

PP25: Unlocking the potential of unstructured data in Real-World Evidence

Leveraging Text for Enhanced Patient Classification and Regulatory Validation

Authors

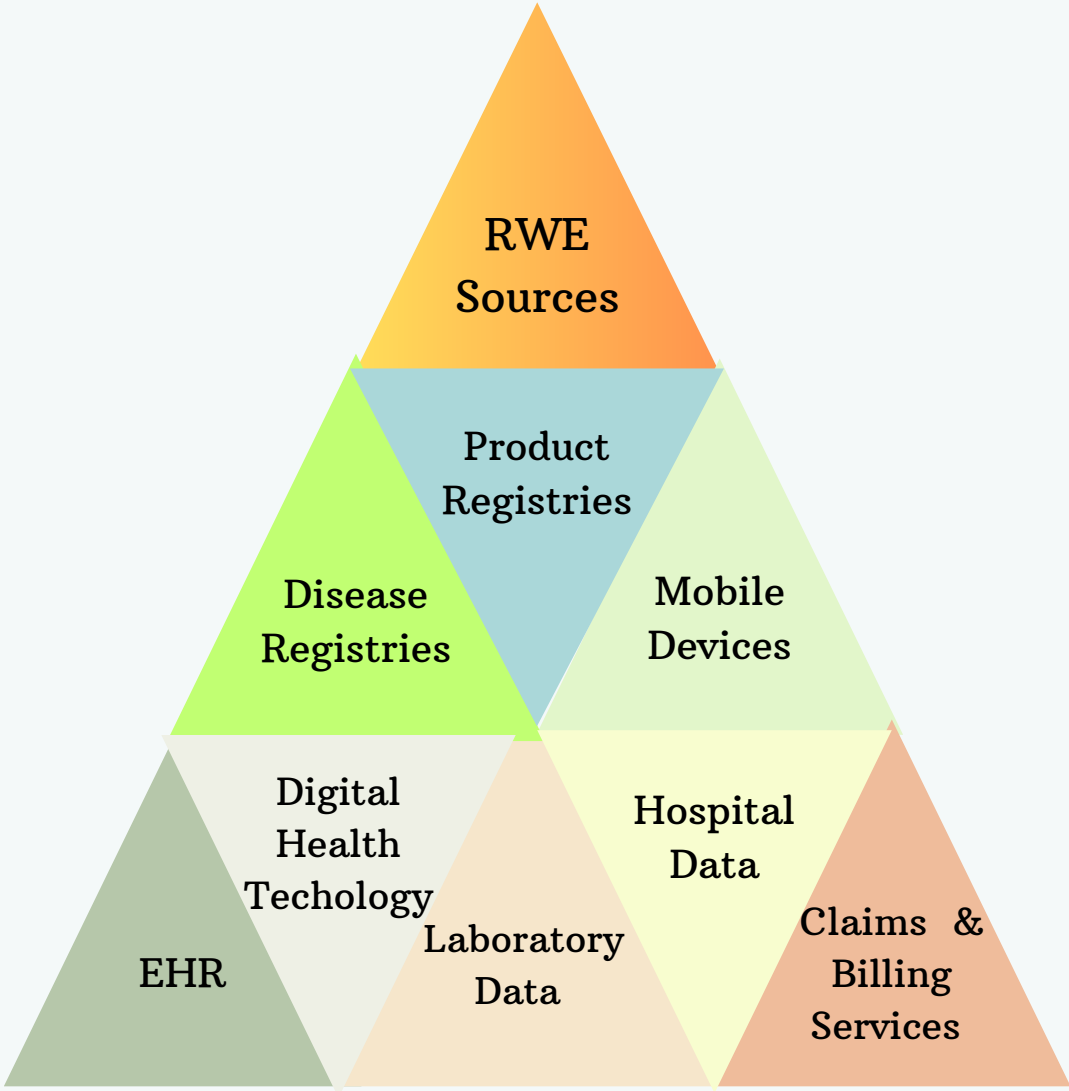
Maria Puri | Maria.Puri@pfizer.com
Nikolaos Kountouris | Nikolaos.Kountouris@pfizer.com
Georgios Panagiotakis | Georgios.Panagiotakis@pfizer.com

Pfizer Inc.

