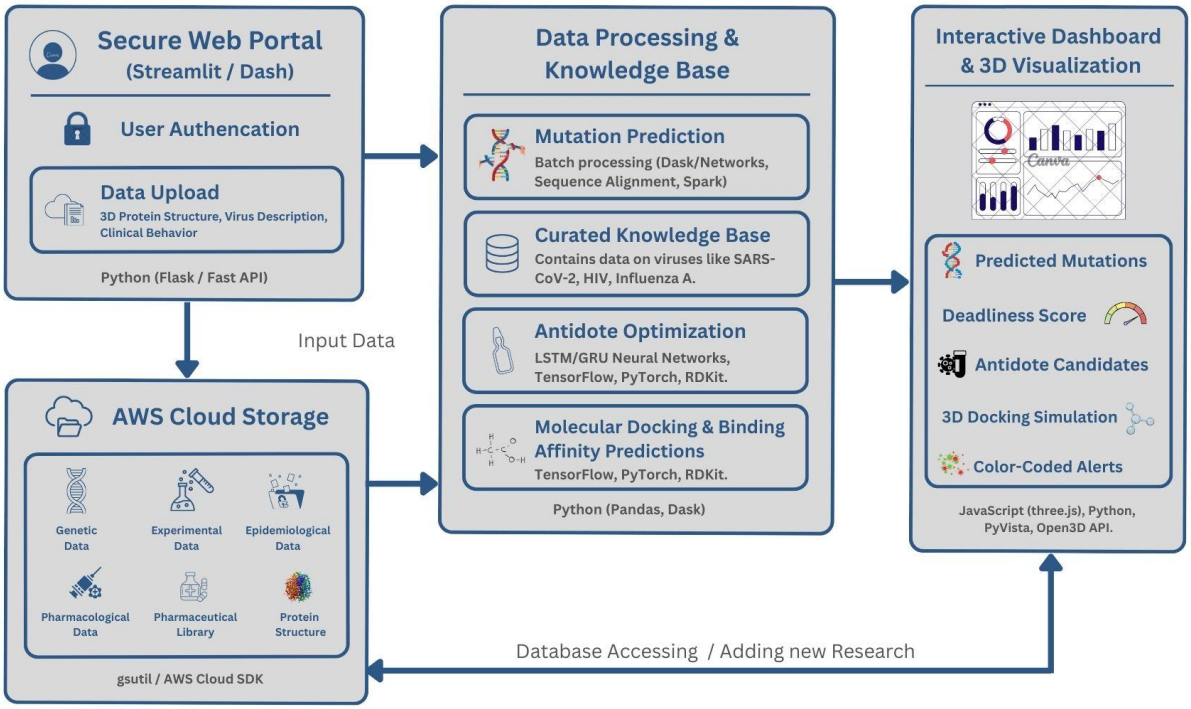
System Architecture



1. Secure Web Portal (User Entry Point)

- Users (researchers, doctors, pharma partners) log into a secure web interface.
- User Authentication ensures that only authorized users access sensitive data.
- Through this portal, users can upload input data, such as:
  - 3D protein structures,
  - virus descriptions,
  - clinical behavior datasets.
- The portal is built on Streamlit / Dash for interface + Python Flask / FastAPI for backend connectivity.

2. AWS Cloud Storage (Central Repository)

- Uploaded data flows into cloud storage (AWS S3), where it is stored securely.
- Types of data stored:
  - Genomic sequences,
  - Experimental results,
  - Epidemiological data,
  - Pharmacological datasets,

    - Pharmaceutical libraries,
    - Protein structures.
- This acts as the central data lake of the system, where raw and processed data are preserved.

3. Data Processing & Knowledge Base (Core Engine)

- Once data is in the system, it is processed in the AI-driven pipeline:
    - Mutation Prediction
    - Uses deep neural networks, sequence alignment, and batch processing.
    - Predicts possible viral mutations and their biological impact.
    - Curated Knowledge Base
    - Stores and organizes all known virus-related information (SARS-CoV-2, HIV, Influenza, etc.).
- Provides reference material for comparisons with new data.
- Antidote Optimization
- Runs AI/ML models (LSTM/GRU neural networks, PyTorch, TensorFlow).
- Suggests optimized drug candidates or modifications for higher effectiveness.
- Molecular Docking & Binding Affinity Predictions
- Simulates how potential drugs bind with viral proteins.
- Uses cheminformatics tools (RDKit, docking simulations) to calculate binding strength and stability.

4. Interactive Dashboard & 3D Visualization (Output Layer)

- After processing, results are presented to the user in a visual and interactive format:
    - Predicted Mutations: Displays possible future viral mutations.
    - Deadliness Score: Quantifies the severity and risk of a mutation.
    - Antidote Candidates: Lists optimized drugs or molecules predicted by the system.
    - 3D Docking Simulation: Shows how drugs interact with viral proteins in real-time 3D.
    - Color-Coded Alerts: Highlights high-risk mutations or critical findings for easy interpretation.
- This visualization layer is powered by JavaScript libraries (Three.js, PyVista, Open3D API), ensuring rich 3D molecular interaction graphics.

5. Feedback Loop (Continuous Learning)

- New research data and validated results are continuously added back into the AWS Cloud Storage and Knowledge Base.
- This feedback loop ensures the system evolves with new viral genomes, clinical findings, and drug trial outcomes, making the predictions more accurate over time.