

Research data repository for students' own results



In the Department of Physics of University of Helsinki, the masters level training of physicists have traditionally included extensive laboratory experiments, their documentation and reporting. This pilot is for including data publication and curation of the experiment results in the laboratory courses: storing the observations, together with relevant metadata into a repository, where the course assistants would have access. The students would then learn to publish and document their data as a normal part of scientific workflow. Naturally, it would be needed also to include methods to “cite” the data sets using a PID offered by the system.

The Scientific Challenge

One of the biggest issues in data sharing is a cultural one. Current research paradigm does not necessarily consider data sharing and curation as a part of normal scientific process. Teaching this as a part of normal course work is the way to get the message across to new scientists, thus creating possibility of new generation of scientists who do not need to be taught and forced to publish their data - no more than they are to publish their results. Taking data publication as a normal part of course work is the key. The data publication should however be realistic, easy and flexible. The system should be similar which are used in long-tail of the research data applications and the overall system should also serve the overall course work. For this reason, the system should be capable of authentication, team work, controlled evaluation and to be completely citable. The annual number of students in the initial phase is less than hundred and the data amounts are small.

Who benefits and how?

One of the biggest issues in data sharing problematics is the research culture, where data sharing is not considered as a part of normal activity. Teaching this as a part of normal course work is the way to get the message through to new scientists.

Technical Implementation

As for what regards B2SHARE, work has been carried out towards the B2Share 2.0-version, which is expected to be the improvement from the B2SHARE 1.x. We currently have demo-site up (and working with the scripts for migrating old data to the new system).

It was originally planned to use the B2SHARE 2.0 for Helsinki University piloting, so in that sense there has been progress (although finishing the second version has taken slightly longer than expected). B2DROP is planned to be used as the initial data storage system, and a specific implementation B2DROP Finland has been installed to CSC. Authorization and student grouping in B2ACCESS are being evaluated for technical requirements and practical implementation.

Preliminary Results

The initial requirements were collected from the teachers, University technical staff and selected students. The planning meetings included these stakeholders and the technical implementation team from CSC. Combination of B2DROP and B2SHARE were selected as the main technological base of the use case, with B2ACCESS providing authorization routines if needed. Key issues were identified, particularly the need of specific B2SHARE metadata for the course use, and need of group authorization, with different levels of access. The authorization changes were seen as important, as they provide course teachers and assistants access to the student results, and to make it possible for the students to submit and access their datasets as a group.

The first draft of template is available at the B2SHARE training site, and it is intended to be used by the first test versions of the repository when the course starts. Forward testing with the demo-site is continuing, to get look-and-feel of that version, the plan is then to continue with the production version when it is ready.

Contacts

- Ari Asmi, University of Helsinki, Finland. [ari.asmi\(a\)helsinki.fi](mailto:ari.asmi(a)helsinki.fi)

Further Information

[Scientific Challenges behind the Pilot](#)

Category:

[Data Pilot Communities](#)

[Read more](#)