

DevMtg 2011-06-06 - OR11 Meeting

- [Developers Meeting before OR11 on Mon, June 6, 2011](#)
 - [Who is invited?](#)
 - [Logistics](#)
 - [General Agenda](#)
 - [Sign Up to Attend!](#)
 - [Meeting Notes](#)

Developers Meeting before OR11 on Mon, June 6, 2011

Face-to-face developer's meeting on topic of DSpace RoadMap just before [Open Repositories 2011 in Austin, Texas](#).

Who is invited?

- All Committers,
- All [DSpace Community Advisory Team \(DCAT\)](#) members,
- Any other interested DSpace developers or technology-savvy individuals

If you don't fall into one of the above categories, you are still welcome to attend. However, be warned that discussion will likely get very technical at times (which is why we recommend you be a developer or have a technology background).

Space is limited. Please sign up on the [Sign Up sheet](#) below. If we begin to achieve capacity, preference will be given to Committers or DCAT Members. However, we welcome any interested DSpace developers to join us and take part in the discussion!

Logistics

- Day: Monday, June 6th, 2011
- Time: **9am-4pm**
- Where: **Perry-Castañeda Library (PCL), Room 1.339**
 - [Walking Directions \(< 1/2 mile\) from AT&T Conference Center](#)
 - More information from OR11 Organizers on [Getting to your PreConference Building/Room](#)
- Attendance Limit: 30 people maximum
- Lunch is NOT provided, but there are lunch options in the area: <http://sites.tdl.org/openrepositories/or11preconference/ut-austin-information/pre-conference-lunch-options/>

Note on Accommodations

 If you're planning on staying at the OR conference hotel, be advised that **the conference rate of \$129/night is only available from June 7th onward**. The OR11 website accommodations page has [a list of alternative hotels](#) you may wish to choose from for the nights of June 5th and 6th.

General Agenda

Format of Meeting

 This meeting will be organized as a group discussion. Although at times one or more of us may lead discussion sections, it is meant to be a group discussion and **not** a series of presentations (in fact, at this time there are no plans for any presentations / slides).

DSpace Meeting Notes

 Community Notes page: <http://piratepad.net/or11dspacesmeeting>

Agenda: (NOTE: This agenda is flexible, and based on discussions throughout day is subject to change)

9:00am: Introductions, Overview of Day

- Introductions
- Suggestions for agenda changes?

9:20ish : Brainstorming Activity

- What does the "Modern Repository" look like? What are the requirements/needs of its users, etc?

10:15ish : Break

10:30am : Brief updates/discussions on Upcoming/Ongoing Work

- [Google Summer of Code \(GSOC\)](#) updates
- Updates on "Fedora Inside" work
- Updates on [DSpace 1.8.0](#) planning
- Other brief updates?

11:15ish : Final Decision on a Version Numbering Scheme

- Discussion & Final Decision on whether we want to change our [version numbering scheme](#) (i.e. Do we want to continue with incrementing version numbers or move to something else?) |

Noon: Lunch Break & Recommended Discussion

- Thoughts on interaction/collaboration between [DCAT](#) & Committers? Community outreach in general?

1:30pm: Open Discussion Period ([Doodle Poll of Topics](#))

- We will discuss "hot topics" decided by all of us
- A portion of this period will be left completely "unscheduled", so that we have flexibility to discuss other topics that come up.

2:45ish: Break

3pm: Summarizing & Bringing it all Together

- Revisiting 1.8.0 Planning.
- What do we see as our Technology RoadMap/Goals for next few years? What can we 'vote on' and approve?
 - **Approve a general DSpace 2011-2012 Technology RoadMap**

4:00ish: End Meeting & Depart for [Dog and Duck Pub](#)

—

Possible Topics for "Open Discussion" period:

- Creating a Common Business Logic Tier (share business logic between each of the UIs, rather than each UI building its own separate business logic)
- Enhancing Metadata Support (how can we enhance metadata support? what metadata should we work to support more in future?). Related discussions:
 - Updating to current DCMI standards: [DS-805](#)
 - Creating a new "dspace" internal usage schema (for administrative or non-DC metadata)? Formalize the "dc" schema to be strictly valid QDC/DC?
 - [Metadata on all DSpace objects \(not just Items\)](#)
- Modularization activities & Async discussions (how can we better modularize DSpace and move towards a "plugin" framework, what would be requirements, etc.) Related Discussions:
 - [Restructure Trunk Projects](#)
 - [Refactoring the DSpace Domain Model](#)
 - [Asynchronous Release](#)
 - [Asynchronous Release - Alternative Option](#)
 - Also discussions about whether to "split" up our "Modules SVN area" into three groups: "Sandbox Modules", "Core/Fully Supported Modules", "Extension Modules"
- Documentation Management Team & ongoing Documentation mgmt in general
 - Related discussions around whether we should create a "Documentation Management" team or a team of 'editors' to help clean up docs and improve management of them. This team of "editors" may provide some basic approval for Docs changes, and would also "lean on" committers when necessary to ensure new features are fully documented, etc.
- Anything "annoyingly obsolete" (technology wise) we really need to address soon?
- REST API or other important projects that may need "one last push" for official release.
- More General Discussion/Brainstorming of DSpace RoadMap (though entire meeting is generally related to this overarching topic)
 - See [Proposed RoadMap to 2.0](#) for some very early RoadMap ideas/brainstorms.

