

PROJECT TITLE : PRISM Gemonics : Decentralized AI-Powered Genomic Data Ownership & Risk Intelligence Platform

PROJECT OVERVIEW

Team Name: Mangzing

Team Members:

1. Khumbongmayum Yaiphaba Singh – Blockchain Developer & UI/UX Designer
2. Chingkheinganba Thoudam – AI/ML Engineer
3. Thongam Gripson Singh – Backend and Frontend

Link the Concept Video: <https://youtu.be/a8waELYlxZ0>

THE PROBLEM LANDSCAPE

Problem Statement: Genomic data is one of the most sensitive categories of personal data. However, current healthcare systems:

- Store genetic reports in centralized databases
- Offer limited patient control over access
- Are vulnerable to breaches and data tampering
- Lack transparent audit mechanisms

Patients do not truly “own” their DNA data.

The core pain point:

No secure, patient-controlled, tamper-proof system exists for genomic data sharing with AI-driven preventive insights.

Target Audience

- Individuals who have undergone genetic testing
- Hospitals & diagnostic laboratories
- Precision medicine researchers
- Healthcare startups

“Why Now”

- Rise of precision medicine
- Increased healthcare data breaches globally
- Growing adoption of consumer genomics
- Expansion of blockchain-based digital identity systems

Current systems cannot ensure:

- Data sovereignty
- Tamper-proof storage
- Transparent access control

PROPOSED SOLUTION & USP

Solution Overview

PRISM Gemonics is a decentralized genomic intelligence platform where:

1. Files are AES-256 encrypted & BLAKE3 hashing ensures integrity
2. Encrypted files are stored on IPFS & Hash fingerprints are stored on blockchain
3. AI predicts disease and risk using Polygenic Risk Scores
4. Doctors request access via smart contracts
5. Patients approve/reject access
6. All actions are logged immutably

Unique Selling Proposition (USP)

- True patient data ownership
- Modern cryptographic integrity
- Zero raw genome stored on-chain
- AI-powered disease Prediction and Risk
- Smart contract-based access governance
- Immutable audit trail

Unlike traditional EHR systems, PRISM Gemonics combines predictive intelligence + decentralized security + patient sovereignty.

Core Logic

The system integrates:

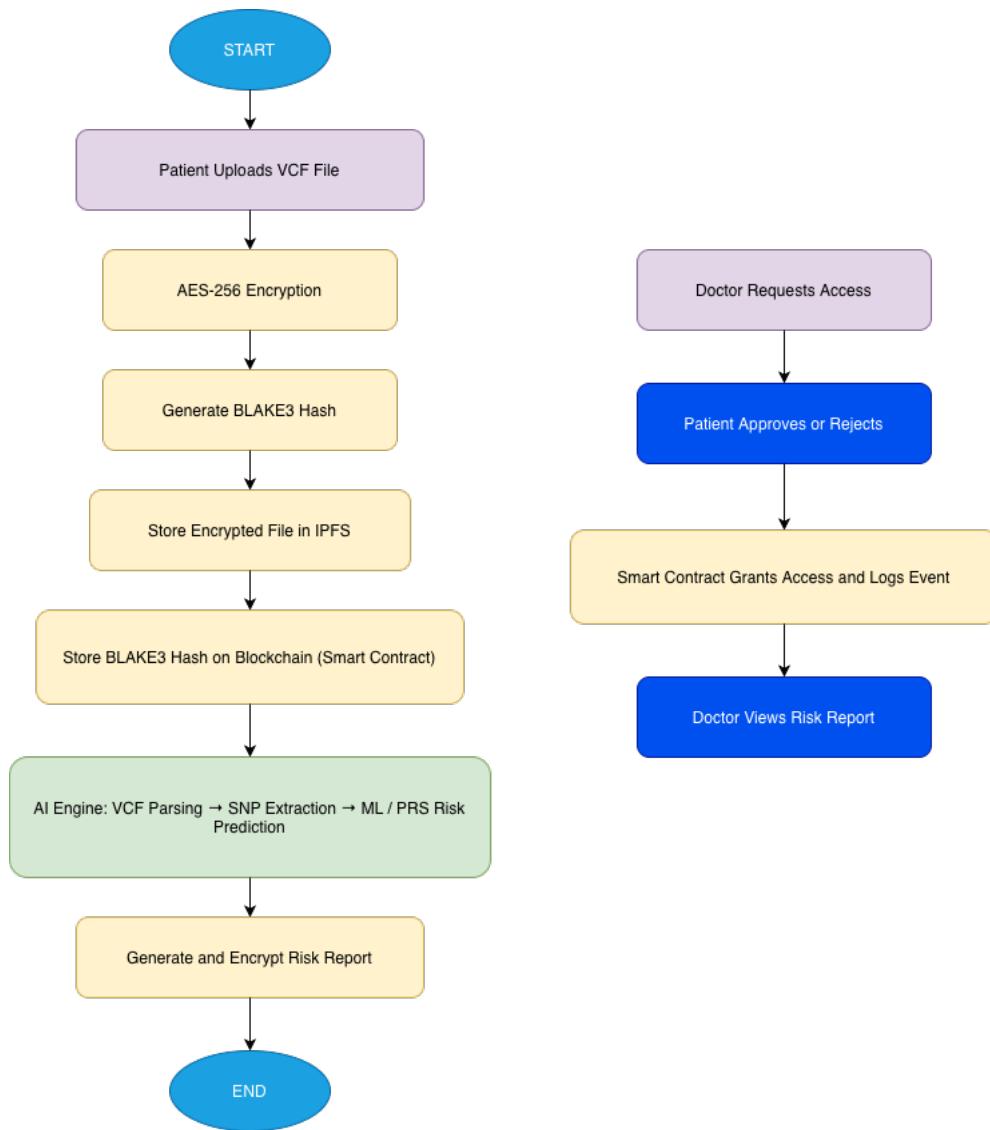
- AI models for disease risk prediction
- AES-256 symmetric encryption for file protection & BLAKE3 hashing for fast, tamper-proof integrity & IPFS for decentralized storage
- Solidity smart contracts for access control

