

DSpace Summer of Code Ideas

- 1 [Google Summer of Code Ideas List for DSpace](#)
 - 1.1 [Past Ideas Lists for DSpace GSoC projects](#)
 - 1.2 [Past DSpace GSoC Projects](#)

Google Summer of Code Ideas List for DSpace

Add your ideas here!



Please add your suggestions for GSoC 2015 projects related to DSpace! If you are interested in mentoring, please let us know! Also, be sure to visit the listing of [Past GSoC Project Ideas](#) below, to see if anything there is still relevant. Also, please note that Google has recommendations on what to minimally include in each "Idea", see: [What is an Ideas List?](#)

NOTE: The below DSpace-specific ideas table is automatically embedded into the global DuraSpace [Google Summer of Code Ideas](#) page.

Please add your own ideas to the table below, and feel free to volunteer as a mentor for any existing idea.

Project Title	Description	Mentor Volunteers
File format validation with DROID and JHOVE	When a file is currently uploaded in DSpace, its file format is guessed by looking at the file extension. This means if I change the extension ".exe" of a file containing a virus to ".pdf", DSpace will not have a clue what's going on and identify the file as a PDF. You can make this more robust by working on the class FormatIdentifier. The student is expected to work on this class to handover the task for format identification to a library like JHove, DROID, or both. Minimal viable work: - Implement JHove OR Droid, and demonstrate that it can perform identification on a number of text and image formats when uploaded to DSpace. Extra kudos: - The student implemented BOTH JHove and Droid, and produced a performance analysis on which of the two (or both) should be preferred, leading up to a more optimized solution. GODLIKE: - Aside from the inclusion in the submission form, the student also creates a DSpace Curation Task that allows a repository manager to run file format identification against existing items, or entire collections. The administrator receives a report on those items that are differently identified by JHove/DROID, compared to what's already stored as the bitstreams format. <i>Difficulty level: Easy</i>	Philip Vissenaekens (Atmire) Andrea Schweer
Translate Wiki Integration	The DSpace community has approached TranslateWiki.net (TWN), the mediawikibased platform for interface translation of open source projects. The initial discussions are promising and the TWN community is currently building support for the Apache Cocoon message format that is used by XMLUI. We need an ambitious GSOC student to connect the dots, working with both communities in extending the integration and ensuring that in the end, the threshold is lowered for external translators to provide interface translations for DSpace. <i>Related Links:</i> JIRA Ticket describing the current status of the dialogue between DSpace and TWN <i>Related DSpace Components:</i> dspace-api-lang : JSPUI and API level language packs (in Java Properties file format) dspace-xmlui-lang : XMLUI language packs (in Apache Cocoon XML format) dspace-xmlui : XML-based interface based on Apache Cocoon dspace-jspui : Custom JSP-based interface <i>Recommended Skills:</i> - Multilingual and/or translation experience - Familiarity with Java web application Internationalization <i>Difficulty Level: Medium</i>	Bram Luyten (Atmire)

Virtual Sets: Separate the internal repository structure from the navigation structure	Currently, the hierarchical structure used in DSpace allows sharing items between collections by explicitly declaring these relations on each item. However, DSpace does not allow to relate a collection or a sub-community between two or more communities. Virtual Sets are arbitrary aggregations of DSpace Objects and criteria, composed of: • arbitrarily selected communities, collections and/or items, • dynamic results from criteria/queries (logical expressions; e.g. Solr queries) • other declared Virtual Sets (initially, cycles are not allowed) Virtual Sets in DSpace would allow the creation of complex navigation structures regardless of the hierarchical (perhaps administrative) structure of communities and collections. Structures, as described above, are supported in at least another repository software such as Fedora-based Hydra due to its generic object model. The initial implementation would affect the dspace-api component and either of the web UIs, since both now use Discovery (Solr) by default. Virtual Sets should be implemented at the dspace-api level for DSpace Objects to offer more orthogonal features like Virtual Sets backup, export and exposure through OAI-PMH. Stretch goals would include making use of Virtual Sets in other interfaces like REST and/or OAI. <i>Related links:</i> • Sets, sets or sets? <i>Related DSpace Components:</i> • dspace-api : DSpace Java API • dspace-xmlui : XML-based interface based on Apache Cocoon • dspace-jspui : Custom JSP-based interface <i>Recommended Skills:</i> • Java programming experience <i>Difficulty Level:</i> High	Ivan Masár
Create a puppet-dspace module	Develop a Puppet-DSpace module for deploying DSpace, and provisioning DSpace-related services. The goal would be to deploy on at least one popular Linux OS, and likely the target will be Debian/Ubuntu since testing will be done with Vagrant-DSpace, which currently uses only Ubuntu. The final product should be useful for deploying DSpace to a cloud infrastructure, or really any server configured to run Puppet (in other words, the module should make no assumptions that rely upon paths used by Vagrant). A stretch goal for this project will be to make the puppet module OS-agnostic, and have it run on both Debian/Ubuntu as well as CentOS/RHEL. But just meeting the Ubuntu target would be sufficient for this project. The puppet module already built for use in Vagrant-DSpace would be a great starting point. The module, when complete, should enable an operator to go from a standard OS base image to a running instance of DSpace, complete with a container to host the application, and (optionally) a PostgreSQL database for metadata. <i>Related DSpace Components/Links:</i> • vagrant-dspace : A Vagrant setup for DSpace development. It includes the "starting point" for a puppet-dspace module under "/modules/dspace/" • Installing DSpace : DSpace 4 Installation instructions <i>Recommended Skills:</i> • Experience with Puppet or similar tools (e.g. Chef). At a minimum some base familiarity and even some Ruby experience (which Puppet is based on). • Familiarity with Vagrant, or willingness to learn. <i>Difficulty Level:</i> Medium	

Next-gen UI	MDS is an experimental offshoot of DSpace in which new ideas may be prototyped and examined. Recently a REST API (with CRUD operations, etc) has been added to MDS. A valuable 'proof of adequacy' is building an entire functional web UI backed only by the API. The goal of this project is to construct such an admin UI for MDS, using a modern, agile web application framework. A proof of concept using AngularJS already begun could serve as a basis for further work. <i>Related DSpace Components:</i> • mds : "Modernized DSpace". An attempt to refactor/redesign the DSpace API to make it more simplistic/modernized. This is a side-project of long-time Committer, Richard Rodgers • mds/webapi : the REST API for MDS project. This API supports CRUD operations. • dspace-rest (loosely related): the official REST API which now ships with DSpace 4 may also provide a possible integration point. However, it is currently read-only. <i>Recommended Skills:</i> • Experience or familiarity with one or more agile web frameworks • Experience or familiarity with buiding agile interfaces against a REST API <i>Difficulty Level:</i> Medium to High	Richard Rodgers Philip Vissenaekens (Atmire)
-------------	--	--

Past Ideas Lists for DSpace GSoC projects

We have archives of all our Past GSoC Ideas Pages still available for reference/ideas. *However, you should check with the available mentors before suggesting any of these older project descriptions. In some cases these projects may require rethinking to bring them up to date.*

- [DSpace Summer of Code Ideas 2012](#)
- [DSpace Summer of Code Ideas 2011](#)
- [DSpace Summer of Code Ideas 2010](#)
- [DSpace Summer of Code Ideas 2009](#)
- [DSpace Summer of Code Ideas 2008](#)

Past DSpace GSoC Projects

- 2011 Projects listed on [GSoC 2011 Projects](#) page.
- For years prior to 2011, please visit [Past DSpace Summer of Code Projects](#).