HealthVision AI

Leveraging Multimodal Large Language Models for Streamlining Health Insurance Claims and Expense Reimbursements

Andrea Lolli
Supervisor: Prof. Matteo Zignani, University of Milan

Academic Year 2023/2024

Bachelor of Science in Artificial Intelligence

Company overview

01. Accenture Overview

- Founded: 1989 as a **technology consulting** division of Arthur Andersen.
- Employees: ~750,000 serving clients in over **120 countries**.

02. Key Services

- **Strategy, consulting, digital, technology, and operations.**
- Expertise in **digital transformation**, **cloud computing**, and **artificial intelligence (AI)**.

03. Recent Investments

- $3 billion investment announced in 2023 to accelerate **AI-driven reinvention**.

04. Awards and Recognitions

- **Fortune's 100 Best Companies to Work For**.
- FTSE Diversity & Inclusion Index.

iOCR

Technology used	**Azure-based solution** used by a primary client for **reimbursement and insurance practices.** Traditional machine learning techniques
Document handling	**Difficulty with diverse formats and handwritten content**
Model complexity	Multiple models complicate analysis and reduce efficiency
Weak points	• **Struggles** with **non-standard document formats** • Significant **issues** with **handwritten text recognition** • Complex system architecture leading to inefficiencies

HealthVision Ai

Technology used	Advanced OCR with Multimodal, Large Language Models (LLMs), and Azure AI Services
Document handling	**Efficient** with diverse formats, including **handwritten content**
Model complexity	Centralized, multimodal approach simplifies analysis and improves efficiency
Strenghts	• **High precision** in processing sensitive and personal data • **Enhanced accuracy and efficiency** in document classification, Q&A, and field extraction • **Reduced complexity** and resource demands in document processing

HealthVision AI

Advanced OCR Multimodal LLM-centric solution leveraging Azure AI Services for streamlining Health Insurance Claims and Expense Reimbursement

OCR

Mutlimodal LLM

Azure AI Services

OCR !

OCR technology

It converts scanned documents, PDFs, or images into editable and searchable data

Document Scan

Paper documents, PDF files or images captured by a digital camera are scanned by the system

Pre-processing

Enhancing the quality of the scanned document or image.

Segmentation

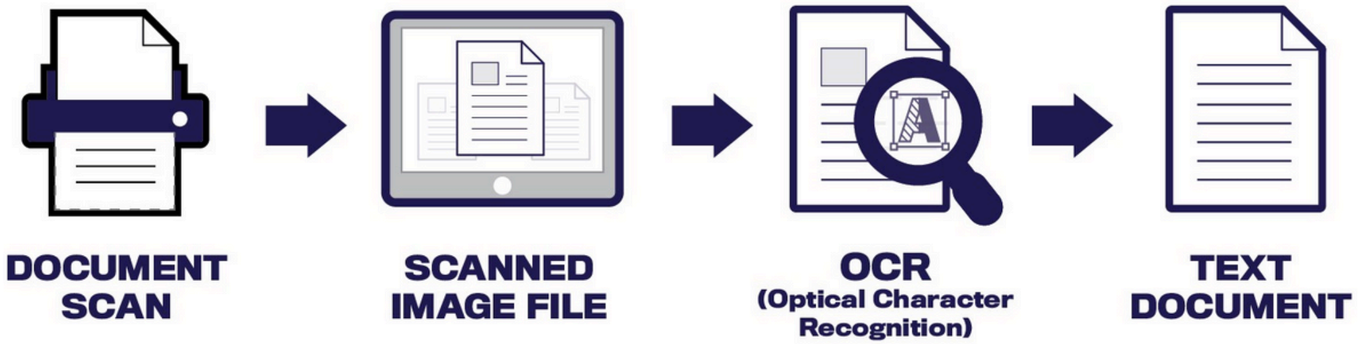
Locating and extracting the text from the document.

