black hat ASIA 2022

ReconPal: Leveraging NLP for Infosec

Jeswin Mathai, Shantanu Kale and Sherin Stephen

About US

Jeswin Mathai

- Senior Security Researcher @ INE
- Published Research at Black Hat US/Asia Arsenal, DEF CON USA/China Demolabs
- Gave research talk at DEF CON China and Rootcon Philippines
- Co-Trainer in Training:

  - Black Hat Asia

  - HITB AMS, GSEC

  - NZ OWASP day

  - Rootcon 13

About US

Shantanu Kale

- Cloud Developer @ INE
- Experienced with AD and Cloud Security
- Led team for Smart India Hackathon organized by GoI

Sherin Stephen

- Cloud Developer @ INE
- Experienced in Building and maintaining reusable code and robust cloud services

OWASP
NEW ZEALAND DAY
2019

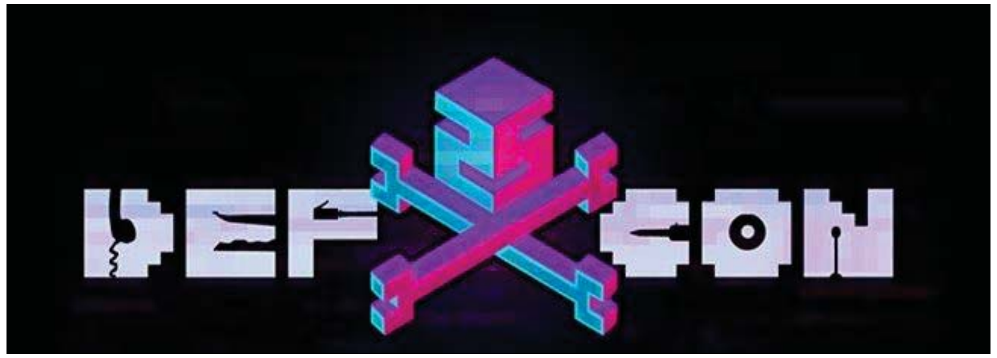
百度安全
DEF CON
CHINA
1.0

black hat

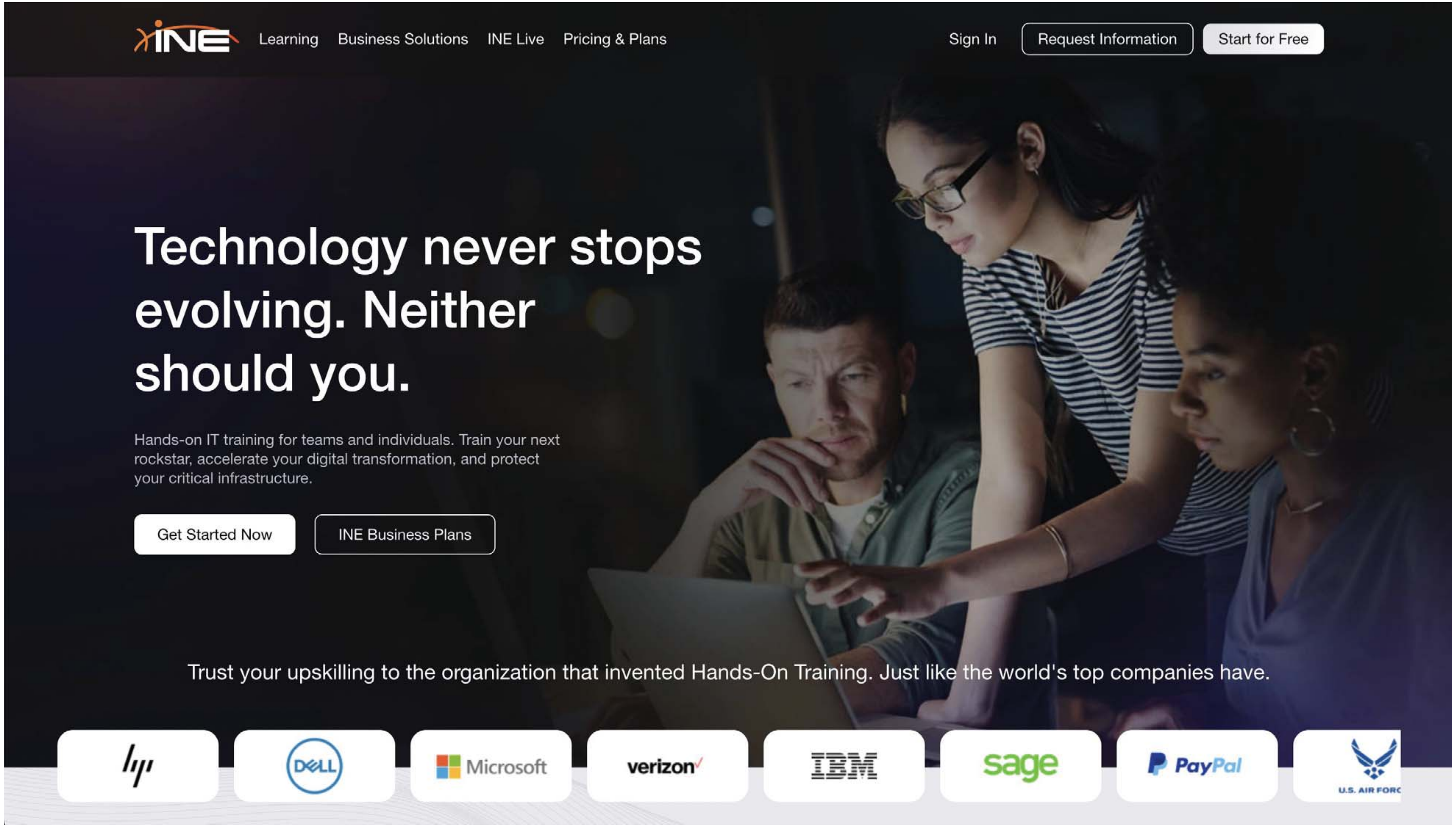
WIRELESS
VILLAGE

WALL OF SHEEP

HITB


DEF CON

ROOTCON
Hacker Conference



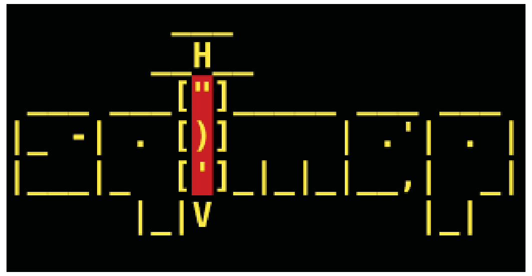
Technology never stops evolving. Neither should you.

