WITH STANDARDS – UNLOCK THE POWER OF DATA

cdisc

2022
JAPAN
INTERCHANGE
13-14 JUNE | VIRTUAL EVENT

Use of Artificial Intelligence (ML/NLU/NLP/NLG) in Regulatory (Scientific) Documents Authoring

FARHA FEROZE

Meet the Speaker

Farha Feroze

Title: Product Manager

Organization: Symbiance Inc

Farha Feroze is a Product Manager at Symbiance. She has been in the path of Product Management and Business Analysis for a good few years. She is known and appreciated for her sharp acumen and logical mind that easily grasp new business concepts. Her current focus is in providing business solutions that use advanced technologies such as Machine Learning (ML), Natural Language Processing (NLP)& Natural Language Generation (NLG) backed by her experience in Life Sciences Software Applications. She also specializes in designing and documenting business models, rules, workflows, reports and process workflows.

Disclaimer and Disclosures

- *The views and opinions expressed in this presentation are those of the author(s) and do not necessarily reflect the official policy or position of CDISC.*

Agenda

1. Background of CSR and automation
2. Process ad steps
3. Benefits of AI
4. Conclusion

Background of CSR and automation

What are we trying to solve?

Clinical Study Report

- Creating Clinical Study Report (CSR) is highly manual and time consuming.
- The content of CSR is constructed from Protocol, SAP, Safety narratives, in-text tables and medical writer's interpretation of study results.
- No automotive way to handle Manual Errors and multiple iterations

Why automation ?

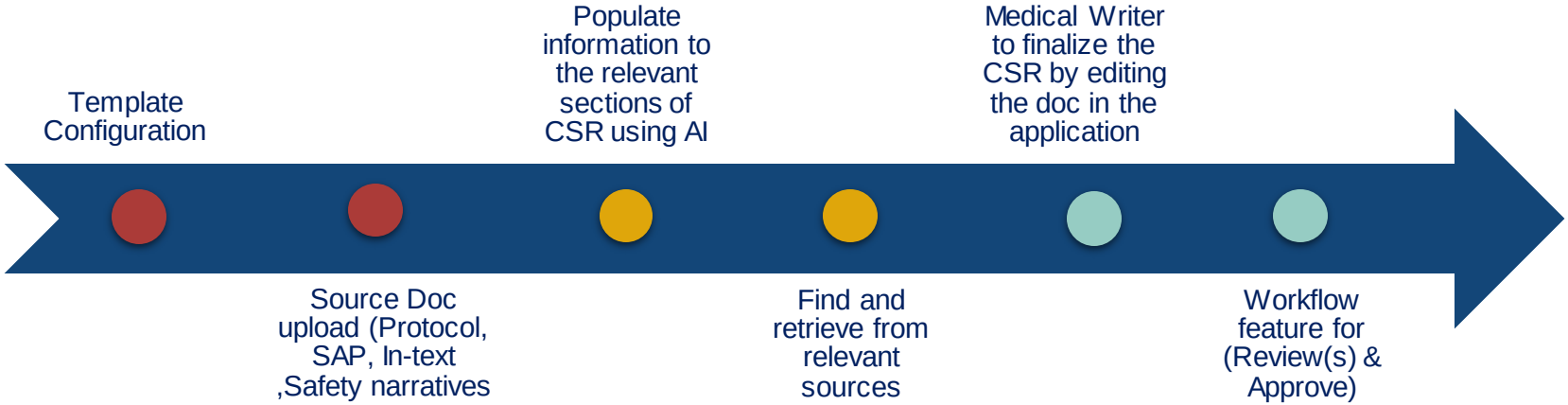
- Significant manual effort can be saved using AI techniques so MW can focus on the Interpretation and discussion point.
- Effective use and reuse of the template for an organization or therapeutic area or compound.
- The ML prediction algorithm along with NLP/NLG is used to identify/effectively used to read/write the texts from other documents such as Protocol, SAP, In-text Table & etc.,
- The CSR Template will be created based on ICH –E3 guidelines
- The medical writers can focus more on discussion points and interpretation of study results.

