

Single Day Event

**Natural Language Processing (NLP) aid in
clinical trial and real-world analysis**

Rashmi G Jain,
Statistical Specialist,
Novo Nordisk

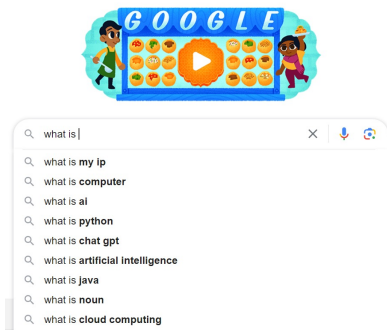


Disclaimer

The opinions expressed in this presentation and on the following slides are solely those of the presenter and not necessarily those of Novo Nordisk.

Novo Nordisk does not guarantee the accuracy or reliability of the information provided herein.

phuse Have you used any NLP powered tools?

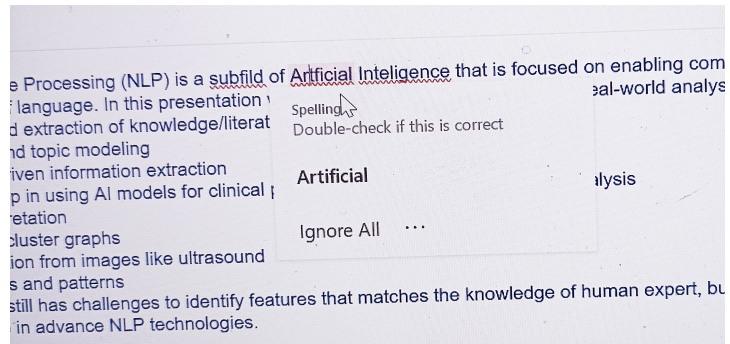


Google suggestion

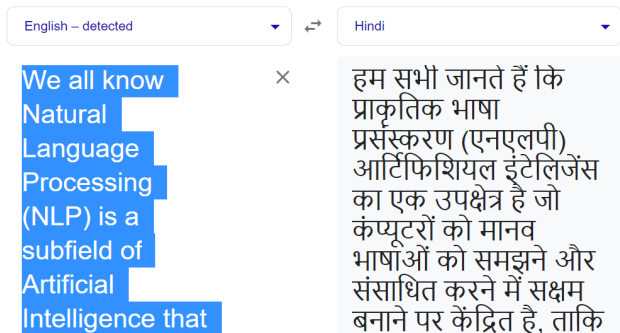


This Photo by Unknown Author is licensed under [CC BY-SA-NC](#)

Virtual assistant devices



Spelling correction



Translation

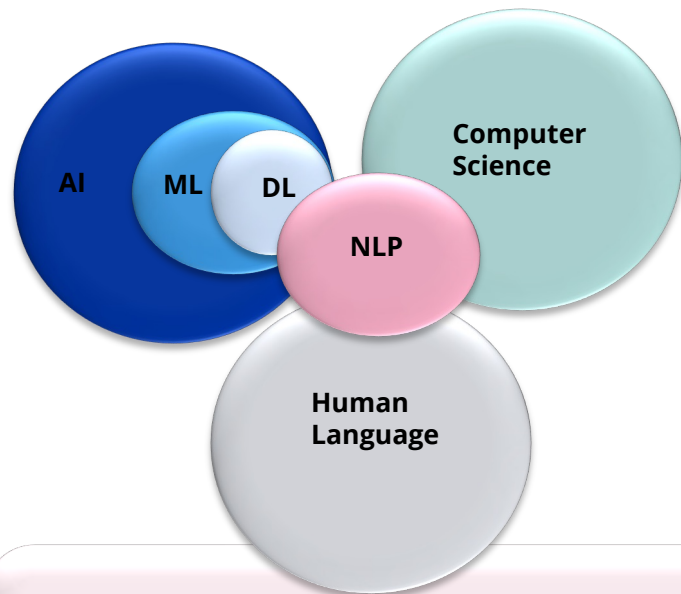


This Photo by Unknown Author is licensed under [CC BY-NC-ND](#)

Spam mail identifier



What is NLP?



NLP refers to the branch of computer science, linguistics and artificial intelligence concerned with giving computers the ability to understand text and spoken words in much the same way human beings can. Its rule-based modeling of human language with statistical, machine learning, and deep learning models.

