

# CurrentAI

Zukang Yang, Jaipaul Mann, Clara Zhu

August 6, 2024



CurrentAI

# Intro to the Team



**Zukang Yang**  
Data Scientist @ Eonum Inc



**Clara Zhu**  
Data Engineer @ Wonderfi



**Jaipaul Mann**  
Data Scientist @ Clari

# Problem



How do we stay up-to-date in AI - the fastest developing area of technology?



## In short:

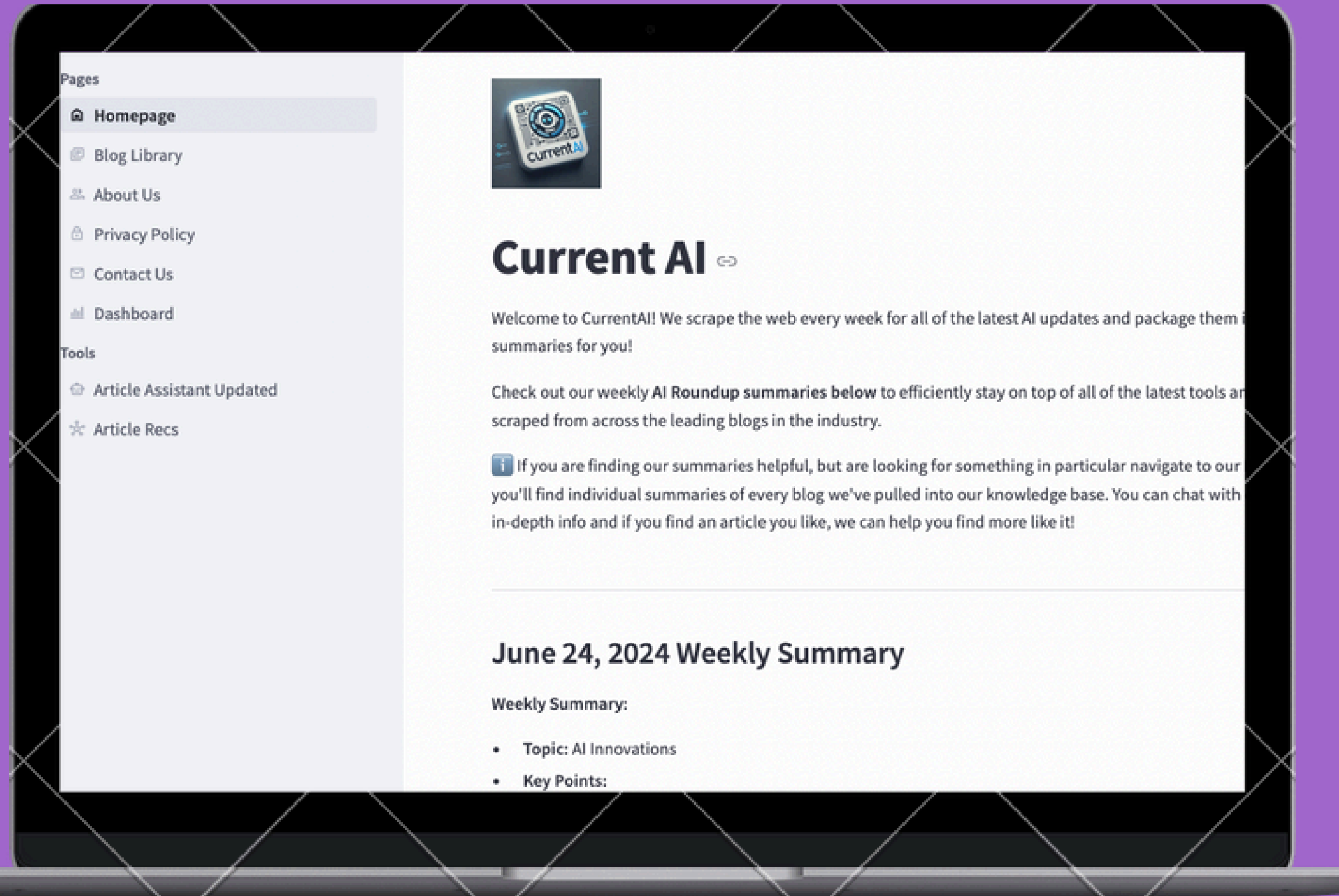
The **volume** and **velocity** of AI-related news and technology is too great for the average AI professionals & enthusiasts to stay current.



