

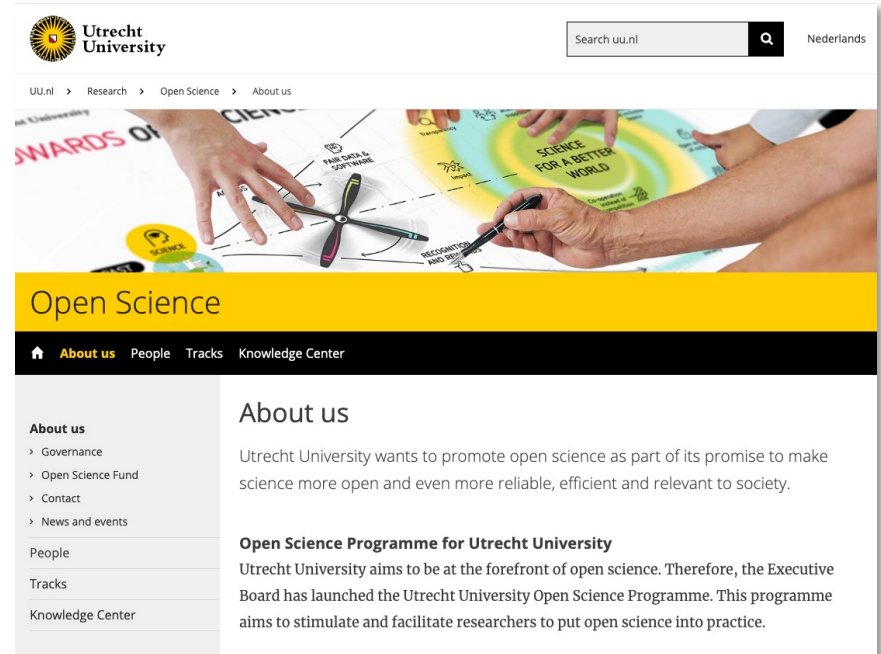
Analysing research software repositories to understand better the current state of the practice in research software development (or: the SWORDS project)

Anna-Lena Lamprecht, Jonathan de Bruin, Keven Quach, Akshay Devkate, Alexander Riedel

IRSER Meeting 16. & 17. January 2024, online

How this work started

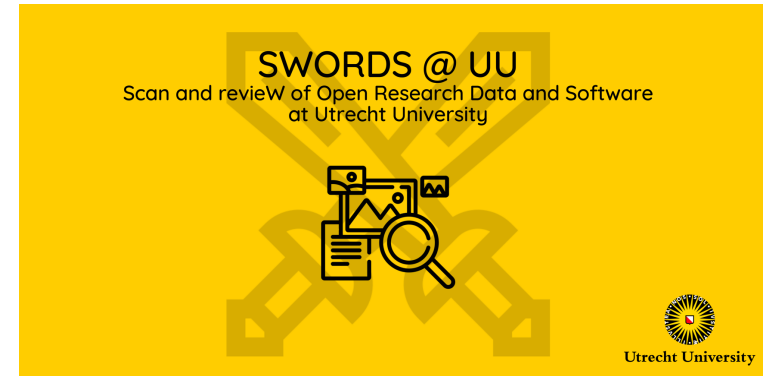
- 2020 at Utrecht University
- Discussion with J. de Bruin (Information and technology services)
- *“What kind of software development support and training should the university offer to its researchers?”*
- *“Well, what do they need?”*
- *“Hmm, what do they actually do currently?”*



