Contribution ID: **80**   Type: **POSTER&PITCH**

From Schema to Questionnaire: Humanizing data description

Monday 4 November 2024 16:00 (1 hour)

Enriching data with metadata is a key concept for the data output of scientific research to be FAIR. Data processing software and custom code often do not support the annotation with metadata out-of-the-box or the usage process does not mandate it. This confronts data creators and maintainers with challenges to annotate their data. From a Human Machine Interface (HMI) perspective, metadata forms are a superior way to support this process compared to manually editing the respective data.

Depending on the use case and the user's role, different ways to generate and use forms for metadata may be helpful. A researcher without programming skills does benefit from ready-to-use services when entering relevant metadata. Software developers might instead prefer a software library that provides forms based on the metadata schema. Managers and those responsible for processes are especially interested in solutions that fit into existing workflows.

For each of the above scenarios we will present a list of criteria to categorize available solutions. For example, a developer might be interested in the particular front end technology that is used by a library, while for a data provider the export formats of a given service are much more relevant. Of course several features do apply to all the use cases, like the capability to allow different profiles of metadata or a real time validation of inputs.

We present a review of available services, tools and software libraries that help with form based metadata annotation. The list will be published on Zenodo to summarize the capabilities in regard to a collection of features that will help to find the most suiting solution. We hope this can empower more people to make their data FAIR.

One solution is Research Data Management Organiser (RDMO), a web application to assist in the structured planning and administration of the data in a scientific project. The interaction with RDMO is based on a set of questionnaires, which can be attached to a project and filled in by data providers.

The so gathered information can be cast into textual forms suitable for funding agencies'report guidelines, or it can be used as metadata for generated datasets.

The integration of RDMO in a data providing and curation workflow starts with generating a questionnaire from a given metadata schema (e.g. the Open Energy Platform metadata standard). Data providers willing to publish on this platform have to answer specific questions concerning their dataset. RDMO additionally applies existing super-ordinate metadata (author and affiliation, funding agency, etc.). Those information sets are inherited from a designated project.

After gathering this information in a schema-compliant file, the provider is assisted in publishing them on a generic metadata store with a download link to the actual data.

The main advantage is that the data provider can focus on answering questions and needs not struggle with machine readable files. Additionally, the data

provider is guided through the whole process. File system monitoring can also be used to increase the reliability by motivating the provider to publish newly detected data.

Please specify "other"

In addition, please add 3 to 5 keywords.

FAIR, Metadata, Questionnaire, Process Reliability

Please specify "other"

For whom will your contribution be of most interest?

Data professionals and stewards

Please assign yourself (presenting author) to one of the following groups.

Researchers

Primary authors: Mr KOUBAA, Mohamed Anis (Institute for Automation and applied Informatics); SCHMURR, Philipp (KIT IAI)

Co-authors: SCHMIDT, Andreas (IT4EDM/IAI); STUCKY, Karl-Uwe (KIT); LIU, Nan; SUESS, Wolfgang (KIT)

Presenters: Mr KOUBAA, Mohamed Anis (Institute for Automation and applied Informatics); SCHMURR, Philipp (KIT IAI)

Session Classification: Poster Session C

Track Classification: Connecting research data: 4. Metadata annotation and management