Character recognition

Analyzing the extracted text to identify individual characters

Post-processing

Converting the recognized text into a machine-readable format like plain text



LLMS AND MULTIMODALITY

Large Language Models (LLMs) are advanced AI systems that process and generate human-like text. They function like sophisticated autocomplete systems.

- **Training**: LLMs are trained on massive amounts of text data learning patterns and relationships between words and phrases.
- **Prediction:** When given a prompt, they predict the most likely next words based on their training.
- **Generation:** By stringing together these predictions, LLMs can generate coherent paragraphs and even entire articles that appear to be human-written.

Multimodal LLMs are advanced language models that work with various data types, not just text.

This capability allows multimodal LLMs to:

- **Understand the content** of an image and generate captions.
- **Create images** from text descriptions.

In essence, LLMs use deep learning and statistics to create highly accurate language models, enabling them to generate and understand text similarly to humans

Microsoft Azure and Azure AI services for OCR!

Microsoft Azure is a **cloud computing platform** offering various services for building and running applications online instead of on local computers.

Azure AI services provide AI capabilities that developers can **easily integrate** into their applications **without needing to be AI experts.**

Azure AI Vision is a key service that analyzes and understands images. A notable feature is **optical character recognition** (OCR), which extracts printed or handwritten text from images and documents.

Key Features and Benefits

User Accessibility:
No extensive training needed; natural language commands for easy interaction.

Proven Effectiveness:
Optimized for medical documents; tested with real data from an Italian client.

Versatile Solution:
Adapted for diverse document categories; robust and scalable.

Multilingual Capabilities:
Accurate processing in various languages.

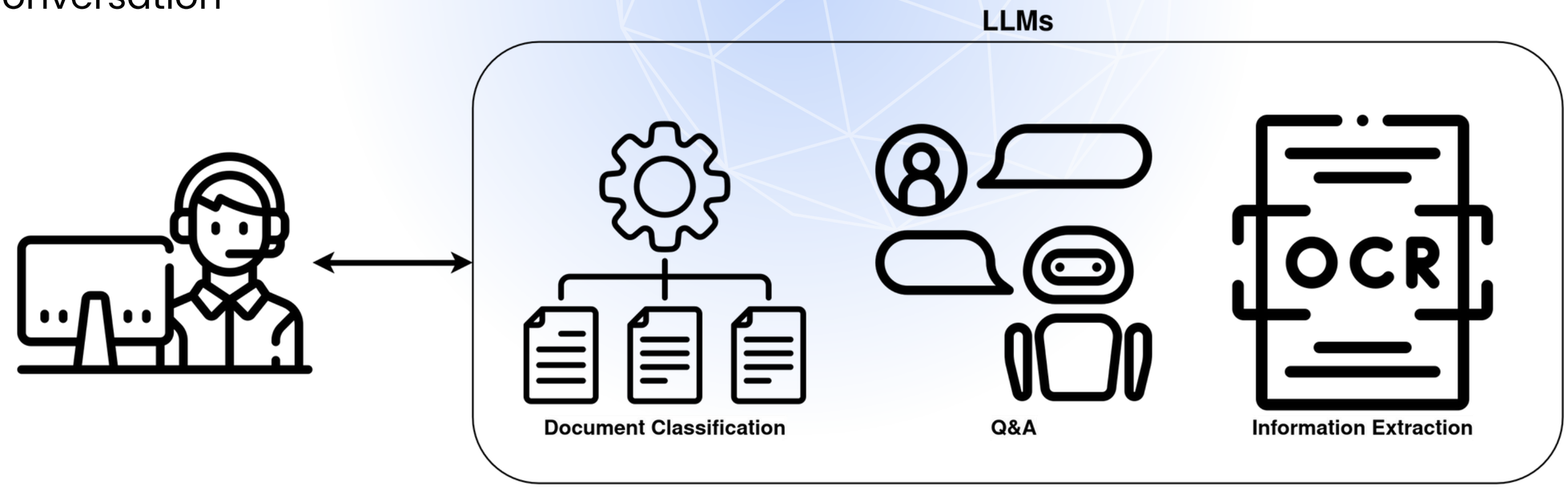
Cost Efficiency:
Improved extraction accuracy; streamlined processes; efficient resource allocation; reduced manual effort

LLM-Centric Approach

Document Classification: The LLM verifies if the chosen category matches the uploaded file, using specific critical points associated with each document type and extracting contextual information.

