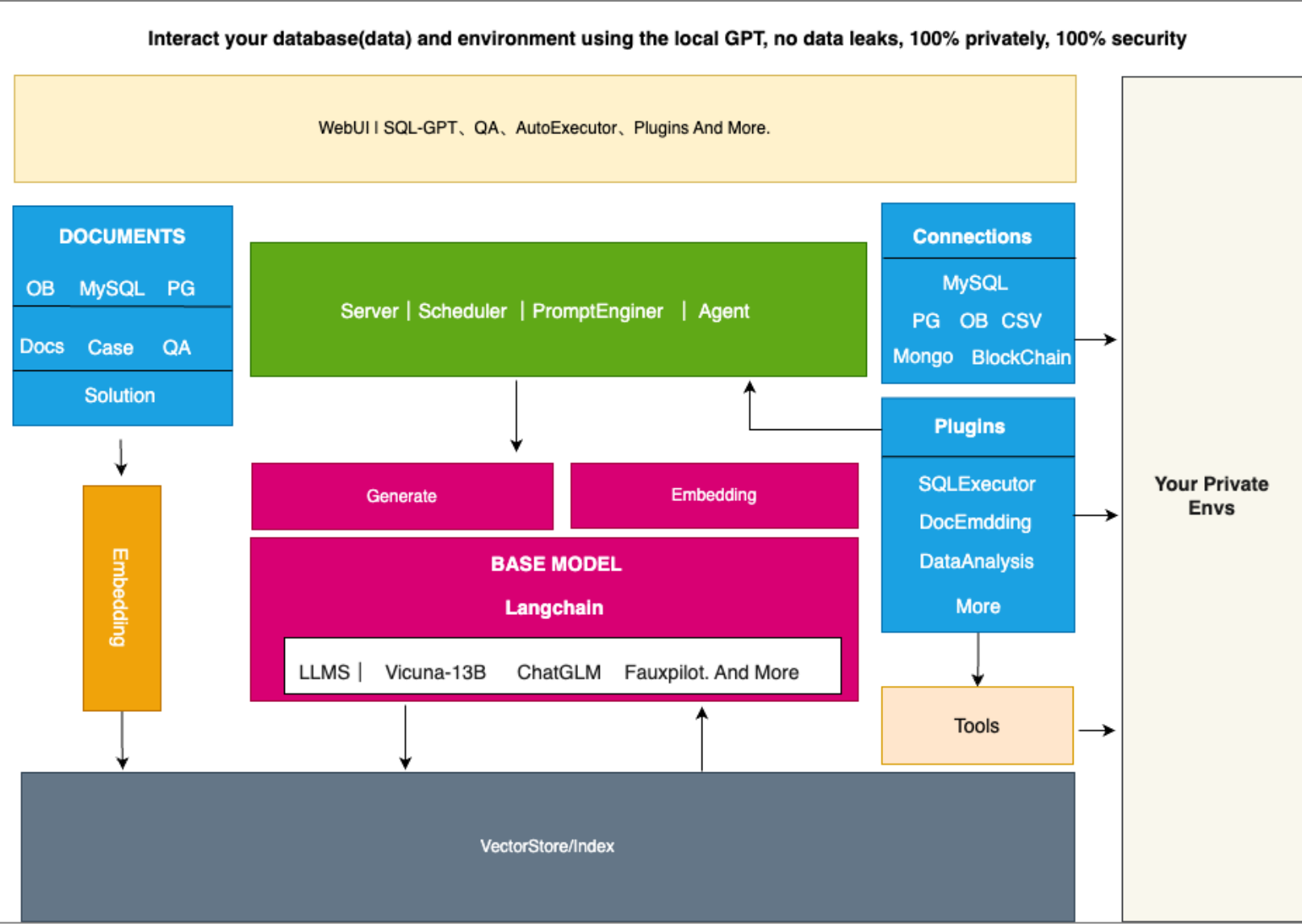
Helyi LLM infrastruktúra

What is DB-GPT?

