



Co-funded by  
the European Union



## HEALTHCHAIN STORIES

### OPHTHALMOLOGY NLP: HARMONISED DATA. ENHANCED DECISIONS AT CLINICAL SCALE

#### THE NEED

Clinical registries and information are dispersed across multiple data sources, created by different teams using various instruments. This results in a high volume of unstructured and unstandardized data that prevents automatic extraction and complicates its use for research or evidence generation.



#### THE HEALTHCHAIN SUPPORT

HealthChain supported Healthcare Organisations in identifying their innovation challenges and selecting companies to address them. They worked closely as an interregional team to co-create, test, and validate a solution aligned with real clinical workflows, patient needs, and organisational constraints. The project provided financial and business support to boost the solution's market-readiness and commercialisation.



## THE SOLUTION

Ophthalmology NLP is a solution that standardizes and harmonizes health data, integrating Natural Language Processing (NLP) that allows the extraction of valuable insights from unstructured data sources, improves clinical decision-making, and streamlines clinical workflows. A key technical advantage is its use of a federated infrastructure that extracts and harmonizes data from various sources (internal hospital systems, external databases, and manual sheets) into the OMOP common data model.



## IMPACT

The ophthalmology NLP-based system generated significant technical and organizational impact by transforming unstructured clinical information into structured, actionable data.

- Developed a robust framework to securely access, extract, harmonize, and analyse data directly from the data owner's infrastructure.
- Successfully trained and deployed an NLP algorithm capable of converting unstructured clinical notes into standardized, usable formats.
- Enabled systematic reuse of clinical data for decision-making, research, and evidence generation.
- Demonstrated the feasibility of implementing advanced AI tools within secure, real-world hospital environments.
- Established a validated technical pipeline supporting future expansion into additional clinical domains.

# OUTCOMES

The pilot successfully evaluated the performance of the NLP-driven matching system, demonstrating measurable progress in structuring and aligning clinical questionnaire data.

- **Solid Semantic Matching Performance:** The NLP model achieved a cosine similarity score of 69.1%, indicating strong alignment between predicted and reference responses.
- **Reliable Numeric Question Matching:** Achieved a 68.4% match rate for numeric-based questions, confirming consistent interpretation of quantitative inputs.
- **High Accuracy in Single-Choice Questions:** Reached an 80.9% match rate, demonstrating strong performance in structured categorical data alignment.
- **Improvement Opportunity in Multiple-Choice Matching:** Achieved a 34.4% match rate in multiple-choice questions, identifying a clear area for further optimization and model refinement.

# SUSTAINABILITY

The long-term sustainability strategy focuses on scaling data partnerships, expanding clinical applications, and embedding the solution within institutional strategy and broader healthcare ecosystems.

- Promptly aims to reach 102 active data partners by 2026 and 276 by 2028, significantly expanding data reach and strengthening the predictive power and clinical value of the NLP models.
- The validated methodology will be replicated in additional hospital departments and extended to new clinical conditions, ensuring scalability and cross-specialty impact.
- Continued joint initiatives with the hospital will build on the pilot's outcomes, reinforcing long-term partnership and iterative innovation.
- Expansion into additional clinical domains will further enhance data standardization and support more consistent, evidence-based decision-making across healthcare systems.

# TESTIMONIALS

---

*“The Ophthalmology NLP solution showed strong potential in transforming unstructured clinical text into structured, usable data. By automating data extraction and supporting PROMs and PREMs collection, the solution reduced manual workload and improved data quality for clinical and research use. The co-creation methodology ensured close alignment with ophthalmology workflows, while maintaining full compliance with data protection and regulatory standards.”*

- OPHTHALMOLOGY NPL Team

