Medical Research GenAI Processor

Enhancing Medical Research Analysis through Generative AI

ProcDNA

Presented by:

Kanishka Raj
Contact: +91-9612223176
Email: thisisraj.57@gmail.com

Contents

Objective

- **Purpose:** To simplify and enhance the analysis of complex medical research papers.
- **Goal:** Automatically extract critical entities, summarize insights, and evaluate document coherence.
- **Target Audience:** Medical researchers, healthcare providers, research organizations, and academic institutions.

Approach

Data Extraction & Processing: Utilize NLP-based entity recognition to extract key entities like drugs, genes, proteins, and efficacy/safety metrics.

Query and Document Summarization: Use retrieval-based and LLM models to retrieve relevant documents and generate summaries.

Evaluation Metrics: Include BLEU Score for coherence, readability scoring, and accuracy of entity linking.

Privacy Guardrails: Redact patient information and sensitive identifiers from outputs to maintain privacy compliance.

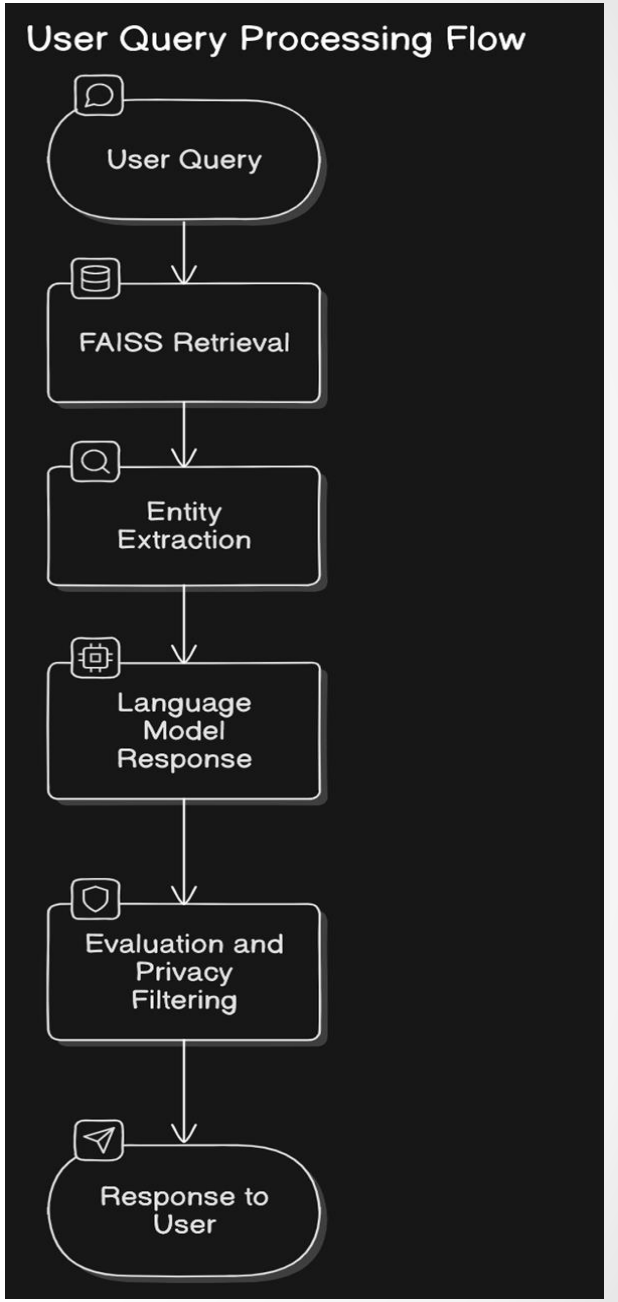
System Architecture

Data Extraction & Processing: Utilize NLP-based entity recognition to extract key entities like drugs, genes, proteins, and efficacy/safety metrics.

Query and Document Summarization: Use retrieval-based and LLM models to retrieve relevant documents and generate summaries.

Evaluation Metrics: Include BLEU Score for coherence, readability scoring, and accuracy of entity linking.

Privacy Guardrails: Redact patient information and sensitive identifiers from outputs to maintain privacy compliance.



System Architecture Contd.

Data Preprocessing Pipeline

- Data Ingestion: Secure upload of research papers, clinical notes, or other documents from various sources.
- Preprocessing Steps: Text cleaning, tokenization, entity extraction (drugs, genes, proteins, etc.), de-identification for privacy.
- Storage: Preprocessed data stored in a scalable cloud database (e.g., Amazon S3 or Azure Blob Storage).

