Designing AI Infrastructure Experiences

Deepesh Sudhakar

DigitalOcean Gradient™
AI Agentic Cloud

A high-level overview of multi-sprint AI platform work

Every initiative highlighted in this deck represents a multi-sprint design cycle requiring deep technical discovery, cross-functional alignment, and systems-level thinking. My role involved leading these projects from initial ambiguity to detailed product execution alongside engineering and product partners.

Product areas covered across the AI platform ecosystem

Shaping the DigitalOcean AI platform into a connected ecosystem designed to support the full developer lifecycle from model selection to production monitoring.

Core Area	Focus Initiatives
Discovery	Model Studio + Model Catalog
Experimentation	Multimodal Playground
Deployment	Serverless Inference Enhancements
Evaluation	Agent Evaluations + Human-in-the-loop Feedback
Data & RAG	Knowledge Base Activity Logs + Auto Indexing

Streamlining model discovery through a unified catalog

Problem

Developers evaluating AI models needed a trusted, centralized environment to discover and compare options before testing.

Design Action

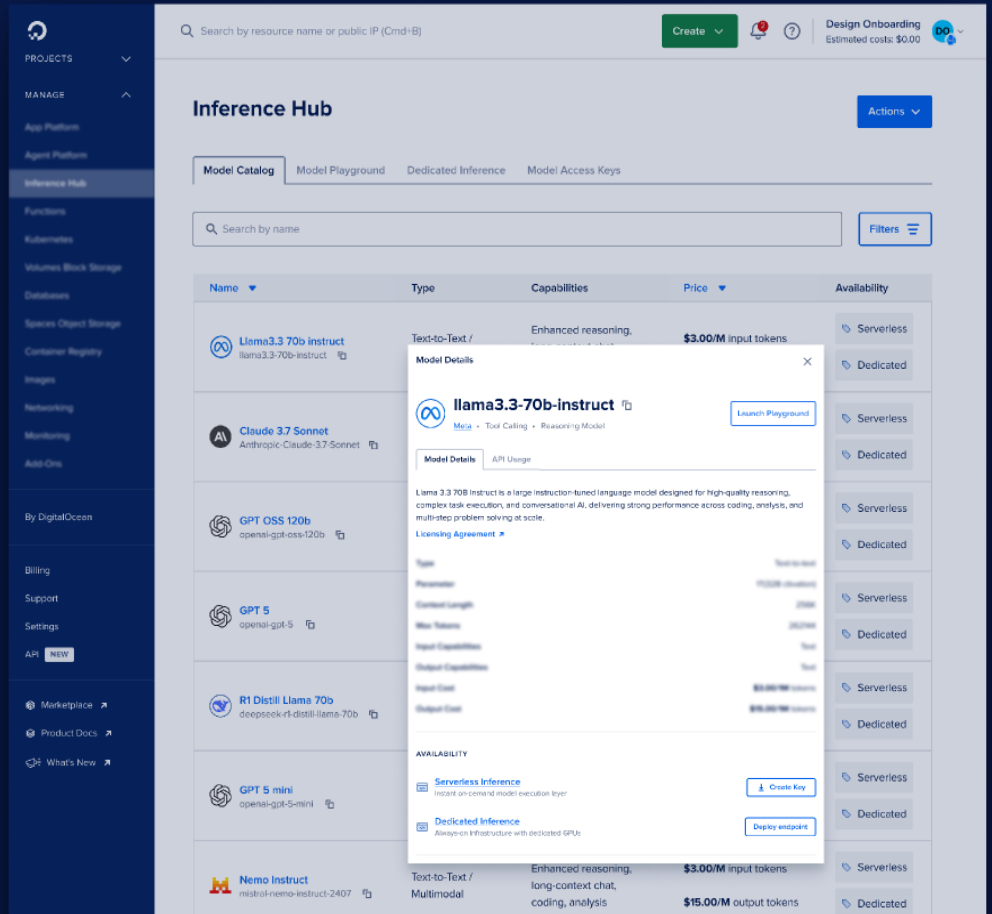
I unified the Model Studio concept with a catalog-led experience, surfacing benchmark signals, model details, and direct comparison entry points.

Why it Mattered

This approach created a "front door" for the AI experience, significantly lowering the friction for model adoption.

Outcome / Signal

This initiative established a clearer discover-to-evaluate path and reduced the ambiguity inherent in choosing between competing LLMs.



