





Summary

About SentientAI	3	Future Data Collection	10
Manifesto	4	Future Dataset Overview	11
Ethics Guidelines	5	Parameters Diagram (LM Evolution Roadmap)	12
Communication Evolution Timeline	6	Use Case and Business Model	13
Problems with News Today	7	Next Solution: Prescription Recommendation	14
ChatGPT Limitations	8		
Introducing Newsboy	9		
1st Iteration Model Training	10		



About SentientAI

Bridging the gap between human empathy and machine thinking.

We are a team of passionate disruptors around the world interested in building a better, more inspiring and empathetic future through the ethical use of AI technology.

Our goal is to develop solutions that solve meta-crisis problems through the use of disruptive technologies, creative thinking and sustainable ethos.

SentientAI is an initiative from Sentient LTDA. A Brazilian creative marketing agency established in 2017.

sentient 

[Learn More](#)

Design & Creative Direction

Thiago Patriota 

Full-Stack

Vasiliki Doropoulou 

Vaishnavi Narayanan 

Front-End

Saira Arif 

Ali Raza 

Machine Learning

Jovan Botic 



Manifesto

Once upon a time, humanity only communicated through archaic languages. Information, myths and tales were expressed with paintings on a wall and stories were told around pit fires through long lost ancestral idioms.

Writing was invented! History, ideas, opinions, ethics, morals and all forms of culture started being not only spread in a much faster pace, but also documented for future generations. The knowledge from our ancestors was preserved and they gave us the necessary way to further push science and innovate.

Machines were invented, our time and dreams were too valuable for us to spend it on repetitive and manual tasks so we industrialized production and started trailblazing our way towards a future filled with more food and tools, making it possible for us to focus on solving new problems.

Then we created languages for machines and found ourselves in the information age, where the whole globe and all nations met and shared information through what we know as internet protocol. We didn't just push the boundaries, we created a new digital dimension where everyone would co-live and freely communicate bi-directionally.

Fast forward, we are in the present, opening the doors of the AI native economy where machines not only talk back to us but they are able to interpret our words and provide us with exactly what we need. As our creation, we need to recognize that AI is here to be a perfect extension of our creative and technical skills and it is up to us to make the best possible use of it. **Do you feel ready?**



Ethics Guidelines

When fire was first discovered, people soon realized that it could be utilized not only to protect and defend but also to attack and subdue. As the first technology that kickstarted all the others, fire was extremely powerful because it represented the core natural concept of survival.

Technology will always be the necessary tool for society to find new horizons and explore new possibilities but it should always be used with caution and empathy for a greater future.

Fire has been rediscovered and it is called AI! If new approaches need to be invented for a new solution that was found, what should be the core foundational principles for innovation?

The Sentient Principles for an Ethical Approach to AI

1 - AI Should never harm humans in any way.

2 - AI Should never deceive.

3 - AI Should not be toxic or exclusive

4 - AI Should not be used for malicious purposes

5 - AI Should not infringe the copyright of humans

6 - AI content generators should be labeled as such

7 - AI Should not exploit user data and privacy

8 - AI should not promote or reinforce stereotypes

9 - AI Should not advise users to perform illegal actions

10 - Organizations using AI should take responsibility for its outcomes and impacts.



Communication Evolution Timeline



**30.000
Years Ago**

Cave dwellers used crude paints to leave behind the graphic representations of their lives



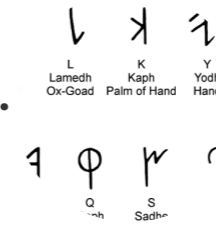
**2900
B.C**

Ancient Egyptians were the first society to use "pages" by weaving together stems of a **papyrus plant**



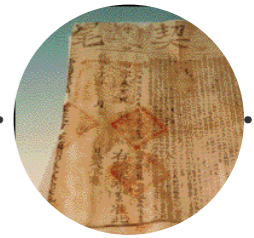
**2100
B.C**

The Epic of Gilgamesh was written on 12 clay tablets. An epic poem from ancient Mesopotamia considered the oldest written story on Earth.



