

PHUSE EU Connect 2024

DH05: Artificial Intelligence Potential Use Cases in Clinical Data Engineering

Mounika Ramesh Pillamarapu, Senior Clinical
Data Engineer



parexel®

Agenda

- › Handling missing data in clinical trials
- › Deduplication
- › Data augmentation
- › Error correction
- › Real world data sanitization



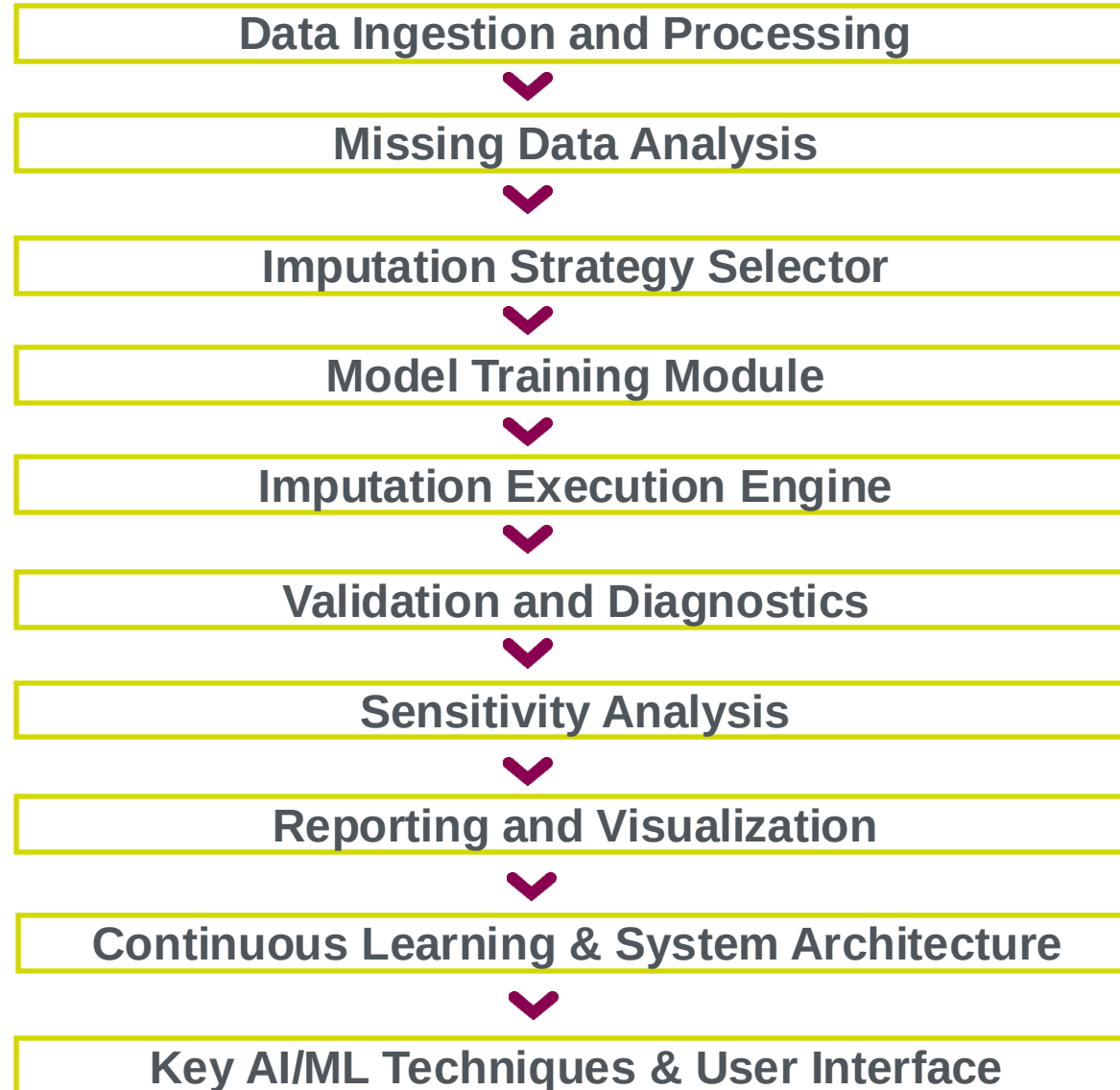
Introduction

- AI is revolutionizing clinical trials, enhancing efficiency, accuracy, and patient care
- Key applications
 - Data analysis
 - Remote monitoring in decentralized trials
 - Extracting information from unstructured sources

Handling missing data in clinical trials

- Different AI models to handle missing data
 - Advanced Machine learning models (imputation methods)
 - Predictive modeling for missingness
 - Natural Language Processing (NLP) for Unstructured Data
- Framework for AI models

Integration of AI and EDC



Deduplication

Category	› Structured Numerical and Categorical Data
AI models	› ML Classifiers (Random Forest)
Data type	› Patient demographics, lab results, vital signs
Example	› Random Forest can handle mix of numerical and categorical data, e.g.: age, weight

Data augmentation

- Set of techniques used to increase data quantity without collecting new patient data
- AI models
 - Generative Adversarial Networks (GANs)
 - Recurrent Neural Networks (RNNs)

Error correction

- Data entry mistakes, protocol deviations
- AI models
 - Machine Learning algorithms : predictive error flagging
 - Natural Language Processing (NLP) : free-text field analysis

Real-world data sanitization

- Use of real-world data for research use while maintaining patient privacy and data integrity
- AI models
 - NLP
 - Pattern Recognition

Challenges in using AI models

- › Regulatory approval
- › Data privacy and protection
- › Computation resources
- › Auditability and reproducibility
- › Generalization of models across different studies

Conclusion

- › AI enables accurate and consistent processing of high volume of data
- › Addresses complex privacy challenges
- › Technical, ethical and regulatory domain
- › Robust validation of the models
- › In future likely to have hybrid approach of combining AI's power with traditional methods

Q&A ?

› Mounika Pillamarapu
mounikaramesh.pillamarapu@parexel.com

› Deepak Pakalapati
deepak.pakalapati@parexel.com

With Heart™

parexel®



Thank you

parexel[®]