The approach of IQVIA to ontologies for healthcare and life sciences

Andrea Splendiani[1], Emiliano Reynares[2], and Anthony Reckard[3]

[1] IQVIA, Kirschgartenstrasse 14, Basel, Switzerland
[2] IQVIA, Provença 392, Barcelona, Spain
[3] IQVIA, Mission Street 123, San Francisco, CA, United States of America

Abstract

IQVIA is a healthcare data company that processes yearly over 100B health records from 1M+ distinct data feeds, addressing use cases from clinical development to real world evidence and market access. Ontologies are in use extensively to structure and unify this data space.
In this contribution, we intend to present how the company uses ontologies: what standard ontologies are integrated, what gaps exist and what custom solutions need to be developed.

Keywords

Ontology, knowledge graph, healthcare, life sciences

1. Introduction

Ontologies are broadly used across healthcare and life sciences. As a healthcare data company covering pharma, from early clinical development to market access as well as healthcare delivery, IQVIA makes an extensive usage of ontologies: for data harmonization (e.g.: OMOP [1]), regulatory submission support (e.g.: CDISC [2]), to structure data through NLP pipelines and to drive analytics (e.g.: population definitions).

The "ontology architecture" of IQVIA can be considered made of three parts, as illustrated below (for a sample of geographies and domains).

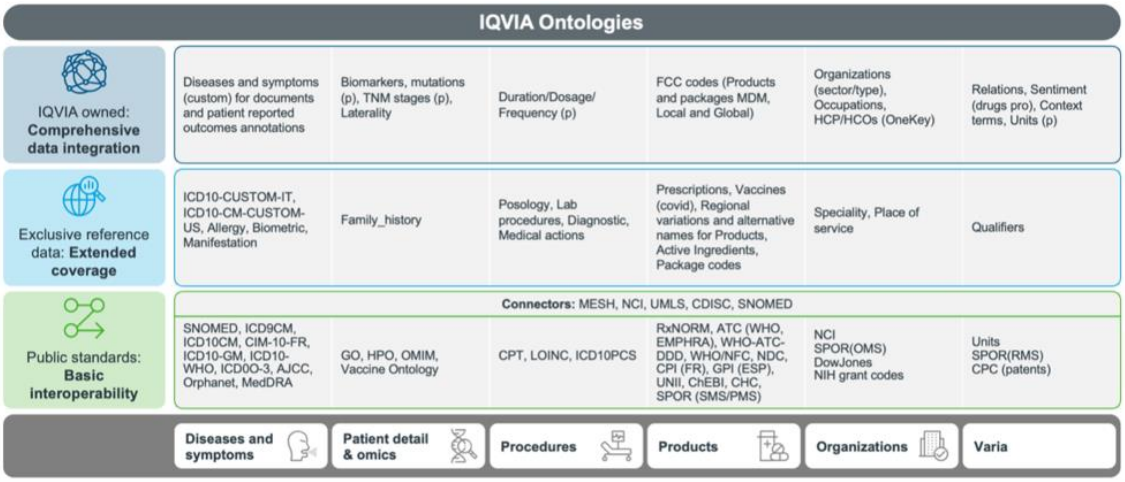

Figure 1: Examples of types of IQVIA ontology assets for selected resources and domains

SWAT4HCLS 2024: The 15th International Conference on Semantic Web Applications and Tools for Health Care and Life Sciences, February 26–29, 2024, Leiden, The Netherlands
andrea.splendiani@iqvia.com (A. Splendiani); emiliano.reynares@iqvia.com (E. Reynares); anthony.reckard@iqvia.com (A. Reckard)
0000-0002-3201-9617 (A. Splendiani); 0000-0002-5109-3716 (E. Reynares)

1.1. Standard ontologies

At the bottom, IQVIA makes use of ontologies that are "standard": either public or provided by third parties, these are shared among different actors of the healthcare ecosystem and constitute the basis for interoperability. The breadth of operations of IQVIA consider ontologies (or reference data) from many different perspectives. For instance, products may make use of as diverse resources as ChEBI [3] (research), ATC [4] (therapeutic classification), RxNorm [5] (products) and more.

1.2. Ontologies and reference data in use

In addition to this, IQVIA collects and integrate reference data from its massive health record ingestion processes. This results in extended resources, such a regional name variation for products, or typical references in prescriptions, or even products for which standard nomenclatures are not present (e.g.: lab supplies).

It is interesting to note that while ontologies represent the standard "as proposed", this extensive collection results in a view of the standard "as used", including an understanding of what granularities exist, what is updated when and so on.

1.3. The last mile

Finally, as an analytics company, IQVIA goes the "last mile" to complement current ontological resources to support analytics. This involves developing knowledge graphs that allow for navigation across ontology versions, even those that are over a decade old, to support the creation of longitudinal patient data. Additionally, custom ontologies are on development to mine patient-reported outcomes or social media for sentiment analysis.

This includes for instance the development of knowledge graphs enabling the navigation across ontology versions (sometimes beyond a decade) to support patient's longitudinal data creation, or the development of custom ontologies to, for instance, mine patients reported outcome or social media for sentiment analysis.

1.4. Future developments

With this contribution we intend to present our assets and expertise in the ontology space, and invite the interested parties to kick off a discussion on how to leverage such resources (e.g.: extensive observation of ontology use in the field) and advance the current state of the art.

References

[1] OHDSI, Standardized Data: The OMOP Common Data Model. Accessed in December 2023. URL: https://www.ohdsi.org/data-standardization/.

[2] Clinical Data Interchange Standards Consortium (CDISC). Accessed in December 2023. URL: https://www.cdisc.org/.

[3] Chemical Entities of Biological Interest (ChEBI). Accessed in December 2023. URL: https://www.ebi.ac.uk/chebi/.

[4] World Health Organization (WHO): Anatomical Therapeutical Chemical Classification (ATC). Accessed in December 2023. URL: https://www.who.int/tools/atc-ddd-toolkit/atc-classification.

[5] National Library of Medicine (NLM): RxNorm. Accessed in December 2023. URL: https://www.nlm.nih.gov/research/umls/rxnorm/overview.html.