ROKI SEYDI

roki@rokiseydi.com · rokiseydi.com · github.com/RokiSeydi · London

Researcher working on computational human context, behavioural agents, and sociotechnical AI safety, with healthcare as a primary application. Investigates how lived experience can be encoded without flattening people into static attributes, and uses context-grounded agents to evaluate AI systems under realistic human conditions. Published work on cultural confabulation and deployment-context failure modes. Spärck AI Scholar (UCL MRes); MPH (Imperial). Engineering background spanning regulated healthcare and financial systems used by millions.

EDUCATION

University College London — MRes AI-Enabled Healthcare *2026 – current*

Inaugural government backed Spärck AI Scholarship Awardee - awarded 1 out of 100 scholarships for exceptional talent in AI

Imperial College London — MPH Global Public Health *2020 – 2024*

International Distinction Award. Dissertation (First Class): *Integrating Lived Experience in NHS Services to Improve Digital Adoption.*

King's College London — BA Global Health and Social Medicine *2016 – 2019*

First Class Dissertation on idioms of distress and culturally situated understandings of health, presented at Oxford anthropology conference as sole undergraduate presenter.

AWARDS & SCHOLARSHIPS

- **International Distinction Award** — merit-based funding for MPH, Imperial College London
- **King's Sanctuary Award** — awarded for excellence in research on migration and health
- **Bertelsmann Technology Scholarship** — fully funded scholarship in product and AI
- **Macquarie Group Scholarship** — fully funded scholarship in cloud computing

RESEARCH EXPERIENCE

Independent Researcher — Healthcare AI Evaluation *2025 – Present*

- Designed benchmark datasets, scoring rubrics, clinical quality dimensions, and validation protocols for evaluating medical reasoning, clinical decision support, and deployment-context performance across frontier language models.
- Identified and operationalised *Cultural Confabulation* — a previously uncharacterised failure pattern where models misapply contextual assumptions in patient-facing healthcare scenarios.
- Developed multi-domain evaluation corpus grounded in real-world community health, fragile health system, and culturally-situated patient scenarios; established cross-model grading methodology quantifying systematic grader bias across frontier model families.
- Distinguished *structural* failure modes (addressable only by frontier developers) from *recoverable* failure modes (resolvable by context engineering at deployment) — a distinction with direct implications for clinical deployment and governance allocation.
- Built evaluation infrastructure on the UK AI Security Institute's Inspect Evals framework with automated cross-model grading and reproducible scoring pipelines.

Research Collaborator — Community Health Impact Coalition (CHIC) *2026 – Present*

- Member of an international research network spanning 20+ countries. Contributing to a Lancet Primary Care Delphi-style consensus publication on AI governance in community health.

- Collaborating with clinicians, ministry-of-health stakeholders, WHO advisors, and policy researchers on evaluation and deployment standards for healthcare AI in low-resource settings.

Lived Experience Advisor — Bromley Healthcare NHS *2024 – Present*

- Advise NHS healthcare leaders on digital adoption, health inequalities, and social determinants of health in primary and community care services.
- Contribute to multi-stakeholder service evaluation and improvement initiatives.

ACADEMIC SERVICE & LEADERSHIP

- **Reviewer** — ICML 2026 Trustworthy AI Workshop
- **Founder & Organiser** — Clinical AI Safety Network, BlueDot Impact alumni community
- **Moderator** — Skoll World Forum 2026, AI Monitoring, Evaluation & Evidence Roundtable
- **Speaker** — Technical AI Safety Conference 2026, University of Oxford
- **Panellist** — King's College London, NHS Reform & Disability Rights

PUBLICATIONS

- Seydi, R. (2026). *Cultural Confabulation: A Structural Evaluation Gap in Large Language Model Reasoning for Healthcare.* Technical AI Safety Conference Proceedings. tais2026.cc/proceedings/seydi-cultural-confabulation

Working papers

- *The Gap Proof: Measuring Deployment-Context Failures in Frontier Language Models for Healthcare.*
- *The Auditor's Blindspot: Systematic Bias in Cross-Model Evaluation of Healthcare AI.*

Public scholarship: *"AI Doesn't Hallucinate My Mother's Culture. It Rewrites It."* Artificial Intelligence in Plain English (40,000+ readers).

ENGINEERING EXPERIENCE

Good Boost — Founding Backend Engineer *2020 – 2022*

Sole backend engineer for a digital health platform serving 20,000+ patients. Built secure clinical data systems handling sensitive patient information under healthcare data-governance requirements. The platform received five industry awards during my tenure.

tell.money — Full Stack Founding Engineer *2022 – 2024*

Delivered open-banking infrastructure operating under FCA regulation; systems used by 5M+ retail banking customers across major UK banks.

Antler UK — Founder in Residence *2025*

Selected for Antler's venture-building programme. Developed AI-enabled healthcare prototypes and conducted clinical user research.

LANGUAGES

English · Italian · Spanish · Mandinka