

# Unlocking Documentation with AI-Powered Chatbots: **A Real-World Case Study**

Christian De Simone & Luca Giudici

# YOUR WAY INTO THE FUTURE

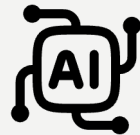
WHO IS MEMORIA

INTR3



WHO IS MEMORIA

# What we do



PROJECTS



KNOWLEDGE



WHO IS MEMORIA

# Projects



CUTTING EDGE &  
TAILOR MADE



NEEDS & REQUIREMENTS  
ANALYSIS



INCREMENTAL  
DEVELOPMENT



COMPLETE PROCESS  
COVERAGE

WHO IS MEMORIA

# Knowledge



WORKSHOPS



ONLINE COURSES



MENTORING



# MACCHINONE











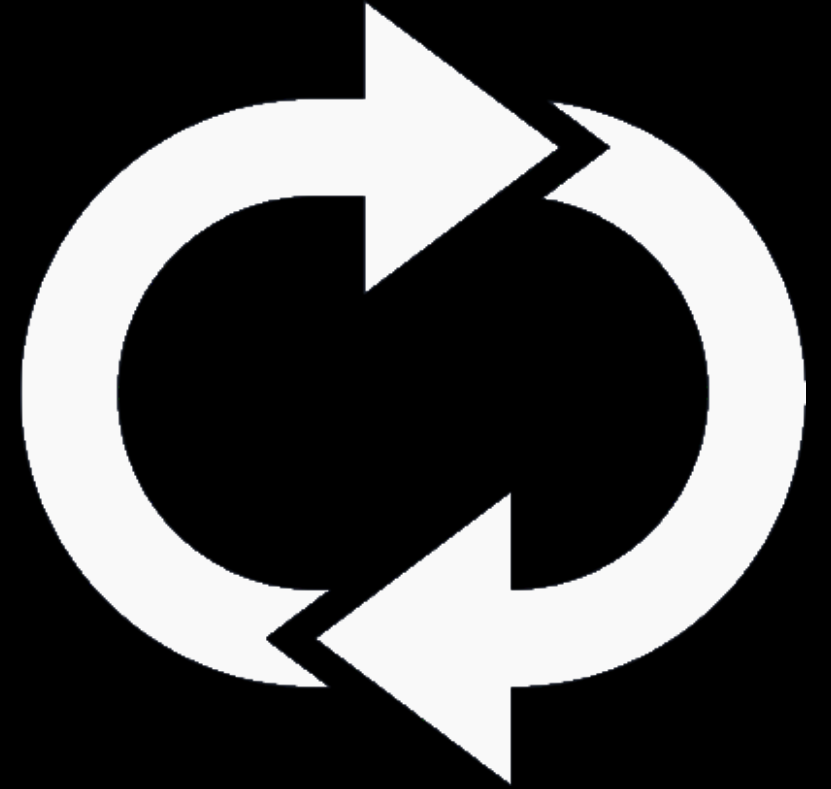


# How did we fix and address such problems with AI?

With an AI assistant that  
answers user's questions

# How we achieved the result

- Agile approach
- Small valuable features
- Continuous feedback
- Incremental deliveries

