

# DevMtg 2018-09-12

## Developers Meeting on Weds, September 12, 2018

### Today's Meeting Times

- DSpace Developers Meeting / Backlog Hour: 20:00 UTC in [#duraspace IRC](#) or [#dev-mtg Slack channel](#) (these two channels sync all conversations)
  - Please note that all meetings are [publicly logged](#)

### Our IRC logging bot has been blocked from Freenode (as of July 27).

Discussion logs are no longer available at <http://irclogs.duraspace.org/>. As our current IRC log bot (based on [PircBot](#)) is unmaintained and doesn't align with Freenode policies (around requiring SASL authentication), Tim has reached out to <https://botbot.me/> to see if they could log our [#duraspace IRC](#) channel. In the meantime, full logs of meeting discussions will be copied into the Wiki notes below.

## Agenda

### Quick Reminders

*Friendly reminders of upcoming meetings, discussions etc*

- [DSpace 7 Working Group](#): Next meeting is tomorrow, Thurs, September 13 at 14:00 UTC
  - **Signup for our next Community Sprint (Oct 1-12): [DSpace 7 Community Sprints](#)**
- [DSpace 7 Entities Working Group](#)
  - Latest meeting (including video recording): [2018-09-11 DSpace 7 Entities WG Meeting](#)
  - Early prototype writeup / docs: <https://docs.google.com/document/d/1wEmHirFzrY3qgGtRr2YBQwGOvH1luTVGmxDldnqvwxM/edit>
- [DSpace Developer Show and Tell Meetings](#): Yesterday's recording is now posted to [DSpace On DockerHub](#)

## Discussion Topics

*If you have a topic you'd like to have added to the agenda, please just add it.*

1. (Ongoing Topic) [DSpace 7 Status Updates](#) for this week (from [DSpace 7 Working Group](#))
  - a. [DSpace 7 Development Status spreadsheet](#)
2. (Ongoing Topic) [DSpace 6.x Status Updates](#) for this week
  - a. 6.4 will surely happen at some point, but no definitive plan or schedule at this time. Please continue to help move forward / merge PRs into the [dspace-6.x](#) branch, and we can continue to monitor when a 6.4 release makes sense.
3. [DSpace Release 5.10 Status](#)
4. Brainstorms / ideas (*Any quick updates to report?*)
  - a. [Bulk Operations Support Enhancements](#) (from [Mark H. Wood](#))
  - b. [Curation System Needs](#) (from [Terrence W Brady](#))
    - i. [PR 2181](#) implements per-run task parameters. Ready for review.
    - ii. [PR 2180](#) improves reporting. Needs a little more testing.
  - c. (*No updates to report*) How to encourage / credit folks who do Code Reviews? ([Tim Donohue](#))
    - i. No major updates this week. Investigating [GitHub GraphQL](#) to query for "PR Creators", "PR Reviewers", etc. and then parsing /compiling those JSON results into "Top 10" lists (currently playing with "jq", a [json command-line parser](#)). Final display of results yet to be determined.
    - ii. Good promise here, but nothing fully "working" yet (haven't had as much time as hoped to dig into this). Should (hopefully) have code/queries to share by next week?
5. Tickets, Pull Requests or Email threads/discussions requiring more attention? (*Please feel free to add any you wish to discuss under this topic*)
  - a. Quick Win PRs: <https://github.com/DSpace/DSpace/pulls?q=is%3Aopen+review%3Aapproved+label%3A%22quick+win%22>

## Tabled Topics

These topics are ones we've touched on in the past and likely need to revisit (with other interested parties). If a topic below is of interest to you, say something and we'll promote it to an agenda topic!

1. Management of database connections for DSpace going forward (7.0 and beyond). What behavior is ideal? Also see notes at [DSpace Database Access](#)
  - a. In DSpace 5, each "Context" established a new DB connection. Context then committed or aborted the connection after it was done (based on results of that request). Context could also be shared between methods if a single transaction needed to perform actions across multiple methods.

