

Connecting EHRs and LLMs: Improving Clinical Decision Making with PHI Masking for Agentic AI Implementation



Bradley Joseph Sweeney

PhD Candidate in Biomedical Engineering, Summer 2026

AI in Digital Health Certification, Spring 2026

Ohio State University

05.11.2026

INTRODUCTION

Healthcare systems generate substantial amounts of patient data every day through every interaction between providers and patients. This data gets captured into Electronic Health Record (EHR) systems in both structured fields, like lab results and diagnoses, and unstructured clinical documentation known as free text. Structured data provides easily searchable and analyzable records, but a large proportion of clinically relevant information exists within these free-text narratives. It has been estimated that unstructured data comprises up to 70-80% of healthcare data, generated in forms such as physician notes, operative reports, discharge summaries, pathology reports, radiology interpretations, and more. All this information comprises the entire health profile of patients; providers must synthesize each piece to create comprehensive treatment strategies.

The reliance on free-text documentation provides a significant challenge for providers. Manual review of patient records is time-intensive and open to subjective variability. As health systems continue to expand, the size and complexity of this information continue to increase posing even greater demands and challenges on health professionals. Advancements in machine learning and natural language processing (NLP) provide the opportunity for systems to be developed capable of analyzing and classifying unstructured data to increase information retrieval and workflow efficiency.

While tools exist for AI to review and summarize unstructured EHR data, there has remained a large risk of maintaining Health Insurance Portability and Accountability Act (HIPAA) compliance with these tools due to prompt injection. Prompt injection involves overriding instructions to large language models (LLMs) through creating input text that models a prompt.¹ This injection causes the LLM to ignore initial instructions of the user to process unauthorized and malicious commands. This means that healthcare systems require an approach to protect data before utilizing AI tools to prevent malicious actors from operating patient information.

Vertebrai developed and implemented a novel HIPAA compliant model architecture to address these challenges enabling health care providers a framework for further AI tool integration and development. Vertebrai utilizes an Agentic AI / Model Context Protocol (MCP) approach enabling the LLM to access the specific patient clinical electronic health record (EHR), redacts structured PHI elements and reviews unstructured text using a Named Entity Recognition (NER) model to identify and mask protected health information (PHI) and personally identifiable information (PII) before it reaches the LLM. This approach enables clinicians to seamlessly integrate their electronic health records (EHR) with AI tools capable of accessing and evaluating the full patient record and providing clinical decision support, link with medical guidelines, treatment history evaluation, and protocol recommendations. Vertebrai enables the LLM to rapidly scan the relevant patient data in the EHR, respond to clinician prompts to deliver summaries of the patient history, identification of historically recorded conditions, chart and compare lab results over time, highlight critical health, Social Determinants of Health (SDOH), personalized care recommendations, while using Retrieval-Augmented Generation (RAG) to link to specialty and organizational protocols, published research studies located in PubMed and other specialty guidelines.

Enabling the LLM to securely access the EHR, while redacting PHI, reduces the clinical documentation burden, increases speed for clinician decision making, improves care quality and outcomes, reduces the cost of care, and reduces provider strain.