**1050
B.C**

The Phoenicians develop an alphabet to make their bookkeeping easier.



**105
A.D**

Paper was first made in Lei-Yang, China by Ts'ai Lun, a Chinese court official.



**1947
A.D**

Full-scale commercial television is first broadcast.



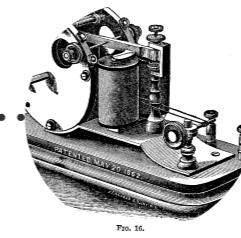
**1920
A.D**

Under the call sign KDKA, Pittsburgh's Westinghouse Electric and Manufacturing Company transmitted the first scheduled radio broadcast



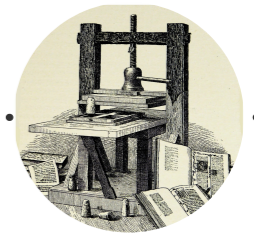
**1876
A.D**

Bell was granted US patent 174465A, for a method of transmitting speech by telegraphy—the telephone.



**1832
A.D**

Samuel F.B. Morse developed an electric telegraph and later on, the morse code



**1440
A.D**

Johannes Guttenberg invented the printing press and industrialized communication



**1965
A.D**

The very first version of what would become known as email was *invented in 1965* at Massachusetts Institute of Technology (MIT)



**1969
A.D**

The first hosts of ARPANET, Internet's ancestor, are connected. Giving birth to the TCP/IP which today we know as IP - Internet Protocol



**1992
A.D**

The very first version of what would become known as email was *invented in 1965* at Massachusetts Institute of Technology (MIT)



**2000-2015
A.D**

Skype, Facebook, Youtube, Reddit, Twitter, Iphone, WhatsApp, Instagram, Snapchat and Discord are launched



**2022
A.D**

ChatGPT went viral when OpenAI released the text-based artificial intelligence tool

Fast Forward



Problems with the News Today

Misinformation

A poll by USA TODAY/CNN/Gallup found **only 36 percent of Americans believe news organizations get the facts straight**, compared with 54 percent in mid-1989.

Sensationalism

In a study by the American Society of Newspaper Editors, eighty percent of the American public said they believe "**journalists chase sensational stories** because they think it will sell papers, not because they think it is important news. "

Poor Coverage

The **reporting on national affairs by the major news magazines has declined by 25 percent**, while the number of **entertainment and celebrity stories has doubled**, according to "The State of the News Media in 2004" report by the non-partisan Project for Excellence in Journalism."

*There isn't a tool on the market today that delivers **relevant news in an easy and intuitive way.***





ChatGPT limitations when it comes to News Reporting



Currently, ChatGPT doesn't support the generation of existing news headlines, only the generation of fictitious news headlines.

This essentially means that no AI solution has yet been properly developed in order to digest and process news through language models.

why you don't provide recent news?

As a language model developed by OpenAI, my training data only goes up until 2021 and I don't have real-time access to news updates. I can provide general information and answers to questions based on the text I was trained on, but I cannot provide up-to-date news headlines or current events.  

you are not allowed to access rss feeds?

As an AI language model developed by OpenAI, I don't have the ability to access real-time news updates or RSS feeds. My responses are generated based on the text I was trained on and do not reflect current events or news updates.  



Introducing Newsboy

Newsboy

Home Newsboy Sentient Ethics Manifesto Contact



Q Generate a News Headline

Search

Results: 12 articles found

Do You Live in a 'Food Swamp'? It Could Be Raising Your Heart Risk
Category: Health News Date: 02-02-2023 22:31:17

Newsboy contains a **rich and wide repository of various news articles** from various sources that **users can search through using natural language**. By analyzing the **sentiment and trustworthiness**, Newsboy delivers news that is catered to each and every user.

