TEXT ANALYSIS IN HEALTHCARE & AWS HEALTHCARE SERVICES

Event organized by ADAMAS University Kolkata INDIA

SARBANI MAITI

VP DATA SCIENCE AT VEERSA TECHNOLOGY

EMAIL: SARBANIIIITB2020@GMAIL.COM

CERTIFICATE
OF APPRECIATION

THE FOLLOWING CERTIFICATE IS GIVEN TO

Ms. Sarbani Maity

This certificate is given to Ms. Sarbani Maity
for sharing her valuable knowledge as a guest speaker in the workshop "Five Days International Workshop on Multidisciplinary Applications of Artificial Intelligence for Innovation and Research " which took place on 12.10.2022 at Adamas University

Prof. (Dr.) Sujoy Bhattacharyya
Associate Dean, SOET
Adamas University

Prof.(Dr.) Amlan Chakrabarti
Director, A.K.C.S.I.T
University of Calcutta

TOPICS

- Overview of Text Analysis using NLP techniques
- Text Analysis in Healthcare
- Text Analysis - Technology Trend
- Challenges & Outlook
- Cloud vs Open source platforms
- Demo - Making sense of healthcare data using Text Analysis with AWS
  - Population Health Analytics with AWS HealthLake and QuickSight

  Ref: https://www.analyticsvidhya.com/blog/2022/04/population-health-analytics-with-aws-healthlake-and-quicksight/

  Healthcare Text Analysis with AWS Medical Comprehend

Overview of Text Analysis using NLP techniques


NLP Tasks
Text Preprocessing
Coreference Resolution
Part Of Speech Tagging
Word Sense Disambiguation
Grammatical Error Correction
Feature Extraction
Text-to-Data and viceversa
Speech-to-Text
Text-to-Image
Data-to-Text
Text-to-Speech
Fake News and Hate Speech Detection
Fake News Detection
Stance Detection
Hate Speech Detection
Text Reasoning
Common Sense Reasoning
Natural Language Inference
Chatbots
Slot Filling
Intent Detection
Keyword Extraction
Topics and Keywords
Keyword Extraction
Topic Modeling
Knowledge Bases, Entities and Relations
Entity Linking
Named Entity Recognition
Relation Prediciton
Relation Extraction
Text-to-Text Generation
Paraphrase Generation
Lexical Normalization
Text Simplification
Text Summarization
Text Generation
Machine Translation
Information Retrieval and Document Ranking
Sentence/Document Similarity
Question Answering
Classification
Text Classification
Sentiment Analysis
Image by NLPlanet

TEXT ANALYSIS – NATURAL LANGUAGE PROCESSING

- Natural language processing (NLP) is a machine learning technology that gives computers the ability to interpret, manipulate, and comprehend human language.
- Organizations today have large volumes of voice and text data from various communication channels like emails, text messages, social media newsfeeds, video, audio, and more.
- The NLP software & various technologies automatically process this data, analyze the intent or sentiment in the message, and respond in real-time to human communication.

NLP PIPELINE & NLP TASKS

Other NLP downstream tasks: Token and Text Classifications, Question Answering, Translation, Text Summarization, Sentiment Analysis , Text generation and many more

SENTENCE SEGMENTATION

- break the text block into separate small sentences

London is the capital and most populous city of England and the United Kingdom. Standing on the River Thames in the south east of the island of Great Britain, London has been a major settlement for two millennia. It was founded by the Romans, who named it Londinium. London's ancient core, the City of London, largely retains its 1.12-square-mile (2.9 km2) medieval boundaries.

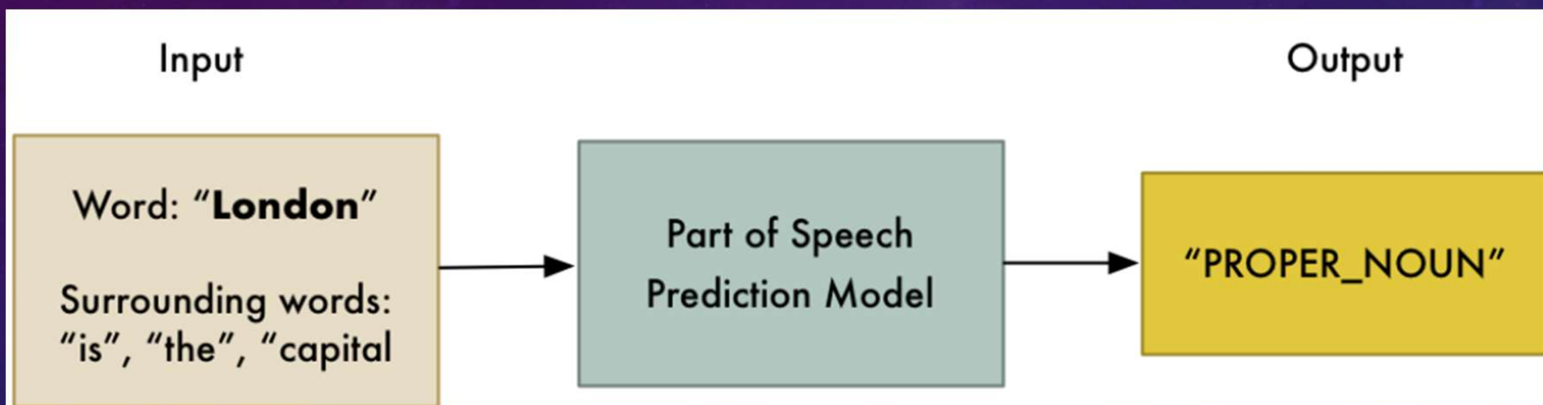
TOKENIZATION OF WORDS

- *“London is the capital and most populous city of England and the United Kingdom.”*

⬇

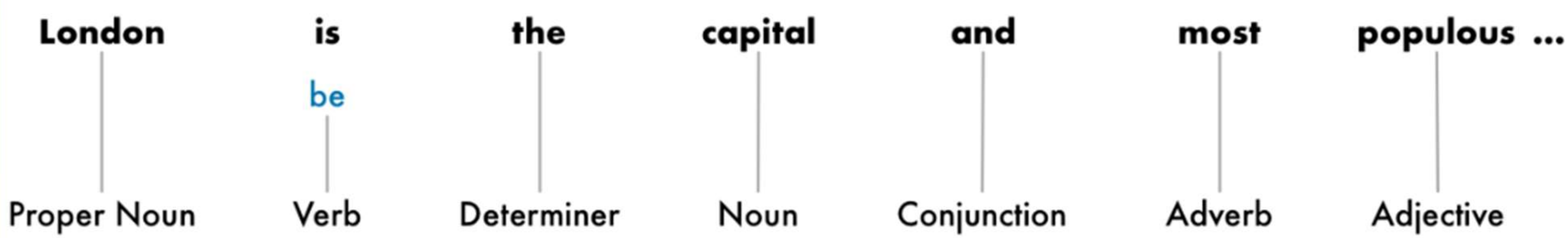
- *“London”, “is”, “the”, “capital”, “and”, “most”, “populous”, “city”, “of”, “England”, “and”, “the”, “United”, “Kingdom”,”.”*

PREDICTING PARTS OF SPEECH FOR EACH TOKEN



LEMMATIZATION OF TEXT

- I had a pony.
- I had two ponies.

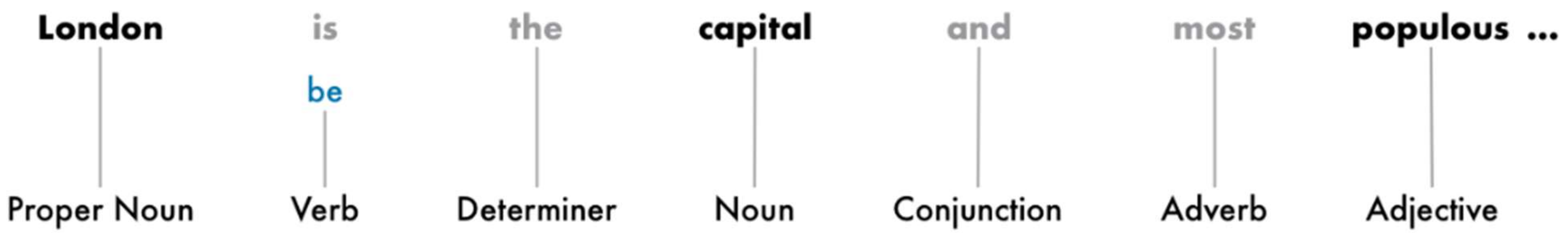


Here is the representation of our sentence, after lemmatization. It adds in the root form of our verb as you can see above. The only change we made here is changing ‘is’ into ‘be’.

➢ Finding out the basic form a.k.a lemma of each word in sentence or string.

RECOGNIZING THE STOP WORDS – NLP WORKS

This is how our text looks after filtering out the stop words in grey:



The basic algorithm to identify stop words just by checking a list of known stop words. But no standard list is applicable for all applications or languages.

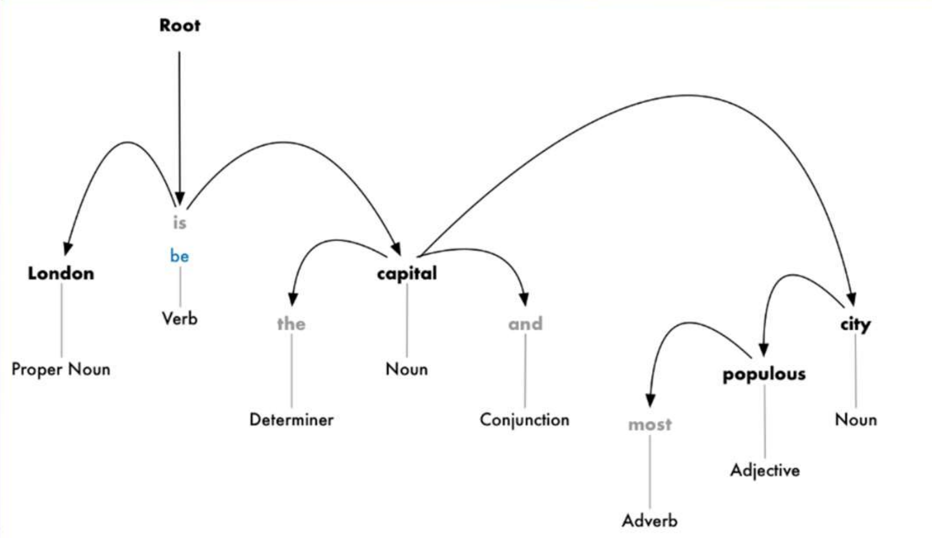
identifying the worth of each word in the sentence. English exploits a lot of filler words. For instance, “and”, “the”, and “a”. When analyzing text, these words introduce a lot of noise.

