

SALAH-EDDINE BENBRAHIM, PhD, MBA, B.Eng.

AI Researcher — Trustworthy, Evaluation & Reproducible ML · Multimodal · Open to international relocation

• benbrahim.salah.eddine@cineinfinite.com • +1 647-875-0826 (CA) · +33 7 81 64 82 93 (FR) • [LinkedIn](#)

cineinfinite.com • github.com/CineInfini/Video_Quality_Auditor • huggingface.co/benbrahimsalaheddine777 • orcid.org/0009-0005-7050-4014 • [CineInfini Studio \(live\)](#) · publications.polymtl.ca

RESEARCH SUMMARY

AI researcher working on the trustworthiness, evaluation, and reproducibility of machine-learning systems, with a PhD in Computer Engineering and 10+ years of applied practice. My current research asks how we can know that an AI system — or the human data used to train and judge it — is reliable, and how to make that judgement reproducible. Sole author of two papers currently under review on the measurement layer beneath AI evaluation. Creator of CineInfini, an open-source evaluation framework (>90% test coverage, <5% measurement variance, DOI-archived). I publish, build in the open, and am eager to contribute to an internationally oriented research environment.

RESEARCH INTERESTS

Trustworthy & responsible AI · evaluation and benchmarking of AI systems · reproducibility and metrology of ML · reliability of human annotation and preference data · LLM-as-a-judge · multimodal (video / AIGC) models · safety and robustness evaluation.

SELECTED PUBLICATIONS & RESEARCH OUTPUT

Index terms: inter-rater reliability; variance components and identifiability; human reference panels; LLM-as-a-judge evaluation; annotation budget allocation; calibrated diagnostics; reporting protocols. Sole author. *Under review* (2026).

Index terms: inter-rater reliability; construction choices in reliability estimation; rater screening and aggregation; panel size; multiverse analysis; reproducibility of published coefficients. Reproduction package: doi.org/10.5281/zenodo.22698801. Sole author. *Under review* (2026).

Seven peer-reviewed papers (3 journal articles, 4 conference papers) spanning cloud computing and distributed systems, network optimization under constraints, and information-systems project management. Citation record: 41+ citations, h-index 4 — most cited: IEEE Trans. Cloud Computing (15), IJISPM (11), Procedia Computer Science (6) (IEEE Transactions on Cloud Computing, Procedia Computer Science, NBIS, Asia Modelling Symposium; Int. Journal of Information Systems & Project Management). Full profile: publications.polymtl.ca

RESEARCH & PROFESSIONAL EXPERIENCE

Founder & Open-Source AI Research Lead — CineInfini 2023 – Present

- Designed and operate CineInfinite, a live multimodal media governance platform: 14 governed evaluation corpora (1.4M+ rated assets, seven distinct rating scales), 24 scoring modules across 6 categories, 134 features. Includes trustworthiness scoring, causal-reasoning validation, physics-plausibility checking, world-model surprise detection, multi-modal safety assessment, AIGC origin detection, and Shapley-like attribution; integrates DOVER (ICCV 2023), FAST-VQA (ECCV 2022), DINOv2 and ArcFace as licence-tracked reference scorers.
- Built the reproducible measurement pipeline underlying it: >90% test coverage executing in under 30 seconds, <5% variance across repeated measurements across 14 governed corpora, computation-level audit logging, semantic versioning, automated CI/CD with go/no-go release gating, PyPI packaging, and Zenodo DOI archival.
- Implemented the multimodal analysis stack: hierarchical shot detection, DINOv2 and ArcFace embeddings, CLIP semantic scoring, optical flow, and Dynamic Time Warping for temporal coherence.
- Released CineInfini Studio, a public live environment computing perceptual quality scores (SRCC, PLCC, KRCC) and running an LLM judge over the results.

Postdoctoral Research Fellow — Polytechnique Montréal 2017 – 2018

- Research on project control and information-systems processes: network analysis of project-control processes, published at CENTERIS 2018 and in the International Journal of Information Systems and Project Management.

PhD Researcher & Lecturer — Polytechnique Montréal 2009 – 2016

- Doctoral research on real-time migration of interdependent virtual machines: mixed-integer programming and Tabu Search metaheuristics validated on Google traces, reducing service penalties by 13.5% (avg.) and physical-machine usage by up to 37%; published in IEEE Transactions on Cloud Computing, NBIS, and AMS.
- Taught MATLAB programming to first-year engineering undergraduates at Polytechnique Montréal, in a bilingual academic environment.

Technical Lead / AI Architect — Gov. of Quebec (2023 – 2025) & JokeJoy (2023 – 2025)

- Led end-to-end architecture for government digital platforms serving 2M+ citizens: reduced latency by 40% (2.5s to 1.5s), maintained 99.9% uptime, and delivered eight major releases on schedule with zero critical incidents.

- Engineered a multi-LLM production system (GPT-4o primary, Gemini 1.5 fallback) with automatic provider switching on latency and cost thresholds and dynamic context-injection prompting.
- Earlier, as Senior BI Data Analyst in government (2019–2020): data governance and quality assurance, reducing data-quality issues by 35%, and mentoring a team of 8 analysts — direct experience of how measurement and evaluation fail in practice at scale.

EDUCATION

PhD, Computer Engineering — Polytechnique Montréal, Université de Montréal (February 2016). Dissertation on real-time migration of interdependent virtual machines in cloud networks.

OPEN-SOURCE & PUBLIC ARTIFACTS

- CineInfini framework — open-source evaluation engines with reproducible pipelines, audit logging, semantic versioning, CI/CD gating, PyPI packaging, and Zenodo DOI archival.
- Designed the methodology, architecture and statistical framework myself, and implemented the platform with AI-assisted development under a verification protocol — the same discipline it applies to the systems it evaluates.
- CineInfini Studio — a public, live evaluation environment computing perceptual quality scores (SRCC, PLCC, KRCC) and running an LLM judge over the results.

TEACHING & MENTORING

- Taught MATLAB programming to first-year engineering undergraduates at Polytechnique Montréal; mentored a team of 8 analysts in a government setting.
- Able to teach and supervise in English and French, across computer science, machine learning, data science, and evaluation methodology.

TECHNICAL ENVIRONMENT

Languages & frameworks: Python (async/concurrent), PyTorch, ONNX, pytest, GitHub Actions, PyPI.

Models & multimodal: GPT-4/4o, Gemini 1.5, Claude, DeepSeek; CLIP, DINOv2, ArcFace, optical flow, Dynamic Time Warping; RAG and prompt engineering.

Methods & statistics: inter-rater reliability, multiverse and sensitivity analysis, SRCC/PLCC/KRCC, mixed-integer programming, metaheuristics.

Infrastructure: AWS, GCP, distributed systems, microservices and REST, CI/CD with release gating, Zenodo DOI archival.

ADDITIONAL INFORMATION

- I can work in Canada, France and across the European Union and EEA. Eligible for TN status under the USMCA for the United States. Available in Montréal, Toronto and Paris.
- Proven self-direction: founded, architected, funded, and shipped a complete research programme and its open-source tooling end to end — the autonomy expected of a faculty member or fellow, already demonstrated.
- Interested in collaboration across evaluation methodology, trustworthy AI, AI for science, and applied multimodal machine learning; available for research-institute, faculty, teaching-track, or postdoctoral appointments.

LANGUAGES & MOBILITY

English (fluent, professional) · **French** (native). **Open to international relocation** (Gulf, Asia, Europe, Africa) and to research-institute, faculty, or postdoctoral appointments.