DevMtg 2015-06-08 - OR15 Meeting

Developers / DCAT Meeting before OR15 on Mon, June 8, 2015

Face-to-face developer's meeting on topic of the DSpace RoadMap at OR2015 in Indianapolis.

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 may be limited. Please sign up for this "working meeting" as follows...

- Open Repositories conference asks that all attendees fill out the Workshop signup form at: http://www.or2015.net/workshops/
- We ask that you also add your name to the Sign Up sheet below. This list will be used to gage the anticipated audience, but it's not a formal attendance form.

Logistics

- Day: Monday, June 8th, 2015 (Workshop / Tutorial day)
- Time: 1:30pm - 5:00pm
    - NOTE: This working meeting is considered a "workshop", so it is listed as part of the Workshop/Tutorial day schedule.
    - **PLEASE be sure to SIGN UP for this workshop using the OR2015 Workshop signup form linked off of the Workshop/Tutorial day schedule**
    - A lunch is provided for all workshop attendees from 12:30pm-1:30pm. See the Conference Program for more details:
- Where: Hyatt Regency D
    - Hotel Floorplan at http://www.indianapolis.hyatt.com/hyatt/images/hotels/indri/floorplan.pdf
- Attendance Limit: 25-35 people maximum
- Post-Meeting:
    - After the meeting, there is an "Ideas Challenge / Developer Reception" at the Tomlinson Tap Room from 6-8pm. All are welcome!

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.

1:30pm : Introductions

- Name, Institution, How are you using DSpace?, How have you customized DSpace (or anything you'd be willing to contribute back)?
- Assign a note taker!

1:45pm - 3:00pm : DSpace Strategic Plan & Technical RoadMap

- How we got here (see Strategic Planning activities)
- Overview & Discussion of the 2015-18 Strategic Plan (especially concentrating on the "Technology" section) - Maureen Walsh (DCAT co-chair and Steering Group member, Ohio State)
    - *Please review DSpace 2015-18 Strategic Plan (final draft), and be prepared to share your (or your institution's) thoughts, questions, or concerns*
- Overview & Discussion of the Technical RoadMap (for 6.0, 7.0 and beyond) - Tim Donohue
    - *Please review RoadMap (final draft), and be prepared to share your (or your institution's) thoughts, questions, or concerns.*
    - *In addition, we are looking for individual or institutional volunteers to help on projects listed in the RoadMap.*

3:00pm - 3:30pm : Coffee Break

- Coffee to be provided in the Regency Foyer

<u>3:30pm - 5:00pm: RoadMap Continuation, 6.0 & Open Discussion</u>

- Continuation of above Strategic Planning & RoadMap discussions (if necessary)
- DSpace 6.0 Release planning:
  - Rough Timeline (*Proposed:* "Feature PR deadline" of Mid-to-Late September -> "Feature Freeze" Mid-to-Late October -> "Testathon" Mid-November -> "Final Release" Early-to-Mid December)
  - What features/fixes are at your institution that you'd like to offer for possible inclusion in 6.0?
    - Reminder: new features need a Pull Request before "PR deadline", and require documentation
  - Discussion of any anticipated "bigger changes" (i.e. large features), and when they will be ready for early testing/analysis
  - Conscript / gather volunteers for 6.0 Release Team (RT)
- *Time Permitting:*
  - Open Discussion Topics - Any topics that attendees have to discuss
  - Compare notes on any integration or acceptance testing work under-way (see DS-2288 and DS-2397)

Sign Up to Attend!

If you're planning to attend this meeting, please add your name to the sign up sheet. This will allow us to determine a proper headcount and ensure we have enough space for everyone.

Sign Up Sheet - Will Be Attending (If you are only able to attend part of the meeting, please note your approximate arrival)

1. Tim Donohue
2. Hardy Pottinger
3. Mark H. Wood
4. Ivan Masár
5. Andrea Schweer
6. Claire Knowles
7. Lisa Johnston
8. Terrence W Brady
9. Elin Stangeland
10. Graham Triggs
11. Monika Mevenkamp
12. Maureen Walsh
13. Sarah Potvin
14. Andrea Bollini (4Science) (Cineca)
15. Roberto Suardi (Cineca)
16. Sarah Molloy
17. Bram Luyten (Atmire)
18. Ignace Deroost (@mire)
19. Richard Rodgers (not sure of exact times)
20. Bill Ingram
21. Tara Stephens-Kyte (UBC)
22. Roeland Dillen
23. Leila Sterman (MSU)
24. Susan Borda
25. Suzi White
26. Jyrki Ilva
27. Erika A. Juarez Ibarra
28. Anis Moubarik
29. Catherine Cronin (NASA Langley)
30. Nicholas Webb (Mount Sinai Health System)
31. Daniel Draper (Colorado State Univ.)
32. Jose Carvalho (University of Minho)
33. Veronique Lecat
34. Maura Carbone (Brandeis University)

