AI Study Buddy — Technical Architecture Report

Date: 2025-10-08 23:22 UTC

# Executive Summary

Analysis of current architecture; proposed modular architecture; user-flow summaries; risks & NFRs.

# Current State

React (Vite) SPA; Node.js/Express API; Prisma ORM; Postgres; pgAdmin; Docker Compose.

# Issues

* Dual run modes cause drift
* No object storage for large uploads
* No queue/cache layer
* Observability gaps
* Vector search undecided

# Proposed Architecture

* Express API + WebSocket
* Redis for cache/rate limits/queue
* Object Storage (MinIO/S3)
* Postgres + pgvector
* Worker(s) for OCR/ASR/summarise/quiz
* Observability (OTel→Prom/Grafana)

# User Journeys

* Flow Onboarding Registration
* Flow Upload Processing
* Flow Quiz Generation
* Flow AI Tutor QA
* Flow Community Discussion