

Real-time FAIRification of rare disease patient registry data

Martijn G. Kersloot^{1,2}, Annika Jacobsen³, Karlijn Groenen⁴, Denis Rakov², Leo Schultze Kool⁴, Ronald Cornet¹, Derk L. Arts^{1,2}

1. Amsterdam UMC, University of Amsterdam, Department of Medical Informatics, Amsterdam Public Health Research Institute, Amsterdam, The Netherlands
2. Castor EDC, Amsterdam, The Netherlands
3. Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands
4. Department of Radiology and Nuclear Medicine, Radboud university medical center, Nijmegen, The Netherlands

Corresponding author: Martijn G. Kersloot (m.g.kersloot@amc.uva.nl)

Abstract

Rare diseases (RDs) affect fewer than 5 people in 10,000. Due to this low prevalence, individual European Union (EU) member states and registries have difficulties constructing databases large enough to perform valid research. Interoperability (I) between registries would support the construction of these databases; therefore, the EU defined 16 common data elements to be registered by all RD registries. We aimed to make this common set FAIR for the registry of vascular anomalies (VASCA), to also ensure findability (F), accessibility (A), and reusability (R). FAIR datasets allow for pooling of data from multiple sources, providing researchers with more data for their research. In this demonstration, we show how an online data capture tool (Castor EDC) allows real-time FAIRification of collected registry data. A semantic model and Twig template are used to automatically transform data, upon entry into the eCRF (electronic Case Report Form), into RDF (Resource Description Framework) triples, a machine-readable format, to make the data FAIR. This implementation can be re-used and adapted for eCRFs of other registries or clinical studies.

Keywords

electronic data capture, FAIR Data Point, metadata