**Using Open Source Software and Open Data to Support Clinical Trial Protocol Design**

Nikolaos Matskanis, Joseph Roumier, Fabrice Estiévenart

CETIC, Rue des Freres Wright, 29/3, Charleroi 6041, Belgium
Tel: +32 71 490 761
nikolaos.matskanis@cetic.be; joseph.roumier@cetic.be; fabrice.estievenart@cetic.be

Clinical trials for drug repositioning aim at evaluating the effectiveness of existing drugs for new treatments. Their protocol design requires a clear specification of the trial hypothesis and is followed by the challenging step of selection and recruitment of eligible subjects. Both steps have complex information requirements and involve managing many interdependent parameters and details that are distributed over several data sources. The PONTE project (Efficient Patient Recruitment for Innovative Clinical Trials of Existing Drugs) has developed a platform that assists authoring, fast navigation and management of the Clinical Trial Protocol documents, provides services of translation of medical terms between different standard and open vocabularies and enables querying of the hospital Electronic Health Record (EHR) systems. In order to support the platform information model, several ontologies have been developed following a multi-layered and modular approach integrating external ontologies, nomenclatures, and vocabularies. Throughout its course the project relied on the availability of Free libre Open Source Software (FlOSS) to achieve its goals. The biomedical domain provides numerous FlOSS and Open Data resources for ethical and pragmatic reasons. Such licensing scheme is also very relevant for clinical trial design. The Linked Open Data initiatives offer access to a variety of medical and healthcare information including - but not limited to - drug profiles, diseases and their mechanisms, gene data sources and past trial results. The platform consumes this data, by providing mechanisms that assist navigation and querying of these open data sources. Additionally it provides intelligent aggregation of available information through semantic technologies and reasoning and presents the query results in a coherent and structured manner, assisting the clinical researcher's hypothesis investigations.

Keywords: semantic, FlOSS, LOD, EHR, clinical