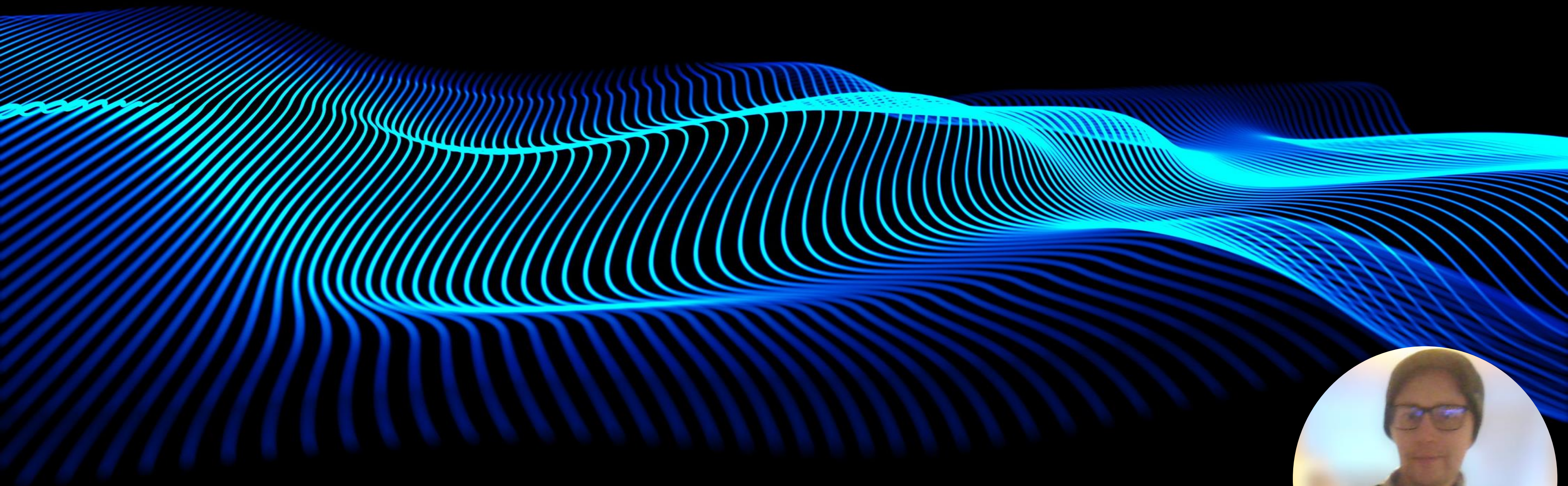


ArcBlok



Background



Disinformation, propaganda and conspiracy theories

Disinformation is a growing problem in today's digital age.

Social media platforms have made it easier than ever to spread misleading information.

From election interference to the spread of conspiracy theories, disinformation is a threat to our democracy, our society, and even the public health.

In response to this problem, we have created an AI product to fight disinformation.



Our problem statement



Disinformation poses a threat to the integrity of public discourse and our democracies. There is a pressing need to develop effective tools for detecting, analyzing, and countering disinformation.

Artificial intelligence is a powerful tool in the fight against disinformation by enabling more efficient and accurate identification of false information and malicious actors.

By analyzing large amounts of data and detecting patterns, AI can quickly identify sources of disinformation. This allows for a more proactive approach to combating disinformation.



