

MOLGENIS catalogue - towards a unified application for data and sample catalogues

Alexander Kellmann^[0000-0001-6108-5552], David van Enckevort^[0000-0002-2440-3993], Marije van der Geest^[0000-0002-1919-6547], Aneas Hodselmans^[0000-0002-2924-918X], MOLGENIS Team, and Morris A. Swertz^[0000-0002-0979-3401]

Genomics Coordination Center, Department of Genetics,
University Medical Center Groningen, Groningen, The Netherlands
{a.j.kellmann, m.a.van.der.geest, m.a.swertz}@umcg.nl
{a.w.hodselmans, d.van.enckevort}@rug.nl
<https://www.systemsgenetics.nl/group/swertz>

Abstract. To implement the ‘Findability’ of ‘FAIR’, many are now creating catalogues to promote reuse of valuable data and samples. However, isolated development of these different catalogues is a large burden and hinders interoperability to pool information together in more comprehensive catalogues. Therefore we compared existing models from many different sources and present MOLGENIS/catalogue, a unified catalogue application that generalizes best practices from many biobank, population, (rare) disease, and cohort catalogue communities, e.g. BBMRI, RD-connect, LifeCycle, Lifelines, Maelstrom. The resulting model seeks to provide an extensible model to quickly create a new catalogue based on the MOLGENIS software.

Keywords: FAIR · Molgenis · Catalogue · Biobanks · Cohort studies

1 Materials & Methods

We compared catalogue data models from many existing initiatives and standards (MIABIS) and consulted our many stakeholder projects to identify common data structures and elements. Then we implemented a unified but modular data model to accommodate the needs of different communities using MOLGENIS software suite, optimised for flexible data models and independent development of web user interface ‘apps’, to allow ‘agile development’ with frequent releases.

2 Results

A first version of the unified catalogue data model has been created for evaluation. In addition we have defined a list of user requirements, based on broad community consultation. This ‘backlog of user stories’ is now being prioritized by

the stakeholders to create a community supported roadmap for implementation. We designed MOLGENIS/catalogue to be highly configurable and extensible to attract community contributions in data model extensions, tools and user interfaces. An online Demo of the MOLGENIS catalogue model can be found here:

<https://testmodel.test.molgenis.org>
user: demo
password: Feedback

3 Conclusions

We believe that focus of efforts into one unified catalogue application will greatly improve development speed and promote interoperability, and invite the biobank community to feedback and steer its development to meet all their needs. This will help us, to make the model even more useful. We're working with various semantic communities, such as SWAT4HCLS, CINECA, EJP-RD, to improve the semantic enrichment and make cross domain federated searches a reality.