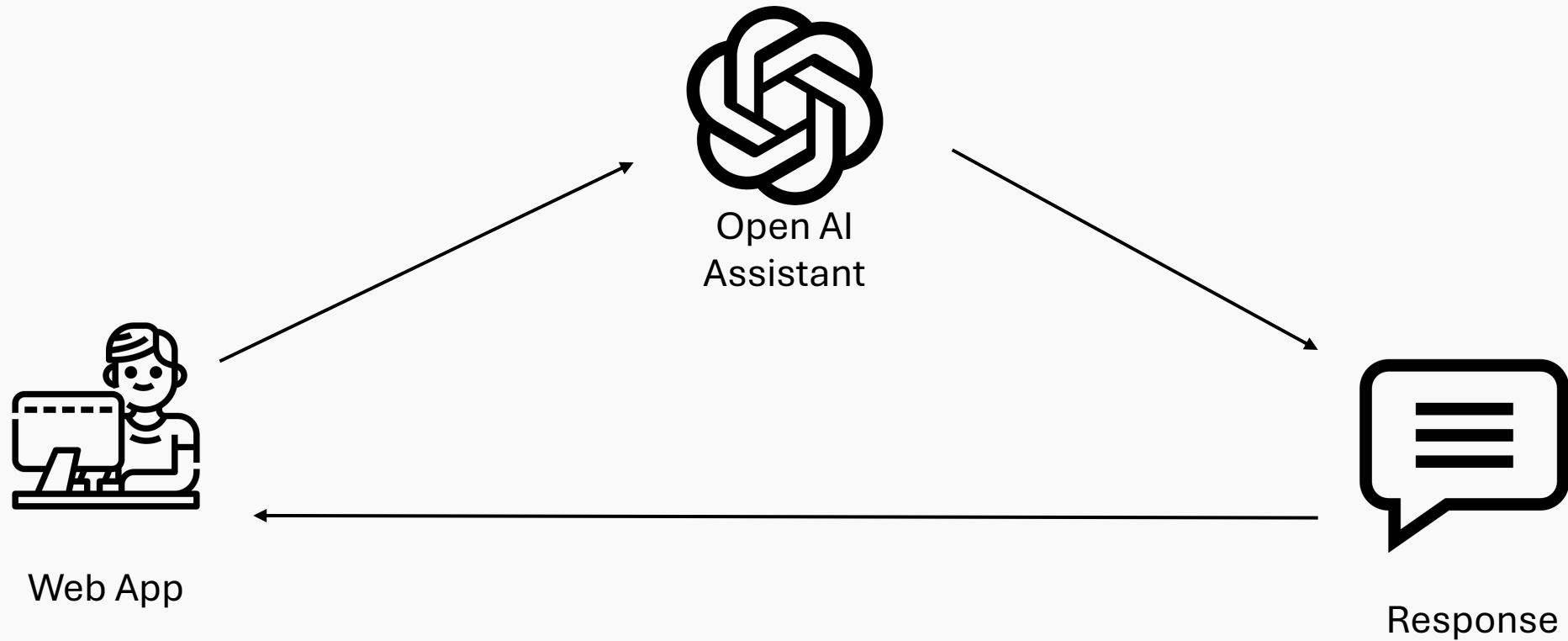




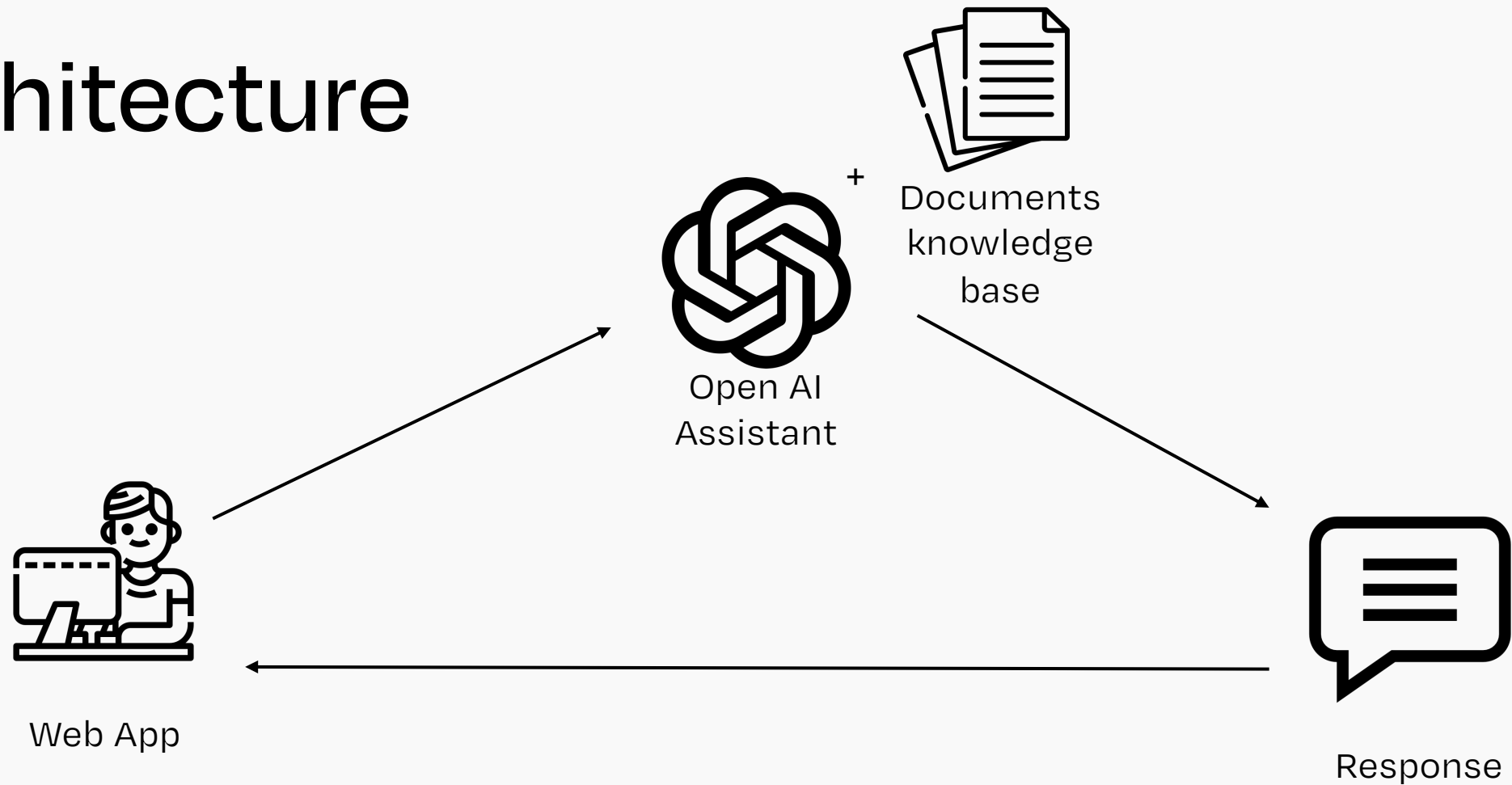
# First Sprint

Building the web app

# Architecture



# Architecture



# Open AI or a custom model ?

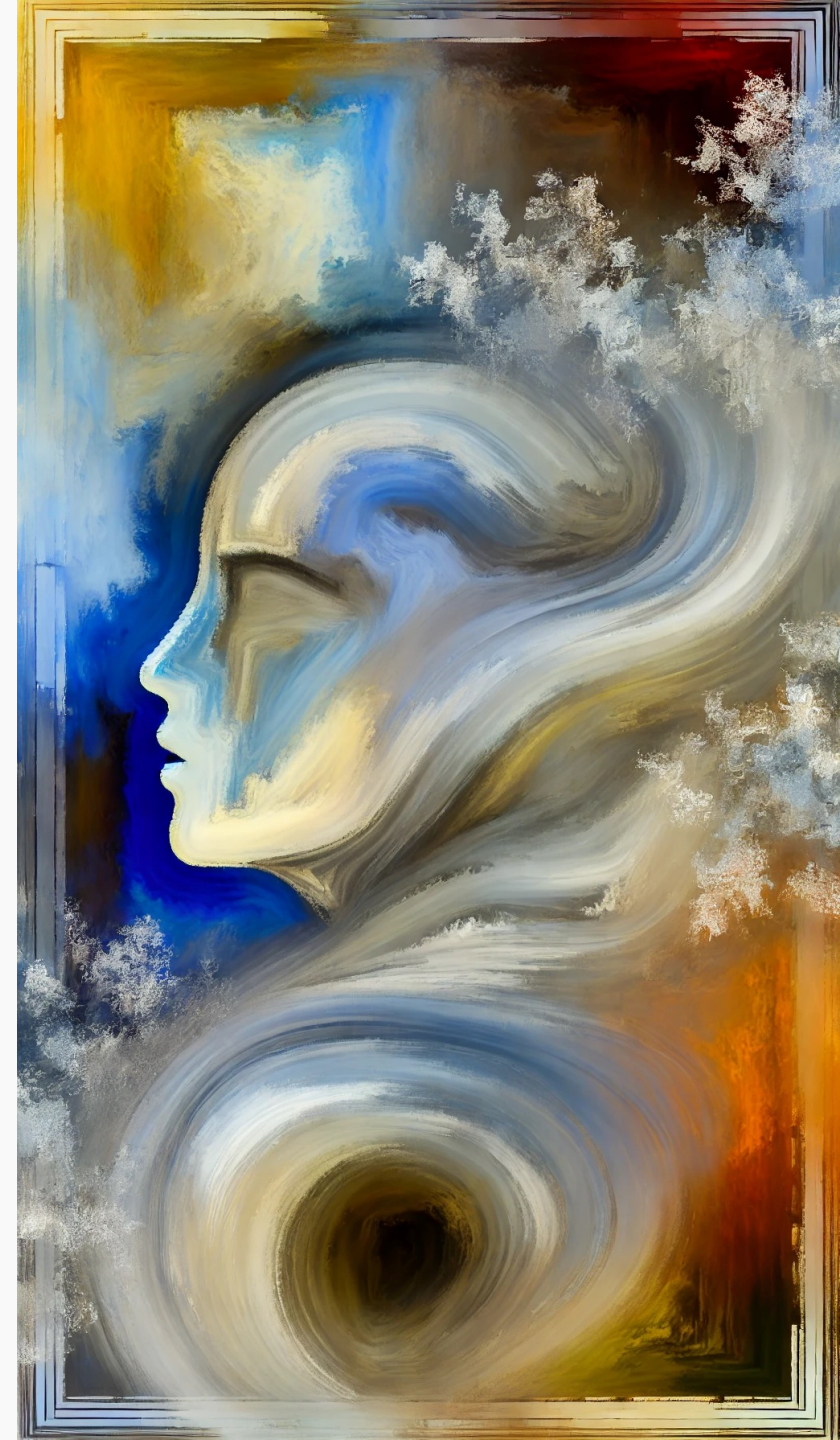
# Open AI or a custom model ?