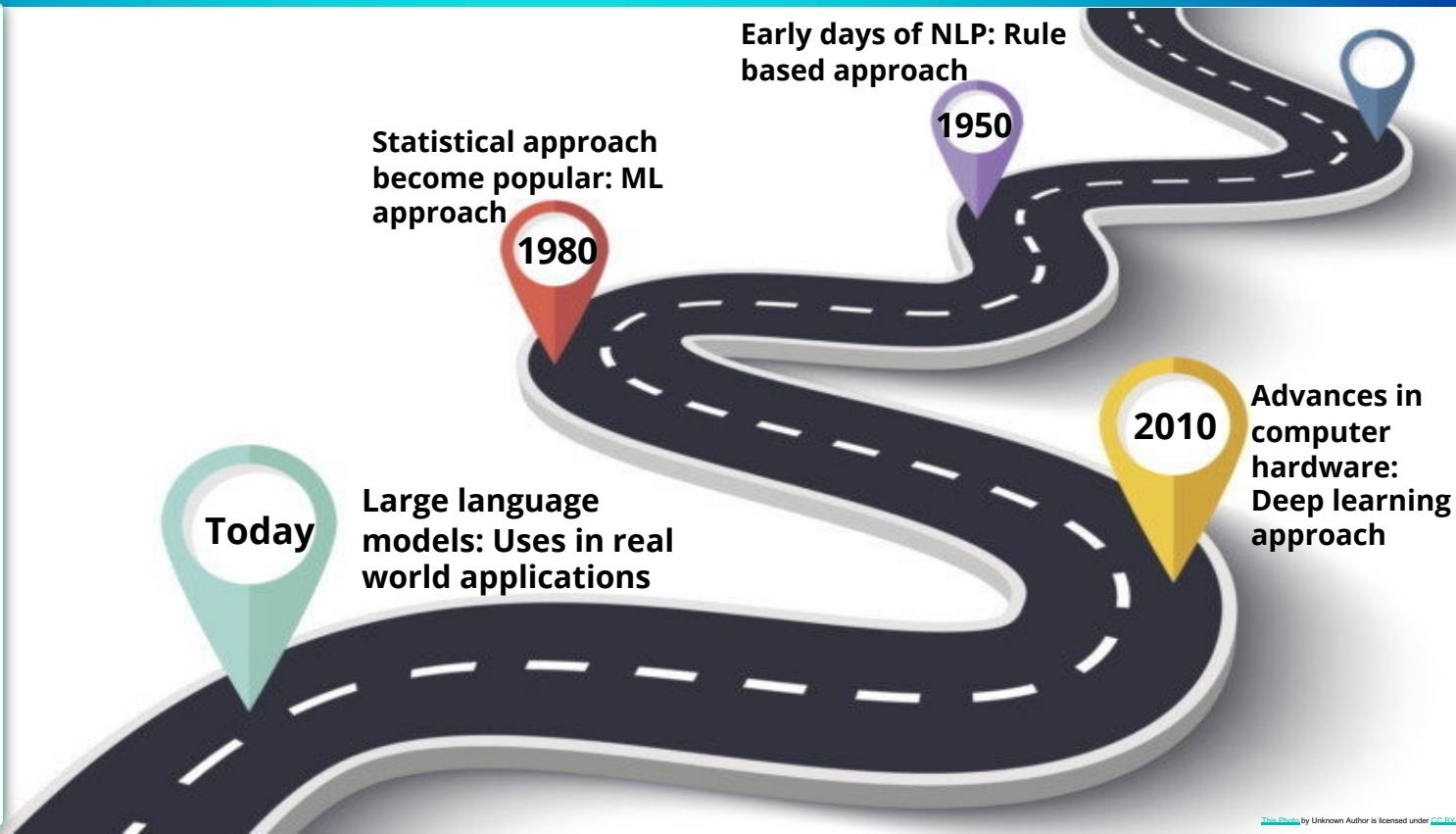




History of NLP

In 1947, after world war II, importance of language translation by a machine increased.

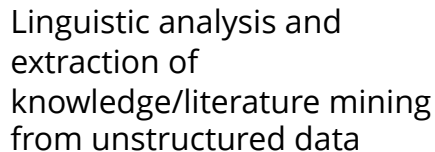
Present day NLP is focused more on information extraction and generation from big data. Uses the information to improve language processing.





clinical trials and
real-world
analysis
benefited from
NLP

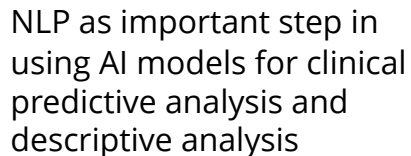
Let's explore NLP aid examples!



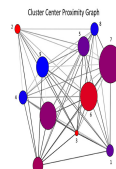
Sentiment analysis and topic modeling



System integration driven information extraction



Patient record interpretation



Creating symptoms cluster graphs



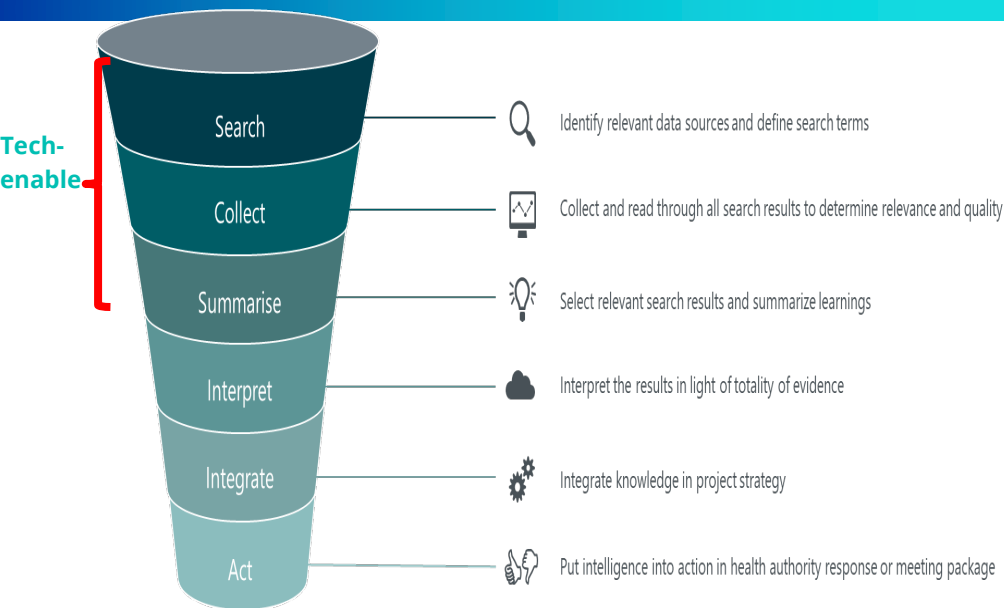
Automate interpretation from images like ultrasound



Detect critical trends and patterns



Extraction of literature from unstructured data



Extraction of knowledge from free text clinical reports, trial metadata, clinical trial descriptions, consumer generated text, biomedical literature, clinical guidelines etc.



