

Migration to GitHub

This page provides notes on the [DSpace SVN](#) to [GitHub](#) Migration that took place in April/May 2012. It has now all been completed. The final announcement of this migrate is at: <http://www.mail-archive.com/dspace-general@lists.sourceforge.net/msg01268.html>

GitHub Migration Steps

- 1 [Codebases to migrate to GitHub](#)
 - 1.1 [Main Codebases](#)
 - 1.2 [Dependency Codebases](#)
 - 1.3 [Additional Active DSpace projects](#)
 - 1.4 [Inactive projects \(won't be moved to GitHub\)](#)
- 2 [Archive old SVN \(for posterity\)](#)
- 3 [Additional Cleanup tasks](#)
- 4 [Final Steps](#)
- 5 [Author File](#)

Codebases to migrate to GitHub

At this time, we will only migrate codebases which are currently under active development (we can always move others later on, as we will keep an archive of SVN). If any are missing from the below list, feel free to add them.

GitHub import hints

 <http://help.github.com/import-from-svn/>

- Using [svn2git](#) is recommended, just make sure to also pass it the [#Author File](#) below.
 - Note that "svn2git" doesn't seem to work on Windows. Works fine on Mac OSX though (Tim)
 - Sample walkthrough of using "svn2git": <http://blog.orfjackal.net/2009/05/converting-confused-svn-repositories-to.html>

Main Codebases

-  [Primary DSpace Codebase](#) - The Official [DSpace/DSpace](#) GitHub repository is now active! All DSpace development is now occurring in GitHub instead of SVN.
 1. The "DSpace/DSpace" GitHub is a **BRAND NEW** copy of SVN, with SVN author mapping using the [#Author File](#) below.
 2. WARNING: The old, "unofficial" "DSpace/DSpace" GitHub repository was renamed to [DSpace-SVN-Deprecated](#)
 - Anyone who had forked the "unofficial" GitHub migrate as-soon as possible to the new GitHub repo, as the old one will **not** be maintained any further.
 - Once everyone has migrated (no more forks exist), we will delete the "DSpace/DSpace-SVN-Deprecated" repo.
-  [dspace-api-lang](#) - migrated to GitHub:[DSpace/dspace-api-lang](#)
 - perhaps migrate to a single 'dspace-lang' GitHub project? (Mark: I think move back into DSpace API and use minor releases to update languages.)
-  [dspace-xmlui-lang](#) - migrated to GitHub:[DSpace/dspace-xmlui-lang](#)
 - perhaps migrate to a single 'dspace-lang' GitHub project? (Mark: I think move back into DSpace-xmlui and use minor releases to manage.)
-  [dspace-services](#) - migrated to GitHub:[DSpace/dspace-services](#)
-  [dspace-pom](#) - consolidated into Official [DSpace/DSpace](#) GitHub repository
-  [License/License Header files](#) (used by 'dspace-pom') - consolidated into Official [DSpace/DSpace](#) GitHub repository
-  [dspace-assembly-plugin](#) (used by 'dspace-parent' to assembly DSpace)- consolidated into Official [DSpace/DSpace](#) GitHub repository

 **QUESTION:** The following codebases are also standalone. Do we want to keep them that way? Do they all merit their own repository in GitHub?

Dependency Codebases

The following codebases are strictly releases of DSpace dependencies. Some of these dependencies just needed to be pushed to Maven Central (i.e. we don't "own" any of the code – all we did was release it to Maven Central because we needed it available there). Others need minor tweaks for DSpace.

-  [dspace-solr](#) (dependency of 'modules/solr') - migrated to GitHub:[DSpace/dspace-solr](#)
-  [dspace-geoip](#) (dependency of 'dspace-stats') - migrated to GitHub:[DSpace/dspace-geoip](#)
-  [dspace-cocoon-servlet-service-impl](#) (dependency of dspace-xmlui-wing) - migrated to GitHub:[DSpace/dspace-cocoon-servlet-service-impl](#)

The below dependencies may no longer be needed, as it seems like we have a valid way to replace them in the future.

-  [dspace-tm-extractors](#) (used by dspace-api to do text extraction of Word docs) - Can we just replace this with POI (<http://poi.apache.org/>), which is used for PPT text extraction (but also looks to work for Word docs)? See [DS-1140](#)
-  [dspace-jmockit](#) (dependency of dspace-parent) - [JMockit is now in Maven Central](#), as long as we can upgrade to 0.999.11 or above. So, we may no longer need this separate project.

Additional Active DSpace projects

-  [dspace-replicate](#) - Already migrated to <https://github.com/DSpace/dspace-replicate>

-  [dspace-rest](#) - Already migrated to <https://github.com/DSpace/dspace-rest>

Mark: separate repo because they have separate releases cycles, though it might be time to put rest into the main DSpace repo.

Inactive projects (won't be moved to GitHub)

No Inactive Projects will be migrated to GitHub at this time

 At this time, we do not plan to migrate any inactive projects to the central DSpace GitHub. Instead they will all be archived in a READ-ONLY SVN. They can always be migrated at a later time, as necessary.

Individual developers/committers are welcome to move any of these projects into their own personal GitHub account. They just won't be moved into the central DSpace GitHub.