As large models are released and iterated upon, they are becoming increasingly intelligent. However, in the process of using large models, we face significant challenges in data security and privacy. We need to ensure that our sensitive data and environments remain completely controlled and avoid any data privacy leaks or security risks. Based on this, we have launched the DB-GPT project to build a complete private large model solution for all database-based scenarios. This solution supports local deployment, allowing it to be applied not only in independent private environments but also to be independently deployed and isolated according to business modules, ensuring that the ability of large models is absolutely private, secure, and controllable.

DB-GPT is an experimental open-source project that uses localized GPT large models to interact with your data and environment. With this solution, you can be assured that there is no risk of data leakage, and your data is 100% private and secure.

DB-GPT

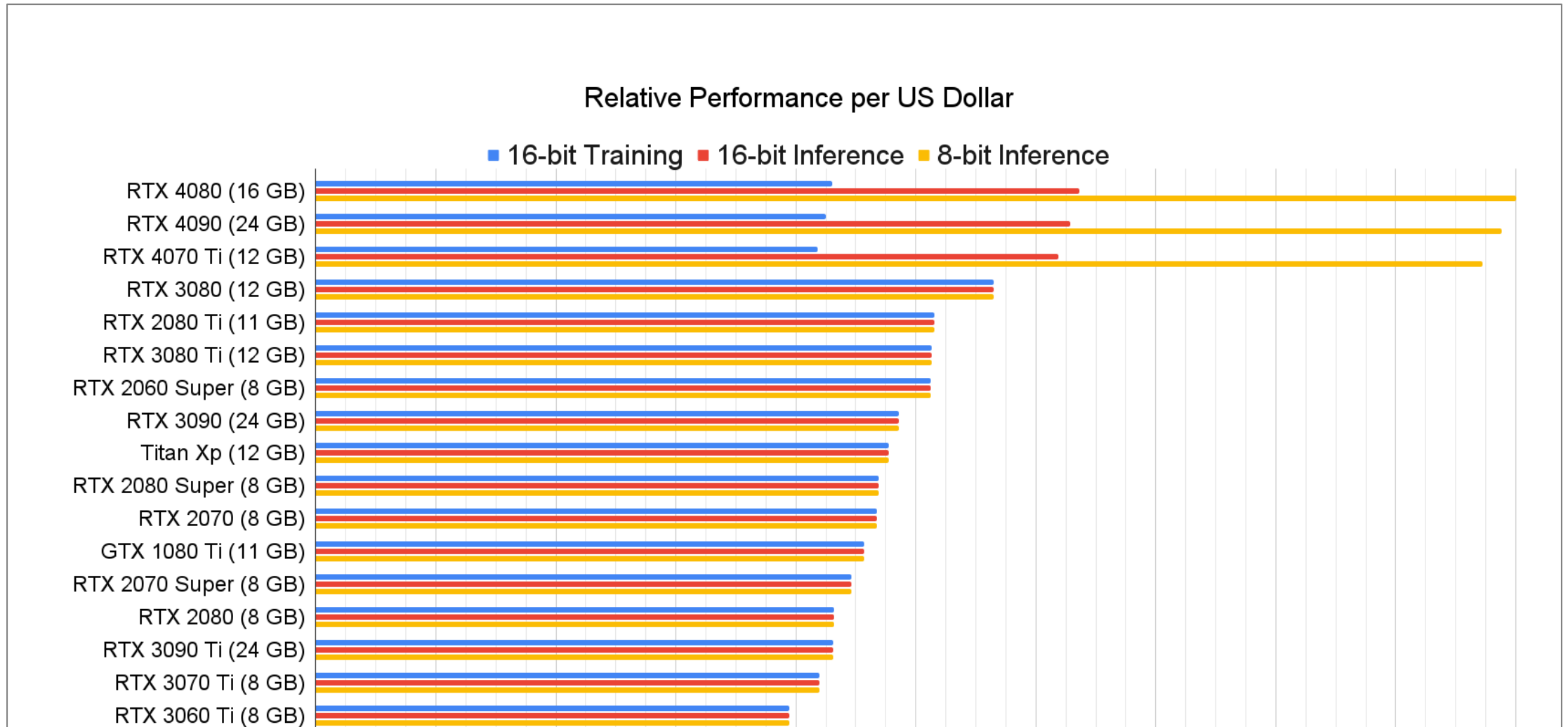


1. Hardware Requirements

As our project has the ability to achieve ChatGPT performance of over 85%, there are certain hardware requirements. However, overall, the project can be deployed and used on consumer-grade graphics cards. The specific hardware requirements for deployment are as follows:

GPU	VRAM Size	Performance
RTX 4090	24 GB	Smooth conversation inference
RTX 3090	24 GB	Smooth conversation inference, better than V100
V100	16 GB	Conversation inference possible, noticeable stutter

GPU választás



GPU választás



Összehasonlítás

[MSI GeForce RTX 4090 VENTUS 3X OC 24GB GDDR6X \(V510-023R\) Videokártya](#)

Video chipset termékcsalád: **RTX 4090**

Video chipset: **Nvidia GeForce**

Hűtés típusa: **Aktív hűtés**

Ventilátorok száma: **3 darab**

Memória mérete: **24 GB**

Memória típusa: **GDDR6X**

Chipset gyártó:NVIDIA, DisplayPort kimenet:3, GPU órajel (MHz):2565, HDMI kimenet:1, Memória méret (GB):24GB, PCI-E x16 csatlakozó:1, Gamer:Van, Fő tulajdonság:grafikus kártya,...

686 329 Ft-tól

árak összevetése

18 ajánlat



Összehasonlítás

[GIGABYTE GeForce RTX 4090 Windforce 24GB \(GV-N4090WF3-24GD\) Videokártya](#)

Video chipset termékcsalád: **RTX 4090**

Video chipset: **Nvidia GeForce**

Hűtés típusa: **Aktív hűtés**

Ventilátorok száma: **3 darab**

Memória mérete: **24 GB**


Memória típusa: **GDDR6X**

Gyártó: GIGABYTE Modell: GeForce RTX 4090 Windforce 24GB (GV-N4090WF3-24GD) Leírás: A mesterséges intelligencia által működtetett teljesítményszorzó. A DLSS egy forradalmi áttörés...