Hands-on IT training for teams and individuals. Train your next rockstar, accelerate your digital transformation, and protect your critical infrastructure.

Trust your upskilling to the organization that invented Hands-On Training. Just like the world's top companies have.

ReconPal

- Text/Speech to Go
- Beginner Friendly
- Age of automation
- A generic way of using multiple tools
- Over thousands of tools
- Enormous number of options and syntax



Components

- OpenAI GPT-3
- Shodan API
- Speech-to-Text
- Telegram Bot
- Docker Containers

SHODAN

OpenAI

Speech-to-Text

docker

OpenAI

- AI research and deployment company.

- Founded By

  - Elon Musk

  - Sam Altman

- Popular Projects

  - DALL-E

  - GPT-3

  - Jukebox

  - GPT-2

GPT-3

- One of the Most Powerful Language Model Ever
- 175 Billion Parameters
- Applications
  - Write Code
  - Complete text (write poems etc)
  - Generate complex mathematical expressions
  - Fill Missing data (spreadsheet)
  - Generate Simple text from Legal Text and vice-versa
  - Language Translation

OpenAI

GPT-3 Comparison

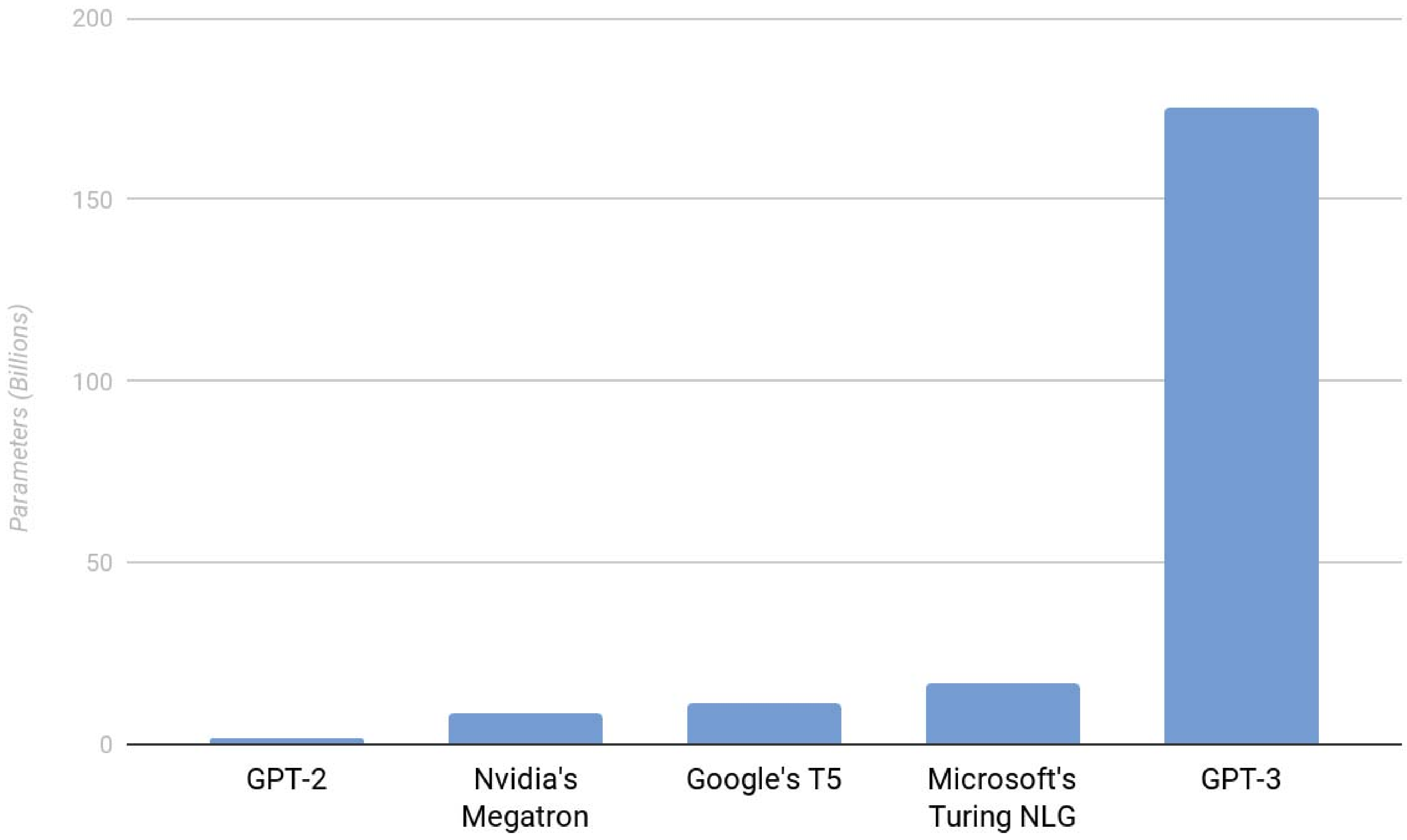


Image Source: https://www.sigmoid.com/blogs/gpt-3-all-you-need-to-know-about-the-ai-language-model/

