## **Molecule Research and Discovery Architecture Augnik Banerjee (RA2211028010006)**

## **Aniket Singh (RA2211028010014)**

## **1. Introduction**

Molecular visualization and generative modeling have become essential in computational chemistry and drug discovery. However, existing platforms lack seamless collaboration, real-time interaction, and AI-driven molecule generation. This project aims to bridge these gaps by building an advanced molecular visualization and collaboration platform, integrating NVIDIA's molecule generation model with real-time 2D/3D visualization, collaborative messaging, and custom molecular structure generation.

## **2. Methodology: Hybrid Research & Implementation**

### **2.1 Rationale for Methodology**

### Since our project is both research-based and application-driven, we employ a Hybrid Research & Implementation Methodology to:

### Analyze existing platforms and identify research gaps.

### Develop a scalable, interactive molecular visualization system.

### Implement AI-based molecule generation and collaborative features.

### Validate usability through performance testing and user feedback.

### **2.2 Stages of Methodology**

**Research & Literature Review** 1 Week

Examine prevailing frameworks & areas of research lacunaeProblem scope defined & mitigation strategy

**Architecture Design & Planning**1 Week

Identify architecture, security levels & cloud service providers System design & workflow charts.

**Prototype Development & Testing** 2-3 Weeks

DDoS detection & mitigation implementationFunctional prototype in a test cloud environment

**Documentation & Evaluation** 1-2 Weeks

Performance analysis, report writingFinal report, presentation & results

## **3. Architecture: Event-Driven Microservices Model**

### **3.1 Justification for Architecture**

Our system follows a modular architecture with real-time collaboration and AI-driven features:

* Scalability: Efficient handling of molecular data and AI computations.
* Real-Time Interaction: Live molecule visualization and chat-based collaboration.
* Cloud-Native Deployment: Deployment on AWS/GCP with containerization.
* Extensibility: Support for new AI models and additional visualization features.
* Cloud-Native Deployment: Facilitates deployment across hybrid cloud environments (AWS, Azure, GCP) with containerization.

### **3.2 Detailed Architecture Components**

We have multiple integrated components in our architecture:

1. **Molecule Visualization Engine**
   * **Function: 2D/3D molecular structure rendering.**
   * **Tools: RDKit.js, Three.js, React ApexCharts.**
2. **AI-Powered Molecule Generator**
   * **Function: Generate new molecular structures using NVIDIA’s molecule generation model.**
   * **Tools: NVIDIA AI Platform, PubChem API, RDKit.**
3. **Real-Time Collaboration Module**
   * **Function: Group messaging and discussion features.**
   * **Tools: Ably, NextAuth.js, Resend.**
4. **Data Storage and Retrieval**
   * **Function: Store molecular data and AI-generated structures.**
   * **Tools: MongoDB, Mongoose.**

## **4. Filling Research Gaps with Our Solution**

**Identified Gaps & Proposed Solutions**

1. Limited Interactive Molecular Visualization
   * Gap: Existing visualization tools lack interactive editing features.
   * Solution: Implement a real-time editable 2D/3D molecular workspace.
2. AI-Driven Molecule Generation Lacks Explainability
   * Gap: AI-generated molecules lack clear justification for their structure.
   * Solution: Implement attribution visualization for AI-generated molecules.
3. Inefficient Real-Time Collaboration in Molecular Research
   * Gap: Most platforms lack built-in messaging and versioning.
   * Solution: Integrate group messaging and version tracking for molecule edits.

## **5. Conclusion**

This project presents a novel molecular visualization and collaboration platform, combining real-time interaction, AI-driven molecule generation, and cloud scalability. By addressing key research gaps, our solution enhances the efficiency and accessibility of computational chemistry tools. The modular architecture ensures flexibility for future enhancements, supporting further advancements in molecular research and drug discovery.

