LexaScan

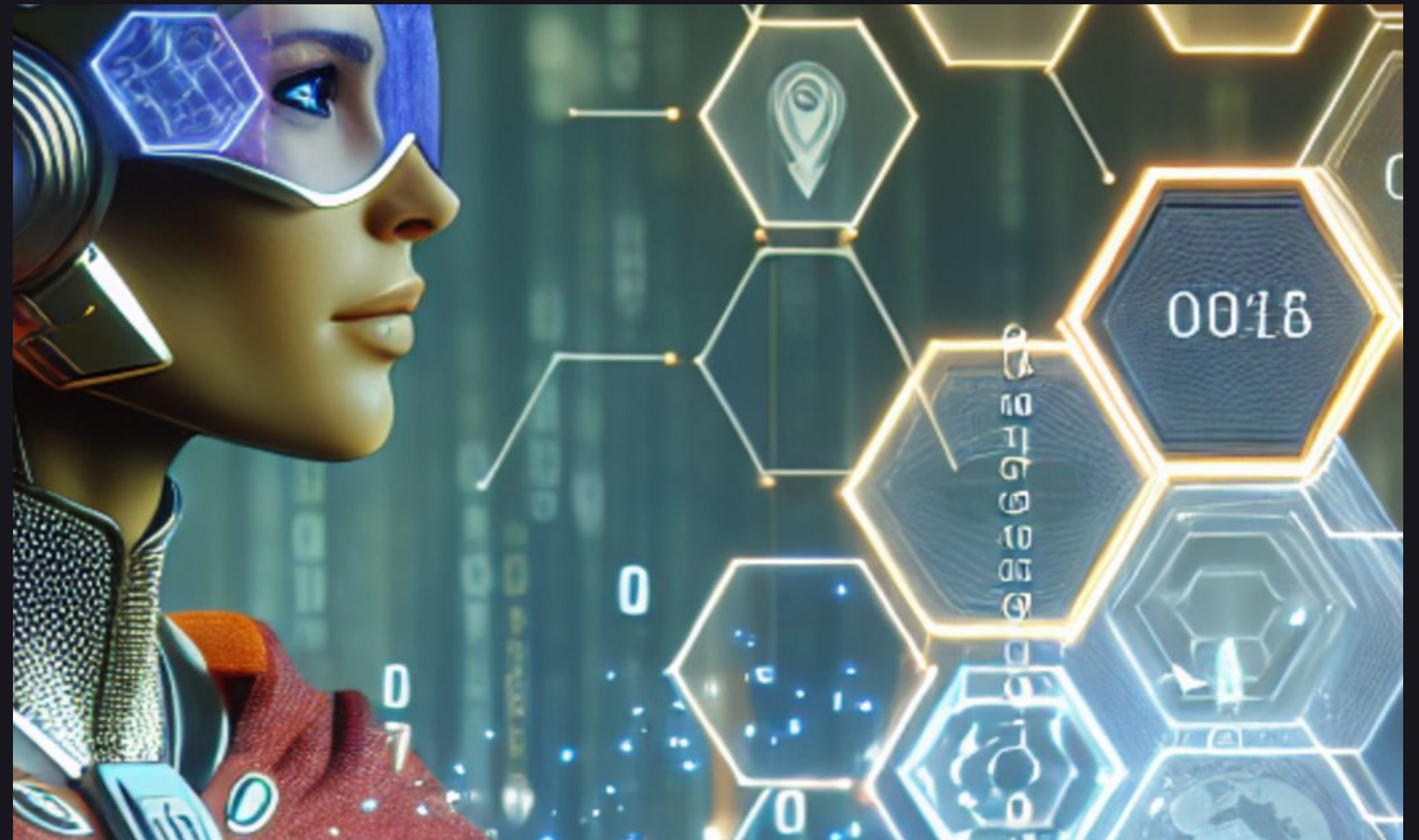## Data extraction from image and text

Information Gathering and Searching

According to a study by McKinsey, employees spend almost 2 hours per day searching and gathering information. Manual extraction processes cannot scale to handle large volumes of data.

Problem

LexaScan

Automated data extraction solves the inefficiencies of manual information gathering. LexaScan is an AI solution that extracts data from text and images, delivering structured JSON format.

