Contribution ID: **57** Type: **Talk (15min + 5min)**

Research software versus mundane software: The gap between software usage and software mentioning

Thursday 27 February 2025 09:40 (20 minutes)

For the vast majority of researchers across disciplines, software use is an everyday practice, and data analysis is not the only way of scientific sensemaking with software. The talk presents survey results showing that research pipelines are populated with diverse types of software –among them software tailored for research purposes („research software“) as well as software covering broader tasks which are not specific to science („mundane software“). Laying the results against recent literature about software mentions in science reveals a gap between using software for research and mentioning it in publications: mundane software is widely used in science but goes often unreported.

The assumption that software use is epistemically relevant is embedded in the discourse on research software engineering. However, speaking from a socio-epistemological perspective, the role of software in general for *research as a process* has been overlooked so far: Depending on the field of research, the roles of research software versus mundane software in achieving results may vary. For example, while data driven research might depend more on the software used for analysis (e.g. an R package), non-empirical research might depend more on software for literature management (e.g. Citavi). These considerations led to the following research question: *Which software do researchers use for which purpose?*

The research question is approached with data from a trend survey on research data infrastructures which was conducted in May 2024 in the context of Base4NFDI (Measure 4.3). The representatively invited survey yielded 1033 answers from German researchers across disciplines and fields. The participants were asked with open text fields about which software and platforms they use for the purposes of a) text production and literature management, b) data collection and analysis, c) knowledge organization and search engines and d) communication and project management. The answers were manually coded into ~ 800 software/platform names in ~3400 software/platform-purpose pairs.

Besides the gap between software usage and software mentioning, the survey shows that in all but the data parts of the research pipeline, most researchers use proprietary software and platforms. As proprietary software has become more networked in the recent decade, this poses a danger of lock-in and a challenge to FAIRification of the research process in general which cannot be overestimated.

I want to participate in the youngRSE prize

no

Primary author: HARTSTEIN, Judith (Deutsches Zentrum für Hochschul- und Wissenschaftsforschung)

Presenter: HARTSTEIN, Judith (Deutsches Zentrum für Hochschul- und Wissenschaftsforschung)

Session Classification: RSE research

Track Classification: Research Software: RSE research