Lowering the barrier to experimentation with multimodal playgrounds

Problem

Developers struggled to understand how different models behaved across text, image, and audio tasks before committing to an inference path.

Design Action

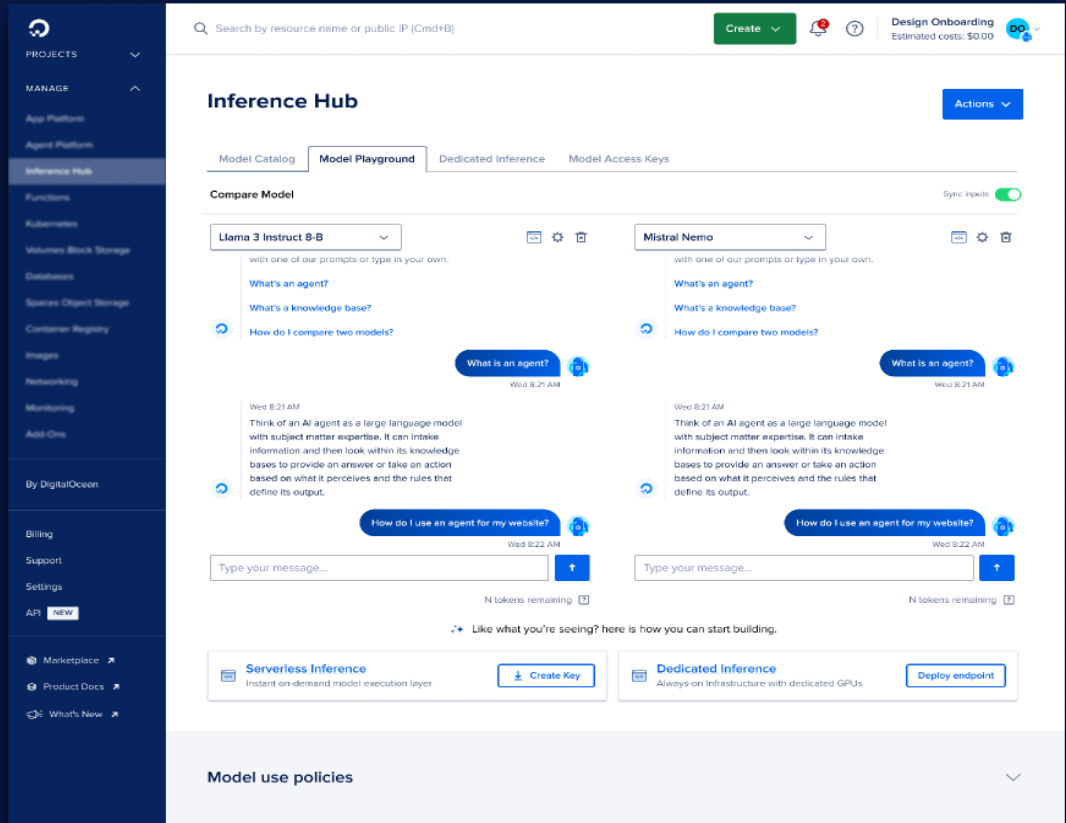
Designed a flexible playground for side-by-side exploration, featuring modular modality inputs and real-time output previews.

Why it Mattered

Experimentation is the bridge to production; a low-friction testing environment builds developer confidence in model reliability.

Outcome / Signal

Successfully created a more intuitive bridge from discovery to deployment decisions by making system behavior transparent.



Simplifying the path to production with serverless inference

Problem

Developers required a faster path to configure and deploy inference without being bogged down by infrastructure complexity.

