Project Title	Global cooperation on FAIR data policy and practice
Project Acronym	WorldFAIR
Grant Agreement No	101058393
Instrument	HORIZON-WIDERA-2021-ERA-01
Topic, type of action	HORIZON-WIDERA-2021-ERA-01-41 HORIZON Coordination and Support Actions
Start Date of Project	2022-06-01
Duration of Project	24 months
Project Website	http://worldfair-project.eu

D7.2 Population health resource library and training package

Work Package	WP07 Population Health
Lead Author (Org)	Jim Todd (LSHTM)

Contributing Author(s) (Org)	Keith Tomlin (LSHTM) Tathagata Bhattacharjee (LSHTM) David Amadi (LSHTM) Jay Greenfield (CODATA) Doug Fils (CODATA) Dorothy Mailosi (CODATA) Chifundo Kanjala (independent expert)
Due Date	30.11.2023
Date	30.10.2023
Version	1.0 DRAFT NOT YET APPROVED BY THE EUROPEAN COMMISSION
DOI	https://doi.org/10.5281/zenodo.10010936

Dissemination Level

X	PU: Public
	PP: Restricted to other programme participants (including the Commission)
	RE: Restricted to a group specified by the consortium (including the Commission)
	CO: Confidential, only for members of the consortium (including the Commission)

Versioning and contribution history

Version	Date	Authors	Notes
0.1	27.10.2023	Jim Todd, Keith Tomlin, Tathagata Bhattacharjee, David Amadi, Jay Greenfield, Doug Fils, Dorothy Mailosi, Chifundo Kanjala	Draft for internal review
0.2	31.10.2023	Laura Molloy, Simon Hodson	Reviewed draft

'Global cooperation on FAIR data policy and practice' (WorldFAIR) has received funding from the European Union's Horizon Europe project call HORIZON-WIDERA-2021-ERA-01-01, grant agreement 101058393.

1.0	30.11.2023	All authors	Content ready

Disclaimer

WorldFAIR has received funding from the European Commission's WIDERA coordination and support programme under the Grant Agreement no. 101058393. The content of this document does not represent the opinion of the European Commission, and the European Commission is not responsible for any use that might be made of such content.

'Global cooperation on FAIR data policy and practice' (WorldFAIR) has received funding from the European Union's Horizon Europe project call HORIZON-WIDERA-2021-ERA-01-01, grant agreement 101058393.

Abbreviations and Acronyms

ALPHA	Analysing Longitudinal Population-based HIV data from Africa
APHRC	African Population Health Research Centre
CDM	Common Data Model
CDIF	Cross Domain Interoperability Framework
CODATA	Committee on Data (International Science Council)
CORDIS	Community Research and Development Information Service
DDI	Data Documentation Initiative
DSD	Data Structure Definition
EHDEN	European Health Data Evidence Network
EOSC	European Open Science Cloud
ETL	Extract Transform Load
FAIR	Findable, Accessible, Interoperable, Reusable
FIP	FAIR Implementation Profile
FER	FAIR Enabling Resource
HEIs	Higher Education Institutions
HDSS	Health and Demographic Sentinel Site
IDSR	Infectious Disease Surveillance and Response
INSPIRE	Implementation Network for Sharing Population Information with Research Entities
JSON-LD	Javascript Object Notation – Linked Data
LSHTM	London School of Hygiene and Tropical Medicine
OHDSI	Observational Health Data Sciences and Informatics program
OMOP	Observational Medical Outcomes Partnership

'Global cooperation on FAIR data policy and practice' (WorldFAIR) has received funding from the European Union's Horizon Europe project call HORIZON-WIDERA-2021-ERA-01-01, grant agreement 101058393.

SDG	Sustainable Development Goals
SDMX	Statistical Data and Metadata Exchange
WHO	World Health Organisation
UN	United Nations

‘Global cooperation on FAIR data policy and practice’ (WorldFAIR) has received funding from the European Union’s Horizon Europe project call HORIZON-WIDERA-2021-ERA-01-01, grant agreement 101058393.

Executive Summary

This project, WorldFAIR – Global Cooperation on FAIR Data Policy and Practice, is funded by the European Commission's WIDERA coordination and support programme under the Grant Agreement no. 101058393. The project consists of 14 work packages, of which work package 7 (WP07) focusses on Population Health. WP07 is led by London School of Hygiene and Tropical Medicine working under the INSPIRE network[1]. The work builds on the delivery of the Observational Medical Outcomes Partnership (OMOP) common data model (CDM) which includes funding by Wellcome (formerly Wellcome Trust) and IDRC Canada.