Informal Evening Meet-Up



At the [Dog and Duck Pub](#) (17th and Guadalupe)

Sign Up to Attend!

If you're planning to attend this meeting, please add your name to the sign up sheet. Also, please add a note as to your general availability that day (e.g. "available all day", "available PM only", etc.).

Sign Up Sheet - Will Be Attending

1. [Tim Donohue](#) – available all day
2. [Valorie Hollister](#) -- available all day (although going between DSpace/Fedora committer mtgs)
3. [Scott Phillips](#) - available all day
4. [Imma Subirats](#) - not sure of availability yet, most probably PM
5. [Mark Diggory](#) - should be available all day at this point.
6. [Mark H. Wood](#) – available all day
7. [Keith Gilbertson](#) -- available all day
8. [Richard Rodgers](#) - available all day
9. [Stuart Lewis](#) - available all day
10. [Kim Shepherd](#) - available all day
11. [Elin Stangeland](#) - available all day
12. [Graham Triggs](#) - available all day (ish)
13. [Jonathan Markow](#) - available from late morning on
14. [Benoit Pauwels](#) - available all day
15. [Andrea Schweer](#) - available all day
16. [Jonathan Amburn](#) - available all day
17. [Robin Taylor](#) - available all day

18. [Sarah Shreeves](#) - available all day
19. [Richard Jones](#) - available all day
20. [Bram Luyten \(Atmire\)](#) - available all day
21. [Bradley McLean](#)
22. [Adam Field](#) - most of the day, depending on jetlag
23. [Hardy Pottinger](#) - available all day
24. [Kazu Yamaji](#) - available all day
25. [Kei Kurakawa](#) - available all day

Want To Attend / May Be Attending

- *add your name*

Will Not Be Attending

- [Peter Dietz](#) - at home, expecting a baby any time.

Meeting Notes

Initial Meeting Notes

Most of the below notes are just copied from this Community Notes page: <http://piratepad.net/or11dspacemeeting> During the meeting we had encouraged everyone to take notes on that PiratePad page.

Brainstorming Activity:

- [Modern Repo Brainstorming Activity](#) - What does the "Modern Repository" look like? What are the needs of its users?

GSoc 2011 Brief Updates:

- SKOS Authority Controls
 - SOLR Based authority control plug for DSpace. Term completion against stuff you put in your SOLR.
 - Always lacking: authority control is always a list, but controlled vocabularies are graphs/networks. Usecase Nescent: SKOS Hive (built on sesame) authority source, now exposing this through SOLR in DSpace submission forms.
 - Feedback on enhancing authority control in DSpace:
- WebMVC (Freemarker UI development) - Talk on this Saturday
 - Instead of newer features, project is mainly focussed on completing the existing UI functionality. (a modern JSP UI).
 - Feedback on WebMVC
- Submission Enhancements in DSpace
 - In-browser interface to edit the submission interfaces
 - Feedback on submission enhancements in DSpace:
 - (Elin noted we need to be careful about our usage of the term "Workflow".. developers sometimes have a different interpretation to community)
 - Reviewer workflow - Talk on this Friday
 - Framework for more flexible workflows (different edit steps, etc)
 - Feedback on reviewer workflow:
 - Elin: need for a common terminology on Workflow vs Submission. Is Submission part of the workflow according to the developers?
- New UI built over RESTful services
 - Yes, it's kind of a new UI. But more oriented over building widgets, over the REST API.
 - Aimed at better testing the REST API
 - Returning student (Bojan) now serving as mentor, which is a 'GSoc first' for DSpace
 - Licensing (strict GPL requirements) has become a big factor in decisions re: toolkits/frameworks/etc to use

DSpace with Fedora-Inside Brief Updates:

- Long term project.
- 1st step already passed: being able to move DSpace objects around. AIP exporter for the whole object (metadata, content, access rights, ...)
- Goal: using these AIP packages to feed a fedora with.
- Next step: how will this exported data be well represented in Fedora?
- After that: how to put business logic of DSpace on top of Fedora. Problem: not enough technical staff in Duraspace, so it will take a long while. In addition, DuraSpace doesn't want to do this work alone, as individual institutions better know the needs/requirements. DuraSpace is looking for other stakeholders to jump in, participate and speed things up.

