White Paper for Clinical Healthcare AI

Use Case for Healthcare: EHR De-Identification & FHIR Standardization

Healthcare organizations must frequently analyze, exchange, or operationalize electronic health record (EHR) data. However, EHR data typically arrives in inconsistent formats, often containing sensitive Protected Health Information (PHI) or Personally Identifiable Information (PII). Manually cleaning, de-identifying, and standardizing this data is time-consuming, error-prone, and risky.

Overview

This white paper outlines how a two-agent SUPERWISE® system automates HIPAA-compliant de-identification and transforms raw EHR datasets into fully standardized FHIR (Fast Healthcare Interoperability Resources) bundles. The workflow uses synthetic data for demonstration, but mirrors real-world challenges in safe, governed health data processing.

By separating compliance logic (Agent 1) from data transformation logic (Agent 2), this system provides a safe, auditable, and scalable solution for healthcare organizations seeking to modernize their data pipelines.

Introduction

Healthcare interoperability and data analytics depend on high-quality, standardized, and privacy-protected data. Yet organizations face several challenges:

Inconsistent and messy EHR data

Different providers and systems produce varying schemas, naming conventions, and formats.

High privacy risk

EHR datasets often contain identifiers such as:

- Names
- Addresses

- Contact information
- Birthdates
- Social security numbers
- Full date of service events

All of which are regulated under HIPAA.

Manual de-identification is slow and unreliable

Human review introduces risk of oversight, inconsistent data handling, and operational bottlenecks.

Lack of standardization limits interoperability

Non-FHIR data cannot be efficiently shared, integrated, or used for downstream analytics or AI pipelines.

The **SUPERWISE®** agentic approach solves these challenges by automating schema-level compliance checks and safe data transformation using governed, explainable, multi-agent intelligence.

The Solution

The EHR De-Identification & FHIR Standardization system consists of two specialized SUPERWISE® agents:

Agent 1 — Schema Identifying Agent

Responsible for PHI/PII detection and schema-level HIPAA compliance.

Agent 2 — Schema Standardization Agent

Responsible for transforming clean, safe data into FHIR-compliant JSON bundles.

This separation of responsibilities ensures:

- Safety
- Auditability
- Explainability
- Compliance
- Reduced model risk
- Role-based agent logic in line with modern AI governance frameworks

Synthetic (non-real) EHR data is used in this demo to illustrate real-world workflows without exposing any sensitive patient information.

SUPERWISE® Two-Agent Workflow

Step 1 — PHI/PII Detection & Schema Cleaning

Agent 1: Schema Identifying Agent

Agent 1 analyzes only the schema of the incoming EHR dataset—field names and structures—not the patient values.

This significantly reduces risk and aligns with HIPAA Safe Harbor rules.

Key Responsibilities:

- Detect fields containing PHI/PII indicators (e.g., name, address, SSN, phone number, MRN).
- Flag fields for removal or transformation.
- Identify date fields requiring year-only masking (HIPAA-compliant de-identification).
- Produce a field-level compliance map marking:
  - Safe fields
  - Unsafe fields
  - Fields needing partial transformation (e.g., date → year)

Output:

A sanitized, enforceable schema blueprint that guarantees no unsafe elements reach the second agent.

Step 2 — De-Identification & FHIR Standardization

Agent 2: Schema Standardization Agent

Using the safe schema map from Agent 1, Agent 2 transforms the cleaned dataset into standardized FHIR resources, such as:

- Patient
- Encounter
- Condition
- Medication Statement
- Observation (lab results, vitals, measurements)

Key Responsibilities:

- Convert filtered fields into FHIR-compliant JSON structures
- Normalize clinical events into structured resource types
- Convert dates to year-only format
- Remove any remaining sensitive fields
- Package all resources into a FHIR Bundle

Output:

A fully compliant, interoperable FHIR dataset ready for secure exchange, analytics, or downstream AI workflows.

Why This Matters

The shift toward interoperability, analytics, and AI in healthcare requires:

Reliable, de-identified, shareable data

Healthcare institutions cannot risk PHI exposure, regulatory fines, or audit failures.

FHIR standardization for interoperability

FHIR is the dominant data format across EHR vendors, payers, research networks, and digital health platforms.

Scalable, automated pipelines

Manual processes cannot keep up with the volume and complexity of modern healthcare data.

The **SUPERWISE®** two-agent framework ensures:

- HIPAA Safe Harbor compliance
- Consistent de-identification practices
- Standardized FHIR formatting
- End-to-end governance and observability

This enables healthcare organizations to confidently operationalize EHR data at scale.

Business Impact

Healthcare Providers

Accelerate data sharing for care coordination and population health analytics.

Payers & Insurance Organizations

Transform clinical and claims data for analytics and risk modeling.

Digital Health Companies

Automate ingestion of messy EHR data into clean, FHIR-based systems.

Healthcare Researchers

Gain access to de-identified datasets suitable for studies and AI model development.

AI/ML Teams

Construct reliable training datasets free of PHI and aligned with standard clinical ontologies.

Governance & Responsible AI with SUPERWISE®

SUPERWISE® provides:

Guardrails - Prevent unsafe operations and PHI exposure.

Policies - Define transformation rules, date-masking logic, and schema restrictions.

Observability - Monitor agent behaviors, policy violations, and pipeline performance.

Auditability - Track every transformation step for full regulatory compliance.

Multi-Agent Intelligence - Role-based agents ensure separation of duties and safer, more reliable AI workflows.

Conclusion

The **SUPERWISE®** EHR De-Identification & FHIR Conversion use case demonstrates a powerful, safe, and explainable approach to health data processing. By dividing compliance and transformation responsibilities across two agents, organizations gain:

- Automated HIPAA-compliant de-identification
- FHIR-standardized, interoperable datasets
- Reduced operational burden
- Scalable, governed AI pipelines

This workflow showcases the future of healthcare data preparation—automated, safe, interoperable, and powered by enterprise-grade agentic governance.