

Evaluate AI Models for Cultural Understanding

The first research-backed framework for comprehensive cultural AI evaluation.
Identify blind spots before they become production issues.

47

DIMENSIONS

8

CULTURAL PERSPECTIVES

42

MODELS BENCHMARKED

7,410+

ANNOTATIONS

The Challenge

HIDDEN CULTURAL RISKS

As AI models grow more powerful, they produce sophisticated yet subtly biased cultural content. Single metrics fail to capture multi-dimensional cultural understanding.

VULCA SOLUTION

Reproducible 47-dimension evaluation across 8 cultural perspectives with explainable diagnostics and deliverable reports.

Key Capabilities

- ✓ **L1-L5 Cognitive Framework** — From visual perception to philosophical aesthetics
- ✓ **8 Cultural Perspectives** — Western, Eastern, African, Latin American, Middle Eastern, South Asian, Oceanic, Indigenous
- ✓ **47D Expanded Diagnostics** — 6 core dimensions expanded to 47 fine-grained metrics
- ✓ **Reproducible & Citable** — Version-controlled evaluation with full provenance
- ✓ **Export-Ready Reports** — PDF reports, JSON/CSV data, BibTeX citations
- ✓ **Batch Evaluation** — API access for continuous monitoring

Who It's For



AI Labs & Companies

Pre-release cultural audits, model selection, fine-tuning validation, 47D regression tracking



Research Institutions

Academic benchmarking, reproducible methodology, citation-ready reports, version control



Museums & Galleries

Multi-cultural AI validation, interpretive AI quality, cross-cultural accuracy assurance

Transparent Pricing

Free

\$0

forever

- Public leaderboard
- Limited demo evaluations
- Citation export

Pilot ★

\$2,500

per engagement

- Full 47D evaluation report
- 8 cultural perspectives
- PDF + evidence samples
- **Scope:** 1 model, ≤50 images
- **NDA available** upon request

Enterprise

From \$15K

annual

- Unlimited evaluations
- Custom benchmarks
- API/SDK access
- SLA & priority support

What You Get (Pilot)



47D Evaluation Report



8 Perspective Analysis



PDF + JSON Export



Risk Identification



Evidence Samples



Model Comparison



Citation Ready



48-72h Delivery

EMNLP
2025

Findings of
ACL

WiNLP
2025

Widening NLP
Workshop

DOI

10.18653/v1/2025.findings-
emnlp.103

GitHub

Open-source
benchmark code


Ready to Evaluate Your Models?

Book a 15-minute demo to see VULCA in action. No credit card required.

 demo@vulcaart.art

 [Book Demo](#)

 [Download Sample Report](#)

 **Data Privacy:** NDA available. Your data is never used for training.