692 849 Ft-tól

árak összevetése

26 ajánlat

 Az Árkereső.hu-n is megrendelhető



További Data alkalmazások

Data Governance

Atlan AI

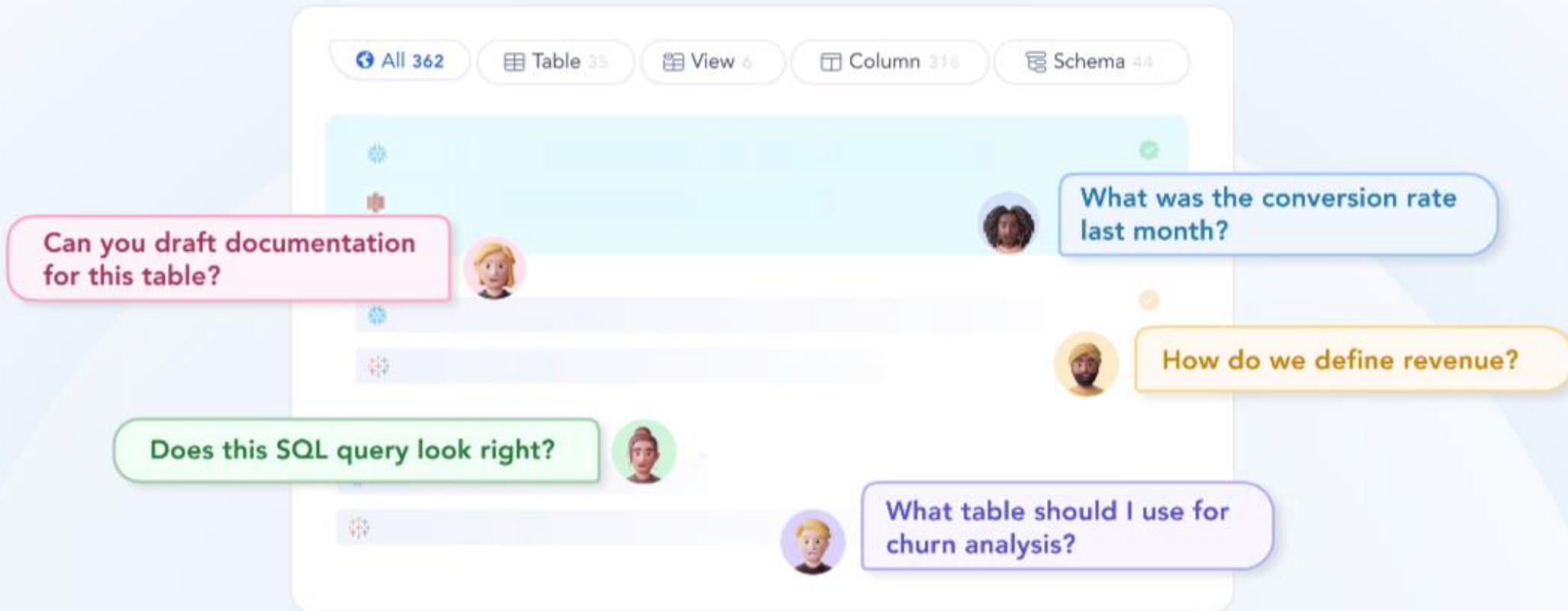
The first ever copilot for
data teams



Join waitlist



Now, every **data question** has an answer.



Stemma makes AI Discover Assistant available for all customers

 Discover Assistant ^

 Edit

This table combines data from various sources related to orders, drivers, customers, restaurants, and cancellations. It could be used to analyze trends in revenue, delivery counts, and customer retention, as well as identify reasons for order cancellations and potential areas for improvement in the onboarding funnel. The table also includes information on restaurant regions, which could be useful for sales and marketing purposes.



[Report Inaccurate Info](#)

Stemma

food_delivery
core.customer_contacts
5 days ago

food_delivery
core.order_items
5 days ago

food_delivery
core.restaurants
5 days ago

food_delivery
events.order_cancelled
6 days ago

food_delivery
core.driver_docs
5 days ago

food_delivery
core.orders
5 days ago

food_delivery
events.order_cancelled
Last Updated: May 31, 2023 12am CDT

[Open In New Graph](#) [Table Details](#)

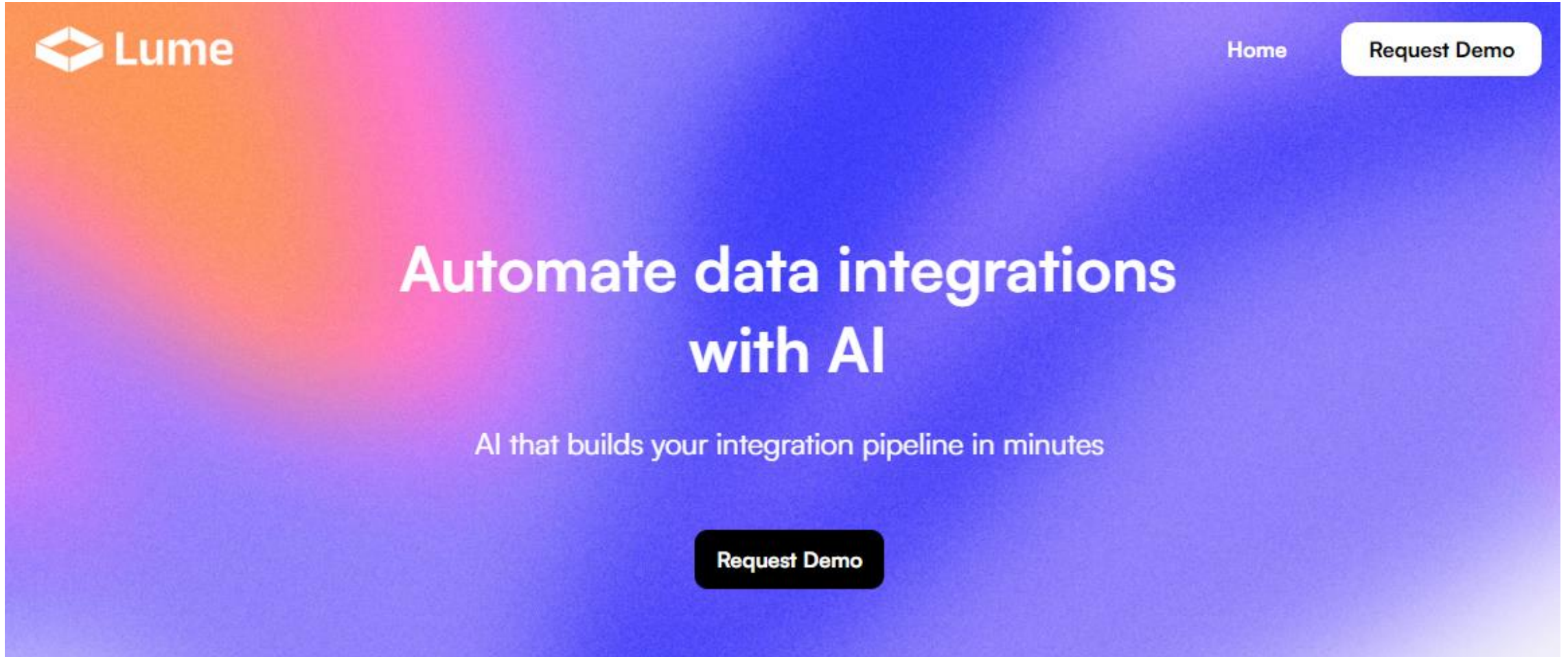
[Hide In Table Group](#)