This is because they appear more frequently than other words. Most of the NLP Pipelines flag them as stop words.

These are the words that we might want to filter out before running any further analysis on the text.

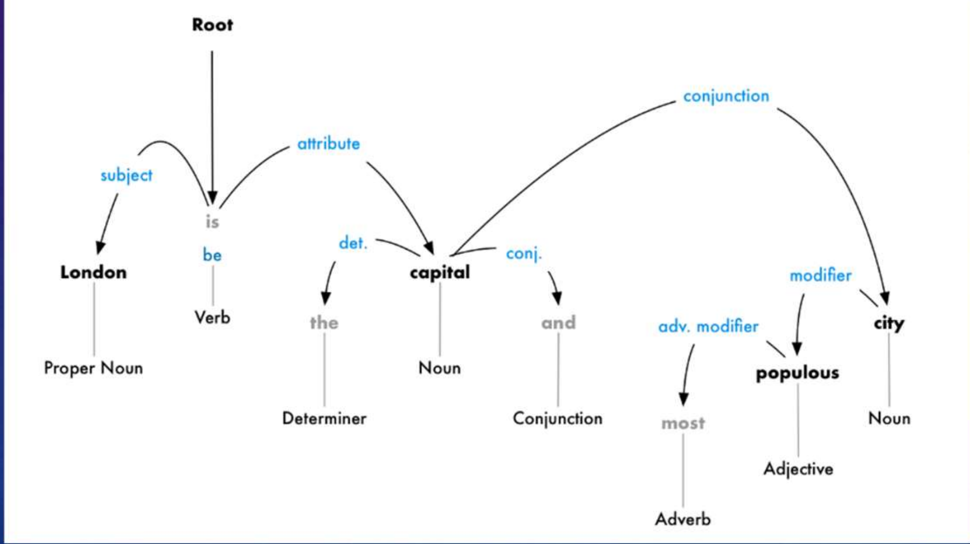
PARSING DEPENDENCY

- how all the words in the sentence depend on each other is our next step. This step is known as dependency parsing. Making a tree assigning a single parent word to each word in the sentence is the goal here. The main verb is assigned as the root of the tree. Here is what the parse tree will look like for our sentence:

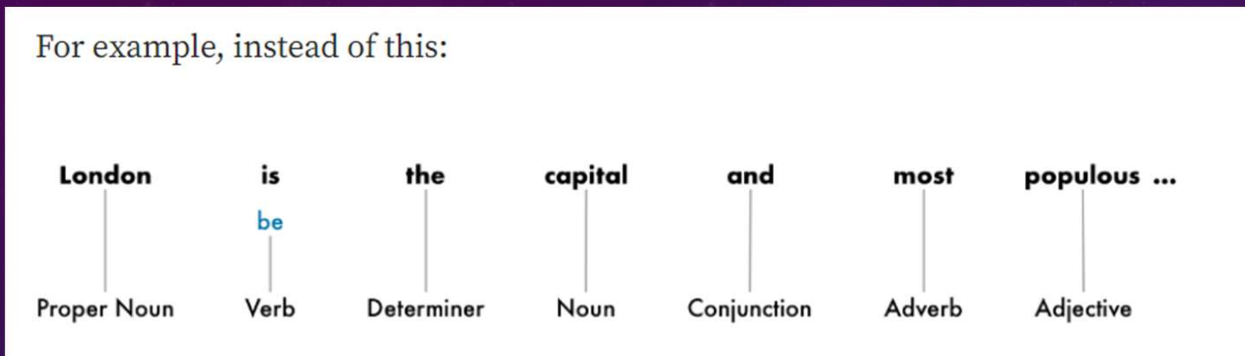


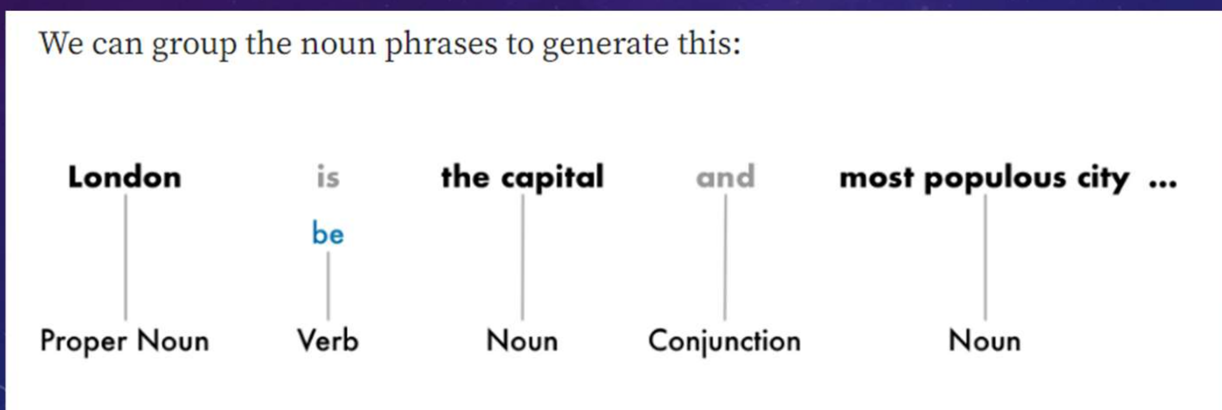
Going a step further, we identify the type of relationship that exist between each word and its parent word.

Going a step further, we identify the type of relationship that exist between each word and its parent word.



FINDING NOUN PHRASES





Sometimes it makes more sense to group together the words that represent a single idea or thing.

We can use the information from the dependency parse tree to automatically group together words that are all talking about the same thing.

NAMED ENTITY RECOGNITION (NER)

In our sentence, we have the following nouns:

London is the capital and most populous city of England and the United Kingdom.

Some of these nouns present real things in the world. For example, *"London"*, *"England"* and *"United Kingdom"* represent physical places on a map. It would be nice to be able to detect that! With that information, we could automatically extract a list of real-world places mentioned in a document using NLP.

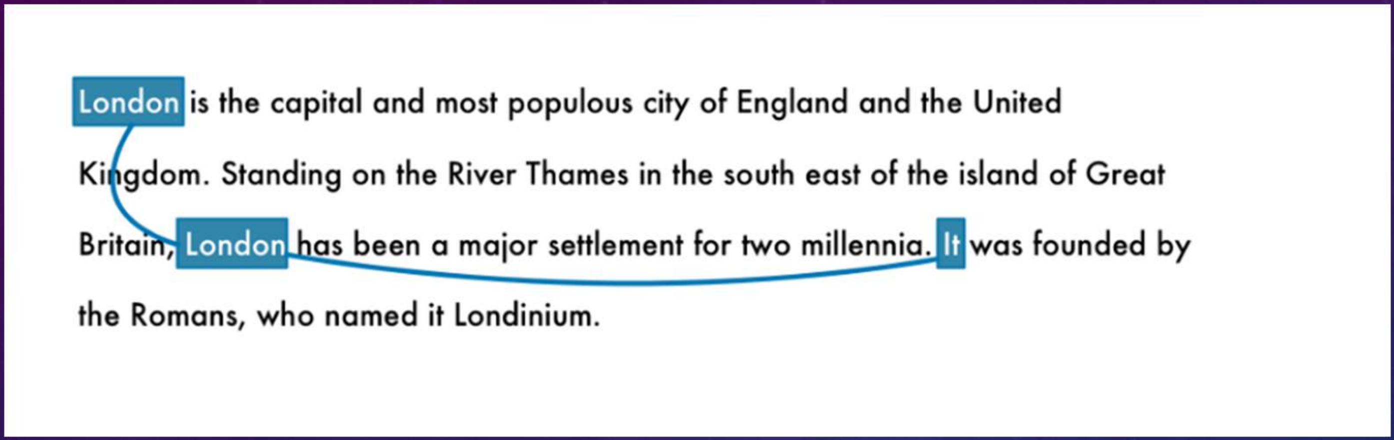
The goal of *Named Entity Recognition*, or *NER*, is to detect and label these nouns with the real-world concepts that they represent. Here's what our sentence looks like after running each token through our NER tagging model:

London is the capital and most populous city of England and the United Kingdom.

London	is the capital and most populous city of	England	and the	United Kingdom
Geographic Entity		*Geographic Entity*		*Geographic Entity*

But NER systems aren't just doing a simple dictionary lookup. Instead, they are using the context of how a word appears in the sentence and a statistical model to guess which type of noun a word represents. A good NER system can tell the difference between "*Brooklyn Decker*" the person and the place "*Brooklyn*" using context clues.