The objective of WP07 is to develop a suite of methods and standards to provide the framework for the Go-FAIR principles[2] for population health data. These standards form the basis of an AI-Ready description of data suitable for use by population health scientists, and understandable across domain and institutional boundaries. The first deliverable (D7.1) identified the Implementation Guide that could be used for population health data, and how it can be developed. This deliverable (D7.2) provides a step-by-step guide as to how to achieve the standards. The deliverable is aimed at population health scientists in low-resource settings, who know their own data and want to make those data FAIR.

Population health uses many different tools to collect and manage data. One set of tools includes the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM)[3] and an OHDSI[4] data analysis workbench that runs on top of it. The OMOP common data model has been used to harmonise and share data, and previous work has shown the tools needed to make OMOP data FAIR. Beyond the data themselves, the results from the analyses conducted on OMOP data can be used as indicators for the success of development goals, including the United Nations Sustainable Development Goals (SDGs)[5]. At each stage the tools, data, models and activities need to be described in a way that can be understood by other scientists and by computer search algorithms.

This deliverable provides an introduction to the processes involved in making population health data FAIR in a pipeline that spans data collection through data analysis into an SDMX[6] indicators database, and gives seven tutorials on what is needed at each step in this pipeline. It outlines the need to describe the study and the study context, how to use DDI Codebook[7] and DDI Lifecycle[8] with study data and how to use repositories like GitHub[9] to make the metadata available. The next tutorials describe the extract-transform-load (ETL) process for putting the data into an OMOP CDM and the role of JSON-LD in preparing the data for machine searching in Schema.org in line with DDI-CDI[10].

1 https://aphrc.org/inspire/
2 https://www.go-fair.org/fair-principles/
3 https://www.ohdsi.org/data-standardization/
4 https://www.ohdsi.org/
5 https://sdgs.un.org/goals
6 https://sdmx.org/
7 https://ddialliance.org/Specification/DDI-Codebook/
8 https://ddialliance.org/Specification/DDI-Lifecycle/
9 https://github.com/
10 https://ddialliance.org/Specification/ddi-cdi

'Global cooperation on FAIR data policy and practice' (WorldFAIR) has received funding from the European Union's Horizon Europe project call HORIZON-WIDERA-2021-ERA-01-01, grant agreement 101058393.

Together these tutorials give an overview of the steps in the OMOP processes which are a pipeline for the data, and how these steps can be performed and documented. Finally the tutorials show how predictive and causal analysis can be conducted and documented using the OMOP CDM and the OHDSI data analysis workbench and how the results can be integrated into an SDMX data cube, which would align with UN standards for SDG indicators.

The deliverable does not provide detailed training for each step, but rather introduces the topic and clarifies the practical knowledge and skills that are needed to make this type of health data more FAIR.

Note: The tutorials are hosted on the WorldFAIR Vimeo channel https://vimeo.com/user/91439529/folder/18642763, which provides functionality for playing videos. Alternatively you can download the tutorials and play in your local environment to experience full functionality.

'Global cooperation on FAIR data policy and practice' (WorldFAIR) has received funding from the European Union's Horizon Europe project call HORIZON-WIDERA-2021-ERA-01-01, grant agreement 101058393.

Table of contents

'Global cooperation on FAIR data policy and practice' (WorldFAIR) has received funding from the European Union's Horizon Europe project call HORIZON-WIDERA-2021-ERA-01-01, grant agreement 101058393.

1. Introduction and scope

Population health data, including data from health and demographic surveillance sites (HDSS) provide an important insight into the health of communities, which often are representative of the nation or region from which they are drawn. The longitudinal follow-ups of the population produce increasing amounts of data, but often create ethical, logistical, and institutional barriers to the FAIR principles for research data[11]. WorldFAIR Work Package 7 (WP07) aims to define the standards for making population health data FAIR, and create the structures which can be used by population health scientists to achieve FAIR data (EHDEN, 2020).

1.1. Using OMOP to share and harmonise population health data

Through the INSPIRE network[12], we identified the means by which the Observational Medical Outcomes Partnership (OMOP) common data model (CDM)[13] could be purposed for population health data, including longitudinal data (OHDSI, 2020). This includes the use of federated databases, where instances of OMOP can be extracted-transformed-loaded (ETL-ed) in a secure way to ensure the confidentiality of the data, while shared results are obtained through ATLAS[14] executable analyses using the shared metadata.