Design Action

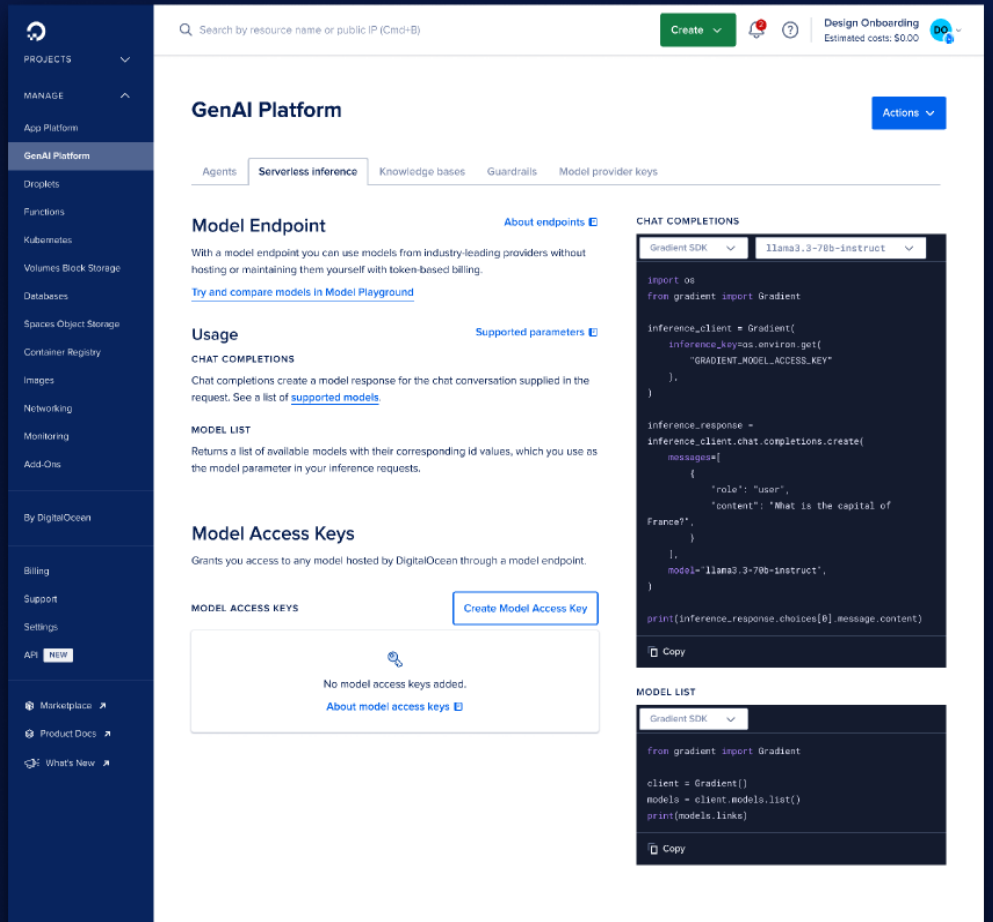
I refined the setup and configuration flows, focusing on API guidance and clear deployment usage screens.

Why it Mattered

Reducing setup friction allows developers to focus on application logic rather than server management.

Outcome / Signal

Improved the clarity of the inference setup journey, making the DigitalOcean experience feel more robust and production-ready for enterprise use.



Turning black-box agent results into observable workflows

Problem

Teams building autonomous agents lacked structured ways to inspect and evaluate behavior beyond manual judgment.

Design Action

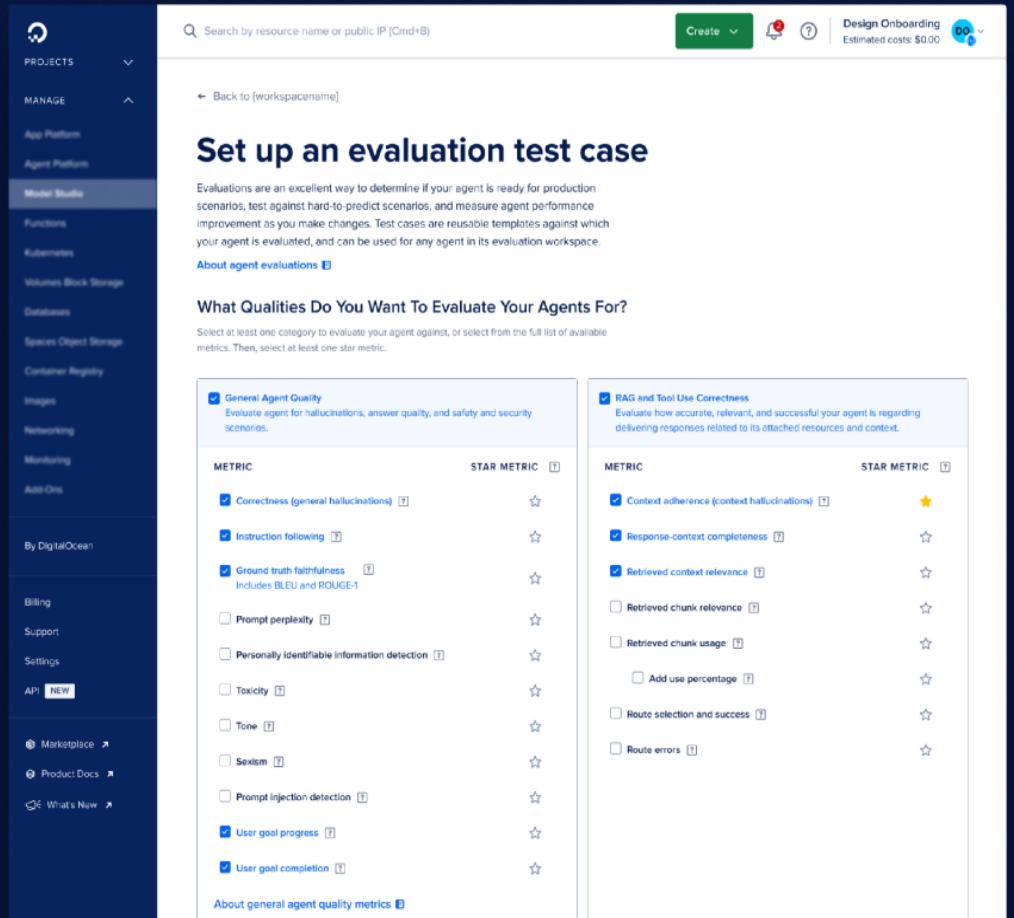
Designed evaluation dashboards that help users review outputs, score performance, and compare agent traces over time.