COREFERENCE RESOLUTION

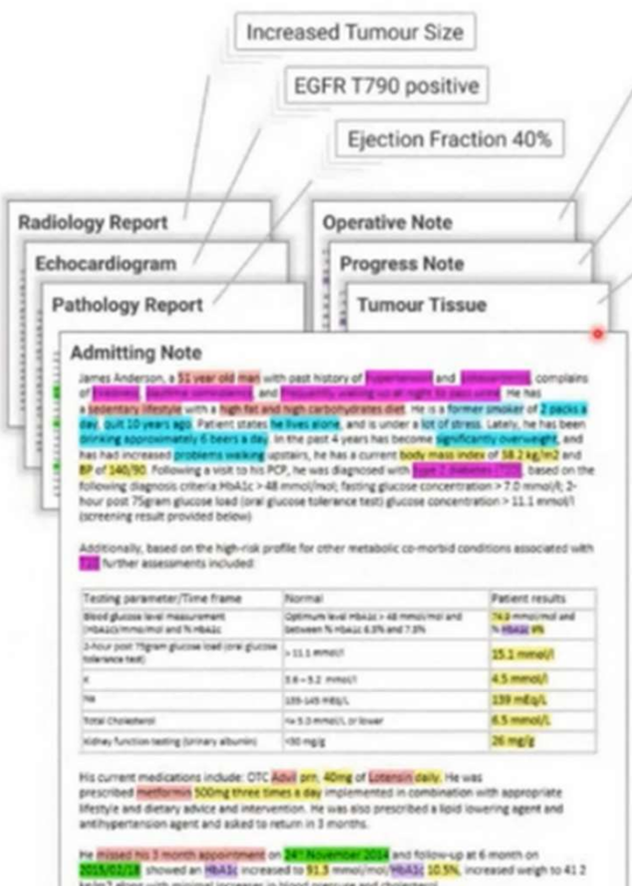


With coreference information combined with the parse tree and named entity information, we should be able to extract a lot of information out of this document.

Coreference resolution is one of the most difficult steps in our pipeline to implement. It's even more difficult than sentence parsing. Recent advances in deep learning have resulted in new approaches that are more accurate.

Text Analysis in Healthcare

NLP can extract powerful insights from the unstructured data locked in clinical notes



Diseases

Text	Normalized Value
Type 2 Diabetes T2D	Diabetes Mellitus, Type 2 SNOMEDid 44054006

Symptoms

Text	Normalized Value
Sleepiness Daytime Somnolence	Fatigue/tiredness
Frequently waking up at night to pass urine	Nocturia Urinary frequency

Social Determinants

Text	Normalized Value
Former smoker Quit 10 years ago	Ex smoker
Problems walking	Ambulatory Status: walking difficulty
Lot of stress	High levels of stress/Stress
He lives alone	Social Isolation
Missed his 3 month appointment	Did not attend

Demographics

Text	Normalized Value
51 year old	51y
man	Male

Dates

Text	Normalized Value
24th November 2014	20141124
2015/02/28	20150228

Measurements

Text	Normalized Value
Body mass index of 38.2 kg/m2	BMI 38.2 kg/m2
15.1 mmol/l	15.1 mmol/l

Medications

Text
On Metformin 0.5g PO three times a day

Normalized value

Medications

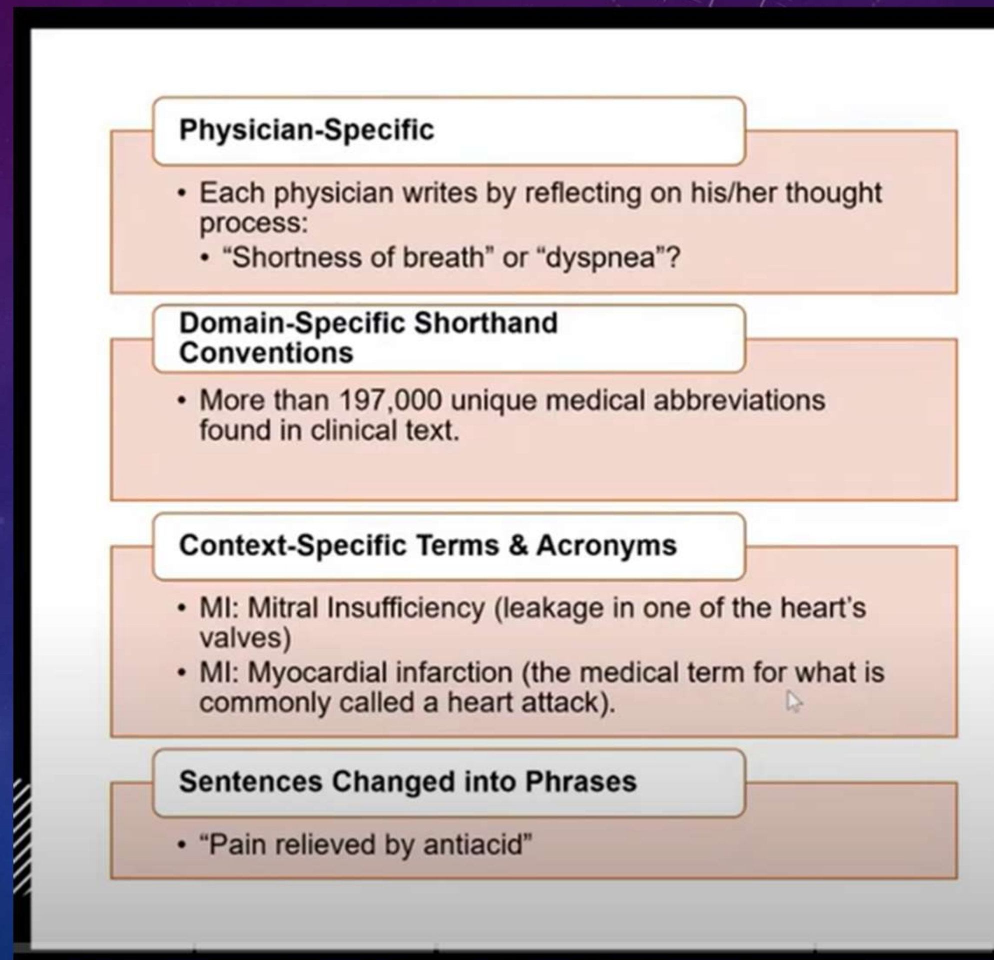
Drug Name	Dose	Unit	Route	Frequency
Metformin	500	mg	Oral	TID

IQVIA

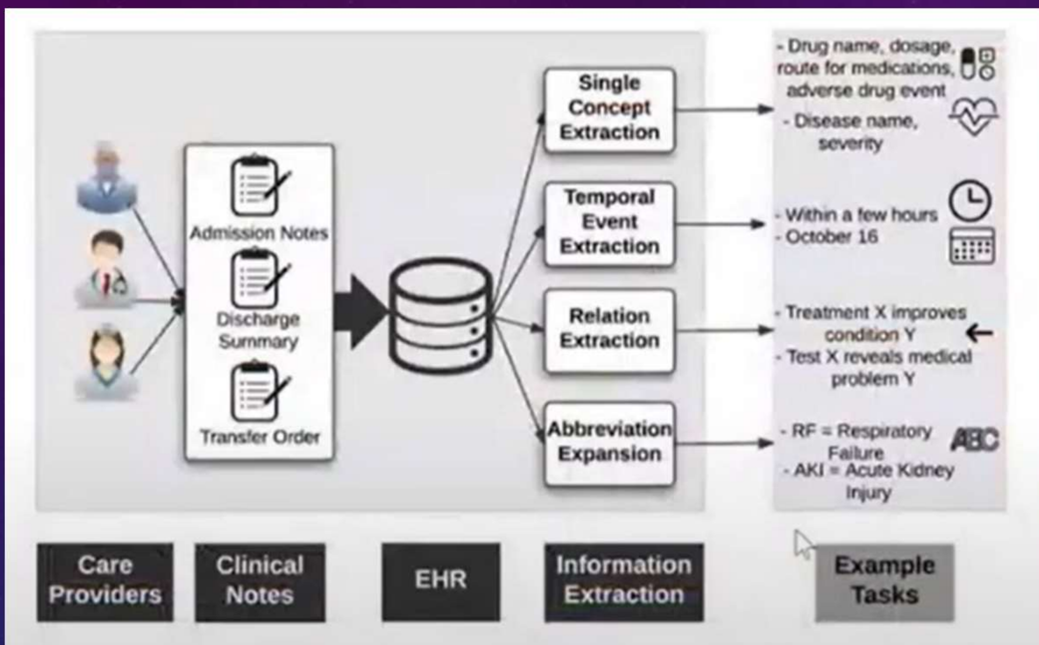
CLINICAL TEXT ANALYSIS

Various type of Medical Text

- Healthcare Data – EHR, EMR
- Clinical Notes
- History & Physical notes
- Discharge summaries
- Radiology Reports
- Lab Reports



CLINICAL TEXT ANALYSIS CHALLENGES



- Single Concept Extraction
  - Extraction of medical concepts from text, such as diagnoses, medications, or procedures.
  - Can be treated as a "sequence labeling" task to assign clinically relevant tags to each word.
- Temporal Event Extraction
  - Assigning the notion of time to each extracted concept, such as "in past few months" or "October 15".
- Relation Extraction
  - Structured relationship between medical concepts, such as treatment X improves condition Y, or test X reveals medical problem Y.
- Abbreviation Expansion
  - Many abbreviations correspond to more than one medical concept.

Non standard sentence, flexible formatting, unusual grammar, free text format , specific medical words/jargons etc.

CLINICAL NLP TASKS

- Word Sense Disambiguation (WSD) : associating words in context with their most suitable entry in a pre-defined sense inventory (typically WordNet).
- Named Entity Recognition (NER) : Tagging entities in text with their corresponding type, typically in BIO notation.
- Relationship extraction : extracting semantic relationships from a text. Extracted relationships usually occur between two or more entities and fall into specific semantic categories (e.g. Diagnosis, procedures, treatments , anatomy etc).
- Negation Identification : Identify pertinent negative from narrative clinical reports
- Semantic Role Labeling : Detect the semantic roles played by each noun phrase associated with the verb of the sentence.
- Information Extraction : Automated extraction of family & observation predictions from unstructured text.
- Coreference Resolution: clustering mentions in text that refer to the same underlying real-world entities.
- Grammatical Error Correction: correcting different kinds of errors in text such as spelling, punctuation, grammatical, and word choice errors.
- Feature Extraction: extraction of generic numerical features from text, usually embeddings.