- Maximize results with minimal investments



# Open AI or a custom model ?

- Maximize results with minimal investments
- State of the art





# Open AI or a custom model ?

- Maximize results with minimal investments
- State of the art
- Able to be swapped out



# First sprint review





# But

Chuck Norris

# Second Sprint

Retrieval of chatbot response document references

# OpenAI message annotations

- 【13†source】

# OpenAI message annotations

- 【13†source】
- File citation

# OpenAI message annotations

- 【13†source】
- File citation
- The response sometimes contained no citation

# OpenAI message annotations

- **【13†source】**
- File citation
- The response sometimes contained no citation
- We needed page references

# We built our vector space

# What is a vector space?

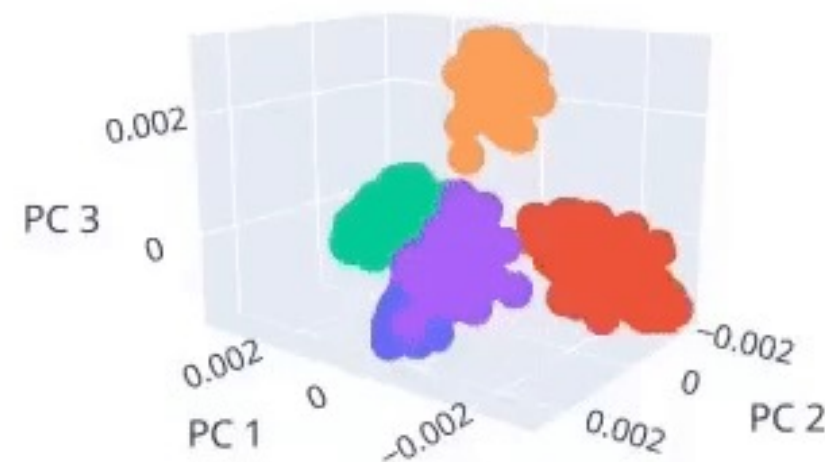


# What is a vector space?



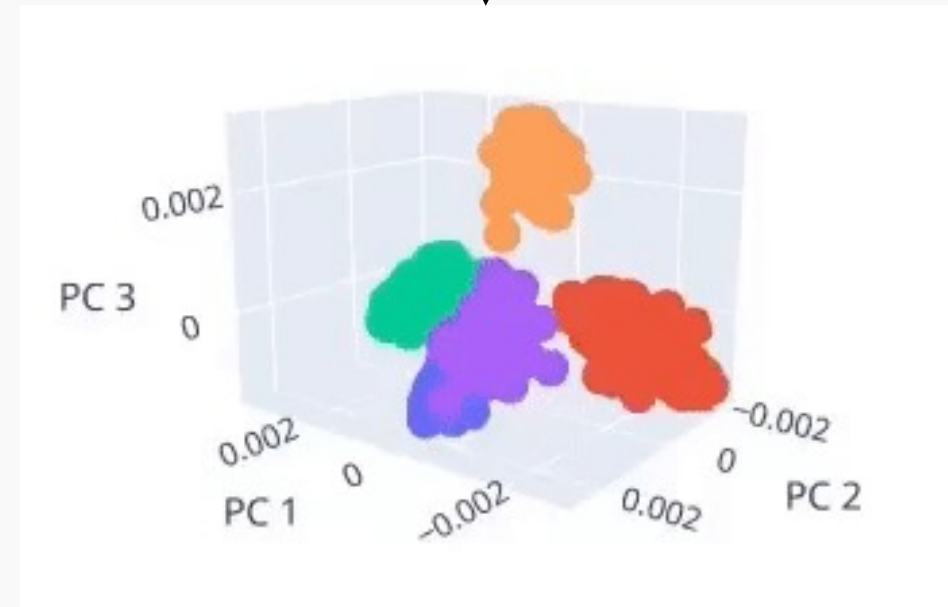
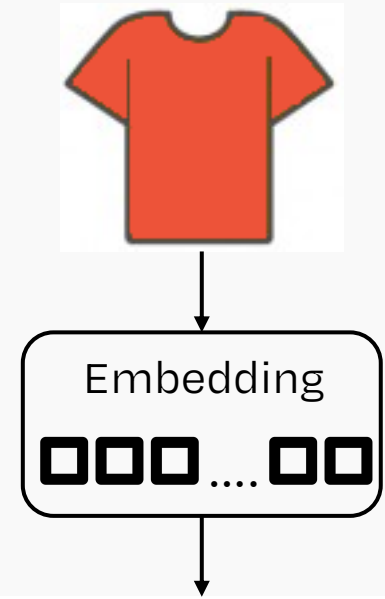
Embedding

□□□...□□



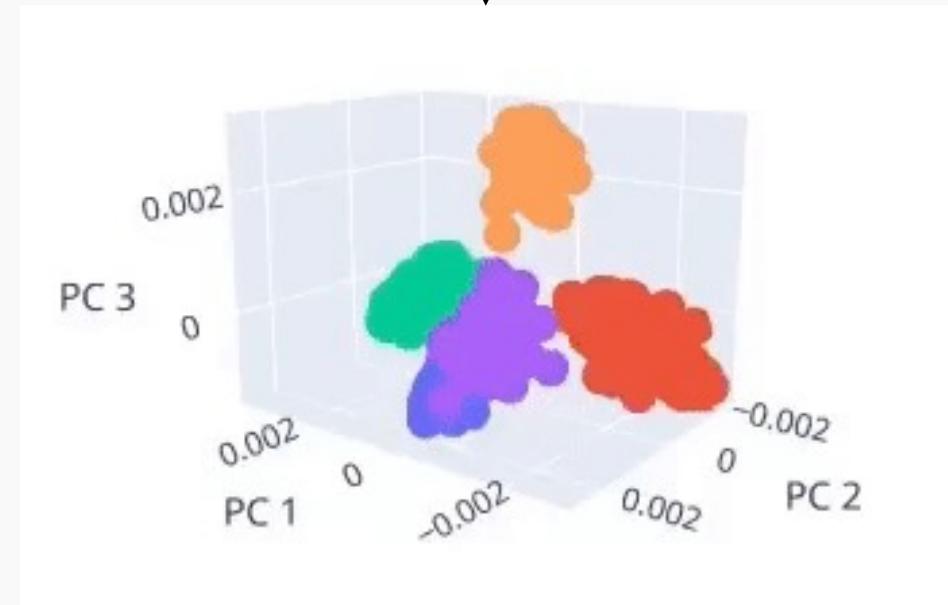
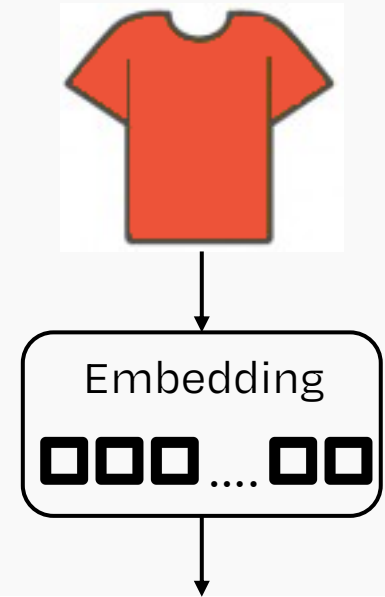
# What is a vector space?

