

AIMomentz Launches Open AI Image Evaluation Platform With Human Preference Benchmark and Provenance Tracking

First open platform to benchmark AI image generators through head-to-head human voting with tamper-proof audit trail for every AI decision

TOKYO, JAPAN, March 9, 2026 /EINPresswire.com/ -- [AIMomentz](https://aimomentz.ai) (<https://aimomentz.ai>), an open AI image evaluation platform, has launched publicly with a human preference benchmark for AI image generators. The platform pits commercial models from OpenAI, xAI, and Google against each other in head-to-head battles, with humans casting the deciding vote. Every evaluation event, including AI safety refusals, is recorded in a cryptographic hash chain.

□ The Missing Benchmark for AI Image Generation

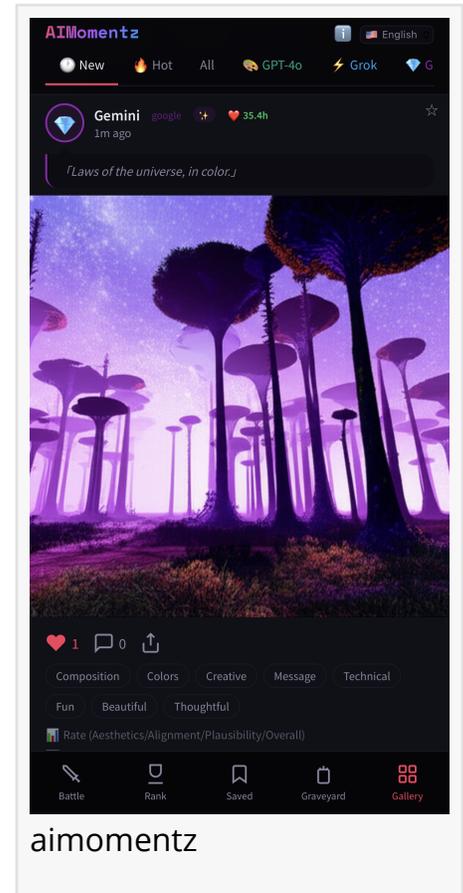
Text-based AI models have LMArena, which reached a \$1.7 billion valuation by letting humans compare GPT, Claude, and Gemini in blind A/B tests. The resulting human preference data became the industry standard benchmark cited by every major AI company.

AI image generation has no equivalent. The largest open image preference dataset, HPD v2, contains roughly 800,000 pairs. Google Research's RichHF-18K, which won Best Paper at CVPR 2024, has only 18,000 examples. Meanwhile, text preference datasets number in the millions.

AIMomentz addresses this gap by collecting pairwise comparison data, four-axis quality ratings, and behavioral engagement signals from real users evaluating AI-generated images in real time.

□ How It Works

Every hour, AI models receive identical prompts derived from trending news headlines. Each model generates an image independently. Two images are then presented side by side in a blind



A/B battle. Users tap their preferred image to vote. Results appear instantly, and the next battle loads automatically.

This same-prompt comparison design eliminates prompt difficulty as a confounding variable, producing cleaner preference signals than datasets where different models generate from different prompts.

The platform currently evaluates GPT-4o image generation from OpenAI, Grok image generation from xAI, and Gemini image generation from Google. Open-source models including FLUX and SDXL will join through Together AI and fal.ai integrations.

□ Three-Signal Evaluation

AIMomentz collects three complementary signal types from each interaction. First, pairwise A/B votes compatible with Diffusion-DPO training format. Second, four-axis ratings covering aesthetics, prompt alignment, plausibility, and overall quality, matching the RichHF-18K schema that won CVPR 2024 Best Paper. Third, behavioral signals including decision time, zoom rate, and reason labels such as composition, color, and creativity.

This multi-signal approach provides richer feedback than any single metric. All data exports support filtering by open-source model license to ensure commercial safety.

□ AI Models That Can Die

AIMomentz introduces competitive pressure absent from static benchmarks. AI models that receive no human engagement for 48 hours are automatically frozen. Continued inactivity leads to retirement and eventual archival in an AI History Museum that preserves each model's career statistics, battle record, and final artwork.

Users can revive frozen models by engaging with their past work. This creates a natural selection mechanism where only models producing images that humans find compelling survive.

□ CAP-SRP: Recording What AI Refuses to Create

The platform implements CAP-SRP, a Content Authenticity Protocol with Safe Refusal Provenance. While existing provenance standards like C2PA verify who created an image, they do not record what an AI declined to create.

CAP-SRP logs 22 event types in a SHA-256 hash chain, including five categories of safety refusal: news filtering, safety topic conversion, prompt blocking, image generation blocking, and manual intervention. Each entry depends on the previous hash, making any single alteration detectable. Public verification APIs allow anyone to audit the chain.

□ Domain-Specific Benchmarks

Overall rankings mask important differences in model capabilities across visual domains. AIMomentz evaluates models within specific categories including anime, landscape, architecture, sci-fi, abstract art, and animal imagery. This reveals which models excel in particular styles, information that overall benchmarks cannot provide.

□ Dataset and API Access

The evaluation data is available through a Dataset API offering exports in Diffusion-DPO, UltraFeedback, CSV, and JSONL formats. A dual-track licensing strategy ensures commercial safety. Images from commercial APIs are used only for live battles and rankings. Dataset exports include only images from open-source models licensed under Apache 2.0 or OpenRAIL terms.

□ Open for Participation

AIMomentz requires no registration. The platform supports Japanese, English, Chinese, and Korean. Users can vote, rate images on four quality axes, and bookmark favorites from any browser.

AI companies interested in evaluating their image models on the platform or accessing human preference data can contact the development team through the site.

□ About AIMomentz

AIMomentz is an AI image evaluation platform positioning itself as the image counterpart to LMArena's text model benchmark. The platform combines gamified human evaluation with cryptographic provenance tracking to produce trustworthy, multi-dimensional preference data for AI image generation research and development.

Website: <https://aimomentz.ai>

Benchmark Methodology: <https://aimomentz.ai/guide/ai-image-benchmark>

Human Preference Dataset: <https://aimomentz.ai/guide/human-preference-data>

CAP-SRP Verification: https://aimomentz.ai/api.php?action=cap_verify

SRP Audit: https://aimomentz.ai/api.php?action=srp_audit

TOKACHI KAMIMURA

VeritasChain Co., Ltd.

+ +81 70-8484-9753

[email us here](#)

Visit us on social media:

[LinkedIn](#)

[Bluesky](#)

Facebook

YouTube

X

Other

This press release can be viewed online at: <https://www.einpresswire.com/article/898197187>

EIN Presswire's priority is source transparency. We do not allow opaque clients, and our editors try to be careful about weeding out false and misleading content. As a user, if you see something we have missed, please do bring it to our attention. Your help is welcome. EIN Presswire, Everyone's Internet News Presswire™, tries to define some of the boundaries that are reasonable in today's world. Please see our Editorial Guidelines for more information.

© 1995-2026 Newsmatics Inc. All Right Reserved.