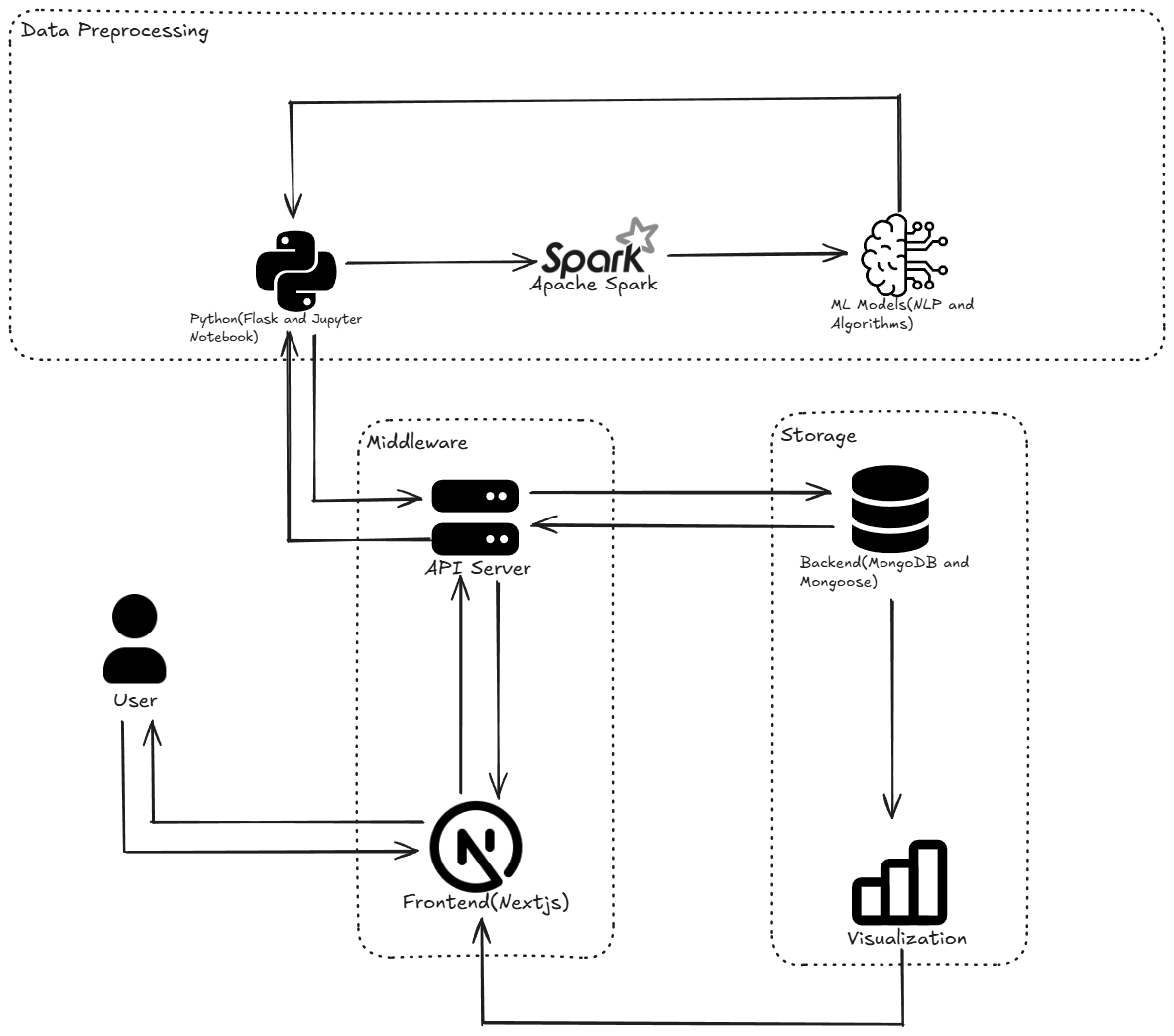


Fig 1. System Architecture

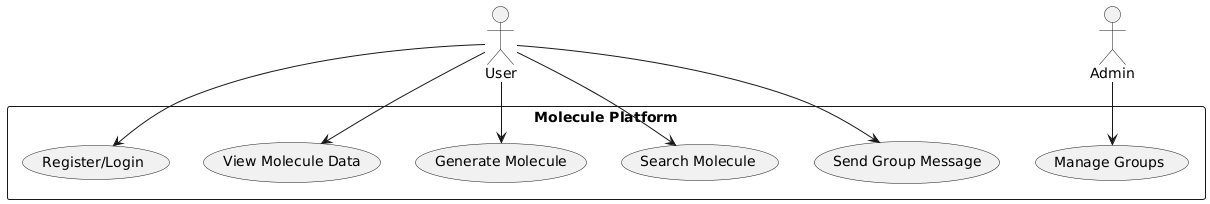


