

ScrapeGraphAI

You Only Scrape Once

GPT-4 Powered App Creation and Evals Hackathon



intel.



Pinecone

Our Team



Marco Vinciguerra

MSc Computer Engineering

mvincig11@gmail.com



Marco Perini

MSc Mechatronics Engineering

perinim.98@gmail.com



Lorenzo Padoan

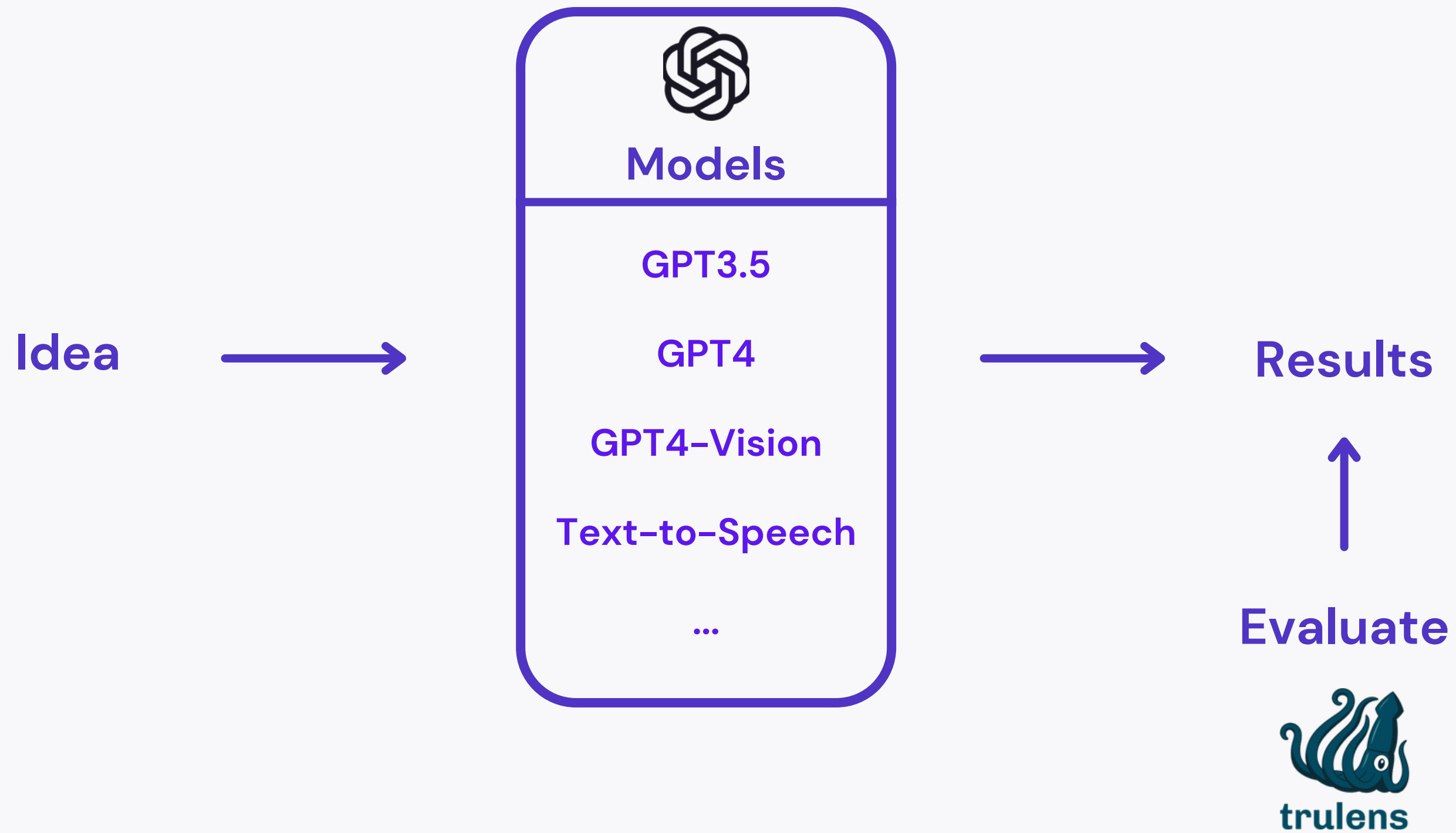
Research Fellow at UNIPD

–

Data Engineer @
Motion Analytica

lorenzo.padoan977@gmail.com

Challenge



Our Question

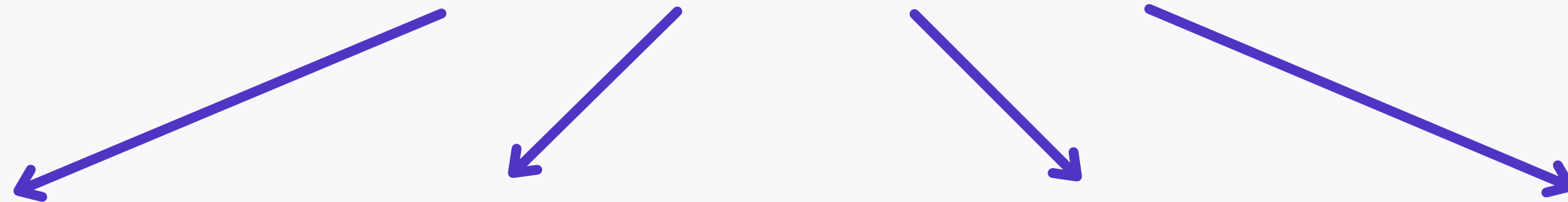


Is it possible to **scrape** websites **without** any knowledge of **HTML**, just by **writing** what I want and **how** I want it?