Final Decision on a Version Numbering Scheme

- Background:
 - "DSpace 2" as a concept/version number has been floated for ~5 years
 - The definition's changed over time
 - Are we moving towards Dspace 2? Fedora-inside-DSpace == DSpace 2?
- Some Possible Version Numbering Options:
 - 1.8, 1.9, 1.10, 1.11 2.0 (we decide when to make the major version shift to "2")
 - 8, 9, 10, 11 ... (drop the "1.", like Java)
 - YY.MM (eg. 12.11), like Ubuntu
 - YY or YYYY (e.g. DSpace 11 = 2011, or DSpace 2011)
- Scott: What's the problem we're trying to solve?
- Richard R: Expectations around 2.0, how close we are to it (as we approach 1.9, etc.), what we do to make it clear where 1.x is in relation to "2.0"
- Richard J: Shouldn't this be a marketing issue, not an issue for developers to figure out?
- Bram pointed out that encouraging sites to upgrade to the latest versions, and making those upgrades easy and painless, will result in more feedback and more interest from developers, and therefore more good code.

- Scott pointed out that even though 2.0 was "revolutionary" in its initial conception, in practice we've been evolutionary/incremental when it comes to actually releasing new features, versions, etc.
- Stuart pointed out that in recent years, we've been getting much better in terms of how often we release, restricting 1.x.1 releases only to bugfixes, etc.
- Over half the room raised their hand to "We don't ever want to release a 2.0". A good alternative being floated is to jump straight to 3.0 and then continue as normal, with major version reflecting major architectural change, and we know for the future that if we're planning big projects, we should use codenames instead of numbers!
- People who are currently happy users of DSpace and following current release will not care about new versioning schemes, and will upgrade in any case.
- There's a major opportunity in winning people again, who decided in the past that the DSpace product didn't fit their usecase.
- The Final Vote:
 1. 3 people voting for going from 1.x --> 3.0 (i.e. Skip over '2.x' version)
 2. 11 people voting for dropping major, eg. DSpace 3, 4, 5, 6 OR 9, 10, 11
 - By "dropping major" we mean moving from a "x.y.z" versioning to just a "y.z" versioning. So, rather than a 1.9.0, we'd just potentially call that "9.0" or we'd name it "3.0" instead
 - There was some indecision whether we should move to "3.0" after 1.x is done, or just drop the "1." and move right to "9.0", "10.0", etc.
 3. 4-5 people voting for date-based versioning (either YY or YY.MM)
- Results: 2.0 is gone, and we will move to a "[major].[minor]" numbering scheme (instead of [major].[minor].[subminor]).
 - Do a strawman poll to find out exactly how the branding will work, whether we co-brand date and major number. (e.g. DSpace 3.0 (2012))
 - Feedback from others: We need to document our version numbering very well. What does it take to reach a "major" release, or do we just release a new "major" release each year. Needs to be clear to users what these version numbers "mean".
 - But the 1.x.y will definitely go, at some point. And 2.0 is definitely gone. **RIP 2.0**

Although a decision was made to retire the name "2.0", this has not yet been announced in a more public fashion (on mailing lists, etc). The reason is that DSpace 1.8.0 will **still be called 1.8.0**. After the 1.8.0 release, we will make a decision on what the following 2012 release will be named (likely either DSpace 3.0 or 9.0). At that point in time, we will announce the retirement of the name "2.0" (though many features of the DSpace [2.0 Prototype](#) will continue to be added to DSpace little by little), while also announcing the new version of DSpace.

So, in summary: DSpace will likely change its version numbering scheme **after** 1.8.0 is released in Oct 2011. At this point in time, we are undecided what that new version numbering scheme will be. However, we will likely skip over the number "2.0", as that version has too much past history and assumptions built up around it. In the end, we hope to simplify and clarify our numbering scheme. As always, DSpace will continue to release at least one new major version each year – it's just that after 1.8.0 the DSpace version numbers may look slightly different than in the past.