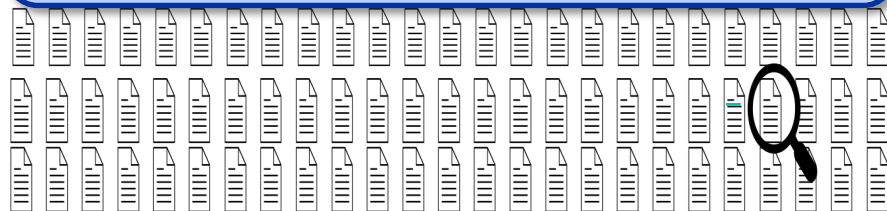
Improves the way we make Trial designs, product strategies and respond to questions from Health Authorities

Reduce time spent on writing new content or searching for information across lengthy files

Give consistent answers

Re-use content to provide quick answers

Avoid increase of resources or delays of approvals



Identify similarities across trials with trial metadata

Can flowchart be helpful?

- ❖ A lot of assessments are common across trials and may not be that useful to describe a trial rational / objective
- ❖ Can be harder to find what is important for a similarity score

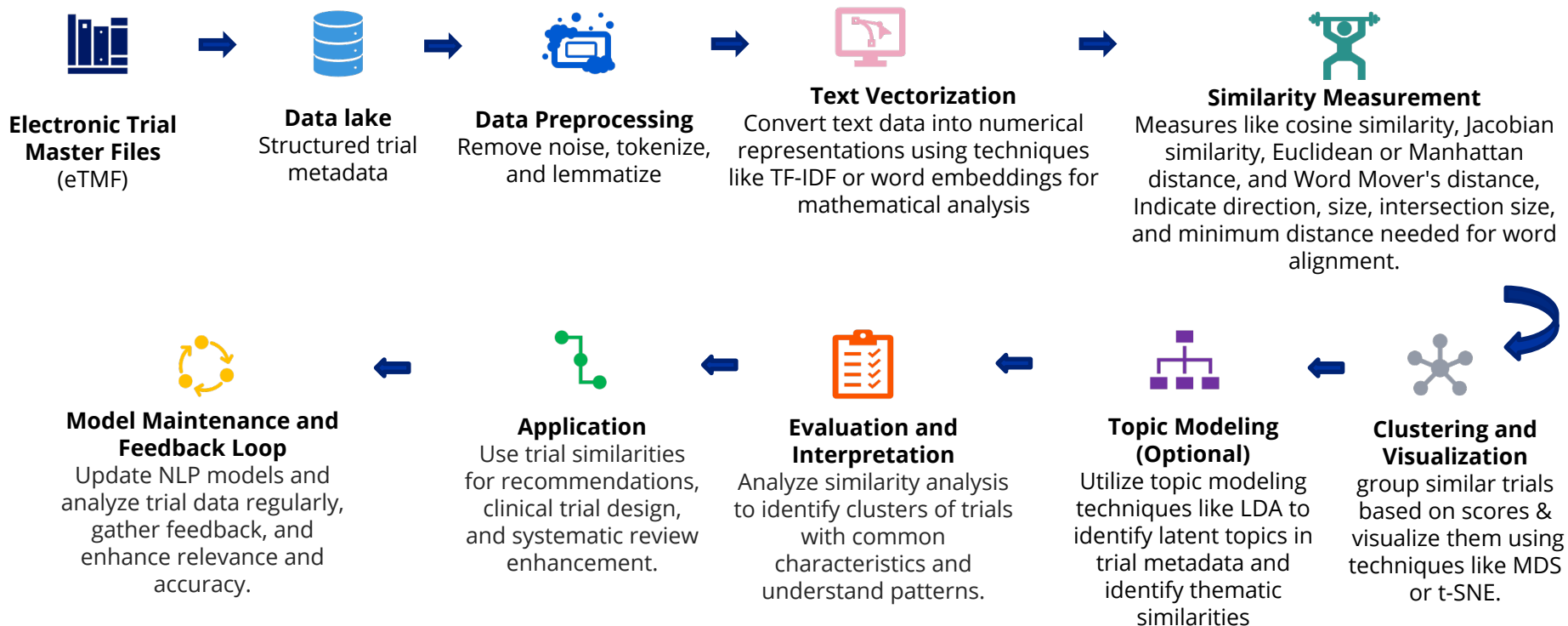
Protocol title, primary & secondary objectives, compound text

- ❖ This text represents the rationale of the trial (why / what / how/ when)
- ❖ These fields are long text, and it is not easy to search key information across, NLP can help

Trial 1	Trial 2	Key features	Scores
Abcd100	Xyz2020	Superiority, placebo, treatment1, ratio, difference, glycaemic control, exclusion criteria x, inclusion criteria x, hypoglycaemic, estimand....etc.	0.78
Abcd100	Xyz2001	Superiority, placebo, treatment1, exclusion criteria x, difference, hypoglycaemic, safety, estimand...etc.	0.69
Abcd100	Efg2301	Superiority,treatment1, difference, hypoglycaemic, safety, Relative risk..etc.	0.56



Process flow



Sentiment analysis & topic modeling

With the pervasiveness of social media platforms, people nowadays share opinions, critics, reviews, forum debates on social networks, which made sentiment analysis a domain of interest.

