AVATAR LAB

Group: G331 | Presenter: Akula.Yashwanth

Tech Stack & Reasoning

Frontend Framework

Next.js: Built-in server-side rendering enables top performance and SEO excellence.

Database Solutions

PostgreSQL: Robust asset storage with reliable relational integrity.

MongoDB: Flexible document storage for unstructured data.

API & ORM

Prisma: Type-safe ORM ensures efficient, error-free database access.

Queue & Cache

Redis: High-speed in-memory caching to reduce latency.

BullMQ: Robust async job queue for handling concurrent processing.

Deep Learning Models

Small-E: Low latency text-to-speech synthesis.

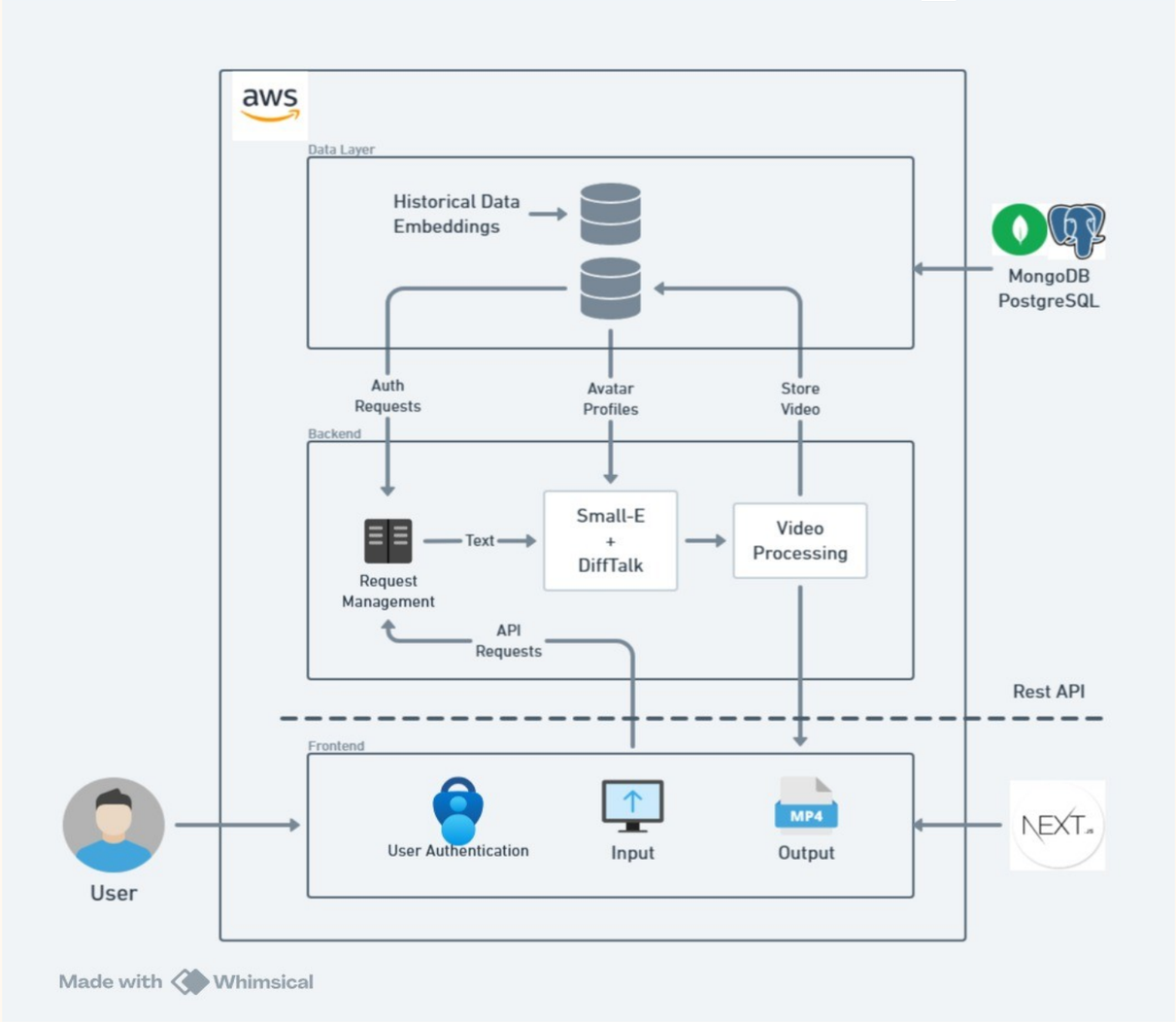
DiffDub: Realistic lip-sync and talking head generation.