Possibly Attending / Unsure

- (*add your name here*)

Meeting Notes

DSpace Developer and DCAT Meeting Notes

Wiki page for the meeting: DevMtg 2015–06–08 – OR15 Meeting

- Building the ground work for strategic planning
- DSpace is eleven or twelve years old – this is a good time to take inventory and begin to do strategic planning based on the last few years of planning
- Planning has been done in a transparent way: all notes are on the wiki
- Feedback needed

Maureen Walsh: Overview of 2014–2015 Strategic Plan

- Community and Sustainability plans are the least developed; if you want to get more involved, they could use your help
- Goals focus on the next three years, but extend longer

Discussion on Community Goals

- Historically, DSpace has grown organically: if someone wants a new feature, it gets developed and contributed. It should be made clear that this roadmap is a departure from the way development has been done in the past.
- We do have a very active userbase, a lot of people are contributing. How do we sustain that?
- We are proposing that the model for DSpace development will have to change. We are proposing a much more managed process. We will be coordinating the work around releases and asking institutions to commit resources toward development sprints.
- Why we think this will work: Fedora 4
- We will need to change the rules for committers: as it is now, if enough committers agree on a feature, it gets in. Going forward, features should also be analyzed for whether they align with the roadmap.

Discussion on Sustainability Goals

- We need to be better at getting the message out: what is DSpace, what does it do, etc, what does it give you that others don't?
- The DSpace Marketing Group is working on putting together a draft marketing plan, which will be completed in a couple months
- Product Manager role is promotion of the product overall: outreach, gathering use cases, fund raising, etc.
- Currently looking for resources to fund a full-time Product Manager
- If you are in a position that you can make a personal appeal to a non-member, DSpace-using institution to join, then please do

Discussion on Technology Goals

- No plan for a complete DSpace rewrite, but a UI rewrite would be OK
- What are the fundamentals of the modern IR use case? What do we need to focus on?
- 130 or more use cases have been gathered and are being analyzed and prioritized
- These use cases drove the development of the road map
- Agility and flexibility are paramount goals
- Feedback is needed
- Non-developer DSpace end user comment: excited to see the proposed changes
- DCAT meeting is 2nd Tuesday of every month; all are welcome to join the call
- UI rewrite is the most significant: it is the most customized locally and will require a lot of support for adoption
- A lot seems to be riding on the future Product Manager; we shouldn't rely on them to solve these problems for us; we need to support them
- The earliest we can get the PM would be 2016
- There is no way to get around the problem of having two UIs
- Need to be reassured that the new interface will
- Who will do the UI work? There needs to be a full-time developer paid by DuraSpace.
- Several institutions have already promised resources to get this done. We think there are enough people lined up now to accomplish our goals
- What features will be moved out of the core functionality?
- We have some use cases that we have deemed high priority. One feature that will be moved out of the core is one of the duplicate statistics modules

Tim gives a brief version of his roadmap slide presentation. More will be discussed at the Thursday session.

Notes on roadmap presentation

- Concentration will be on the next two releases: 6 and 7
- c.f. RoadMap wiki page
- Anonymous feedback can be given on the wiki
- DSpace 6.0 will be the last ad-hoc-development release, but with an eye to the roadmap; 7.0 will be roadmap driven
- 7.0 will be late 2016 / early 2017
- Need to start planning immediately!

Discussion

- Idea: Virtual collections based on a search query. What is a collection other than a tag? What does it mean for something to be in a collection?
- There are wiki pages for each of the priority use cases
- API seems to be conspicuously missing. Is the idea that the new user interface will interact with a Web API that supports the entire suite of - - DSpace interactions
- Part of the agenda is to figure out what is missing in the current REST API. We will need to fill in those gaps.
- This is the time to make the API the canonical way to interact with DSpace
- Single UI will be a massive job; API will be another massive job; if I were to focus on one massive job, it would be the API
- Top down or bottom up?
- Developers agree that the core business logic is where we want to be.
- What is the difference between refactor and rewrite?
- Refactor can be done incrementally, whereas a complete rewrite would have to happen "behind closed doors" until it is completed
- Fedora did a lot of new development, which was a risky venture albeit a successful one
- Difference between us and Fedora: The Fedora group had enough resources, with committed institutions contributing $30,000. Fedora has about 250 institutions; DSpace has over two thousand. We need to figure out how to best harness that large user group, particularly from non-North-American institutions
- How can we reach out to the non-USA institutions?
- Mostly around what's going on in the UK: everything we are doing around repositories is in preparation of the open access ref
- Need to come back to the definition of *contribution*; in-kind contributions is a way forward; suggest we don't close the gate on random, ad-hoc improvements
- Can DSpace try to work more with the European funding models
- Repositories have changed since DSpace stared. We use it for a lot more things nowadays. DSpace has to be a place for other kinds of content
- Is DuraSpace able to accept smaller contributions? Full membership is a challenge.
- Is there marketing material for how to convince your boss to support DSpace?