GPT-3 Comparison

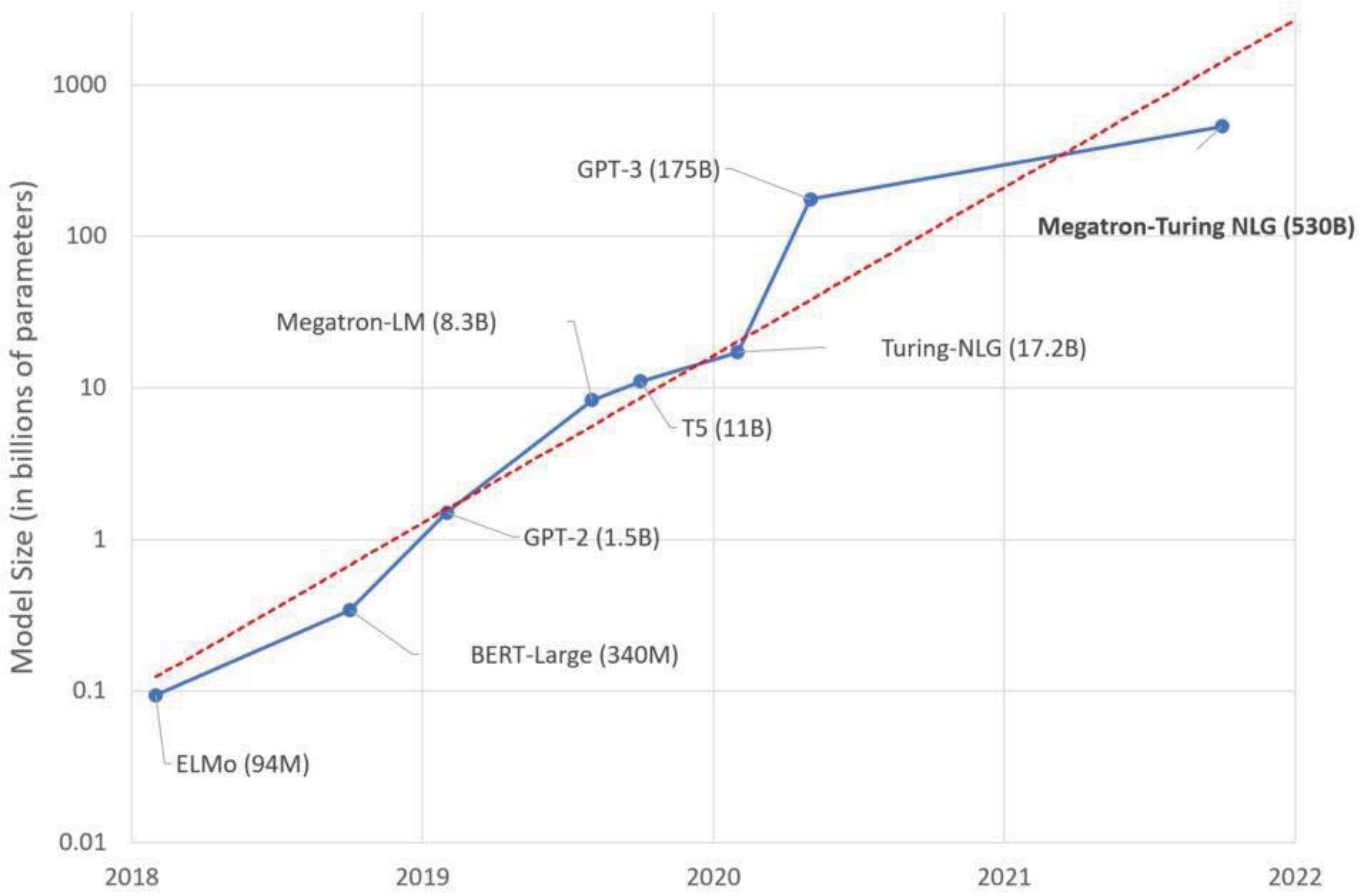


Image Source: https://www.microsoft.com/en-us/research/blog/using-deepspeed-and-megatron-to-train-megatron-turing-nlg-530b-the-worlds-largest-and-most-powerful-generative-language-model/

