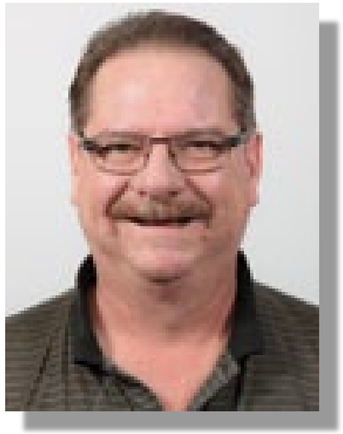
Health Informatics

Jayashree Kalpathy-Cramer, PhD
Director Health Informatics

cctsi.cuanschutz.edu

Leadership Team

Director of Health Informatics



Director of Research Informatics

Strategic Goals

1. **Goal:** Integrate clinical, biological, and public health data; deploy robust AI analytics environments
   - **Roadblocks addressed:** Enhance scientific rigor and reproducibility, Minimize infrastructure roadblocks for conducting CTR, Fragmented research workflows,
2. **Goal:** Improve accessibility and interoperability of regional and national data sharing
   - **Roadblocks addressed:** : Minimize infrastructure roadblocks for conducting CTR, Data silos, lack of standardization
3. **Goal:** Develop, validate, and deploy clinical decision support and trial execution tools
   - **Roadblocks addressed:** Enhance scientific rigor and reproducibility, optimize workflows, operational inefficiency
4. **Goal:** Expand hands-on translational informatics education across career stages
   - **Roadblocks addressed:** AI literacy gap, training accessibility

Year 3 Impact: The Problem — Fragmented Research (Goal 1)

- **The Problem:** Researchers context-switch 8–10 times per workflow across fragmented tools
- Separate logins and data silos for every tool: PubMed, Google Scholar, Zotero, PowerPoint, Word, ChatGPT
- Results trapped in disconnected platforms with no unified search or retrieval
- Email-based collaboration creates version control chaos
- No audit trail for AI-assisted research outputs

Year 3 Impact: The Solution — Apogee (Goal 1)

- **Apogee:** One unified portal for the complete research lifecycle
- **Privacy-First:** Cloud AI (openAI GPT-5 through institutional approved API) AND locally-hosted models for sensitive data
- **Multiple LLM Backends:** Ollama/LM Studio with many local models, DGX Spark(if available), API calls to commercial APIs endpoints, researcher/agent chooses per task
- Institutional deployment with HIPAA-compatible governance and audit trails
- **Modular Adoption:** Start with Daily Digest, expand to full platform
- **8+ Integrated Tools:** Daily Digest, Literature Search, Grant Review, Slide Generator, Summarizer, Reformatter, Writing Tools, Response to Reviewers

Year 3 Impact: Literature Discovery & Corpus

- **1M+ Papers Indexed** across 5 data sources: PubMed, arXiv, Semantic Scholar, OpenAlex, HuggingFace
- **Hybrid Search:** Vector + full-text with PubMed-style field tags (author[au], MeSH) and natural language
- **Embeddings:** vector embedding in Qdrant for semantic similarity and retrieval
- **Automated Ingestion:** Daily trending papers (6 AM), weekly bulk (Sundays), monthly journal backfill
- 7,723 papers added [ophthalmology and/or AI] in 2026 alone

Year 3 Impact: Daily Research Digest

- AI-curated daily briefing: early deployment in Ophthalmology AI division, next ophthalmology division, then CCTSI broadly
- **Multi-Source:** Pulls from PubMed, arXiv, Semantic Scholar, OpenAlex, HuggingFace overnight
- **Thematic Clusters:** Papers grouped into themes: e.g. Retinal AI, Foundation Models, Clinical NLP, Segmentation
- **AI Summaries:** LLM-generated editorial overviews for each theme, scan 50 papers in 2 minutes
- **Personalized:** Users customize topics; bookmarking behavior improves relevance over time