Fig 2. Use Case

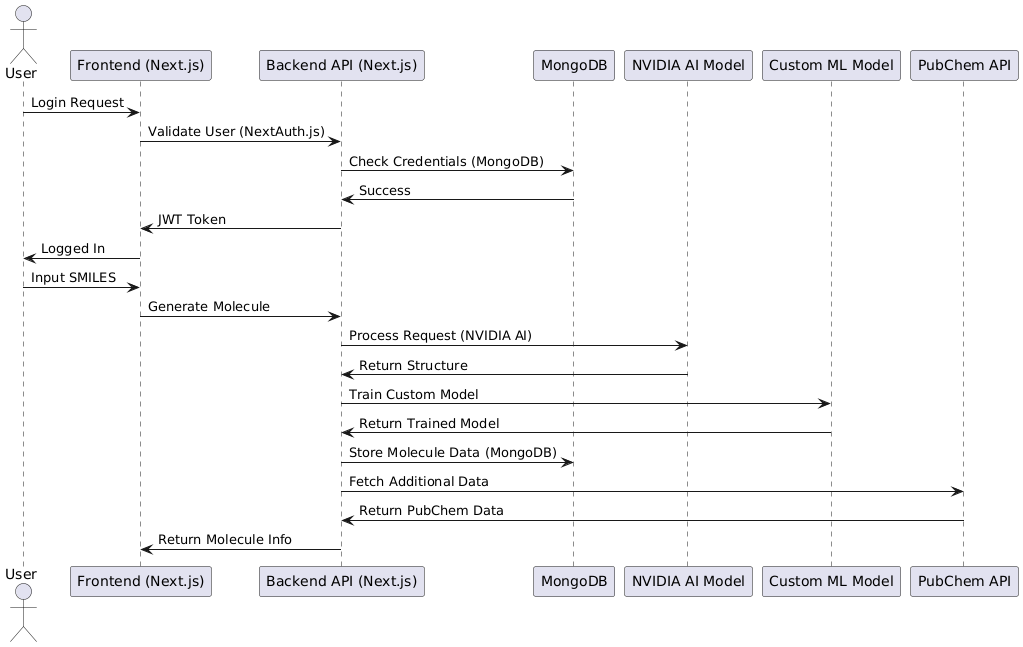


Fig 3. Sequence

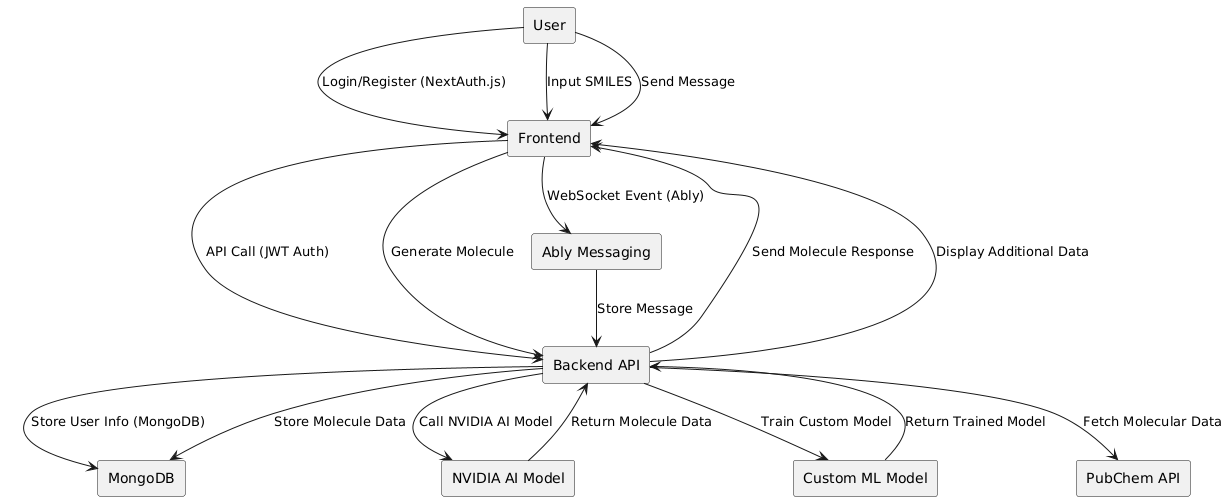