Early work within WorldFAIR identified that OMOP data are not intrinsically FAIR. The OMOP tables are based on standard ontologies and vocabularies that can be reused in different settings. However there is no integrated mechanism for describing FAIR implementation profiles (FIPs)[15] and FAIR enabling resources (FER) available in any particular OMOP CDM. Researchers looking for data using the OMOP CDM and OHDSI tools need to be able to search for the metadata, the aggregate data used in analyses, the microdata from individual responses, and the processes and relationships between the microdata. The nature of the protocols between the different microdata and the analyses are important to understanding the results.

The first step (WorldFAIR Deliverable 7.1)[16] was to identify and describe the Implementation Guide for making OMOP data FAIR by describing the metadata, data and processes used for each data source. The implementation guide needs to be built up alongside the data in the OMOP CDM, so that each step of the process is explicitly described and can be used to understand the data and results.

[11] https://www.nature.com/articles/sdata201618
[12] https://aphrc.org/inspire/
[13] https://www.ohdsi.org/data-standardization/
[14] https://atlas.ohdsi.org/
[15] https://doi.org/10.5281/zenodo.7378109
[16] https://zenodo.org/records/7887385

'Global cooperation on FAIR data policy and practice' (WorldFAIR) has received funding from the European Union's Horizon Europe project call HORIZON-WIDERA-2021-ERA-01-01, grant agreement 101058393.

1.2. Defining the implementation guide for describing OMOP data

WorldFAIR Deliverable 7.1 used data from the Infectious Disease Surveillance and Response (IDSR) case-based surveillance forms[17] to demonstrate how to make the OMOP data FAIR. The IDSR forms are recommended by WHO (WHO 2019) and used in many countries for notifiable diseases and to identify potential pandemic response. The IDSR data provide an example of how data can be managed from the raw data collected by the study team, through the ETL process into the OMOP pipeline.

Figure 1 shows the processes by which data are moved through the OMOP pipeline from data collection through to analysis and results which can be used by policy makers. The pipeline can be repeatedly used with new data to provide real-time updates useful for policy makers tracking an epidemic.