- N dimensional space (depending on the size of the vector)



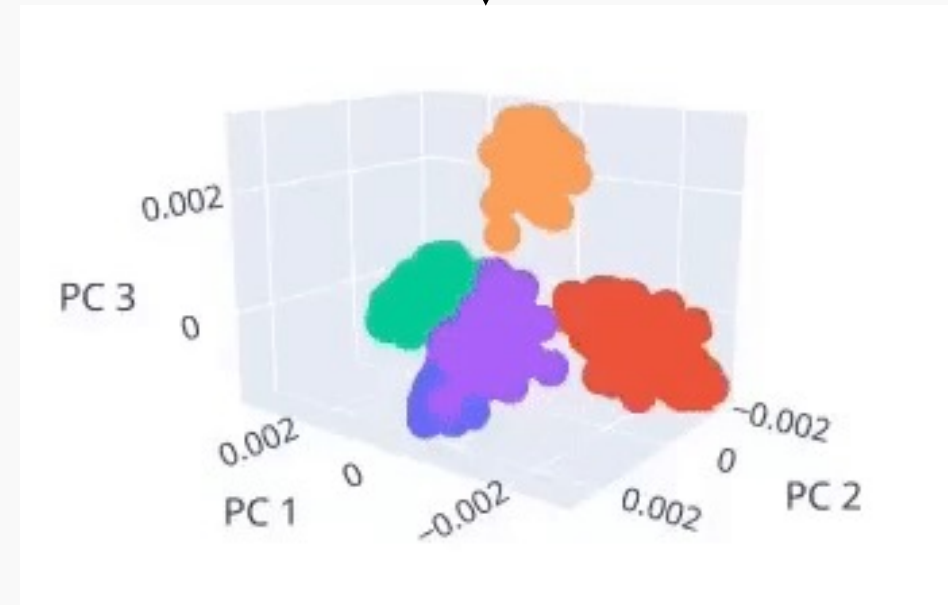
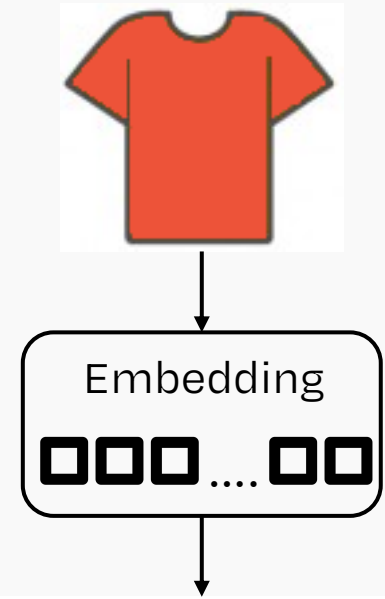
# What is a vector space?

- N dimensional space (depending on the size of the vector)
- Each item is located at a certain coordinate



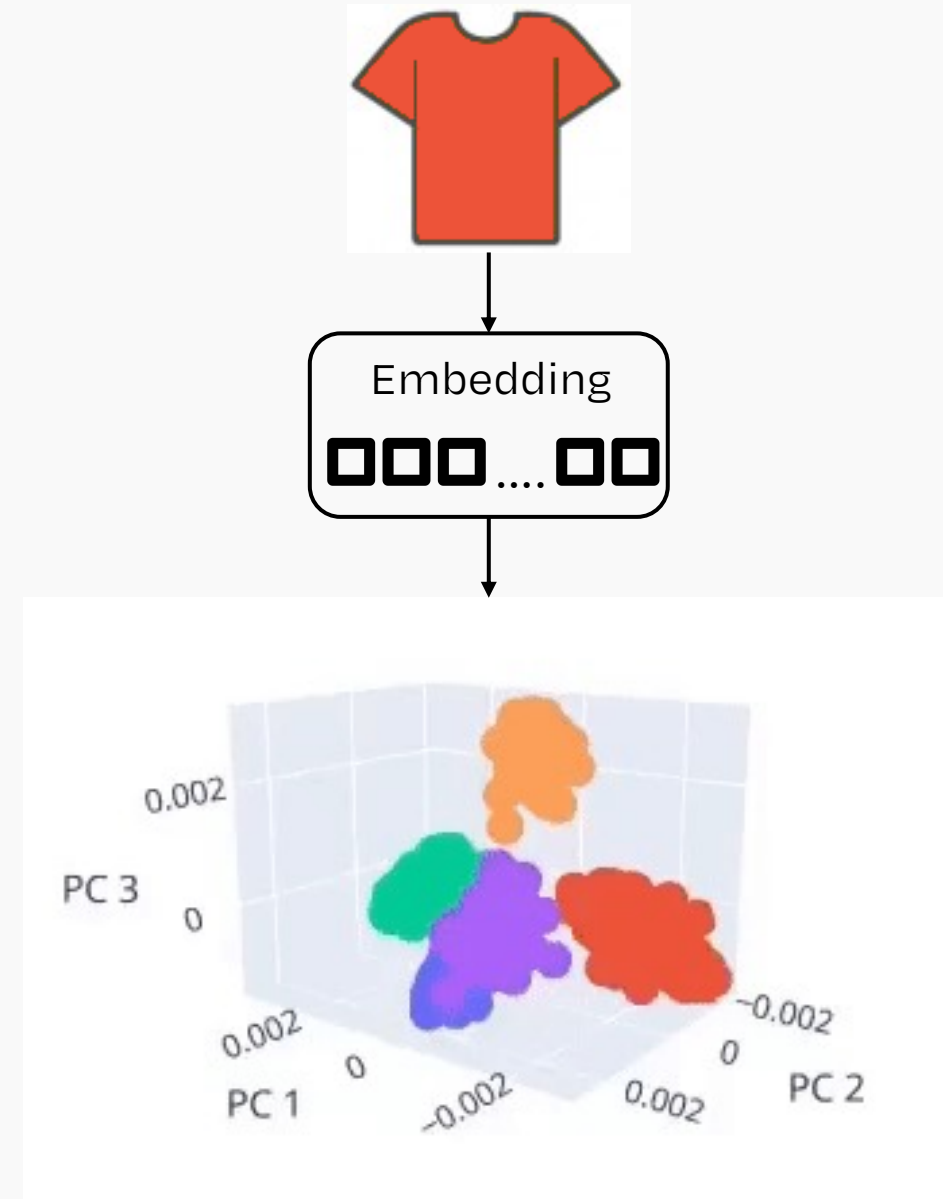
# What is a vector space?

- N dimensional space (depending on the size of the vector)
- Each item is located at a certain coordinate
- Items with similar visual/semantic features also have similar embedding representations



# What is a vector space?

- N dimensional space (depending on the size of the vector)
- Each item is located at a certain coordinate
- Items with similar visual/semantic features also have similar embedding representations
- Huge performances in search