DCAT Update / Feedback

- They've been taking new feature requests from JIRA, members pick issues that are of interest to their institution or that they've heard interest for, figure out prioritisation that way
- They've been through about 5 new feature requests this way so far
- Feedback from committers/developers is useful – coming back with further requirements, etc.
- Some issues require participation from the broader community, DCAT will be working to enable this
- Also working with DSpace ambassadors to try and reach repository managers that might not follow dspace-tech or be involved via usual channels
- Has been a bit slow to start, but it has been really helpful having Robin (in capacity as 1.8 release coordinator) and Tim join in the DCAT discussions
- Encouraging members with a particular interest in new features/issues to join in on IRC committer meetings
- Robin: One benefit has been giving repository managers a voice and making sure developers don't just follow their own pet projects, remind them of real-world priorities
- Tim: gives us a perspective outside of our own institution
- Stuart: Can we get a new flag or status in JIRA so we can get DCAT review of issues we've written code for (or issues we don't feel are clearly defined... instead of them ending up stale because we couldn't figure out how to move on them in a committer meeting)
 - **TODO:** Tim will investigate ways to "flag" an issue in JIRA as "needing a DCAT Review"
- Bram: How many of the 1000+ repository managers are involved in meetings/conferences/any kind of communication where they can express their requirements, issues, ideas? Eg. I use MS Word but I don't turn up to MS focus groups on word processor design. Would be really cool to cross-ref the list of dspace.org listed instances, and people registered on the mailing lists, to find out about those listed on dspace.org who are NOT on the mailing lists. They might need some extra attention to get involved.

Open Discussion Period

Topic 1: Enhancing metadata support (DC cleanup and/or adding metadata to all objects)

- Mark gave overview of how Fedora handles metadata as datastreams (eg. RELS-EXT and RELS-INT RDF, the DC datastream, representation in METS, etc).
- Breaking out from a "relational table" structure - schema.field.qualifier
- Use cases:
 - SWORD2 - have to do a mapping from regular dublin core to DSpace "dublin core"
 - Other Research Data systems - richer metadata than can be supported in dublin core (hierarchical schema needs - MODS or disciplinary based metadata schemas)
 - Theses & Dissertations - very specific ETD-MS metadata schemas (and sharing that metadata with aggregating systems)
- Quick Wins:
 - Remove any Hardcoding assumptions on Dublin Core
 - Cleanup our DC to be real Qualified Dublin Core
 - Provide some existing Schema files for other schemas so that you can import them easily
 - PRISM (page numbers) etc.
 - Redirect some of the Admin Metadata into a "DSpace" internal schemas (or local namespace)
 - Get input from DCAT on metadata schemas
 - What is the set(s) of metadata that should come "out-of-the-box" DSpace?
 - Can we improve Admin UI to help 'guide' people.
 - Training for users - Don't just stuff new things into DSpace metadata, look for a schema that it really fits into.

Topic 2: Configuration changes (dspace.cfg)

- Use Cases:
 - RichardR: Can already start to 'split' out dspace.cfg into 'config/modules' directory files
 - Scott: I want a separate "local" config for more non-default settings
 - Stuart: Should we just 'carve up' these configs for 1.8? (Some agreement)
 - But, is that a "painful" upgrade process?
 - Stuart: Could we do a "reload" method for Admin UI (avoid tomcat restart)
 - May need to avoid reloading some files (like DB settings)
- Quick Wins:
 - Split up dspace.cfg already for 1.8?
 - "Local Overrides" for 1.8?
 - Investigate "Reload" button for Admin UI
- Mark - Concerns on how Configuration Service also gets updates (via Reload)

Topic 3: REST API?

- Briefly discussed. General agreement that we want to release a 'beta version' in 1.8.0 release.

DSpace 1.8 Discussion

- Patches planned to go in, owner, status
- Tim: Backup/restore from Admin UI
- Mark: Configurable Reviewer Workflow
- Refactoring ConfigurationManager/PluginManager? Some more impl
- Stuart: Some batch metadata changes
- Stuart: Global changes (Kim: with regexes!)
- REST API? We should complete/polish the spec and include
- Tim: EasyInstaller... plan has changed since inception (see DS-802), might not be ready
- Robin: SWORD Client. Yes, sort of.
- Richard J: SWORD2 server. Probably not.
- Richard R: Context Guided Ingest: Cautiously optimistic.
- Mark: Discovery enhancements. Lots of new code in the module, needs committing to trunk
- Richard R: Creative Commons licensing rewrite. Still on schedule for inclusion in XMLUI. Needs volunteer for JSPUI. Robin pointed out we need to keep the Edit Item stuff in sync with this. Bonus: Keeping license text in metadata makes it all searchable, etc.
- Would be nice to get volunteers for: the quick wins in DC refactoring, dspace config split-up. Will throw them into JIRA.
- New Curation tasks:
 - Richard R: Format identification, virus scanning
 - Stuart: Link checker