GPT-3 Comparison

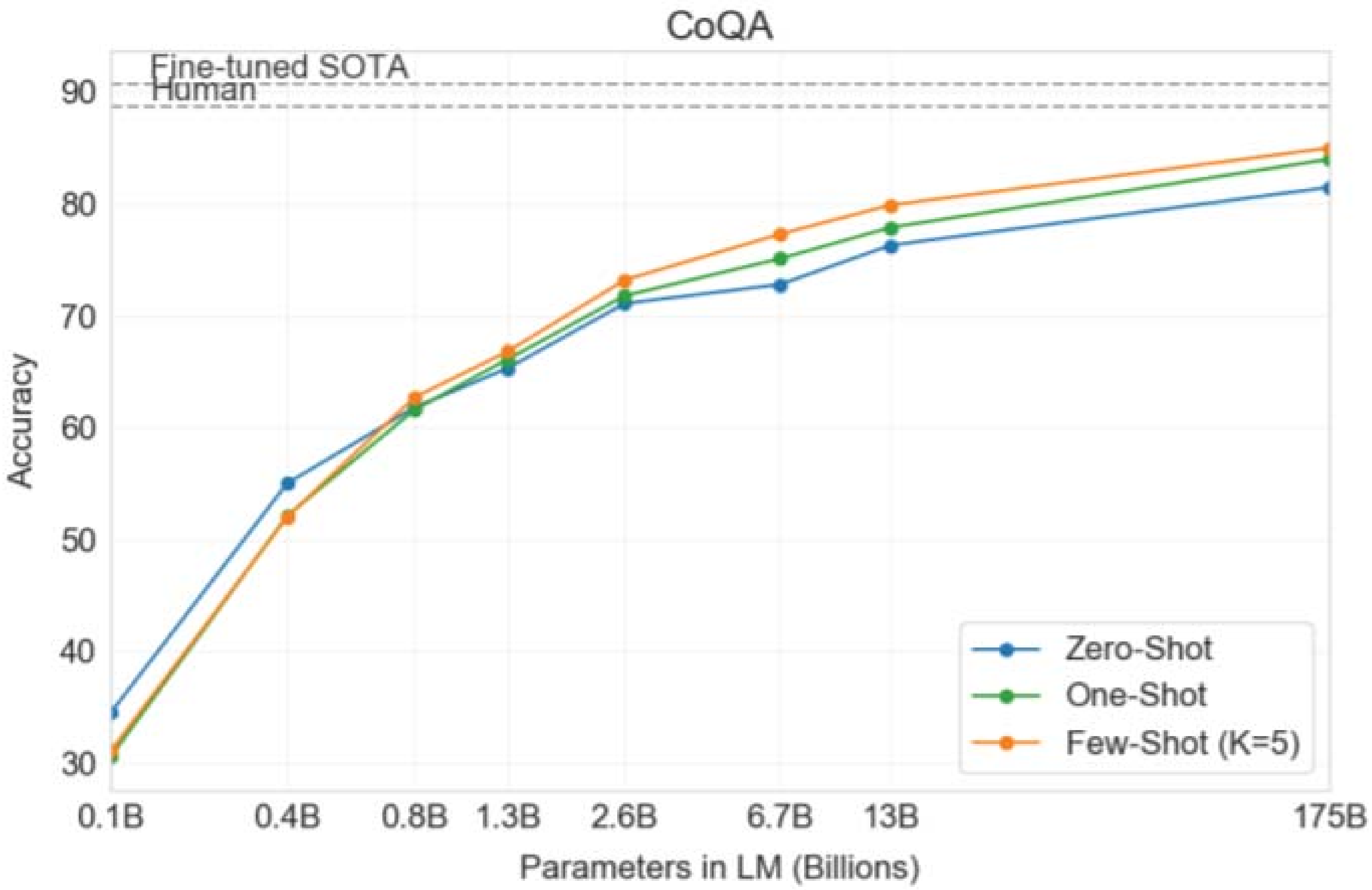
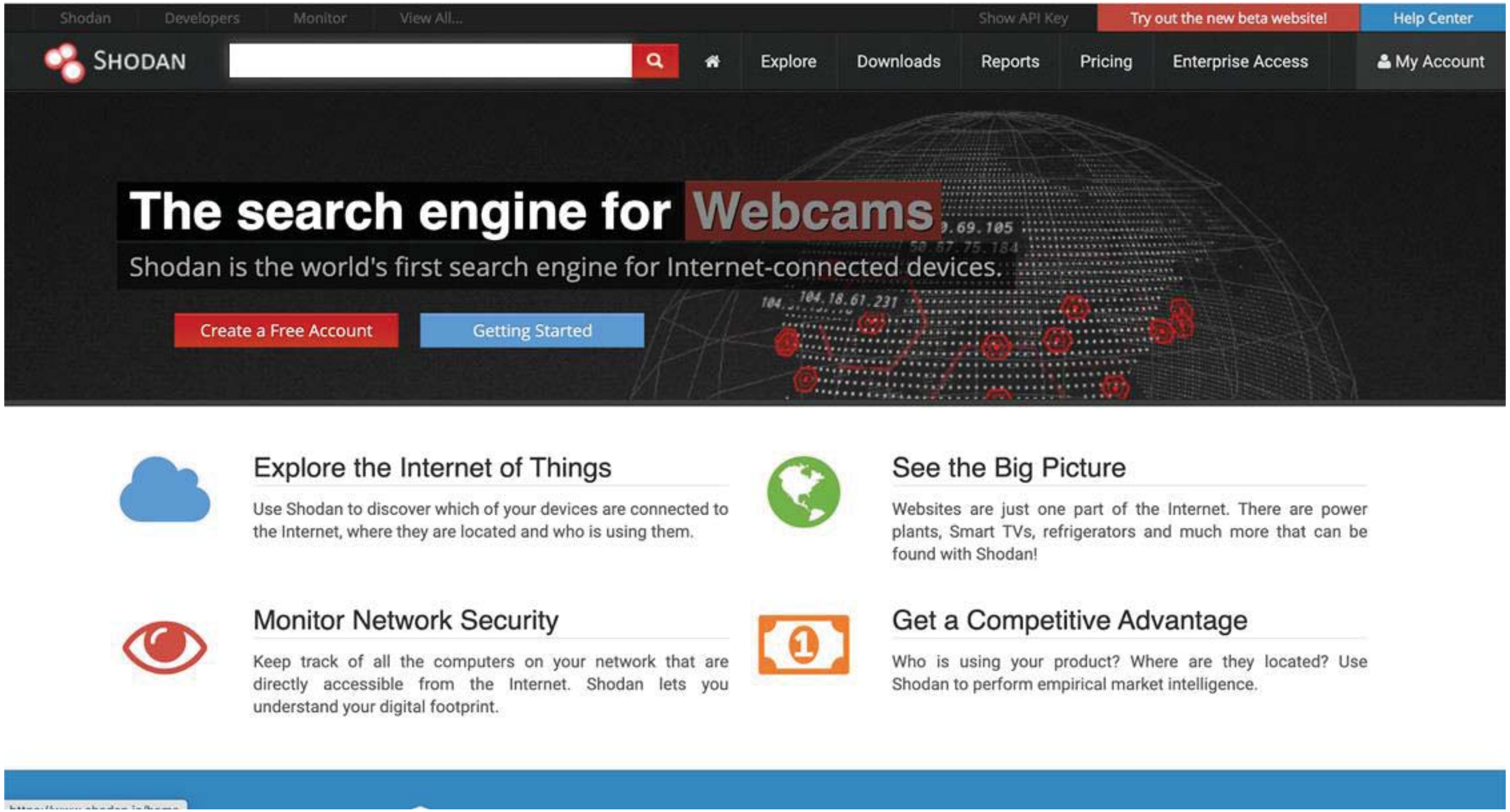


Image Source: https://lex.substack.com/p/long-take-openai-backed-with-1b-by

Shodan



The search engine for Webcams

Shodan is the world's first search engine for Internet-connected devices.

Explore the Internet of Things

Use Shodan to discover which of your devices are connected to the Internet, where they are located and who is using them.

See the Big Picture

Websites are just one part of the Internet. There are power plants, Smart TVs, refrigerators and much more that can be found with Shodan!