Q&A: The LLM allows for document content investigation and disambiguation

Information Extraction: The LLM in instructed to extracted field of interest mentioned during conversation



System Architecture

HealthVision AI employs a modular architecture inspired by the **Model-View-Controller (MVC)** design pattern.
The architecture, built for **flexibility**, ensures that the system can be tailored to meet your specific needs by separating the application logic into three mains components:

- **Model:**
  Encapsulates core AI functionalities and business logic for document classification, content investigation, and information extraction using Azure OpenAI's LLMs.
- **View:**
  Represents the user interface built with Streamlit, handling UI rendering, user interactions, and application flow.
- **Controller:**
  Acts as an intermediary between Model and View, processing user inputs, orchestrating interactions, and abstracting AI service complexities.

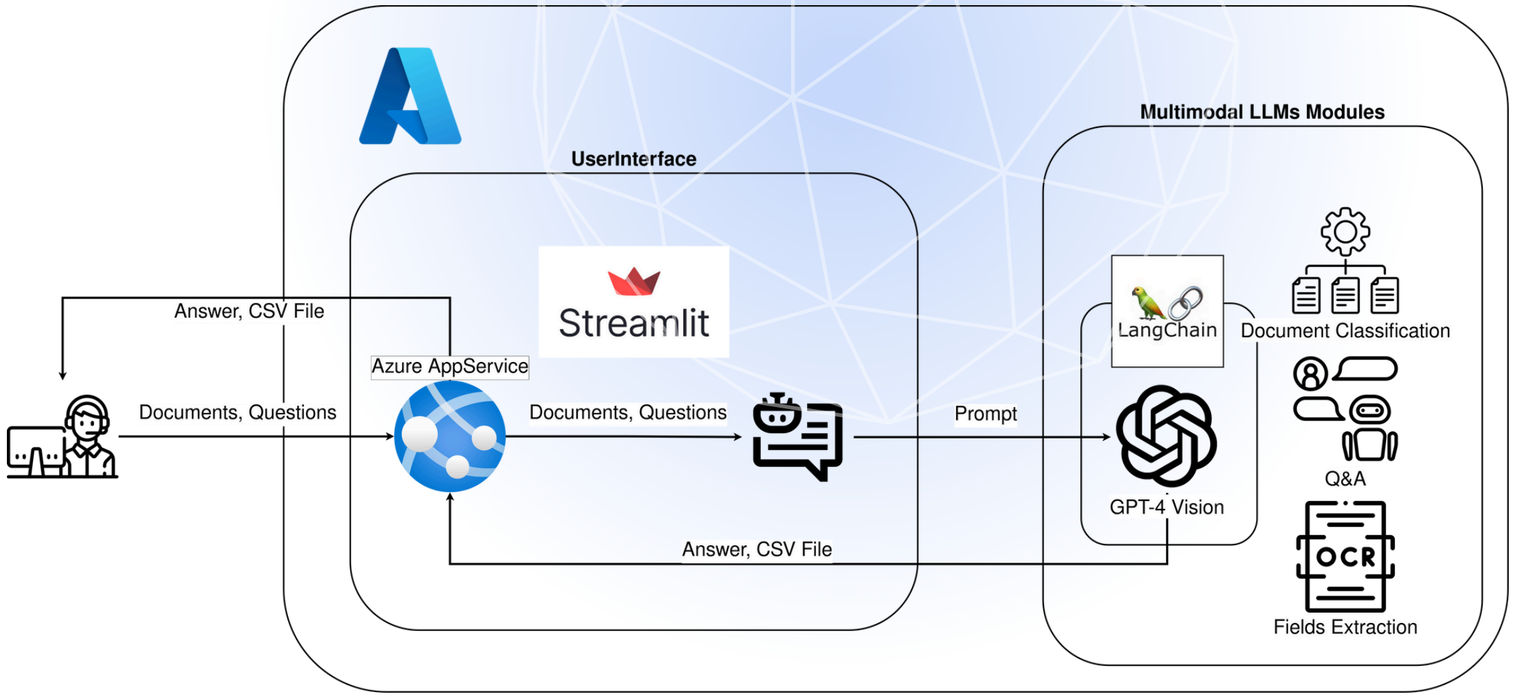
Technology Stack

Streamlit:
Interactive web app, minimal front-end expertise, unified Python development

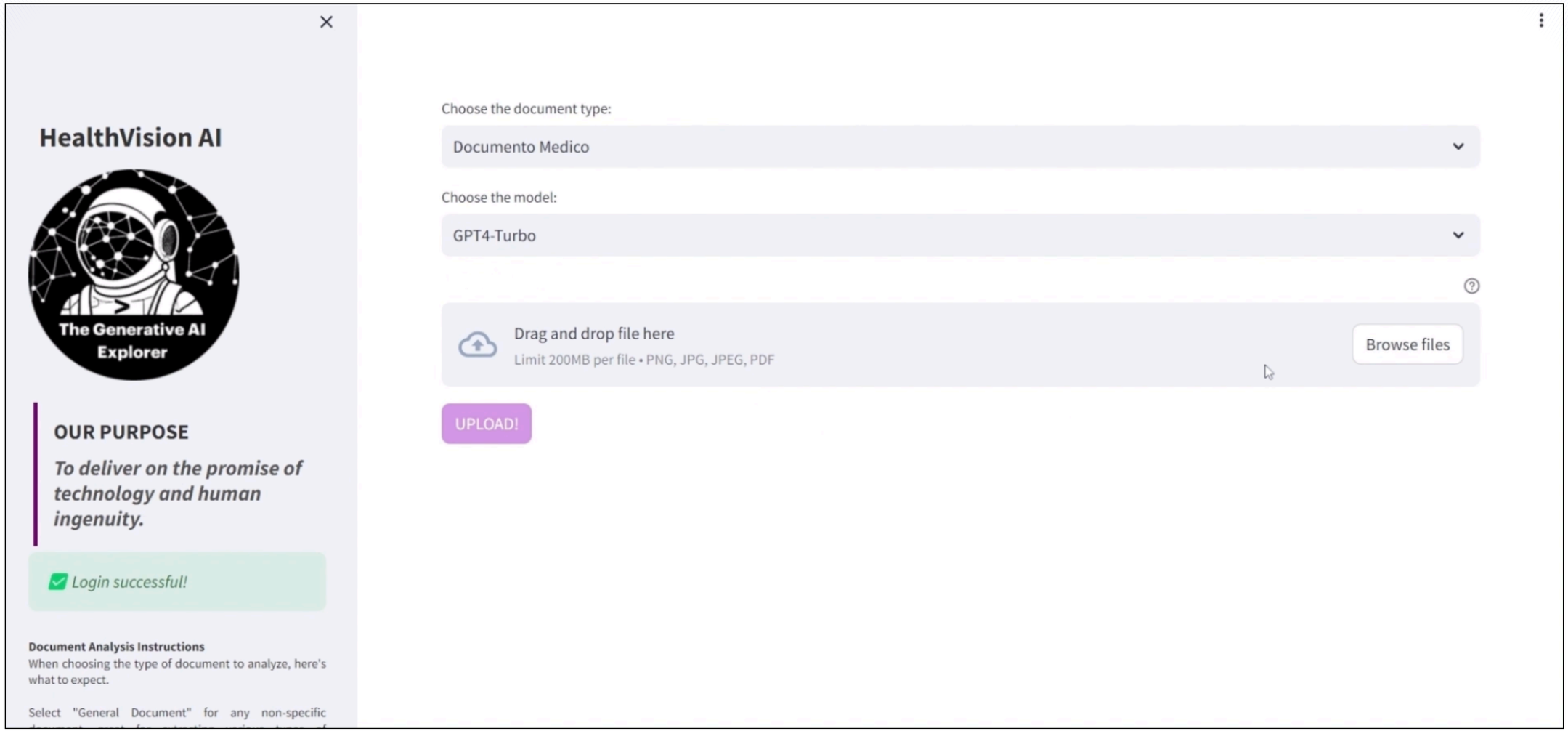
LangChain:
Simplifies LLM integration, modular and scalable, extensible framework

