

**NOVEMBER 30, 2022**

# AI agents are changing the game

*"What was our best performing retail branch in California last year?"*

Jane, Head of Sales

*"Remind me whenever a distribution agreement needs to be renewed?"*

John, Legal

*"Give me a summary of the marketing section of our last annual report?"*

Kartik, Marketing

*"What was the sentiment of all customer feedback submissions of last week?"*

Bob, Customer Success

*"What were the exact words of our first ever TV commercial?"*

Elliot, CEO

*"Here's a picture of a product, show me something similar!"*

Amrita, Online-Shopper



Object

Vector

Task



IMAGE



IMAGE  
TRANSFORMER



[1.3, 0.6, 1.2, -1.3, ...]



Object recognition, deduplication,  
scene detection, product search, ...



TEXT



NLP  
TRANSFORMER



[0.3, -0.4, 1.2, 0.3, ...]



Translation, understanding, Sentiment,  
Question Answering, Semantic Search, ...



AUDIO



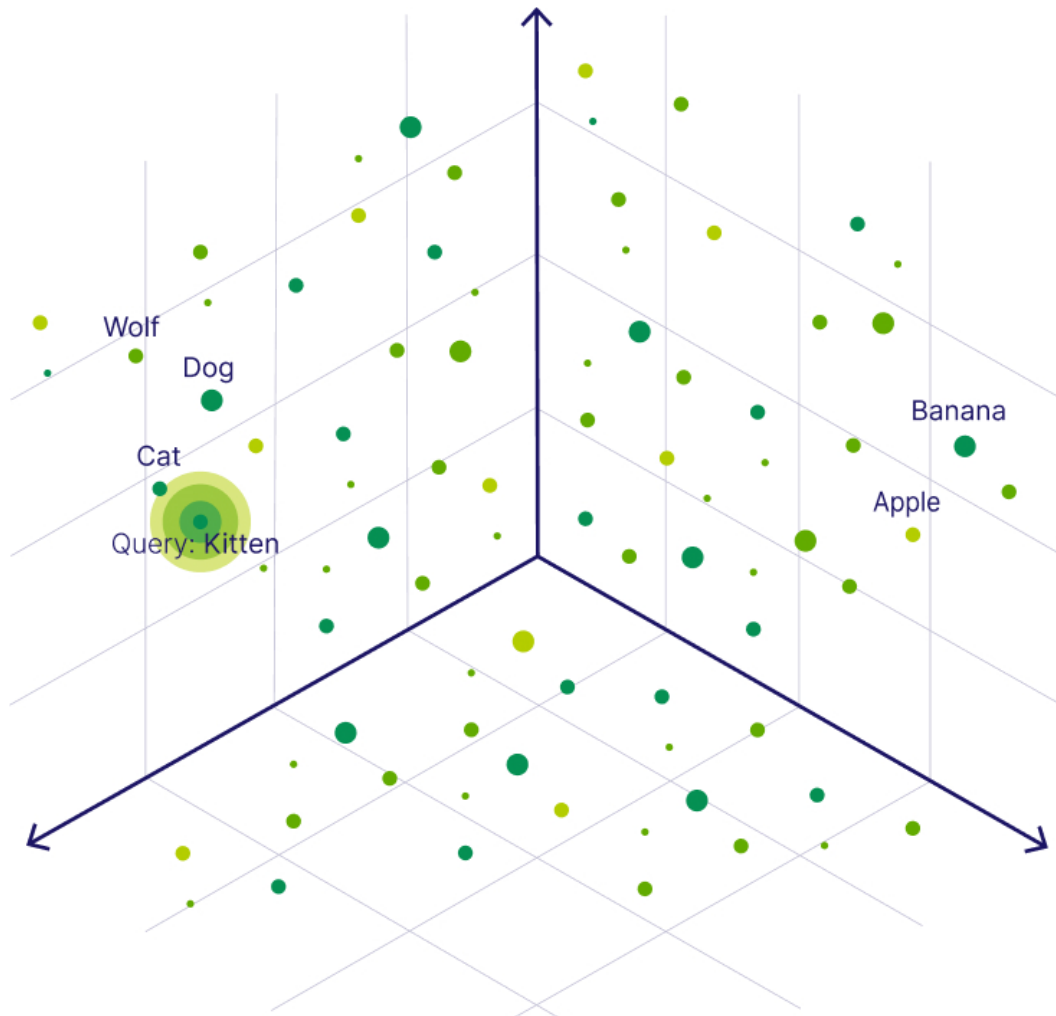
AUDIO  
TRANSFORMER



[1.2, -0.3, 0.7, -1.8, ...]