Projects that seem to be "inactive" at this time include:

- [DSpace 2.0 SVN codebase](#) (This is the code from the 2009 [DSpace 2.0](#) prototype)
- Modules directly related to the 2.0 Prototype, including:
 - The 'dspace-storage' modules and all "storage-*" modules listed at: <http://scm.dspace.org/svn/repo/modules/>
- All Sandbox projects at: <http://scm.dspace.org/svn/repo/sandbox/>
- All old DSpace Maven Tools at: <http://scm.dspace.org/svn/repo/tools/maven/> (EXCEPT the 'dspace-assembly-plugin' which is listed above)
- The following 'demo/experimental' modules:
 - [dspace-core](#) module
 - [dspace-database](#) module
 - Old copy of 'dspace-discovery' at: <http://scm.dspace.org/svn/repo/modules/dspace-discovery/>
 - [dspace-imsdp](#) module
 - [dspace-schedule](#) module
 - Old copy of 'dspace-stats' at: <http://scm.dspace.org/svn/repo/modules/dspace-stats/>
 - [dspace-sync](#) module
- Any other SVN projects we haven't explicitly listed as one we are planning to move to GitHub.

Mark: Some of these can be considered for future, I want to take back those that I put considerable work into and will put these up in my repo or in our atmire repo. DSpace-imsdp was an example of a crosswalk/packager addon, it could stand to be in the DSpace repo. Stats and discovery can remain behind.

Archive old SVN (for posterity)

-  Make existing SVN READ-ONLY
-  Archive the existing SVN (scm.dspace.org hosted by OSUOSL) to <http://svn.duraspace.org> for posterity.
 - DSpace READ-ONLY SVN archive is at: <https://svn.duraspace.org/dspace/>
-  Post a notice in the archived SVN (at svn.duraspace.org) that codebase is now at GitHub

Additional Cleanup tasks

-  Turn off any syncing between SVN and GitHub
 - Notify the Committers before making this change!!
-  Ensure Maven 'License Header' check still works when running 'mvn install' (this is setup in 'dspace-pom')
 - Likely will need to move [LICENSE HEADER](#) to GitHub, maybe into same project as 'dspace-pom'?
 - Also will likely require a new release of 'dspace-pom'
-  Ensure Maven [Release Procedure](#) still works - likely will require a POM update or two.

Mark: I've run a few releases via sonatype and github, it might be good to run a test release or two just to verify. We just need to update the source management section of the maven parent Pom. I have examples of these settings in my repo.

Final Steps

-  Notify community when development has moved completely to GitHub.
-  Redirect old SVN URLs:
 -  Redirect scm.dspace.org to svn.duraspace.org (if possible, setup this redirect so that old URLs just go directly to that SVN archive)
-  Contact OSUOSL and let them know we no longer need to use the old SVN server.
- **(In process/ongoing)** Build some "GitHub Best Practices" documentation @ [Development with Git](#)
 - Need to also update or delete [Tracking your source code with Git](#)

Author File

The Author File provides a mapping of SVN users to GitHub users.

Format:

```
svnuser = gituser_full_name <gituser_email>
```

KevinVandeVelde = Kevin Van de Velde <kevin@mire.be>
ScottPhillips = Scott Phillips <scott@scottphillips.com>
StuartLewis = Stuart Lewis <stuart@stuartlewis.com>
ahkim = Austin Kim <ahkim@users.sourceforge.net>
aschweer = Andrea Schweer <schweer@waikato.ac.nz>
azeckoski = Aaron Zeckoski <azeckoski@gmail.com>
benbosman = Ben Bosman <ben@atmire.com>
bollini = Andrea Bollini <bollini@cilea.it>
bradmc = Brad McLean <bmclean@duraspace.org>
cjuergen = Claudia Juergen <claudia.juergen@tu-dortmund.de>
dchud = Dan Chudnov <daniel.chudnov@gmail.com>
dstuve = David Stuve <david.stuve@hp.com>
gam = Greg McClellan? <gam@users.sourceforge.net>
gam5 = Greg McClellan <gam5@users.sourceforge.net>
gcarpent = Grace Carpenter <gcarpent@users.sourceforge.net>
grahamtriggs = Graham Triggs <grahamtriggs@gmail.com>
jimdowning = Jim Downing <jim.downing@gmail.com>
jrutherford = James Rutherford <jrutherford@users.sourceforge.net>
jtrimble = Jeffrey Trimble <jatrimble@ysu.edu>
keithg = Keith Gilbertson <keith.gilbertson@library.gatech.edu>
kshepherd = Kim Shepherd <kim.shepherd@gmail.com>
lcs = Larry Stone <lcs@users.sourceforge.net>
mdiggory = Mark Diggory <mdiggory@atmire.com>
mirceag = Gabriela Mircea <gabriela.mircea@utoronto.ca>
mwoodiupui = Mark Wood <mwood@IUPUI.Edu>
pbreton = Peter Breton <pbreton@users.sourceforge.net>
peterdietz = Peter Dietz <dietz.72@osu.edu>
pottingerhj = Hardy Pottinger <pottingerhj@umsystem.edu>
pvillega = Pere Villega <pere.villega@gmail.com>
richard-jones = Richard Jones <richard@cottagelabs.com>
robintaylor = Robin Taylor <robin.taylor@ed.ac.uk>
rrodgers = Richard Rodgers <rrodgers@mit.edu>
rtansley = Robert Tansley <roberttansley@google.com>
sands = Sands Fish <sands@mit.edu>
scottphillips = Scott Phillips <scott@scottphillips.com>
stuartlewis = Stuart Lewis <stuart@stuartlewis.com>
syeadon = Scott Yeadon <scott.yeadon@anu.edu.au>
tdonohue = Tim Donohue <tdonohue@duraspace.org>