## What's needed:

Means for getting info on new tech, packages, companies, and trends that are **relevant** to busy professionals and packaged to **reduce cognitive load**

# MVP and Our Approach







# Demo

solving problems for our users



## Takeaways from Target Users

*“need a balance of high-level notes and the ability to double-click into articles for a deep dive if it seems interesting”*



## Key Questions

- What new tech has come out that I need to learn about?
- How does this tech work and how can I use it?



## CurrentAI Impact

proposed workflows start at the macro-level and then guides user into deep-diving into the articles that matter to them



# Key Features

solving problems for our users



## **Weekly AI Trends & Highlights**

Summarizes info from across the most recently posted blogs and gives users the convenience to tap into recent trends



## **Current Knowledge Base + LLM Chat**

Giving users the ability to consume a ton of information quickly, with a boost from LLM-powered summaries and chat



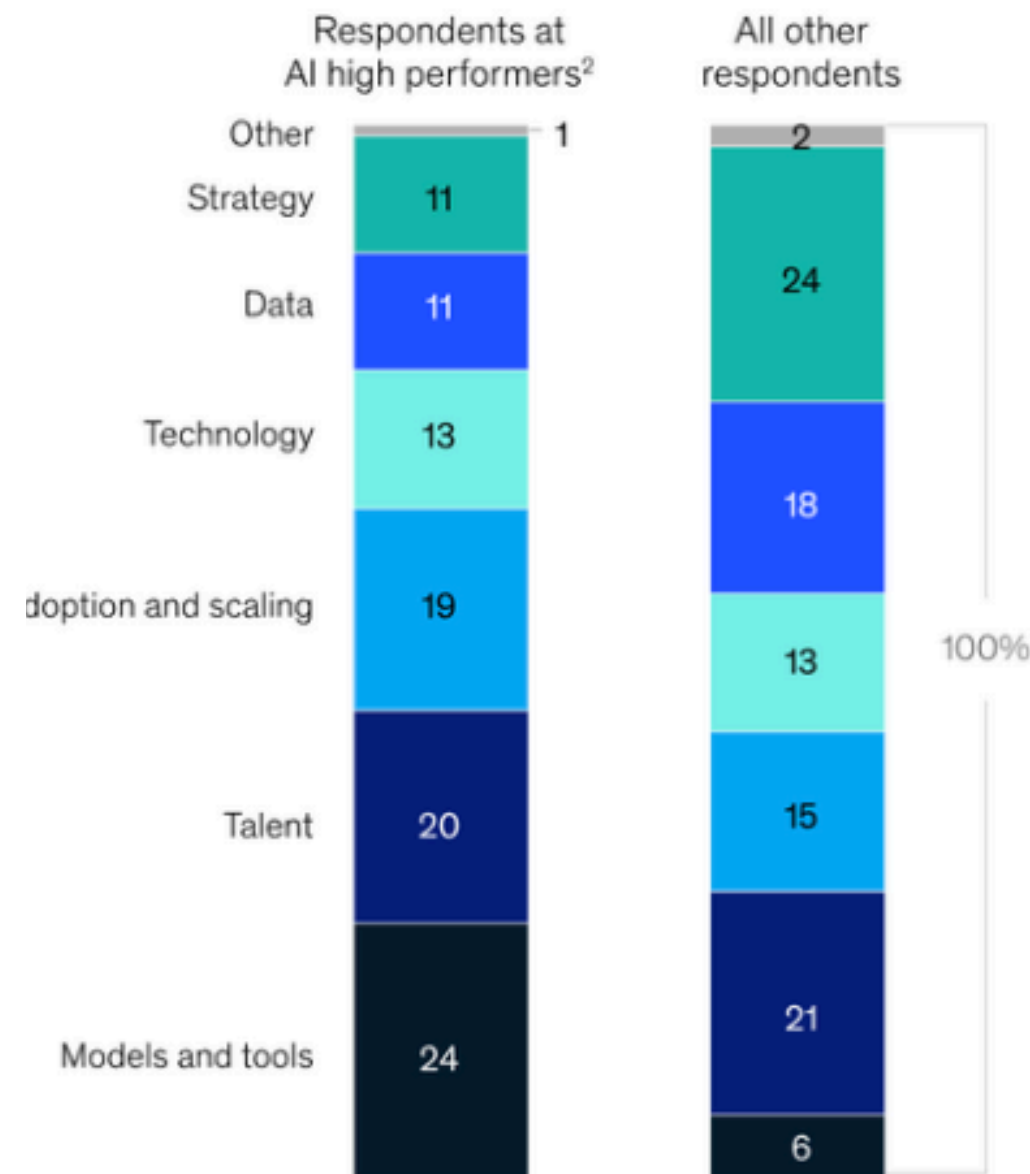
## **Article Recommendation**

Helping users find additional information, news, and articles relevant to what they care about

