R A K U A I

Spatial AI for every assistant

Scan a room with your phone. Get a 3D Gaussian-splat world that Claude, ChatGPT, and Gemini can see, measure, and reason about — over MCP.

NVIDIA Inception Member

rakuai.com · Investor & partner deck · 2026

AI assistants are blind to physical space

Today's LLMs can read your code, your email, your data — but they can't see your room.

0

rooms

have any 3D awareness in today's AI assistants by default

3+

billion

phones already pack the camera and compute to capture one

1

MCP

is the standard protocol every modern assistant now speaks

RakuAI = phone → 3D world → AI sees it

A capture-first spatial AI platform built on Gaussian splats and MCP.

1. CAPTURE

Guided phone walkthrough captures the space in 30–60 seconds. PWA — installs from any mobile browser, works offline.

2. RECONSTRUCT

COLMAP → GLOMAP → Brush → splat-transform. A real Gaussian-splat .spz lands on CDN with a metric scale solved from a credit-card reference.

3. REASON

Scene is exposed to LLMs over MCP. Claude/ChatGPT/Gemini can query geometry, run measurements, run spatial reasoning prompts.

Three layers, all real today

Production at raku-api.fly.dev. ~9,700 endpoints. C++ engine, Python orchestration, browser PWA.

raku-runtime

C++ ENGINE · MCP SURFACE

C++ spatial engine. Exposes 9,481 functions over MCP. Smart-glasses + Android XR launchable web-app entry.

raku-api

PYTHON BACKEND · ~9,700 ENDPOINTS

FastAPI orchestration. Job pipeline. Capture lifecycle + accounts + connect-flow + GPU upload routing. Redis-backed multi-instance state.

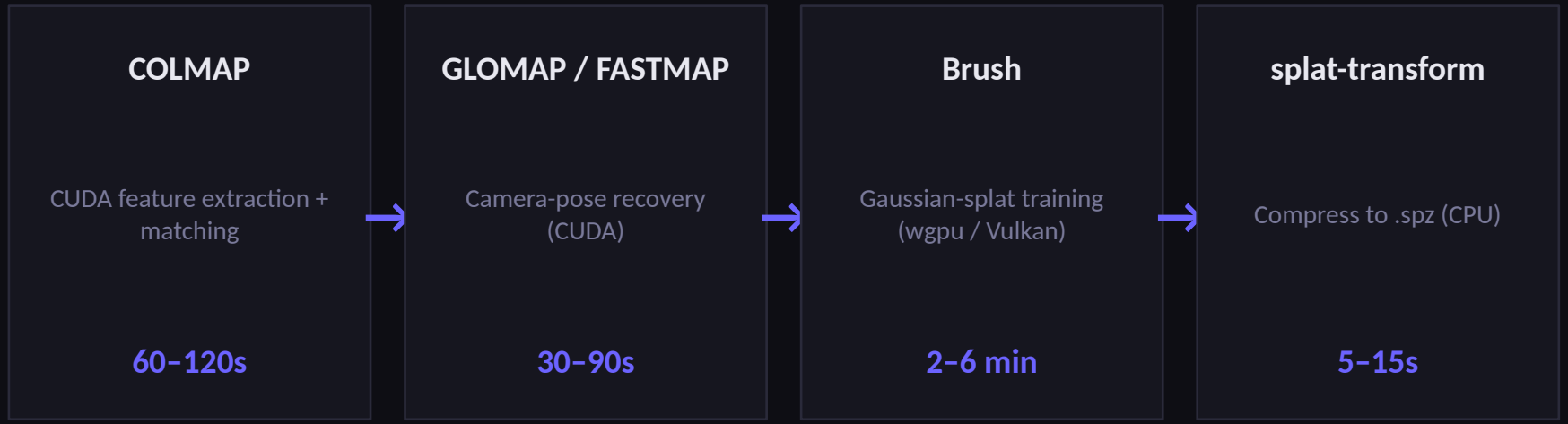
raku-capture PWA

PHONE · OFFLINE SHELL · WALK-THE-ROOM

Installable phone PWA. Guided capture, live coverage heuristic, splat viewer (Spark + three.js). Hosted at rakuai.com.

GPU-accelerated reconstruction pipeline

Reconstruction pipeline is GPU-bound from camera-pose recovery to splat training.