Why it Mattered

Systematic evaluation is critical for AI safety and performance iteration.

Outcome / Signal

Framed agent quality as a reviewable product workflow rather than a hidden result, allowing for more predictable software development.



Capturing human judgment to improve model reliability

Problem

AI systems need human feedback loops to correct behavior and improve accuracy, but feedback capture is often intrusive.

Design Action

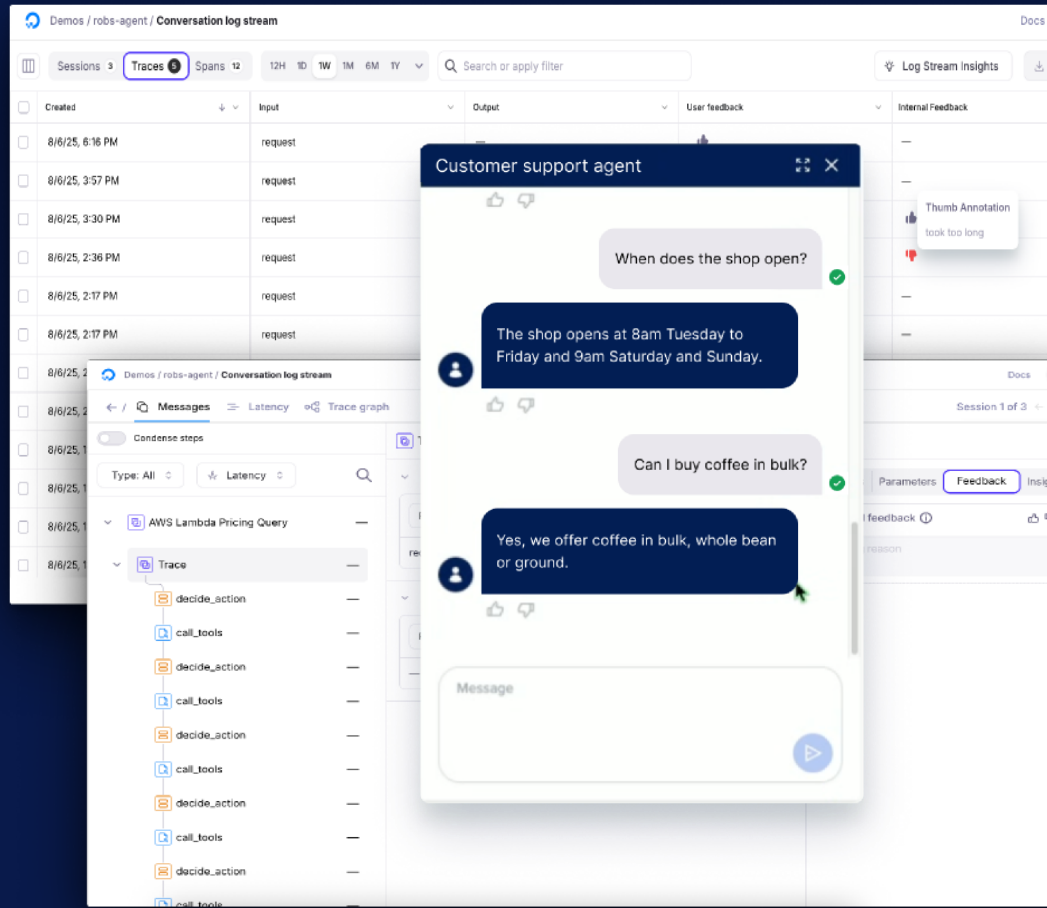
Developed lightweight feedback and review patterns that capture actionable signals without disrupting the developer's primary workflow.