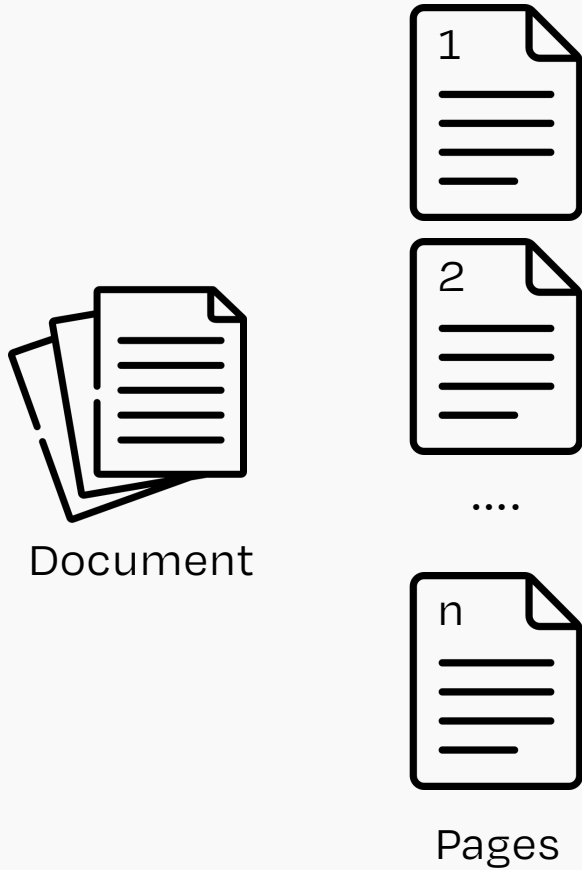


# Our vector space

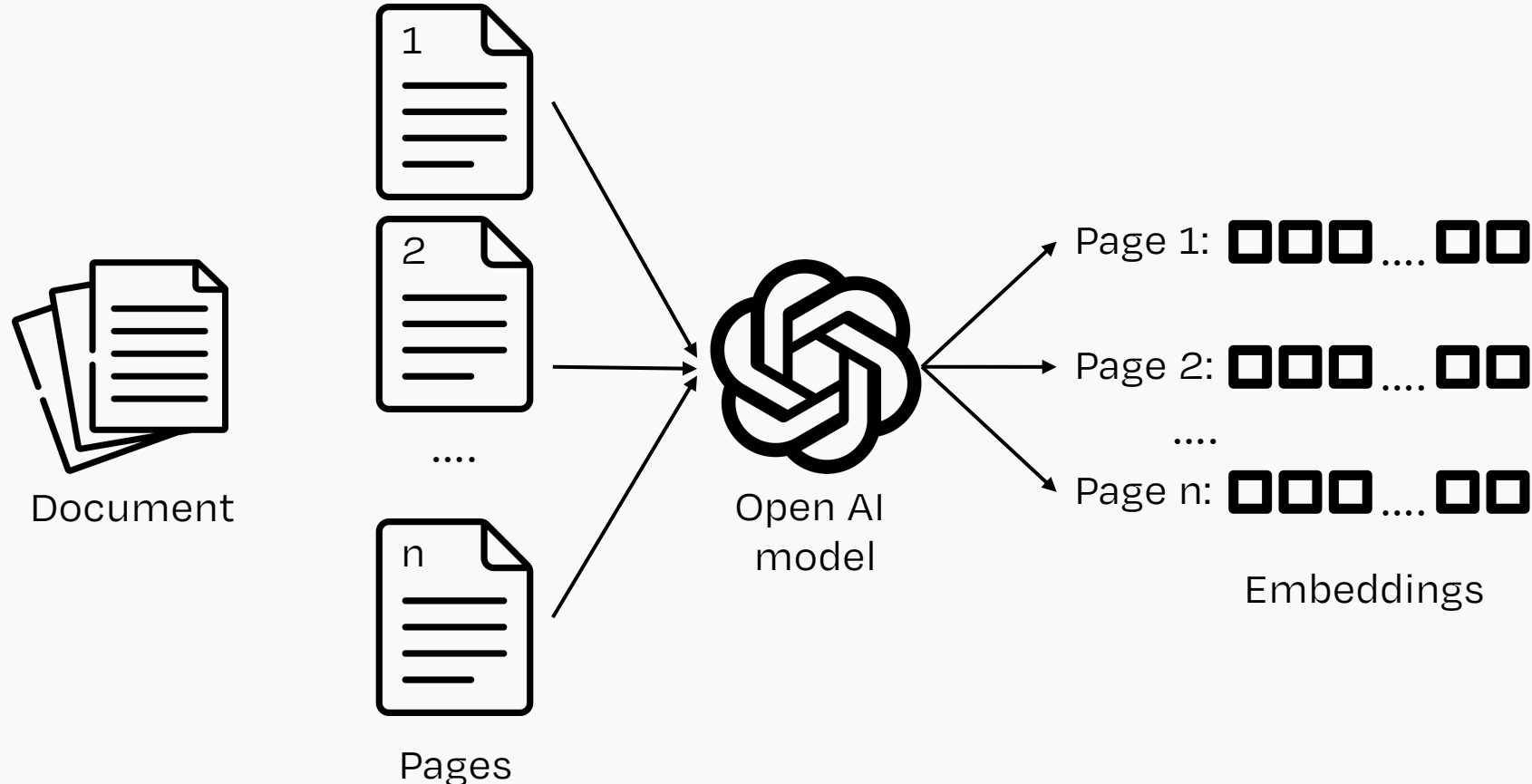


Document

# Our vector space

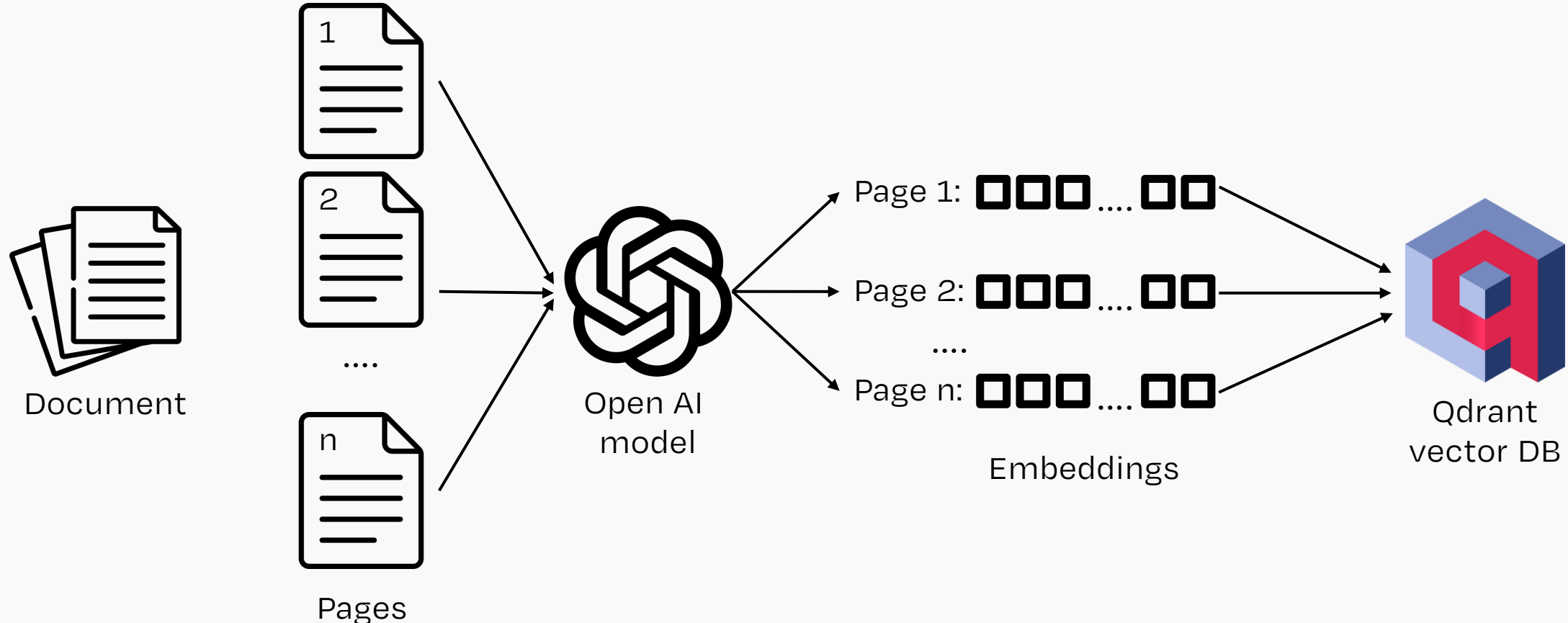


# Our vector space





# Our vector space



# Qdrant

- Vector Database



# Qdrant

- Vector Database
- Performances in vector search



# Qdrant

- Vector Database
- Performances in vector search
- Possibility to choose granularity