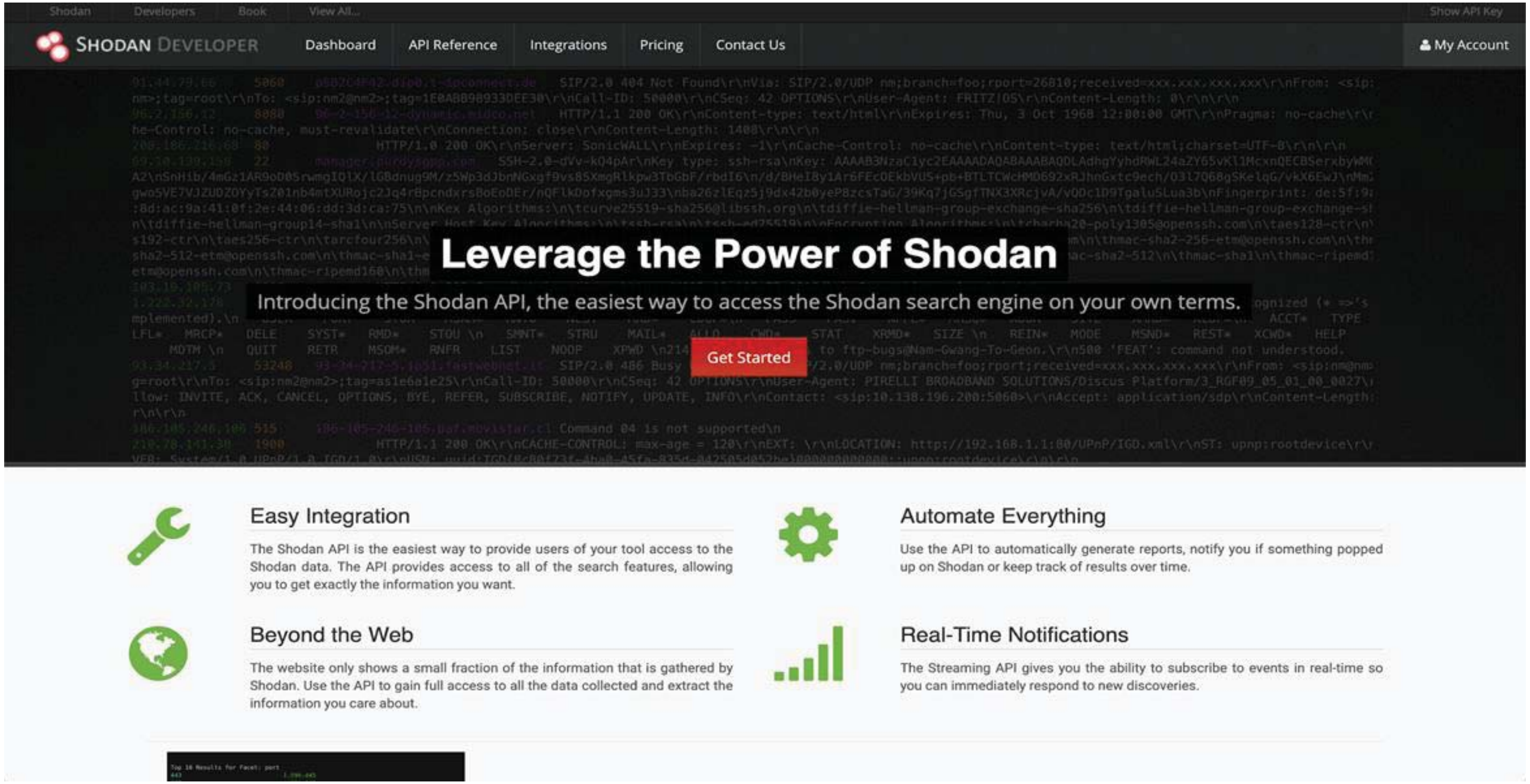
Monitor Network Security

Keep track of all the computers on your network that are directly accessible from the Internet. Shodan lets you understand your digital footprint.

Get a Competitive Advantage

Who is using your product? Where are they located? Use Shodan to perform empirical market intelligence.

Shodan API



Leverage the Power of Shodan

Introducing the Shodan API, the easiest way to access the Shodan search engine on your own terms.

Easy Integration

The Shodan API is the easiest way to provide users of your tool access to the Shodan data. The API provides access to all of the search features, allowing you to get exactly the information you want.

Automate Everything

Use the API to automatically generate reports, notify you if something popped up on Shodan or keep track of results over time.

Beyond the Web

The website only shows a small fraction of the information that is gathered by Shodan. Use the API to gain full access to all the data collected and extract the information you care about.

Real-Time Notifications

The Streaming API gives you the ability to subscribe to events in real-time so you can immediately respond to new discoveries.

Google Speech-to-Text

- Managed by Google
- State-of-the-art accuracy
- API Access
- Affordable

Speech-to-Text

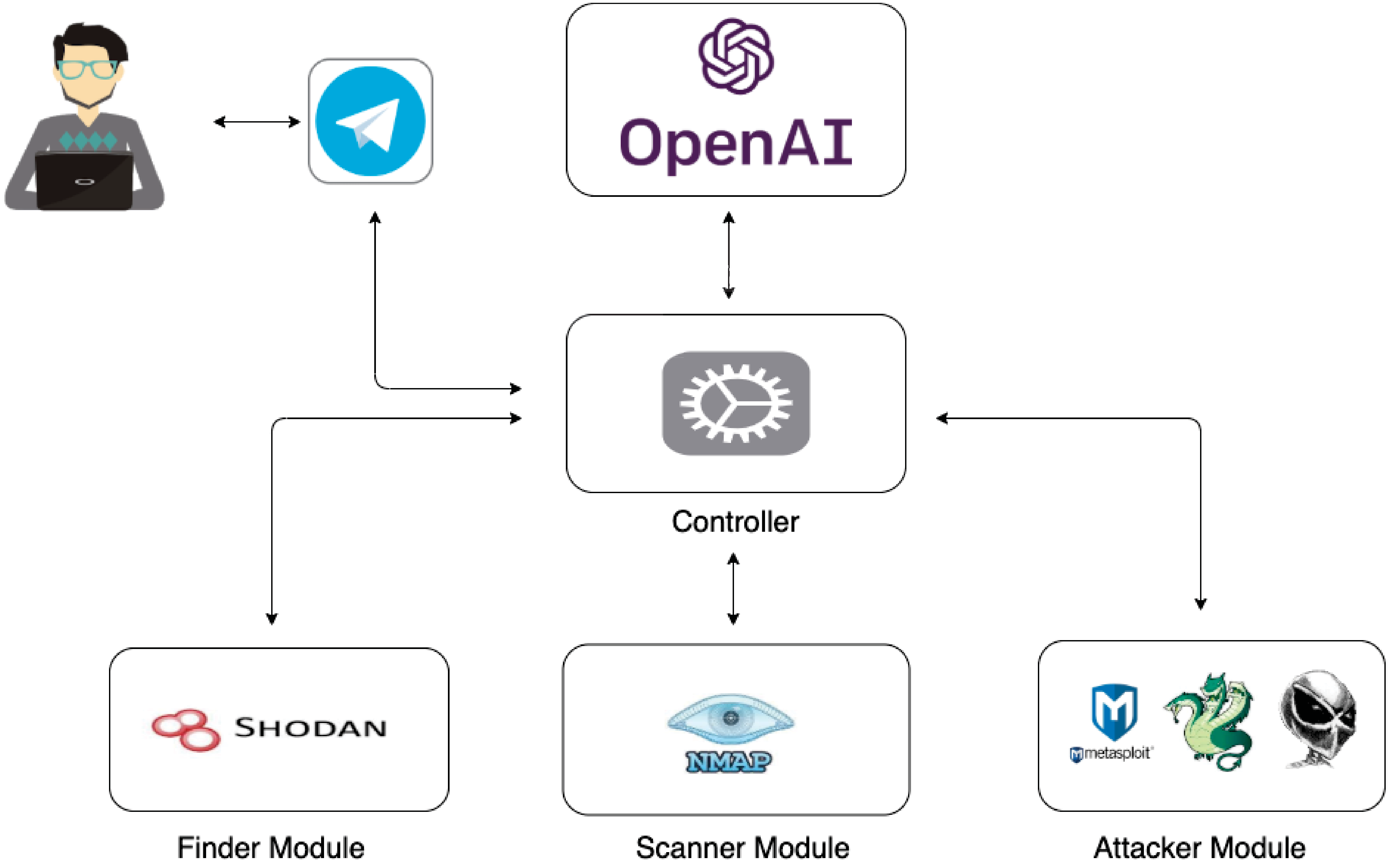
Telegram Bot

- Desktop application for all platforms
- Python API Support
- Free
- Easy to Setup