How my product
can solve the
problem and how
it works



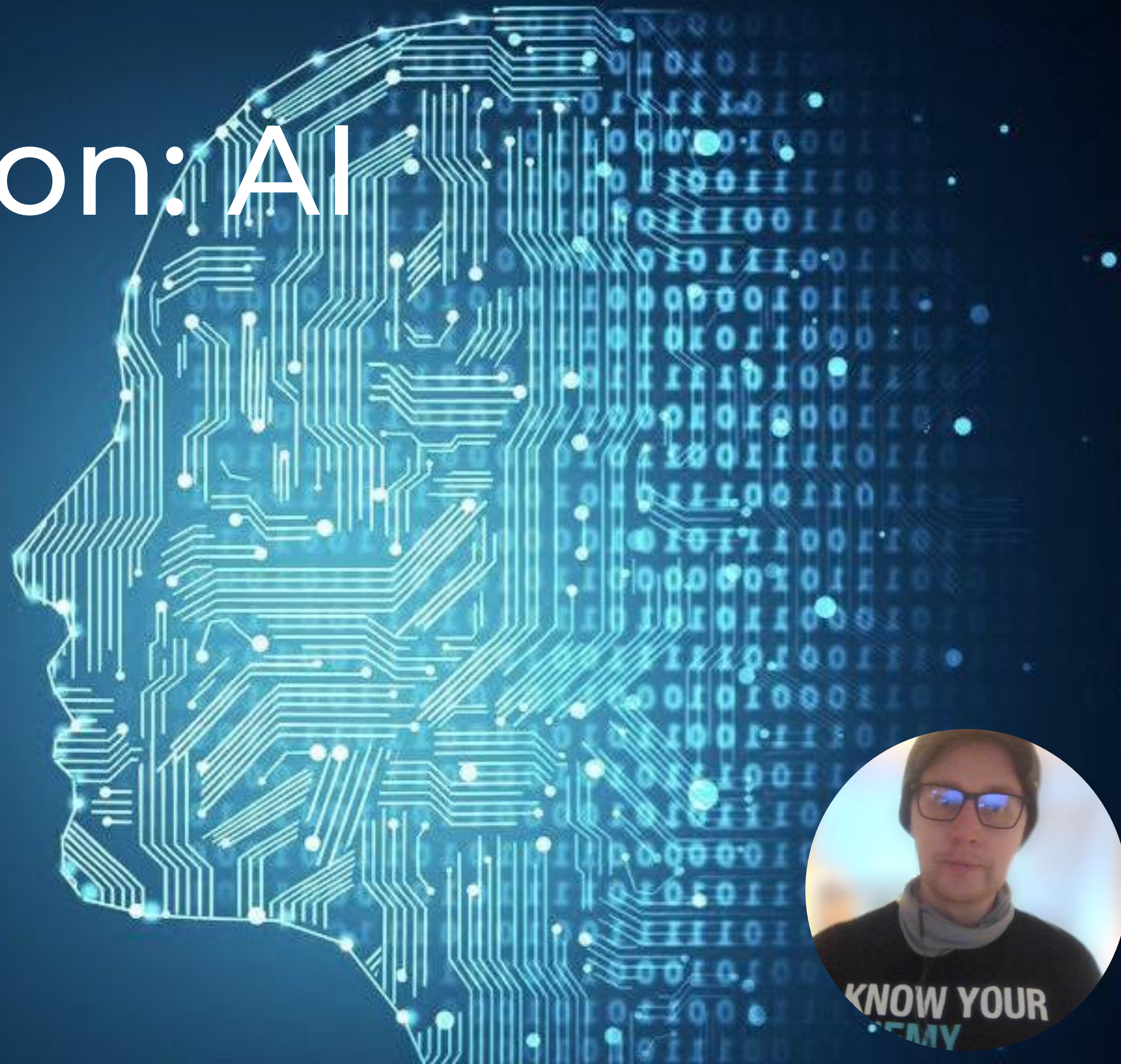
Our product uses AI and machine learning to provide a scalable, efficient, and effective solution to the problem of disinformation.

By analyzing data from a wide range of sources, including traditional and unknown social media, news outlets, and other online platforms, our product can detect patterns of disinformation and allow analysts to track its spread across the internet.

It can also identify the sources of disinformation and provide real-time alerts to users.



The solution: AI



Our product is designed to tackle the problem of disinformation using advanced machine learning algorithms and AI using a combination of natural language processing and network analysis to identify patterns and trends in online disinformation.