Fig 4. DFD Diagram

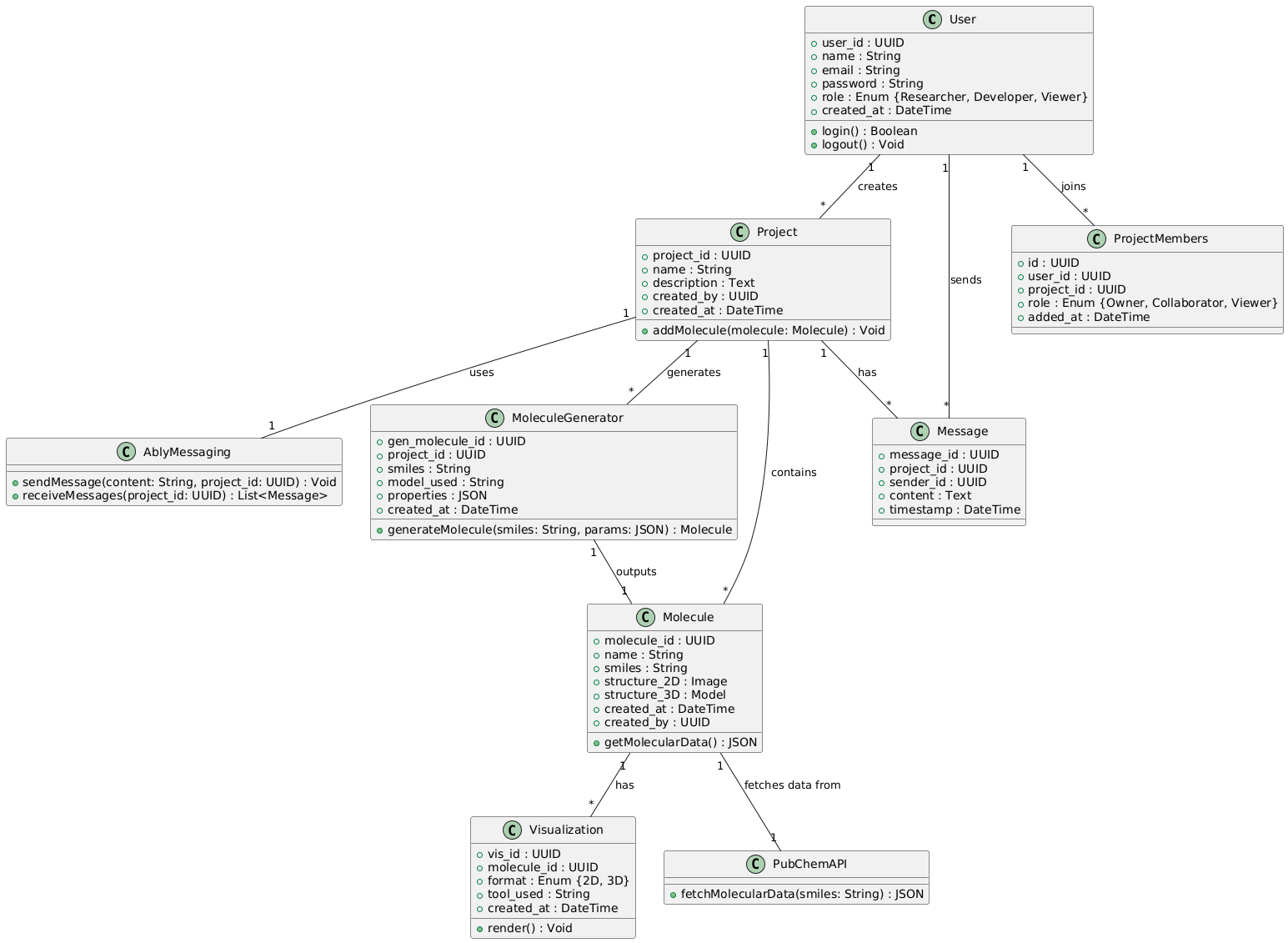


Fig 5. Class diagram

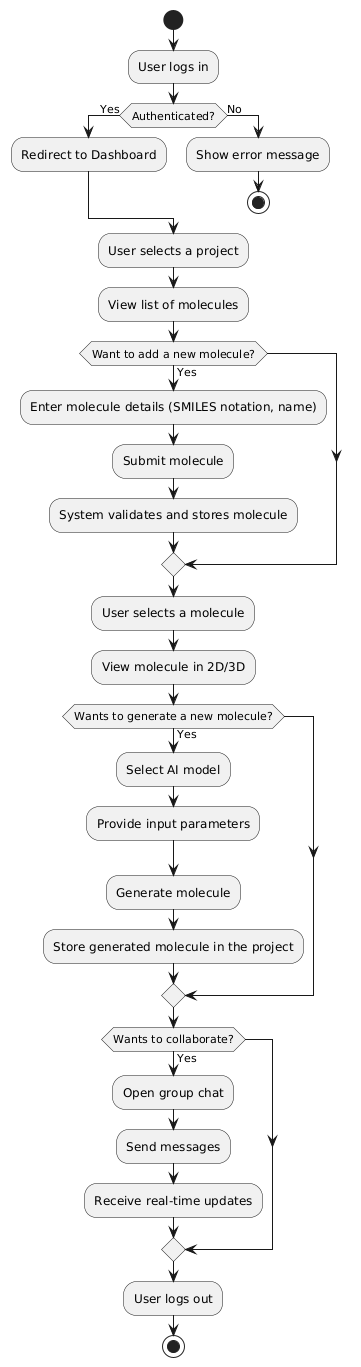