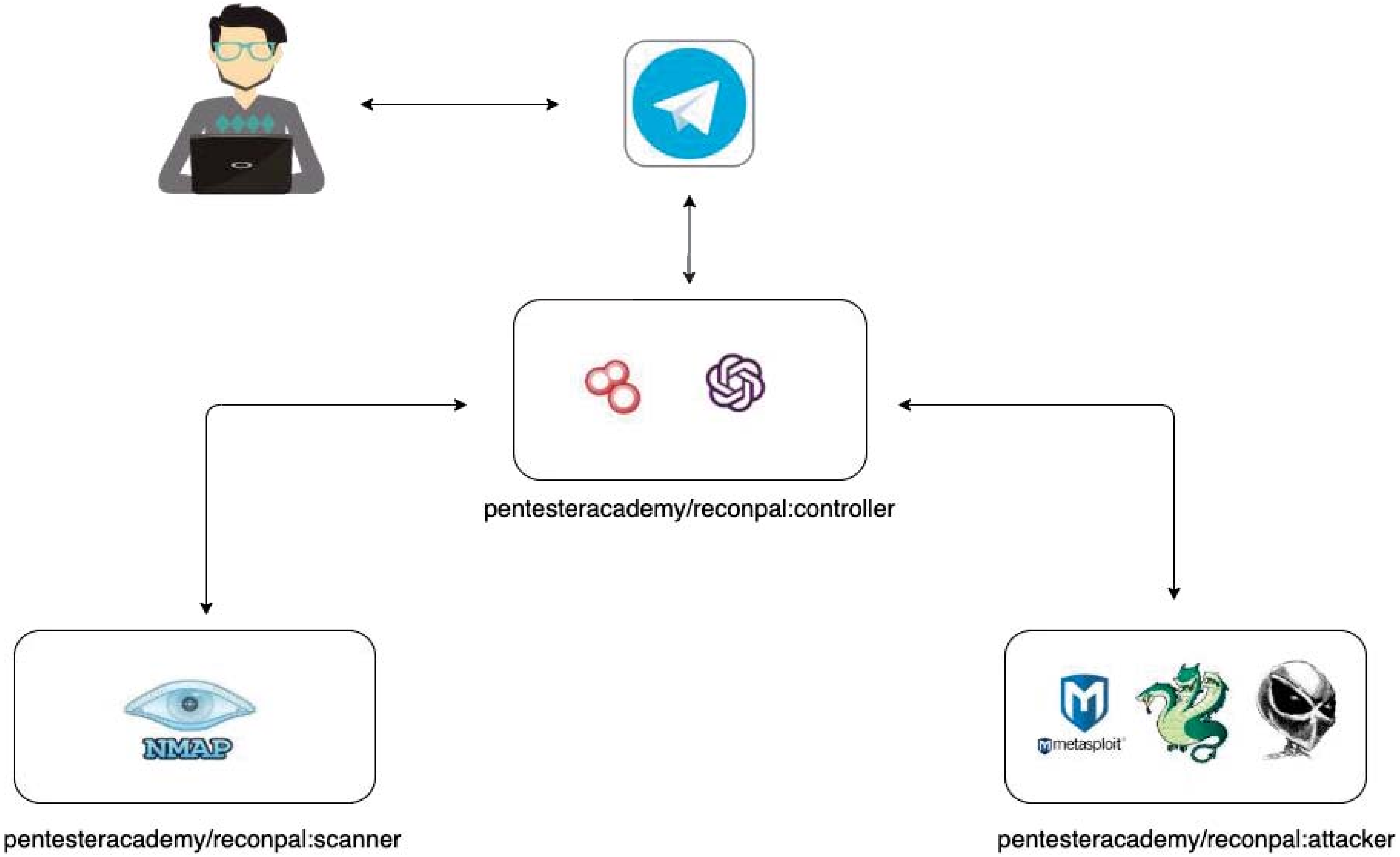
Docker

- Easy to setup and manage
- Images
  - pentesteracademy/reconpal:controller
  - pentesteracademy/reconpal:scanner
  - pentesteracademy/reconpal:attacker
- Use docker-compose to setup everything
- Future updates will only require pulling updated images

ReconPal Architecture



ReconPal Packaging


pentesteracademy/reconpal:controller
NMAP
pentesteracademy/reconpal:scanner
metasploit
pentesteracademy/reconpal:attacker

Requirements

- OpenAI API Key
- Shodan API Key
- Google Cloud Credentials (Speech-to-Text Support)
- Packages
  - docker.io
  - docker-compose

Installation

- Create a Telegram Bot

- Install Packages

```
- sudo apt-get updates
- sudo apt-get install docker.io
- sudo curl -L "https://github.com/docker/compose/releases/download/1.26.0/docker-compose-$(uname -s)-$(uname -m)" -o
  /usr/local/bin/docker-compose
- chmod +x /usr/local/bin/docker-compose
```

Installation

- Fetch the docker-compose.yml

```
- git clone https://github.com/pentesteracademy/reconpal.git
- cd reconpal
```