KEY PROCESSES

Architecture Diagram

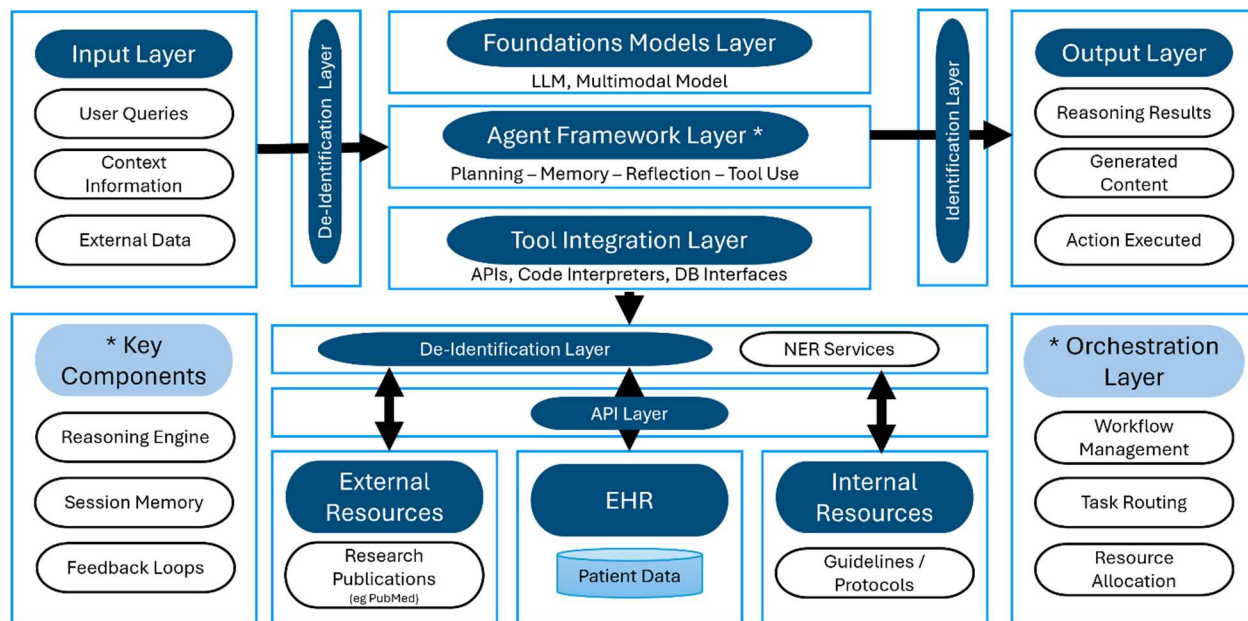


Figure 1: Vertebral Framework. Details of the key components of the Vertebral architecture

Standard Processing flow

Clinician enters a prompt to query the Vertebral Clinical Decision Assistant. Prompt text is analyzed by the Named Entity Recognition (NER) using Natural Language Processing (NLP). The NLP scans text to identify and classify key information into predefined categories. It acts as a foundational AI tool for organizing unstructured data, Protected Health Information (PHI) recognized entities. Dates are identified, and a random shift for the session is established to shift dates. Entered text is highlighted when the NER model identifies PHI prior to the user entering “send”. De-identified (and date-shifted) message is passed to the Tool Orchestrator agent. PHI is redacted with a masked value which is established and exists for the session.

The Tool Orchestrator agent accesses the list of available tools, retrieves data from the EHR, external resources, and internal resources, evaluates the clinician prompt, determines which tools return relevant information, and accesses them in parallel. Internal resource retrieval is supported through a Retrieval-Augmented Generation (RAG) pipeline, where client-provided documents are indexed within Azure Cognitive Search and vectorized using text embedding models to enable semantic search and contextual retrieval during execution. The Tool Orchestrator can dynamically generate search queries to retrieve relevant medical reference material and client-specific documentation from these indexes. During the session, conversational context and retrieved information are retained within secure session memory, enabling the LLM to leverage prior prompts, responses, and retrieved documents to support subsequent clinician interactions. The Tool Orchestrator has secure, authenticated access to the EHR following FHIR R4+ standards with OAuth 2.0 authentication. Prior to response delivery, redacted entities are securely re-identified within the session context to present properly identified patient information back to the clinician.

The de-identified clinician prompt, prior session history, and de-identified tool responses retrieved from the EHR and supporting systems are passed to the Response Creation Agent. The Response Creation Agent generates a response using the retrieved data and session context. If insufficient information is available, the workflow returns to the Tool Orchestrator Agent to retrieve additional information before continuing response generation. As the response is generated, it is streamed to the frontend application. The front end retrieves the session-specific mapping table from the backend and re-identifies redacted entities in real time for display to the authorized clinician.

All tool calls are logged as audit events, and the exact data retrieved, along with the de-identified data handed off to downstream agents, is fully auditable. All clinician prompts (both identified and de-identified) and all AI-generated responses (both de-identified and re-identified) are securely stored for auditing and compliance purposes. End-to-end agent activity during orchestration, including inter-agent message passing, retrieval operations, and LLM processing steps, is logged to support traceability, monitoring, and post-session review.

EXPECTED OUTCOMES

Vertebrai's design provides the necessary architecture for improving scalability and efficiency of clinical information review. The workflow seamlessly integrates provider's EHR data with access to external medical resources (e.g. PubMed), free-text collection and summarization, and external tools while protecting PHI and PII through model masking important data prior to reaching the LLM. This implementation leads to increased physician efficiency, increased diagnostic accuracy, increased treatment efficacy, reduced readmissions, and reduced total cost.

Physician Efficiency and Time Savings

Vertebrai reduces the clinician chart review time by up to 90%. Analysis of large volumes of free-text data from EHR data requires substantial time from clinicians, leading to slow processing and inaccuracies. Vertebrai's NLP capabilities significantly decrease the processing time while maintaining equivalent or even superior accuracy to summaries from medical experts. LLM's have the capability to quickly scan and process large amounts of unstructured data. In identification of immune related events, LLM processing time was recorded at 9.53 sec/chart in comparison with an estimated 15 min/chart for manual adjudication.² This level of time reduction allows clinicians the opportunity to gain both the time to interpret and determine more individualized treatment strategies while also providing the potential to visit more patients in a single day. In primary care, testing an Artificial Intelligence-based Clinical Decision Support System, the immediateness of model chart synthesis reduced clinician's median time for EHR review by 7.8 min.³ Vertebrai's immediateness to chart review and synthesis of unstructured data reduces the time needed for treatment identification and increase the quality time physician's spend with patients.

