



# Leveraging Language Models for Technology Landscape Monitoring

**PRESENTER: André Charneca (CYD/EPFL)**

Angelika Romanou (EPFL)

Dr. Ljiljana Dolamic (CYD)

Prof. Karl Aberer (EPFL)

# What do we want to build?



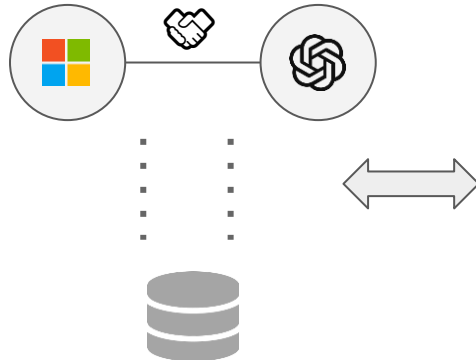
## News Articles

- Most recent
- Multiple connections
- Complex information

MICROSOFT / TECH

**Microsoft extends OpenAI partnership in a 'multibillion dollar investment'** / Microsoft says it has invested billions into OpenAI in a multiyear deal that will see the software giant become OpenAI's exclusive cloud provider.

By TOM WAREEN / @tomwareen  
Jan 25, 2023, 5:14 PM GMT+1 | 0 Comments / 9 Views



## NLU system

Understand **Company & Technology** entities

Understand **relations** between entities

**Reason** over multiple relations

**Infer** requested information to representations & data



- ✓ Who is working on NLP & AI ?
- ✓ Who is investing in OpenAI ?
- ✓ Who started investing on generative models when OpenAI launched ChatGPT?

# Project Goal

## Objective and Motivation



### Objective

Extract company-company, company-product and product-product relationships from news articles.



### Why?

Structured data enables:

- Clearer representation
- Cross-article connections
- Easier storage & querying
- Exploration of more complex reasoning structures

ARTICLE

Unstructured Data



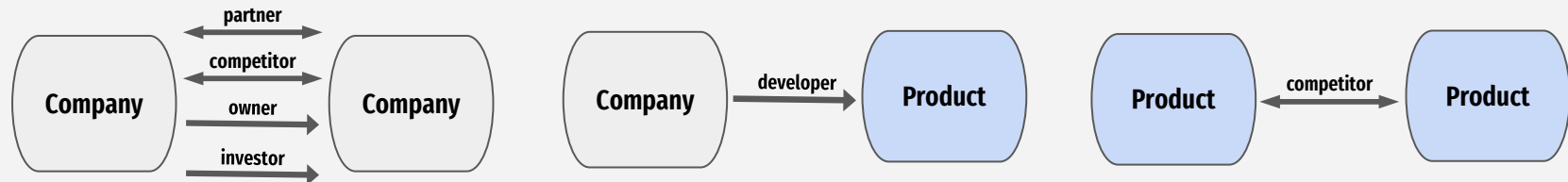
GRAPH



Structured Data

# Project Goal Illustration

## Possible Relationship Types

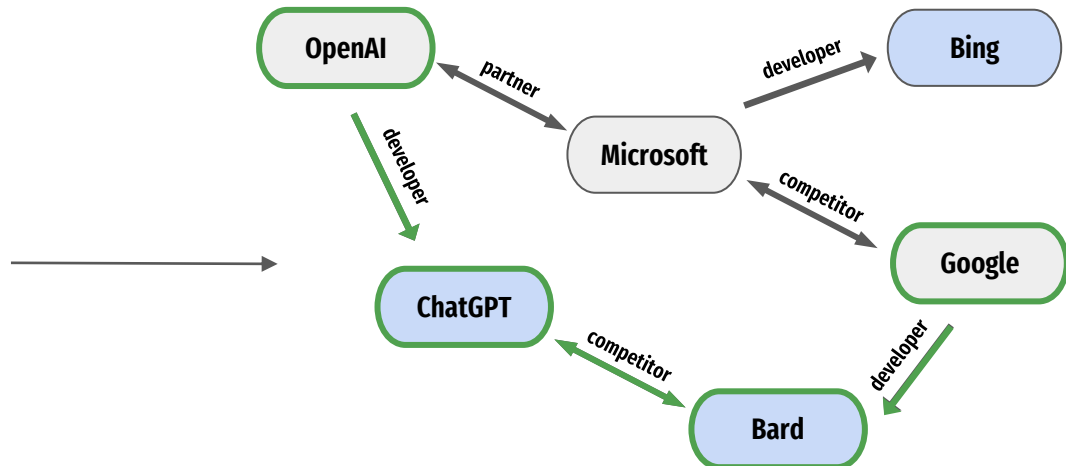


## Q: Who is OpenAI's competitor?

### Microsoft brings Bing chatbot to phones after curbing quirks

By MATT O'BRIEN February 22, 2023

"Powered by some of the same technology behind the popular writing tool ChatGPT, built by Microsoft partner OpenAI, the new Bing AI is part of an emerging class of AI systems that have mastered human language and grammar after ingesting a huge trove of books and online writings. The Microsoft Corp's Bing is one of the major Western products in the AI field. Google's working on it's response, Bard."