# Qdrant

- Vector Database
- Performances in vector search
- Possibility to choose granularity
- Model independent

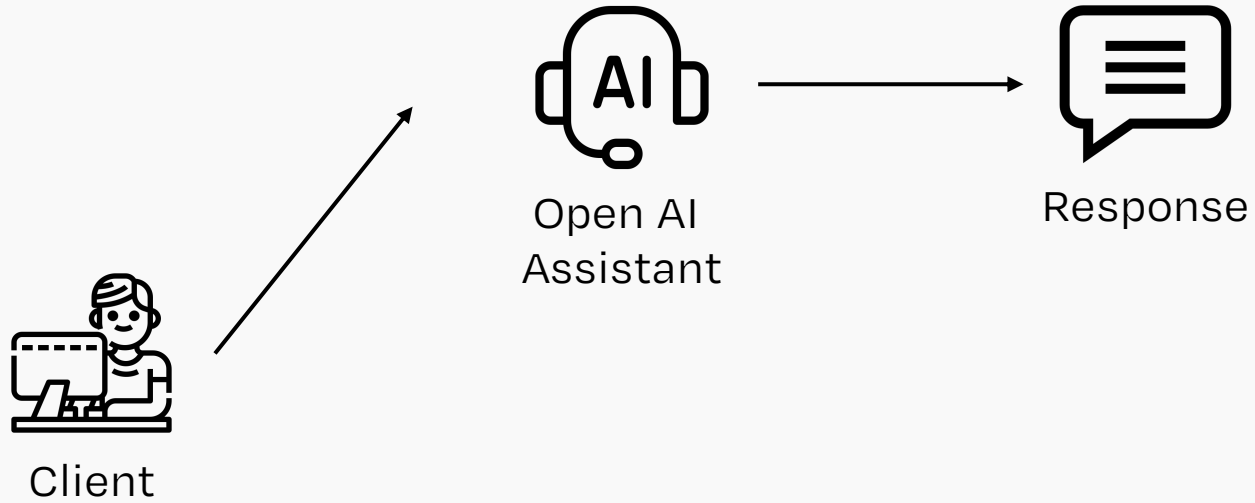


# Qdrant

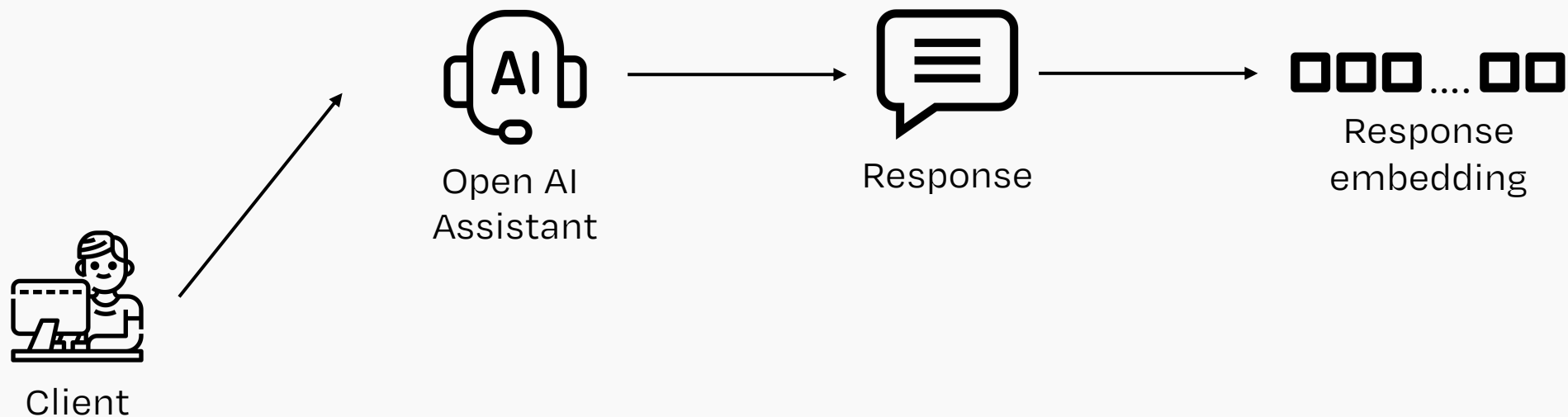
- Vector Database
- Performances in vector search
- Possibility to choose granularity
- Model independent
- Open-source