Generative AI Processing Engine

- Vector Store (e.g., FAISS): Embeddings for document similarity and retrieval.
- Generative Model (e.g., Google Generative AI): Processes queries with fine-tuned prompts, extracts insights, and generates summaries.
- Privacy & Safety Layer: Enforces guardrails (removing patient identifiers and dates).

Knowledge Base Integration

- External Knowledge Bases: Real-time APIs for medical data enrichment (e.g., PubMed, clinicaltrials.gov).
- Semantic Search: Retrieves contextually relevant information to supplement model responses.
- Entity Linking: Matches extracted entities to known data in the knowledge base to improve accuracy.

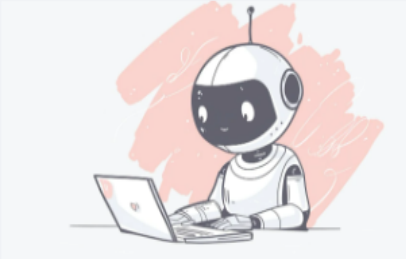
Result Storage and Retrieval

- Results Database: Stores processed summaries, entity data, and session history in a NoSQL database (e.g., MongoDB, DynamoDB).
- User Query History: Retains past queries and responses for easy retrieval.
- Data Access: Secure API for authorized retrieval, supporting data exports and downloadable reports**.**

Infrastructure Costs

- Cloud Services: AWS/Azure/Google Cloud for storage, compute (AI/ML processing), and API integrations.
- Cost Management: Pay-per-use models for compute, storage, and API requests to optimize expenses.

Chat Bot Interface



Empowering medical research through AI

Status: Online

About

This AI-powered application analyzes and summarizes complex medical research papers, extracting critical entities and providing enriched context.

Medical Research GenAI Processor

Document store loaded successfully!

Ask your Query

Type your question here...

Submit Query | Clear Input | Delete Last Query

Key Challenges & Solutions

Challenge: Handling diverse and extensive biomedical vocabularies.

Solution: Use customized regex-based entity recognition, focusing on medical terminology patterns.

Challenge: Maintaining high summarization accuracy with minimal deviation.

Solution: Utilize sentence-level BLEU scoring and manual tuning of the prompt.

Challenge: Data privacy in generated responses.

Solution: Implemented regex-based anonymization to redact identifiers.

Challenge: Real-time document loading and response generation.

Solution: Cache FAISS vector store to optimize retrieval times and reduce latency.

Cloud Infrastructure

Cloud Provider: Google Cloud Platform, Azure or AWS

Core Services:

- **Compute**: VMs for model hosting and compute-intensive tasks.
- **Storage**: Persistent storage (Cloud Storage/Azure Blob) for document data.
- **Database**: FAISS vector store for optimized search.
- **APIs**: Google Generative AI API for language model responses.
- **Frontend Deployment**: Streamlit server (App Engine or Azure App Service).

Infrastructure Costs:

- **Compute Costs**: VMs with GPUs
- **Storage**: Persistent storage for document files
- **Generative AI API Costs**: Varies per usage

Go-To-Market Strategy

TARGET SECTORS:

Academic research, biotechnology firms, pharmaceutical R&D, medical journals, and healthcare providers.

PRICING MODEL:

- Freemium: Basic access for small research groups.
- Premium: Subscription-based for advanced analysis and increased document limits.

MARKETING CHANNELS:

- Partnerships with academic institutions.
- Conferences, webinars, and medical research forums.
- Targeted outreach to biotechnology and pharma R&D departments.

MARKETING CHANNELS:

- Partnerships with academic institutions.
- Conferences, webinars, and medical research forums.
- Targeted outreach to biotechnology and pharma R&D departments.

Future Scope

- **Enhanced Entity Recognition:** Expand to cover more medical entities, like biomarkers, clinical trial outcomes, and disease-specific indicators.
- **Knowledge Base Integration**: Integrate with external medical knowledge bases (e.g., PubMed, clinicaltrials.gov) to enhance context and accuracy.
- **Advanced Privacy Features:** Implement AI-powered de-identification to improve privacy protection in medical and clinical documents.
- **Multi-lingual Support:** Extend capabilities to analyze research documents in multiple languages, targeting global medical research communities.
- **Real-Time Data Updates:** Develop a continuous updating feature for real-time analysis of newly published research papers and medical reports.
- **Contextual User Feedback Loop**: Incorporate user feedback for continuous improvement in model accuracy and relevance, refining responses to meet research-specific needs.

THANK YOU