By using topic modeling, dominant themes of given texts from social media platforms, news articles, and purchase behavior can be detected

Sentiments – Flu vaccine

The vaccine is available now



I received the first vaccine.



Thanks

It's been 7 days and my arm still
hurts



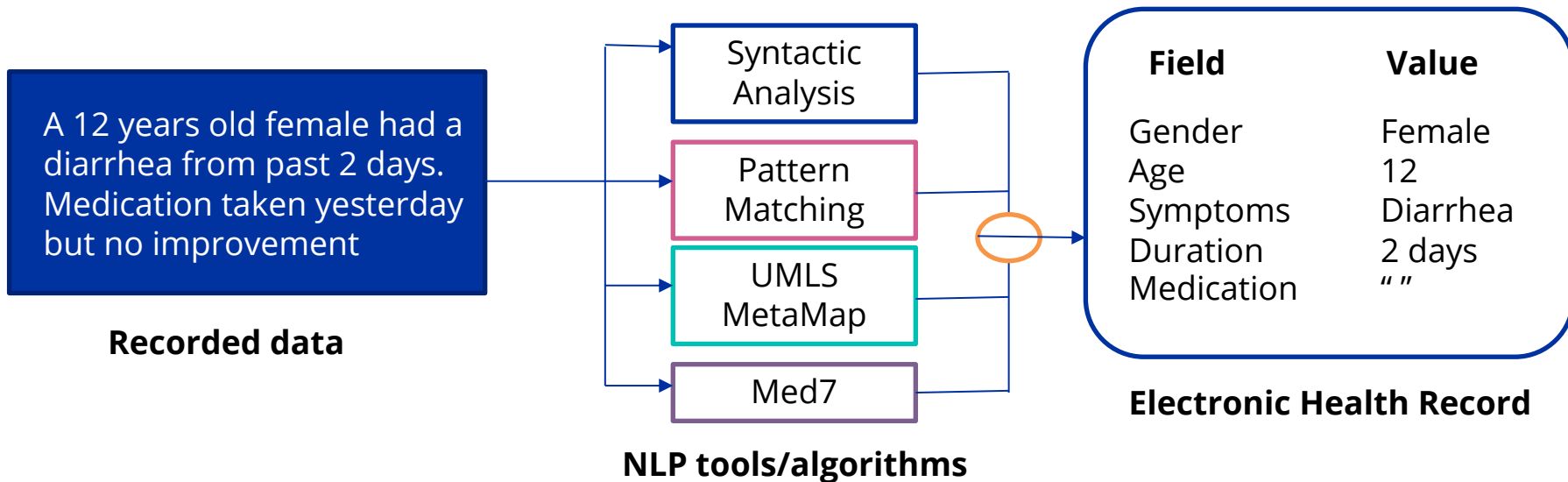
Topics – Flu vaccine

Opinions and emotions around
vaccines

Knowledge about vaccines

Vaccine administration

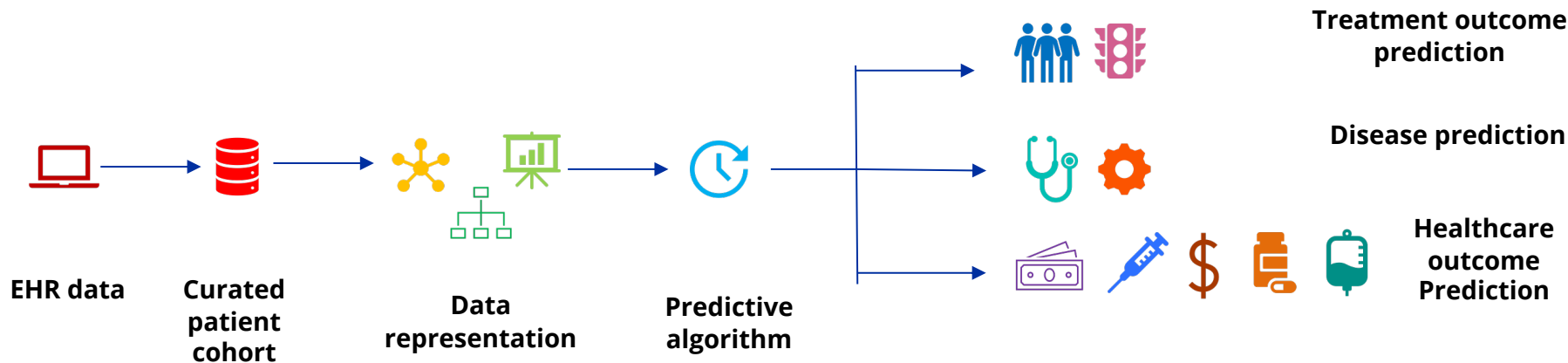
phuse System integration driven information extraction



Automating unstructured data to structure data. Reduce missing data in real world.

phuse NLP as important step in using AI models (1/2)