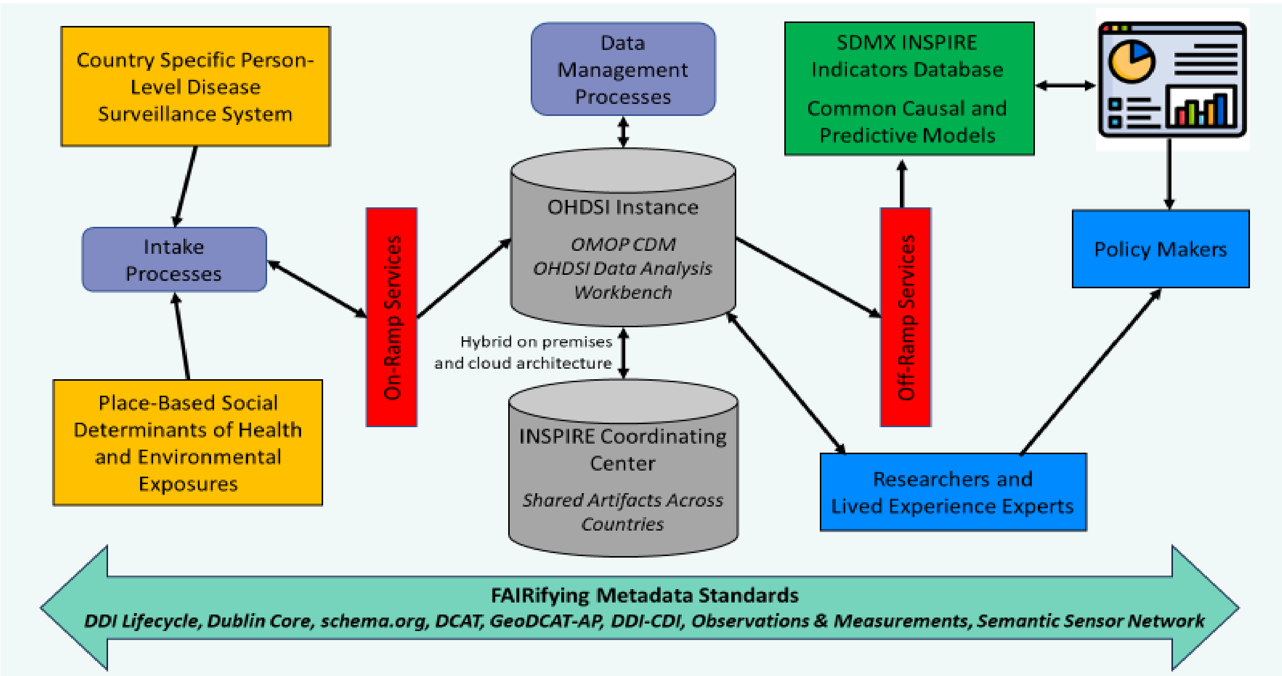

Figure 1. Processes for sharing IDSR data using the OMOP CDM

The implementation guide (see Deliverable 7.1) outlines how the metadata model in OMOP can be described using Schema.org (Schema.org 2022). As part of the Cross-Domain Interoperability Framework (CDIF)[18], WorldFAIR WP02[19] proposed that the implementation guide for the use of Schema.org with OMOP tables can be expressed using JSON-LD. The implementation guide

[17] https://iris.who.int/bitstream/handle/10665/331234/WHO-2019-nCoV-SurveillanceCRF-2020.2-eng.pdf?sequence=1
[18] https://worldfair-project.eu/cross-domain-interoperability-framework/
[19] https://worldfair-project.eu/synthesis/

'Global cooperation on FAIR data policy and practice' (WorldFAIR) has received funding from the European Union's Horizon Europe project call HORIZON-WIDERA-2021-ERA-01-01, grant agreement 101058393.

describes how different tables, processes and relationships can all be described using JSON-LD and this would provide a resource that is both machine readable and understandable to researchers.

However, as shown in Fig 1, the ability to describe and FAIRify the metadata standards across the whole data pipeline requires multiple tools and integration of resources (Wilkinson, 2023). Population health researchers may be focused on the individual study data, and the exposomes (that is, the social determinants of health and environmental exposures) affecting the population . To make the data and metadata FAIR we need to identify and describe the tools used for intake processes (on-ramp and ETL processes), for the analysis within OMOP (using OHDSI tools), and for how the results can be found and used by researchers and policy makers.

The objective of this deliverable is to provide a step-by-step guide to each of these processes so that population health scientists in low resource settings can easily describe their data and metadata through the whole OMOP pipeline.

2. Describing the D7.2 tutorials

This Deliverable consists of a series of seven tutorials which take the participant through the OMOP pipeline, alongside this explanatory document. Six of the tutorials are videos, and one of the tutorials is an interactive MindMap. The videos are found in:

https://vimeo.com/user/91439529/folder/18642763

(Note: The tutorials are hosted on a Vimeo channel which provides functionality for playing videos. Alternatively you can download the tutorials and play in your local environment to experience full functionality.)

In each of the tutorials, a different perspective on successful use of the OMOP pipeline is shown with the accompanying tools for describing the process. In this way, the tutorials extend the DDI data pipeline[20] currently used by many researchers in health settings to include the harmonisation process and the analysis and outcomes that can be obtained from the extended OMOP pipeline (Fig 2).

[20] https://ddialliance.org

'Global cooperation on FAIR data policy and practice' (WorldFAIR) has received funding from the European Union's Horizon Europe project call HORIZON-WIDERA-2021-ERA-01-01, grant agreement 101058393.