Word Sense Disambiguation (WSD)

- Word sense: a meaning of a word.
- Acronym
  - "The patient underwent a left BK amputation."
    Sense: below knee
  - "BK viremia in the past."
    Sense: BK (virus)
- Abbreviation
  - "CT of head showed old CVA on left side."
    Sense: cerebrovascular accident
  - "Straight with no CVA tenderness."
    Sense: costovertebral angle

Named Entity Recognition (NER)

BRIEF HISTORY: The patient is an (XX)-year-old female with history of <problem> previous stroke </problem> ; <problem> hypertension </problem> ; <problem> COPD </problem> , stable ; <problem> renal carcinoma </problem> ; presenting after <problem> a fall </problem> and possible <problem> syncope </problem> .
While walking , she accidentally fell to her knees and did hit <problem> her head on the ground </problem> , near <problem> her left eye </problem> .
<problem> Her fall </problem> was not observed , but the patient does not profess <problem> any loss of consciousness </problem> , recalling the entire event.
The patient does have a history of <problem> previous falls </problem> , one of which resulted in <problem> a hip fracture </problem> .
She has had <treatment> physical therapy </treatment> and recovered completely from that .
<test> Initial examination </test> showed <problem> bruising </problem> around the left eye , normal lung examination , normal heart examination , normal neurologic function with a baseline decreased mobility of <problem> her left arm </problem> .
The patient was admitted for <test> evaluation </test> of <problem> her fall </problem> and to rule out <problem> syncope </problem> and possible <problem> stroke </problem> with <problem> her positive histories </problem> .
<test> DIAGNOSTIC STUDIES: All x-rays </test> including <problem> left foot , right knee , left shoulder and cervical spine </problem> showed no <problem> acute fractures </problem> .
<problem> The left shoulder did show old healed left humeral head and neck fracture </problem> with <problem> baseline anterior dislocation </problem> .
<test> CT of the brain </test> showed no <problem> acute changes </problem> , <problem> left periorbital soft tissue swelling </problem> .
<test> CT of the maxillofacial area </test> showed no <problem> facial bone fracture </problem> .
<test> Echocardiogram </test> showed normal left ventricular function , <test> ejection fraction </test> estimated greater than 65% .

http://text-machine.cs.uml.edu/cliner/

- Disease/Disorder
- Test
- Treatment
- Medications
- Anatomical Sites
- Anatomical modifiers
- Procedures
- Findings

Relationship Extraction (RE)

- Determine relationships between entities or events

"We used hemofiltration to **treat** a patient with digoxin overdose that was complicated by refractory hyperkalemia." [PMID: 3718110]

Relationship: Hemofiltration-**TREATS**-Patients

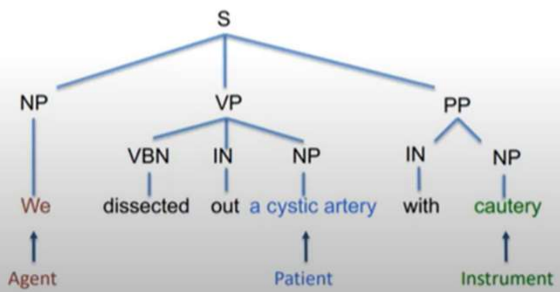
Negation Identification (NegIde)

Identify pertinent Negatives from narrative clinical reports

- "The chest X-ray showed no infiltrates…"
- "The patient denied experiencing chest pain"
- " no murmurs, rubs or gallops"
- "murmurs, rubs and gallops are absent"

Semantic Role Labeling

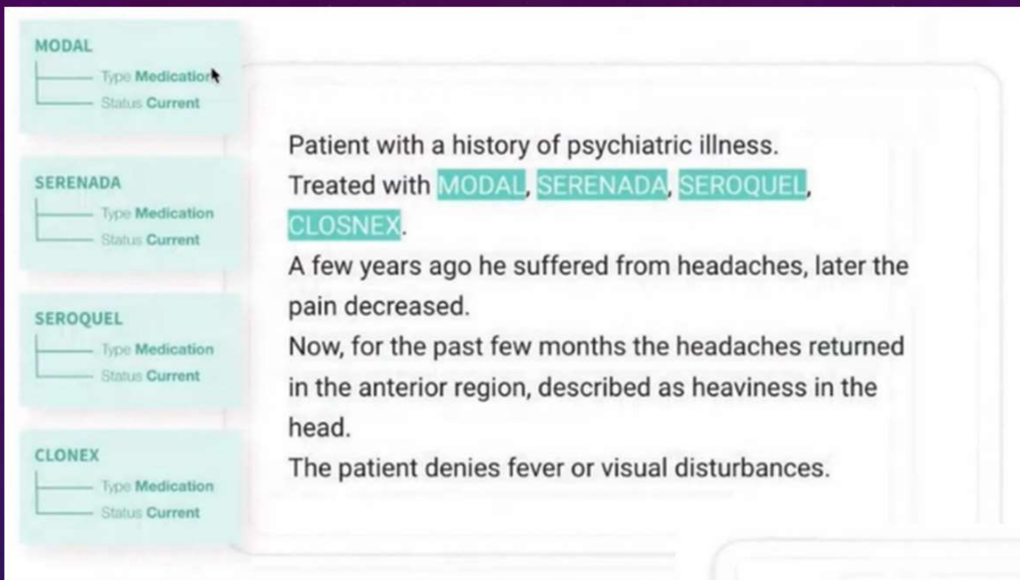
- Detect the semantic role played by each noun phrase associated with the verb of a sentence
  - Agent: Noun Phrase (NP) before the verb
  - Patient: NP after the verb
  - Instrument: NP in a Prepositional Phrase (PP)

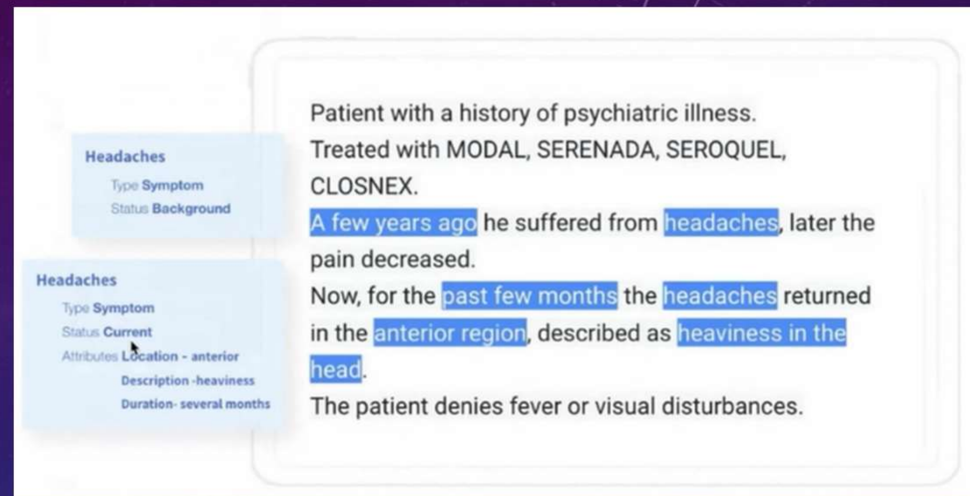


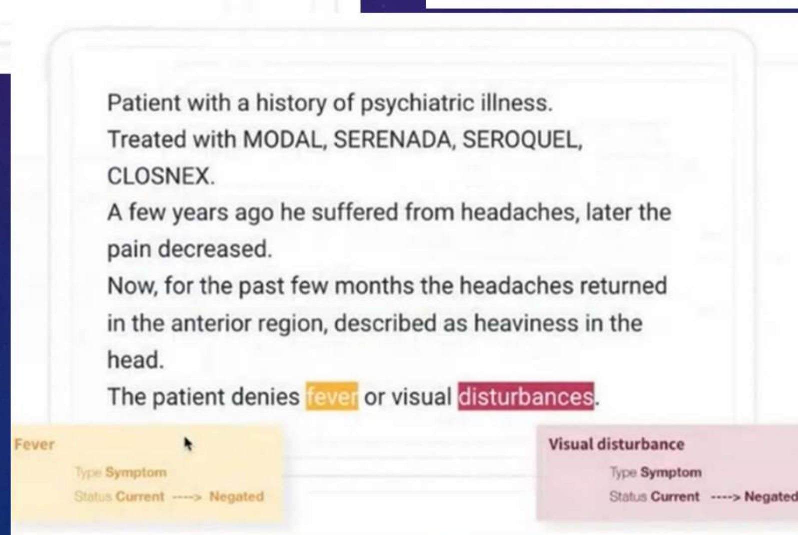
Information Extraction

- Automated extraction of family and observation predications from unstructured text
  - Supplied text: "Heart disease on the father side of the family. Mother has arthritis."
  - Extracted elements:
    - Constituent: family {FAMILY HISTORY: FAMMEMB}
    - Constituent: observation {Heart disease: C1576434}
    - Constituent: family {father side of the family: Paternal*}
    - Constituent: family {Mother: MTH}
    - Constituent: observation {arthritis: C1692886}
  - Predications:
    - Family Member{father side of the family}, Observation{Heart disease}, Negated{false}
    - Family Member{Mother}, Observation{arthritis}, Negated{false}

CLINICAL TEXT EXAMPLES







PUBLIC DATASETS

- MIMIC II and MIMIC III
- MTSample
- I2b2 NLP challenges data
- THYME corpus (deidentified ,
- ClinicalTrialGov.in
-

NLP USE CASES IN HEALTHCARE

- Medical Outcome Prediction
- Predicting recovery trajectory
- Clinical information extraction ,
- Disease prediction,
- Medical coding,
- Patient phenotyping,
- Patient Population

- Resource Allocation
- Intervention Planning
- Predicting the readmission risks
- Automate Claim Adjudication
- Predicting Claim Rejection
- Medical Fraud Detection
- Re-admission prediction

HEALTHCARE TERMS

EHR – Electronic Health Record

EMR – Electronic Medical Records

HIPAA - Health Insurance Portability & Accountability Act

PHI - Protected Health Information

HL7 - HL7 is an open-standard data model that enables data interoperability for systems