Development Tools

ngrok: Secure local server exposure for model testing.

RobustVideoMatting: Professional background segmentation for realism.

Architecture Changes & Milestone Status

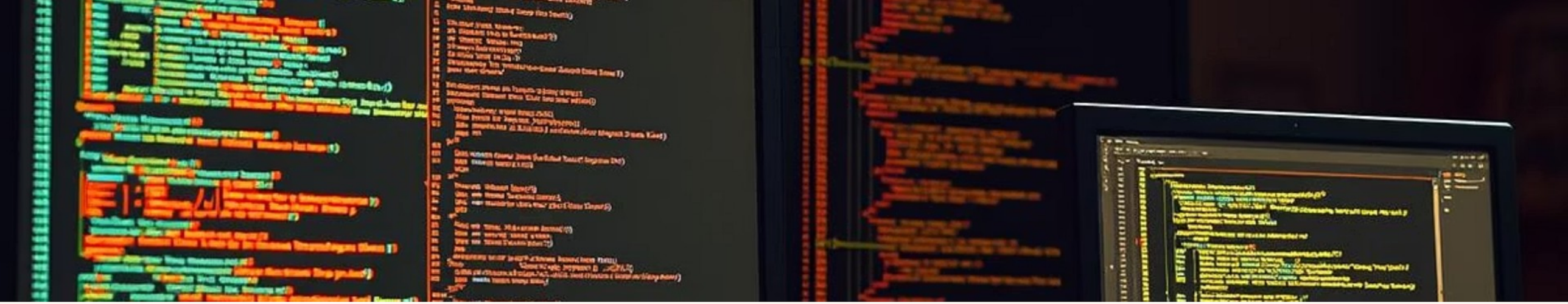


Architecture Changes

- Replaced DiffTalk with DiffDub model
- Enabled ngrok for local access
- Added RobustVideoMatting for background removal
- Adopted Prisma ORM and customized API