Increases diagnostic accuracy

Clinicians using Vertebrai experience up to a 30% reduction in diagnostic errors. AI tools display the capability to increase diagnostic accuracy over trained physicians, and empowering clinicians to gain access to these tools in a HIPAA compliant manner leads to the reduction of clinical errors. Vertebrai provides a model that allows clinicians to make final decisions while leveraging tools that have shown improved ability over trained physicians at diagnostic accuracy. LLMs have shown increased sensitivity over ICD codes for detection of immune-related adverse events.² In tests with AI combining structured and free-text EHR data, the model was capable of identifying almost 4 million additional true cases of NVAf showing the power of connecting the entire EHR into clinical models.⁴ In summarizing clinical notes, LLMs rated equivalent or superior to expert summaries in over 80% of cases, displaying capacity for improving EHR interpretation.⁵ The integration of EHR models to LLM with Vertebrai provides healthcare groups and providers superior accuracy in diagnosis leading to improved patient outcomes and satisfaction.

Improves clinical detection of underlying risk

Vertebrai increases detection of missed EHR social risk factors, SDOHs, up to 40x more than analysis from ICD codes and structured data alone. Integrating AI models provide clinicians with the ability to identify high risk patients and disease potential earlier, allowing for earlier intervention. Vertebrai synthesizes important data for clinician review, finding trends in health data and key points from unstructured data that were previously overlooked. In NLP-enhanced identification of NVAf, high risk patients could be identified sooner which was estimated to lead to prevention of over 176,000 strokes and 10,000 deaths annually.⁴ AI models were able to capture poor social support from physician notes at a much higher rate than previously documented, reflecting a prevalence of 16.0% compared to a previous 0.4%.⁶ Physicians provide countless pieces of patient health data that are not able to be represented by simple ICD codes and quantified data; Vertebrai captures the entire scope of patient history and lifestyle leading to earlier clinical decisions and interventions.

Reduces Readmissions

Utilizing Vertebrai, healthcare systems experience a 25% reduction in readmissions. Identifying patients with the highest risk of readmission remains a constant goal and strategy of health professionals to reduce the burden on the health system. Vertebrai provides clinicians with access to predictive modeling tools and synthesis of free-text metrics not captured in quantitative tests to increase their ability to determine patient centered interventions that reduce risk. It has been shown that artificial intelligence tools assessing hospitalizations decreased readmission rates from 11.4% to 8.1% compared to only a 0.5%-point decrease in control hospitals.⁷ Vertebrai allows for this integration between the EHR data and artificial intelligence tools to extend across entire networks of hospitals to reduce the total readmission rate.

Increases Recapture of Annual Reporting Gaps

Vertebrai by surfacing previously documented but uncoded conditions for clinician validation, enables recapture of a substantial portion of missed diagnoses, improving Hierarchical Condition Category (HCC) recapture rates by up to 20–40% in environments where documentation gaps are common. A substantial number of chronic conditions are incompletely documented within structured fields and result in missed diagnoses and failure for annual HCC recapture. The Medicare Payment Advisory Commission’s June 2023 report highlighted that many chronic conditions fail to reappear in documentation cycles, citing the need to track these more accurately for patient risk adjustment and reimbursement accuracy.⁸ Studies in Medicare Advantage Populations show that additional diagnoses are frequently identified through targeted documentation processes such as Health Risk Assessments, demonstrating that clinically present conditions are often absent from claims-based coding workflows.⁹ Commercially available AI-powered tools show that HCC recapture rates can be improved by 24% alone.¹⁰ Through utilizing Vertebrai, healthcare systems correctly captures and carries over key HCC diagnoses and metrics for patients each year.