Anomaly detection, speech-to-text, music  
transcription, machinery malfunction, ...



Object

Vector



IMAGE



IMAGE  
TRANSFORMER

[1.3, 0.6, 1.2, -1.3, ...]



TEXT



NLP  
TRANSFORMER

[0.3, -0.4, 1.2, 0.3, ...]

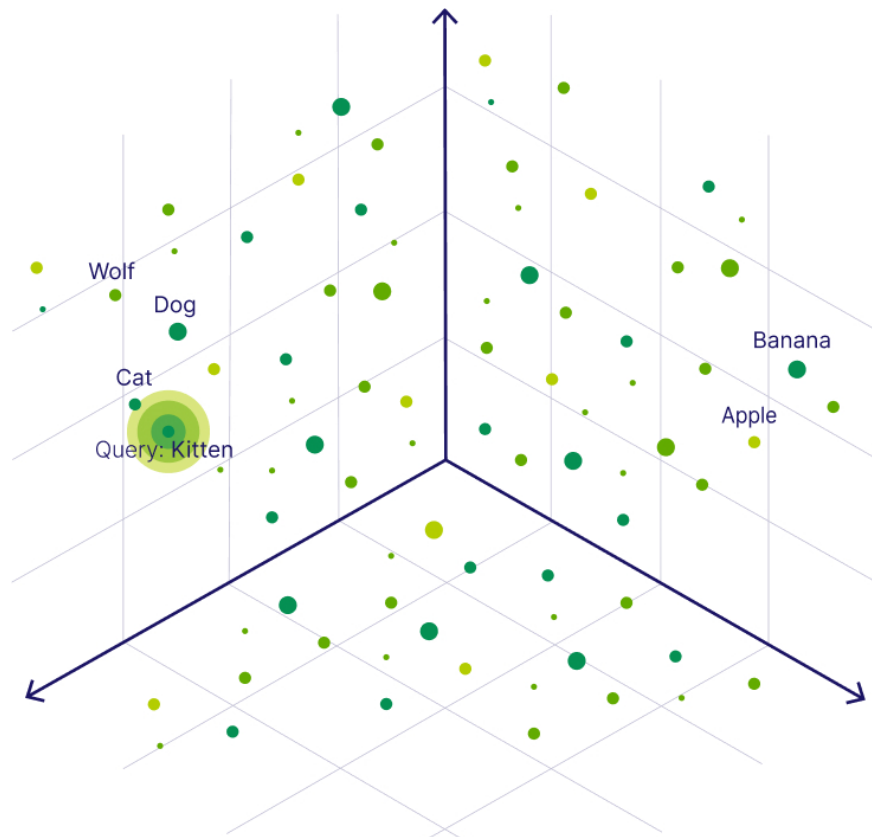


AUDIO

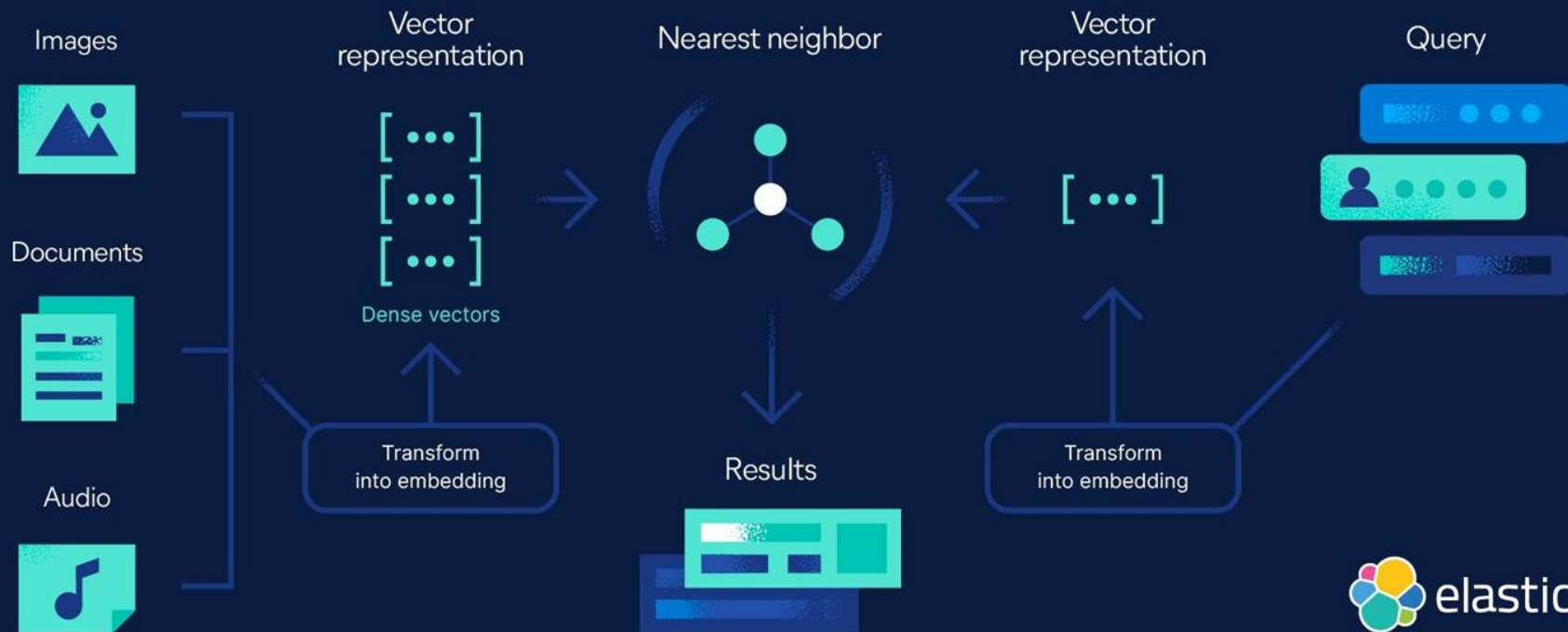


AUDIO  
TRANSFORMER

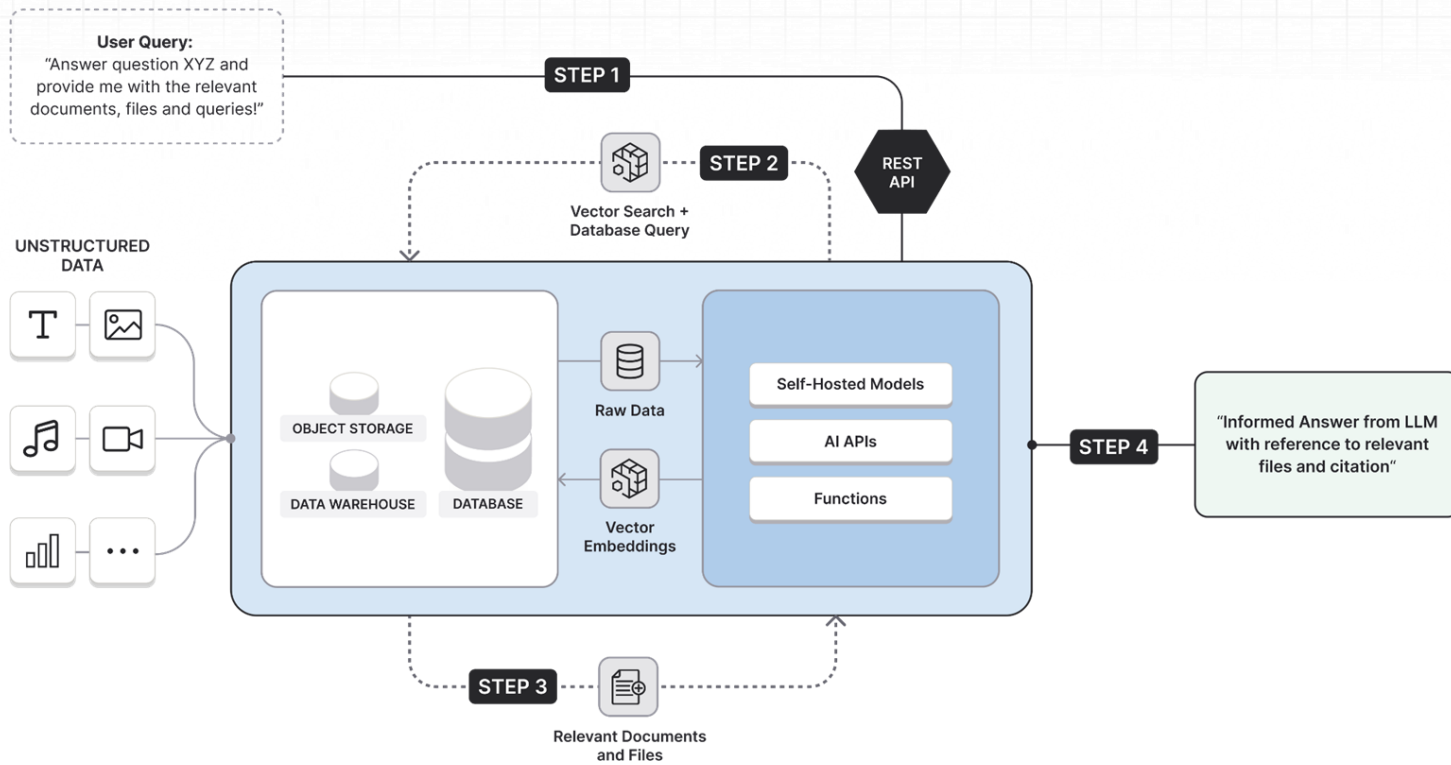
[1.2, -0.3, 0.7, -1.8, ...]