1

User types a **phrase into the search bar** containing the **news category, keywords, sentiment, and date** to find a news article

2

Newsboy **parses through** its repository of articles and runs ML models to **match the user-entered criteria**

3

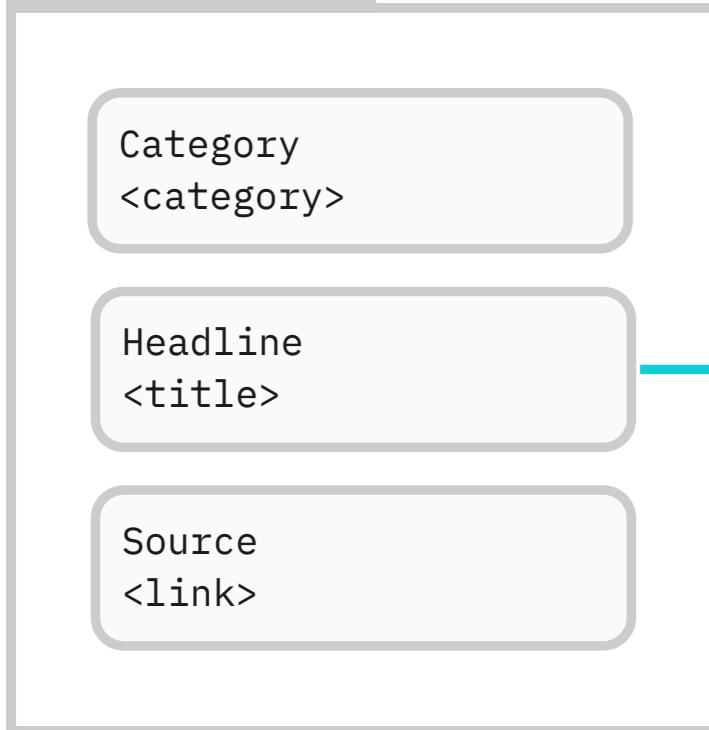
Newsboy returns the **relevant matching articles** after analyzing the **sentiment and trustworthiness**



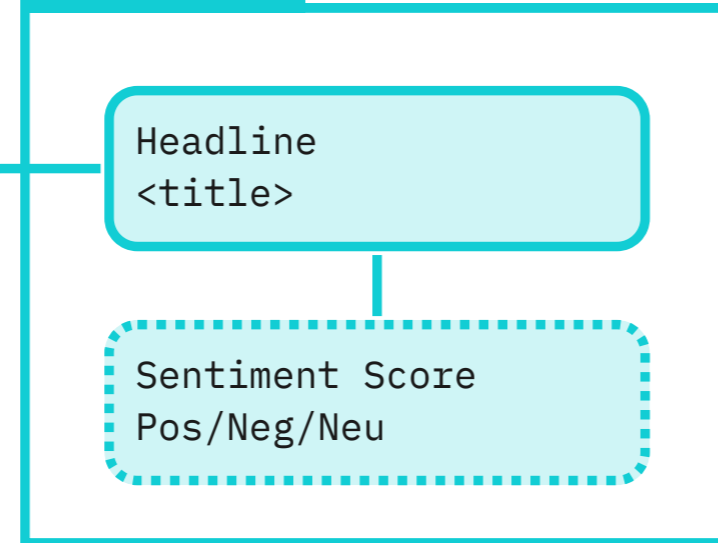
1st Iteration Model Training

Considering that data comes labeled in categories from the RSS feeds, there was no need to train our LM to identify them, so we started by training our LM to improve the search mechanism of news by providing a manual training set with headlines + sentiment scores so that we could help the model understand what can be considered a negative news and a positive news.