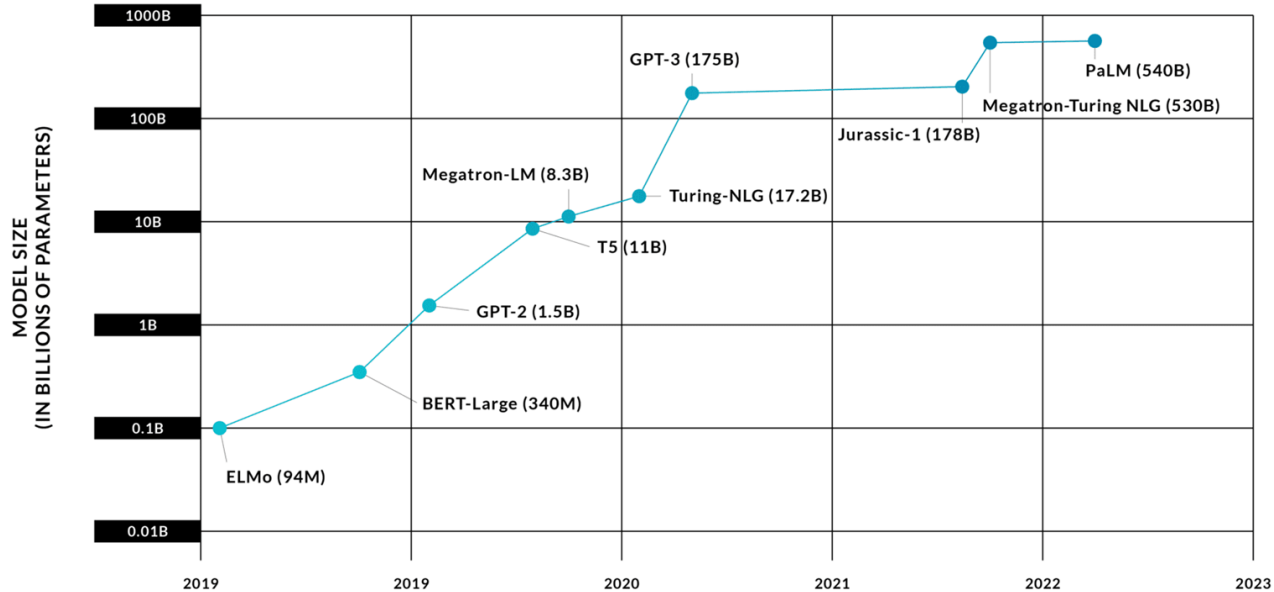
# Current Landscape in NLP

## Pre-trained LMs (PLM)

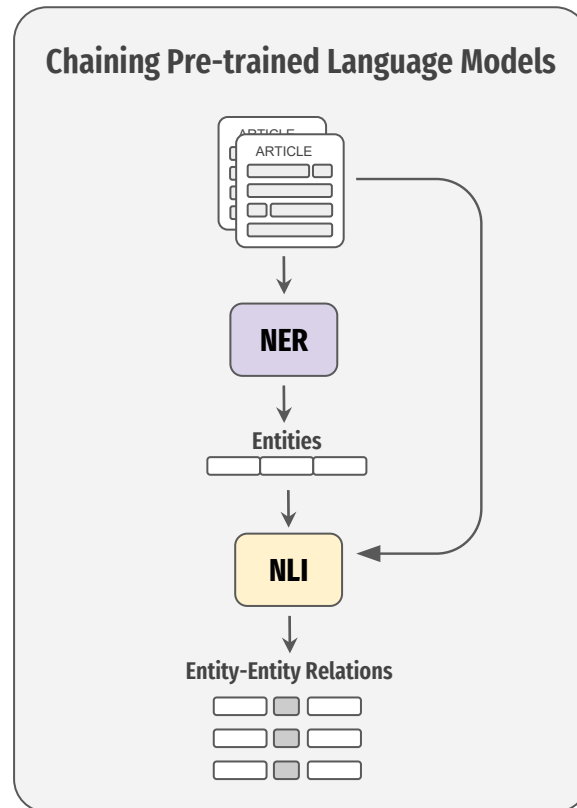
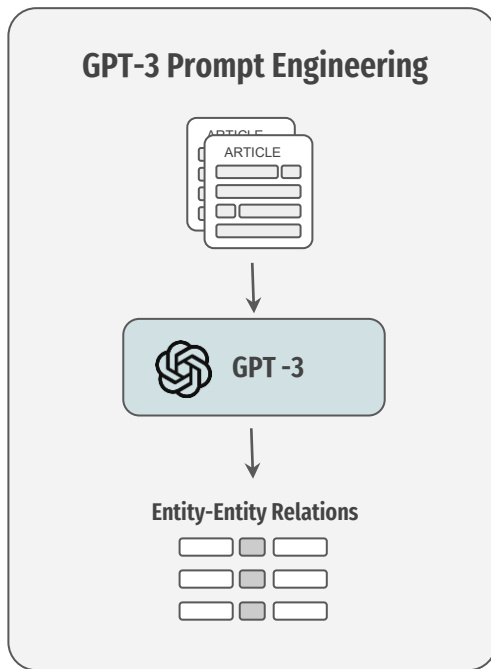
- Parameters: 350M - 2B
- Open-Source
- Fine tuning with additional training of params

