Adoption and use of common standards for Scotland’s data strategy

Solmaz Eradat Oskoui

5th April 2024

Overview

1	Data Harmonise Ltd
2	Interoperability & Information Standards
3	Scotland’s Health Data Sharing Landscape
4	OMOP CDM Overview
5	PoC for Data Harmonisation Strategy

Professional

Applied Mathematics, PhD

St.Andrews University

2014

Consultant/Senior Data Analyst/Data Engineer

Academia/IT/Financial Services/Health Informatics

2014 - 2023

Founder

Data Harmonise Ltd

2024

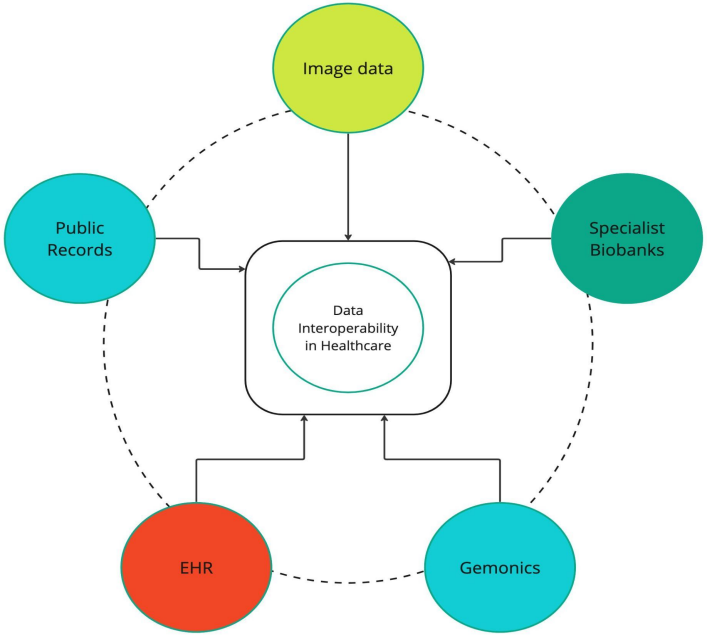
Company Mission

*Advance **interoperability** and **quality** within healthcare data by providing **data standardisation** and **engineering solutions** leveraging the Observational Medical Outcomes Partnership Common Data Model (**OMOP-CDM**). Our commitment is to facilitate data links, foster collaboration, accelerate research outputs, and drive innovation.*

Interoperability & Information Standards

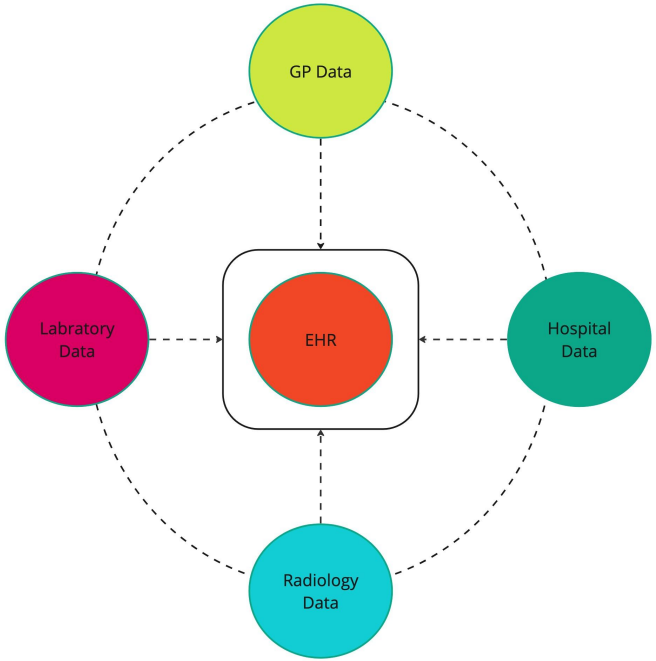
What is data Interoperability?

- Healthcare data can vary greatly both in structure and terminologies from one organisation to the next.
- Data are collected for different purposes, stored in different formats using different database systems and information models.
- To improve the interoperability of health data requires standardising both the structure (**syntactic interoperability**) and the terminologies (**semantic interoperability**).