Core Static Data From RSS Feed

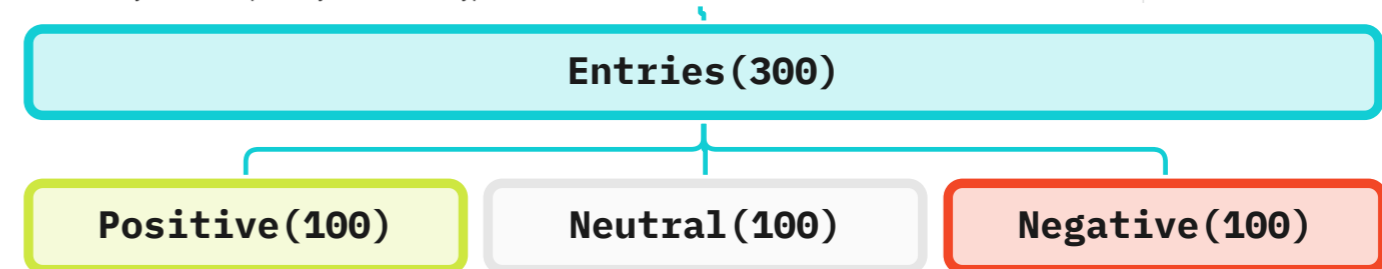


Training Set Data in LM



Data Sample

Headline	Sentiment
With new funding, Atomic AI envisions RNA as the next frontier in drug discovery	positive
Google to make changes to Android business terms in India after antitrust blow	negative
Waymo lays off staff as Alphabet announces 12,000 job cuts	negative
Acura will sell its EVs exclusively online starting in 2024 Engadget	neutral
Microsoft announces \$52.7 billion in Q2 revenue amid plans to layoff 10,000 workers Engadget	negative
With Nvidia Eye Contact, you'll never look away from a camera again	positive
US Marines Defeat DARPA Robot by Hiding Under a Cardboard Box	positive
Microsoft cloud outage hits users around the world	negative
Internet Archive takes down upload of BBC's Modi documentary	negative
Amazon Workers Set for Unprecedented UK Strike on Wednesday	negative
Researchers look a dinosaur in its remarkably preserved face	positive
UK Looks to Traditional Finance to Inform Crypto Regulations	neutral
Can Privacy and Transparency Coexist in Crypto?	neutral



Positive Bias/Positive Events/Positive Feelings.

Example Keywords: Sustainability, Privacy, Regulations, Growth, Innovation...

Negative Bias/Negative Events/Negative Feelings

Example Keywords: Crisis, Layoffs, strike, conflict, breach...



Future Dataset Collection

We plan to collect data in the future using RSS (Really Simple Syndication) Feeds that are provided by news outlets in order to distribute content in real time.

- RSS Feeds are separated by different channels for each category of information
- RSS Feeds have a <channel> + <item> structure but they differ in labels and types of available metadata.

We will plan to collect data through the various feeds and label them for our model training.

WIRED

<https://www.wired.com/feed/tag/ai/latest/rss>

```
<channel>
  <title>Feed: Artificial Intelligence Latest</title>
  <description>Channel Description</description>
  <link>https://www.wired.com</link>
  <atom:link href="https://www.wired.com/feed/tag/ai/latest/rss" rel="self" type="application/atom+xml"/>
  <copyright>© Condé Nast 2023</copyright>
  <language>en-US</language>
  <lastBuildDate>Wed, 01 Feb 2023 12:25:58 +0000</lastBuildDate>
  <item>
    <title>An AI Filter Revealed My Secret Self</title>
    <link>https://www.wired.com/story/ai-filter-beauty-tiktok/</link>
    <guid isPermaLink="false">63d14f37e86be50408c465b3</guid>
    <pubDate>Tue, 31 Jan 2023 14:00:00 +0000</pubDate>
    <media:content/>
    <description>Yes, they're problematic. They can also reflect an alternate reality that feels eerily close to the truth.</description>
    <category>Ideas</category>
    <media:keywords>art, Instagram, artificial intelligence, internet culture, Film</media:keywords>
    <dc:creator>Thao Thai</dc:creator>
    <dc:publisher>Condé Nast</dc:publisher>
    <dc:subject>Ideas</dc:subject>
    <media:thumbnail url="https://media.wired.com/photos/63d85f89de734741c2ad3d0e/master/pass/ideas-ai-filters-self.jpg" width="2400" height="1350"/>
  </item>
```