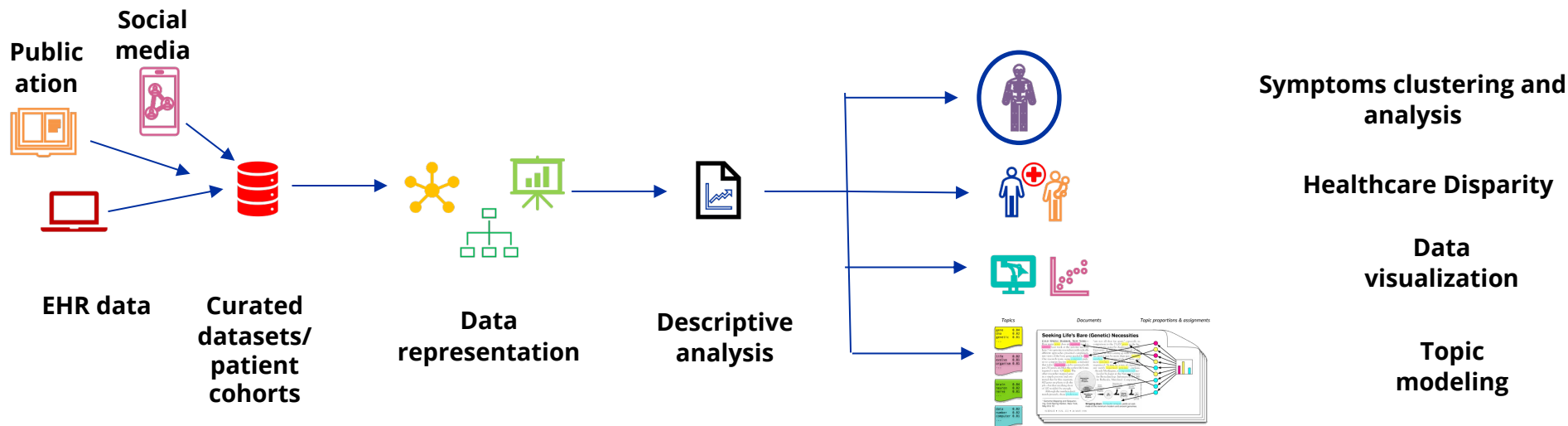
AI model for clinical Predictive Analysis



Trying to predict any clinical events

phuse NLP as important step in using AI models (2/2)

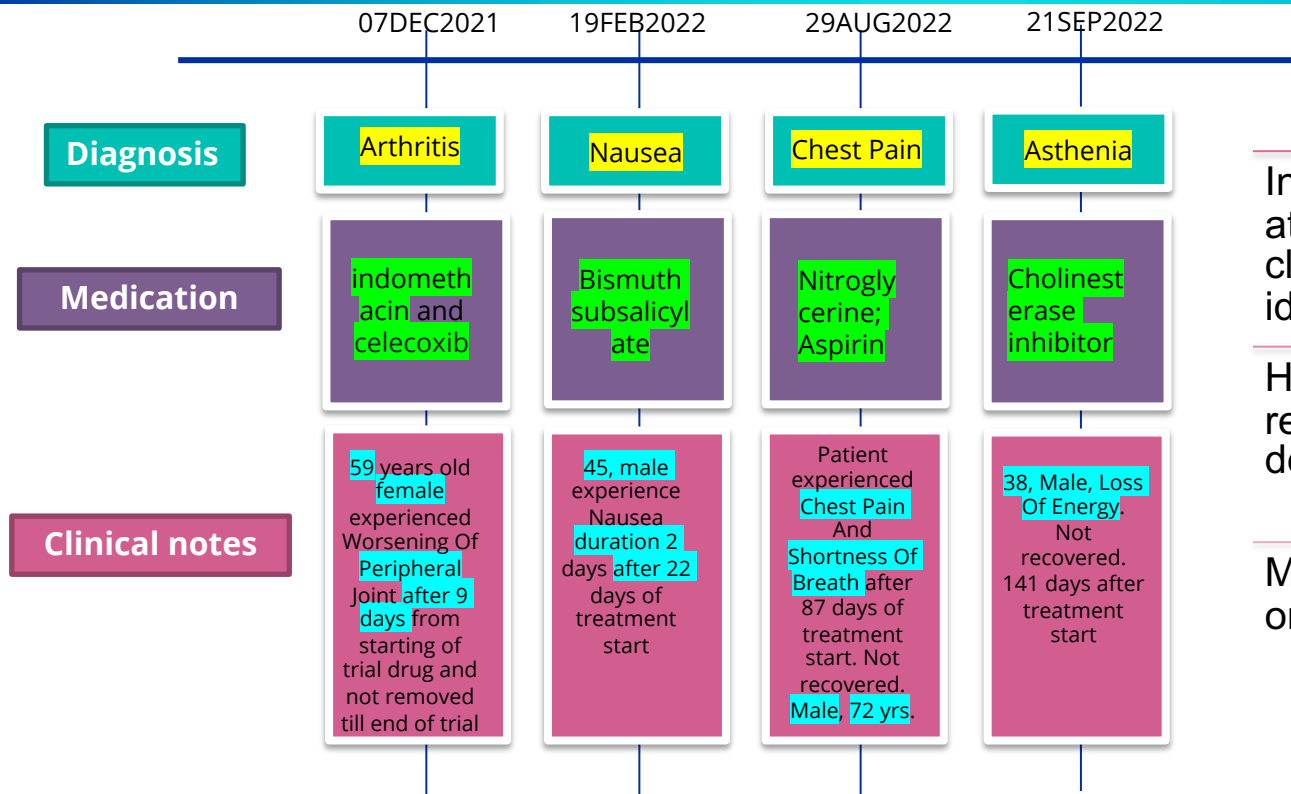
AI model for clinical Descriptive Analysis



