



Contribution ID: 44

Type: **Workshop or Hackathon**

Software metadata extraction for Software Management Plans

Thursday, March 7, 2024 1:30 PM (1h 30m)

Software metadata extraction for Software Management Plans

A hands-on workshop session at DE-RSE 2024

Software Management Plans (SMPs) help Research Software Engineers (researchers who develop code as part of their research or software engineers who support research activities) to oversee some of the activities during the software development lifecycle. Such activities could support (i) researchers in developing better software by following some minimum good practices, and (ii) software engineers in adopting some practices that might not be common outside research (e.g., archiving releases, providing citation information).

In this workshop, we will introduce (research) software metadata and its connection to SMPs, a tool for metadata extraction from GitHub repositories (GitHub, paper, website), and the Software Management Wizard (SMWizard—a tool to facilitate filling in ELIXIR SMPS, see preprint and web page). After the introduction, we will have a hands-on session to work in small groups to try and improve the SMW and the metadata extractor (e.g., suggesting improvements to the SMP/SMW/metadata extraction, using the tools to improve your own GitHub repo machine-readability, implementing a new data integrator for the Wizard, improving the metadata extraction). The hands-on will finish with feedback from participants, followed by a wrap-up from the workshop organizers.

Organizers

- Leyla Jael Castro. ZB MED Information Centre for Life Sciences. ORCID:0000-0003-3986-0510
- Stephan Ferenz. Department of Computer Science at University Oldenburg. ORCID:0000-0001-9523-7227
- Daniel Garijo. Ontology Engineering group at Polytechnic University of Madrid. ORCID:0000-0003-0454-7145
- Marek Suchánek. Department of Software Engineering at Czech Technical University in Prague. ORCID:0000-0001-7525-9218

Tentative agenda

| Time | Activity | Responsible |

| — | — | — |

| 10' | Welcome | Leyla Jael Castro |

| Introductory session | | |

| 10' | Software metadata | Stephan Ferenz |

| 10' | Machine-actionable SMPs | Leyla Jael Castro |

| 15' | Software metadata extraction | Daniel Garijo |

| 15' | SMWizard | Marek Suchánek |

| Hands-on session | | |

| 90' | In groups, work on one of the following topics: (i) Create your SMP with the SMWizard and brainstorm on improvements, (ii) create an integrator for the SMWizard, (iii) use the metadata extraction tools to produce Codemeta and/or Bioschemas metadata files and brainstorm on improvements, (iv) extend or develop new functionality for the metadata extractor, (v) your own idea | All participants |

| 20'| Feedback from groups | One participant per group |
| 10'| Wrap-up | Leyla Jael Castro |

Slot length

Workshop (3h)

Primary authors: GARIJO, Daniel (Ontology Engineering group at Polytechnic University of Madrid); CASTRO, Leyla Jael (ZB MED Information Centre for Life Sciences); SUCHÁNEK, Marek (Department of Software Engineering at Czech Technical University in Prague); FERENZ, Stephan (University Oldenburg)

Presenters: GARIJO, Daniel (Ontology Engineering group at Polytechnic University of Madrid); CASTRO, Leyla Jael (ZB MED Information Centre for Life Sciences); FERENZ, Stephan (University Oldenburg)

Track Classification: Research Software (legacy): Research Data Management/ Research Software Management