# Vector search data workflow



# Advanced RAG Architecture





## What to consider?

- + **Customization**
- + **Control**
- + **Integration**
- + **Competitive Edge**
- + **Knowledge Retention**
- + **Vendor Lock-in**
- + **Availability**
- + **Data Privacy**
- + **Security**
- + **Costs**

## Problem

Implementing custom  
AI agents and workflows that  
have access to your  
data is a huge challenge

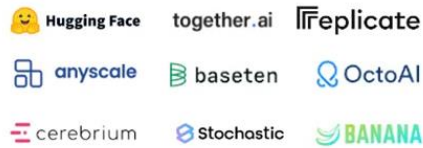


# Open-source AI mapping

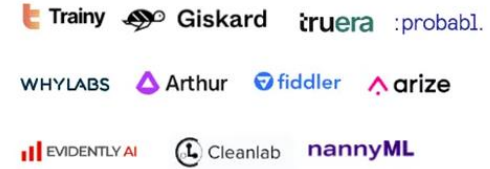
## Generative AI - LLM developers



## LLM hosting & deployment



## AI Observability & monitoring



## Generative AI - Image model developers



## Vector databases



## AI application development & integration



## Privacy, governance & risk management



## Platforms to train open source vision AI algorithms



## Synthetic training data



## Bonus : Most popular AI frameworks



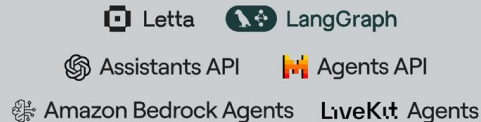