## Large Language Models (LLM)

- Parameters: 150-500B
- Private & Open-Source
- Fine tuning with prompting

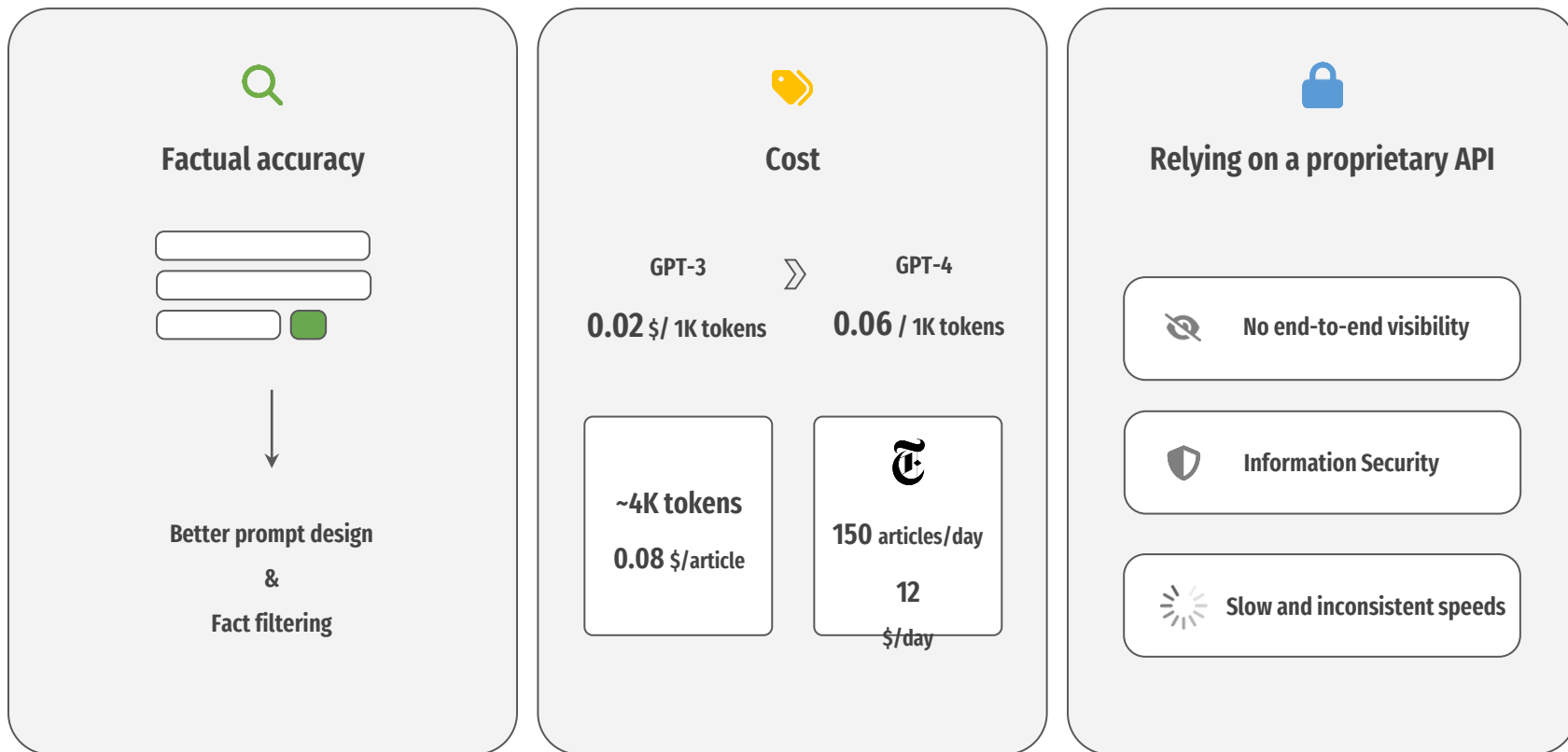


# Two Different Approaches



# Relationship Extraction with GPT-3

## Challenges with GPT variants



# Relationship Extraction with GPT-3 Prompt Engineering

## Target Article

### Chinese apps remove ChatGPT as global AI race heats up

By Jessie Young and Wayne Chang, CNN  
Published 3:41 AM EST, Thu February 23, 2023



The ChatGPT app displayed on a screen on February 3, 2023.

**Hong Kong (CNN)** — Several popular Chinese apps have removed access to ChatGPT, the artificial intelligence chatbot that has taken the world by storm even as major Chinese tech companies race to develop their own equivalent.

ChatGPT, developed by the American research lab OpenAI, is not officially available in China, but several apps on the Chinese social media platform WeChat had previously allowed access to the chatbot without the use of a VPN or foreign mobile number.

Those doors now appear shut. Earlier this week, the apps ChatGPTRobot and AIGC Chat Robot said their programs had been suspended due to "violation of relevant laws and regulations," without specifying which laws.

Two other apps, ChatGPTAI and Chat AI Conversation, said their ChatGPT services went offline due to "relevant business changes" and policy changes.

## Few-Shot Prompt

Each article below has an associated list of entities and a table with relationships. The table links the mentioned entities to other mentioned entities. The list has items of the form "entity (type)"...

...

Here are some examples:

x 3  
Other articles

Article:

Entity list = [...]  
|Entity|Relationship|Entity|Passage|  
|...|  
|end|

Article:

Entity list {OpenAI (ORG), ChatGPT (PROD), Microsoft (ORG), Google (ORG)}  
|Entity|Relationship|Entity|Passage|  
|OpenAI|developed|ChatGPT|"ChatGPT, developed by the American research ...  
|Microsoft|investor|OpenAI|"Microsoft plans to invest billions in the ...  
|Google|competitor|OpenAI|"Google announced it will soon roll out Bard, ...  
...