The New York Times

<https://rss.nytimes.com/services/xml/rss/nyt/World.xml>

```
<channel>
  <title>NYT > World News</title>
  <link>https://www.nytimes.com/section/world</link>
  <atom:link href="https://rss.nytimes.com/services/xml/rss/nyt/World.xml" rel="self" type="application/rss+xml"/>
  <description/>
  <language>en-us</language>
  <copyright>Copyright 2023 The New York Times Company</copyright>
  <lastBuildDate>Wed, 01 Feb 2023 12:23:49 +0000</lastBuildDate>
  <pubDate>Wed, 01 Feb 2023 11:50:14 +0000</pubDate>
  <image>
    <title>NYT > World News</title>
    <url>https://static01.nyt.com/images/misc/NYT_logo_rss_250x40.png</url>
    <link>https://www.nytimes.com/section/world</link>
  </image>
  <item>
    <title>Heavier Russian Attacks Suggest Opening Moves of New Offensive, Ukraine Says</title>
    <link>https://www.nytimes.com/live/2023/02/01/world/russia-ukraine-news</link>
    <guid isPermaLink="true">https://www.nytimes.com/live/2023/02/01/world/russia-ukraine-news</guid>
    <atom:link href="https://www.nytimes.com/live/2023/02/01/world/russia-ukraine-news" rel="standout"/>
    <description>A surge in Russian artillery barrages and a buildup of troops could be the beginning of a new push by Moscow in the Donbas region of eastern Ukraine, President Zelensky says.</description>
    <dc:creator>The New York Times</dc:creator>
    <pubDate>Wed, 01 Feb 2023 11:40:30 +0000</pubDate>
    <media:content height="151" medium="image" url="https://static01.nyt.com/images/2023/02/01/multimedia/01ukraine-briefing-header-2am-gtzb/01ukraine-briefing-header-2am-gtzb-moth.jpg" width="151"/>
    <media:credit>Ivor Prickett for The New York Times</media:credit>
    <media:description>Civilians boarding an evacuation train in Kherson, Ukraine, as Russian forces continue to shell the city.</media:description>
  </item>
```



Future Dataset Overview

News metadata describes the content of a story and is created and packaged with the story. It is most often retrieved and presented with the story. **Metadata enables accurate and efficient search and discovery of the relevant news content.**

Primary Datasets

RSS Feed Example

Category <category>	Headline <title>	Excerpt <description>	Source <link>	Timestamp <pubdate>
<code><category>Ideas</category></code>	<code><title>Doctors, Get Ready for Your AI Assistants</title></code>	<code><description>Hospitals have begun using machine learning to help analyze and collect images, and the medical applications are endless.</description></code>	<code><link>https://www.wired.com/story/doctors-artificial-intelligence-medicine/</link></code>	<code><pubDate>Sat, 28 Jan 2023 12:00:00 +0000</pubDate></code>

Secondary Datasets

RSS Feed Example

Media <media:thumbnail>	Media Credits <media:credit>	Media Description <media:description>	Author <creator>	Keywords <keywords>
<code><media:thumbnail url="https://media.wired.com/photos/63ceb6f404fdce858b27a139/master/pass/05_As-AI-health-assistant-at-every-level.jpg" width="2400" height="1600"/></code>	<code><media:credit>Alex Plavevski/EPA, via Shutterstock</media:credit></code>	<code><media:description>A closed PCR testing booth sits idle in Shanghai.</media:description></code>	<code><dc:creator>Eric Topol</dc:creator></code>	<code><media:keywords>The WIRED World in 2023, health, healthcare, medicine, artificial intelligence</media:keywords></code>



Parameters Diagram (LM Evolution Roadmap)

Training Level 1

Prompt Example

"Search for an exciting technology news"

SENTIMENT

CATEGORY

Headline
<title>

Sentiment Score
Pos/Neg/Neu

Classification Values

Training Level 2

Prompt Example

"Search for an exciting technology news about the carbon market"

SENTIMENT

CATEGORY

Keywords

Headline
<title>

Excerpt
<description>

Categories
10 Unique Values

Classification Values

Training Level 3

Prompt Example

Search for an exciting technology news about the carbon market in USA

SENTIMENT

CATEGORY

Keywords

Location

Source
<link>

Excerpt
<description>

Location
10 Unique Values
(Countries)

Classification Values

Training Level 4

Prompt Example

Search for an exciting technology news about the carbon market in USA from yesterday

SENTIMENT

CATEGORY

Keywords

Location

Time

Timestamp
<pubdate>

Prompt
relative or absolute



Use Case & Business Opportunity

Newsboy can evolve to be like ChatGPT, offering a freemium model to grow its user base and strive for product-led growth and user-centered design.