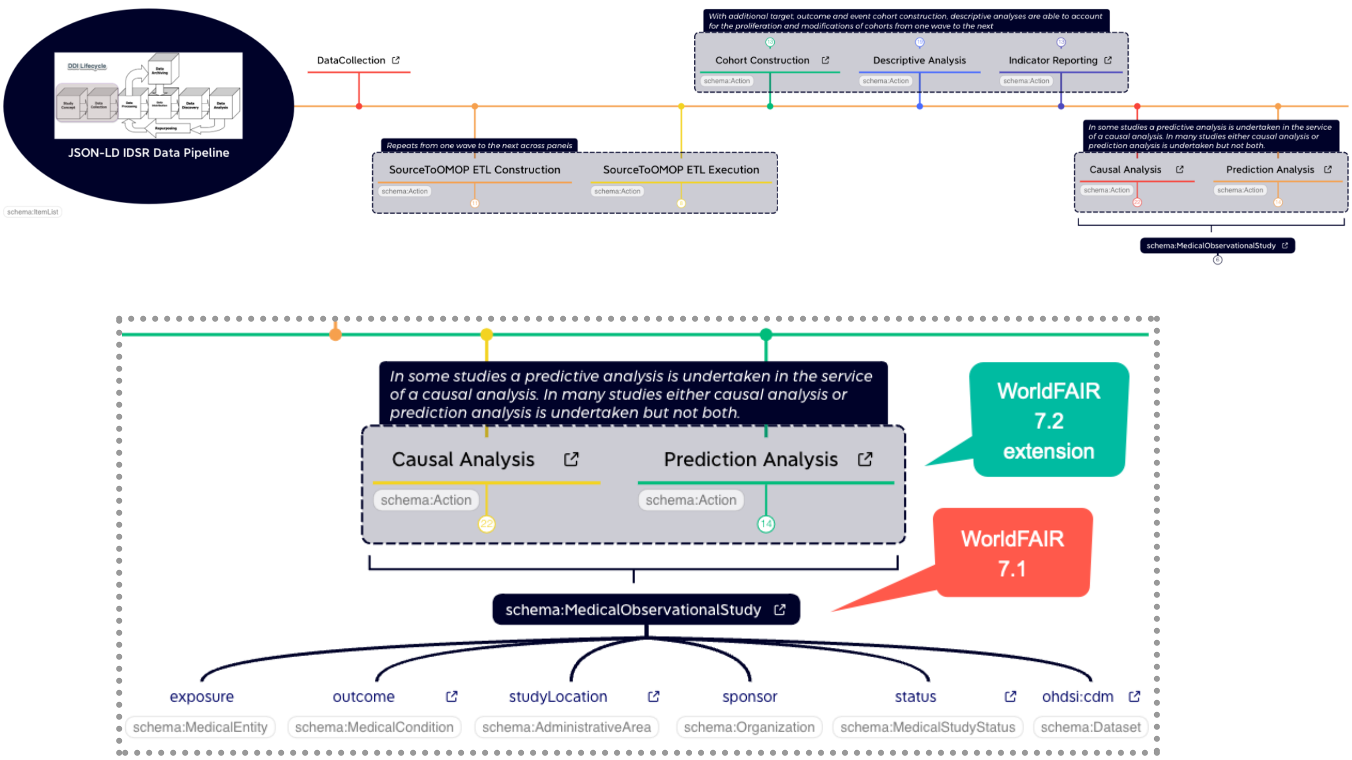

Figure 2. Extension of the DDI data pipeline for harmonised OMOP data.

The work uses the same case-study as in Deliverable 7.1. By using IDSR data to demonstrate the tools for making the data FAIR, the links with the previous deliverable are strengthened. The seven tutorials from Deliverable 7.2 cover the following topics:

1. Rationale for making population data FAIR and overview of the tutorials
2. Describing the study and the context for mapping the data to OMOP
3. Making the study metadata FAIR
4. An introduction to JSON-LD and how to express structured data on the web
5. An overview of the OMOP pipeline, including a description of the data collection and analysis
6. Using the ETL processes to take data from the IDSR into the OMOP CDM
7. Sharing descriptive statistics and integration with UN indicators.

In the following subsections, we describe the scope and intended learning outcomes for each of these tutorials.

2.1. Rationale for FAIR data

https://vimeo.com/user91439529/review/889039921/37b3c0edc5

'Global cooperation on FAIR data policy and practice' (WorldFAIR) has received funding from the European Union's Horizon Europe project call HORIZON-WIDERA-2021-ERA-01-01, grant agreement 101058393.

This tutorial highlights the unique nature of population health data with ethical, legal, social, and institutional (ELSI) barriers, which prevent the timely sharing of data. The ethical issues are not only to preserve the confidentiality of individuals in the database, but also recognize the value of the data to data producers. Until recently there has been limited infrastructure available to population health studies in low resource settings, to make their data FAIR. There is an urgent need for governance and policy on data sharing before the processes can be developed.

The tutorial introduces the OMOP common data model which is available through the INSPIRE network. We highlight the use of repositories as the first step in describing our data. The WorldFAIR cross-domain model provides the tools for making population health data FAIR, and highlights the need for the step-by-step understanding of each of the tools in the process. Ultimately the purpose of making our population health data FAIR is to ensure our results can be used as evidence for good policies to improve the health in low and middle income countries. The tools we introduce are useful in achieving that objective.