It works by analyzing large datasets of news articles, social media posts, and other online content, identifying patterns and trends that suggest the presence of disinformation campaigns.



We further augment our AI model by utilizing the DISARM framework, a qualitative collection of adversary & defender TTPs for describing and identifying disinformation threats.

We can apply this framework to provide advanced reasoning about social media/news article and other data it is given.

DISARM Objects

The disarm frameworks contain many object types, including tactic stages (steps in an incident), and techniques (activities at each tactic stage). We also have data objects to show how the frameworks are used in practice, and to make our datasets on tools and responders available.

Framework objects					
Frameworks	Phases	Tactics	Techniques	Tasks	Countermeasures
Detections	Responsetypes	Metatechniques	Playbooks	Resources	
Data objects					
Incidents	Examples	External Groups	Tools		



The product is highly customizable, allowing users to specify the type of disinformation they are looking to detect and the sources they want to monitor. This flexibility makes it well-suited to a variety of different applications and use cases.



With the rise of fake news and disinformation campaigns, our product is uniquely positioned to address a growing need in the digital age.

- Media organizations: Our product will be used by media organizations to verify the authenticity of news stories and to combat the spread of fake news and disinformation campaigns.
- Government agencies: Our product will be used by government agencies, such as intelligence agencies and law enforcement, to monitor and track disinformation campaigns by foreign actors.
- Social media platforms: Social media platforms will find it necessary to use our product to detect and remove fake news and disinformation campaigns from their platforms.
- Corporations: Our product can also be utilized by corporations to detect and remove false information about their products or services.



Our company offers a range of tiered access options for our product, each designed to meet the specific needs and budgets of our clients.

Our basic tier provides essential features at an affordable price point, while our premium tiers offer advanced features for more complex analysis.

With each tier, clients can customize their experience and access more comprehensive data and analysis tools.

We are committed to providing flexible and cost-effective solutions to help our clients combat disinformation and make informed decisions based on reliable information.

Noob – free tier with limited access for individual practitioners

Normie – Entry level tier for basic access

Chad – Significant features and up to 5 log-ins for mid-sized companies

Based – Significant features and up to 100 log-ins for mid to large sized companies

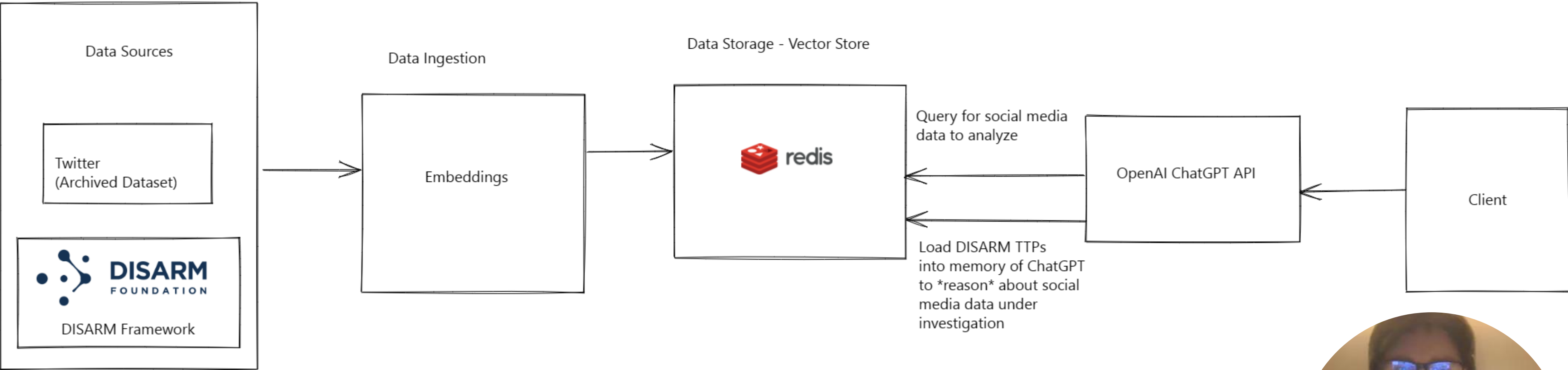
Pilled - Unlimited features and log-ins for governments, military, and large global companies