Solution

Technology

We used GPT-4 Vision from OpenAI to extract data from text and images. We also used TruLens to ensure that the output is relevant and harmless.
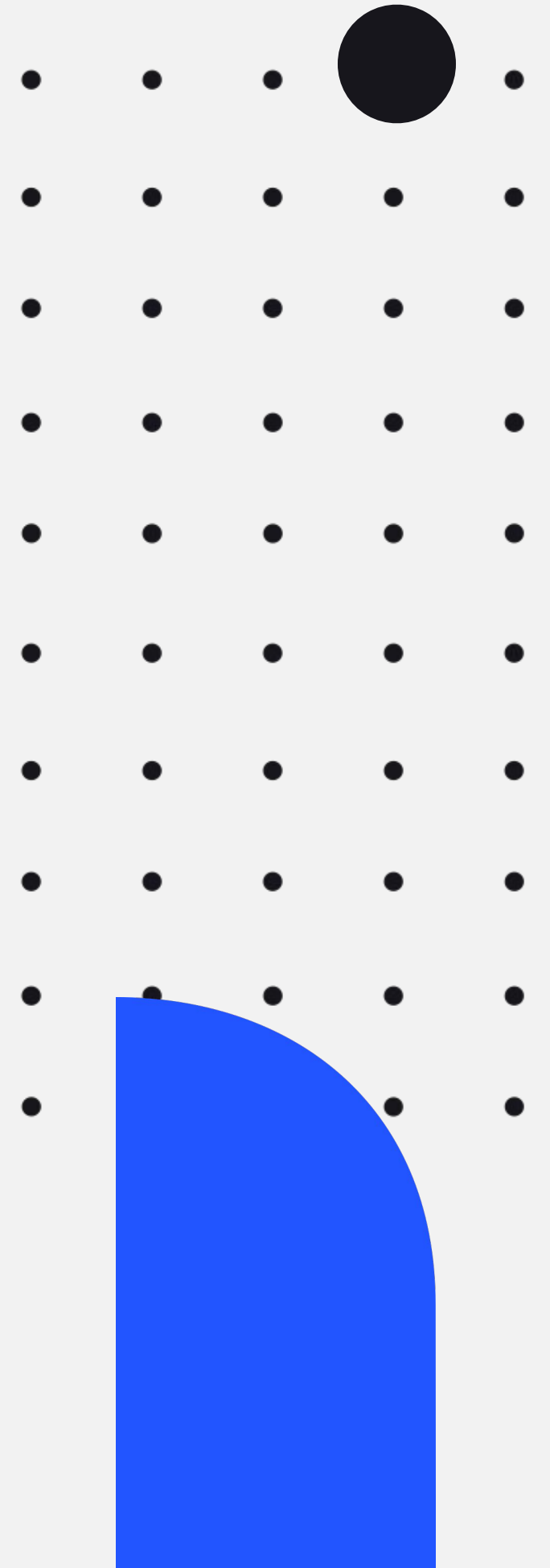
Market

SAM

10 percent of
TAM = $140
million

TAM

Estimated at $1.4
billion in 2022
and is expected
to reach $3.8
billion by 2028

## SOM

30 percent of
SAM = $42
million

Demo

ms

car get there? Police p... ▼

Extract

Extracted Info:

```
▼ {
    "title" : "How did a car get there? Police probe 'unusual' crash"
    "description" :
    "The driver of the lodged car, who was briefly trapped in the vehicle, w
    rescued with minor injuries."
    ▼ "entities" : {
        ▼ "people" : [
            0 : "driver"
        ]
        ▶ "places" : []
        ▶ "companies" : []
        ▼ "organizations" : [
            0 : "Police"
        ]
    }
    ▶ "dates_and_times" : []
    ▼ "image_objects" : [
        0 : "car"
        1 : "crash site"
```

