

# Swiss Health MetaData Repository

Eliane Maalouf, Alessio De Santo, Paul Cotofrei

Information Management Institute

September 13, 2019



# Context

- ▶ Health Information System (HIS)
  - ▶ Goal: generate **information** to improve health management decisions at all levels of the health system
  - ▶ Functionalities : **data** collection, **data** processing and reporting.
  - ▶ Strengthening HIS : promoting **data integration** through a well-described and coherent set of best practices.

HIS design needs to favor flexibility

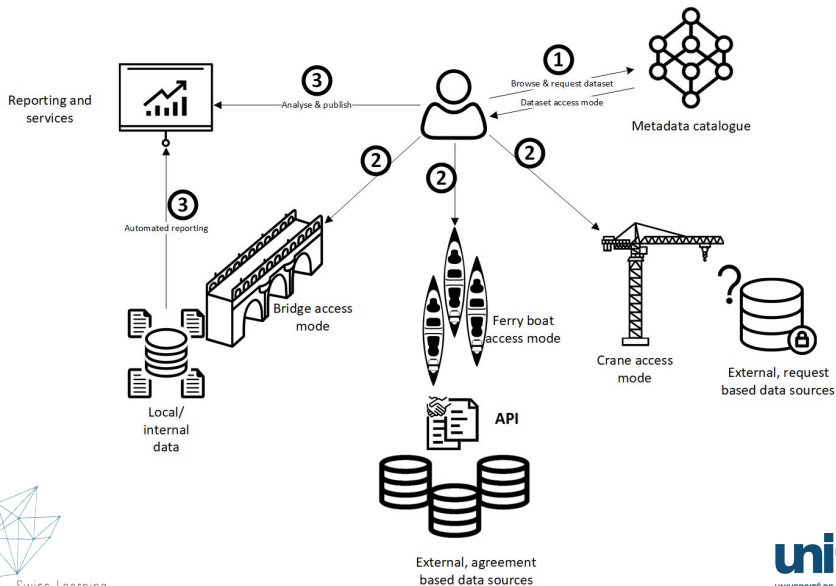
# Data Management

Data integration infrastructure:

- ▶ **Centralized approach:** all relevant data collected in a central location
  - ▶ Data Warehouse - uniform data model
  - ▶ Data Lake - no single data model
- ▶ **Decentralized approach:** accessing data at their original location
  - ▶ Federated database accessed using Web Services

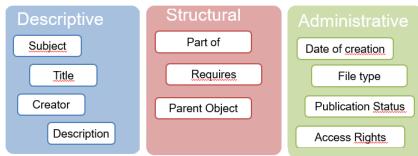
Choosing best approach for Swiss health Information system by considering technical, legal, cultural and political implications

# Hybrid infrastructure



# MetaData Repository

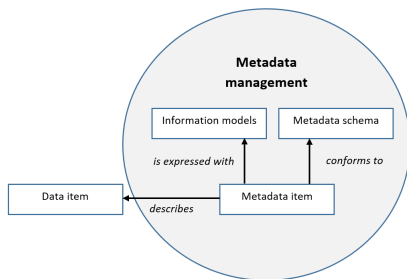
- ▶ Hybrid approach - necessity to have information (**metadata**) about data sources
- ▶ Metadata : a structured set of information used to describe a resource



- ▶ A central metadata repository - main element of the **metadata management infrastructure**

# Metadata management

- ▶ Metadata schema - provides consistency for description of digital resources (How)
- ▶ Information models - provides consistency for the meaning of description terms (What)
  - ▶ Syntax encoding scheme
  - ▶ Semantic (vocabulary) encoding scheme
- ▶ Metadata management - process of using metadata schema and information models in order to structure metadata items related to data
  - ▶ guarantees that the needed data are retrieved and accessed from the right databases and respecting each dataset specificities