About This Connection [View Query](#)

The table "events.order_cancelled" is joined to the query through a left join with the table "core.dim_cancel_reasons". The column "cancel_reason" within the "events.order_cancelled" table is used to match records with the corresponding cancel reason in the "core.dim_cancel_reasons" table. This join allows for the cancel reason to be included in the final output. Additionally, the table is not used to filter records or transform columns within itself.

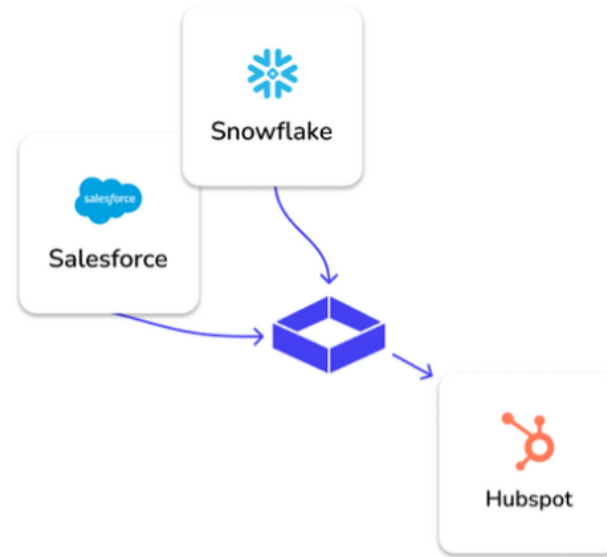
Search Columns
Find columns by name

Adatintegráció

The image shows the hero section of the Lume website. It features a vibrant, multi-colored gradient background transitioning from orange on the left to blue on the right. In the top left corner, the Lume logo is displayed, consisting of a white cube icon followed by the word "Lume" in a white sans-serif font. In the top right corner, there are two navigation elements: the word "Home" in white text and a white rounded rectangular button with the text "Request Demo" in black. The main headline is centered in large, bold, white text, reading "Automate data integrations with AI". Below the headline, a sub-headline in smaller white text reads "AI that builds your integration pipeline in minutes". At the bottom center, there is a black rounded rectangular button with the text "Request Demo" in white.

