Executive Summary

The EMIF Platform architecture has been built around two main tracks: the EHR (Electronic Health Record) track; and the Cohort track. The rationale behind this separation is inherited from the Description of Work (DoW) and during the project the advantages for maintaining these two tracks in parallel have been higher than the ones for merging both. The most important reason was to avoid creating tools that are not the most adequate for a specific problem or research goal. But at the same time, several tools are shared by both tracks. The first is the EMIF Catalogue which is currently the entry point in the EMIF platform ecosystem.

The EMIF Catalogue aims to create a marketplace where data custodians can publish and share information regarding their clinical databases, and where biomedical researchers can search and find data sources based on their study requirements. This web application aggregates clinical databases of the same type into communities, providing a set of tools for end users (researchers) to find, filter and display information based on their specific requirements. At a deeper level, the platform can extract data from sets of databases and provide information regarding the suitability and feasibility of such studies.

Beside the Catalogue, this deliverable describes the work performed up to now in all technical components that are part of the EMIF platform architecture, such as Jerboa Reloaded (EHR), the Patient Selection Tool (Cohort), the Variable Selection Tool (Cohort), ATLAS integration (EHR), Knowledge Objects and tranSMART (Cohort), Taska (both tracks), and the PRRE (both tracks).

Moreover, we detail in this document the key points that are planned for the final year of the project and that will lead us to the final version of the EMIF Platform. These tasks are mostly related to study feasibility and study management, i.e. the layers 3 and 4 of the architecture (D14.7). Besides the continuous enrichment of the already developed tools, to cope better with upcoming user requirements, this effort will include, e.g., the capability to process distributed queries over harmonised databases complying with the Common Data Model (CDM), the integration of a study management system in the Catalogue, and a final solution for the Private Remote Research Environments (PRRE).

With this infrastructure in place, researchers will be able to: 1) publicly access the EMIF Catalogue; 2) join specific communities of their interest; 3) search and explore over a set of databases' fingerprints; 4) assess data suitability, for a considered study, in selected databases; 5) then go deeper with study feasibility based on selection criteria for patients and variables; 6) finally, conduct the study in a secured analysis environment.

Contacts

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