Input

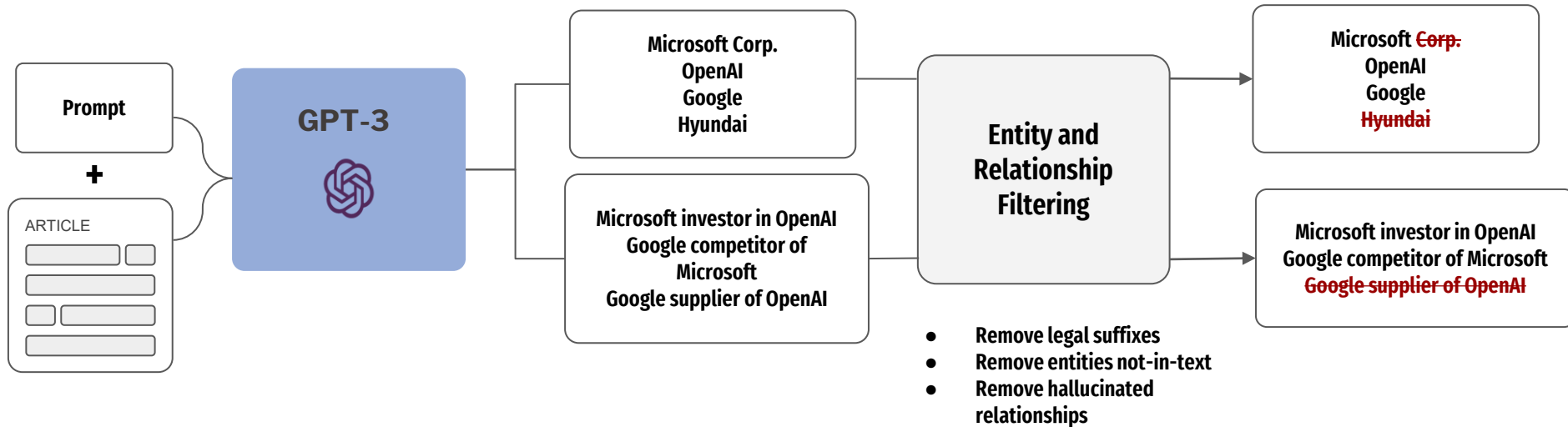


Output



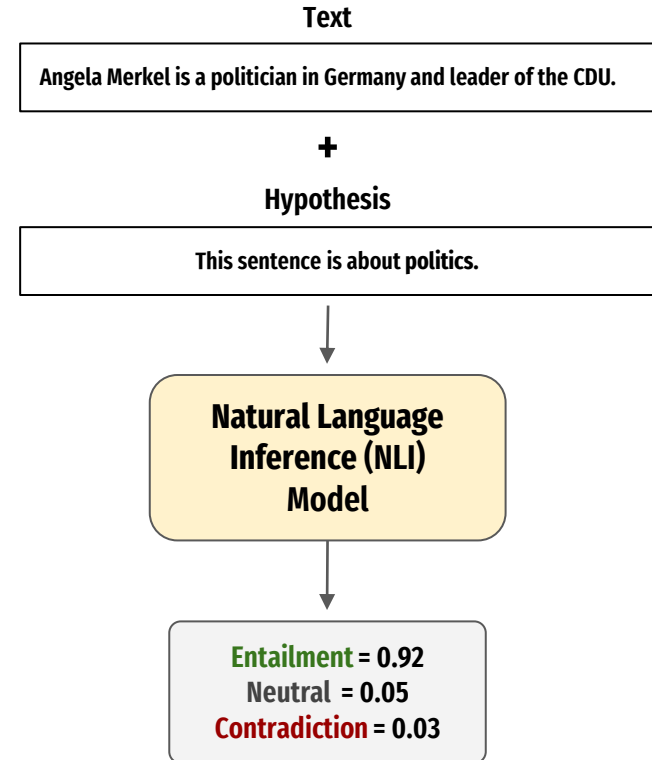
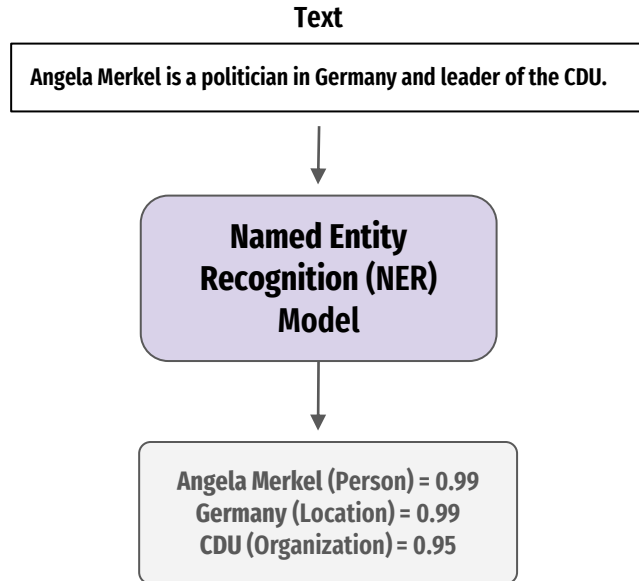
# Relationship Extraction with GPT-3

## Output Processing



# Relationship Extraction with PLMs

## NER and NLI: What are they?



# Relationship Extraction with PLMs

## Entity Extraction

### Article Paragraph

Powered by some of the same technology behind the popular writing tool ChatGPT, built by Microsoft partner OpenAI, the new Bing AI is part of an emerging class of AI systems that have mastered human language and grammar after ingesting a huge trove of books and online writings. The Microsoft Corp's Bing is one of the major Western products in the AI field.

# Relationship Extraction with PLMs

## Entity Extraction

**NER**  
Orgs and Products

Powered by some of the same technology behind the popular writing tool ChatGPT, built by Microsoft partner OpenAI, the new Bing AI is part of an emerging class of AI systems that have mastered human language and grammar after ingesting a huge trove of books and online writings. The Microsoft Corp's Bing is one of the major Western products in the AI field.

**Entity Resolution**

- [Text]. According to this, ChatGPT is a product.  
- [Text]. According to this, OpenAI is a company.  
- [Text]. According to this, Western is a company.  
...

The Microsoft  
Microsoft } Microsoft

Bing AI  
Bing } Bing

# Relationship Extraction with PLMs

## Entity Extraction

**NER**  
Orgs and Products

Powered by some of the same technology behind the popular writing tool ChatGPT, built by Microsoft partner OpenAI, the new Bing AI is part of an emerging class of AI systems that have mastered human language and grammar after ingesting a huge trove of books and online writings. The Microsoft Corp's Bing is one of the major Western products in the AI field.