Take EHRs for example ...

- **Centralised Patient Data:** Electronic Health Records (EHRs) store a patient's comprehensive medical history in one place, including demographics, medical history, clinical notes, radiology data, laboratory results, and more.
- **Information Exchange:** EHRs can receive and exchange patient data with various healthcare providers: GPs, hospitals, radiology departments, and laboratories.
- **Standardised Communication:** To ensure seamless communication and data exchange between different healthcare systems (often from different vendors), there is a need for common standards!



Focus of each Standard

- **DICOM (Digital Imaging and Communications in Medicine):** A standard focusing on data exchange and data persistence (structure, storage, maintenance) for medical images (X-rays, MRIs, CT, ultrasound, etc.).
- **HL7 FHIR (Fast Healthcare Interoperability Resources):** A standard focusing on data **exchange** and data persistence (structure, storage, maintenance) of EHRs in real time between healthcare systems.
- **openEHR (open Electronic Health Record):** A standard focusing on **data persistence** (structure, storage, maintenance) and data exchange for EHRs.
- **PRSB (Patient Reported Symptoms Burden):** A standard focusing on **capturing and coding** patient-reported symptoms and their impact on daily life, typically gathered through standardised questionnaires or surveys.
- **OMOP CDM (Observational Medical Outcomes Partnership Common Data Model):** A standard focusing on structure and terminology to facilitated the integration of data from different EHR systems for the purpose of large-scale research studies.

Comparison between standards

Feature	DICOM	FHIR	openEHR	PRSB	OMOP CDM
Primary focus	Data exchange for medical imaging	EHR data exchange	EHR data persistence	Quantify patient symptom and experience	Facilitation of large-scale research studies between EHR systems
Also does	Data persistence for medical imaging	Data persistence	EHR data exchange	Supporting research	Common language for healthcare
Strengths	Comprehensive, data consistency	Flexibility, adaptability	Comprehensive, data consistency	Improve patient care and treatment	Comprehensive, data consistency, adaptability
Use Cases	Integration with EHR	Point-of-care applications, data integration	Building EHR systems, long-term data storage	Clinical practise/clinical trials, public health	Building research data warehouses for fast research and insight
Technical Expertise	**Tags** implementation guides	**Resources** implementation guides	**Archetypes** implementation guides	May require with complex data interpretation	**CDM tables** & **Vocabularies** implementation guides
Data Complexity	Large file size with comprehensive meta data files	Flexible, may require additional effort for complex data	Built-in support for complex clinical data	Qualitative elements (eg, descriptive responses from patients about their experience with symptoms)	Complexity dealing with data heterogeneity, multimodality, transformation, quality, and missing data coming from different EHR systems

DICOM works a little differently with OMOP compared to FHIR, openEHR, and PRSB.

How do common data models like DICOM, FHIR, openEHR & PRSB work with OMOP for research?

1. FHIR, openEHR & PRSB:

These three standards work collaboratively to **provide data for OMOP research**:

- **FHIR (Fast Healthcare Interoperability Resources):** Enables real-time clinical data exchange between healthcare systems. Data from FHIR can be used to populate openEHR-based Electronic Health Records (EHRs).
- **openEHR (open Electronic Health Records):** Provides a standardised way to manage and store comprehensive clinical data in EHRs. This data captured during routine care is a valuable source for research using OMOP.
- **PRSB (Patient Reported Symptoms Burden):** PRSB questionnaire/survey designed with FHIR and openEHR standards in mind, potentially can be integrated with OMOP to enrich research on symptom impact.

How they work with OMOP:

- **Data Source:** Data from openEHR and FHIR enabled systems becomes a source for OMOP research.
- **Transformation and Mapping:** After addressing heterogeneity and data quality issues in the EHR data, data is transformed and mapped to the OMOP CDM.
- **Fast Research:** Once transformed, the data becomes usable for large-scale research analysis using OMOP tools. Researchers can then investigate trends, treatment effectiveness, and patient experiences in a quicker and collaborative way.