# Web App

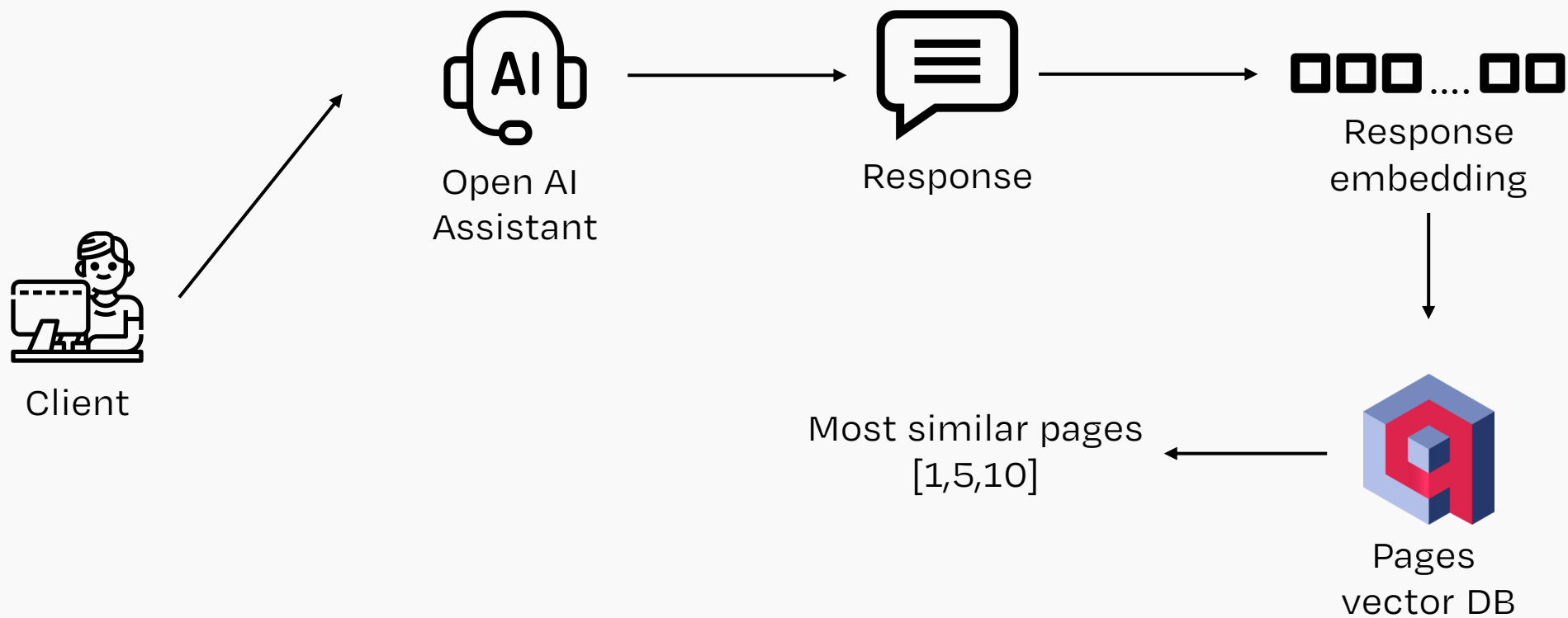


# Web App

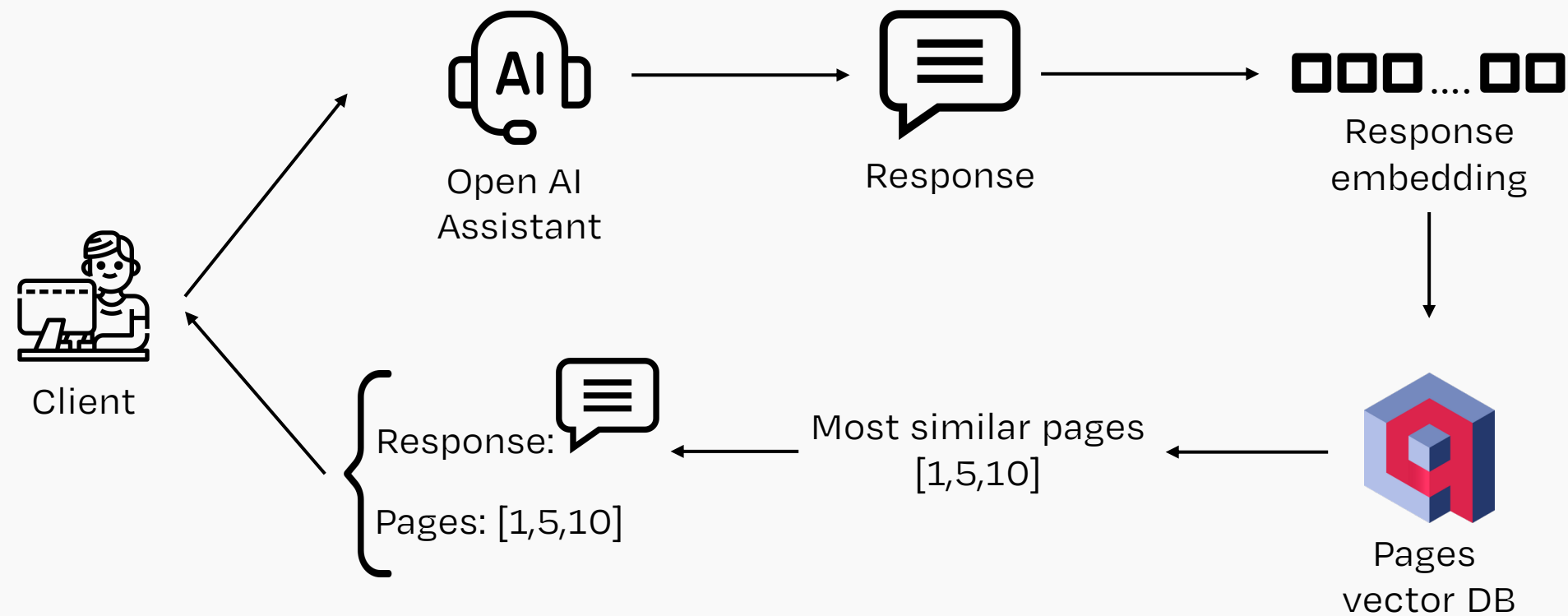




# Web App



# Web App



# Second sprint review

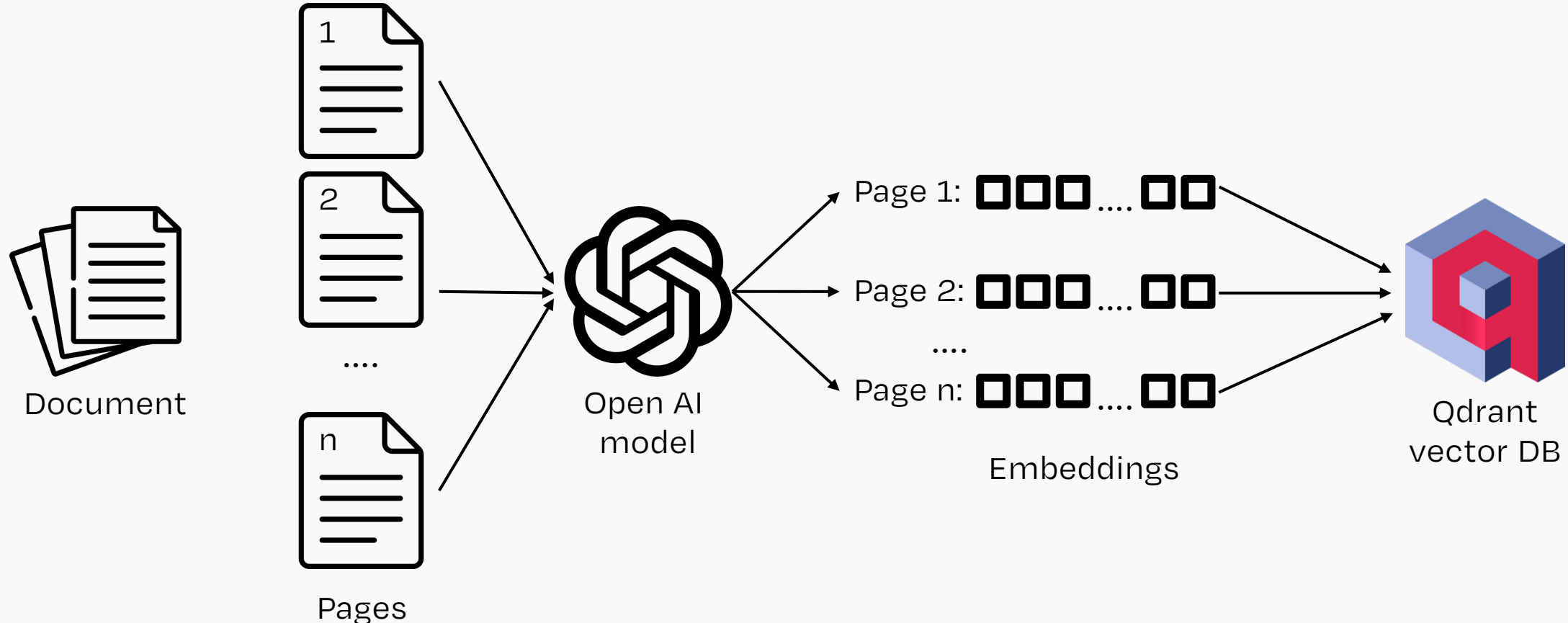
# But

Chuck Norris

# Third Sprint

Automatic generation of vector space

# Our vector space



# Haystack

- Pipeline



# Haystack

- Pipeline
- Multi-model support





# Haystack

- Pipeline
- Multi-model support
- Flexibility



# Haystack

- Pipeline
- Multi-model support
- Flexibility
- Scalability



# Haystack

- Pipeline
- Multi-model support
- Flexibility
- Scalability
- Open-source

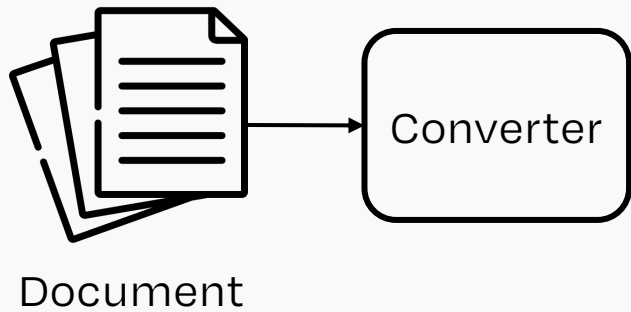


