sciecominfo
Nordic-Baltic Forum for Scientific Communication

# KUBIS DATAVERSE NETWORK

## A repository for primary data at the The Royal Library / Copenhagen University Library and Information Service

Thea M. Drachen, Asger V. Larsen, Bertil F. Dorch

*The KUBIS Dataverse Network (https://data.kb.dk/dvn) is a digital archive for scientific primary data for use by researchers at The University of Copenhagen. KUBIS Dataverse Network is available to researchers, departments and institutes at the University and research groups with an affiliation to the University of Copenhagen.*

### Some of our researchers expressed a demand for this kind of service

The KUBIS Dataverse Network was established on the basis of the results of a "Report of a qualitative study of social sciences collection, use and sharing of primary data" (in Danish) (http://www.hprints.org/hprints-00451000). This report concluded, among other things, that there was an urgent need among a considerable part of the surveyed researchers for backup procedures and safe archiving of research data. Furthermore that there was a need to be able to archive data in many different formats: audio, images, text, numbers, etc. In addition there was a desire amongst scientists to share data with others and that the researcher him or herself could control who had access to these data.

### What is the Dataverse Network?

The KUBIS Dataverse Network serves as an additional backup system, which archives data in a structured form for both short and medium term preservation. It can also serve as a way of sharing data. Each researcher/group can create his/her own Dataverse in the KUBIS Dataverse Network and can store and process the data, and if he/she chooses to share his/her data. The open source application was developed by The Institute for Quantitative Social Science (IQSS) at Harvard University (http://thedata.org/, link to the IQSS Dataverse Network: http://dvn.iq.harvard.edu/dvn/). The KUBIS Dataverse Network is not hosted in the cloud, but on secure servers at the Royal Library / Copenhagen University Library and Information Service, using 256-bit encryption.

The Dataverse Network is hierarchically organized with a number of dataverses (each created by e.g. individual researchers, university institutes, or a research project groups). Each dataverse creator has the option to create one or more “studies” in that particular dataverse (as a top level organization of the contents). Files can be public or restricted as default, but each file can be set to public, restricted to all, or restricted to all but the persons or groups granted access by the uploader.

Each study is given a persistent identifier (Handle) at upload and version history is saved, enabling the author to give permanent links/handles to the uploaded data. When creating a study and uploading a file, the uploader will be asked to fill in metadata for that file, which greatly enhances the possibility of reusing data at a later stage by the same researcher or others he/she might give permission. These metadata are typically filled in at the study-level, so the amount of metadata one has to fill in for individual files is limited.

The amount of metadata that can be attached is vast, but just a few fields are mandatory (such as name and creator). There are no restrictions to the file format in a Dataverse, what you put in is what you get out (only SPSS file formats SAW/POR will change to TAB).

### … and what is it not

The Dataverse is not a virtual research environment or an analysis tool. Apart from some statistical analyses (using R), you need to download the file(s) you want to analyse.

You cannot bulk upload to the Dataverse Network. Each file needs to be uploaded separately. For files you want to share in a collective group, you might want to zip the files to cut down the number of times you need to upload.

Accessing the Dataverse Network the first time can be cumbersome. Once you get the hang of the very hierarchical organization of the Dataverse Network it is fairly easy to navigate in it, but it is not very web 2.0.

### Derived services

We got a request from a research group at the Department of Cross-Cultural and Regional Studies at the University of Copenhagen. They had data they wanted to share and research they wanted to promote, they asked us for help and we ended up giving them a

combination of a webpage, finding and uploading to repositories the publications for which we could get copyright clearance, and setting up a dataverse for the Department for their data. Read more here (http://shkaratmsaied.tors.ku.dk/).

This was such a success, that we started up Research@KUBIS (http://libguides.culis.kb.dk/researchatkubis), which uses the Dataverse Network as one of its three components. It is a help towards 1) building a project site, 2) promoting data and 3) promoting publications.

The project site platform (LibGuides) is especially interesting for researchers in cross-institutional collaborations to use, since project collaborators are not required to be affiliated with a certain University or Institute to be given editor-rights to a project page. And the system is WYSIWYG (What You See Is What You Get) and very easy to work in.

We have only just started promoting the KUBIS Dataverse Network and the Research@KUBIS services, but an increasing number of researchers and research groups have shown an interest in using the systems.

**Thea M. Drachen,** Research Support Service, The Royal Library / Copenhagen University Library and Information Service, forskerservice@kb.dk

**Asger V. Larsen** Research Support Service, The Royal Library / Copenhagen University Library and Information Service, forskerservice@kb.dk

**Bertil F. Dorch,** Research Support Service, The Royal Library / Copenhagen University Library and Information Service, forskerservice@kb.dk