Blank CSR Template

1. TITLE PAGE
2. STUDY SYNOPSIS
3. TABLE OF CONTENTS
4. LIST OF ABBREVIATIONS & DEFINITION OF TERMS
5. ETHICS AND REGULATORY APPROVAL
6. INVESTIGATORS AND STUDY ADMINISTRATIVE STRUCTURE
7. INTRODUCTION
8. STUDY OBJECTIVES
9. INVESTIGATIONAL PLAN
10. STUDY POPULATION
11. RESULTS
12. SAFETY EVALUATION
13. DISCUSSION AND OVERALL CONCLUSIONS
14. TABLES, FIGURES AND GRAPHS
15. REFERENCES
16. APPENDICES

Process and Steps

Steps



Use of ML/NLP/ NLG

Prediction Accuracy :

- The text from PDF document such as Protocol/SAP are extracted using ML/NLP and the texts are understood by the system engine.

Identifying Individual Sections:

- ML model is used to predict the best matching content from various source documents(Protocol, SAP, In-text etc.,) for all the sections in a CSR.
- Named Entity Recognition (NER) which is a subprocess of NLP is used to identify the drug names, dosages, duration of drug, sponsor name and protocol number

Title Page & Synopsis:

- Customized ML Algorithm used in title page and study synopsis.

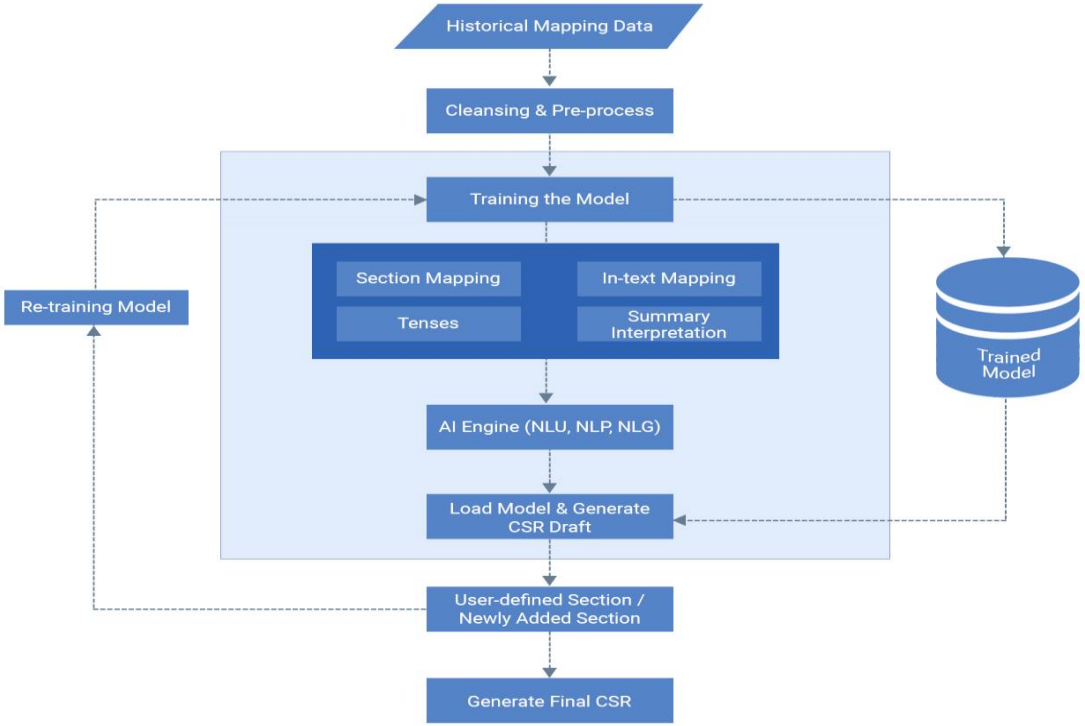
Deep Learning Model usage:

- The section 10 (study population), section 11 (results) and section 12(safety evaluation) is more of In-text tables that utilizes Deep Learning Model to find the best matching In-text tables.

