EMIF Deliverable 11.2: First Version of Common Data Model and Associated Terminology Mapping

Executive summary

Executive Summary

Deliverable D11.2 builds on work reported in EMIF deliverable D11.1, to develop a centralized approach to storing, maintaining and using the formalized representations of clinical concepts across multiple data sources. The proposed approach is known as the EMIF Knowledge Object: each such object defines a small and re-usable data structure for a discrete clinical concept, along with constraints on the value domains (e.g. term lists) each fragment of the structure may hold, and which components of each structure are mandatory or optional. Each eventually envisaged Knowledge Object will also contain mapping information to each relevant data source registered within the EMIF federation, to facilitate efficient, scalable and reproducible query distribution of data set combination from multiple heterogeneous data sources to address agreed research questions. Because these definitions will be held in a centralised and singular representation, it is also intended for the repository or library infrastructure to support a variety of expert formats, so that the data dictionaries of tools such as tranSMART or Jerboa – which may be used for physical data integration – can be configured in a nearly automated and reproducible way from predefined and robust knowledge models.

The work reported here extends ideas described in D11.1, and should therefore be read in conjunction with that earlier deliverable. The core activities reported here are a pilot validation of the EMIF Knowledge Object approach using the example of the Victoria Stroop Test, which is a dementia psychometric assessment used by one of the EMIF vertical teams, and a proposal for brokering multiple terminologies using an illustrative example of diabetes mellitus which is a core concept for the other vertical team.

This deliverable reports an update of work in progress, and considerably more work is needed before this approach can be properly assessed as a suitable approach for semantic harmonization for EMIF. This work will be undertaken during year 2, and reported in deliverable D11.3.

Contacts

EMIF-Platform: Johan van der Lei – Nigel Hughes
j.vanderlei@erasmusmc.nl - nhughes@its.jnj.com