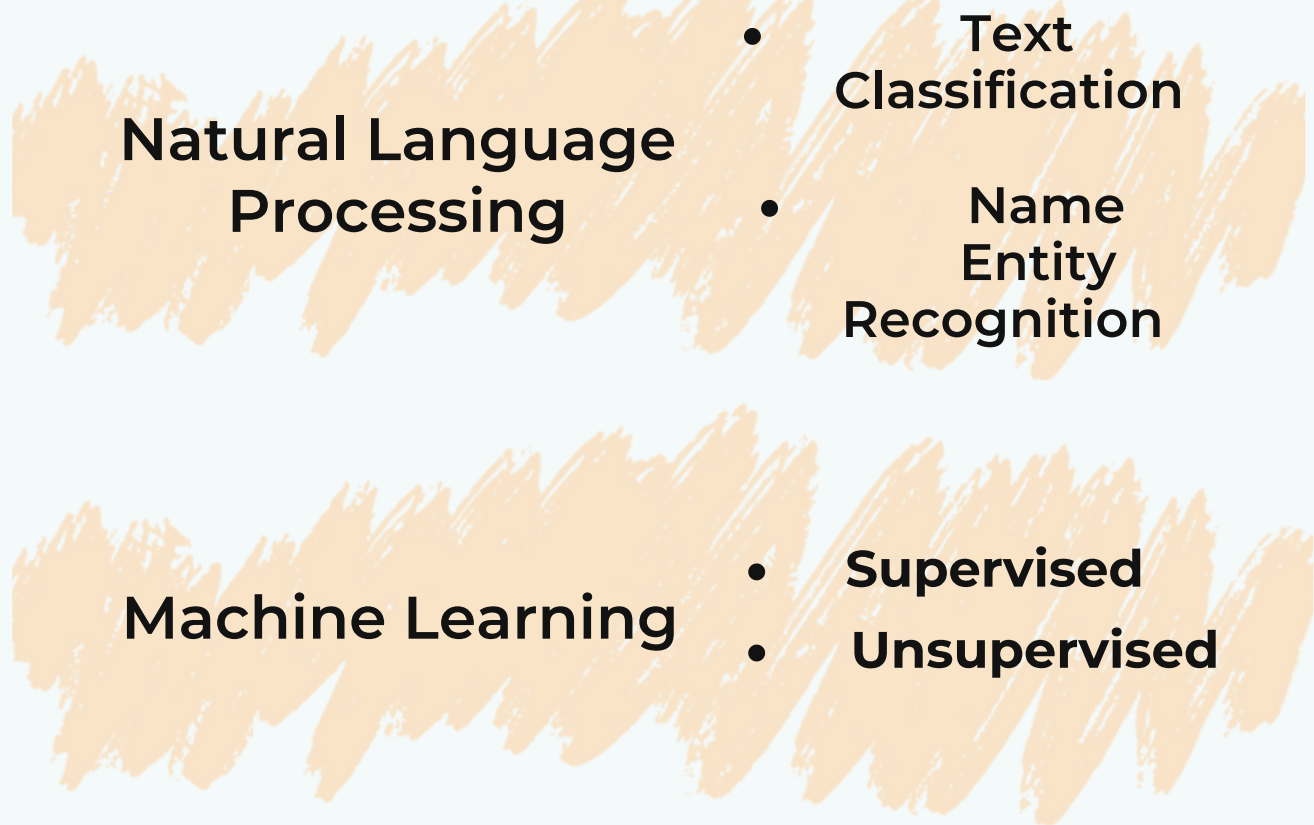
01 Introduction

Real World Data (RWD) has been commonly utilized for safety reasons for a while. Recently the Food & Drug Administration (FDA) created a framework to make use of RWD as evidence of effectiveness. Real World Evidence (RWE) is defined by the FDA as evidence obtained from studies, on RWD. The goals include enhancing the quality of RWD understanding research designs for RWE and creating tools to evaluate RWE. The utilization of RWD aids researchers and healthcare professionals in examining treatment trends and outcomes leading to improved decisions personalized healthcare practices and policy development. Technologies, like Natural Language Processing (NLP) can be applied to analyze data effectively