Azure App Services:
Scalable online interface, high availability, global access

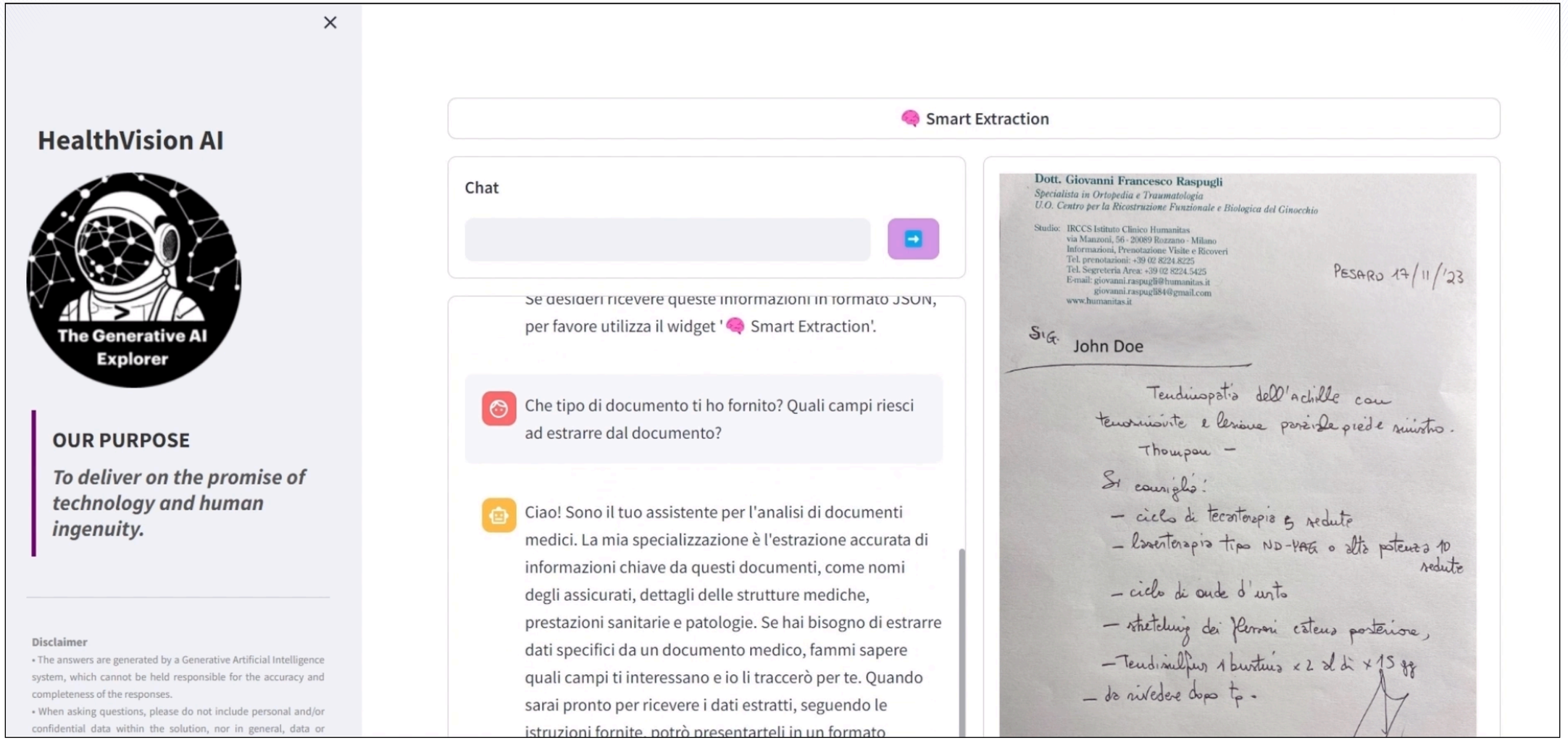
Docker:
Containerized environment, cross-platform compatibility, consistent deployment