BREAK

Notes on the roadmap presentation (cont.)

- Integrations to lower effort to deposit -
- Priority selection for 7.0 - Top priority items depend on whether there are developers available, not feasible to make everything priority 1
- Priority 2 - recommended for 7.0 but dependent on contributors
- Hierarchical metadata, MODS
- Enhanced basic statistical reports - based on data already available, but not already surfaced; use cases have been collected
- Basic preservation - format recognition, checksum tracking; potential advanced integration use cases, such as migration
- Easy and intuitive deposit UI - still collecting use cases for this, high priority but not developed enough yet
- Beyond 7.0 - priorities
- Simplify installation and upgrade - some work done on this, but more work to do; may be dependent on modularisation of DSpace hence lower priority right now
- Theme management - once single UI has been developed, focus could
- Replacing "bundle" with metadata flags on bitstreams
- Prioritised integrations - dependent on UI framework

How will we achieve this?

- Managed projects - based on part time, donated developers with scheduled (1-2 week) development sprints
- More volunteers/contributors needed to achieve this; some organisations have stepped forward already but get in touch

DSpace Single UI - 7.0 release late 2016

- Prototype (at least) 2-3 UI platforms
  - at least 1 Java-based
  - at least 1 non-Java
- By the end of 2015 intention to approve a single platform/prototype and finalize the developer team/schedule
- Early 2016 - developer sprints begin in readiness for a 7.0 release

Discussion

- 3 UIs in prototype but only one will go forward as the new UI for 7.0
- Clear path for moving to new UI for users
- Need input from the DSpace community early
- Make the process as open as possible
- Goal at OR2016 would be to present some kind of beta
- Resources for testing will be stretched and a lot of effort from committers with testing for 6.0 candidate, plus work on prototype
- Prototyping may have to be done by people who are not committers
- Need to scope now what 6.0 release contributions are with an earlier cut off
- Suggestion that 6.0 could be a larger bug fix release and hold off major changes to UI based on potential single UI in later release
- Does there need to be a 6.0 release this year?
- There was some discussion about this in the roadmap meetings
- Feeling was that there are things that institutions would like to get into 6.0 as soon as possible
- There is some development in process around the API ❓ on the 6.0 status page - Atmire have been developing this
- High risk not to do a 6.0; roadmap is speculative at the moment, and targeted for end of 2016 but this cannot be promised potential for a big gap between major releases
- More people really needed, not just for members, but also developers/contributors who can help and also test
- More contributors fundamental to the roadmap

Other 7.0 priorities

- Projects will be established for these with a similar development model as the single UI project, not as developed at the moment
- Mentoring may be possible to build on the committer team and bring on developers/contributors to become committers

How you can help

- Become a DuraSpace member
  - Helps toward hiring a Product Manager
  - Voice in DSpace governance & roadmap

- Donate a developer (part time)
  - choose a project
  - get in touch
  - sprints get scheduled by DuraSpace

Discussion

- Acknowledgement of the amount of work that has gone into developing the roadmap and strategies
- Support for open discussion and use of DCAT - this would be continue for future projects

DSpace 6.0 planning

- 6.0 status page - DSpace Release 6.0 Status
- Highest priority is to clean up search/browse

- Basic priorities have been identified
- Atmire - a few open issues on JIRA that Atmire are working on, should be resolved through work being undertaken with Sydney at the moment
- Collection quality control management has been requested by Georgetown
- Rough list at the moment
- Pull request nearly ready for the OAI harvester - CottageLabs
- 'Code Review Needed' - flagged items mean that something is ready for testing and developers would be interested in feedback
- Texas A&M - SAF (Simple Archive Format) builder working with batch loading, added capabilities for special licences
- DSpace Labs project - cool tools that have been developed/under developed where people can go and test these and also work collaboratively
- Licences - access control licences, specifically for datasets where need special licence request that prompts login
- Special agreements page
- Command line utility for reporting
- More contributions welcome for the REST API - gaps to be identified
- More use cases very welcome - there are lots there already
- None of the use cases cover 'Review'
- Add new use cases using the template here: Use Cases, also add to and comment on current use cases

Discussion

- What is the migration status of DCTerms? - DCTerms schema came with version 5.0 (in addition to Dublin Core)
- What is the plan for completing m
- Ticket for this work open in JIRA
- DSpace-general - list is very busy

Unable to locate Jira server for this macro. It may be due to Application Link configuration.

om/

Questions

- IRC - is there intention to move to web/audio meetings?
- IRODs - is there current integration for this for Dspace? EUDAT have been working on something in this area
- Integrations - is there somewhere that we can survey and record who is working on specific integrations? Such as a Google Spreadsheet?
Suggestion to contact DCAT in the first instance and then look at ways to survey the community