**Entity Resolution**

The Microsoft  
Microsoft } Microsoft  
  
Bing AI  
Bing } Bing

**NLI Entity Filtering**

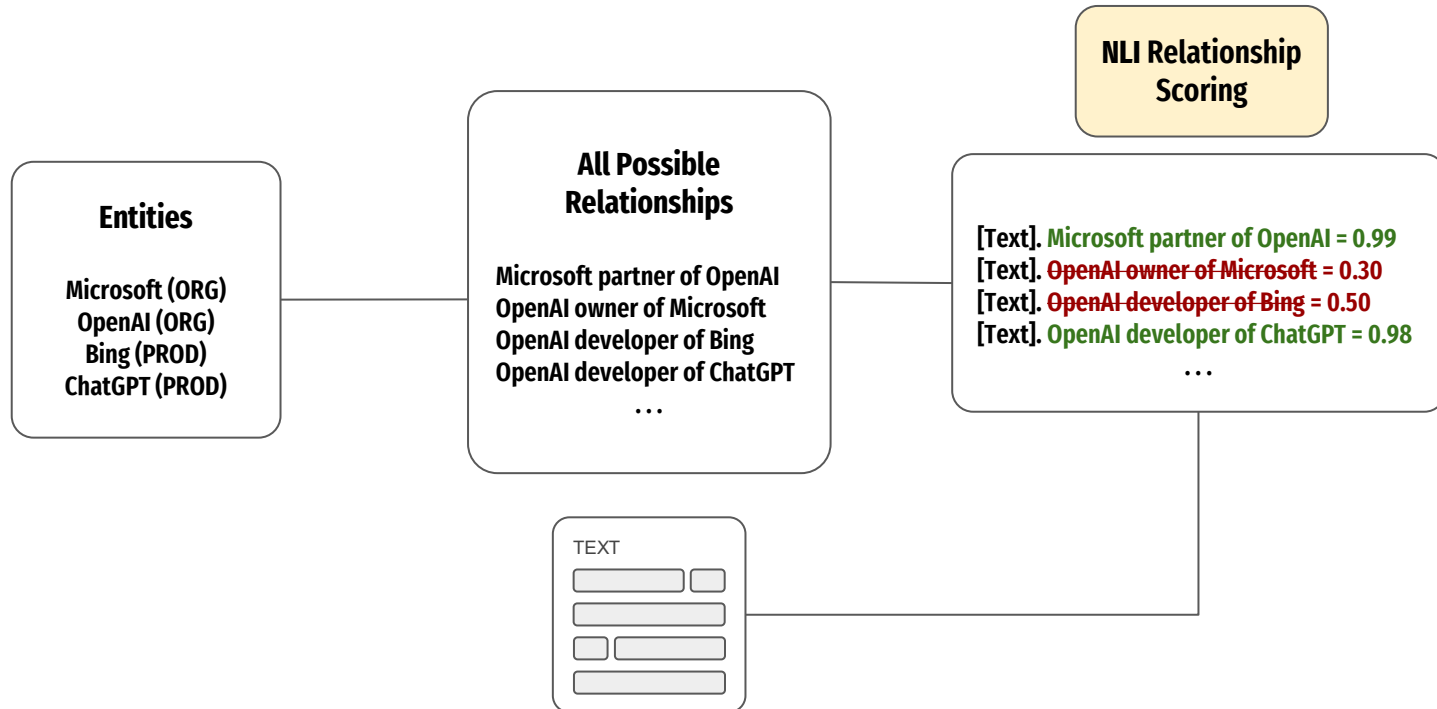
- [Text]. According to this, ChatGPT is a product. ✓
- [Text]. According to this, OpenAI is a company. ✓
- [Text]. According to this, Western is a company. ✗
- ...

**Entities**

Microsoft (ORG)  
OpenAI (ORG)  
Bing (PROD)  
ChatGPT (PROD)