Home Page


HealthVision AI
The Generative AI Explorer
OUR PURPOSE
To deliver on the promise of technology and human ingenuity.
Login successful!
Document Analysis Instructions
When choosing the type of document to analyze, here's what to expect.
Select "General Document" for any non-specific
Choose the document type:
Documento Medico
Choose the model:
GPT4-Turbo
Drag and drop file here
Limit 200MB per file • PNG, JPG, JPEG, PDF
Browse files
UPLOAD!

Q&A Page


HealthVision AI
The Generative AI Explorer
OUR PURPOSE
To deliver on the promise of technology and human ingenuity.
Disclaimer
• The answers are generated by a Generative Artificial Intelligence system, which cannot be held responsible for the accuracy and completeness of the responses.
• When asking questions, please do not include personal and/or confidential data within the solution, nor in general, data or
Smart Extraction
Chat
Se desideri ricevere queste informazioni in formato JSON, per favore utilizza il widget '🧠 Smart Extraction'.
Che tipo di documento ti ho fornito? Quali campi riesci ad estrarre dal documento?
Ciao! Sono il tuo assistente per l'analisi di documenti medici. La mia specializzazione è l'estrazione accurata di informazioni chiave da questi documenti, come nomi degli assicurati, dettagli delle strutture mediche, prestazioni sanitarie e patologie. Se hai bisogno di estrarre dati specifici da un documento medico, fammi sapere quali campi ti interessano e io li traccerò per te. Quando sarai pronto per ricevere i dati estratti, seguendo le istruzioni fornite, potrò presentarteli in un formato
Dott. Giovanni Francesco Raspugli
Specialista in Ortopedia e Traumatologia
U.O. Centro per la Ricostruzione Funzionale e Biologica del Ginocchio
Studio: IRCCS Istituto Clinico Humanitas
via Manzoni, 56 - 20089 Rozzano - Milano
Informazioni, Prenotazione Visite e Ricoveri
Tel. prenotazioni: +39 02 8224.8225
Tel. Segreteria Area: +39 02 8224.5425
E-mail: giovanni.raspugli@humanitas.it
giovanni.raspugli84@gmail.com
www.humanitas.it
PESARO 17/11/'23
SIG. John Doe
Tendinopatia dell'Achille con
tendinosi e lesione parziale piede sinistro.
Thompson -
Si consiglia:
– ciclo di tecarterapia 5 sedute
– onde d'urto
– stretching dei flessori catena posteriore,
– Tendinsulfur 1 bustina x 2 al dì x 15 gg
– da rivedere dopo tp -