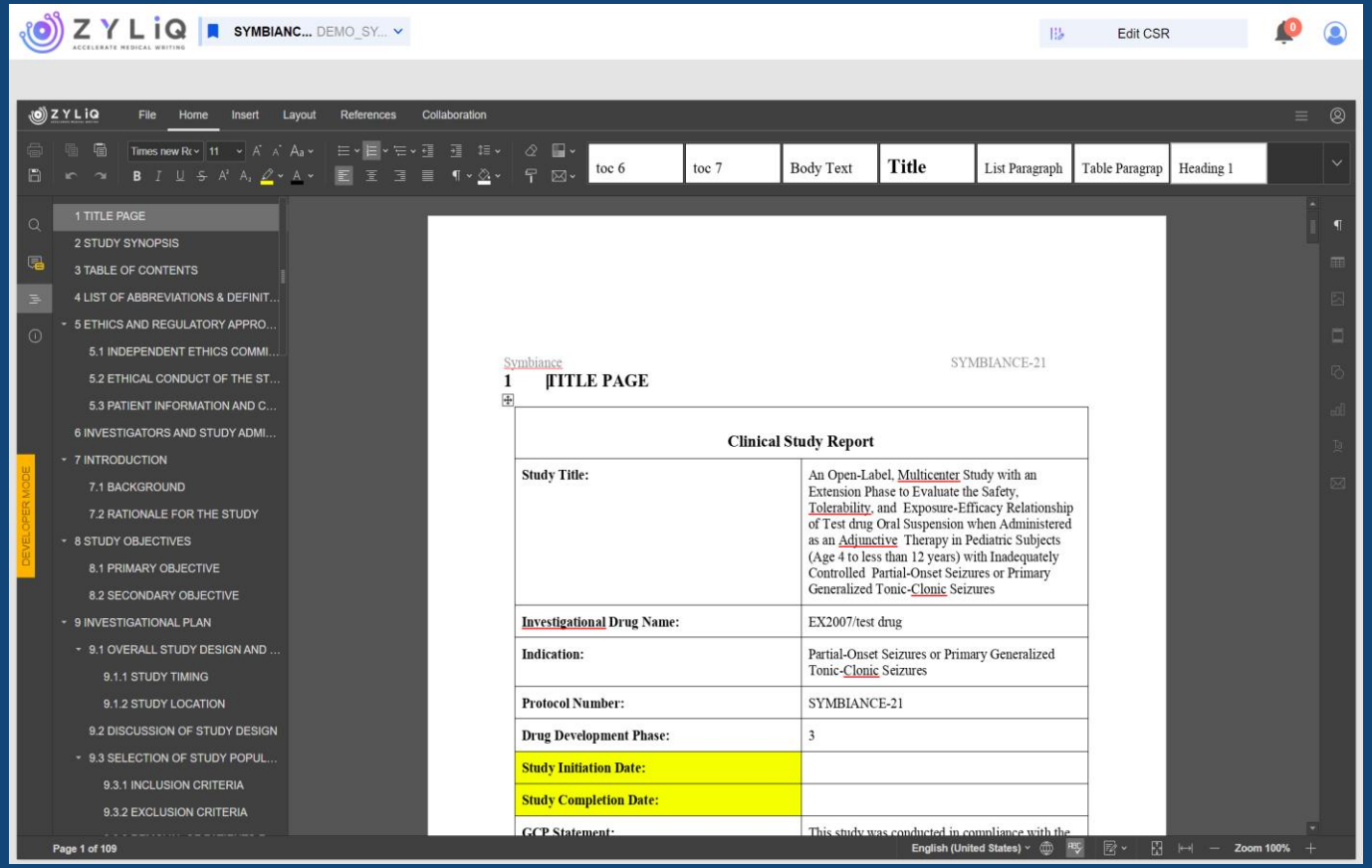
Handling of Un-structured data

- Extracting data from source documents (PDF/Word, formatting and special symbols, subscript & superscripts and diagrams)
- Summarizing large paragraphs using NLU & NLG for CSR synopsis
- Reading the data from the unstructured TLFs & table details extraction of data and interpretation in simple English
- Tense conversion
- Section mapping challenges
- Optimization for performance in ML models