2.2. Describing the IDSR study and context

https://vimeo.com/user91439529/review/889058769/747ab388bd

This tutorial explains the infectious disease surveillance and response (IDSR) dataset and its implementation worldwide. Understanding the context of the IDSR studies is important to making the data FAIR. The tutorial takes the case-based form and explains how the source data can be input into the OHDSI common data model.

The tutorial explains the OMOP CDM and how it works. It provides a simple explanation of the tools available from OHDSI which are available in an OMOP CDM. It introduces the steps for enabling the source data to be extracted, transformed and loaded (ETL) into the OMOP CDM. It highlights the importance of concepts and vocabularies in ETL of the data and their subsequent use.

2.3. FAIR study metadata

https://vimeo.com/user91439529/review/889062010/27e1663cd4

This tutorial explains the role of metadata for FAIR data. It defines metadata as “data that describes other data”, which makes it important to understand any set of data and essential to the production of FAIR data. Metadata starts with the knowledge of the domain, and various tools can be used to describe the domain and the data in ways that allow others to search, find, access and use the data.

‘Global cooperation on FAIR data policy and practice’ (WorldFAIR) has received funding from the European Union’s Horizon Europe project call HORIZON-WIDERA-2021-ERA-01-01, grant agreement 101058393.

The tutorial goes through the DDI Lifecycle[21] which can be used by every researcher to describe the collection, processing and analysis of the data in their study (DDI 2023). DDI Codebook[22] and DDI Lifecycle can be mapped to other standards, including standards which are used in the OMOP CDM. Most importantly the metadata can be mapped into schemas, including schema.org[23] which allows the metadata and data to be machine readable across the web. The way we use to do this is through JSON-LD programs which can be constructed from DDI files.

2.4. Using JSON-LD

https://vimeo.com/889355108?share=copy

This tutorial introduces how we can use JSON-LD to express structured data on the web[24]. It outlines the advantages of expressing data in this way, and how this makes the data, metadata and process FAIR. JSON-LD encodes the RDF data model, is easily understood by both humans and machines (computers) and is supported by a wide range of tools and resources. The JSON-LD code links or maps the key concepts to descriptions and context of the metadata and data. JSON-LD is a powerful format for representing and sharing structured data on the web, offering a number of benefits, including scalability, expressiveness, ease of implementation and interoperability.

The tutorial describes the components which can be connected to our metadata. This uses standard shared vocabularies which aid understanding of the concepts and resources we use. It describes how to use connections, framing and validation to build JSON-LD code. The code can then be accessed and consumed by others to form Knowledge Graphs which link to many other applications. We demonstrate some of the many applications which can be used once the information is expressed in this way. JSON-LD can be used for a variety of purposes, such as validation, master data catalogue, search engine optimisation, knowledge graph generation, and data integration.

2.5. The OMOP data pipeline

The tutorial uses a MindMap[25]: https://xmind.app/m/ktzbmQ to explain the OMOP data pipeline. Once you access the MindMap, please increase the MindMap size ideally to 100% (and at least 50%).

The MindMap tutorial embraces the **exploration method of learning**, in which you will enter a knowledge graph / mind map that runs both wide (across the pipeline) and deep (details for each

[21] https://ddialliance.org/Specification/DDI-Lifecycle/
[22] https://ddialliance.org/Specification/DDI-Codebook/2.5/
[23] https://schema.org/docs/schemas.html
[24] https://api.thehyve.nl/uploads/Vos-FAIR-metadata-1.pdf
[25] https://xmind.app/m/TYRVvs

'Global cooperation on FAIR data policy and practice' (WorldFAIR) has received funding from the European Union's Horizon Europe project call HORIZON-WIDERA-2021-ERA-01-01, grant agreement 101058393.

step in the pipeline). Users should interact with it (see the users' guide), and can explore the whole pipeline (the 'wide' view) and some of the details within the (meta)data pipeline (the 'deep' view). At each point in the pipeline, it might become a rabbit hole for further investigation, or you can move on to the next item in the pipeline. Mostly, you will discover objects and properties, some of which have a guide on how they are used and some of which do not have a guide, although this may be available on the internet. One guide is the Book of OHDSI[26] which explains the way OHDSI works. Another guide is available through the large language model called Bard[27]. It is possible for anyone (a human) using this project for a specific (meta)dataset, to read the knowledge graph, or mind map and turn it into a machine language called JSON-LD. JSON-LD is a machine-readable representation of the (meta)data pipeline that can be used to populate and semantically annotate databases. The translation of the mind map (knowledge graph) into JSON-LD (and vice versa) is a journey, which is still being developed and we have our early versions of translations for each stage across the data pipeline.

