Richard J. Young, Ph.D.

RY

Applied AI Research Scientist. Foundation models from frontier evaluation to production at national scale, across healthcare, cybersecurity, and decision-support.

richard@deepneuro.ai **916-380-2941** mobile (text ok) · 415-881-1115 office Las Vegas · Portland cv.deepneuro.ai github/ricyoung hf/richardyoung

---

Foundation-model safety is empirical, not philosophical. I run large, replicable evaluations under adversarial, clinical, and decision-support pressure, then ship the benchmarks, datasets, and models openly so the failures can be reproduced and fixed.

96–100%
attack success on 6 of 10 frontier LLMs · TEMPEST replication, 97K queries, 8 vendors

150K → 5M+
members served, sat 0→8.8/10 · UHC Consumer Resolution Center, in *Harvard Business Review*

110M+ lives
70K+ physicians under IRB oversight · Co-Chair, UHG Institutional Review Board

55K+
open-weight downloads across 80+ releases; 80+ data scientists on internal RAG platform

FLAGSHIP AI IN PRODUCTION

CxS Insights — DEPLOYED

Consumer-abrasion analytics platform · non-technical teams query member distress & call metrics without SQL.

47.5M members
260M+ transcripts · 75+ multi-agent LLM tools · secure onshore/offshore access controls

Member Outreach — IN HBR

MLOps distress-scoring pipelines surfacing at-risk members daily for UHC's Consumer Resolution Center.

150K → 5M+ scaled
satisfaction 0→8.8/10 · *Harvard Business Review*, 2026

Ask Richard — DEPLOYED

Upskilling & knowledge platform · daily source of truth for Databricks, RAG, and LLM workflows across the team.

80+ daily users
27,982 docs + 6,687 images · 3-mode RAG · 12+ LLM ecosystem

WHAT I DO

AI Safety & Evaluation

Adversarial multi-turn evaluation at frontier scale, guardrail robustness, instruction adherence, chain-of-thought faithfulness, and code-safety vs. content-safety separation.

Code Safety & Cybersecurity

The "Code as a Weapon" research program with Dr. G. D. Moody (UNLV Lee, Director of Cybersecurity Programs): authored the canonical 1,554-prompt benchmark and a growing body of studies separating executable malicious-code generation from defensive security knowledge: multi-vendor behavioral evaluation, mechanistic abliteration, and systematic review across 13 corpora.

Clinical AI at Deployment Scale

PHI leakage in medical OCR, fairness audits of LLM-based ED triage (EQUITRIAGE), domain-specialized clinical embeddings (CardioEmbed), and clinical-trial extraction with multi-LLM ensembles.

Production AI Systems

Five conversational-analytics platforms on Databricks serving Optum / UHC: LLM direct tool-calling and agentic orchestration over ~150 tools built as Unity Catalog functions, across 47.5M+ production records with hybrid model routing.

Open-Source Ecosystem

95+ open-source releases across Hugging Face and Ollama (open-weight models, datasets, and interactive Spaces), with ~40,000 cumulative downloads and reproducible companion artifacts for most published papers.

SELECTED EVIDENCE

- Refusal-eval systematic review (PRISMA, 13 corpora) · with G. D. Moody, 2026
- TEMPEST adversarial replication · 10 frontier models, 97K queries
- 1,554-prompt code-safety benchmark · Fleiss' κ = 0.876
- CardioEmbed clinical embeddings · 99.6% retrieval, +15.94 pp over SOTA
- 256-LLM instruction adherence eval · largest single eval at release
- EQUITRIAGE gender-bias audit · LLM-based ED triage; NSF SBIR in prep
- ★ **Speaking** · Keynote, John Snow Labs Applied AI Healthcare (50,000+ data scientists) & Commencement, Concorde College, 2026

STACK

Python · PyTorch · Hugging Face · frontier APIs (Anthropic Claude, OpenAI, Qwen, Llama, Gemma) · Databricks (SQL Warehouse, Apps, Vector Search, Unity Catalog, Delta Lake) · Flask · Postgres+pgvector · ChromaDB · Spark · Snowflake · Azure / Azure OpenAI · Slurm HPC · MLX & GGUF quantization · drift monitoring · CI/CD.

TEACHING

Top-rated graduate professor (4.86/5, well above UNLV benchmarks) in business intelligence, big data, machine learning, and analytics at UNLV Lee Business School.

OPEN TO

Senior research / applied scientist roles at frontier AI labs, mission-driven research institutes, and foundations · **research professor or faculty-in-residence** roles in AI and translational data science · **advisory / scientific-board** roles in AI safety governance, clinical AI, and code-safety evaluation.

IN BRIEF

Ph.D. Computational Neuroscience, UNLV · **Senior AI Research Scientist**, UnitedHealth Group · **Co-Chair**, UHG IRB · **Assistant Professor-in-Residence**, UNLV Lee · **30+ publications** · **35+ invited talks** on healthcare & agentic AI (audiences up to 50,000+) · **patent inventor** (U.S. Appl. 18/323,518).