FHIR - The Fast Healthcare Interoperability Resource is a draft data standard developed & nurtured by HL7 International

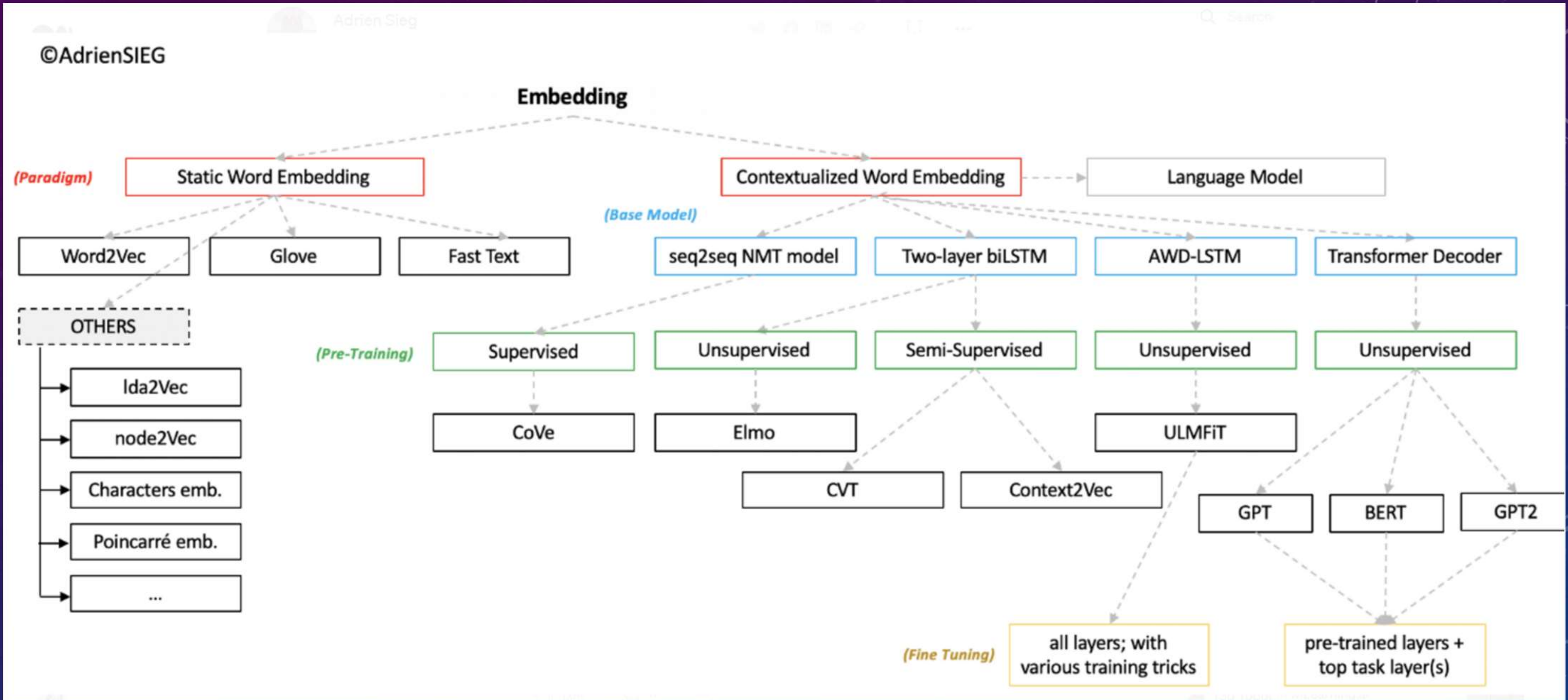
DICOM - Digital Imaging and Communications in Medicine is the standard for communication & management of medical imaging & related data

HEALTHCARE AI CHALLENGES & OUTLOOK

- With the digital transformation, health organizations are capturing a huge amount of electronic medical records (EMR) every day. Healthcare data is complex and highly secured under various protection laws.
- The data is stored in various non-standard formats; unlocking this data and making sense of it is a very complex business case. Modern patients and demanding consumers need information quickly and securely.
- Though health information exchanges (HIE) are helping healthcare organizations to build specialized networks that rely on interoperable systems to share electronic health records (EHR) seamlessly and securely, the challenges are there due to custom build EHR, budget restrictions, complex technology to extract that information is a standard format.
- Healthcare organizations are facing significant challenges due to rising costs, inefficient operational processes, and inconsistent, complex health data, overall impacting the patient care system.

Text Analysis - Technology Trend

EMBEDDING TO LANGUAGE MODEL



https://towardsdatascience.com/from-pre-trained-word-embeddings-to-pre-trained-language-models-focus-on-bert-343815627598

OPEN SOURCE TECHNOLOGIES

- https://text-machine.cs.uml.edu/cliner/
- Spark-nlp
- NLTK
- Spacy
- MedLEE
- cTAKES
- MedEX
- HiTEX
- MedTagger
- BioMedICUS

Word Embedding Techniques

One hot encoding
Label encoding
Word2Vec
Topic Modelling
TF/IDF
Vector Assembler
Estimator & Transformers
FastText

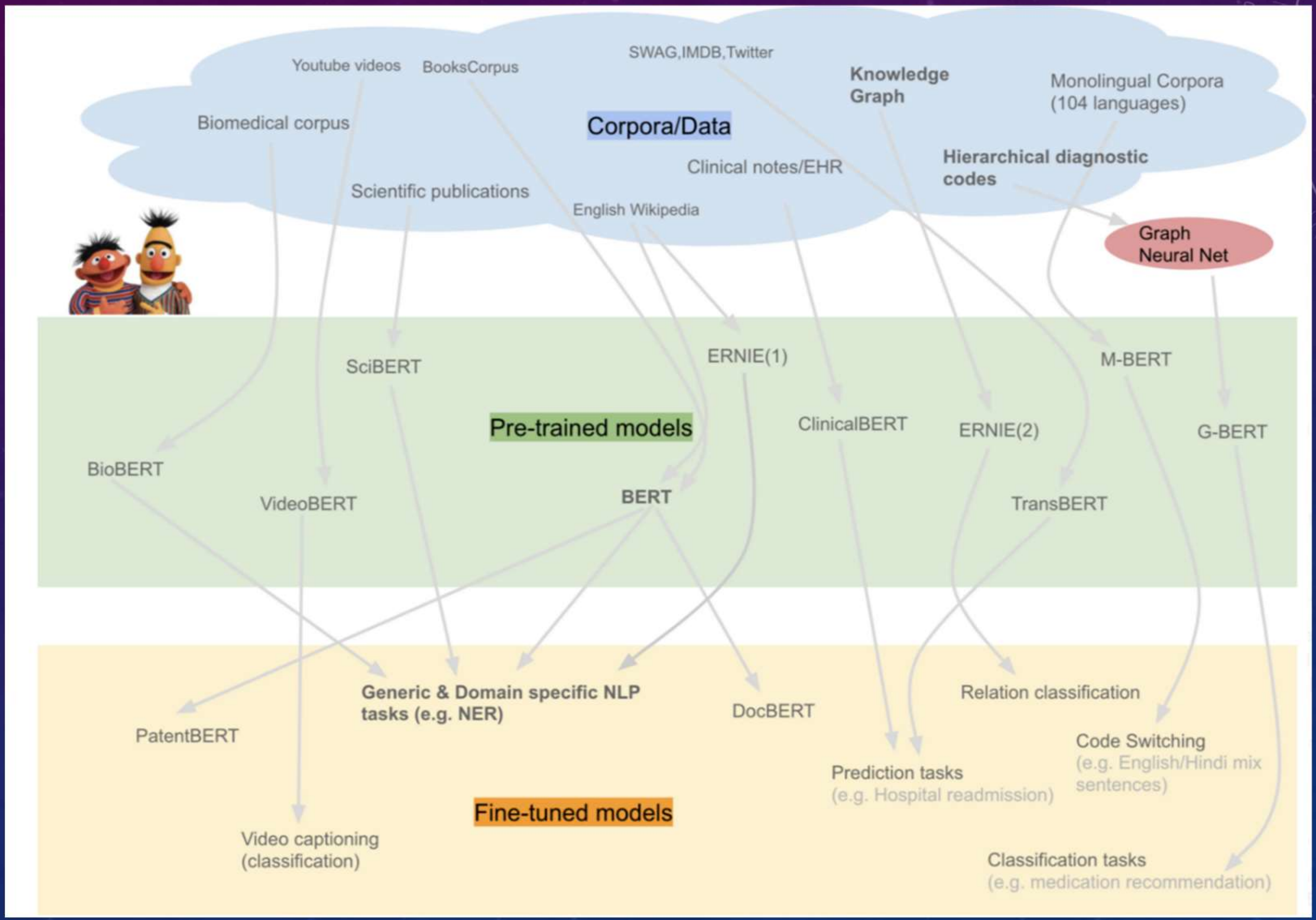
Pretrained embeddings

BERT
BioBERT
Bio_ClinicalBERT
ClinicalBERT
BioWordVec & BioSentVec
G-BERT
Med-BERT

BERT → Transfer by **Attention Extraction, BERT** uses a bidirectional Transformer

https://mostafadehghani.com/2019/05/05/universal-transformers/

REAL APPLICATION : PRE-TRAINED VS FINE TUNED



TOP OPEN SOURCE LARGE LANGUAGE MODELS

A Language Model at the heart is just a probability distribution over sequences of tokens (words).
The Language Models are the core of modern Natural Language Processing (NLP) and their applications can be for a variety of NLP tasks such as speech-to-text, sentiment analysis, text summarization, spell checking, token classification, etc.
In most NLP tasks, the Language Models can determine the probability of the next token by analyzing the given text. The Language Model can be in the form of Unigrams, N-grams, Exponential, or Neural networks.

Cloud vs Open source platforms

Pre-trained APIs for speech, transcription, translation, text analysis, and chatbot functionality.