Extraction Page

HealthVision AI

The Generative AI Explorer

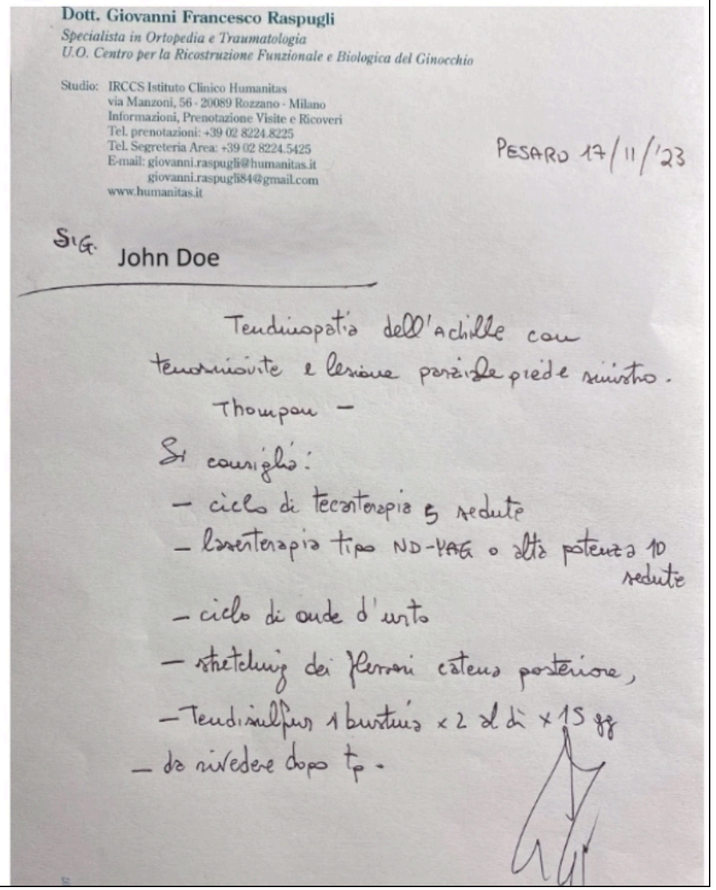
OUR PURPOSE

To deliver on the promise of technology and human ingenuity.

Disclaimer

- The answers are generated by a Generative Artificial Intelligence system, which cannot be held responsible for the accuracy and completeness of the responses.
- When asking questions, please do not include personal and/or confidential data within the solution, nor in general, data or

	0
nome_assicurato	John
cognome_assicurato	Doe
nome_struttura	IRCCS Istituto Clinico Humanitas, Dott. Giovanni Francesc
patologia	["Tendinopatia dell'Achille", 'tenosinovite', 'lesione parzia
prestazione	['ciclo di tecarterapia 5 sedute', 'laserterapia tipo ND-YAG



Dott. Giovanni Francesco Raspugli
Specialista in Ortopedia e Traumatologia
U.O. Centro per la Ricostruzione Funzionale e Biologica del Ginocchio

Studio: IRCCS Istituto Clinico Humanitas
via Manzoni, 56 - 20089 Rozzano - Milano
Informazioni, Prenotazione Visite e Ricoveri
Tel. prenotazioni: +39 02 8224.8225
Tel. Segreteria Area: +39 02 8224.5425
E-mail: giovanni.raspugli@humanitas.it
giovanni.raspugli84@gmail.com
www.humanitas.it

PESARO 17/11/'23

Sig. John Doe

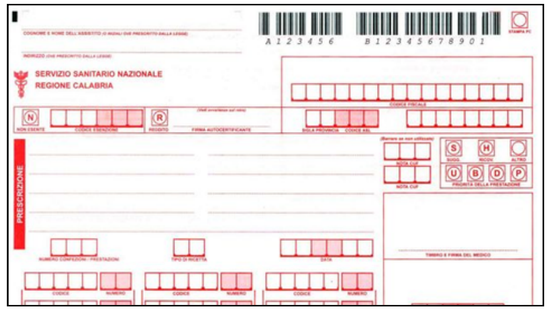
Tendinopatia dell'Achille con tenosinovite e lesione parziale piede sinistro.
Thompson –
Si consiglia:
– ciclo di tecarterapia 5 sedute
– laserterapia tipo ND-YAG o alta potenza 10 sedute
– ciclo di onde d'urto
– stretching dei flessori estensori posteriore,
– Tendisulfur 1 bustina x 2 al dì x 15 gg
– da rivedere dopo tp.