How we made it



Application Architecture





A working demo of
it



disarm_chatgpt.py

disarm_chatgpt.py

```
1  import redis
2  import pandas as pd
3  from sentence_transformers import SentenceTransformer
4  import openai
5  import json
6  import os
7  import numpy as np
8  from langchain.prompts.prompt import PromptTemplate
9  from langchain.llms import OpenAI
10 from langchain.chains import ChatVectorDBChain
11 from arcblok_vectorstore_redis import ArcBlokRedis
12 from langchain.embeddings.openai import OpenAIEmbeddings
13
14
15 class Disarm_GPT(object):
16     NUM_VECTORS = 300
17
18     def __init__(self):
19         embeddings = OpenAIEmbeddings()
20         # Connect to redis instance
21         self.redis_client = ArcBlokRedis(
22             host='redis-14733.c233.eu-west-1-1.ec2.cloud.redislabs.com',
23             port=14733,
24             password='SGY0vbdLGFRtmNGylXGdAiwABVpLUY6Z',
25             index_name='disarm',
26             embedding_function=embeddings.embed_query)
27
28     # Initialize sentence transformer model
```

Ln 1, Col 1 History

Our AI product is a powerful tool for fighting disinformation in the digital age. By leveraging advanced machine learning algorithms and blockchain technology, we can detect and respond to disinformation campaigns in real time, helping to safeguard against the spread of fake news and other malicious content.

We believe that our product has the potential to make a significant impact in the fight against disinformation. We are excited to continue developing and improving it in the years to come. Thank you for your attention, and we look forward to answering any questions you may have.



For this, a screen recording of how you perform a common user's interaction might be perfect



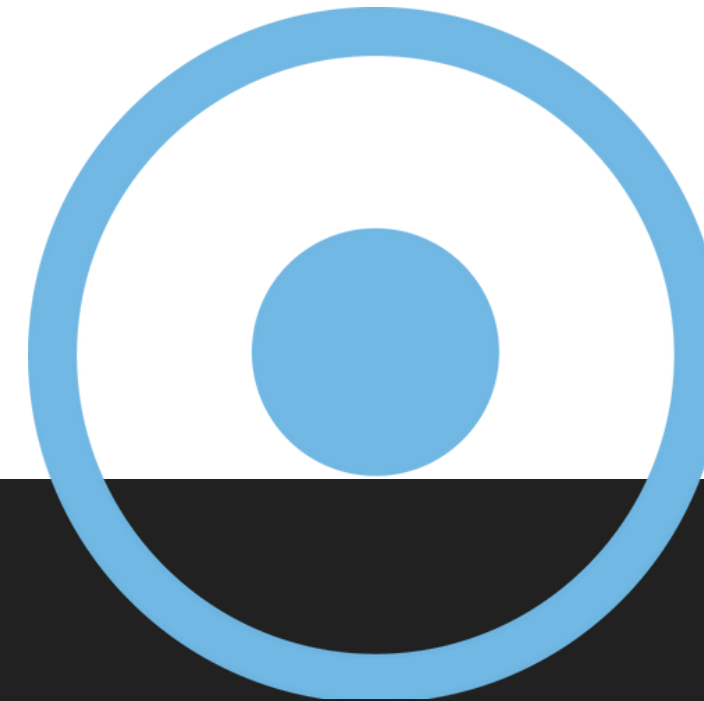
XSplit

[Link](#)



OBS Studio

[Link](#)



Screencast-O-Matic

[Link](#)



MONOSNAP

[Link](#)

Thanks!