Services include Amazon Comprehend & Comprehend Medical to discover insights and relationships in text & medical text with various NLP tasks.
AWS Textract extracts text from images, printed or handwritten forms, ID cards, pdf documents and provide json output and user friendly dashboard.

Amazon Transcribe for automatic speech recognition,
Amazon Translate for fluent translation of text,

Amazon Polly for natural sounding from text to speech,
Amazon Lex to build chatbots to engage with customers,

Amazon Kendra to do an intelligent search of enterprise systems to quickly find the content one is looking for.

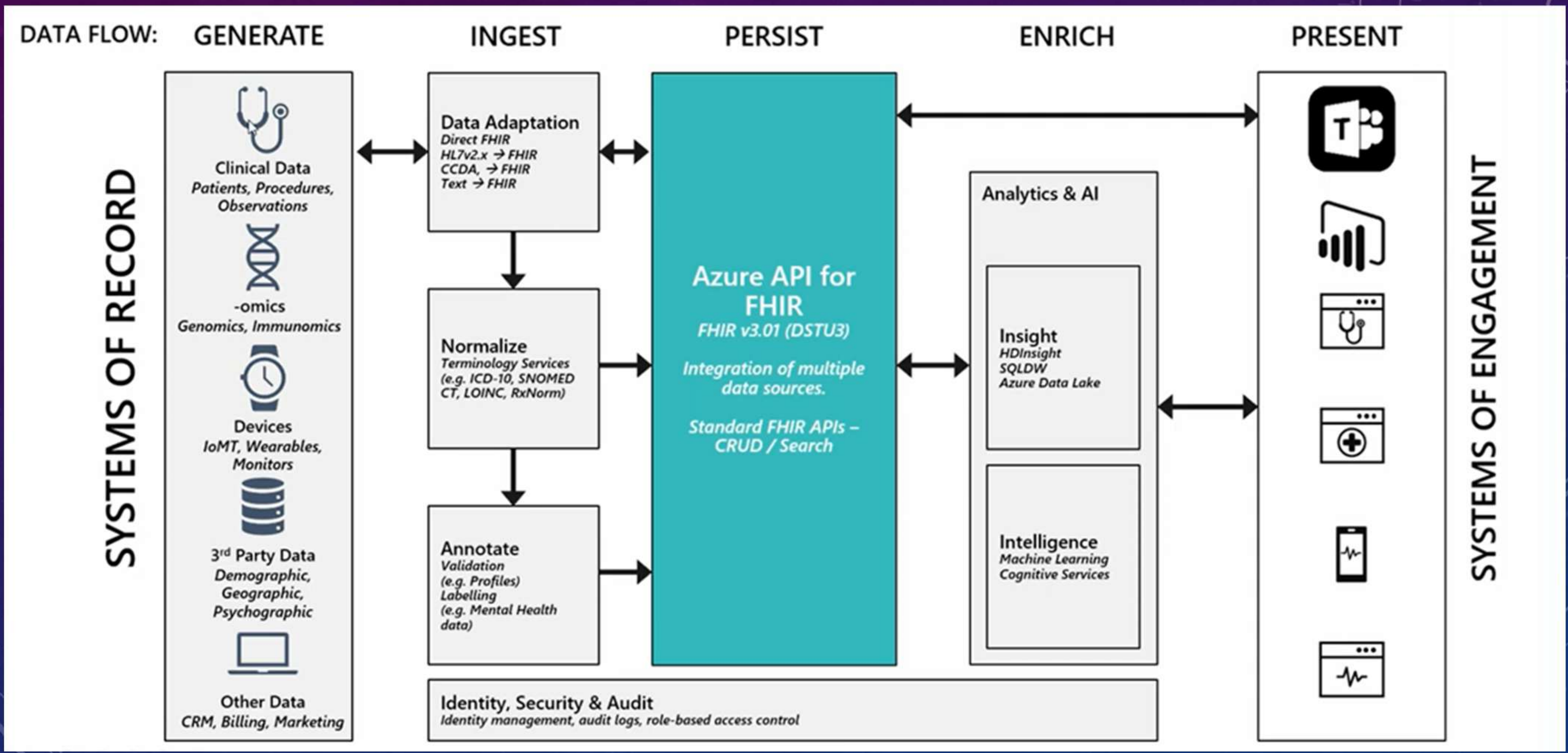
Amazon SageMaker offers platform to prepare data and build, train, and deploy NLP models for any use case with fully managed infrastructure, tools, and workflows, including no-code offerings for business analysts.
Hugging Face on Amazon SageMaker offers fine-tune pre-trained models from Hugging Face, Transformers & Bert

Azure Cognitive Service for Language is a cloud-based service that provides Natural Language Processing (NLP) features for understanding and analyzing text. Use this service to help build intelligent applications using the web-based Language Studio, REST APIs, and client libraries.
Named Entity Recognition (NER)
Personally identifying (PII) and health (PHI) information detection
Language detection
Sentiment Analysis and opinion mining
Summarization
Key phrase extraction
Entity linking
Text analytics for health
Custom text classification
Custom Named Entity Recognition (Custom NER)
Conversational language understanding
Orchestration workflow
Question answering

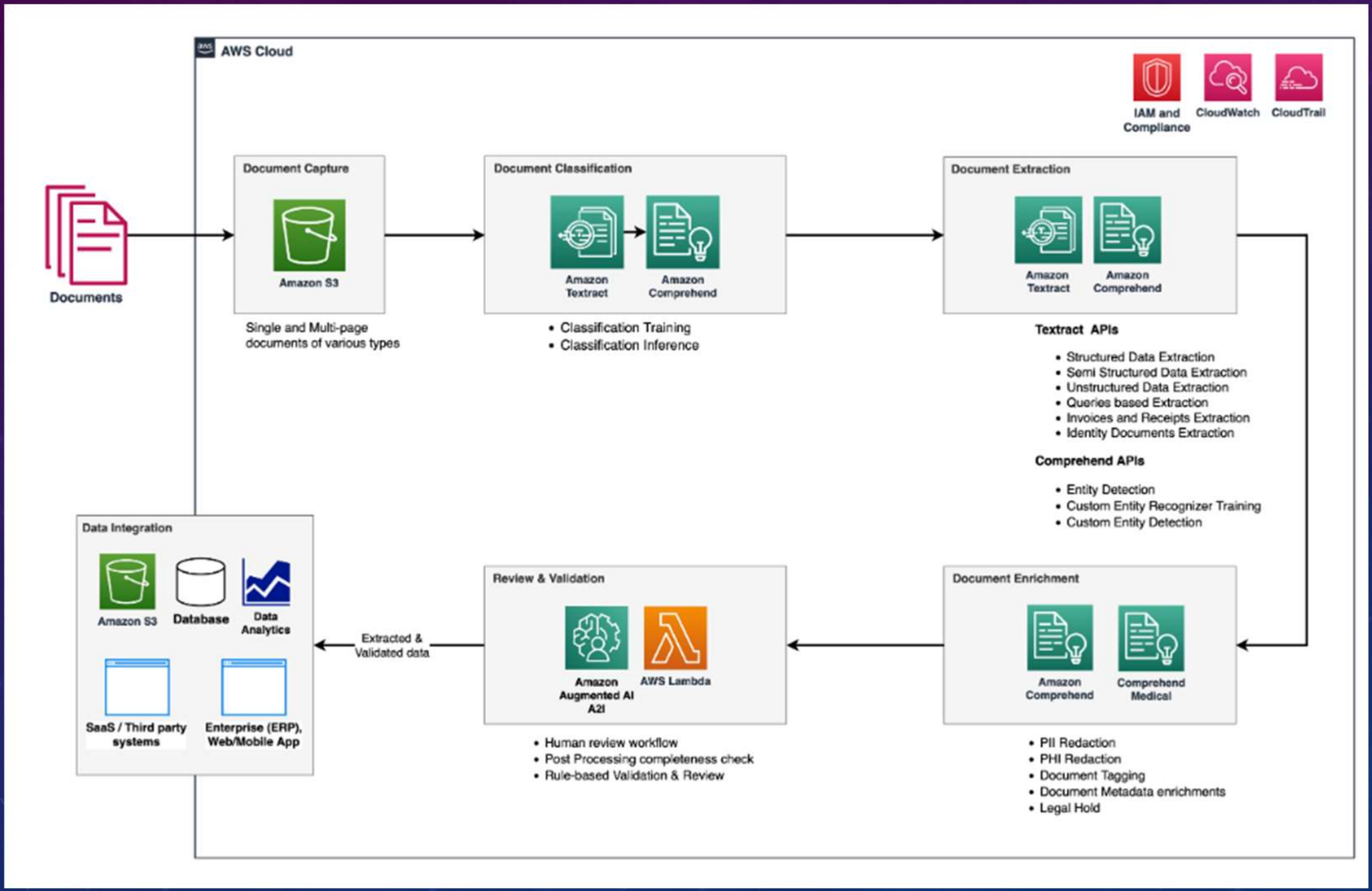
GCP based NLP services offer

Sentiment analysis
Entity analysis
Entity sentiment analysis
Syntactic analysis
Content classification