Fig 6. Activity

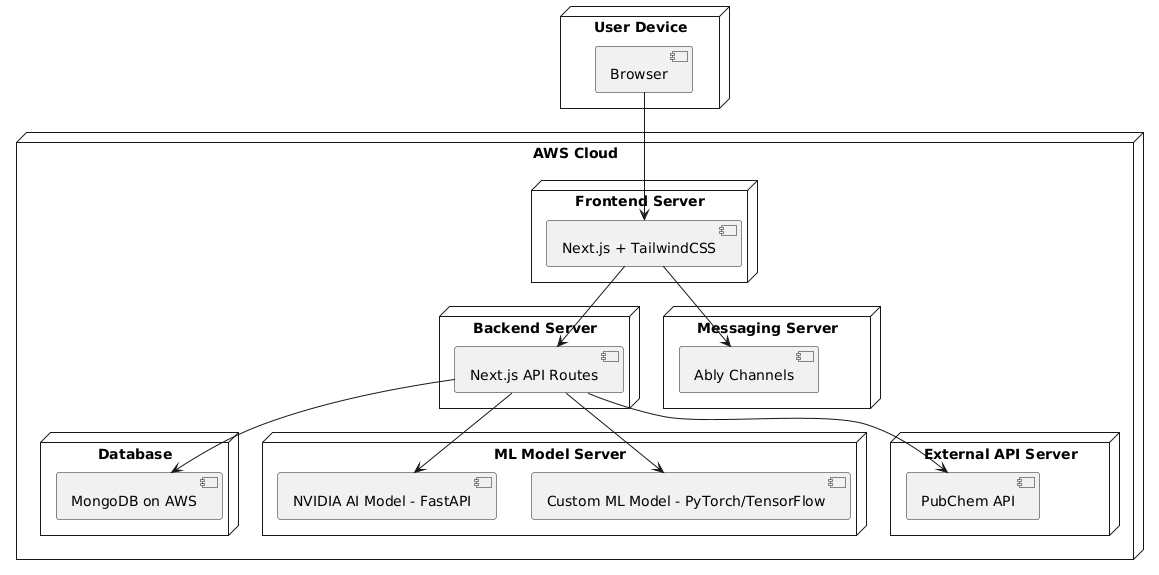


Fig 7. Deployment