How do common data models like DICOM, FHIR, openEHR & PRSB work with OMOP for research?

2. DICOM (Digital Imaging and Communications in Medicine):

- **Focus:** DICOM is specifically designed for exchanging medical images (X-rays, MRIs, CT scans, etc.) and associated data electronically.

How DICOM work with OMOP:

- Integration of imaging data aligned to the OMOP CDM has been purposed and awaiting feedback.
- This extension adds two new tables, designed to capture detailed information about imaging events and features with their provenance.
- Integrating imaging features with OMOP CDM unlocks a wealth of data for researchers, leading to more comprehensive studies.

Standards working together for better local, national and international healthcare understanding …

Clinical data standardised to FHIR, OpenEHR, PSRB and DICOM provide a comprehensive data source for clinicians to provide better patient care.

Transforming this standardised data to the OMOP CDM allows large scale research to be conducted from local, national and international EHR systems, providing better healthcare insights and better decision making.

Scotland’s Health Data Sharing Landscape

Health & Social Care Data Strategy

First Data Strategy for Health and Social care focusing On areas:

- **Data Access**
- **Talent and Culture**
- **Protecting and Sharing Data**
- **Technology and Infrastructure**
- **Information Standards and Interoperability**
- **Creating Insight from Data**
- **Supporting Research and Innovation**

Through the **Cloud Platform Seer,** long term ambition to improve data quality and create data links, enabling frontline and social care professionals to make better patient care decisions in a secure, transparent and ethical manner.

Health and social care: data strategy - executive summary

Safe Havens in Scotland

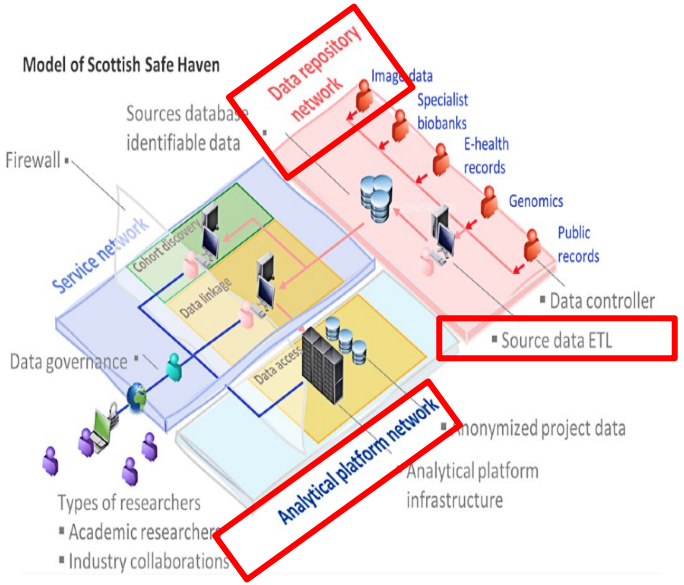
Network of 5 Safe Havens has been established to support:

- **Data discovery** and research feasibility
- Indexing and **linkage**
- **Cohort Building**
- **Analytical support** and archive

Safe Havens make use of **two networks**:

1. **Data repositories** set up mainly on NHS networks.
2. **Analytical platform** set up within university-managed networks

*"Provides **secure analytics platforms** for researchers to access linked, deidentified electronic health records (EHRs) while managing the risk of unauthorised reidentification."*



A National Network of Safe Havens: A Scottish Perspective (Preprint)

ETL = **E**XTRACT, **T**RANSFORM, **L**OAD

Using a Common Data Model in Safe Havens?