Share

https://lexascan.streamlit.app
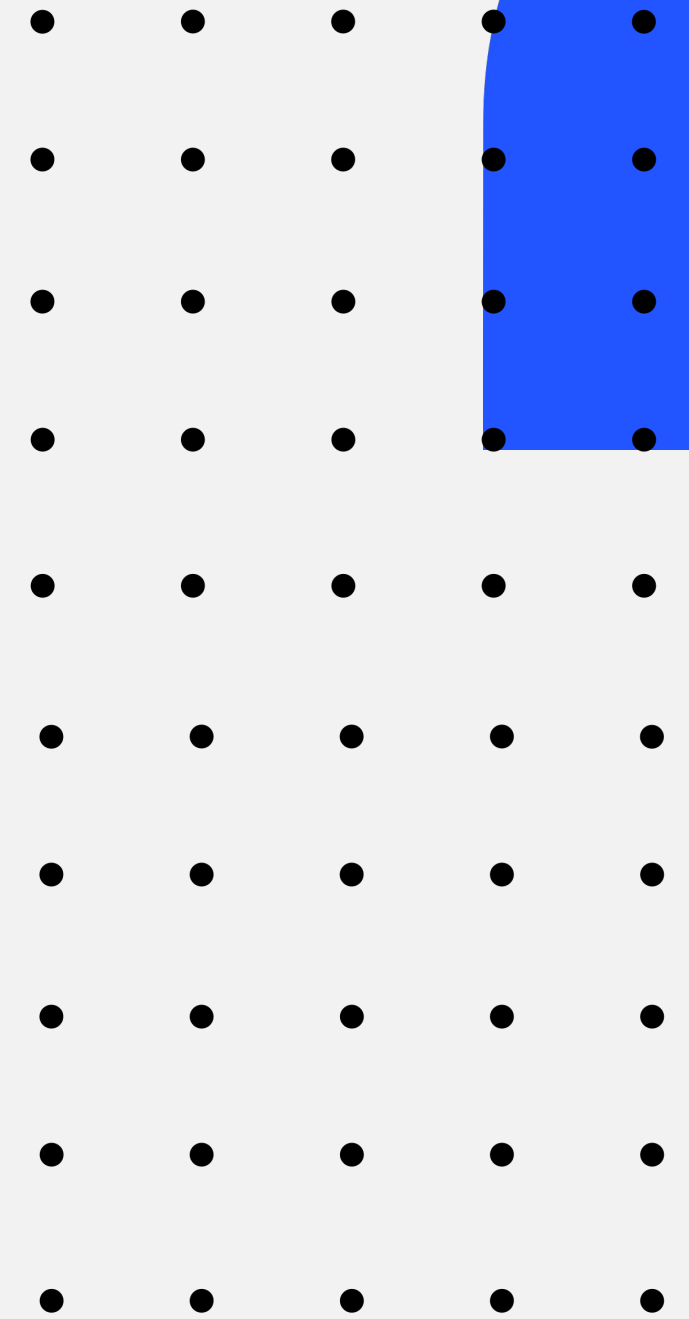
Team

We utilize advanced technology to address issues like health care accessibility and disease prevention, emphasizing creativity, collaboration, and ethical principles among our diverse team of computer science, quantum computing, and AI experts to develop scalable solutions for pressing issues.
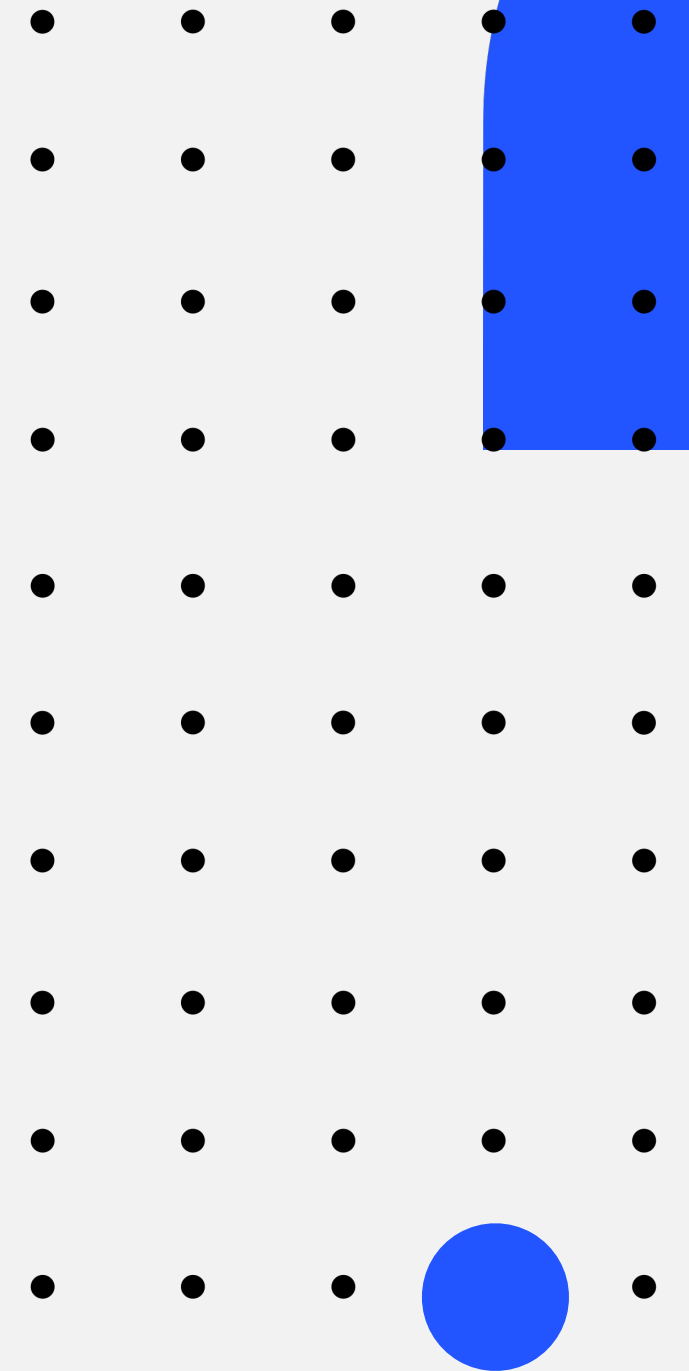
Thank You

Thanks to LabLab for hosting.

Do you have any questions?

https://discord.gg/8CTd8ShKQq