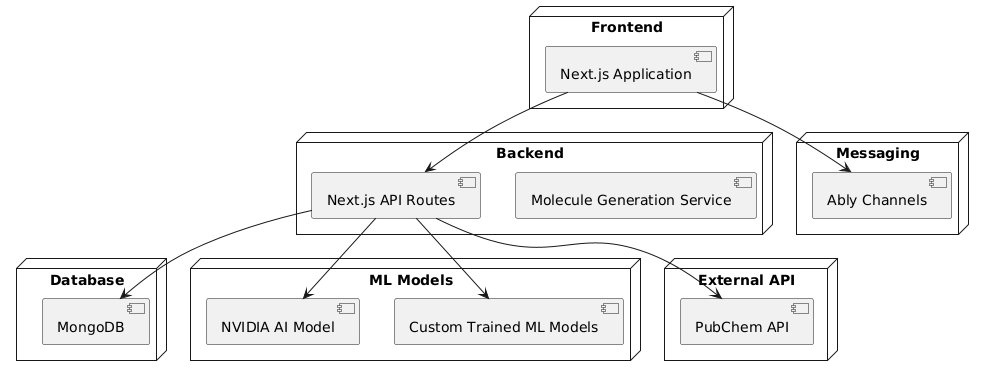


Fig 8. Component

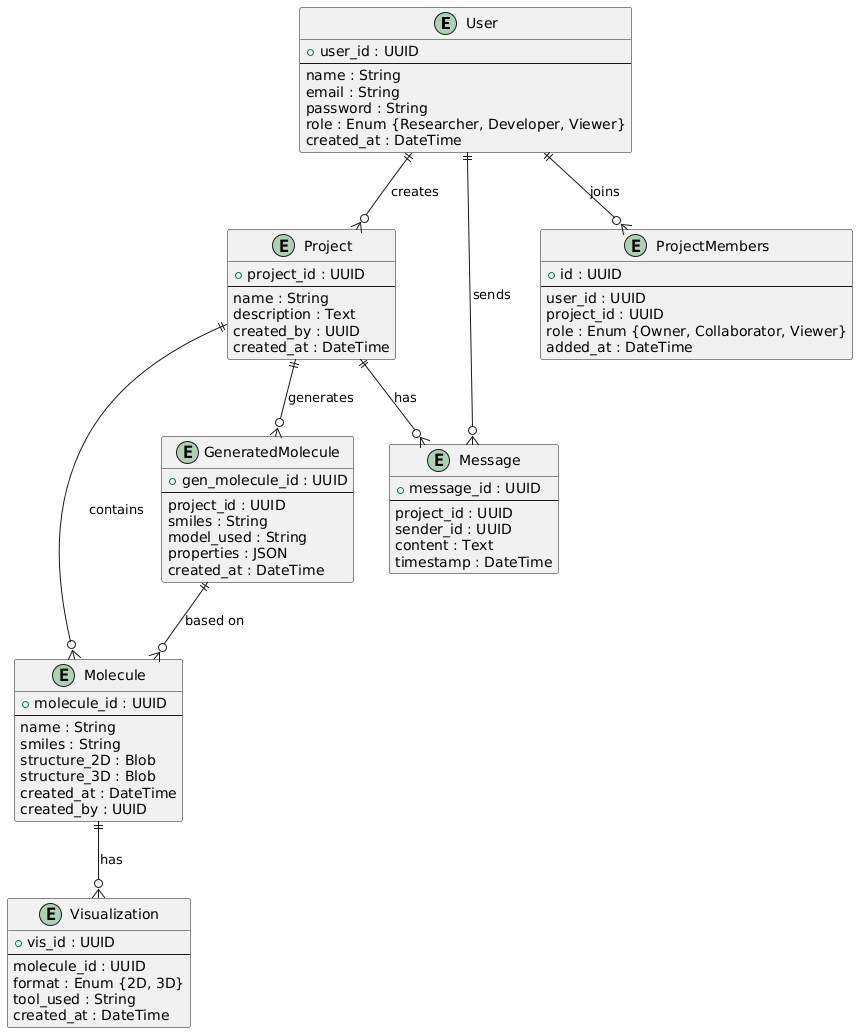


Fig 9. Entity relationship