Avatar Lab:

AI powered Text-to-Avatar Platform

G331
By 23BD1A05BB
HRITHIK RAJ

Tech Stack Choices & Reasoning

Next.js

Chosen for server-side rendering, replacing MERN stack React + Express.

PostgreSQL

Efficient video storage and reliable relational data management.

MongoDB

Flexible document storage for unstructured data.

Prisma

Type-safe database ORM for seamless data access.

ngrok

Enables local inference by exposing local servers securely.

RobustVideoMatting

Advanced background removal for realistic video output.

Key Learnings & Project Hurdles

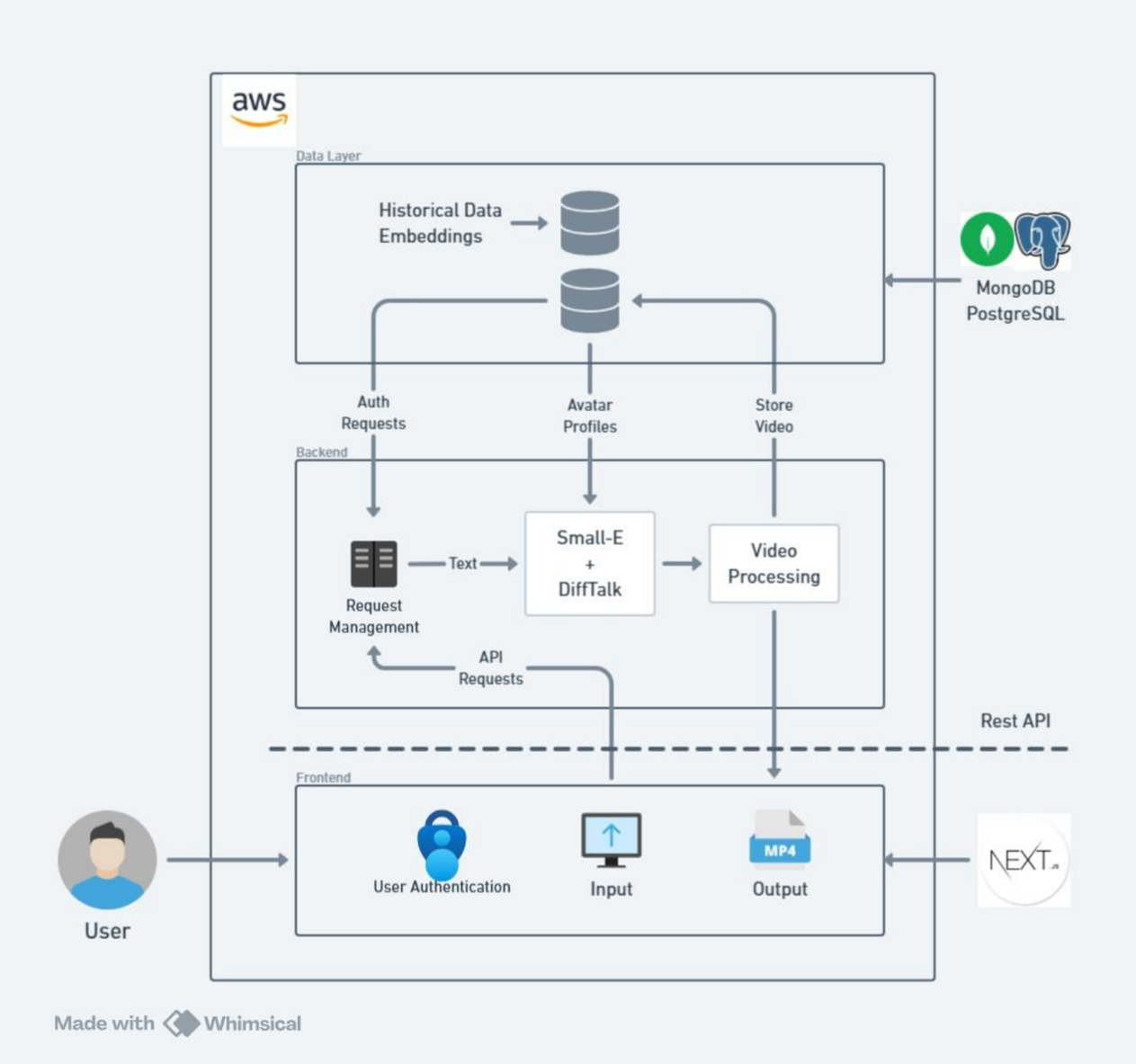
Key Learnings

- Mastered Next.js SSR and API integration
- Database management with Prisma, PostgreSQL, MongoDB
- Using ngrok for secure local inference
- Applied RobustVideoMatting for video processing

Project Hurdles

- Deployment challenges and resource limits
- Steep learning curve for new tech
- Knowledge gaps in cloud deployment
- Time constraints affected full deployment

Architecture Changes & Milestone Achievements



Architecture Changes

- Replaced DiffTalk with DiffDub model
- Integrated ngrok for local server access
- Added RobustVideoMatting for background removal
- Implemented Prisma ORM and customized API

Milestone Status

- **Core functionality:** Achieved successfully
- **Deployment:** Not fully completed

Model Selection & Reasoning

Model	Used	Reason	Alternatives	Notes
SMALL-E	Yes	Efficient, diffusion-based learning	VALL-E	GPU compatible, lightweight
DiffDub	Yes	High-quality talking head synthesis	SyncTalk, SadTalker	Better realism and control
RobustVideoMatting	Yes	Effective background removal	None	Essential for video realism

Summary & Next Steps

Project Successes

Core AI pipeline built with efficient tech stack

Challenges

Deployment and resource constraints remain

Next Steps

Focus on deployment, optimize models, enhance demo