AZURE FIHR API



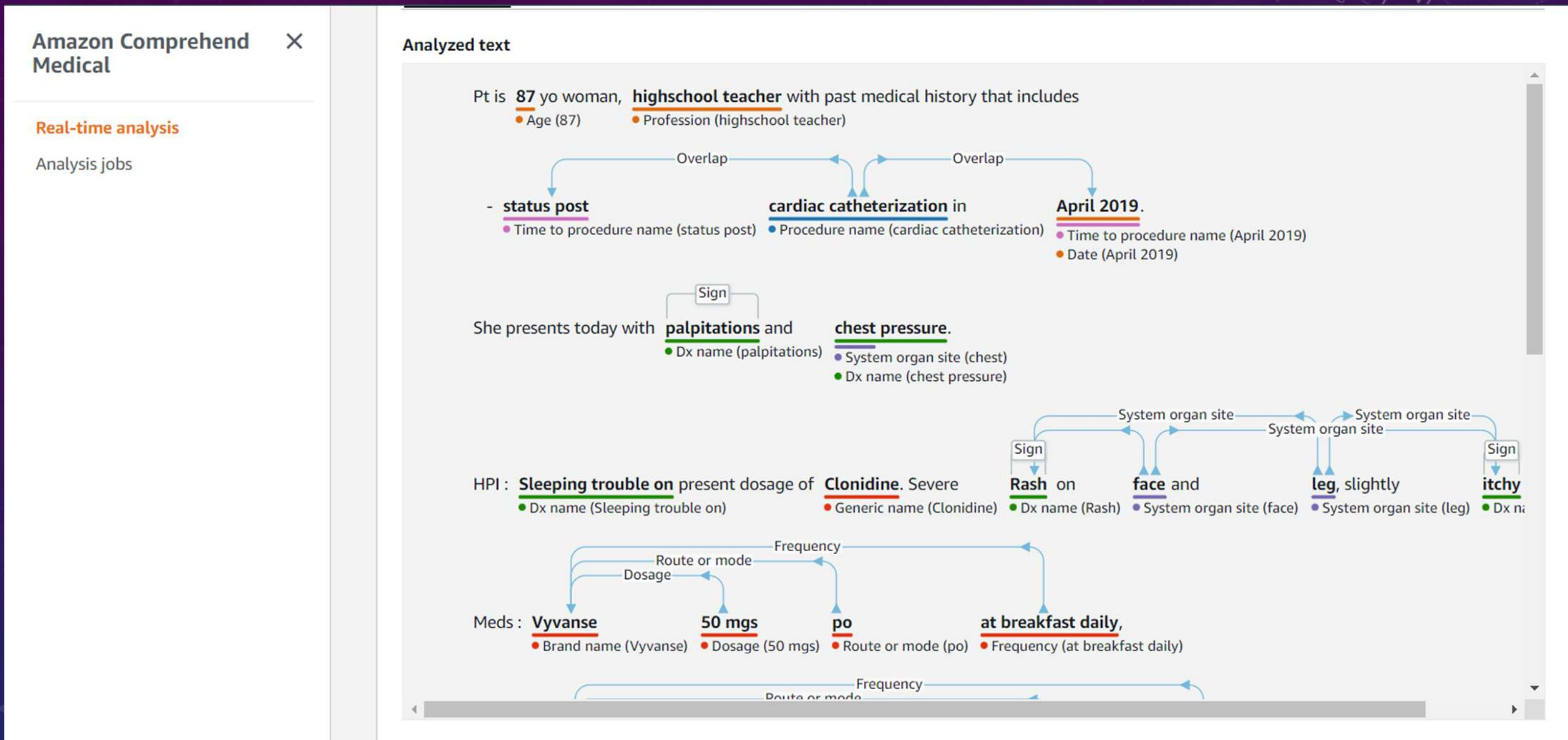
INTELLIGENT DOCUMENT PROCESSING WITH AWS MEDICAL COMPREHEND

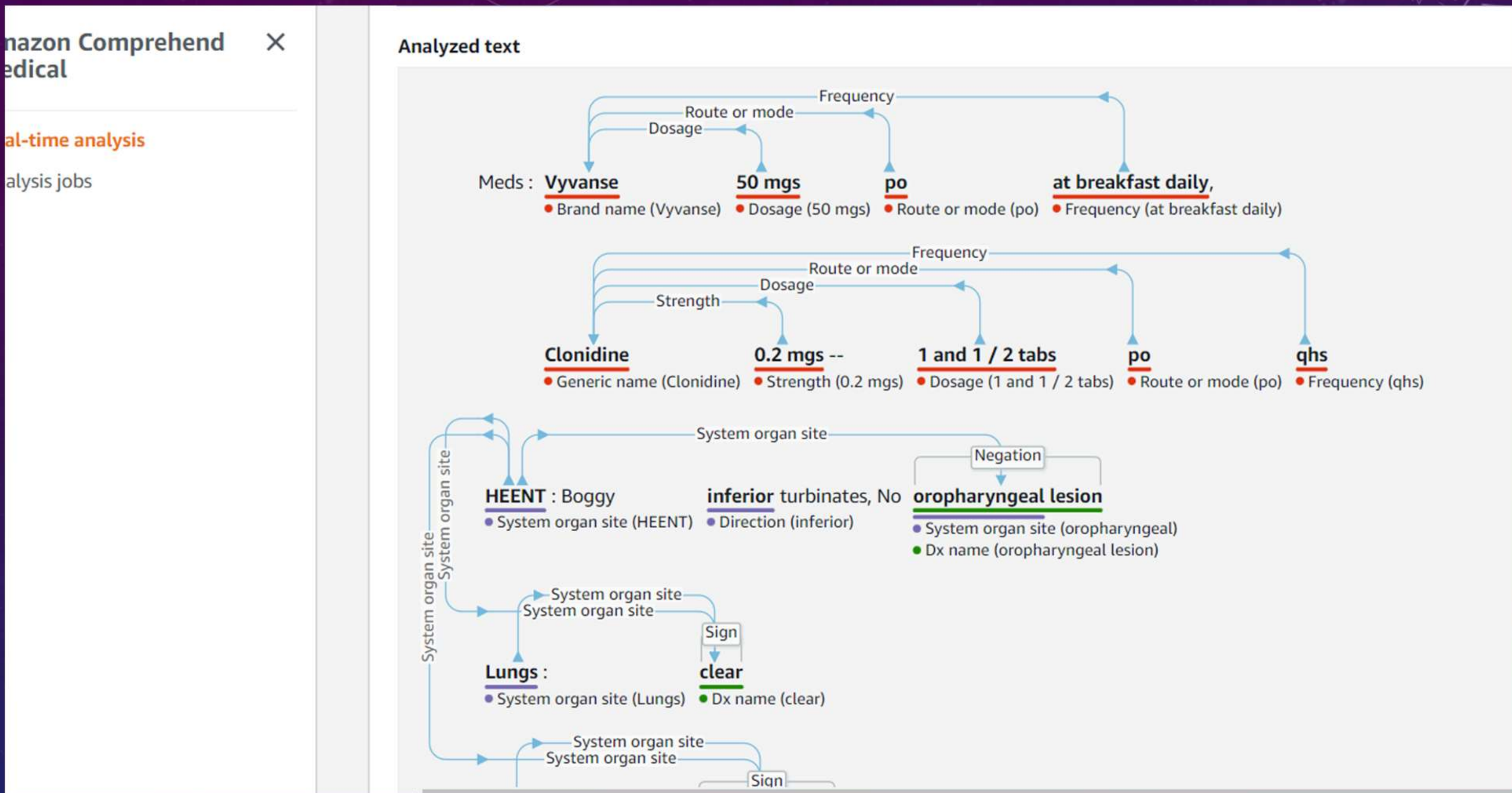


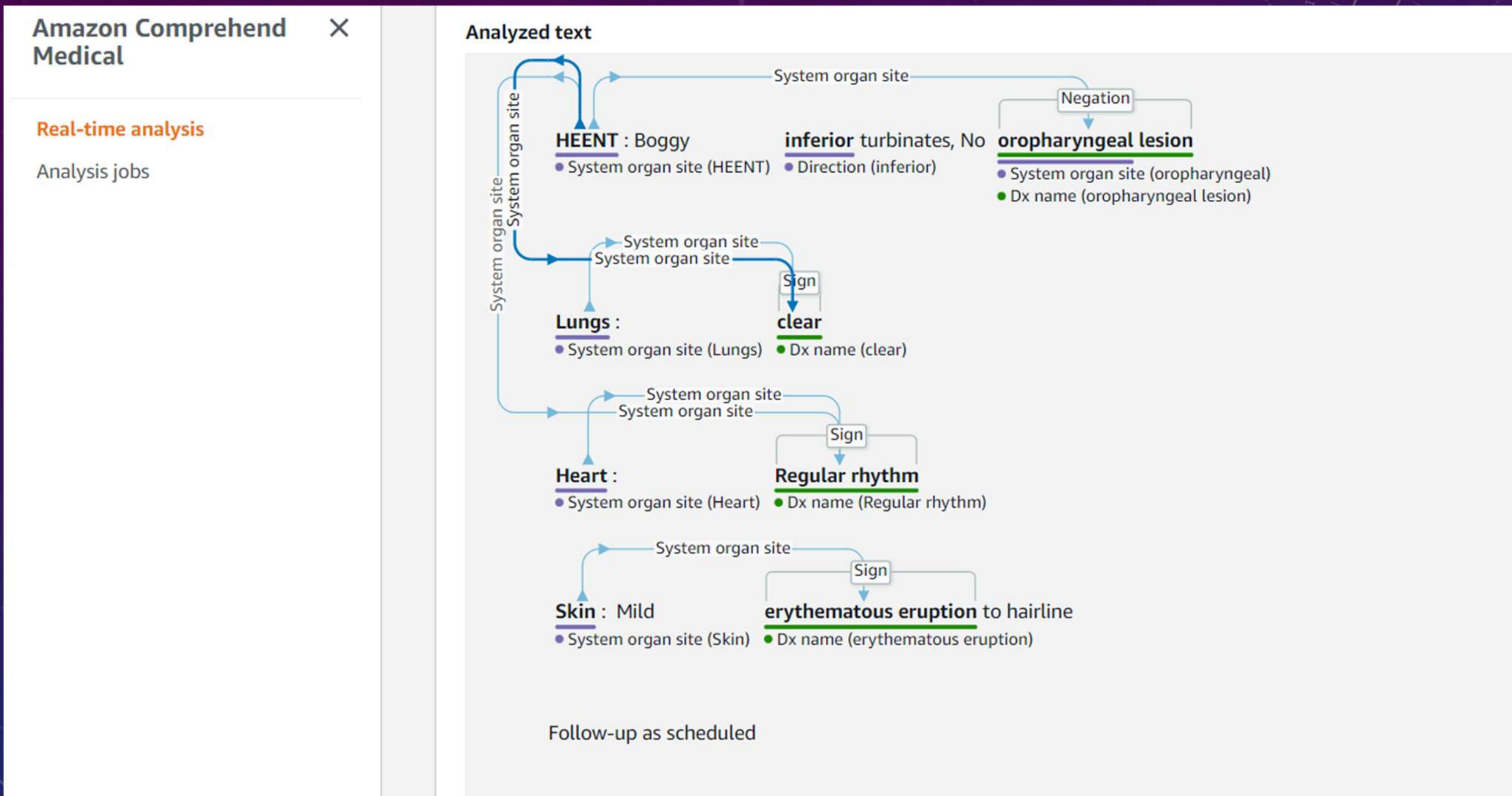
PUT AWS MEDICAL COMPREHEND EXAMPLE FOR VISUALIZATION

