TrialScribe

Automating Clinical Trial Template Creation for Compliance and Efficiency

AI&ML@SJSU

Project Description & Goal

- Develop AI-driven system to generate clinical trial document templates
- Ensure alignment with FDA, EMA, and ICH regulations
- Improve speed, accuracy, and compliance in documentation workflows
- Provide scalable solution adaptable across trial phases

Problem Motivation

- Manual documentation is time-consuming and error-prone
- Regulatory requirements demand precision and standardization
- Growing complexity of multi-site, multi-phase clinical trials
- Need for automation to reduce bottlenecks and human errors

To overcome a common issue in starting clinical trials, such as an insufficient number of patients enrolling, the clinical site location agent selects a clinical trial site based on data from multiple sources (such as Medicare) on disease occurrence.

Then to start the trials, the clinical protocol writing agent creates a clinical trial protocol specific to the patient population and location.

Datasets

- Regulatory guidelines (ICH E6(R2), FDA, EMA)
- Historic clinical trial protocols and reports
- Case report forms (CRFs) and informed consent documents
- Annotated compliance checklists and audit reports
- Centers for Medicare and Medicaid Services
- CDC Data Sources
- Healthcare Cost and Utilization

High Level Plan

- Phase 1: Data collection and preprocessing
- Phase 2: AI model fine-tuning for template generation
- Phase 3: Integration with trial management systems
- Phase 4: Pilot deployment and validation (if time allows)
- Phase 5: Iterative improvement and scaling (if time allows)

Related Work/Research

- AI in regulatory document automation
- Natural Language Processing for compliance checks
- Prior industry solutions (e.g., Medidata, Veeva Vault)
- Research on agentic AI for healthcare and life sciences

Project Evaluation Metrics

- Regulatory compliance accuracy (%)
- Reduction in document preparation time
- Error detection and correction rates
- User satisfaction and adoption metrics
- Scalability across therapeutic areas

Tech Stack & Desired Member Experience

- Python, PyTorch/TensorFlow for AI modeling
- NLP frameworks (Transformers, spaCy)
- Integration with clinical data management systems
- Members gain experience in AI for healthcare compliance
- Exposure to real-world regulatory challenges

Roles Needed

- AI/ML Engineers – model development & fine-tuning
- Clinical Domain – regulatory alignment
- Data Engineers – dataset curation & pipelines
- Software Engineers – system integration
- QA/Validation – compliance testing

Sign Up Here! @Brave123

https://forms.gle/26JyfpKzS17zrZKb8