

Case Study: Building a **Reliable** Data Foundation for AI

The Opportunity

Healthcare organizations are racing to implement predictive models and large language models (LLMs). Yet, many quickly encounter a foundational issue: their underlying clinical data is inconsistent, fragmented, and often unreliable. Incomplete CCDs, cut-and-paste progress notes, and messy PDFs make it difficult for downstream systems to trust or act on the data.

This client, one of the largest health technology and solutions providers in the U.S., saw the chance to fix the root cause—establishing an enterprise-scale clinical data foundation that would support AI, risk modeling, analytics, and better patient care long into the future.

Commercial Impact

By deploying emtelligent's Medical Language Engine, the organization built a scalable pipeline that structured unstructured clinical data and delivered clean, AI-ready input to downstream systems. This improved analytics, model accuracy, and enterprise reporting. Scale was essential: the client manages 285M+ patient records. To cut backlog and accelerate actuarial analysis, emtelligent bulk-processed 25M records in two weeks, moving to a steady-state of ~50K records daily. Key benefits included:



AI-ready data
foundation



Improved model and
analytics performance



Reduced costly rework
from poor input data



Better quality and
compliance reporting

The Approach

emtelligent collaborated with the client to implement a modular, upstream-first approach to clinical data management. Key steps included:

- Parsing CCDs and other clinical documents to expose and fix inconsistencies early.
- Applying NLP to extract and structure data from unstructured clinical notes.
- Integrating structured data outputs into a persistent enterprise data mart to power longitudinal views, analytics, and AI models.



The Result

The implementation established a trusted clinical data foundation that transformed fragmented documentation into governed, AI-ready datasets.

With structured inputs flowing into downstream systems, the client was able to realize the following benefits:

Reduced Rework: Automated classification and indexing minimized manual effort and error correction.

Model Reliability: Cleaner, structured data improved the consistency and accuracy of downstream models.

Stronger Compliance Reporting: Comprehensive, auditable data flows simplify regulatory submissions.

Confident Decision-Making: Clinical and business leaders act on trustworthy insights with reduced uncertainty.



Contact Us

Kim Perry
Chief Growth Officer

 emtelligent.com

 312.307.8402

 kim@emtelligent.com



About emtelligent

Based in Vancouver, British Columbia, emtelligent helps healthcare professionals work more efficiently and improve patient care.

Built by medical experts for medical experts, its Medical Language Engine and suite of clinical and search tools automate the extraction and analysis of data across systems, specialties, and populations, turning complex, unstructured medical text into actionable insights.

Through partnerships with health networks, imaging facilities, research institutions, payer organizations, and technology innovators, emtelligent increases safety, boosts operational efficiency, and elevates care quality across North America.

 emtelligent.com

 877.GO.EMTEL

 info@emtelligent.com