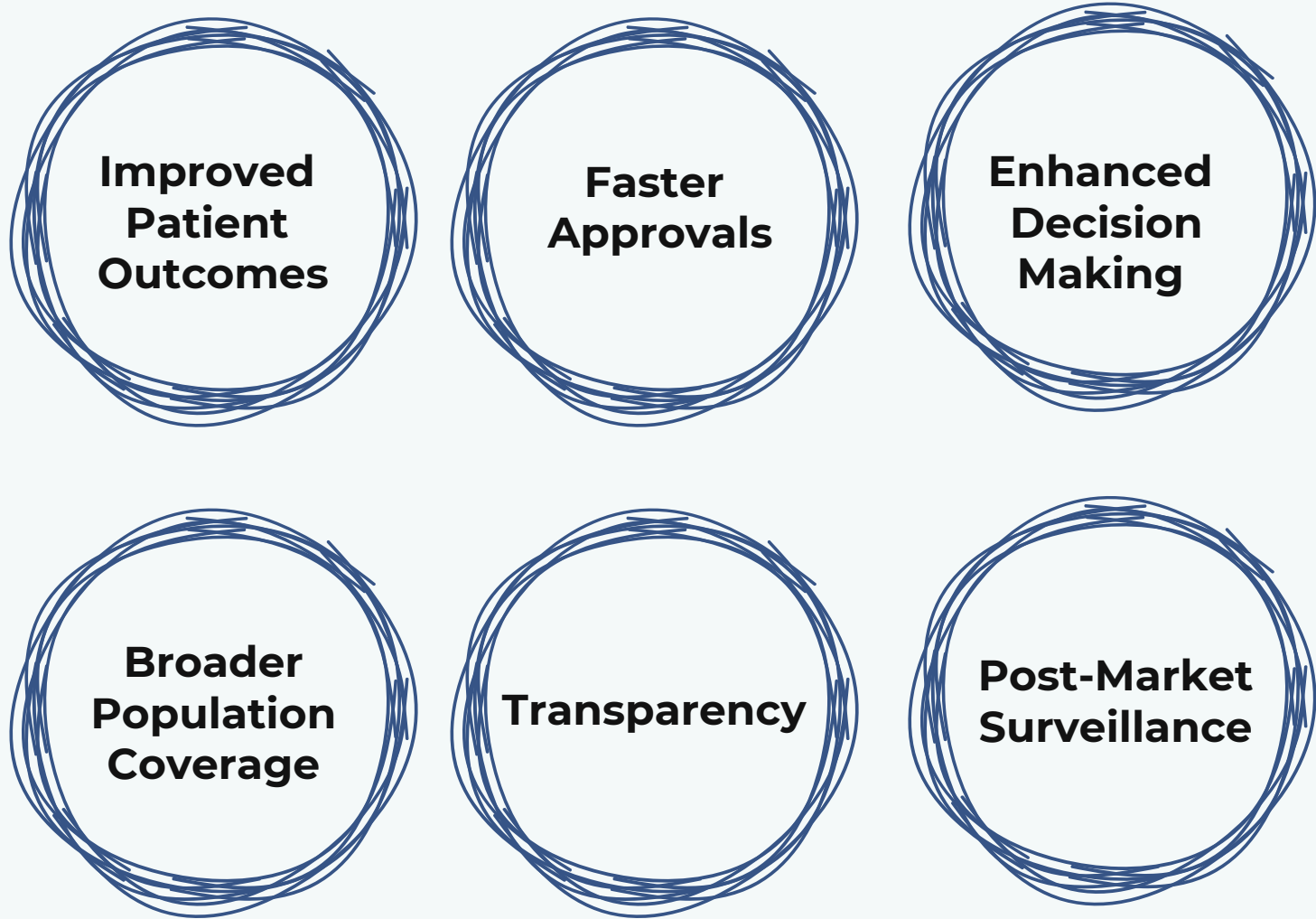
02 Data Sources



03 Techniques



04 Benefits of RWE & AI for Regulatory Decision



05 Case study - Patient Stratification Using Unstructured Clinical Text Data and Large Language Models (LLM's)

1. Electronic Health Records (EHRs)

are collected, including symptoms and physician observations



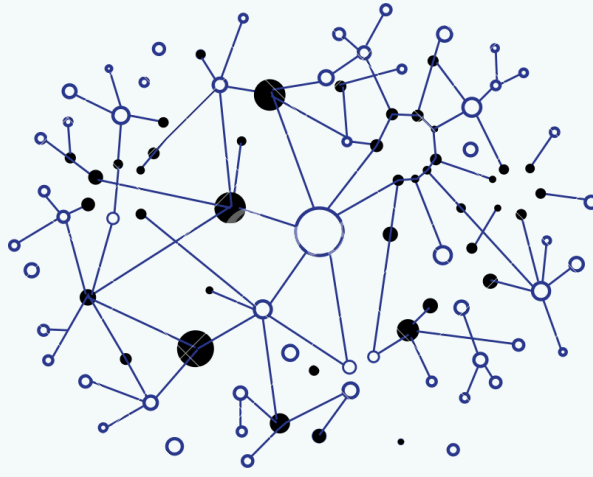
2. NLP Techniques

Clinical Entity Identification: Identifying key clinical entities
Relationship Extraction: Relationships between these entities

Entity Indetification

The patient, **John Doe**, presented with a persistent cough and was diagnosed with **bronchitis** on **January 5, 2024**. He was prescribed **amoxicillin** and instructed to take it twice daily.