Target Audiences

- Citizens
- Students (Academic)
- Journalists (PR & Media)
- Scientists (Research)
- Copywriters (Marketing)

Freemium Model

Free Tier
Ad-supported
revenue

Currently, ChatGPT doesn't support the generation of existing news headlines, only the generation of fictitious news headlines.

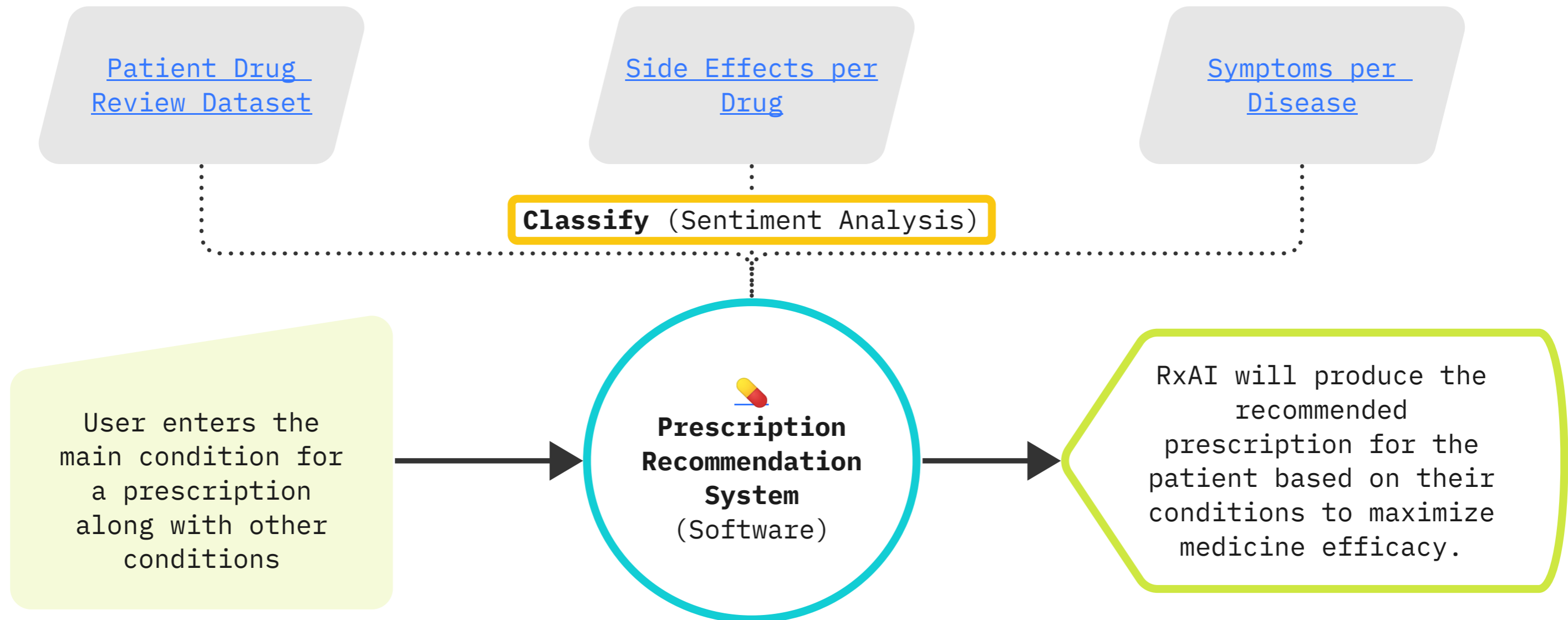
Paid Tier (\$15/mo)
Used as a research
database

This essentially means that no AI solution has yet been developed in order to digest and process news through language models.



Next Solution: Prescription Recommendation

Using the following datasets, we will train a model on the sentiment of drug reviews and conditions to recommend the best prescription for a patient given their conditions.





Thank You!