Trying to discover pattern based on information we extracted. These patterns helpful for generating new clinical hypothesis



Patient record interpretation



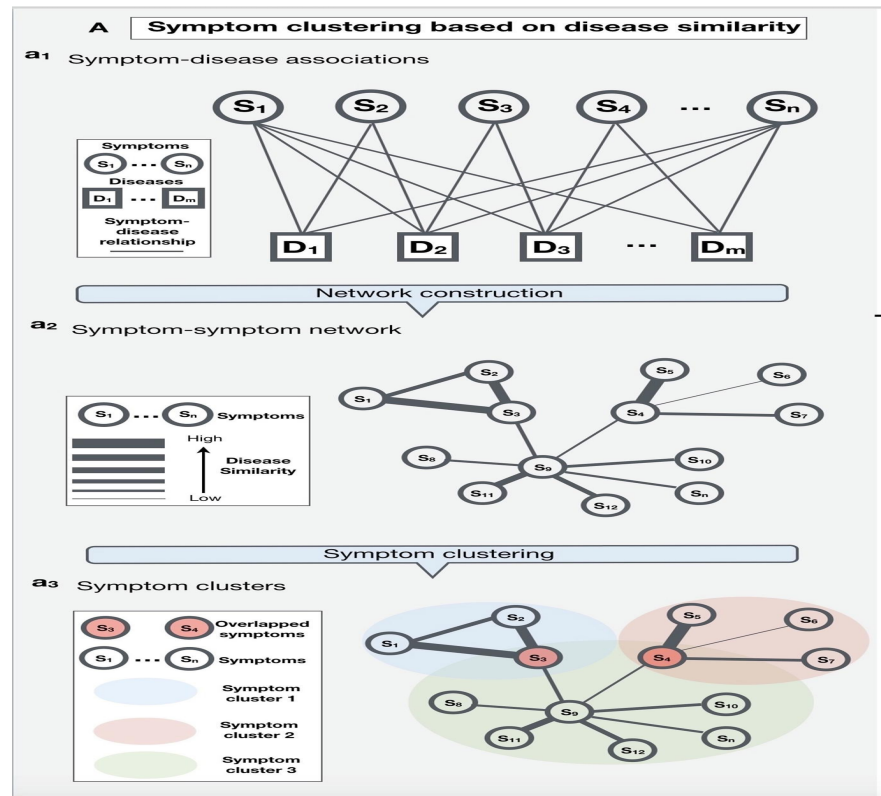
Interpreting chart reviews using attention mechanisms helps clinicians validate drugs and identify relevant features.

High attention weights highlight relevant features, making clinical decisions easier.

Model can be improved based on clinical needs.

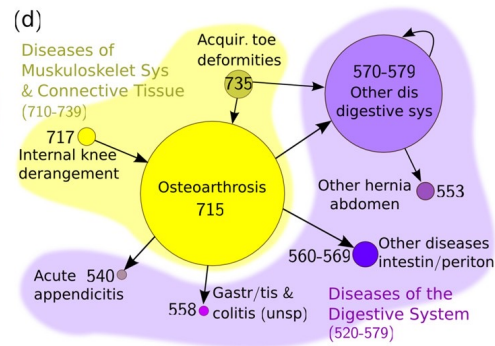
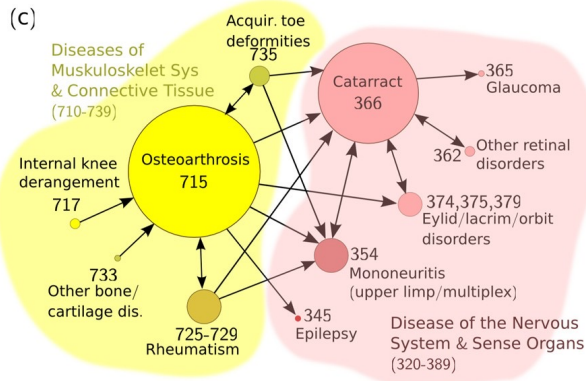
Symptoms cluster graphs

- Symptom clusters are ≥ 3 interrelated symptoms, aiding disease management, early detection, and reducing delays in interventions for patients and clinicians.
- Can be analyzed using t-tests to determine patient variables like age differences
- Chi-squared tests to identify differences in categorical variables like gender, diabetes status, and cancer stage.
- For instance, diabetes patients experience more numbness or tingling in fingers, hands, or lower extremities, while stage III cancer patients experience memory loss.



Symptoms cluster graphs: case

- NLP groups similar documents or sentences based on their content & meaning.
- It facilitates efficient categorization and analysis of large volumes of information.