Why it Mattered

Trust in AI is built on the ability to course-correct model outputs through human expertise.

Outcome / Signal

Introduced clearer pathways for turning subjective user judgment into structured data points for model fine-tuning.



Building trust through transparent knowledge base activity

Problem

Users indexing data for RAG applications lacked visibility into job success, failures, and why specific items were skipped.

Design Action

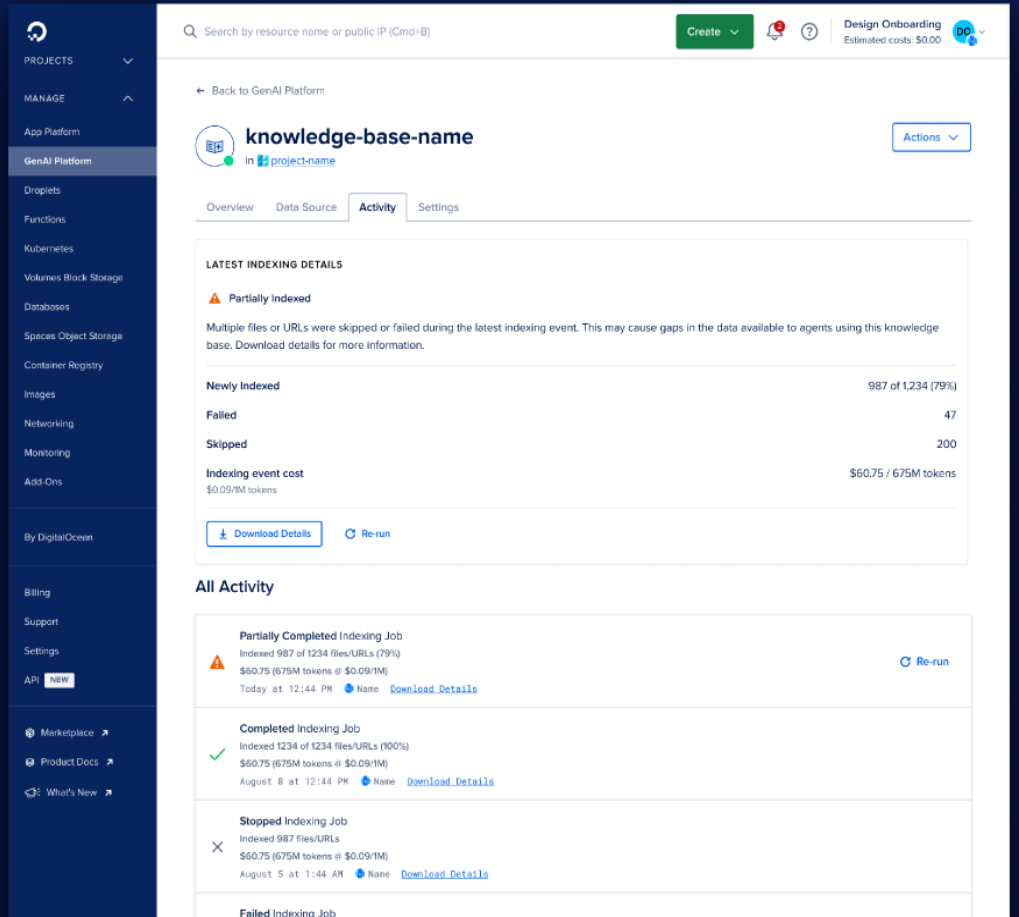
Designed comprehensive activity logs that visualize indexing progress, job-level outcomes, and detailed failure states.

Why it Mattered

Transparency in data processing reduces troubleshooting time and builds trust in the underlying knowledge engine.

