



Contribution ID: 31

Type: Poster

Streamlining of a metadata-based computational and statistical analysis pipeline

Efficient, extendible knowledge representation and reasoning enabling data mining and data visualization is becoming increasingly important in systems biology due to the constant growth, accumulation and availability of vast multi-variate and multi-modal data sets rendered possible through former and recent advantages in e.g., high-throughput experimental measuring techniques, i.e., omics and related technologies. To foster and to make the transfer from raw data to knowledge representation, thus giving the data meaning, ultimately through a contextual knowledge management, additional information describing the data, metadata is required. A prerequisite to allow for an efficient and extendible knowledge representation are a controlled vocabulary and ontology. By following established industry-standard modelling approaches, i.e., by defining a conceptual data model and in turn a logical data model, defining a physical data model resulting in data integration into a graph-based data base for knowledge representation we strive to streamline the computational workflow of metadata management and integration into GUI-based reactive, user-friendly custom statistical analysis tools (R Shiny). Making use of proper metadata representation to extract knowledge from the data sets, derived analysis and visualization tools can be developed in addition. To facilitate the collection of metadata we emphasize our decision for a low-entry barrier solution to input metadata and link data for users through an Excel spreadsheet (but not limited to) integrating readily into previously mentioned analysis tools and demonstrate use-cases of our web-based analysis pipeline for a repeatable reproducible omics analysis based on metadata.

Primary author: SEEP, Lea (Faculty of Mathematics and Natural Sciences, Rheinische Friedrich-Wilhelms-Universität Bonn)

Co-authors: Dr GREIN, Stephan (Faculty of Mathematics and Natural Sciences, Rheinische Friedrich-Wilhelms-Universität Bonn, 53115 Bonn, Germany); Prof. HASENAUER, Jan (Faculty of Mathematics and Natural Sciences, Rheinische Friedrich-Wilhelms-Universität Bonn, 53115 Bonn, Germany)

Presenter: SEEP, Lea (Faculty of Mathematics and Natural Sciences, Rheinische Friedrich-Wilhelms-Universität Bonn)

Session Classification: Posters