# Evidence in the Industry



Knowledge of contemporary **models & tools** poses the biggest AI-related challenge for AI professionals<sup>1</sup>



Industry trends require AI professionals to **swiftly evolve and deliver** new products upon new AI-tech releases

*“The **blistering pace of advancement** in generative AI leaves companies struggling to effectively implement and measure the technology”<sup>2</sup>*

Knowledge of contemporary Models & Tools, Talent, and Technology are all obstacles that CurrentAI has the **ability to impact**

<sup>1</sup> Pulled from *The state of AI in 2023: Generative AI's breakout year report* from McKinsey & Company

<sup>2</sup> Pulled from Maria Korolov's article *“For IT leaders, operationalized gen AI is still a moving target”* for CIO



# Value Across User Segments

high impact for a number of different personas



## Current AI Professionals

“Front-door” for researching new strategy, package, SDK, etc.

**Value:** saves hours of research leg-work



## Prospective Job Hunters

Study the technology and news specific to the companies you are targeting

**Value:** curated interview prep at scale

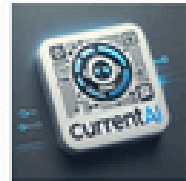


## AI Researchers and Enthusiasts

Summarize trends across the industry quickly, deep-dive into specifics as needed

**Value:** halve the time needed for industry overview abstracts





## Current AI [↗](#)

Welcome to CurrentAI! We scrape the web every week for all of the latest AI updates and package them into weekly summaries for you!

Check out our weekly AI Roundup summaries below to efficiently stay on top of all of the latest tools and news scraped from across the leading blogs in the industry.

**i** If you are finding our summaries helpful, but are looking for something in particular navigate to our search page where you'll find individual summaries of every blog we've pulled into our knowledge base. You can chat with our AI assistant for in-depth info and if you find an article you like, we can help you find more like it!

---

### June 24, 2024 Weekly Summary

Weekly Summary:

- Topic: AI Innovations
- Key Points:

# Technical Approach

---

# Data and Pipeline

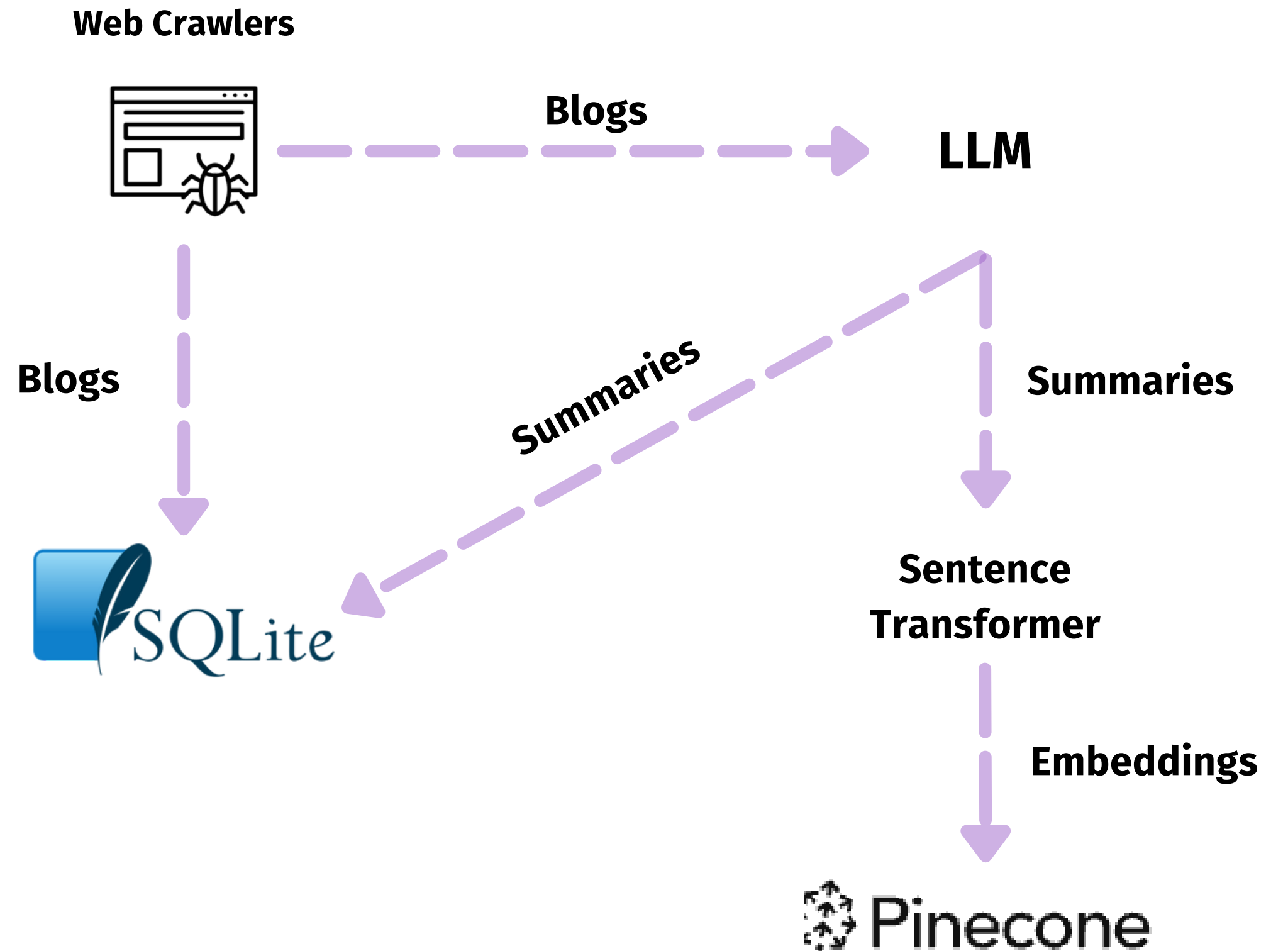
## ◆ Datasets



- Engineering blogs
- Article summaries
- Article embeddings

## ◆ Databases

- **SQLite** (Local)
- **Pinecone** (Cloud)