Year 3 Impact: Grant Peer Review

- **5-Stage Pipeline:** Upload → Reviewer → Devil's Advocate → Merger → Output
- Prompted with example grants and reviews
- NIH-style Specific Aims with Significance, Innovation, Approach assessments
- PubMed literature verification ensures claims are grounded in evidence
- RFA cross-referencing validates alignment with funding opportunity
- Outputs NIH-format scoresheet with consensus review
- Saves weeks of pre-submission preparation and internal review cycles
- Evaluation on internal and external grants (and publication) in progress

Year 3 Impact: Manuscript Preparation Tools

- **Journal Suggestion:** 92,000+ paper knowledge base with RAG-enhanced semantic matching to find best-fit journals (for ophthalmology)
- **Reformatter:** Validates against 15+ journal specs (word limits, sections, citations); outputs formatted DOCX or LaTeX
- **Writing Tools:** Cover letter generator, abstract writer, lay summary generator, all journal-aware
- **Humanizer:** Converts language to researcher's style based on previous publications
- **Response to Reviewers:** AI-drafted point-by-point responses with evidence citations from the corpus

Year 3 Impact: AI Training Curriculum (Goal 4)

- **5 Hands-On Modules:** workshops
- 01: NotebookLM for Literature Grounding & Synthesis
- 02: Practical Copilot Use Cases for Research & Administration
- 03: Parsing & Structuring Information from PDFs
- 04: Introduction to Automation
- 05: Introduction to Apogee
- All career stages: students, postdocs, faculty, staff — no coding prerequisite

Year 3 Impact: Training Philosophy & Deep Learning

- **AI-Fluent, Not AI-Dependent:** Teach researchers to evaluate, direct, and verify AI outputs
- **Bridge the Gap:** From “AI exists” to “I use AI daily in my research”, practical skills, not theory
- **National Leadership:** REDCap Consortium chairs for Training and Security — setting national standards
- **Deep Learning for Medical Imaging:** 3-session course (Foundations, Architectures, Applications)
  - Complete package: interactive notebooks, exam questions, Manim animations — open educational resource on GitHub

Year 3 Impact: Data Sharing & Interoperability (Goal 2)

- **99% Conditions Mapped** to standard terminologies
- **97% Labs Mapped** to standard terminologies
- **96% Medications Mapped** to standard terminologies
- **78% Procedures Mapped** to standard terminologies
- **ISO 27001 certification achieved** — international data security compliance

Year 3 Impact: Clinical Tools & Infrastructure (Goals 1 & 3)

- **Infrastructure Resilience:** Veeam backup/recovery platform; improved security, compliance, and operational maturity
- **Imaging Systems:** Upgraded ultrasound + GE ViewPoint; enhanced data security, RBAC, and ISCV integration
- **REDCap Automation:** Service account model for secure, API-driven study workflows
- **REDCap Knowledge base + Mosio SMS** for Children's Hospital

Year 4 Plans

- **2026 Apogee:** Expand to 3+ departments. Add hybrid semantic search. User feedback loop.
- **2026 Evaluate Apogee tools and AI training**
  - Identify additional features for Apogee
  - plan courses for 2027
- **REDCap Consortium chairs for Training and Security** — setting national standards for research informatics

Year 4 Joint Plans for Health Informatics and BERD

- Joint education pipelines in AI combining the models and tool applications.
- Having CIDA and BERD members pilot AI in research tools for Health Informatics.
- Working together on a process for jointly deciding what data sources can or should CCTSI support with BERD and Health Informatics Resources.

Response to EAC Critiques

Previous year's EAC critique: Concerns were raised about the momentum of the BERD and Informatics partnership with the leadership change in Health Data Compass.

- **We are both on new Enterprise Committee BIDS (Biostatistics, Biomedical Informatics, and Data Science). Nichole is co-lead.**
- **We have met with Health Data Compass to understand the path forward. Considering a Unit Dedicated Resource for CCTSI, but resources are involved.**
- **Shared education programming and piloting tools together.**

Questions for Discussion

Where should we invest in custom, research-specific AI tools versus leveraging commercial platforms and how do we avoid building one-off solutions that don't scale?

How should we measure if our AI efforts are translational?

As AI accelerates certain parts of the research pipeline, where are you seeing new bottlenecks emerge? Is informatics the right lever to address them?