Project-specific RDF

Expanding the SPHN RDF schema

Dr. Vasundra Touré, Scientific Coordinator
Personalized Health Informatics, SIB Swiss Institute of Bioinformatics

A project of SAMW ASSM
Schweizerische Akademie der Medizinischen Wissenschaften
Académie Suisse des Sciences Médicales
Accademia Svizzera delle Scienze Mediche
Swiss Academy of Medical Sciences
SIB Swiss Institute of Bioinformatics

Outline

- Protégé, ontology editor
- SPHN strategy for projects
- Steps to extend the SPHN ontology with an example: FluidBalance
- Summary

Protégé, ontology editor

Protégé is a desktop ontology editor for building ontologies: https://protege.stanford.edu/
It supports the W3C standards.

A web version: WebProtégé is also available.
Enables collaborative development

DCC recommends and uses Protégé as ontology builder and editor.
DCC's web version: http://webprotege.dcc.sib.swiss/

SPHN strategy for projects

SPHN enables the extension of the SPHN ontology to fit the need of a project

Rules have been defined and need to be followed to ensure that **data FAIRness** can still be achieved

The different steps are written in the SPHN documentation:
https://sphn-semantic-framework.readthedocs.io/en/latest/user_guide/ontology_generation.html

(...and will be detailed in the following slides)

1. Load the template ontology in Protégé

Start from the template ontology to generate the project ontology.

The template ontology provides:

- Import of the SPHN RDF ontology extended
- Import of libraries used
- Prefixes already defined
- Pre-filled metadata (to be updated by the project)

Available at: https://git.dcc.sib.swiss/sphn-semantic-framework/sphn-ontology/-/blob/master/template_ontology

Requirements:

- Load in your computer the repository ‘test_project’ (provided in the training materials)
- Increase heap space of Protégé to -Xmx10000M (https://protegewiki.stanford.edu/wiki/Setting_Heap_Size)

Let’s work with a concrete example: Fluid Balance

Information is not in the Dataset yet. What is fluid balance?

Input of fluids, e.g.
infusion, drinking water

Nursing charts to
document input and output

Loss of fluids, e.g.
urine output

What information that needs to be captured?

Dataset guiding principles

- Separate elements with individual meanings instead of mixing pieces of information

~~Input value = 150 ml~~

value = 150
unit = ml
direction = input

Dataset guiding principles

- Separate elements with individual meanings instead of mixing pieces of information

~~Input value = 150 ml~~

value = 150
unit = ml
direction = input

- Use controlled vocabulary for value sets and fix the meaning of concepts

~~direction: input; output~~

direction: 251840008 | Fluid output (observable entity) |;
251992000 | Fluid intake (observable entity)|

Dataset guiding principles

- Separate elements with individual meanings instead of mixing pieces of information

~~Input value = 150 ml~~

value = 150
unit = ml
direction = input

- Use controlled vocabulary for value sets and fix the meaning of concepts

~~direction: input; output~~

direction: 251840008 | Fluid output (observable entity) |;
251992000 | Fluid intake (observable entity)|

- Check if you can re-use existing concepts - a meaning should be defined only once

fluid type identifier: 227566009
fluid type name: Soup (substance)
fluid type coding system and version: SNOMED-CT-2021-07-31

for fluid type, type = Code

re-use of concept Code from the existing SPHN Dataset

Fluid Balance in the Dataset

Find the concept here (as in 'Dataset' form): https://bit.ly/3j1KoM0

concept or concept compositions or inherited	concept name	description	type	standard	value set or subset	meaning binding SNOMED CT
concept	Fluid Balance	intake or output of fluid to calculate the fluid balance				251856003 | Fluid balance status (observable entity) |
composedOf	fluid type	type of fluid, e.g. urin, water, drinking water, blood	Code	SNOMED CT	child of: 105590001 | Substance (substance) |	
composedOf	value	amount of the intake or output	quantitative			
composedOf	direction	intake or output	Code	SNOMED CT	251840008 | Fluid output (observable entity) |; 251992000 | Fluid intake (observable entity)|	
composedOf	start time	datetime at which the fluid intake or output started	temporal			
composedOf	end time	datetime at which the fluid intake or output ended	temporal			
composedOf	unit	unit of the amount of the fluid intake or output	Unit	UCUM		

Next step: update of the ontology, following rules and conventions provided, to incorporate the concept of Fluid Balance and its related properties.

Summary

- We learned in this hands on session how to use Protégé in the context of SPHN with:
  - Loading the template ontology
  - Fitting the ontology annotations to the project
  - Adding Classes, Object and Data Properties
  - Reference to SPHN or other terminology Classes
  - Adding meaning binding to project Classes
  - Adding multiple Domains and Ranges to a property

References

SPHN Semantic Framework documentation - ontology generation user guide
https://sphn-semantic-framework.readthedocs.io/en/latest/user_guide/ontology_generation.html

SPHN ontology and template ontology:
https://git.dcc.sib.swiss/sphn-semantic-framework/sphn-ontology

Contact: dcc@sib.swiss

A project of SAMW ASSM Schweizerische Akademie der Medizinischen Wissenschaften Académie Suisse des Sciences Médicales Accademia Svizzera delle Scienze Mediche Swiss Academy of Medical Sciences
SIB Swiss Institute of Bioinformatics