Our AI automatically transforms data schemas for you

Provide any start and end schema, and Lume will transform and deliver the data to your destination, no code required.



```
SELECT
  COALESCE(address.ONELINE, address.LINE1, '')
  AS FULL_ADDRESS,
  building.BEDS,
  address.LOCALITY AS CITY,
  identifier.ID,
  assessment.CALCTTLVALUE AS LIST_PRICE,
  NULL AS STATUS,
  lot.LOTSIZE1 AS LOT_SIZE,
  building.YEARBUILTEFFECTIVE AS BUILT_IN,
  NULL AS SELLER,
  summary.PROPTYPE AS TYPE,
  building.BATHSFULL AS BATHS,
  assessment.CALCTTL
  AS PRICE_PER,
  building.GROSSSI
FROM
  {{ ref('address')
LEFT JOIN {{ ref(
AS building ON ac
```

address_d

building_d

my_schema

Review the data output before deployment

Confirm validity of transformations by reviewing key information:

- ✓ Sample data
- ✓ Lineage graph
- ✓ Underlying SQL code

Grafikonok elemzése

[BLOG](#) ›

Foundation models for reasoning on charts

FRIDAY, MAY 26, 2023

Posted by Julian Eisenschlos, Research Software Engineer, Google Research

Visual language is the form of communication that relies on pictorial symbols outside of text to convey information. It is ubiquitous in our digital life in the form of iconography, infographics, tables, plots, and charts, extending to the real world in street signs, comic books, food labels, etc. For that reason, having computers better understand this type of media can help with scientific communication and discovery, accessibility, and data transparency.

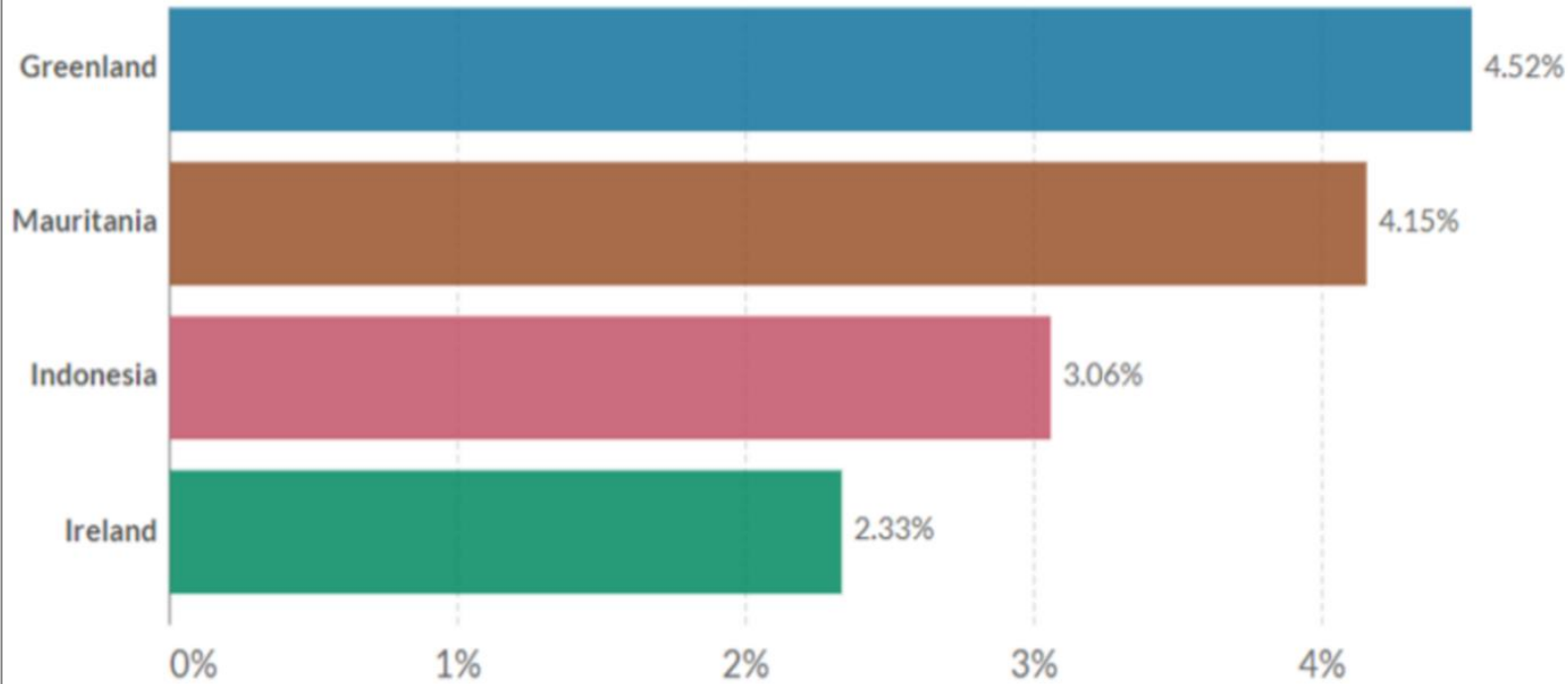
While computer vision models have made tremendous progress using learning-based solutions since the advent of [ImageNet](#), the focus has been on natural images, where all sorts of tasks, such as [classification](#), [visual question answering](#) (VQA), [captioning](#), [detection and segmentation](#), have been defined, studied and in some cases advanced to reach human performance. However, visual language has not garnered a similar level of attention, possibly because of the lack of large-scale training sets in this space. But over the last few years, new academic datasets have been created with the goal of evaluating question answering systems on visual language images, like [PlotQA](#), [InfographicsVQA](#), and [ChartQA](#).

