deRSE25 and SE25 Timetables



Contribution ID: 12

Type: Talk (15min + 5min)

SMECS: A Software Metadata Extraction and Curation Software

Tuesday 25 February 2025 15:10 (20 minutes)

Metadata have shown to be one of the success factors for the so-called FAIRification of research software, especially in improving the findability and reusability of research software [1], [2]. Creating high-quality metadata can be resource-intensive [3]. Moreover, users often find it challenging to utilize metadata effectively for retrieval [4], [5]. To support researchers from various domains in creating metadata for their research software, we developed the Software Metadata Extraction and Curation Software (SMECS).

SMECS lowers the barrier to create metadata for research software by combining extraction from existing sources with easy-to-use curation of metadata. SMECS is a python-based software that is available open-source . As a first step, SMECS can extract existing metadata from different sources. Afterwards, the tool allows the researcher to curate the metadata and add additional information. Finally, SMECS provides the researchers the metadata as a JSON file in the CodeMeta format which is a common standard for research software metadata [6].

For research software, a lot of metadata is already presented in online repositories, e.g., on GitHub or GitLab. Sometimes, even structured metadata is available in CFF [7] or CodeMeta files. Therefore, after the user provides a link to a repository, SMECS extracts as much metadata as possible from the API of the corresponding repository, e.g., name, contributors. All extracted metadata is mapped to CodeMeta.

In the second step, SMECS presents the researcher the extracted metadata in a user interface to allow further curation by the researcher. The researcher can check and change the extracted metadata as well as add additional metadata. Finally, the researcher can export the metadata as a CodeMeta file.

SMECS was designed with an emphasis on User-Centered Design (UCD), ensuring that the needs, preferences, and behaviors of users were prioritized throughout the development process. The objective of User-Centered Design for SMECS is not only to develop useful metadata extraction and curation but also to enhance user satisfaction and task performance within SMECS. To strengthen the overall usability and better meet user needs, SMECS was improved based on feedback from usability experiments during the iterative design process. In the first usability experiment, participants interacted with the software by completing a series of predefined tasks. Following this, each participant filled out the System Usability Scale (SUS) questionnaire to assess their experience with the tool. Finally, semi-structured interviews were conducted with each participant to gather qualitative data regarding their user experience with SMECS, providing deeper insights into their interactions and perceptions of the tool's usability.

Our results reveal that, SMECS includes a good user experience. To simplify the metadata creation for researchers even further, we plan to expand SMECS' capability to extract metadata from a wider range of sources (e.g., CFF and CodeMeta files including functionalities of HERMES [8], and README files by including functionalities of SOMEF [9]).

In our talk, we will present the current state of SMECS and discuss SMECS with the audience.

References

- [1] https://doi.org/10.1016/j.patter.2021.100222.
- [2] https://doi.org/10.3233/DS-190026.
- [3] https://doi.org/10.1002/meet.2009.1450460397.
- [4] https://doi.org/10.1177/0165551513507405.
- [5] https://doi.org/10.1016/j.lisr.2005.01.012.
- [6] http://ssi1.eprints-hosting.org/id/eprint/2/

[7] https://citation-file-format.github.io/

[8] http://arxiv.org/abs/2201.09015

[9] https://doi.org/10.1109/BigData47090.2019.9006447

I want to participate in the youngRSE prize

Primary author: FERENZ, Stephan Alexander (Carl von Ossietzky Universität Oldenburg; OFFIS)

Co-authors: JAFARBIGLOO, Aida (OFFIS); NIESSE, Astrid (Carl von Ossietzky Universität Oldenburg); Dr WERTH, Oliver (OFFIS)

Presenter: FERENZ, Stephan Alexander (Carl von Ossietzky Universität Oldenburg; OFFIS)

Session Classification: Metadata in Research Software

Track Classification: Data and Software Management: software metadata