Screenshot from <https://www.uu.nl/en/research/open-science/about-us>

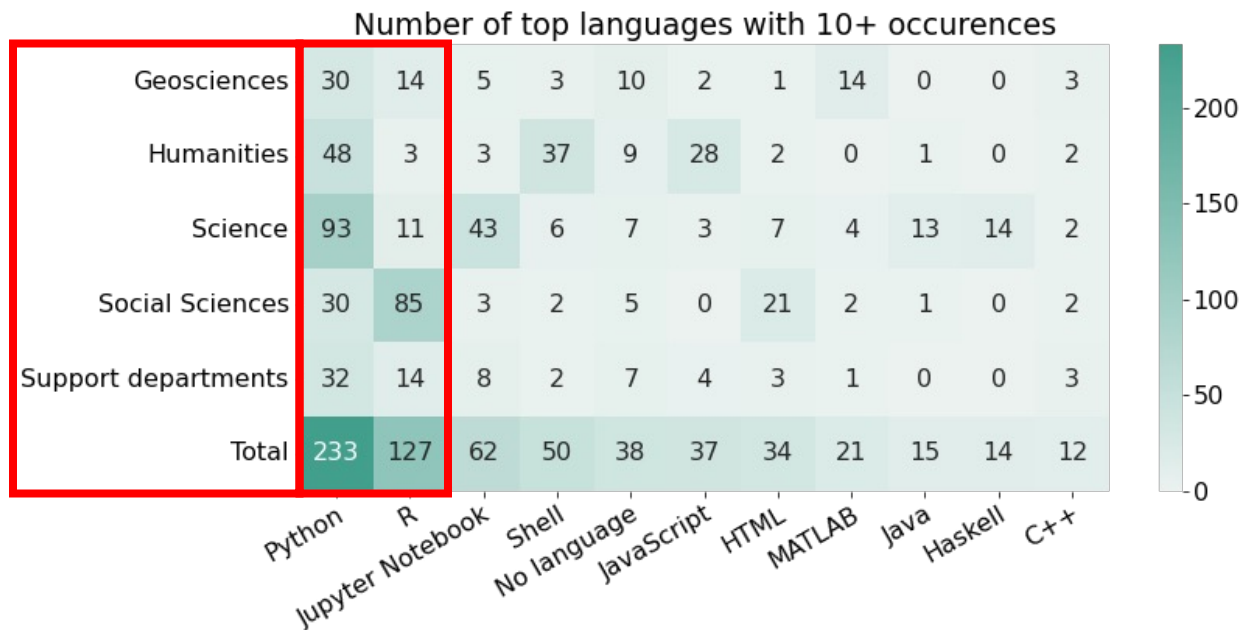
SWORDS@UU

- Start of the SWORDS project at UU
(Scan and review of Open Research Data and Software)
- Pipeline for getting insights about how researchers develop, manage, and publish software on GitHub
- See <https://www.uu.nl/en/news/how-fair-is-the-software-landscape-of-utrecht-university> for a gentle introduction
- Or <https://github.com/UtrechtUniversity/SWORDS-UU> for the full repository