# Metadata management

- ▶ Generic description metadata schema - Dublin Core
  - ▶ Core set of metadata elements : *Creator, Format, Date, Rights, Subject, Identifier, Language, etc.*
  - ▶ Rules and recommendation
    - ▶ Syntax encoding schema for date and time
    - ▶ Universal Resource Identifier (URI)
- ▶ Information models - abstract description expressing the meaning of items, formalized as:
  - ▶ *Nomenclatures* - a set of rules used for forming the names/terms
  - ▶ *Classifications* - concept of classes, with exhaustiveness and mutual exclusiveness
  - ▶ *Terminologies* - vocabulary of terms, including preferred terms, synonyms, definitions
  - ▶ *Taxonomies* - hierarchical representation using super and sub - concepts
  - ▶ *Ontologies* - domain representation, based on logically defined formalism, and using concepts, properties, relationships, instances or axioms as building blocks.

# Health information models (HIM)

## ▶ Health information models

- ▶ SNOMED / SNOMED CT (ontology)
  - ▶ Role/Domain : Lab results, Decision support for specialists in infectious diseases
- ▶ LOINC (terminology)
  - ▶ Role/Domain : Laboratories
- ▶ ICD11, ICF, ICHI (classifications/ontologies)
  - ▶ Role/Domain : Medico-economics, Post-market surveillance, Epidemiology
- ▶ WHO-ATC (classification)
  - ▶ Role/Domain : Post-market surveillance
- ▶ medDRA (terminology)
  - ▶ Role/Domain : Clinical trials, Post-market surveillance
- ▶ openEHR (terminology)
  - ▶ Role/Domain : Management, retrieval and exchange of health data in electronic health records

## ▶ Some criteria for optimal HIM

- ▶ Patient-centred care ontology basis
- ▶ Technology neutral model semantic
- ▶ International use
- ▶ Available in Swiss official languages



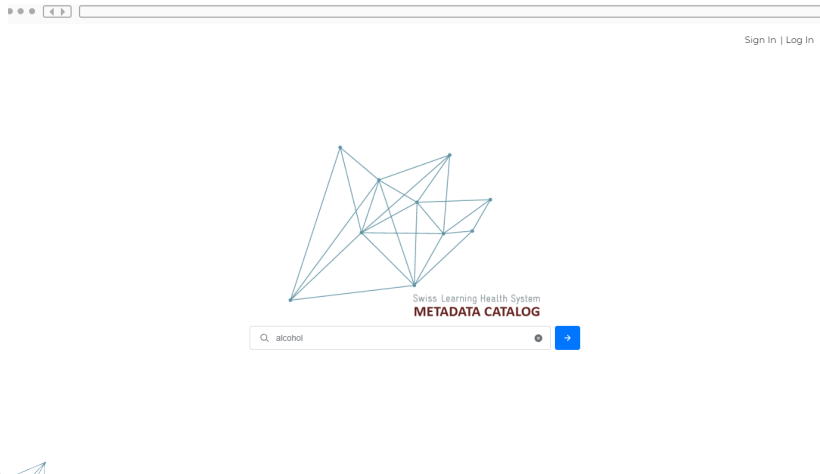
# Design principles of a central metadata repository as a key element of an integrated health information system

- ▶ Policy Brief (review phase)
- ▶ Stakeholders Dialog (in preparation)
  - ▶ Potential participants: data providers, policy makers, health researchers, developers
  - ▶ October - November 2019
- ▶ Prototype implementation (autumn 2020)

