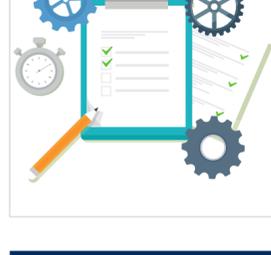


Natural language processing (NLP) in healthcare: Insight generation from unstructured clinical data

What is NLP?

A specialized branch of artificial intelligence focused on the **interpretation and manipulation of human-generated spoken or written data**



How can NLP help us?

- Information extraction
- Unstructured to structured data conversion
- Document categorization
- Summarization

High-potential NLP use cases in healthcare

NON-EXHAUSTIVE

Administrative cost reduction



Efficient billing: Extract relevant information from unstructured physician notes and appropriately assign medical codes to facilitate the billing process



Accurate prior authorization approval: Leverage information from physician notes to alleviate delays and administrative errors

Medical value creation



Effective clinical decision support: Aid members and healthcare providers with decision support at the point of need (e.g., predict post-surgical complications¹)



Streamlined medical policy assessment: Compile and compare clinical guidance from public sources to define the most appropriate care guidelines for care delivery

¹ See example of NLP used to predict diagnoses for specific diseases: Automated Assessment of Patients' Self-Narratives for Post Traumatic Stress Disorder Screening Using Natural Language Processing and Text Mining, March 2017, National Institutes of Health

Examples of NLP approaches and applications

- Doc2Vec: vector representations of clinical documents**
Compare and detect changes in clinical guidelines and lab reports
- Named entity recognition: leveraging unified medical language system (UMLS)**
Extract clinical concepts (e.g., diagnoses, procedures, and symptoms) from electronic medical records, patient discharge summaries, and lab reports
- Sequence to sequence: using stacks of long short-term memory (LSTM) networks**
Map clinical concepts and diagnoses with codified clinical guidelines
- Deep reinforcement learning and contextual bandits: using deep Q-networks²**
Develop human-to-machine natural language instructions (e.g., robot-assisted surgery guided by human instructions, search-oriented conversations)

² Deep Q-networks have applications beyond NLP

Case study: McKinsey used NLP to accelerate benchmarking clinical guidelines

- 1 Unstructured clinical guidelines are used as input data**
Algorithm **aggregates clinical guidelines from numerous sources into a common framework** and loads into the database
- 2 Unstructured text is organized into structured data**
Algorithm **parses the information** for relevant clauses about the clinical appropriateness and coverage of diagnoses **into a structured tabular format**
- 3 Clinical terms are classified into ICD-10 diagnosis codes**
NLP algorithm derives ICD-10 diagnosis codes based on frequency of occurrences in clinical guidelines
- 4 ICD-10 diagnosis codes are aggregated for final output**
NLP algorithm is run at **multiple thresholds for better accuracy³**

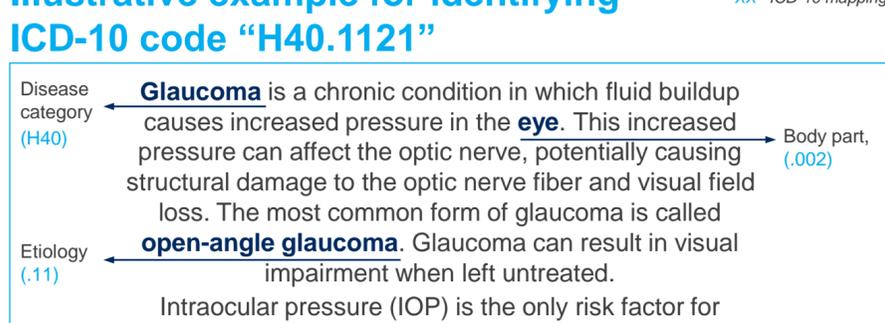
Results of initial prototype process

60% decrease in time required for synthesis of clinical guidelines⁴

³ Manual intervention required to set thresholds and tie outputs with business requirements
⁴ Benchmarked against average time required for similar output using non-automated process

Illustrative example for identifying ICD-10 code "H40.1121"

XX ICD-10 mapping



Text from clinical guidance extract	ICD-10 nomenclature	ICD-10
Glaucoma	Disease category	H40
Open angle	Etiology	0.11
Eye	Body part	0.002
Treatable	Extension	0.0001

= H40.1121 (Glaucoma/Primary open-angle/Left eye/Mild stage)

SOURCE: Multiple public sources on clinical guidelines; International Classification of Diseases, (ICD-10); International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM); expert interviews