

DataSpace Policies and Guidelines

Intent of the Service

DataSpace is a digital repository meant for both archiving and publicly disseminating digital data that is the result of research, academic, or administrative work performed by members of the Princeton University community. The services provided with DataSpace will focus on promoting awareness of the data contained in the repository and on addressing concerns for ensuring long-term availability of the data.

Content Guidelines

1. The work must be a result of work performed by faculty, students, or staff at Princeton University or a work funded or otherwise supported by Princeton University.
2. The work must be of intrinsic interest to Princeton University as an institution of higher education.
3. The work must be of long-term interest (e.g. longer than any individual's tenure at the university).
4. The work must be in digital form.
5. The work should be complete and ready for distribution. The work cannot be modified once it is submitted.
6. The author/owner must be willing and able to grant Princeton the right to preserve and distribute the work via DataSpace subject to any access restrictions agreed upon prior to submission of the work.
7. If the work is part of a series, other works in the series should also be contributed so that DataSpace can offer as full a set as possible.

Community and Collection Administration

Communities and collections are used to organize content hierarchically within DataSpace. Communities can contain sub-communities and collections. Collections contain the Items that are composed of digital files and associated descriptive information.

Top-level communities can be created for long-lived administrative units at Princeton University which have well defined leadership and the ability to assume some administrative responsibilities for the community. Each community will assign a coordinator who will work with OIT and Library staff to ensure that the DataSpace policies and guidelines are applied within the community's collection of content. Each community coordinator will be responsible for:

- Coordinating requests for the creation of new sub-communities, collections, and submitter accounts
- Customize the community's pages with appropriate logos, descriptive text, and licensing information
- Understanding and observing University policies relevant to DataSpace, and educating community submitters regarding these policies
- Developing more detailed policies and guidelines specific to collections within the community where necessary
- Deciding upon a submission workflow for each collection
- Overseeing the submission approval processes
- Notifying OIT's DataSpace administrator of organizational changes affecting submissions

OIT DataSpace system administrators will be responsible for:

- Creating communities, sub-communities, collections, and submitter accounts as approved by community coordinators
- Ensuring the files and metadata contained in the repository are accessible
- Withdrawing items from the repository leaving "tombstone" records indicating that the item has been withdrawn

Princeton University Library staff will provide assistance for determining the most appropriate file formats and descriptive metadata that will help ensure the discoverability and long-term usability of the data.

Access Policy

The general intent is to make all content in DataSpace publicly accessible. However, there may be certain special circumstances under which depositors need to restrict access to content. In these cases, the restrictions should remain in place for a well-defined and reasonably short period of time (e.g. no more than X years) after which the content will be made publicly available. Other repositories services (i.e. WebSpace or SharePoint) should be used to store data that should be shared only with a select set of users.

Charges

OIT will charge a one-time fee for submitting content into DataSpace. Revenue collected will be used to cover the costs of the storage required by the system. Charges will be calculated based on the amount of storage required by each submission. Charges will be applied to the project grant number the depositor records during the submission of an item to the system. Charges will be posted daily.

Beginning on July 1, 2009 the charge per Megabyte of storage will be \$0.006 with a minimum charge of \$0.60 per submission.

System Availability

The DataSpace system will not be treated as a business critical application where availability is concerned. The system will have very limited redundancy or failover capabilities. Preserving the data in the repository is of the utmost importance, but maintaining continuous availability of the system is a much lower priority.

User Accounts

User accounts are required for submitting content to the repository, for managing submission workflow processes, and for managing communities and collections within the repository. They are not required for accessing publicly available content held in the repository.

To obtain an account, a request must be submitted to the OIT DataSpace administrator via e-mail. The requestor must have a University NetID. The relevant community coordinator and DataSpace administrator must approve the creation of the account. DataSpace will store a local copy of the user's NetID and e-mail address. DataSpace will pass the user to the University's Central Authentication System for authentication. User passwords will never be retrieved by or stored within the DataSpace system.

Users whose University NetID have been disabled, typically because they have left the University, will no longer be allowed to login to DataSpace and therefore will not be able to submit any new content.

Users who have submitted content to DataSpace should update the DataSpace administrator with any changes in contact information. This information may be used should we need to have discussions about content at any point in the future.

Deposit License

In order for DataSpace to reproduce, translate and distribute your submission worldwide your agreement to the following terms is necessary. Please take a moment to read the terms of this license.

By clicking through this license, you (the author(s) or creator(s) or copyright owner(s)) grant to Princeton University the non-exclusive right to reproduce, translate (as defined below), and/or distribute your submission (including the abstract) worldwide in electronic format.

You agree that Princeton University may translate the submission to any medium or format for the purpose of preservation.

You also agree that Princeton University may keep more than one copy of this submission for purposes of security, back-up, and preservation.

You represent that the submission is your original work, and/or that you have the right to grant the rights contained in this license. You represent that your submission does not, to the best of your knowledge, infringe upon anyone's copyright. You also represent, to the best of your knowledge, that public distribution of your submission will not violate the personal privacy rights of any group or individual.

If the submission contains material for which you do not hold copyright, you represent that you have obtained the unrestricted permission of the copyright owner to grant Princeton University the rights required by this license, and that such third-party owned material is clearly identified and acknowledged within the text or content of the submission.

IF THE SUBMISSION IS BASED UPON WORK THAT HAS BEEN SPONSORED OR SUPPORTED BY AN AGENCY OR ORGANIZATION OTHER THAN PRINCETON UNIVERSITY, YOU REPRESENT THAT YOU HAVE FULFILLED ANY RIGHT OF REVIEW OR OTHER OBLIGATIONS REQUIRED BY SUCH CONTRACT OR AGREEMENT.

Princeton University will not make any alteration, other than as allowed by this license, to your submission.

Princeton University reserves the right to make unavailable any content which it deems to be in violation of this agreement, or of applicable local, state or federal law. Princeton also reserves the right to make content unavailable for limited periods of time so as to perform maintenance on the repository or its contents.

Princeton will endeavor to maintain this repository and make its contents available indefinitely. However, the University reserves the right to terminate this repository and its services should circumstances warrant. In such a case, Princeton's sole responsibility will be to make a best-faith effort to transfer the content to another repository and/or return the content to its original owner.

Size Limits on Deposits

The maximum size of a file that can be deposited or accessed via the DataSpace web interface is 2 GB. Although DataSpace can manage files larger than 2 GB, transferring them via the web interface is impractical. Files larger than 2 GB can be submitted, but must be deposited manually by a DataSpace systems administrator. If a user requests a copy of a file larger than 2 GB, special arrangements must be made to deliver the file via alternative means (e.g. a more suitable network protocol such as FTP or removable media such as CD, DVD, tape, or other mass storage device).

File Format Recommendations

In order to preserve the long-term accessibility of the information deposited in DataSpace we recommend files be encoded in formats that are expected to remain relatively stable and well supported over the long-term. For instance, PDF/A and ASCII are favored encodings for textual content. TIFF and JPEG2000 are favored for images. While any file format may be deposited in the repository, care should be taken to select formats that will help ensure that future users will be able to access and render the contents of files stored in the repository. Staff from the Office of Information Technology and the University Library can assist in the selection of the most appropriate file formats and metadata vocabularies.