# Relationship Extraction with PLMs

## Relationship Extraction

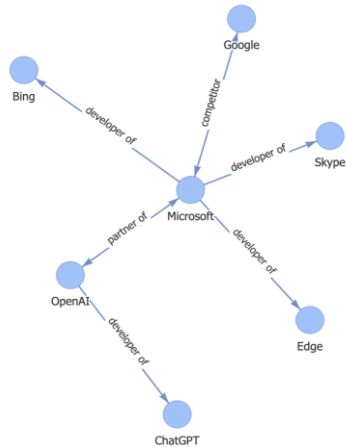


# Output Visualization

## Cross-Article Connections

### Microsoft brings Bing chatbot to phones after curbing quirks

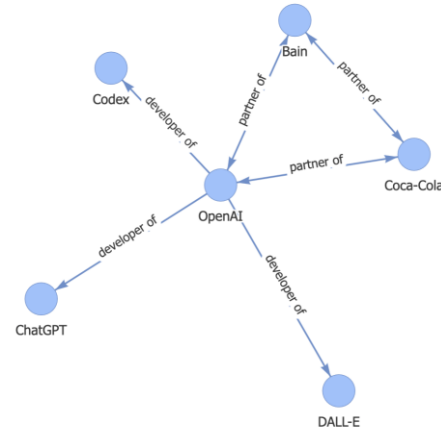
By HEAT GIBBER February 22, 2023



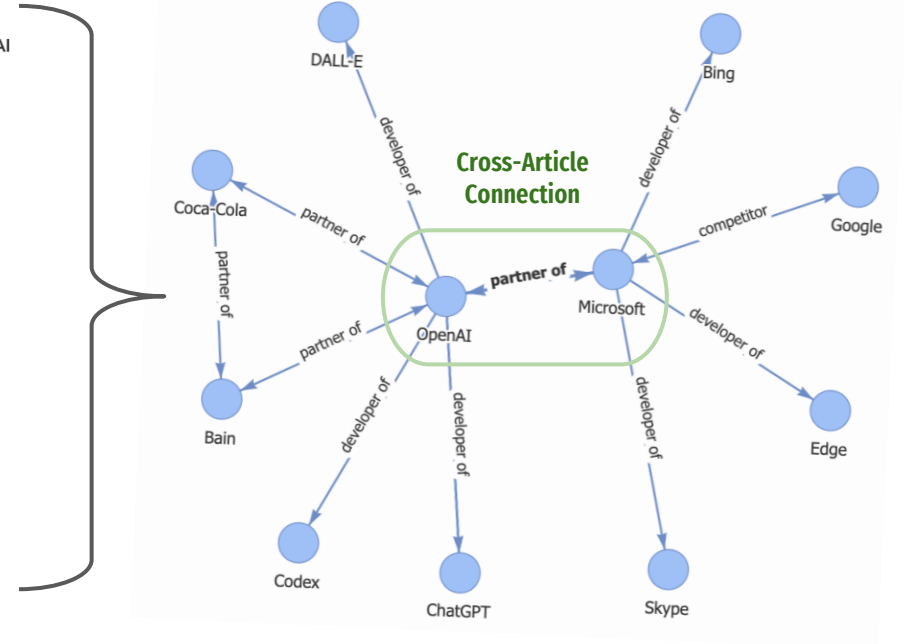
<https://apnews.com/article/technology-mobile-phones-business-software-522a2225391a968ee1d791e32d8b91eb>

### Bain & Company inks partnership with ChatGPT developer OpenAI

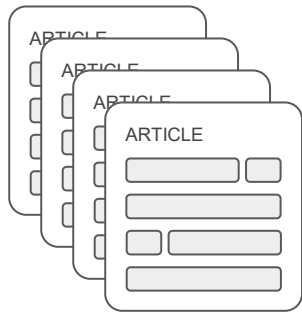
23 February 2023 | Consulting.us | 2 min. read



<https://www.consulting.us/news/8775/bain-company-inks-partnership-with-chatgpt-developer-openai>



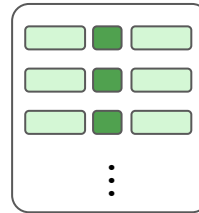
# Evaluation Method



Manual Annotation

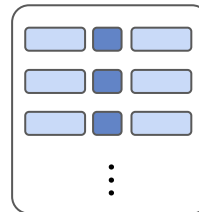
- 50+ news articles
- Reuters, APNews, CNBC, CNN, ...