Milestone Status

- Core functionality achieved successfully
- Deployment not fully completed yet



Challenges & Learnings

Key Challenges

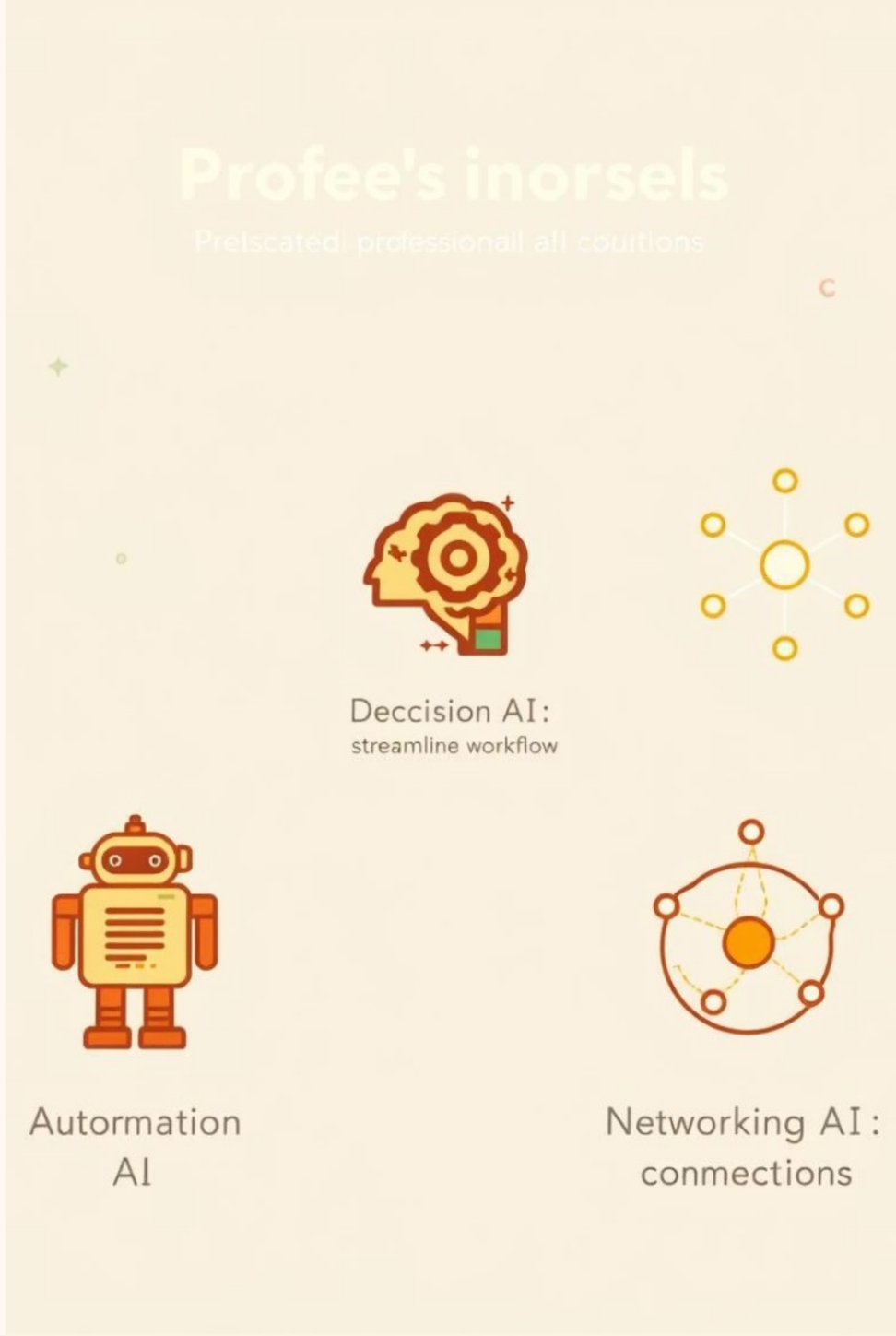
- Slow video processing impacted UX
- Scaling from local ngrok to production
- Resource allocation for concurrent GPU usage
- Efficient background task handling
- Providing real-time processing feedback
- Managing BullMQ queue without overload

Technical Learnings

- Advanced Next.js SSR and API integration skills
- Hybrid DB management with Prisma, PostgreSQL, MongoDB
- Secure local dev using ngrok for model inference
- Robust background removal with video matting

Model Selection & Reasoning

Model	Purpose	Selected Because	Alternatives	Key Notes
SMALL-E	Audio processing	Cost-effective, efficient	VO-VAE 2 Vall-E	GPU compatible, lightweight
DiffDub	Talking head synthesis	High-quality, realistic output	Sadtalker, StyleGAN	Better realism and control
RobustVideoMatting	Background removal	Effective video enhancement	None	Essential for video realism