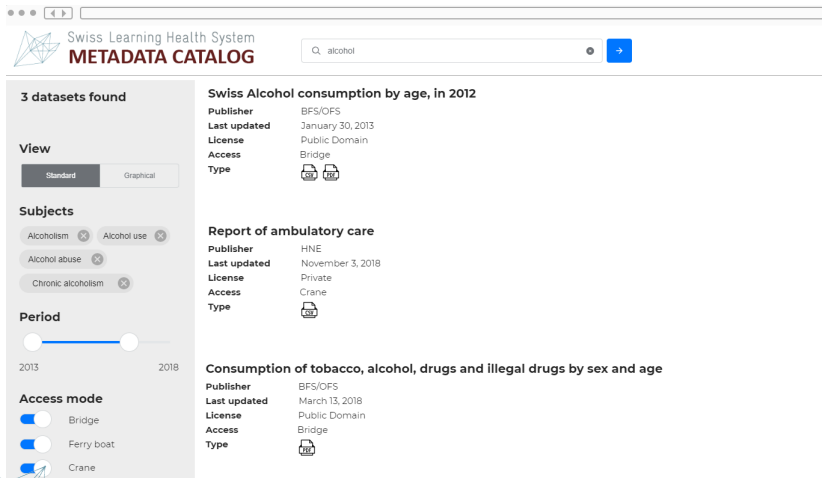
# Metadata repository prototype

- ▶ Implementation of a prototype for Swiss Health Ontological Supported MetaData Repository for creating, editing, controlling, deploying and finding reusable metadata.
- ▶ Components :
  - ▶ Back-end:
    - ▶ Semantic repository
    - ▶ Query translation
    - ▶ Semantic search engine
    - ▶ Semantic integration
    - ▶ Web server
  - ▶ Front-end:  
web-based interactive application
    - ▶ Browsing interface: browsing, searching, and exporting
    - ▶ Input interface: adding, updating and annotating metadata

# Sneak peek



# Sneak peek



Swiss Learning Health System  
**METADATA CATALOG**

Q, alcohol



**3 datasets found**


**View**  
Standard Graphical


**Subjects**  
Alcoholism Alcohol use  
Alcohol abuse  
Chronic alcoholism

**Period**  
2013 2018

**Access mode**  
 Bridge  
 Ferry boat  
 Crane

**Swiss Alcohol consumption by age, in 2012**  
Publisher BFS/OFS  
Last updated January 30, 2013  
License Public Domain  
Access Bridge  
Type  

**Report of ambulatory care**  
Publisher HNE  
Last updated November 3, 2018  
License Private  
Access Crane  
Type 

**Consumption of tobacco, alcohol, drugs and illegal drugs by sex and age**  
Publisher BFS/OFS  
Last updated March 13, 2018  
License Public Domain  
Access Bridge  
Type 

# Sneak peek

The screenshot displays the 'Swiss Learning Health System METADATA CATALOG' interface. At the top, there is a search bar containing the text 'alcohol'. Below the search bar, the interface indicates '3 datasets found'. On the left side, there are several filter sections: 'View' with 'Standard' and 'Graphical' options; 'Subjects' with filters for 'Alcoholism', 'Alcohol use', 'Alcohol abuse', and 'Chronic alcoholism'; 'Period' with a slider from 2013 to 2018; and 'Access mode' with options for 'Bridge', 'Ferry boat', and 'Crane'. The main area on the right features a complex network diagram. The central nodes are 'DS\_1' and 'DS\_2'. 'DS\_1' is connected to 'FSD', 'Bridge', 'Public Domain', and 'Tabular'. 'DS\_2' is connected to 'Crane', 'Treatal', 'Proprietary', 'HNE', 'Dataset', 'Alcohol consumption', and 'Gender'. Various other nodes and labels are scattered around, including 'Federal Statistical Office', 'Neuchâtel Hospitals', and several dates like '+04.06.2019+', '+04.06.2018+', and '+04.06.2019+'. The diagram uses different line styles (solid, dashed, dotted) to represent different types of relationships between the datasets and their associated metadata or sources.

# Thank you

University of Neuchâtel  
Faculty of Business & Economics  
Information Management Institute

## Contacts:

- ▶ Eliane Maalouf, PhD Student, [eliane.maalouf@unine.ch](mailto:eliane.maalouf@unine.ch)
- ▶ Alessio de Santo, PhD Student, [alessio.desanto@unine.ch](mailto:alessio.desanto@unine.ch)
- ▶ Paul Cotofrei, Assoc. Prof., [paul.cotofrei@unine.ch](mailto:paul.cotofrei@unine.ch)