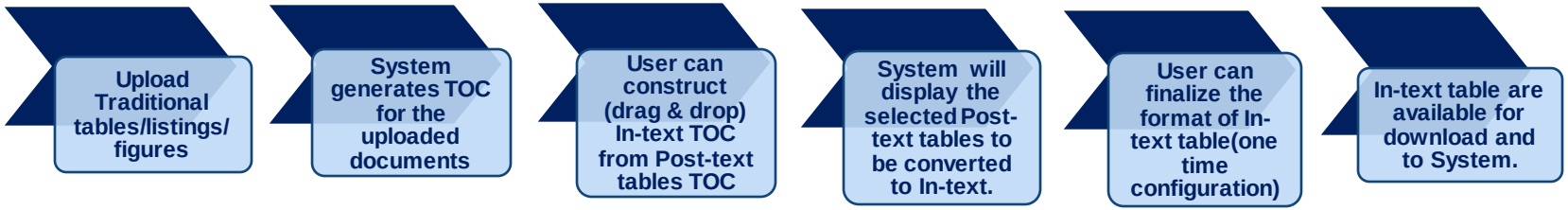
How AI is applied??



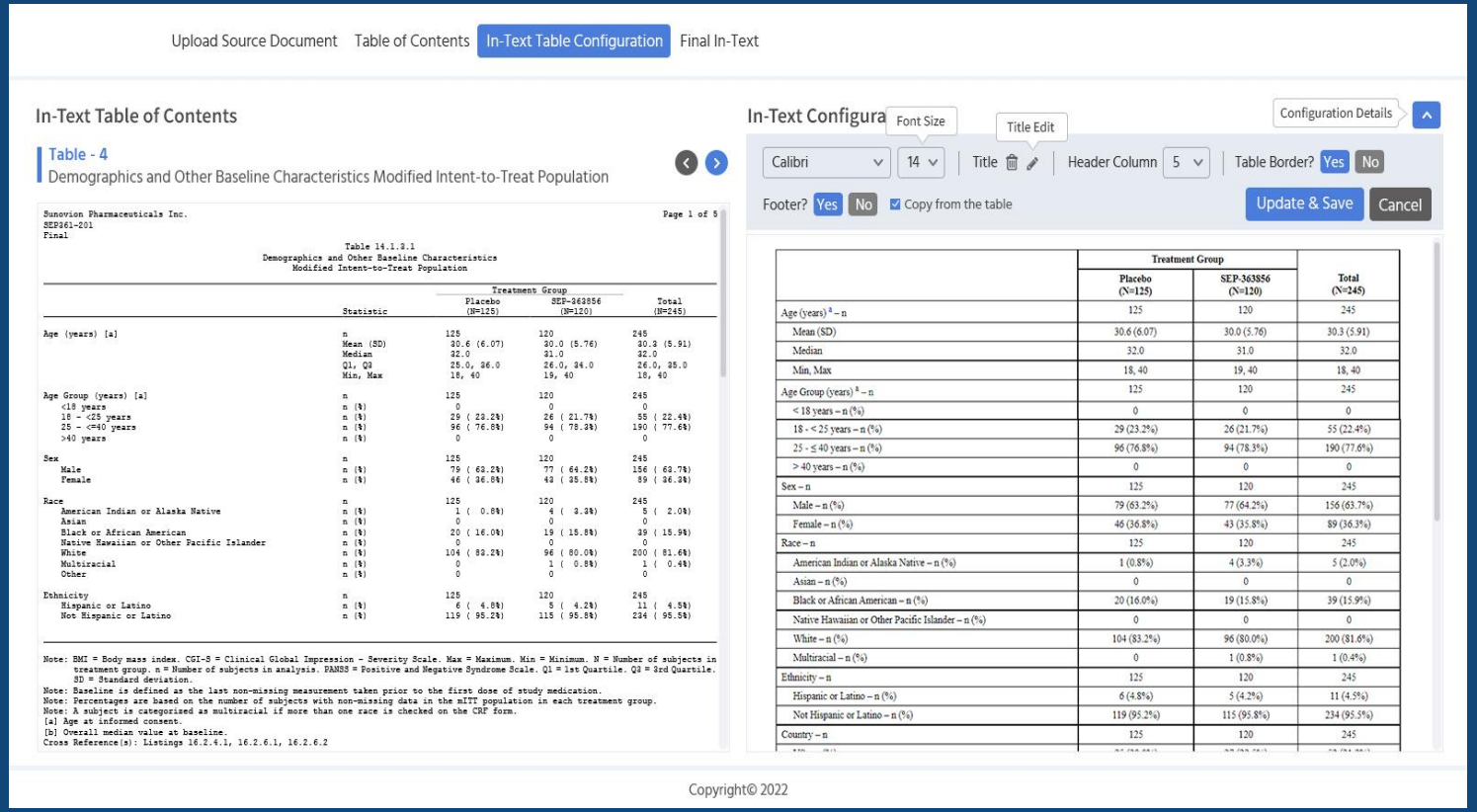
Application Edit Screen

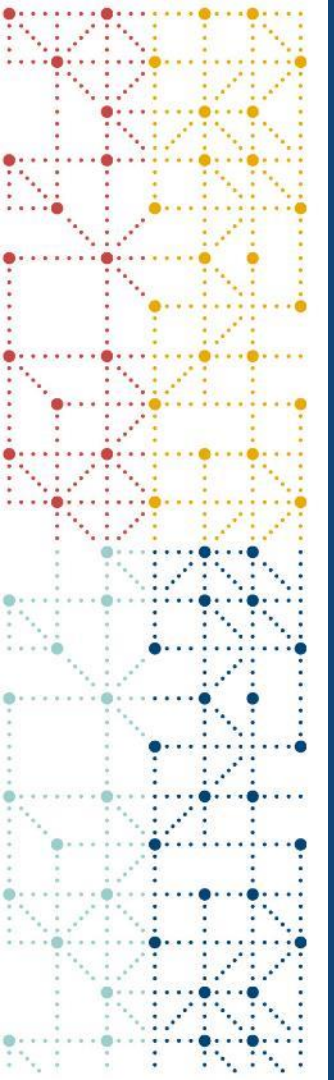


Converting Post-text tables to In-text tables



Application Edit Screen





Benefits of Automation using AI

Benefits of AI

- The system is designed to generate pre-filled CSR with information from Protocol, SAP and other sources as per ICH-E3 guidelines in the respective sections of the CSR
- The system will save 60-70% of time of the medical writers when they write the CSR using the application.
- Medical writers can focus more on interpretation of study results and their discussion points.
- System can be configured as per the sponsor's needs in the workflow integration

Conclusion

Conclusions

- MW can add or modify the content by section within the system.
- Multi-Authoring (can have more than one authors for same study)
- Conversion of post-text TLF to In-text tables
- ZYLiQ editor functions as MS word and have the same look and feel.
- Traceability Report(source file to final CSR mapping), Audit Log & Version History
- Workflow integration(Author(s), Reviewer(s) & Approver)
- Narrative writing from Patient profiles/ Subject Listing .
- In-text table interpretation in simple English.
- Automatic Tense conversion (Present/future tense to past tense)

Thank You!

Contact:

Farha Fathima Feroze

Product Manager

Mobile: +1 (609) 945 7431