

BRIAN GISA

Multilingual AI Evaluation Specialist | Low-Resource Language & Reasoning Quality

Ottawa, Canada • briangisaa@gmail.com • (613) 470-2233 • [linkedin.com/in/brian-gisa/](https://www.linkedin.com/in/brian-gisa/)

Available globally as an independent contractor (Deel / Remote.com / direct invoice)

PROFILE

AI evaluation specialist with 4+ years of experience assessing model outputs for factual accuracy, reasoning quality, instruction-following, and safety. Native Kinyarwanda speaker with strong reading comprehension in French, an unusual combination for a field that struggles to find qualified reviewers for low-resource languages. Builds structured evaluation rubrics, automates research workflows in n8n, and writes detailed comparative justifications that materially shift model behavior in RLHF.

CORE COMPETENCIES

- **Output Evaluation:** Evaluating model responses for accuracy, reasoning soundness, instruction adherence, and harm potential
- **Multilingual Review:** Native Kinyarwanda; native English; French (strong reading & comprehension) — translation faithfulness, cultural register, code-switching
- **Reasoning Audits:** Identifying logical fallacies, hallucinations, unsupported inferences, and chain-of-thought breaks in long-form reasoning
- **Rubric & Feedback Design:** Authoring detailed comparative justifications and rubric refinements that improve inter-annotator agreement
- **Automation & Tooling:** n8n agentic workflows, Google Sheets / Airtable evaluation dashboards, Notion documentation systems

PROFESSIONAL EXPERIENCE

AI Evaluation Specialist (Independent Contractor) — *Multiple AI evaluation platforms · Remote*

Jan 2023 – Present

RLHF, multilingual evaluation, and reasoning-quality projects supporting frontier-model and applied-AI clients.

- Maintained a quality audit score above 95% across multiple high-stakes reasoning and instruction-following projects, with output reviewed in calibration rounds against senior reviewers.
- Specialized in flagging subtle reasoning failures: Unsupported inferences, premise-drift in multi-turn dialogue, and confident hallucinations in administrative and customer-facing scenarios — that automated checks miss.
- Drew on native Kinyarwanda and reading proficiency in French to assess translation fidelity, cultural register, and code-switching behavior on multilingual tasks where qualified reviewers are scarce.
- Authored hundreds of comparative justifications per month explaining why one model response was preferred over another. The structured-reasoning artifact is directly used to fine-tune model preferences.
- Surfaced rubric ambiguities to project leads across four distinct reasoning workstreams, contributing edits that reduced annotator disagreement and tightened guidelines for downstream contributors.

AI Workflow Automation Extern — *Wayfair (Market Intelligence)*

Jan 2026 – Mar 2026

- Built n8n workflows that automated competitor and trend monitoring for category managers, replacing recurring manual research with structured weekly outputs.
- Designed live Google Sheets dashboards that translated raw market signals into decision-ready views for merchandising stakeholders.

AI Outreach Quality Analyst (Extern) — *HorizonX*

Sep 2025 – Dec 2025

- Audited 100+ AI-generated outreach sequences for clarity, intent fidelity, and logical flow — flagging the edge cases where automation produced confident but incorrect follow-ups.
- Stress-tested three sales-automation tools, documenting reproducible logic gaps in lead-scoring and reply-handling pipelines.

Operations & Documentation Coordinator (Independent) — *Hospitality & SMB clients, Kigali* Sep 2021 – Dec 2022

- Standardized 7+ internal SOPs and digital filing systems for distributed hospitality teams — operational discipline that later transferred directly into evaluation-rubric design.
- Coordinated multilingual client correspondence across English, French, and Kinyarwanda — calendars, shared drives, and documentation pipelines.

INDEPENDENT PROJECTS

- **AI Failure-Mode Notes:** Public series of short write-ups on subtle reasoning failures observed in multilingual model output, with anonymized examples and structured analysis of root cause. (Published on LinkedIn.)
- **Kinyarwanda Evaluation Rubric:** Working evaluation rubric for assessing translation fidelity, cultural register, and code-switching in Kinyarwanda model outputs — a published reference for an underserved language pair.

SELECTED EVALUATION WORK — PATTERNS I CATCH

- **Premise drift:** Long-context reasoning tasks where the model contradicts a premise stated 6–8 turns earlier without flagging the change.
- **Multilingual register failures:** Outputs that appear bilingual but quietly substitute English syntax patterns onto French / Kinyarwanda — fluent-sounding, structurally wrong.
- **Tone-mismatched safety behavior:** Refusals or compliance triggered by surface keywords rather than actual intent — frequent in customer-service scenarios where the model misreads tone.
- **Confident hallucination:** Plausible citations or procedural steps that don't trace back to any source — most dangerous in admin and policy contexts where surface fluency masks fabrication.

EDUCATION

MSc, Project Management & Data Analytics (in progress) — *University of London (online)*

Expected Sep 2026

BA, Business Communications — *Southern New Hampshire University*

Graduated 2023

Diploma, Business Management & Entrepreneurship — *Algonquin College*

LANGUAGES

English (Native) • Kinyarwanda (Native) • French (Strong reading & comprehension; written review on request)

Available for English and Kinyarwanda evaluation, translation-quality review, and red-teaming. French: reading-based review and comprehension tasks.

TOOLS & PLATFORMS

n8n • Google Workspace (advanced Sheets) • Airtable • Notion • Slack • Trello • Microsoft Office • Deel • Remote.com

AVAILABILITY

Based in Ottawa, Canada. Open to global remote contract engagements via Deel, Remote.com, or direct invoicing — no employer sponsorship required for contractor work. Standard availability across North American and European time zones.

References available on request.