We present the development of GA4GH Beacon v2 for structured clinical data discovery based on the Observational Medical Outcomes Partnership Common Data Model (OMOP-CDM). Beacon v2 is a standard specification for anonymous federated data discovery of genomic and phenoclinic data. Whereas, OMOP is a widely used standardized model for organizing and harmonizing clinical information mostly coming from Electronic Health Records (EHRs) data, enabling large-scale data analysis across multiple data sources. This Beacon implementation uses the OMOP-CDM to enable querying of clinical data sources for specific phenotypic criteria, such as diagnoses, treatments, and laboratory results. Thus, this application will increase the value of having an OMOP-CDM in the organisations as it will leverage the work of the Beacon implementers and future Beacon Networks (Beacon that query many Beacons).

Methods

- User-friendly interface for querying clinical data sources.
- Address the challenges and opportunities for implementing a GA4GH Beacon for clinical data discovery, including data privacy and security considerations, and the potential impact of such a tool for clinical research and healthcare delivery.
- Developed in Python following the EGA-CRG Beacon v2 RI API.

Software

- This project is based on the Beacon v2 Python reference implementation and has been developed in the context of the IMPaCT-DATA WP3.5.
- The reference implementation for this development can be downloaded from the following repository: gitlab.bsc.es/impact-data/impd-beacon_omopcdm
- Please follow the instructions provided in the README for installation and usage. The software is released under an Apache 2.0 license.
- The main libraries employed are:
  - aiohttp for exposing the REST API methods
  - aiosql to model the SQL real-time queries against the OMOP database.

References


Contact
Sergi Aguilo-Castillo
sergi.aguilo@bsc.es
Alberto Labarga
alberto.labarga@bsc.es
Barcelona Supercomputing Center (BSC)
Plaza Eusebi Güell, 1-3. 08034 Barcelona, Spain

Funding
This work has been funded by the Institute of Health Carlos III (project IMPaCT-Data, exp. IMP/00019), co-funded by the European Union, European Regional Development Fund (ERDF).