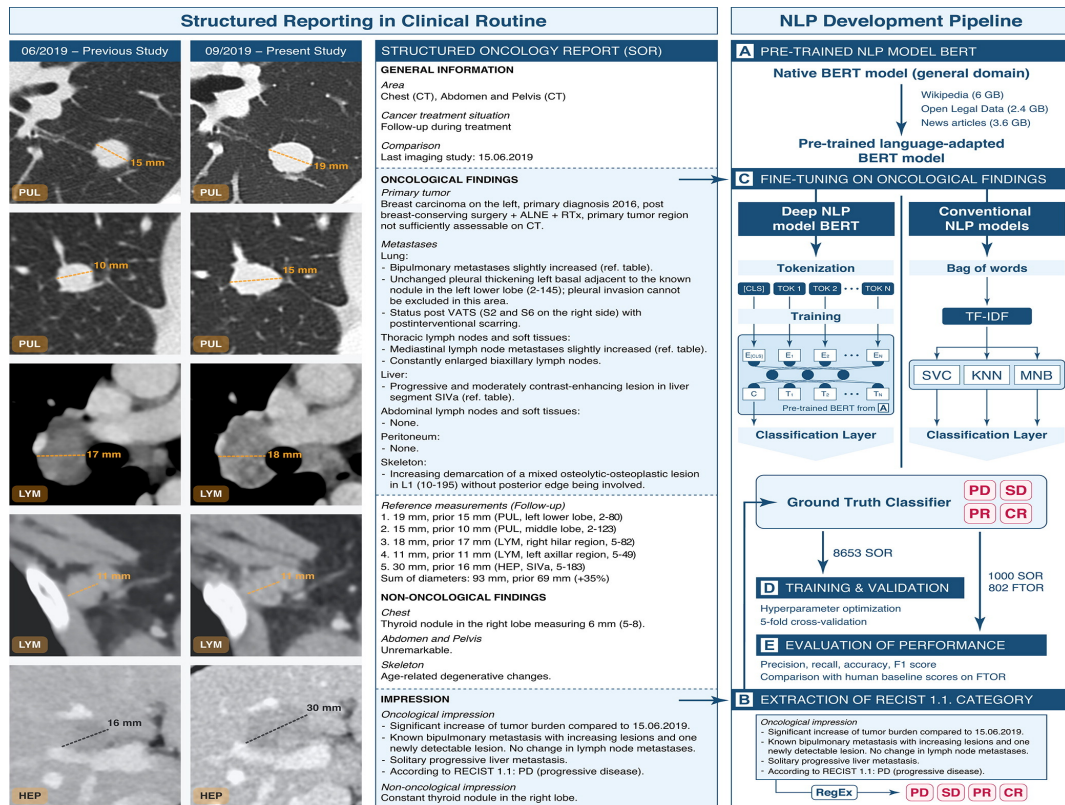


- Osteoarthritis is linked to various diseases, including musculoskeletal system and connective tissue disorders, nervous system and sense organ diseases, is shown in Fig (c) Osteoarthritis → Cataract was the most frequent statistically significant comorbidity found
- A cluster relating to digestive system disorders, including cholelithiasis, pancreas diseases, and gastrointestinal hemorrhage, is shown in Fig. (d)



Automate interpretation from images

- Saving time and decreasing the reading variability between laboratories and interpreters.
- More efficient, saving time in the preliminary classification of disease



Detect critical trends and patterns



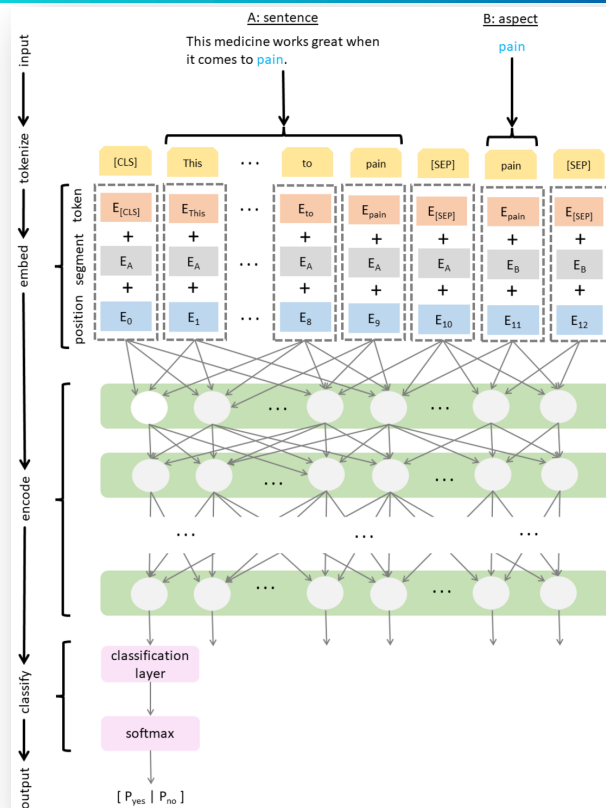
Extracting cause-effect entities from medical literature is an important task in medical information retrieval.



Causality between disease and symptoms between medications and side effects between genes and disease etc.



In medical literature cause and effect phrases in a sentence are not simply noun or noun phrases, rather they are complex phrases consisting of several word





Summary

NLP models still has challenges to identify features that matches the knowledge of human expert, but these models can improve over time.