# AI Agents Stack

NOVEMBER 2024

## VERTICAL AGENTS



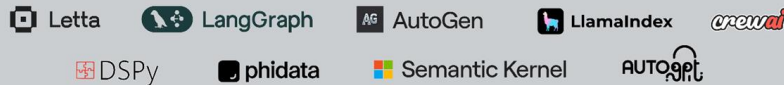
## AGENT HOSTING & SERVING



## OBSERVABILITY



## AGENT FRAMEWORKS



## MEMORY



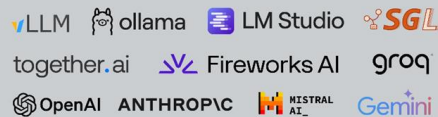
## TOOL LIBRARIES



## SANDBOXES



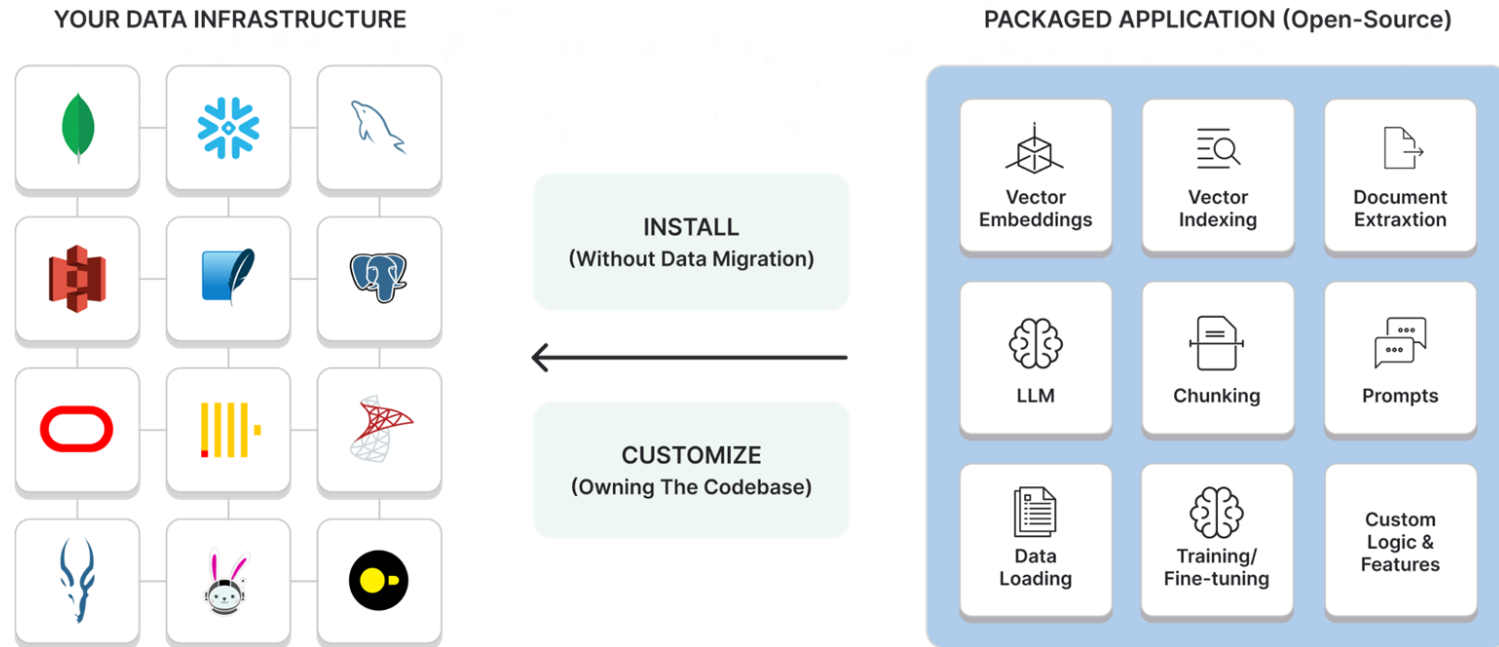
## MODEL SERVING



## STORAGE



# Adopt latest AI today, without migrating your data



## Superduper Installation

# Easily installed and securely hosted on your existing infrastructure, whether cloud or on-prem

### YOUR EXISTING INFRASTRUCTURE



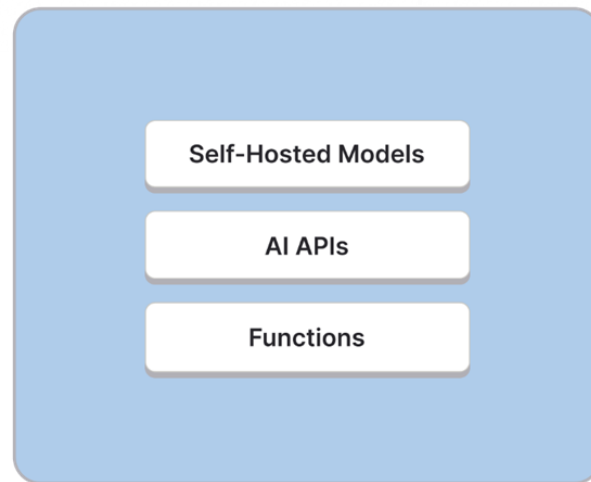
🕒 less than 5 minutes

**INSTALL**  
(Docker, Kubernetes, Terraform)

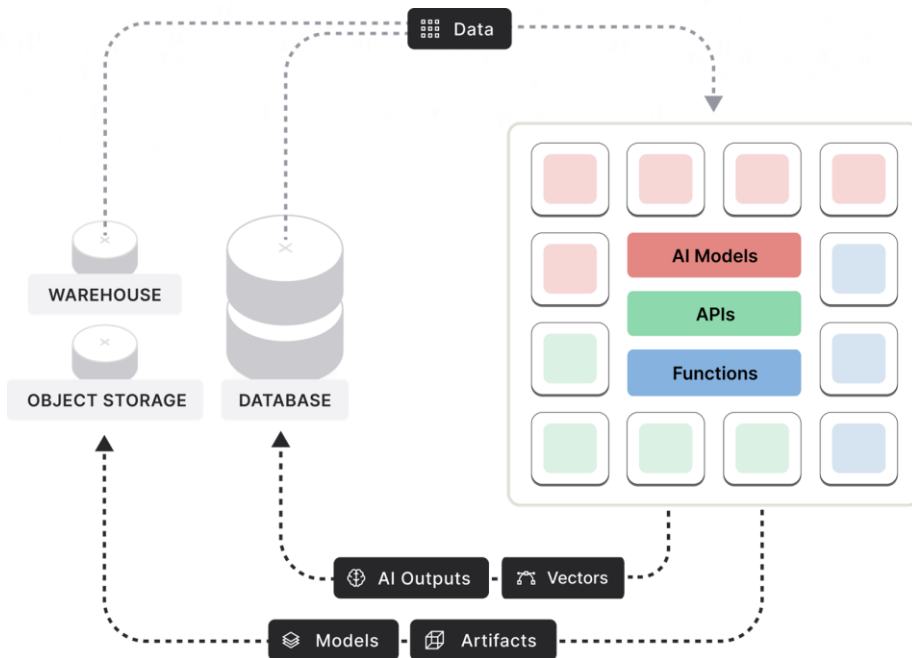


**LEVERAGE**  
(Ray, Grafana, JupyterHub)

### SUPERDUPER PLATFORM + APPLICATIONS



# Eliminate unneeded complexity from your AI-data stack





# Reach out!

[timo@superduper.io](mailto:timo@superduper.io)

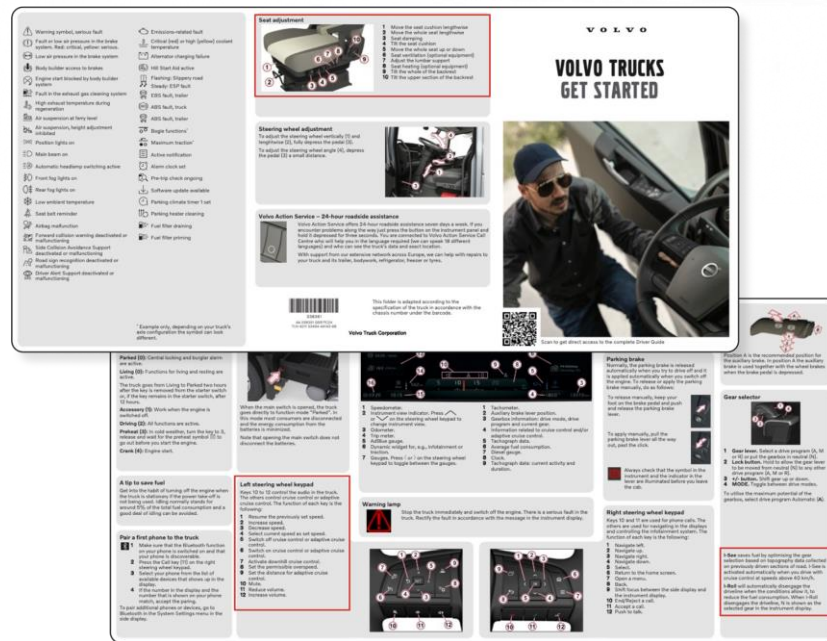


# An AI assistant for Volvo truck drivers

Transforming a static technical document into an interactive expert that answers any question from the document.

This demo showcases:

- **Enhanced document interaction:** revolutionizing the way users engage with user manuals
- **Free-text Q&A:** instantly retrieve any information from the document
- **Accurate citations:** direct references to specific sections, pinpointed with exact coordinates



Demo

# An AI assistant for your legal department

Building an AI assistant that can retrieve information and answer any questions from legal documents locked away in static file types like PDFs.

This demo showcases:

- **Automated information retrieval:** effortlessly extract key details from contracts and agreements
- **Data transformation:** turn unstructured data into structured data for downstream use
- **Automated alerts:** receive notifications for key events or approaching deadlines