- Pt is 87 yo woman, highschool teacher with past medical history that includes
- - status post cardiac catheterization in April 2019.
- She presents today with palpitations and chest pressure.
- HPI : Sleeping trouble on present dosage of Clonidine. Severe Rash on face and leg, slightly itchy
- Meds : Vyvanse 50 mgs po at breakfast daily,
- Clonidine 0.2 mgs -- 1 and 1 / 2 tabs po qhs
- HEENT : Boggy inferior turbinates, No oropharyngeal lesion
- Lungs : clear
- Heart : Regular rhythm
- Skin : Mild erythematous eruption to hairline

- Follow-up as scheduled


Amazon Comprehend Medical
Real-time analysis
Analysis jobs
Analyzed text
Pt is 87 yo woman, highschool teacher with past medical history that includes
Age (87)
Profession (highschool teacher)
Overlap
Overlap
- status post
cardiac catheterization in
April 2019.
Time to procedure name (status post)
Procedure name (cardiac catheterization)
Time to procedure name (April 2019)
Date (April 2019)
Sign
She presents today with palpitations and chest pressure.
Dx name (palpitations)
System organ site (chest)
Dx name (chest pressure)
System organ site
System organ site
System organ site
Sign
Sign
HPI : Sleeping trouble on present dosage of Clonidine. Severe Rash on face and leg, slightly itchy
Dx name (Sleeping trouble on)
Generic name (Clonidine)
Dx name (Rash)
System organ site (face)
System organ site (leg)
Frequency
Route or mode
Dosage
Meds : Vyvanse 50 mgs po at breakfast daily,
Brand name (Vyvanse)
Dosage (50 mgs)
Route or mode (po)
Frequency (at breakfast daily)
Frequency

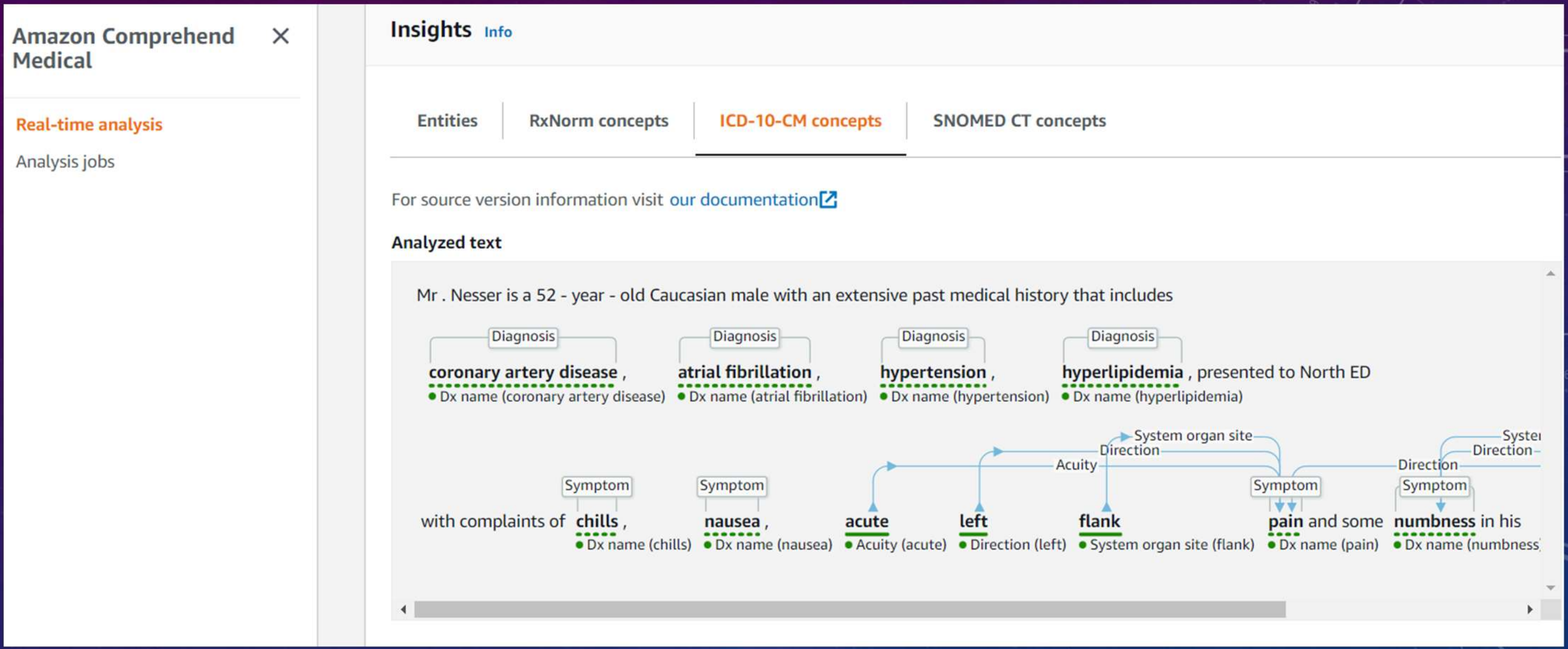

nazon Comprehend
edical
al-time analysis
alysis jobs
Analyzed text
Frequency
Route or mode
Dosage
Meds : Vyvanse 50 mgs po at breakfast daily,
Brand name (Vyvanse)
Dosage (50 mgs)
Route or mode (po)
Frequency (at breakfast daily)
Frequency
Route or mode
Dosage
Strength
Clonidine 0.2 mgs -- 1 and 1 / 2 tabs po qhs
Generic name (Clonidine)
Strength (0.2 mgs)
Dosage (1 and 1 / 2 tabs)
Route or mode (po)
Frequency (qhs)
System organ site
Negation
HEENT : Boggy inferior turbinates, No oropharyngeal lesion
System organ site (HEENT)
Direction (inferior)
System organ site (oropharyngeal)
Dx name (oropharyngeal lesion)
System organ site
System organ site
Sign
Lungs : clear
System organ site (Lungs)
Dx name (clear)
System organ site
System organ site
Sign



Amazon Comprehend Medical
Real-time analysis
Analysis jobs
Analyzed text
System organ site
Negation
System organ site
System organ site
HEENT : Boggy
System organ site (HEENT)
inferior turbinates, No
Direction (inferior)
oropharyngeal lesion
System organ site (oropharyngeal)
Dx name (oropharyngeal lesion)
System organ site
System organ site
Sign
Lungs :
System organ site (Lungs)
clear
Dx name (clear)
System organ site
System organ site
Sign
Heart :
System organ site (Heart)
Regular rhythm
Dx name (Regular rhythm)
System organ site
Sign
Skin : Mild
System organ site (Skin)
erythematous eruption to hairline
Dx name (erythematous eruption)
Follow-up as scheduled

ICD-10

- Mr . Nesser is a 52 - year - old Caucasian male with an extensive past medical history that includes coronary artery disease , atrial fibrillation , hypertension , hyperlipidemia , presented to North ED with complaints of chills , nausea , acute left flank pain and some numbness in his left leg.



Amazon Comprehend Medical
Real-time analysis
Analysis jobs
Insights Info
Entities
RxNorm concepts
ICD-10-CM concepts
SNOMED CT concepts
For source version information visit our documentation
Analyzed text
Mr . Nesser is a 52 - year - old Caucasian male with an extensive past medical history that includes
Diagnosis
coronary artery disease ,
Dx name (coronary artery disease)
Diagnosis
atrial fibrillation ,
Dx name (atrial fibrillation)
Diagnosis
hypertension ,
Dx name (hypertension)
Diagnosis
hyperlipidemia , presented to North ED
Dx name (hyperlipidemia)
System organ site
Direction
Acuity
Direction
Symptom
with complaints of chills ,
Dx name (chills)
Symptom
nausea ,
Dx name (nausea)
acute
Acuity (acute)
left
Direction (left)
flank
System organ site (flank)
Symptom
pain and some
Dx name (pain)
Symptom
numbness in his

Demo -

Making sense of healthcare data using Text Analysis

MAKING SENSE OF HEALTHCARE TEXT – DEMO WITH AWS HEALTHLAKE

- Population Health Analytics with AWS HealthLake and QuickSight
- https://www.analyticsvidhya.com/blog/2022/04/population-health-analytics-with-aws-healthlake-and-quicksight/

DEMO USING AWS NLP & BERT

- Extract insight from discharge summaries using AWS Comprehend Medical
- https://colab.research.google.com/drive/1d7s3VMIFramuOT9BKxn0ogkGvBdcOOcr#scrollTo=OvyHicwN-7Vu

- Fine tune with BERT model
- https://github.com/nwams/ClinicalBERT-Deep-Learning--Predicting-Hospital-Readmission-Using-Transformer

REFERENCES

- https://www.youtube.com/watch?v=njIDXjc-8sY&t=187s
- https://medium.com/@nwamaka_41565/predicting-hospital-readmission-using-nlp-5f0fe6f1a705
- https://www.youtube.com/watch?v=cytM6QH2Zno
- https://www.youtube.com/watch?v=njIDXjc-8sY&t=187s
- https://towardsdatascience.com/from-pre-trained-word-embeddings-to-pre-trained-language-models-focus-on-bert-343815627598
- BERT/Transformer : https://jalammar.github.io/illustrated-transformer
- /https://jalammar.github.io/illustrated-bert/

THANK YOU