"If common data models (CDM) such as ***Observational Medical Outcomes Partnership (OMOP)*** *and i2b2 star schema were used, either for data repositories or analytical platforms, the burden on multi–Safe Haven projects would be* ***reduced****, and* ***operational access to data would be faster and more predictable."***

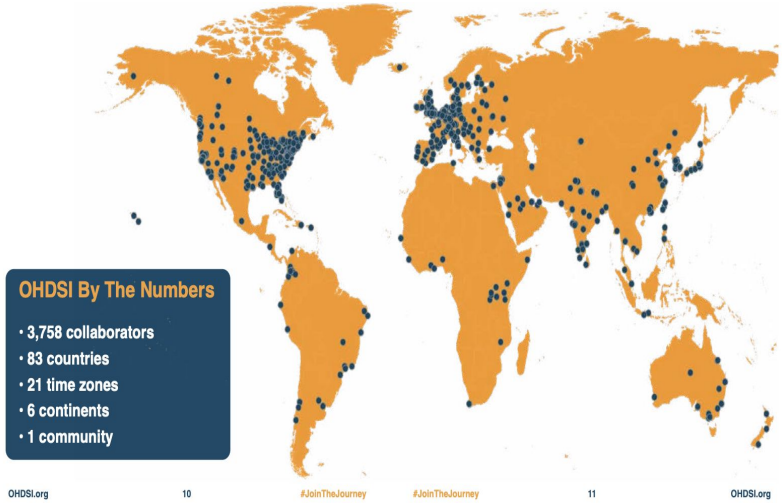
OMOP CDM Overview

What is the OMOP-CDM?

- **Goal:** The Observational Health Data Sciences and Informatics (**OHDSI**) a multi-stakeholder, interdisciplinary **community** that is striving to bring out the value of observational health data through large-scale analytics.

- **Method:** Created a standardised data model the Observational Medical Outcomes Partnership Common Data Model (OMOP-CDM)

- **Impact:**
  - **Growing adoption over 10 years**
  - **534 data sources** from **49 countries mapped**
  - Represents **956 million patients** (**12% of world population)**



Where the OHDSI community has been and where are we going Annual report 2023

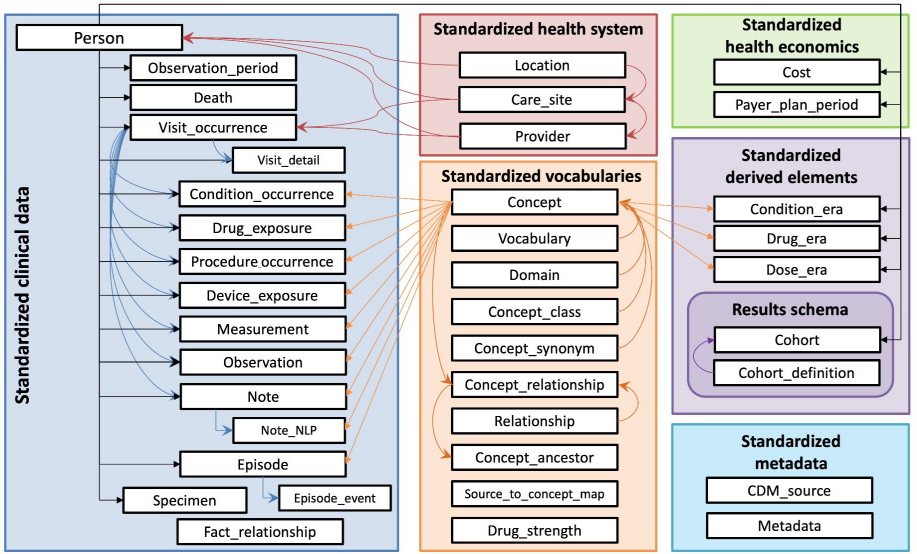
Adopting the OMOP-CDM

OMOP-CDM used to l**ink Observational healthcare databases**.

Key features:

- Patient-centric
- Tabular/Relational Design
- Standardised Vocabulary
- Extendable
- Built for Collaborative Analytics

It's benefits can also **extend to planning and population health management**



OMOP-CDM V.5.4

Benefits of adopting the OMOP-CDM

- **Open-source and Collaborative:**
  - Actively developed and maintained tools on Github with a supportive community forum.
  - Offers guidance and assistance for general and specific data conversion questions.

- **Standardised Representation:**
  - Creates a common language for healthcare data, simplifying researcher understanding.