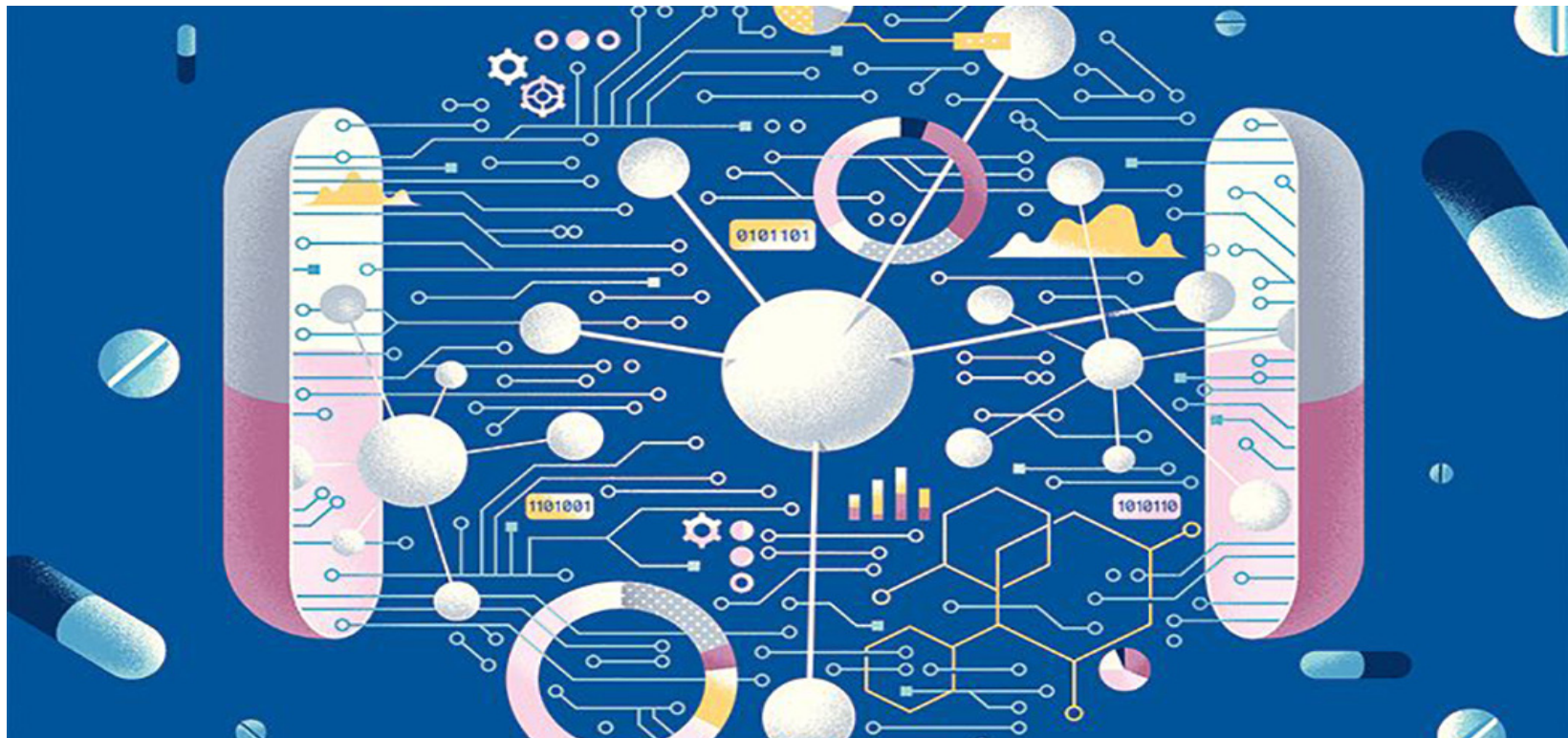
To advancing medical research it's high time that we need to invest more in advance NLP technologies.

Building NLP tools takes significant effort and cost

Limited understanding of technology and scarcity of skilled resources makes journey challenging

Identifying good use cases is challenging

phuse Quality of your natural language processing going to impact the quality of your analysis!



References

- <https://www.sciencedirect.com/science/article/pii/S153204642030109X>
- Sentiment analysis and topic modeling for COVID-19 vaccine discussions <https://link.springer.com/article/10.1007/s11280-022-01029-y>
- <https://saxon.ai/blogs/natural-language-processing-in-ai/>
- <https://www.nlplanet.org/>
- <https://blog.gramener.com/nlp-use-cases-pharma-healthcare/>
- <https://alldus.com/blog/5-applications-of-nlp-in-healthcare/>
- <https://medium.com/nlplanet/two-minutes-nlp-33-important-nlp-tasks-explained-31e2caad2b1b>
- Identifying temporal patterns in patient disease trajectories using dynamic time warping: A population-based study Alexia Giannoula, Alba Gutierrez-Sacristán, Álex Bravo, Ferran Sanz & Laura I. Furlong. <https://www.nature.com/articles/s41598-018-22578-1#Fig2>
- <https://www.sciencedirect.com/science/article/pii/S153204642030109X> Integrated network analysis of symptom clusters across disease conditions
- <https://pubs.rsna.org/doi/full/10.1148/ryai.220055> Deep Learning-based Assessment of Oncologic Outcomes from Natural Language Processing of Structured Radiology Reports
- <https://www.mdpi.com/2504-4990/4/2/21> The Case of Aspect in Sentiment Analysis: Seeking Attention or Co-Dependency?
- [The Effects of Feature Scaling: From Bag-of-Words to Tf-Idf - Feature Engineering for Machine Learning \[Book\] \(oreilly.com\)](#)



The Global Healthcare Data Science Community

Contact Channels

📞 UK +44 1843 609600
✉ office@phuse.global
🌐 phuse.global

Social Media

🐦 [@phusetwitta](#)
📘 [/phusebook](#)
▶ [/phusetube](#)
🏢 [/company/phuse](#)