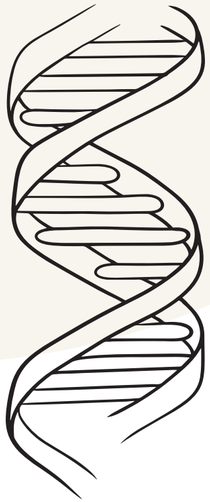
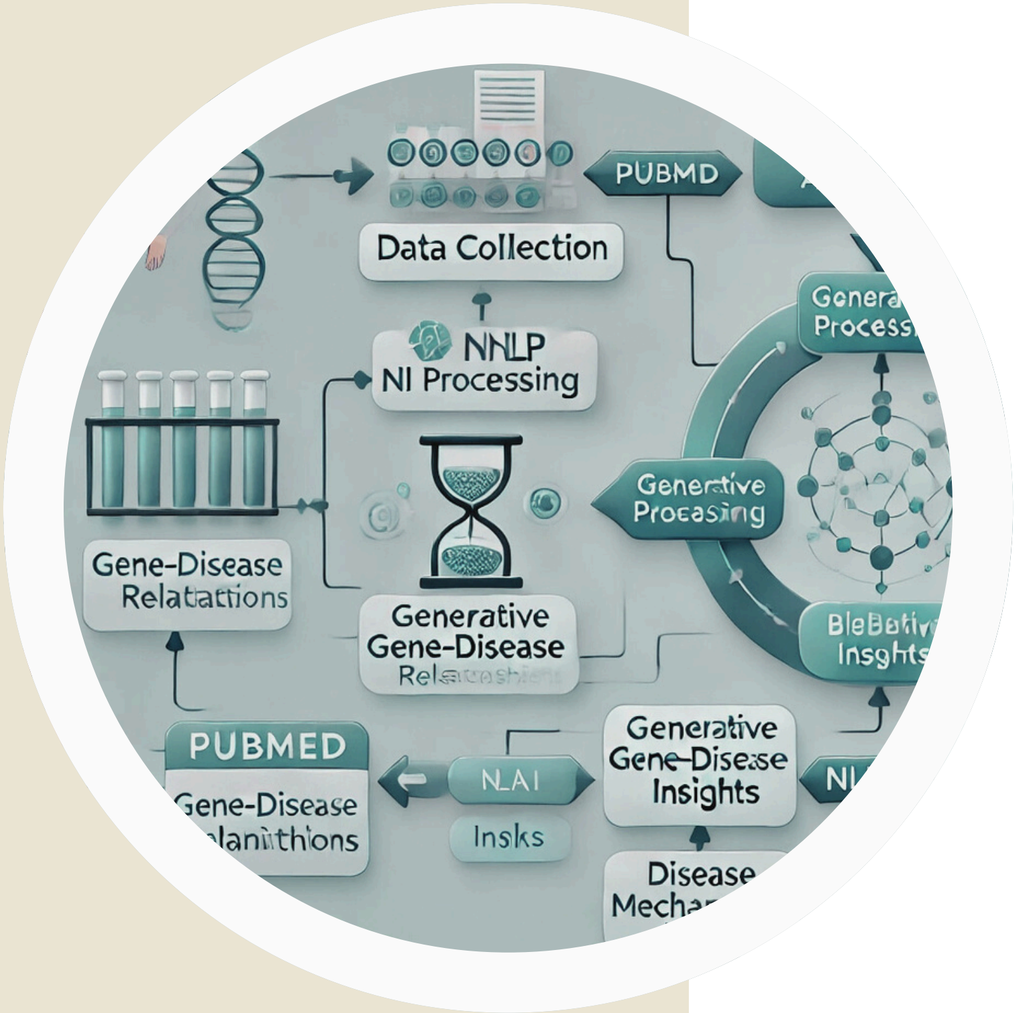


GenomicGPT FRAMEWORK

Presented by:
Kunal Malhan
Praveena Jothsna Pendyala
Keerthana Goka

Team Pumpkin Seeds

Harvard Rare Disease Hackathon 2025



TRACK

GENOMIC DIAGNOSTICS

B. Framework to determine the Mechanism of Disease for Gene-Disease relationships

CHALLENGES IN FINDING EVIDENCE OF GENE-RARE DISEASE RELATIONSHIPS

SCATTERED RESEARCH PAPERS

Research and review papers related to specific gene-disease pairs are spread across multiple platforms. Clinicians struggle to find all relevant information quickly when diagnosing a patient.

DISPERSED GENETIC DATA

Genetic evidence datasets are scattered across different sources, making it difficult for clinicians to access comprehensive data at the time of diagnosis.

SOLUTIONS

PROPOSED

LangChain

01

AI-Driven Literature Extraction, Summarization & Scoring

Leveraging Agentic AI, LangChain, and RAG systems to gather relevant research from multiple sources, generate summaries, and assign a confidence score to gene-disease relationships. This helps clinicians access the latest research in one place, especially for rare diseases.

02

Enhancing Confidence Scores with ML Models

Beyond literature-based scoring, ML models analyze genetic datasets to refine confidence scores, providing stronger evidence for gene-disease links.

PROPOSED FRAMEWORK

GENETIC EVIDENCE

Gene

Disease

PubMed
GeneReview

Exploring existing Research/Review Articles

Agentic AI
Langchain

Highly Relevant

Extract top 5 relevant review articles

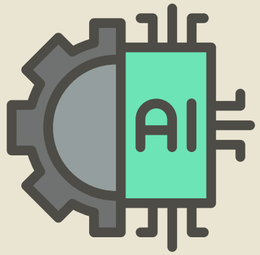
ChromaDB
GPT4

Retrieval Augmented Generation system to get required evidence

The G2P Project
ClinGen Evidence Repository

- Variant
- Mutation
- Mode of Inheritance
- Summary of Interpretation

Machine Learning



Predicts the Evidence and calculates Scoring Matrix

RESEARCH EVIDENCE

TARGET AUDIENCE

Geneticists & Rare Disease Centers

Easily access the latest research on rare diseases in one place, aiding in faster and more informed diagnosis.

Researchers Investigating Disease Mechanisms

Utilize AI-generated literature confidence scores as a starting point for deeper evidence-based research.

Families of Rare Disease Patients

Find authentic, technical literature in a centralized platform to better understand the condition and available research.

CHALLENGES AND FUTURE SCOPE

- Fine-tuning models with domain-specific datasets (e.g., ClinVar, Orphanet).
- Integrating real-time research updates using AI-powered search agents for the general public to access.

THANK YOU

FOR THE OPPORTUNITY