- **FAIR Data Principles:**
  - Supports **F**indability, **A**ccessibility, **I**nteroperability, and **R**eusability of data.
  - Benefits researchers and healthcare technology development for better patient and population health decisions

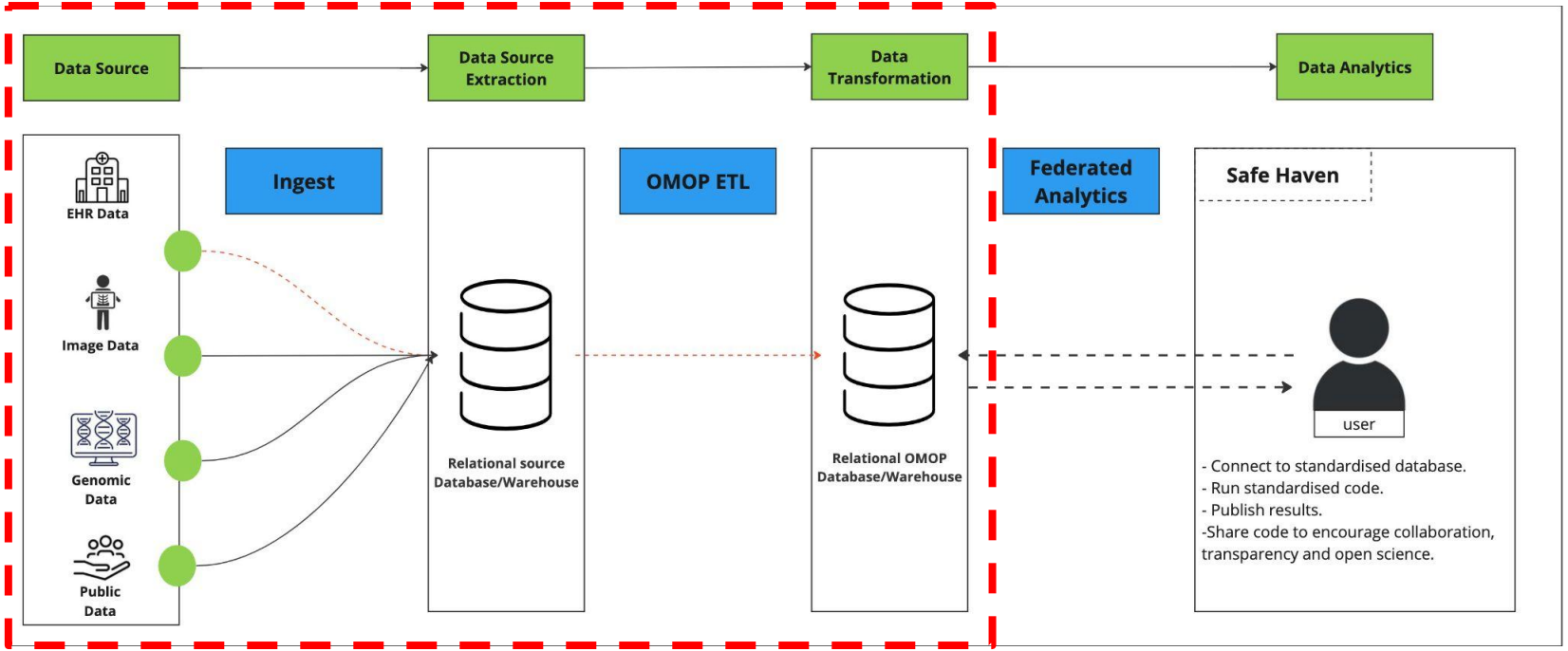
- **Large-Scale Datasets:**
  - Enables creation of large, standardized datasets for national and international federated network studies.

Challenges when adopting the OMOP-CDM

- **Data Dictionary Effort:**
  - Requires well-defined source data dictionaries, which can be resource-intensive to create and maintain.
- **Conversion Time:**
  - First iteration of mapping typically takes 6-12 months to complete.
- **Technical Considerations:**
  - Setting up and governing data, especially with PII (Personally Identifiable Information), can lead to delays.
- **Expert Input:**
  - Conversion requires input from clinical experts familiar with the source data.
- **Ongoing Maintenance:**
  - Regular updates are needed to adapt to changes in source data, data models, and vocabularies.