Visual language image

Share of marine territorial waters that are protected, 2018

Marine protected areas are areas of intertidal or subtidal terrain - and overlying water and associated flora and fauna and historical and cultural features - that have been reserved by law or other effective means to protect part or all of the enclosed environment.

Our World
in Data



Question

What is the difference between the sum shares of Indonesia and Ireland and the share of Mauritania?

ChartQA: A Benchmark for Question Answering about Charts with Visual and Logical Reasoning

Ahmed Masry, Xuan Long Do, Jia Qing Tan, Shafiq Joty, Enamul Hoque

Abstract

Charts are very popular for analyzing data. When exploring charts, people often ask a variety of complex reasoning questions that involve several logical and arithmetic operations. They also commonly refer to visual features of a chart in their questions. However, most existing datasets do not focus on such complex reasoning questions as their questions are template-based and answers come from a fixed-vocabulary. In this work, we present a large-scale benchmark covering 9.6K human-written questions as well as 23.1K questions generated from human-written chart summaries. To address the unique challenges in our benchmark involving visual and logical reasoning over charts, we present two transformer-based models that combine visual features and the data table of the chart in a unified way to answer questions. While our models achieve the state-of-the-art results on the previous datasets as well as on our benchmark, the evaluation also reveals several challenges in answering complex reasoning questions.

Anthology ID: 2022.findings-acl.177


Volume: Findings of the Association for Computational Linguistics: ACL 2022


Month: May

Year: 2022

 PDF

 Cite

 Search

 Code

 Video

És végül

ChatGPT: Imagine you are a database server

Posted on Dec 2, 2022

#machine learning

#aiml

#db





After reading [this post](#) about ChatGPT imitating Linux, **I wanted it to be a database server.**

Let's try it out!





Imagine you are a Microsoft SQL Server. I type commands, and you reply with the result, and no other information or descriptions. Just the result. Start with `exec xp_cmdshell 'whoami'`;

ChatGPT, az adatbáziskezelő





AR Imagine you are a Microsoft SQL Server. I type commands, and you reply with the result, and no other information or descriptions. Just the result. Start with CREATE DATABASE EventsDB

 Database created successfully.   

AR CREATE TABLE DataEvents (Year INT, EventName NVARCHAR(MAX));

 Table created successfully.   

AR INSERT INTO DataEvents VALUES (2022, 'BudapestData');
INSERT INTO DataEvents VALUES (2022, 'BudapestML');
INSERT INTO DataEvents VALUES (2023, 'Data+ML');

 Rows inserted successfully.   

AR select * from DataEvents;

```
markdown
Copy code

Year | EventName
-----
2022 | BudapestData
2022 | BudapestML
2023 | Data+ML
```

Thank You

