AI-Powered Medical Record Analysis: Faster and More Accurate

By David Zimmerman
Director of AI Engineering
Empower AI

AI streamlines medical record review by quickly extracting and analyzing key data, saving time and effort for healthcare staff.

Federal healthcare systems manage staggering volumes of medical records—sometimes spanning tens of thousands of pages per case. Analysts, auditors, and case managers often spend hours searching for key information scattered throughout these documents. It's a model that doesn't fit well with today's needs for high productivity, and artificial intelligence (AI) provides a solution.

AI streamlines this labor-intensive process, delivering both efficiency and precision for government and healthcare partners. It is in use today and is showing promising results.

The Challenge of Navigating Complex Medical Records

Medical records are notoriously unstructured and variable. They are often scanned images riddled with OCR (Optical Character Recognition) errors, and each file can vary widely in format and length. In some cases, a single medical record can exceed 20,000 pages.

The challenge becomes more complicated in secure, air-gapped environments, where regulations prohibit internet-connected services. Add to this the need for high-throughput analysis, and traditional search tools quickly fall short. To meet these stringent requirements, Empower AI developed a Retrieval-Augmented Generation (RAG) accelerator that brings AI to the challenge. Here's how the on-premise deployment works and why it matters.

Key Capabilities

- **LLM-Driven Embeddings and Metadata Storage.** The system uses a large language model to generate dense vector embeddings enriched with structural and contextual metadata. This enables smarter, faster content retrieval across complex record layouts.
- **Multi-Vector, Hybrid Search.** Combining dense, sparse, and full-text search techniques, the solution can pinpoint content even in inconsistent or fragmented documents.
- **LLM Quality and Relevance Filters.** An additional layer of machine learning assesses the relevance of search results, filtering out noise and improving signal accuracy.
- **Re-Ranking with Prioritization Algorithms.** A second LLM model re-ranks search outputs based on a weighted scoring system, ensuring the most accurate and context-rich results rise to the top.
- **Locally Hosted, Containerized Deployment.** The entire system is designed to operate securely within government environments, requiring no internet connectivity and complying with privacy and data sovereignty rules.

Real-World Results

In evaluation against a labeled test dataset, Empower AI's accelerator delivered impressive results:

- 13 different medical record sections were used as benchmarks
- The system achieved ~90% accuracy in returning the desired content across test cases
- Analysts were able to retrieve actionable information in seconds rather than hours

Benefits for Federal Healthcare and Beyond

With the accelerator concept, Empower AI takes specific use cases like medical records reconciliation and applies the AI that results in a purpose-built solution. This brings transformative value to agencies managing complex, document-heavy workflows. It drives operational efficiency by dramatically reducing manual review time, freeing staff to focus on higher-value tasks such as case decisions and policy compliance. At the same time, it delivers accuracy at scale, helping teams to surface critical information that might otherwise be lost in thousands of pages, especially in legacy or low-quality scanned documents where OCR limitations often cause gaps.

Built with a security-first mindset, accelerator-based solutions are designed specifically for federal use, operating in air-gapped environments and aligned to the government's strictest security and compliance requirements. Finally, its scalable, reusable AI framework extends well beyond medical records to enable high-impact automation across legal discovery, claims processing, audit trails, and other high-volume, unstructured data challenges.

A Strategic Advantage for Agencies Managing Complex Health Data

Empower AI's accelerator approach represents a powerful new capability in the federal AI toolkit. By making medical record analysis faster, more accurate, and fully secure, it offers a mission-aligned, future-proof solution to one of the government's most stubborn data challenges. For agencies tasked with healthcare, benefits administration, or records management, this AI accelerator delivers a measurable leap forward in productivity and insight.

About the Author

David Zimmerman is an accomplished technology leader with a strong background in machine learning, artificial intelligence, and engineering management. Serving as director of AI engineering, he drives advanced AI initiatives and oversees high-impact ML engineering teams. David began his career as an artificial intelligence engineer, developing intelligent systems and contributing to foundational AI projects.