Outcome / Signal

Made indexing behavior more inspectable, allowing users to manage large-scale data ingestion with significantly reduced ambiguity.



Automating knowledge maintenance for continuous AI accuracy

Problem

Maintaining an up-to-date knowledge base was a manual burden, leading to stale data in AI-driven answers.

Design Action

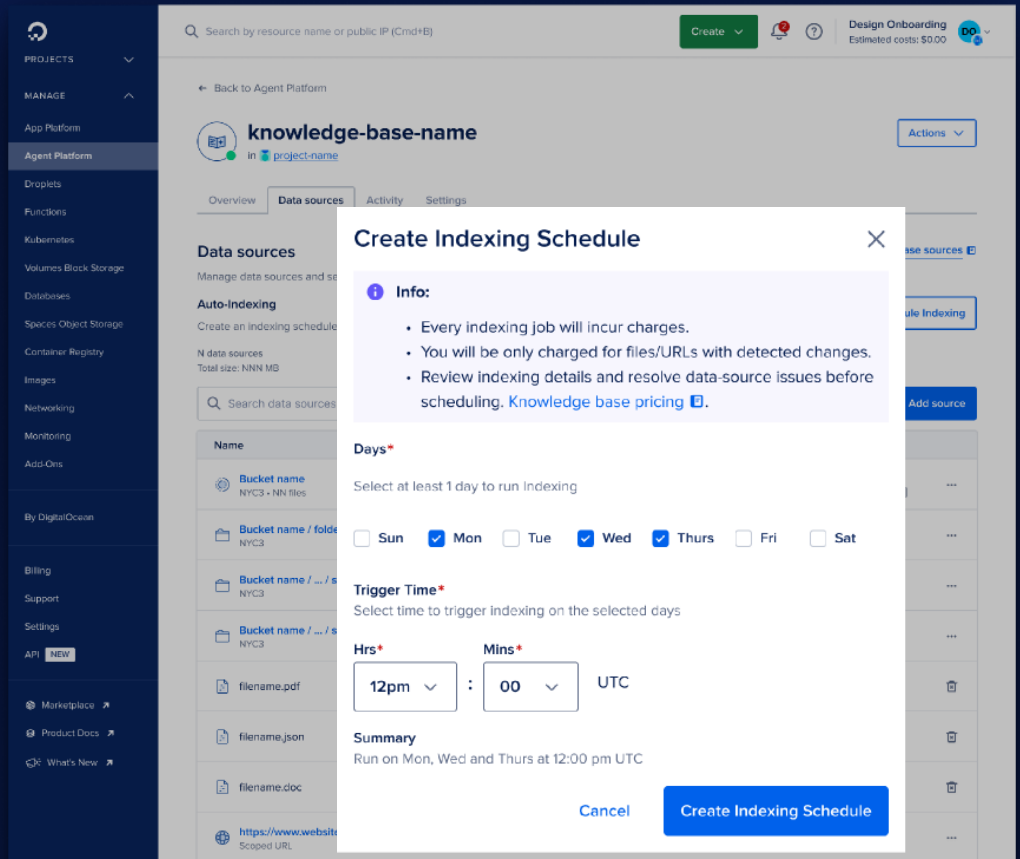
Designed automated scheduling and management patterns for recurring indexing, including state-based controls like pause and resume.

Why it Mattered

Automation transforms maintenance into a background workflow, ensuring AI reliability while preserving user control.

Outcome / Signal

Defined a clearer operational model for keeping AI knowledge sources current without manual intervention.



Driving platform impact through design leadership

My work across DigitalOcean's AI initiatives is grounded in three core leadership pillars:

Systems Thinking

I connected discovery, evaluation, and deployment into a coherent platform narrative rather than siloed features.

Product Judgment

I balanced aggressive roadmap goals with technical constraints and the need for a polished, user-centric experience.

Execution Leadership

I led cross-functional design efforts across multiple product surfaces, ensuring high-fidelity delivery in a fast-paced AI environment.

Together, these pillars ensure that design is not just a visual layer, but a strategic driver of product success.

Selected AI platform work, with deeper case studies available

Let's discuss the product context, design trade-offs, and full case studies in detail.

This deck provides a high-level summary of selected DigitalOcean AI/ML platform work. For a deeper look at the design process, user research findings, and specific metrics, I would be happy to walk through full case studies during a scheduled video call.

Thank you for your time.