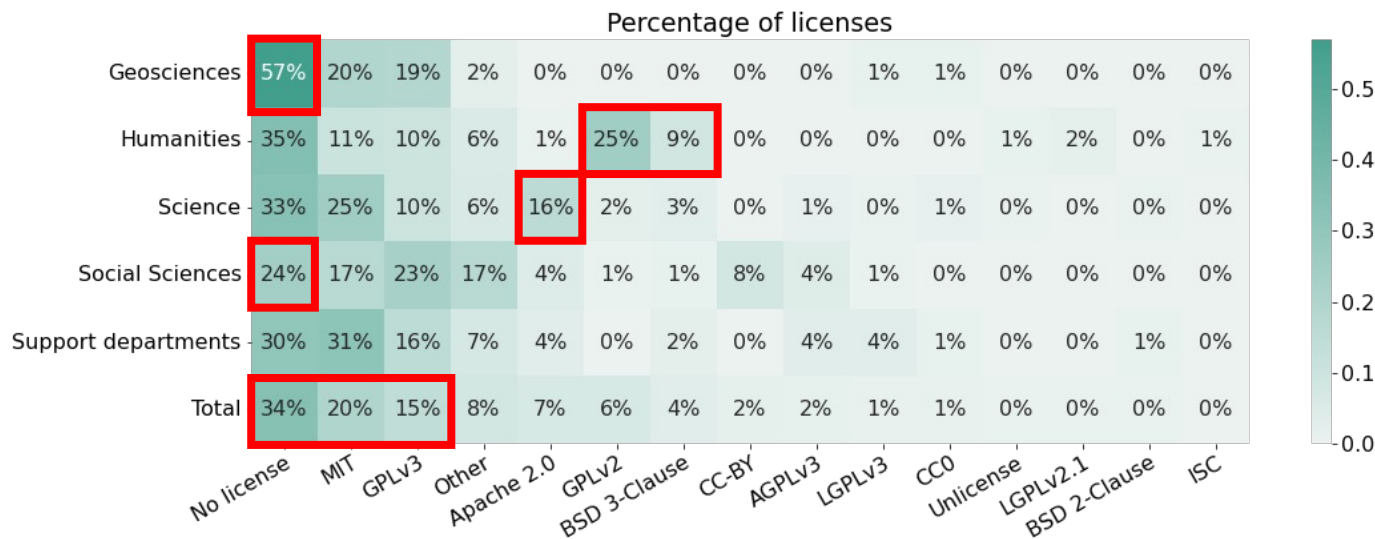
SWORDS@UU: Language usage

Results for 708 research software repositories



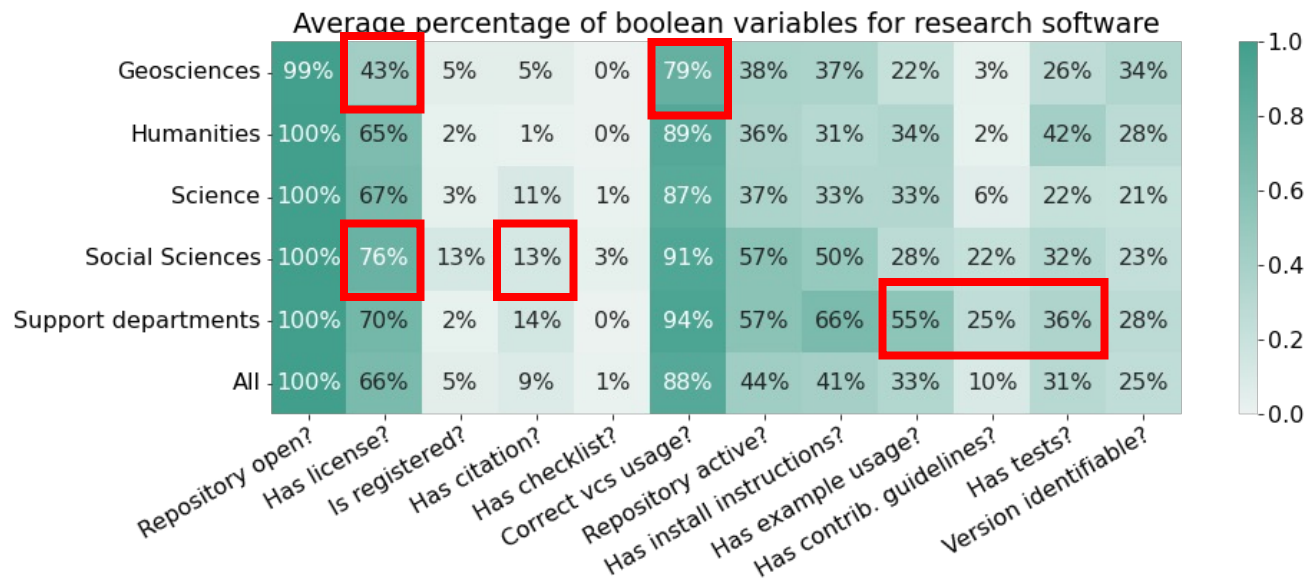
SWORDS@UU: License usage

Results for 708 research software repositories



SWORDS@UU: Average percentage FAIR variables

Results for 708 research software repositories



SWORDS@UU: Derived recommendations

- Foster local programming communities (R/Python Café)
- Two kinds of resource recommendations:
 - Training for software engineering & hands-on usage of relevant tools
 - Best FAIR practices document
- Highlight FAIR & high-impact research software
- Create yearly organisation-specific FAIRness reports of research software