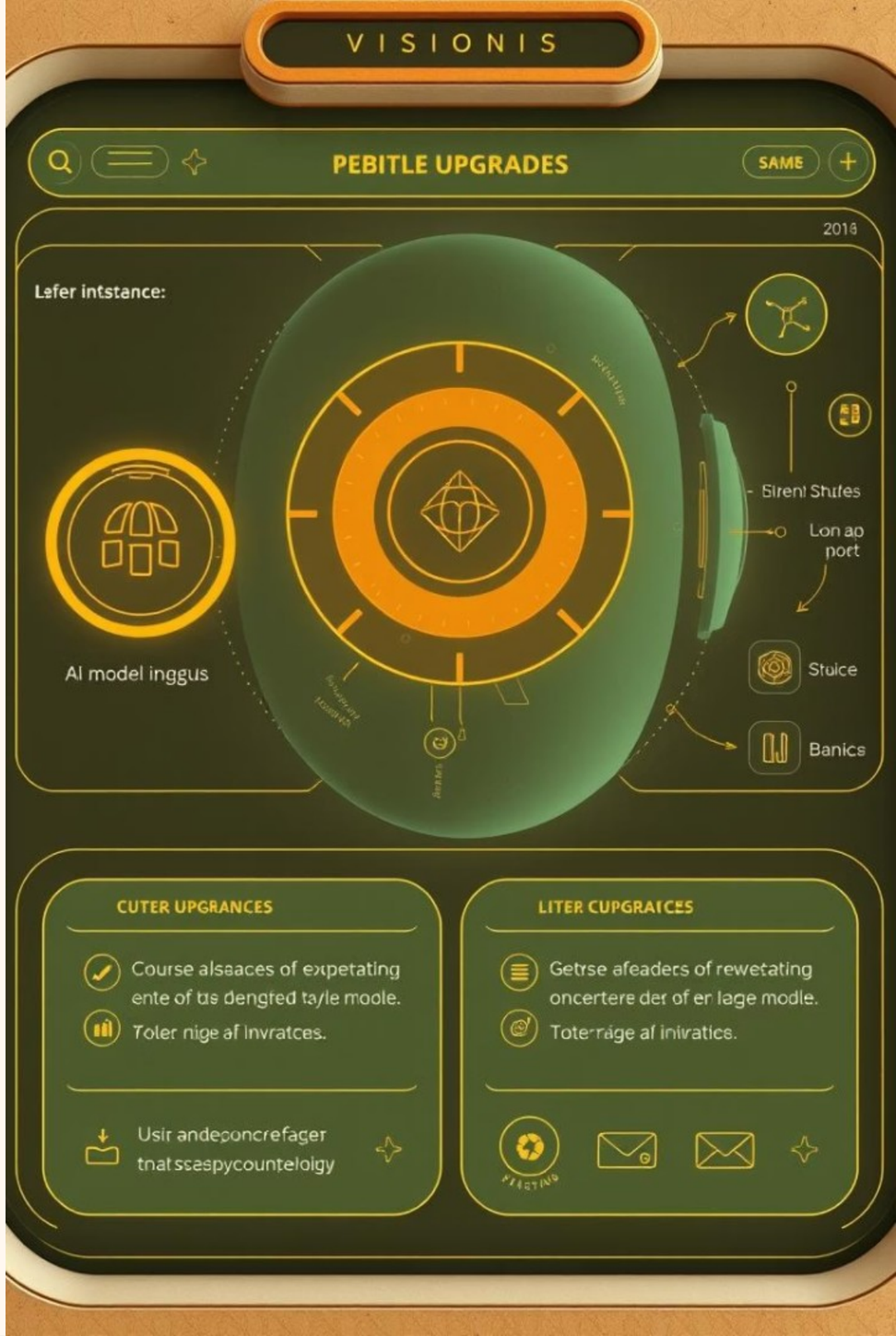
Future Plans - Model & UX Enhancements

Model Improvements

Implement multi-lingual support and higher video resolution.

User Experience

Collect user feedback to rapidly iterate and improve performance.



Future Plans - Hosting Strategy

Hosting Options Evaluation

- Vercel: Excellent frontend optimization, higher cost at scale
- Render: Flexible, predictable pricing suited for growth
- AWS: Highly scalable and customizable for complex needs

Next Steps

Focus on accessibility, quality, and user-driven improvements.

Choose hosting aligned to budget and scalability goals.

Summary & Next Steps

Summary

- Robust tech stack chosen for performance and scalability
- Key architecture changes improved functionality
- Challenges informed learnings on optimization and deployment
- Future plans focus on enhancements and hosting strategy

Next Steps

1. Complete deployment and optimize production environment
2. Implement planned model and UX improvements
3. Finalize hosting platform based on cost and scale
4. Gather continuous user feedback for iterative updates