Executive Summary

The U.S. and Canadian healthcare systems are under growing pressure to achieve seamless data interoperability — not just for compliance, but to enable better patient outcomes, timely public health reporting, and system-wide efficiency.

In response to this challenge, I developed a lightweight ETL (Extract, Transform, Load) pipeline that converts legacy healthcare data formats (like HL7v2 , CSV, flat files, Restful/SQL queries) into modern FHIR R4 resources. This article outlines the thinking behind that strategy, the technology choices I made, and why this matters for healthcare organizations of all sizes.

The Problem: Legacy Data Blocking Progress

Many healthcare providers still operate with:

- Siloed EHR systems
- Inconsistent coding (CVX, SNOMED, LOINC)
- HL7v2 feeds that are rigid, outdated, or poorly documented
- Limited patient matching logic and no validation layer

These realities create major barriers to adopting FHIR — the modern standard for health data interoperability.

My Strategy: A Lightweight, Standards-Based ETL Pipeline

I built a modular ETL pipeline that:

- Ingests HL7v2 , CSV messages . Flat files, SQL and Restful query results from source systems
- Normalizes coding using CVX, MVX, LOINC, and SNOMED
- Transforms inputs into FHIR R4 resources, particularly Immunization, Patient, Encounters, Observations and others
- Validates against structural and value set requirements
- Outputs JSON ready for use with FHIR servers or analytics platforms

Why It Works: Built for the Real World

1 Standards-compliant (HL7 v2.x, FHIR R4, CVX, LOINC) 2. Python-based, minimal dependencies 3. Handles low-resource environments (public health departments,

clinics, labs) 4. Includes validation, logging, and dashboards for monitoring data quality
5. Production-aware: Built to solve current integration pain without overengineering

---

Why This Matters

FHIR transformation isn't just a tech upgrade — it's an enabler of:

- Timely immunization reporting
- Public health surveillance
- Clinical decision support
- Compliance with U.S. ONC and CMS mandates

This ETL pipeline demonstrates how legacy healthcare infrastructure can be modernized in an incremental, practical way — bridging the gap without breaking what already works.