# Haystack pipeline

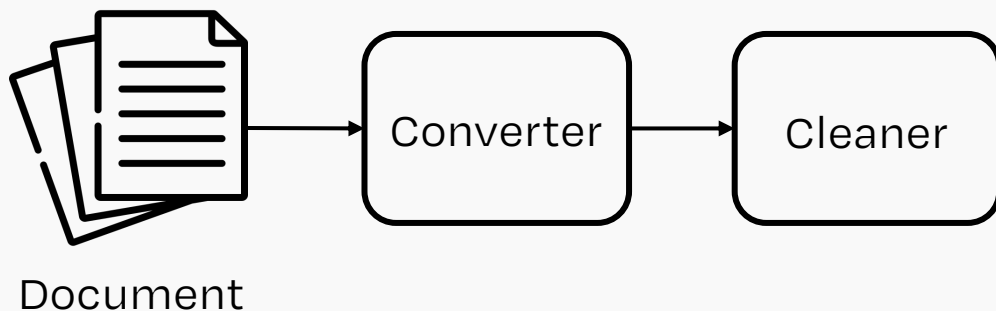


Document

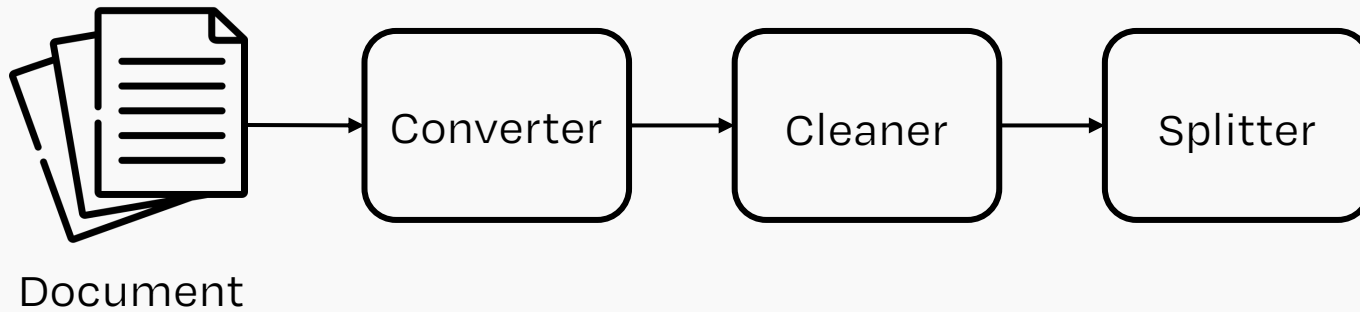
# Haystack pipeline



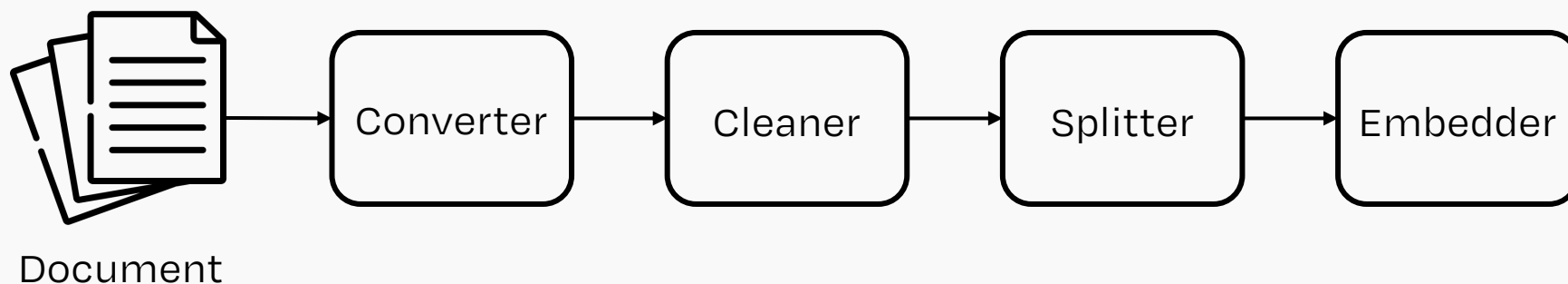
# Haystack pipeline



# Haystack pipeline

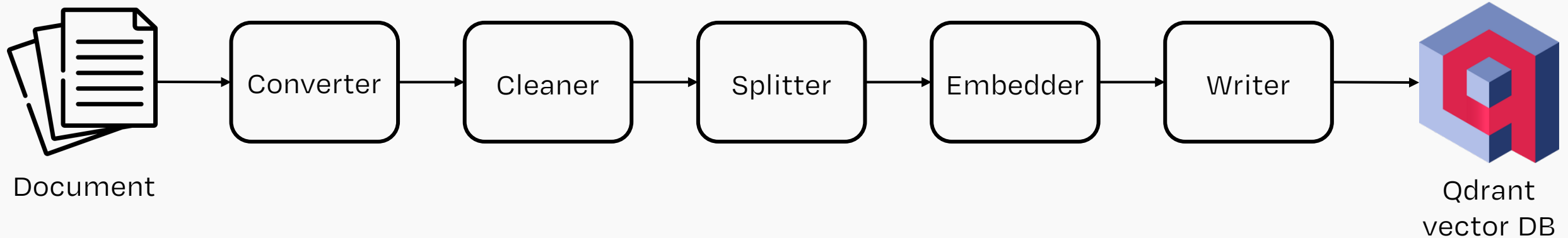


# Haystack pipeline





# Haystack pipeline



# Third sprint review

# DEMO

# YOUR WAY INTO THE FUTURE

# Thanks, Questions?



CHRISTIAN DESIMONE

[christian.desimone@intre.it](mailto:christian.desimone@intre.it)

[349 4462747](tel:3494462747)

LUCA GIUDICI

[luca.giudici@intre.it](mailto:luca.giudici@intre.it)

[340 2557460](tel:3402557460)

SOCIAL

[IG: @Memor\\_\\_ia](#)

[Linkedin: @memor-ia](#)