Check out Keven Quach's Master Thesis "Mapping Research Software Landscapes through Exploratory Studies of GitHub Data" for further details:
<https://studenttheses.uu.nl/handle/20.500.12932/43162>

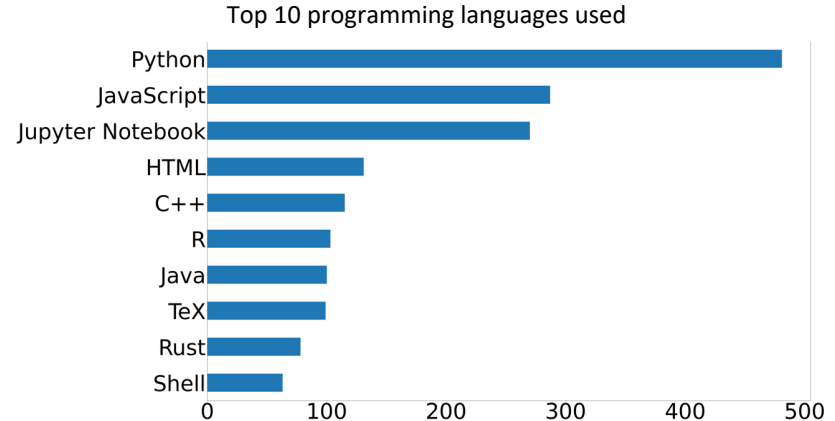
SWORDS template

- SWORDS prepared for use at other organizations
- <https://github.com/UtrechtUniversity/SWORDS-template>
- Mind that the first parts of the pipeline are the hardest
 - Identifying users (how do you know the relevant persons' GitHub profiles?)
 - Collecting relevant repositories (how to tell research software repositories apart from others)
 - Requires some manual work...



SWORDS@UP

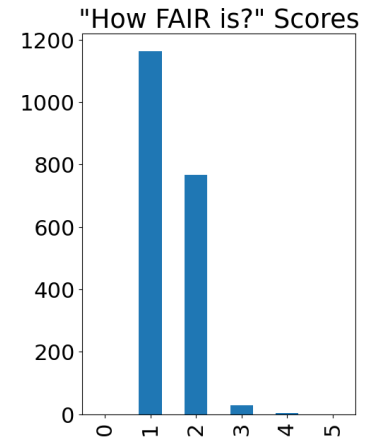
- Part of the PhD research of Akshay Devkate
- Applying the SWORDS template to GitHub repositories of University of Potsdam researchers
- Preliminary results for ca. 2000 GitHub repositories



Five criteria from

<https://fair-software.eu/>:

1. In publicly accessible repository with version control
2. Has licence
3. Registered in domain registry
4. Enable citation
5. Software quality checklist used



SWORDS @ GIT.UP

- Bachelor thesis project of Alexander Riedel
- Adapting the SWORDS pipeline to work with Git.UP, the University of Potsdam's GitLab instance
- Easier to identify users and repositories
- Analysis pipeline needs to be adjusted to work with GitLab



Summary

- Open Science and SWORDS@UU
- SWORDS template
- SWORDS@UP
- SWORDS@Git.UP

Lots of ideas for future work

- Integrate more elaborate FAIRness assessment tools (e.g. BSC's FAIRsoft)
- Create more datasets and compare groups (other organizations, domains)
- Cross-exchange with DLR (analysis methods and datasets)
- Extend analyses (e.g. Akshay's thesis project: assess use of testing techniques and other methods for software quality assurance)
- *Please get in touch if you are interested to collaborate on this topic!*

Thanks for listening!