The tutorial shows how the JSON-LD DDI Lifecycle and schema.org representation of the IDSR population health research data contribute to a (meta)data pipeline. The metadata pipeline for IDSR data includes (i) the OHDSI platform, (ii) an SDMX compliant indicators database and (iii) a GitHub repository. Some of the extra resources used in the MindMap are shown here:

Term	**Description**	**Link**
DDI Lifecycle	Data Documentation Initiative Lifecycle specification	https://ddi-lifecycle-documentation.readthedocs.io/en/latest/
IDSR	Integrated Disease Surveillance and Response	https://www.afro.who.int/publications/technical-guidelines-integrated-disease-surveillance-and-response-african-region-third
JSON-LD	JSON for Linked Data	https://en.wikipedia.org/wiki/JSON-LD
OHDSI	Observational Health Data Sciences and Informatics	https://www.ohdsi.org
schema.org	schema.org	https://schema.org
SDMX	Statistical Data and Metadata eXchange	https://sdmx.org/?page_id=2561

This Mindmap tutorial introduces the data pipeline available in the OMOP CDM, and how the OHDSI tools can be used within the pipeline. The tutorial not only describes the progression of data through the OMOP pipeline from data collection through to analysis, but also provides sample code

[26] https://ohdsi.github.io/TheBookOfOhdsi/
[27] https://bard.google.com/

'Global cooperation on FAIR data policy and practice' (WorldFAIR) has received funding from the European Union's Horizon Europe project call HORIZON-WIDERA-2021-ERA-01-01, grant agreement 101058393.

at certain steps. It links the study hierarchy with the JSON-LD representation of the study, enabling us to see how the study is encoded within the data pipeline.

2.6. The ETL processes for OMOP CDM

https://vimeo.com/user91439529/review/889064160/150a677a72

This tutorial provides more depth to the construction of the ETL pipeline. It provides videos of the process by which the ETL codes are produced. In doing so it picks up the White Rabbit[28], Rabbit-in-a-Hat[29] and Usagi[30] tools which were introduced in Tutorial #2, "Describing the IDSR study and context", for the ETL construction. Those tools enable us to build SQL and R codes to execute the ETL for bringing data into the OMOP CDM. The accompanying videos show how this can be managed using Pentaho Data Integration tools[31]. It also highlights the importance of the GitHub repository both for maintaining the tools used, but also to provide interoperability for tools which can be reused by others.

2.7. Results and United Nations (UN) indicators

https://vimeo.com/889605178?share=copy

This tutorial concentrates on the results from analyses that are conducted on OMOP data. These results are like any other data: they need to be FAIR, which will make them visible and available to others. The UN system has a well-used data cube for SDG indicators using SDMX data structure definition (DSD)[32]. The key question is how we can push the OMOP analysis results into the format for the UN indicators thereby making the results FAIR.

There are several advantages when we integrate the results from the analyses of data in our OMOP CDM with the internationally recognised SDMX indicators. This approach stores local, regional and national estimates of the indicators side by side, where they can be compared and triangulated. The automated re-analysis of the OMOP data can be scheduled on a regular basis, or when new data are collected. This tutorial describes how coding provides the means to relate the indicators to external exposures including climate and environmental data. The UN statistics are widely available not just for the SDG indicators, but also for processes which contribute to the indicators.

[28] https://www.ohdsi.org/analytic-tools/whiterabbit-for-etl-design/
[29] https://ohdsi.github.io/WhiteRabbit/RabbitInAHat.html
[30] https://ohdsi.github.io/TheBookOfOhdsi/ExtractTransformLoad.html#usagi
[31] https://www.hitachivantara.com/en-us/products/pentaho-platform/data-integration-analytics/pentaho-community-edition.html
[32] https://data.un.org/WS

'Global cooperation on FAIR data policy and practice' (WorldFAIR) has received funding from the European Union's Horizon Europe project call HORIZON-WIDERA-2021-ERA-01-01, grant agreement 101058393.

3. Conclusions

These tutorials are just an introduction to the steps by which population data, metadata and data process can be made FAIR. They are not intended to explain the processes completely, but do show the scope of the documentation that is needed. Some institutions and studies may be starting with the initial documentation of their data and metadata using DDI Codebook or DDI Lifecycle. In doing so, it is important to provide a data dictionary for the study and also to describe the study context and processes which are applied to the data. These can be performed through ETL of the data into the OMOP CDM which provides standardised ways to describe concepts used in the study.