# Models and “Modeling”

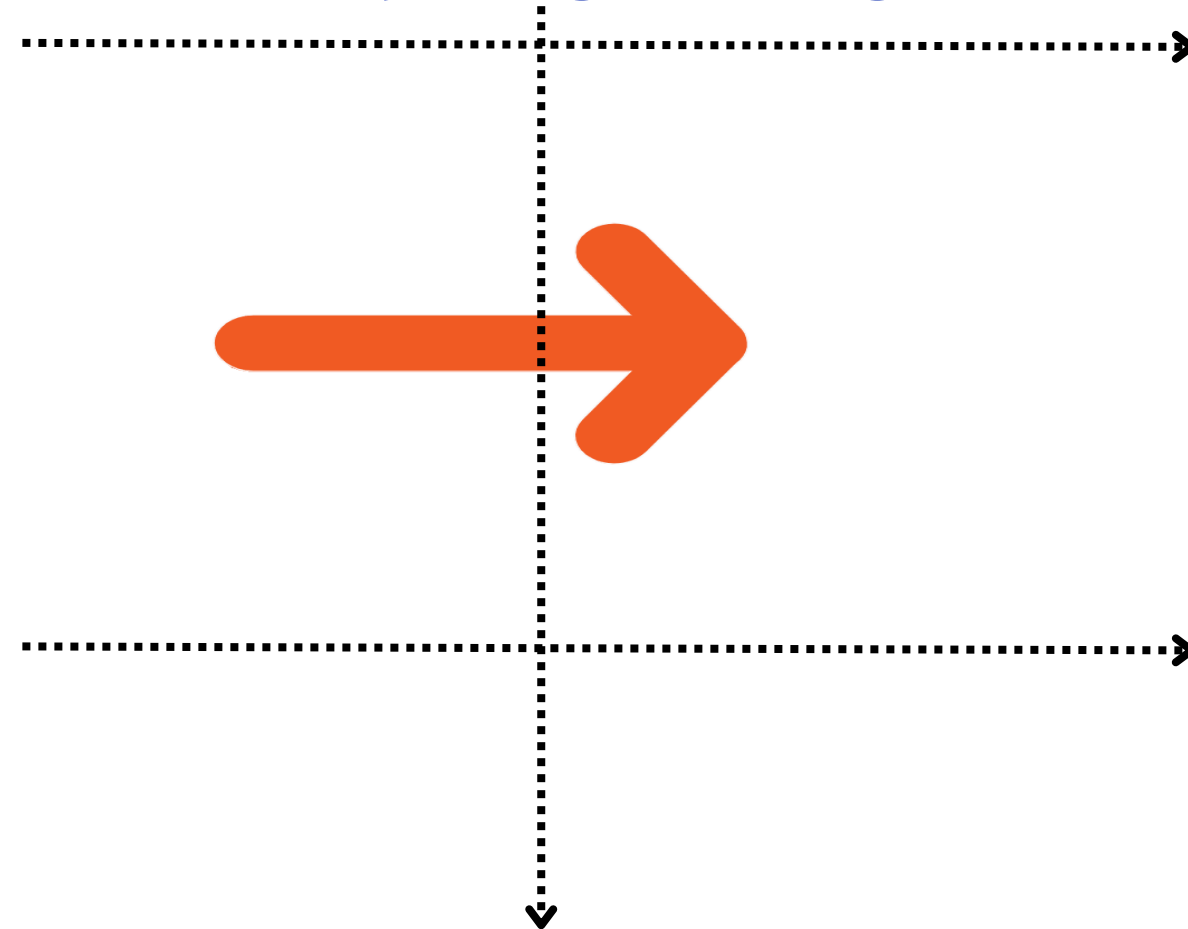
- A LLM

- Summary Generation
- Keyword Extraction
- Conversation

- A Sentence Encoder

- Recommendation based on similarity

Prompt Engineering



- **Claude Sonnet** (Batch)

- Summary Generation
- Keyword Extraction

- **Claude Haiku** (Real-time)

- Conversation

- **gte-large-en-v1.5** (Batch)

## Conversation Prompt:

You are a knowledgeable assistant. Below is the context and chat history.

Use the context to answer the user's question if the information is available.

If the context does not provide the information, rely on your general knowledge to answer.

# Technical Challenges



## Uncontrollable LLM Output

- **Challenge**
  - **Local LLM for Generation**
    - Repetitive
    - Incomplete
    - Irrelevant
    - Slow
- **Solution**
  - Proprietary LLM
  - Prompt engineering

## High Cost from AWS Service

- **Challenge**
  - **AWS Vector and Relational DB**
    - Expensive
    - Latency
- **Solution**
  - **Local SQLite DB**
  - **Pinecone**
    - Free in the current scale
    - Fast query speed

## Challenging Evaluation

- **Challenge**
  - No ground truth
- **Solution**
  - Invented “Tag Score”
  - User Feedback