Relationship Extraction



Unstructured Clinical Notes

Data Extraction

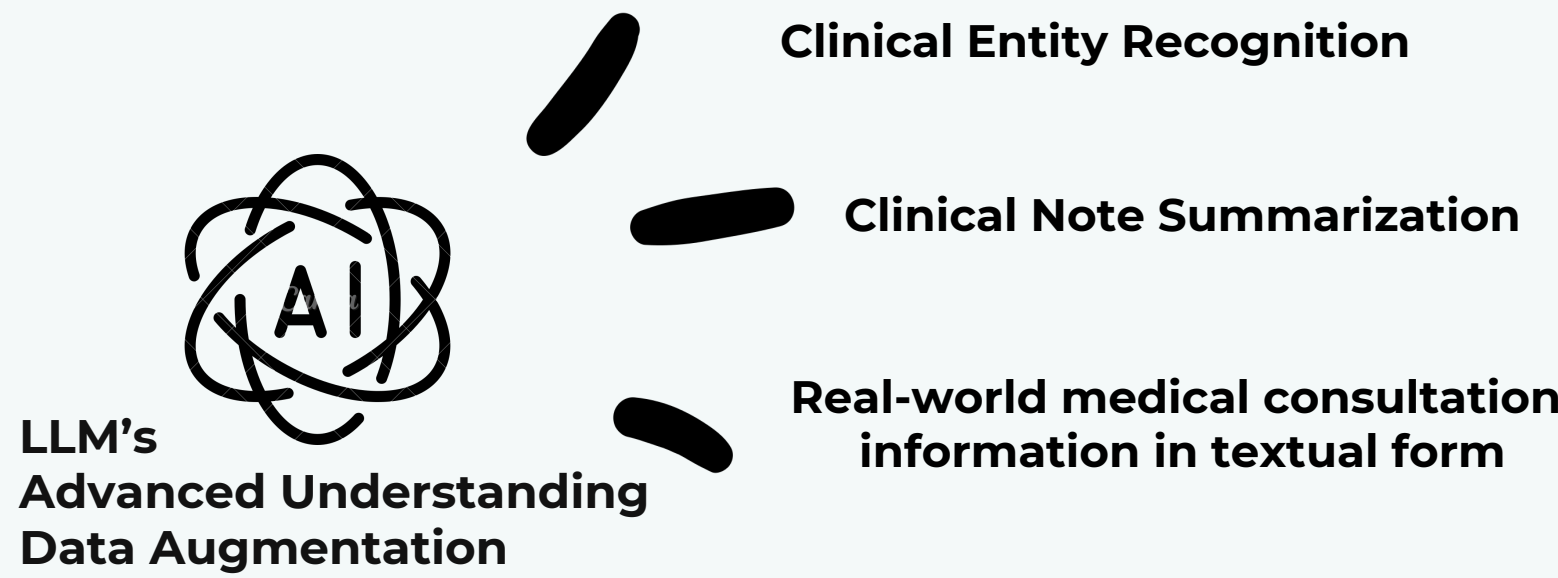
Natural Language Processing

Large Language Models

Database Storage & Integration

3. Role of LLM's

Advanced Text Understanding: LLMs (e.g. GPT-4) can process complex clinical language, providing insights that traditional NLP models may miss. They can predict patient outcomes or identify subgroups based on detailed descriptions
Contextual Analysis: LLMs consider the full context of documents, improving the accuracy of entity and relationship extraction, and distinguishing between similar symptoms with different causes



4. Database Storage & Integration

Data Transformation: Convert unstructured text into a structured format for database storage, using standardized data schemas
Data Warehousing: Store this structured data in clinical data warehouses for efficient querying and retrieval

06 Conclusion

Incorporating unstructured information into real world evidence (RWE) has the potential to enhance the quality of evidence aid, in making healthcare choices and improve patient outcome results. The key to achieving success lies in tackling obstacles and leveraging tools to extract insights, from diverse data outlets

Related Literature

[1] Wang Z, Shah AD, Tate AR, Denaxas S, Shawe-Taylor J, Hemingway H. Extracting diagnoses and investigation results from unstructured text in electronic health records by semi-supervised machine learning. PLoS One. 2012;7(1):e30412.
[2] FDA to Support Potential Use of Real-World Data to Generate Real-World Evidence for Regulatory Decision-Making (biopharminternational.com)
[3] May, 25 (1970) Ai, Big Data and real world evidence - the challenges and opportunities, pharmaphorum. Available at: <https://pharmaphorum.com/views-and-analysis/ai-big-data-and-real-world-evidence-the-challenges-and-opportunities> Available at: <https://hitconsultant.net/2019/02/27/state-of-ehr-data-real-world-evidence-rwe/> (Accessed: 30 July 2024).