Dataset Preparation

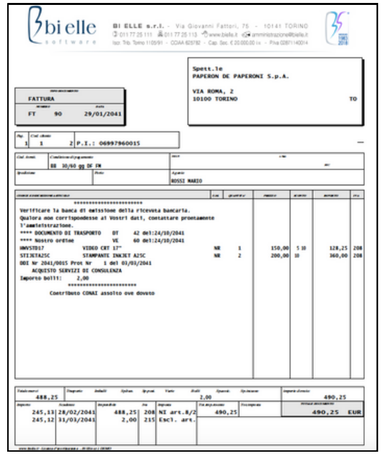
Existing dataset comprising diverse documents obtained from a company's primary client. This ensured that our evaluation was based on **real-world data** encountered in **healthcare** and **insurance** contexts, a factor that directly relates to the practical application of our work in your industries.

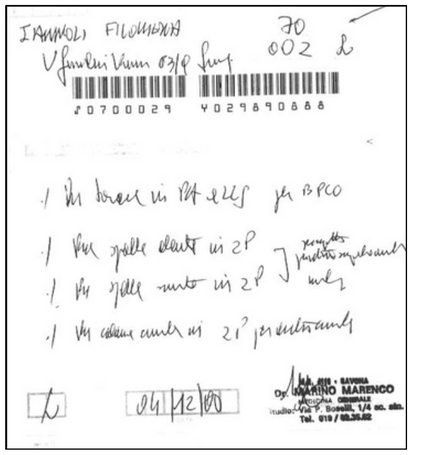
DATASET

HANDWRITTEN INFORMATION DOCUMENTS
red and white prescriptions and reports



COMPUTER-TYPED INFORMATION DOCUMENTS
tickets, invoices, receipts and policy documents





Performance Comparison

Performance measure: Accuracy

Accuracy improvement:

- Computer-typed: **18%**(avg) increase
- Hand-written: **6%** (avg) increase

Computer-typed	Name	Surname	Date	Doc. number	Amount	Structure Vat	Structural Fiscal Code	Structural Name
HealthVision AI	100%	96%	100%	84%	96%	94%	94%	42%
iOCR	80%	77%	80%	71%	68%	73%	91%	25%

Hand written	Name	Surname	Date	Doc. number	Amount
HealthVision AI	83%	78%	82%	76%	62%
iOCR	68%	70%	77%	71%	64%

GPT-4 Vision API Costs

Token-based Pricing:

- Input Tokens: $0.01 per 1,000 tokens
- Output Tokens: $0.03 per 1,000 tokens

Azure AI Services Costs

Transaction-based Pricing:

- $1.5 per 1,000 transactions

Each document processing cycle:

- N+2 transactions (N = number of queries)

CONSOLIDATED COST FORMULA:

$$\$\ \text{Call} = [1406 + 1495(N + 1) + 50(N^2 + 2N)](10^{-5}) + (754 + 50N)(3x10^{-5}) + (N + 2)(1.5x10^{-3})$$

Parameter N: Number of queries in the Q&A phase

Conclusion

Key Achievements and implications for the industry:

- Performance Improvements
- LLM-Centric approach
- User-Centric Design
- Versatility & Scalability
- Efficiency Gains & Error Reduction

Future research directions

- Cost Considerations
- Performance Variability
- Data Privacy and Security

UNIVERSITÀ DI PAVIA

Bachelor of Science in Artificial Intelligence

THANK YOU

for your time and attention

Andrea Lolli

Bachelor of Science
in Artificial Intelligence