“True” Relationships

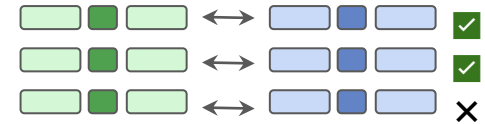


- 250+ relationships
- 200+ companies and products

Predicted Relationships



Compare





# Results



## Metrics

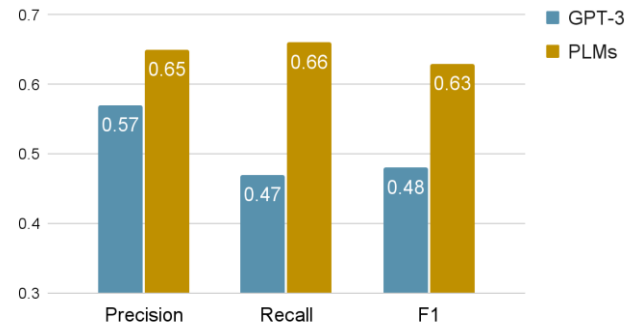
Averaged over Articles

**Precision** = portion of predictions that are correct

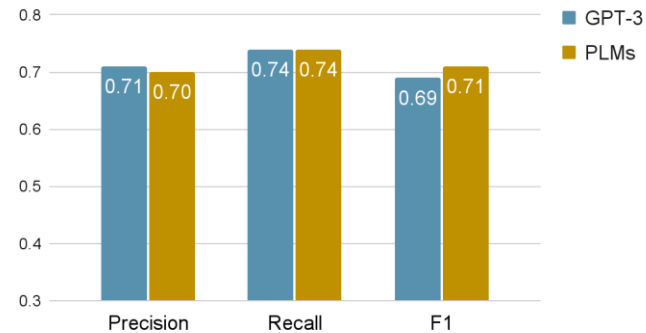
**Recall** = portion of true relationships detected by model

**F1** = overall performance

Relationship extraction



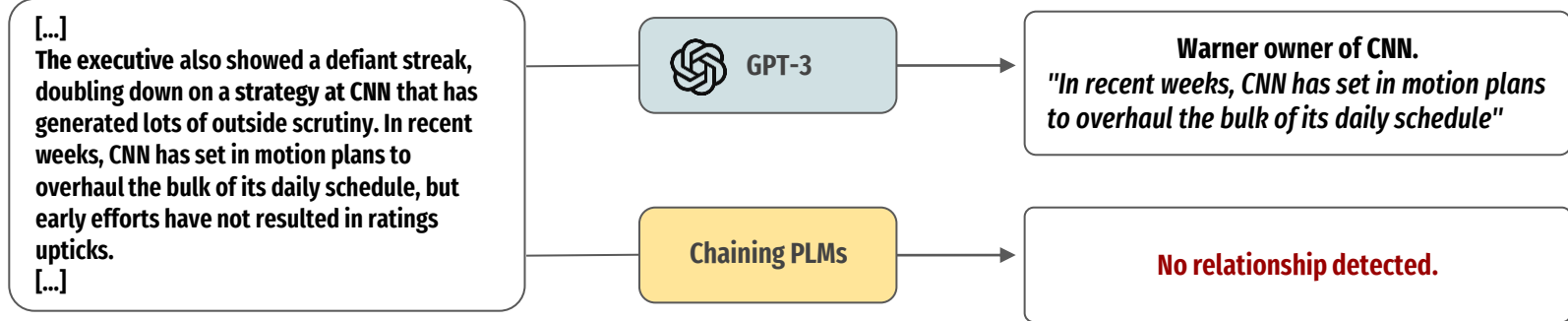
Entity Extraction



# Error Examples

## Warner Bros. Discovery's Zaslav Shows New Combative Streak, Tweaking Media Rivals

By Brian Steinberg ▾



GPT-3 captures long range relationships ✓  
 PLMs context is too short ✗

# Error Examples

**Amazon cuts 9,000 more jobs, bringing 2023 total to 27,000**

By HALELUYA HADERO March 20, 2023

[...]  
 Like other tech companies, including Facebook parent Meta and Google parent Alphabet, Amazon ramped up hiring during the pandemic to meet the demand from homebound Americans that were increasingly buying stuff online to keep themselves safe from the virus.  
 [...]

 GPT-3

Chaining PLMs

**Amazon competitor Meta**  
*"Like other tech companies, including Facebook parent Meta"*

**Alphabet competitor Amazon**  
*"Like other tech companies, including Google parent Alphabet"*

**Meta owner of Facebook**  
**Alphabet owner of Google**

GPT-3 changes original text ✗  
 PLMs capture fine-grained relationships ✓

# Comparison of Both Approaches

## GPT-3 vs PLMs



- + Provides passages from text
- + Task flexibility
- + Bigger context (whole article)
- + Better Natural Language Understanding
  
- Paid proprietary API
- Very slow (20 s/query)
- Very little control

### Chaining PLMs

- + Fine grained control
- + Best performance
- + Fast (<1 s/article)
  
- Requires tuning parameters
- Lesser task flexibility
- Small context (2-3 paragraphs)

# Next steps

## Ideal System

- Provides passage for each relationship
- Good Natural Language Understanding
- Big context
- Running on-premises
- Control over model

Solution



## Open Source LLMs!

Meta's LLaMa  
Stability AI's Stable LM  
TII's Falcon LLM

Question: Can these LLMs compete with GPT's performance?

**Thank you!**

→ **Named Entity Recognition**

- ◆ **xlm-roberta-large-finetuned-conll03-english**
- ◆ **100M parameters**

→ **Zero-Shot Classification**

- ◆ **MoritzLaurer/DeBERTa-v3-large-mnli-fever-anli-ling-wanli**
- ◆ **100M parameters**



**500x smaller**

**GPT-3**



**175B parameters**