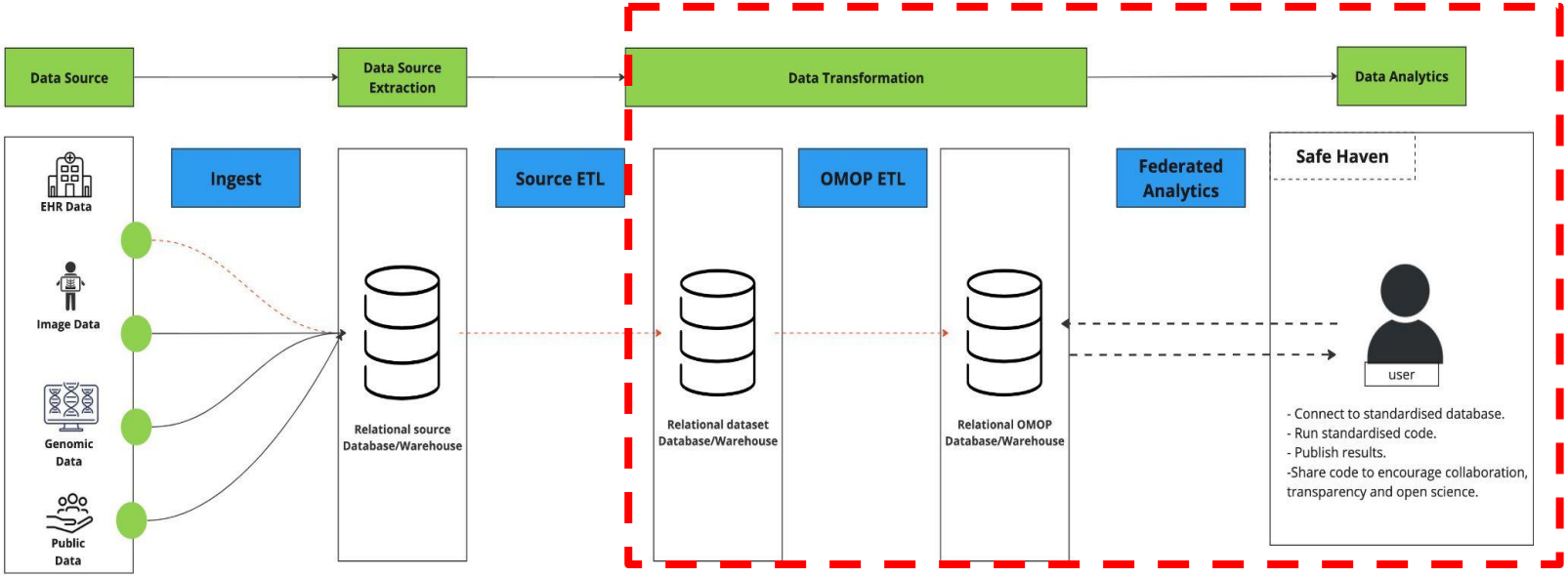
PoC for Data Harmonisation Strategy

Standardising within **data repository**



OMOP ETL carried out on NHS network

Standardising within **analytical platform**



OMOP ETL carried out analytical platform

Possible areas to work Together

1 - Data interoperability workshops:

- Education and upskilling through bringing awareness to the use and benefit of standards such as the OMOP-CDM.

2 - Data engineering services using OMOP-CDM:

- Mapping a subset of data within a NHS network which has not converted data to OMOP-CDM
- Mapping a subset of data within an analytical repository that has not been covered the OMOP-CDM

3 - Testing a Proof Of Concept (PoC) using OMOP-CDM:

- Building ETL pipelines for the subset and sharing learning with directorates to support the data strategy plan

4 - Expanding the concept to other data sources:

- Scaling up the pipelines to look at imaging, genomics, social care, other local authorities data.

Thank you for listening!

Useful Links

- Health and social care: data strategy - executive summary
- NHS Research SDE Network agrees to adopt common data model
- A National Network of Safe Havens: A Scottish Perspective (Preprint)
- Where the OHDSI community has been and where are we going Annual report 2023
- OMOP-CDM V.5.4