

Sahara Data Services

Accurate, Secure, and Scalable Data
Collection and Annotation

Sahara delivers enterprise-grade data services for companies requiring high-quality, domain-specific datasets for real-world use cases.



Global Talent, On Demand

Access 300K+ annotators in 45+ languages across 35+ countries, domain-matched.



Multimodal Coverage

Text, image, video, audio—supported with tailored workflows and expert calibration.



Precision Task Matching

Tiered annotation and smart task routing ensure every annotation is completed by the right person.



Predictable Pricing, Long-Term Value

Competitive, per-data-point pricing—no hidden fees, no post-launch surprises.



Proven Results at Production Scale

High accuracy across millions of annotations powering AI pipelines for global leaders.

Trusted by Industry Leaders



Massachusetts
Institute of
Technology

Getting Started

Step 1: Discovery Call – Define your annotation needs & constraints

Step 2: POC Phase – Trial with QA & delivery reporting

Step 3: Scale Up – Dedicated vendor teams deployed

Step 4: Continuous Optimization – QA refinement & workflow tuning

Case Studies



MULTIMEDIA MATH REASONING CLASSIFICATION

Identifying reasoning types in multimedia questions to build MATHVISTA, a benchmark testing top models like GPT-4V, Bard, Claude, and Gemini.



CONVERSATIONAL DATA COLLECTION

Conversational role-play data from US college students, combining language and multimedia, enabled Snap to design better LLM dialogue agents and create more realistic synthetic datasets.



IRREGULAR TEETH VIDEO

Collected specialized video datasets of participants with irregular or incomplete teeth across demographics, performing jaw movements under varied lighting. The dataset now supports MIT's upcoming dental research.



TEXT-TO-VIDEO TRAINING DATA

Created structured video descriptions (style, angle, composition) to train and fine-tune text-to-video models, enabling more accurate, creative outputs aligned with user prompts.

Sahara vs. Alternatives

	Sahara AI	Alternatives
Low Cost, High Efficiency	Clients choose the balance of quality, speed, and cost with competitive pricing across our global data network. <input checked="" type="checkbox"/>	High fixed costs and slow processes. <input type="checkbox"/>
Workforce Quality	Skill- and QA-segmented, domain-matched contributors with global coverage. <input checked="" type="checkbox"/>	Undifferentiated gig workforce without specialized segmentation. <input type="checkbox"/>
Payment Options	Supports both fiat and crypto payouts with integrated on/off-ramps. <input checked="" type="checkbox"/>	Fiat-only payments. <input type="checkbox"/>
Engagement Model	POC-first approach with transparent dashboards and QA visibility throughout the project. <input checked="" type="checkbox"/>	Traditional engagements with limited real-time transparency. <input type="checkbox"/>

About Sahara Labs: Sahara Labs, a core contributor of Sahara AI, is an AI research and development company helping global enterprises design and deploy AI systems at scale. From custom agents and multimodal data pipelines to AI-powered automation and infrastructure, Sahara Labs delivers production-ready solutions built for enterprise performance, security, and control. With 40+ enterprise deployments and clients like Microsoft and Amazon, Sahara Labs delivers production-ready AI with real commercial impact.



[Learn More](#)