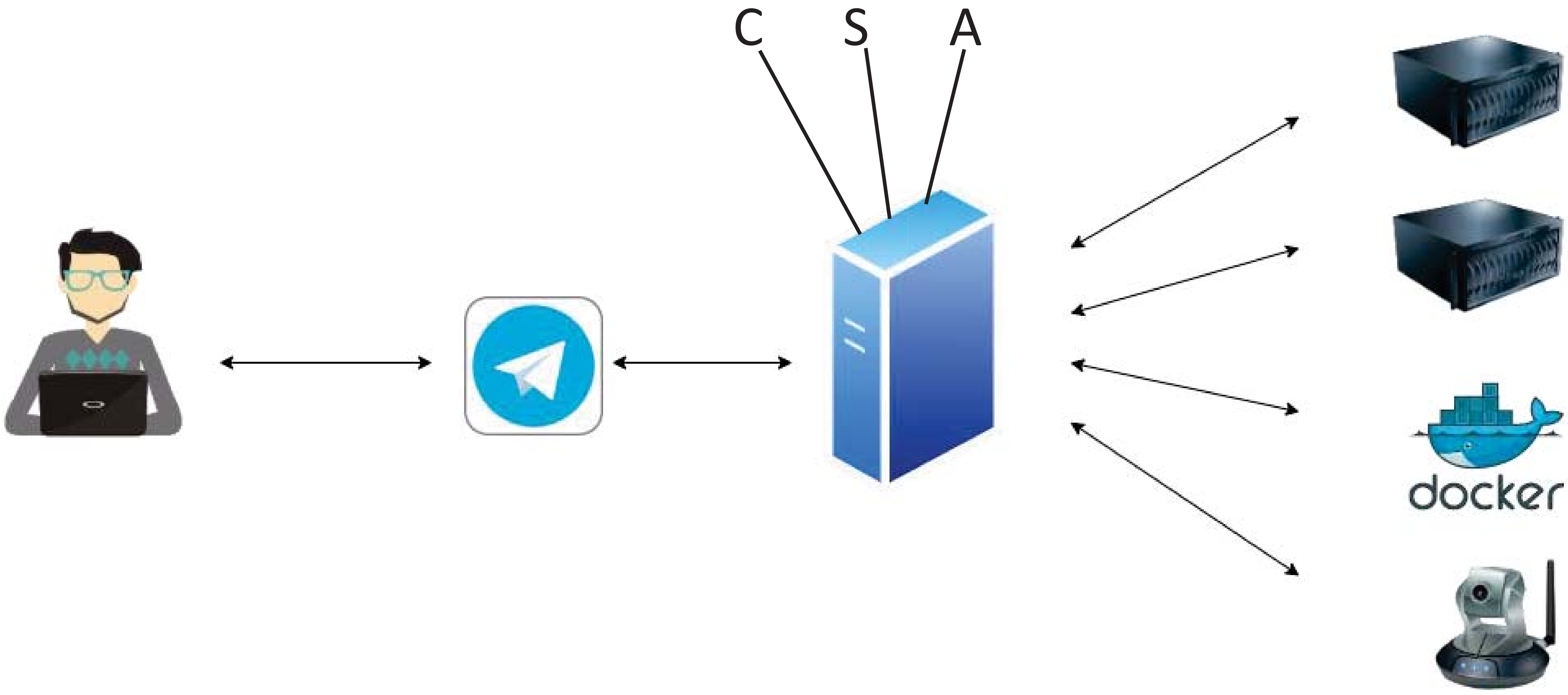
- Edit docker-compose.yml and update the OpenAI, Shodan API Key, GCP credentials file path and Bot token