Biological Logic:

- Extract disease-associated SNPs from VCF
- Encode genotypes (0,1,2)
- Apply weighted risk scoring
- Generate probability output

TECHNICAL ARCHITECTURE & STACK

System Workflow



Tech Stack

Frontend : 1. Next.js
2. MetaMask Wallet Integration

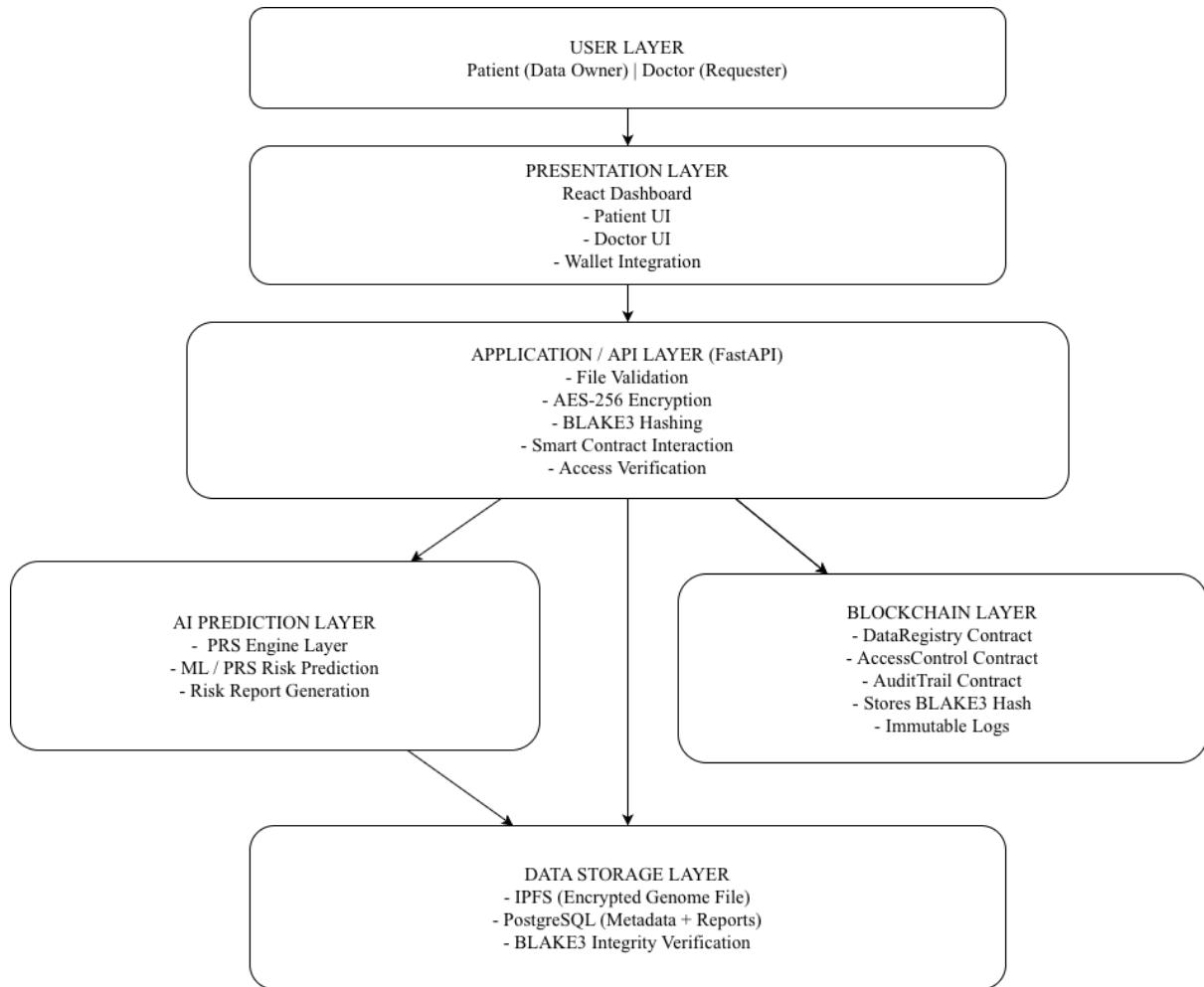
Backend / Database: 1. FastAPI
2. PostgreSQL