Solution



ScrapeGraphAI



**Python
Library**

**Open
Source**

**Direct Graph based
Scraping Pipelines**

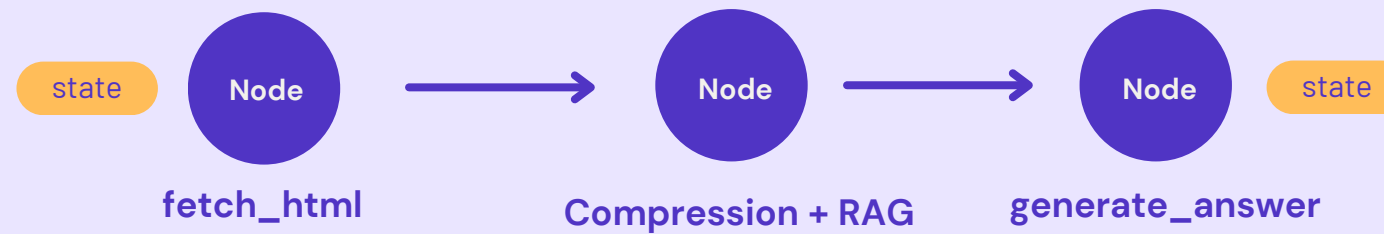
**Highly
Customizable**

Solution

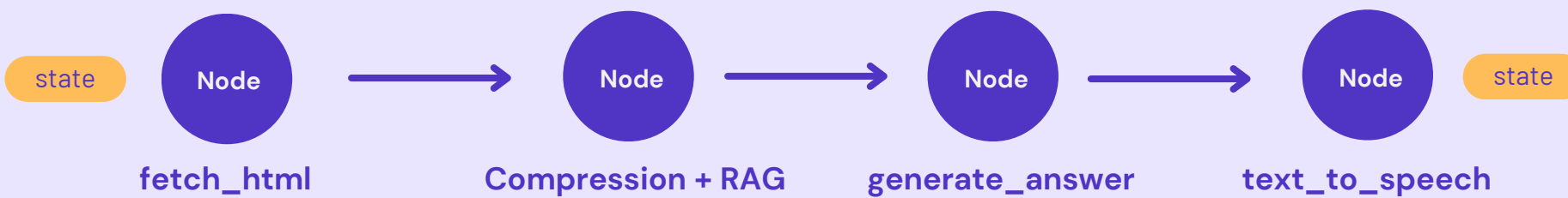


URL
+
User Prompt

SmartScrapper Graph



SpeechSummary Graph



Solution



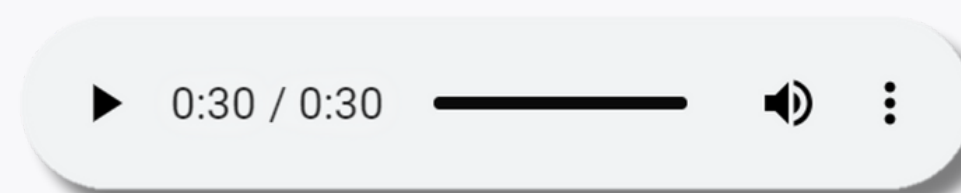
Prompt 1: Give me all the **news** with their **abstracts**

↳ **Answer:** `{'news': [{'title': 'The LabLabAI hackathon deadline is close!',
'abstract': 'Be ready to be shocked by the teams final submissions'}]}`

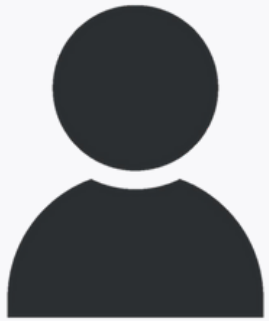
Prompt 2: Make an **audio** summary of the **news**

↳ **Answer:** `{'news': [{'title': 'The LabLabAI hackathon deadline is close!',
'summary': 'Deadlines are coming close so be ready...'}]}`

+



Potential Customers



Developers

+



Companies

+



Researchers

Call for Action



Early Adopters

+

Community

+

Partnerships

Initial
Investments

Demo



Scrapegraph-ai



API key

Select the model

- gpt-3.5-turbo
- gpt-3.5-turbo-0125
- gpt-4
- text-to-speech

Link to scrape

Write the prompt

Run the program



Repositories



ScrapeGraphAI

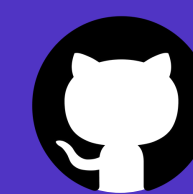
+



Streamlit Website

VinciGit00/Scrapegraph-ai

VinciGit00/Scrapegraph-LabLabAI-Hackathon



ScrapeGraphAI

You Only Scrape Once