- Start containers with docker-compose

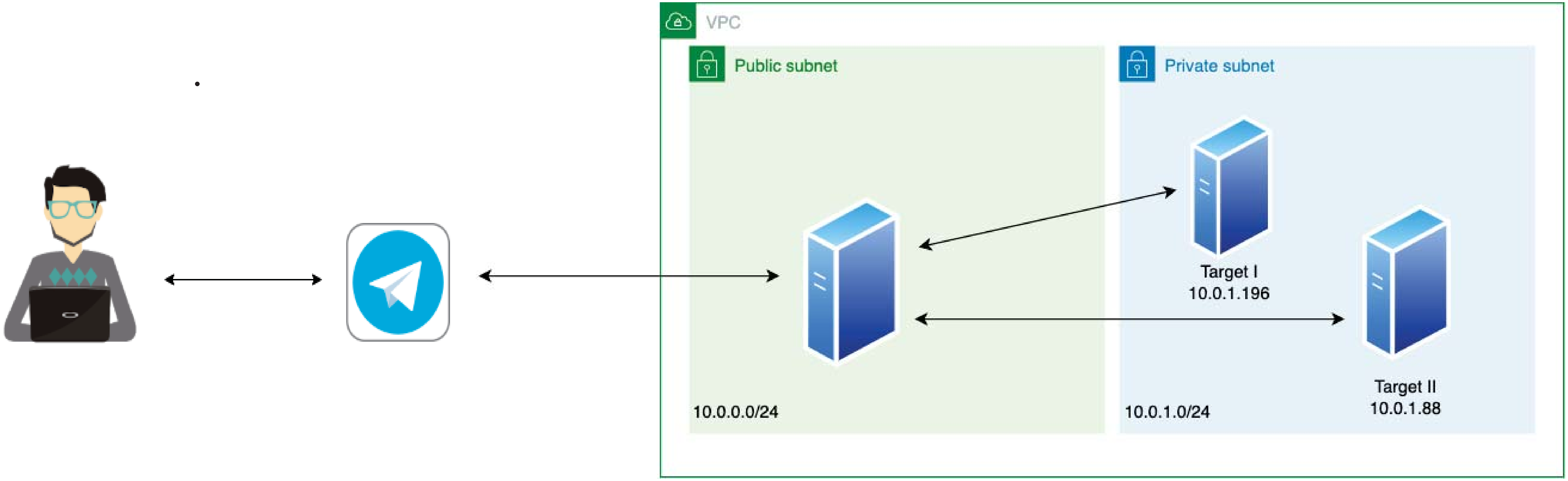
```
- sudo docker-compose up
```

Installation

Setup



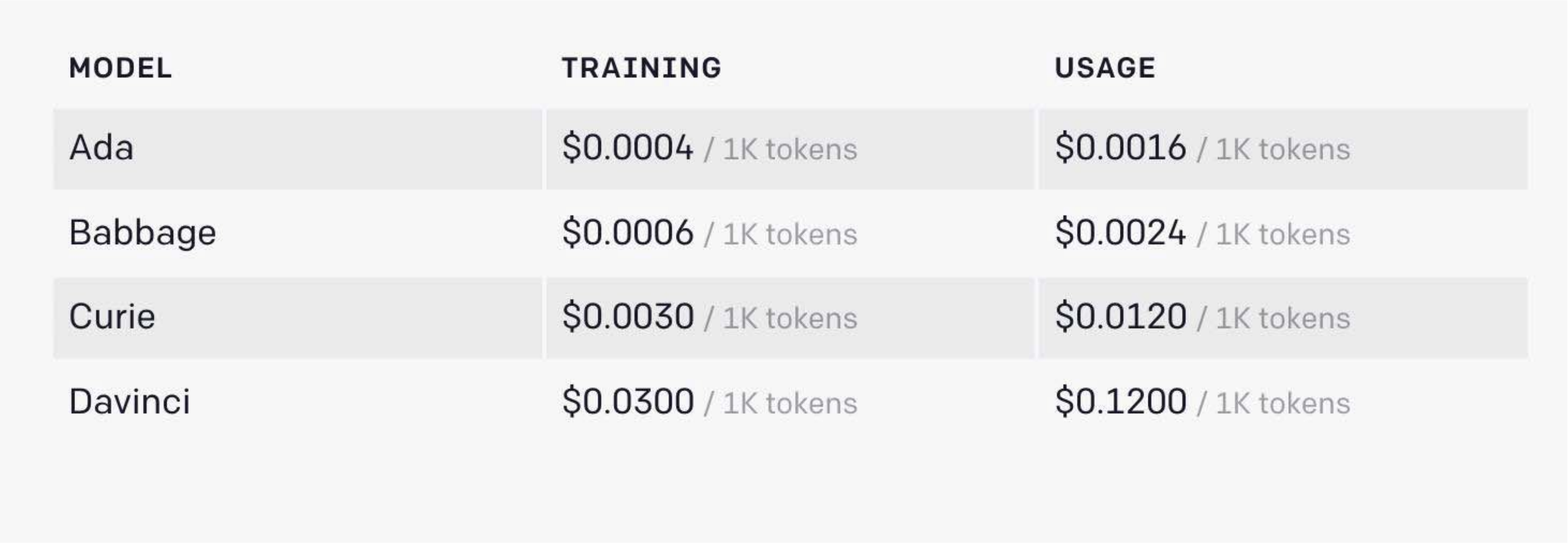
Setup for demo



DEMO

Running Cost

- OpenAI API Pricing



MODEL	TRAINING	USAGE
Ada	$0.0004 / 1K tokens	$0.0016 / 1K tokens
Babbage	$0.0006 / 1K tokens	$0.0024 / 1K tokens
Curie	$0.0030 / 1K tokens	$0.0120 / 1K tokens
Davinci	$0.0300 / 1K tokens	$0.1200 / 1K tokens

Running Cost

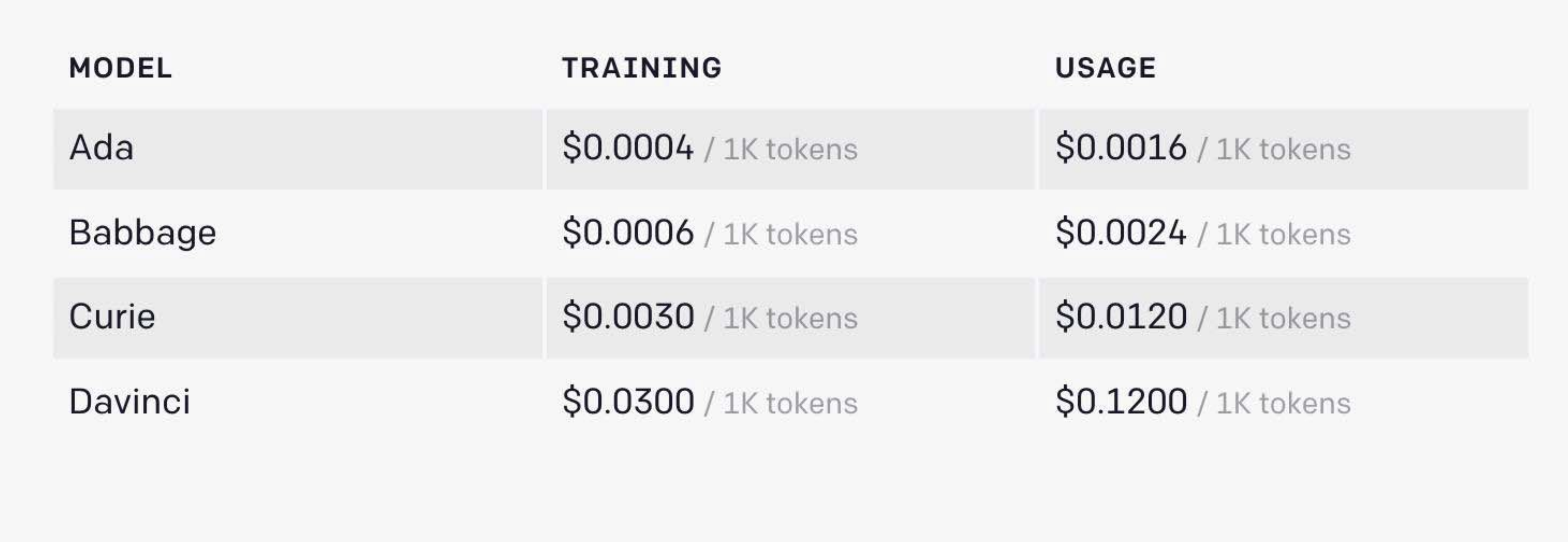
- Shodan API
  - 100 query credits per month (Lifetime membership)
    - 49 $ One-time payment
  - 10,000 query credits to unlimited (API Plans)
    - Freelancer - 69 $ / month
    - Small Business - $359 /month
    - Corporate - $1099 /month

Running Cost

- Google speech to text pricing
  - Speech Recognition (without Data Logging - default)
    - 0-60 Minutes - Free
    - Above 60 Minutes - $0.006 / 15 seconds
  - Speech Recognition (with Data Logging opt-in)
    - 0-60 Minutes - Free
    - Above 60 Minutes - $0.004 / 15 seconds

Limitations

- OpenAI API Pricing



MODEL	TRAINING	USAGE
Ada	$0.0004 / 1K tokens	$0.0016 / 1K tokens
Babbage	$0.0006 / 1K tokens	$0.0024 / 1K tokens
Curie	$0.0030 / 1K tokens	$0.0120 / 1K tokens
Davinci	$0.0300 / 1K tokens	$0.1200 / 1K tokens

Future Plans

- Enhanced input processing
- Generating summary from output
- Interactive attacker sessions
- Session Management

Thanks

jmathai@ine.com