

# OZone

## Institutional Repositories and the Ontario Scholars Portal

The Ontario Scholars Portal is a project of the Ontario Council of University Libraries - a consortium of twenty university libraries in the province of Ontario. Scholars Portal is an initiative to give users a single point of access to high-quality research from a broad range of disciplines, and one of its programs is to explore the use of institutional repositories as a tool for consortia resource sharing.

OZone is a shared IR environment being used to develop archives of electronic resources, to exchange electronic instructional packages among libraries, and to host institutional repositories that can be aggregated at a consortium level for searching. The platform we used to build our OZone institutional repository service is DSpace, which was originally jointly developed by MIT Libraries and Hewlett-Packard (HP) and is now freely available as open source.

DSpace manages and distributes digital items, made of digital files, provides tools for creating the metadata, indexing and searching of associated metadata and full text, and it is designed to support long-term preservation. The digital material enters the repository in any format, deposited by contributors using seven simple web forms. The system describes the digital material using Dublin Core with Qualifiers metadata, distributes the information with necessary access control via secure web server, assigns persistent identifiers to items, and preserves using large scale, stable, managed long-term storage, file migration techniques as technology evolves, and emulation software.

DSpace is hardware and OS platform independent, it runs on UNIX-like OS (Linux, HP/UX etc), but it can also run on Windows with no problem. The tools used to build and run a DSpace instance are: Java 1.4 or later, Apache Ant 1.5 or later - Java-based build tool, PostgreSQL 7.3 or later - open source relational database, Jakarta Tomcat 4.x/5.x or equivalent - servlet container, and Lucene for search and indexing. DSpace also supports the Open Archives Initiative's Protocol for Metadata Harvesting (OAI-PMH) version 2.0 as a data provider, and uses the Handle System from CNRI to assign and resolve persistent identifiers.

The software has a tree layer architecture, the storage layer is responsible for physical storage of metadata and content, the business logic layer manages the content of the repository, users, authorization, and workflow, and the application layer contains components that communicate with the world outside such as the web user interface, the Open Archives Initiative protocol for metadata harvesting service, and web services.

The content is organized in communities and sub-communities. In OZone each community has an administrative unit behind such as a university department, a lab, a research unit, and most often the top level community is an institution. Each community or sub-community can have as many collections as the community wants to define.

Collections provide a way to bundle items that have something in common, and each collection can have a workflow process defined to approve and edit the metadata for all the items submitted.

DSpace is also designed to make submission easy, and can be adapted to meet individual

needs for each community. The communities can manage themselves the permission based submission and the approval process, and with the introduction of the custom submission forms as of DSpace version 1.2.2 we can have as many or as few steps and also use any of the Dublin Core with Qualifiers in any of the web forms. Also very important when dealing with large existing collections, are the batch tools to import and export items in a simple directory structure, where the Dublin Core metadata is stored in an XML file.

The approval process can have up to 3 steps defined for each collection: reviewer, coordinator, and metadata editor. The reviewer can accept or reject new submissions, the coordinator can edit the metadata and accept/reject new submissions, and the metadata editor can only edit the metadata, the item cannot be rejected at this step. Each person participating in the workflow is automatically notified by email when a new submission is waiting for approval.

The platform can accommodate any material type, a few examples are: documents, such as articles, preprints, working papers, technical reports, conference papers, books, theses, data sets, computer programs, visualizations, simulations, administrative records, published books, images, audio and video files, multimedia publications, reformatted digital library collections, learning objects, web pages. Because the system will accept any file format, there are three levels of preservation defined for a given format: supported: we fully support the format; known: we can recognize the format, but cannot guarantee full support; unsupported: we cannot recognize a format; these will be listed as "application/octet-stream", also interpreted by the system as unknown. Supported formats will be functionally preserved using either format migration or emulation techniques. Known formats are those that we can't promise to preserve, such as proprietary or binary formats, but which are so popular that for sure third party migration tools will be developed to assist in the migration process. Unsupported formats are those that we don't know enough about such as some proprietary formats or a one-of-a-kind software program.

Some new enhancements developed by the DSpace community since we implemented OZone, that helped us provide a better service to our users are: thumbnails in item view can now be switched on or off, browse and search thumbnail options, improved item importer, customisable submission forms, configurable number of index terms in Lucene for full-text indexing, the submit button on collection page only appears if user has authorisation to submit to that particular collection, community/collection strengths, and the new role – collection manager .

The main future steps, planned by the DSpace community, that we believe will further help us improve our service are in the area of: modularity, internationalization, and preservation. Improvement in the modularity area will make integrating and maintaining our custom changes to the code a less time consuming task, and it will also allow for a better branding of our communities. By improving the preservation component we can ensure the longevity and accessibility over time of the work deposited in OZone.

CORIL (Cooperative Online Repository for Information Literacy), the first community in OZone, is an initiative to support information literacy instruction among Canadian universities. The aim of the project is to increase the quantity and availability of a range of high quality learning, facilitate the sharing and dissemination of learning resources, create a learning community of Canadian instruction librarians, and provide a mechanism for peer-reviewed publication of instruction materials. By taking advantage of all the features

presented here all these goals can be achieved using our OZone service.

[www.ozone.scholarsportal.info](http://www.ozone.scholarsportal.info)

Gabriela Mircea

[gabriela.mircea@utoronto.ca](mailto:gabriela.mircea@utoronto.ca)

References:

1. DSpace: <http://www.dspace.org/>
2. DSpace System Documentation <http://dspace.org/technology/system-docs/>
3. The DSpace 2.0 Architectural Goals <http://wiki.dspace.org/DspaceTwo>
4. DSpace Wiki: <http://wiki.dspace.org/>
5. OZone [www.ozone.scholarsportal.info](http://www.ozone.scholarsportal.info)