Cost reduction

Modeled analyses suggest that implementation of the Vertebrai platform can provide meaningful operational and financial benefits. These benefits include but are not limited to improvements in clinical workflow efficiency, reduction in manual chart review burden, earlier identification of high-risk patients, and enhanced clinical decision support. With modeled deployment, estimated value generation could approach \$1000 per hospital admission depending on the size of the institution, patient population, and implementation and utilization patterns. Vertebrai’s modeled assumptions include delivering a 30% reduction in high-risk admissions, 10-20% reduction in ICU admissions, a 0.5-day reduction in average length of stay, 219 hours saved per provider per year, with the average cost of inpatient days at \$2500. In total, these estimates show that Vertebrai can provide 50-bed hospitals with \$3.0 million in annual operational value; 100-bed hospitals can realize \$6.0 million annually, and small outpatient clinics may realize approximately \$300K annually. Vertebrai accomplishes these savings through the reduction of provider documentation and chart review time, improving access and synthesis of both structured and unstructured clinically relevant patient information, earlier detection of high-risk patients, and improved workflow efficiency.

CONCLUSION

Vertebrai provides the core tools for the LLM to connect EHR data, in a HIPAA compliant manner, access to external and internal medical references, capable of increasing physician efficiency, increasing diagnostic accuracy, improving clinical detection of underlying risk, reducing readmissions, increasing recapture of annual reporting gaps, and reducing the total cost on health systems. Through masking PHI/PII data, clinicians safely navigate utilization of LLM tools that provide faster information retrieval, improving physician efficiency, and assisting in clinical decision making. This novel approach to healthcare provides health care systems the next step in transforming the medical environment and patient care.

Vertebrai's functionality through its platform design, supporting literature, and preliminary modeling, clinicians should anticipate significant improvement in time savings, patient outcomes, and cost reductions. Prospective validation in real-world clinical environments is necessary for further exact quantification of clinical performance, workflow impact, and financial return. Overall, Vertebrai offers a scalable framework for transforming the healthcare environment through connecting EHR systems with modern processing tools, turning unstructured EHR data into actionable clinical insights in a way that improves both care delivery and health-system efficiency.

BIBLIOGRAPHY

1. Pathade C. Red Teaming the Mind of the Machine: A Systematic Evaluation of Prompt Injection and Jailbreak Vulnerabilities in LLMs. Accessed May 12, 2026. <https://arxiv.org/html/2505.04806v1>
2. Sun VH, Heemelaar JC, Hadzic I, et al. Enhancing Precision in Detecting Severe Immune-Related Adverse Events: Comparative Analysis of Large Language Models and International Classification of Disease Codes in Patient Records. *J Clin Oncol Off J Am Soc Clin Oncol*. 2024;42(35):4134-4144. doi:10.1200/JCO.24.00326
3. Gomez-Cabello CA, Borna S, Pressman S, Haider SA, Haider CR, Forte AJ. Artificial-Intelligence-Based Clinical Decision Support Systems in Primary Care: A Scoping Review of Current Clinical Implementations. *Eur J Investig Health Psychol Educ*. 2024;14(3):685-698. doi:10.3390/ejihpe14030045
4. Elkin PL, Mullin S, Mardekian J, et al. Using Artificial Intelligence With Natural Language Processing to Combine Electronic Health Record's Structured and Free Text Data to Identify Nonvalvular Atrial Fibrillation to Decrease Strokes and Death: Evaluation and Case-Control Study. *J Med Internet Res*. 2021;23(11):e28946. doi:10.2196/28946
5. Van Veen D, Van Uden C, Blankemeier L, et al. Adapted large language models can outperform medical experts in clinical text summarization. *Nat Med*. 2024;30(4):1134-1142. doi:10.1038/s41591-024-02855-5
6. Navathe AS, Zhong F, Lei VJ, et al. Hospital Readmission and Social Risk Factors Identified from Physician Notes. *Health Serv Res*. 2018;53(2):1110-1136. doi:10.1111/1475-6773.12670
7. Romero-Brufau S, Wyatt KD, Boyum P, Mickelson M, Moore M, Cognetta-Rieke C. Implementation of Artificial Intelligence-Based Clinical Decision Support to Reduce Hospital Readmissions at a Regional Hospital. *Appl Clin Inform*. 2020;11(4):570-577. doi:10.1055/s-0040-1715827
8. June 2023 Report to the Congress: Medicare and the Health Care Delivery System – MedPAC. Accessed May 10, 2026. <https://www.medpac.gov/document/june-2023-report-to-the-congress-medicare-and-the-health-care-delivery-system/>
9. James HO, Dana BA, Rahman M, et al. Medicare Advantage Health Risk Assessments Contribute Up To \$12 Billion Per Year To Risk-Adjusted Payments. *Health Aff (Millwood)*. 2024;43(5):614-622. doi:10.1377/hlthaff.2023.00787
10. AI-Powered Risk Adjustment Built for Accurate RAF Scoring and Reimbursement. Accessed May 11, 2026. <https://www.rapidclaims.ai/services/ai-powered-risk-adjustment-built-for-accurate-raf-scoring-and-reimbursement>