Making the OMOP CDM FAIR requires extra work which can be facilitated by JSON-LD. JSON-LD expresses the data in a structured way which encodes the RDF data model making the data and metadata accessible to both humans and computers. The development of knowledge graphs provides the link to other applications and ensures the data and metadata are understandable to researchers in other domains.

Documentation of the whole OMOP pipeline not only describes the data and metadata, but also documents the processes used at each step. These processes and algorithms can be stored in repositories such as GitHub, allowing for Interoperability and Reuse by others. For each step along the pipeline there are tools available for the data scientist, ranging from White Rabbit, Rabbit in a Hat and Usagi for the visualisation of the data through to Pentaho Data Integration tools for managing the ETL programs.

The end product is a set of data that are FAIR and can be used to produce results that integrate with other sharing standards. High quality results can be integrated into international systems for reporting indicators. The WHO SDGs use the SDMX data structure definitions, making the results FAIR and accessible worldwide.

We anticipate that these freely-available tutorials will help those working in the population health sector to embrace the tools and methods introduced, and find them useful in their efforts to increase the FAIRness and interoperability of population health data.

'Global cooperation on FAIR data policy and practice' (WorldFAIR) has received funding from the European Union's Horizon Europe project call HORIZON-WIDERA-2021-ERA-01-01, grant agreement 101058393.

4. Bibliography

DDI (2023). Document, Discover and Interoperate, 24 April 2023, Available at https://ddialliance.org https://ddialliance.org

Delaney, J & Seeger, J (2012). Chapter 11 Sensitivity Analysis, 2012, Available at https://effectivehealthcare.ahrq.gov/sites/default/files/ch_11-user-guide-to-ocer_130129.pdf

EHDEN (2020). FAIRification of observational studies and databases: the EHDEN and OHDSI use case, Available at https://faircookbook.elixir-europe.org/content/recipes/applied-examples/ehden-ohdsi.html Last accessed 29 October 2023.

Gregory, A., Todd, J., Amadi, D., Greenfield, J., Muyingo, S., & Tomlin, K. (2023). WorldFAIR Project (D7.1) Population Health Data Implementation Guide (Version 1). Zenodo. https://doi.org/10.5281/zenodo.7887385

OHDSI (2020). OHDSI Kicks Off International Collaborative to Generate Real-World Evidence on COVID-19 with Virtual Study-a-thon. Available at https://www.ohdsi.org/ohdsi-news-updates/covid19-studyathon/. Last accessed 30 October 2023.

UNAIDS (2022). IXF version 3.0, Available at https://data.unaids.org/pub/manual/2008/sdmx_ixf3_en.pdf. Last accessed 29 April 2023.

US Department of State (2023). PEPFAR Fiscal Year 2023 Monitoring, Evaluation, and Reporting (MER) Indicators, Available at https://www.state.gov/pepfar-fy-2023-mer-indicators/" https://www.state.gov/pepfar-fy-2023-mer-indicators/. Last accessed 28 October 2023.

Schema.org (2022). MedicalObservationalStudy: A Schema.org Type, 25 October 2022, Available at https://schema.org/MedicalObservationalStudy" https://schema.org/MedicalObservationalStudy Last accessed 29 October 2023.

WHO (2019). Integrated Disease Surveillance and Response Technical Guidelines (3rd ed). Available at https://iris.who.int/bitstream/handle/10665/312317/WHO-AF-WHE-CPI-01-2019-eng.pdf Last accessed 22nd November 2023.

Wikipedia (2023). Neural network, 29 October 2023, Available at https://en.wikipedia.org/wiki/Neural_network. Last accessed 29 October 2023.

Wilkinson, M.D., Dumontier, M., Aalbersberg, Ij.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L.B., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J.G., Groth, P., Goble, C., Grethe, J.S., Heringa, J., 't Hoen, P.A.C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M.A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data 3, 160018.

'Global cooperation on FAIR data policy and practice' (WorldFAIR) has received funding from the European Union's Horizon Europe project call HORIZON-WIDERA-2021-ERA-01-01, grant agreement 101058393.

https://doi.org/10.1038/sdata.2016.18

'Global cooperation on FAIR data policy and practice' (WorldFAIR) has received funding from the European Union's Horizon Europe project call HORIZON-WIDERA-2021-ERA-01-01, grant agreement 101058393.