AI / Specialized Tools: Python, gzip, XGBoost, scipy, scikit-learn, pandas, numpy

Blockchain: Solidity, Hardhat, Ethereum-compatible network

Storage & Security: IPFS, AES-256 Encryption, BLAKE3 Hashing

Architecture Visualization



KEY FEATURES & FUNCTIONALITIES

Feature 1 (Primary): AI-based genomic disease risk prediction.

Feature 2 (UX):
1. Wallet-based authentication
2. Patient dashboard with risk visualization

Feature 3 (Reliability & Security):
1. AES-256 encrypted genomic storage & BLAKE3
2. Immutable blockchain audit logs
3. Temporary permission-based decryption keys

IMPLEMENTATION ROADMAP

Phase 1:

Focus: 1. Problem validation
2. Architectural design
3. Smart contract drafting
4. AI prototype (basic model)

Deliverables: 1. Concept video
2. Documentation
3. Initial Prototype

Phase 2:

Focus: 1. Full encryption integration
2. BLAKE3 hashing implementation & IPFS deployment
3. AI Model Refining & UI/UX refinement
4. Final demo video

Deliverables: 1. Working prototype
2. 2-minute demo video

IMPACT & SUSTAINABILITY

Social & Economic Impact

- Empowers individuals with genomic ownership & Promotes preventive healthcare
- Reduces fraud and tampering & Encourages transparent research collaboration

Scalability: The architecture can expand to:

- Electronic Health Records (EHR)
- Medical imaging storage
- Clinical trial management
- Insurance claim validation
- Biomedical research datasets

Designed for distributed large-scale genomic storage.

Risk & Mitigation

Major Risk: AI prediction bias due to limited genomic datasets.

Mitigation:

- Use diverse genomic datasets ()
- Regular model retraining
- Transparency in risk score explanation
- Human-in-the-loop clinical review

Genomics & Genetic Databases & AI/ ML Technology Documentation

1. NCBI SNP Database (dbSNP) – <https://www.ncbi.nlm.nih.gov/snp/>
2. GWAS Catalog – <https://www.ebi.ac.uk/gwas/>
3. 1000 Genomes Project – <https://www.internationalgenome.org/>
4. XGBoost - <https://arxiv.org/abs/1603.02754>
5. VCF Processing - <https://academic.oup.com/bioinformatics/article/27/15/2156/402296>
6. Python Software Foundation - <https://www.python.org/>

Blockchain & Cryptography Documentation

5. Ethereum Solidity Documentation – <https://docs.soliditylang.org/>
6. IPFS Documentation – <https://docs.ipfs.io/>
7. BLAKE3 Official Specification – <https://github.com/BLAKE3/BLAKE3>

Research Papers on Polygenic Risk Scores

8. Polygenic Risk Scores: Genomes to Risk Prediction – <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10681370/>
9. Genome-wide association studies, Polygenic Risk Scores and Mendelian Randomisation – <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC12013552/>
10. Polygenic Risk Score Knowledge Base (PRSKB) – <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9438378/>