Target GPUs: L40S, A10, A100. Target time-to-splat: < 8 min per scan. Cost target: < $0.50 per reconstruction.

Why now

Three independent waves landed in the same 18-month window.

MCP is here

Anthropic shipped the Model Context Protocol. Claude, ChatGPT, Gemini all consume it. A standardized way to give LLMs spatial tools didn't exist before 2024.

Gaussian splats matured

Brush, Spark, three-gs. Real-time splat rendering on phones is now a solved problem. The 3D representation just works.

Smart glasses launching

Meta Ray-Ban Display ships installable web apps. Android XR opens the same path. The capture-and-view loop closes.

Every camera. Every assistant.

One capture-app + MCP runtime — three device classes, three LLM assistants.

DEVICES

Phone PWA

Today

Meta Ray-Ban Display

Web-app path

Android XR

Same web path

ASSISTANTS (via MCP)

Claude

MCP native

ChatGPT

MCP via connector

Gemini

MCP via connector

Four wedges, one platform

Once any room can be queried by an LLM, the use cases stack fast.

Real estate & redesign

List a property in 5 min; let buyers "walk" it with an AI guide. Redesign rooms with photoreal furniture (Meshy.ai integration).

Insurance & inventory

One-pass scan → itemized contents + measurements. Claims, moves, downsizing — replace a clipboard with a 60-second capture.

Accessibility

Spatial assistance for low-vision or wheelchair users. "Is there a step at the front door? Will my chair clear the hallway?"

AR/VR creators

Splat assets for game engines, training environments, OpenXR previews — the same pipeline that powers consumer captures.

Already shipped

Production-grade code on three repos. Phase 1 + Phase 2 + Wave 1–3 merged to main.

- Phone capture PWA · guided walkthrough · PWA installable · offline shell
- Reconstruction pipeline · SimulatedBackend + GlomapBrushBackend · Redis-backed multi-instance job store
- MCP surface · 9,481 functions · spatial query layer · reasoning prompt pack
- Connect flow for Claude/ChatGPT/Gemini · hosted relay + raw API key path
- Self-serve free tier · tier-aware quotas · Cloudflare R2 + CDN for asset delivery
- End-to-end CI on self-hosted runners · 1946/1969 tests passing · 76% coverage

Live: rakuai.com · github.com/RakuXR

The bottleneck is GPU compute

Every other layer of RakuAI is solved or cheap. Reconstruction GPUs are the line-item.

TODAY

- SimulatedBackend on prod
- No real .spz served
- Metric scale stays a documented SEAM
- Users see placeholder geometry

WITH GPU CREDITS

- GlomapBrushBackend live on prod
- Real Gaussian-splat .spz per scan
- Real metric scale (cm-accurate)
- LLMs query true room geometry

Unit economics target: $0.20–$0.50 GPU cost per reconstruction. Investment unlocks multi-cloud scale-out and the team to ship.

Roadmap

Three quarters from credits to real spatial AI in users' hands.

Q2 2026	Q3 2026	Q4 2026	2027
Light up real reconstruction	**Smart-glasses path**	**Scale + multi-cloud**	**Spatial AI everywhere**
Local T2000 validation. Deploy GlomapBrushBackend to Azure / Modal. Flip mode=real on prod.	Meta Ray-Ban Display web-app launch. Capture-from-glasses pilot. Real-time scene query from the headset.	Wave 5 pluggable worker. Multi-vendor inference evaluation for scene-reasoning. Creator-tool integrations (Meshy, Omniverse, etc.).	Android XR launch. SDK partner program. "Talk to any room" becomes table-stakes for assistants.

Join us.

Spatial AI is happening. RakuAI is the platform that connects every camera to every assistant.

PARTNERS

Smart-glasses OEMs, LLM platforms, AI labs. Integrate the MCP capture surface, ship spatial agents to your users.
partners@rakuai.com

INVESTORS

Seed round opening Q3 2026. GPU compute, graphics engineering hire, smart-glasses launch partners.
kevin@rakuai.com

DEVELOPERS

MCP runtime preview, Brush+Spark splat tooling, free tier on rakuai.com.

github.com/RakuXR · rakuai.com

Kevin Griffin · Founder, RakuAI, LLC

kevin@rakuai.com · rakuai.com · github.com/RakuXR