# Technical Model Evaluation



## Our Approach

- ◆ Tag Score
- ◆ Similarity Score



## Evaluation Result

	Score	Examiner	Sample Size
Tag Score	83.44%	AI	50
Similarity Score	87.20%	AI	50



# Model Evaluation - User Feedback



## User Feedback Approach

- ◆ Chatbot quality evaluation
- ◆ User Feedback survey



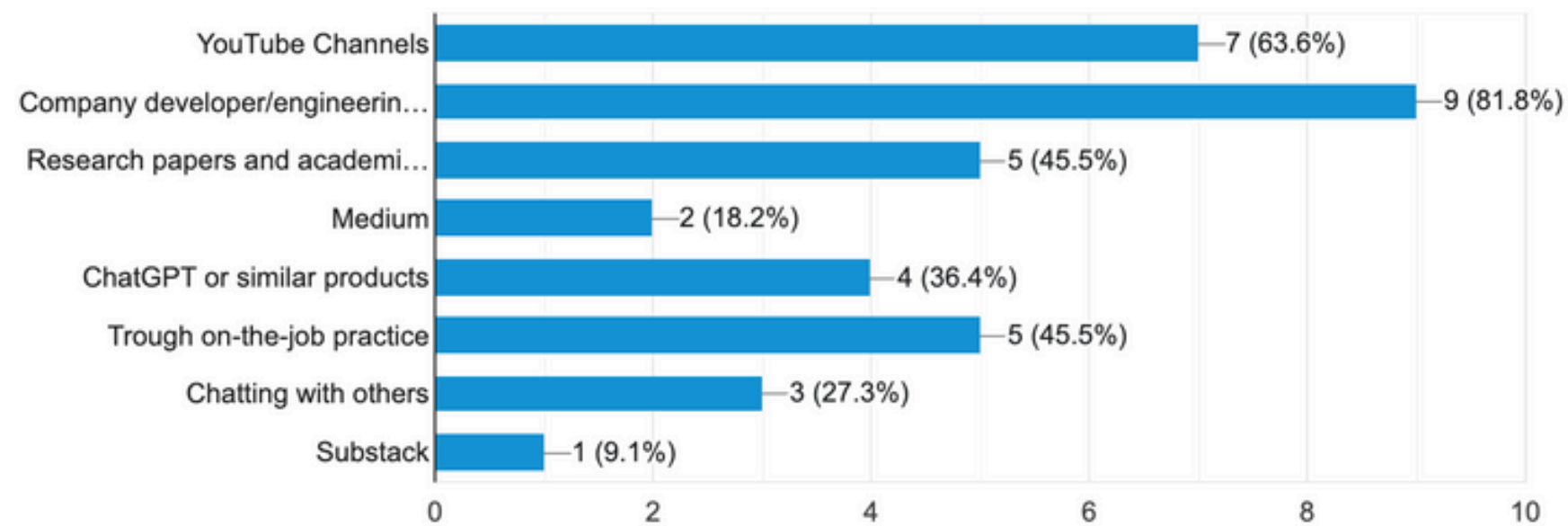
## Evaluation Result

	Score	Examiner	Sample Size
Accuracy	96%	CurrentAI Team	20
Helpfulness	92%	CurrentAI Team	20
Helpfulness	85%	Users	12

# Insights

## Top 3 AI sources

- Company developer/engineering blog
- YouTube Channels
- Research papers & on the job practice

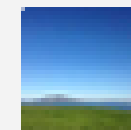
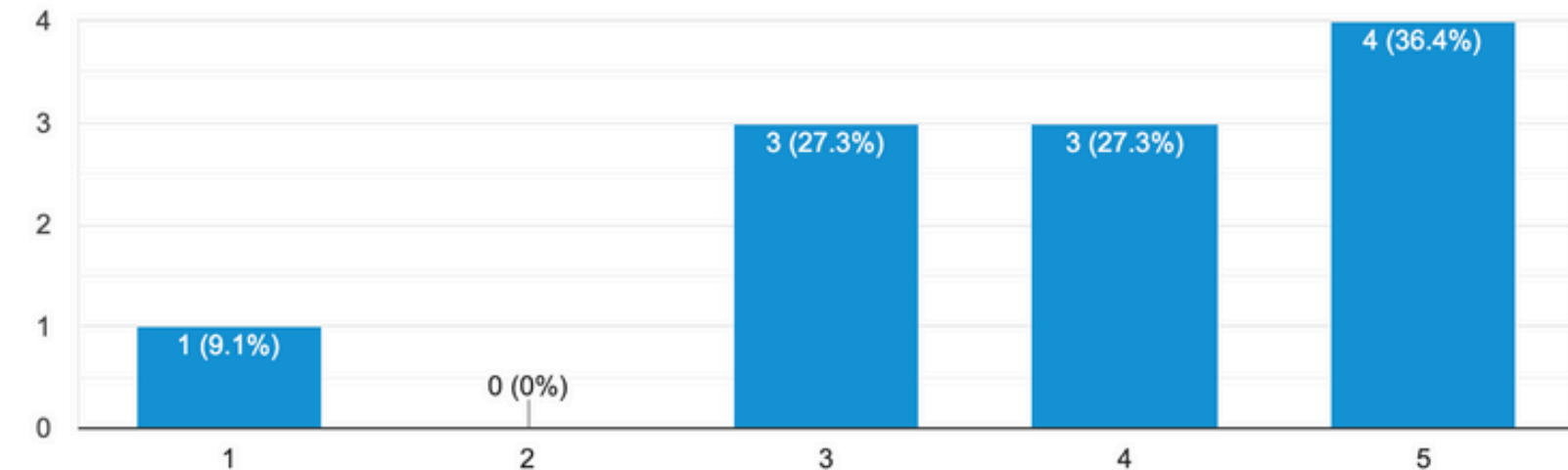


92% of users expressed their intention to reuse our tool

How likely are you to use our tool again?



11 responses



Praise, I compare your chatbot with chatgpt-4. I can see yours are better. I guess you have done some nice prompt engineering work in the backend.

# Roadmap - From MVP to Product

## ◆ Improved UI

- Based on many user feedback

## ◆ Provide email subscription

- Provide email newsletter subscription

## ◆ Improving the Chatbot

- Ensure the content generation is continuously improving based on user feedback loop



# CONCLUSION

Helping busy professionals stay informed about AI effortlessly