- b. In DSpace 6, Hibernate manages the DB connection pool. Each **thread** grabs a Connection from the pool. This means two Context objects could use the same Connection (if they are in the same thread). In other words, code can no longer assume each new `Context()` is treated as a new database transaction.
- i. Should we be making use of `SessionFactory.openSession()` for READ-ONLY Contexts (or any change of Context state) to ensure we are creating a new Connection (and not simply modifying the state of an existing one)? Currently we always use `SessionFactory.getCurrentSession()` in `HibernateDBConnection`, which doesn't guarantee a new connection: [https://github.com/DSpace/DSpace/blob/dspace-6\\_x/dspace-api/src/main/java/org/dspace/core/HibernateDBConnection.java](https://github.com/DSpace/DSpace/blob/dspace-6_x/dspace-api/src/main/java/org/dspace/core/HibernateDBConnection.java)

## Ticket Summaries

1. Help us test / code review! These are tickets needing code review/testing and flagged for a future release (ordered by release & priority)

Key	Summary	T	Created	Updated	Assignee	Reporter	P	Status	Fix Version/s
DS-3962	Improve community-list performance	↑	Jul 23, 2018	Jul 10, 2020	Unassigned	April Herron (Atmire)	☰	CODE REVIEW NEEDED	6.5
DS-4303	Old (v4-6) REST API is deprecated in v7, remove in v8	☰	Jul 17, 2019	Mar 02, 2020	Unassigned	Tim Donohue	⊖	CODE REVIEW NEEDED	8.0
DS-3879	Enable XInclude for some XML configuration documents	↑	Mar 28, 2018	Feb 28, 2020	Mark H. Wood	Mark H. Wood	☰	CODE REVIEW NEEDED	7.0
DS-2670	Restrospectively create DOI's for alle items in archive	↑	Jul 22, 2015	Jun 22, 2018	Mark H. Wood	Claudia Jürgen	☰	CODE REVIEW NEEDED	7.0
DS-3952	Expose Request Copy	+	Jul 12, 2018	Oct 31, 2018	Mark H. Wood	Mark H. Wood	☰	CODE REVIEW NEEDED	7.0
DS-4370	Allow structure builder to import beneath an existing community	↑	Oct 21, 2019	Oct 21, 2019	Kim Shepherd	Kim Shepherd	☰	CODE REVIEW NEEDED	7.0
DS-4145	Curation basic link checker doesn't follow redirects	+	Jan 14, 2019	Jan 14, 2019	Kevin Van de Velde (Atmire)	Kevin Van de Velde (Atmire)	☰	CODE REVIEW NEEDED	7.0
DS-4289	SAF exporter does not write the 'collections' file	↑	Jul 01, 2019	Jul 10, 2019	Mark H. Wood	Mark H. Wood	☰	CODE REVIEW NEEDED	7.0
DS-4118	change_password and register emails have hardcoded sample addresses	⊖	Dec 17, 2018	Feb 05, 2020	Mark H. Wood	Rob Jenson	☰	CODE REVIEW NEEDED	7.0
DS-3385	OAI filter to exclude set from harvesting	↑	Nov 14, 2016	Nov 23, 2016	Unassigned	Ivan Masár	☰	CODE REVIEW NEEDED	7.0
DS-3749	Selective enabling of identifier generation per-container	+	Nov 02, 2017	May 10, 2021	Mark H. Wood	Mark H. Wood	☰	CODE REVIEW NEEDED	7.0
DS-4537	OAI index doesn't update for deleted items	↑	Jul 14, 2020	Jul 17, 2020	Unassigned	Andrew Wood Atmire	☰	CODE REVIEW NEEDED	7.0
DS-4300	Optionally apply given Handles when importing community /collection structure	+	Jul 15, 2019	Jan 06, 2020	Mark H. Wood	Mark H. Wood	☰	CODE REVIEW NEEDED	7.0
DS-3951	Send Request Copy email to multiple recipients, and some new recipient strategies	+	Jul 12, 2018	Jul 12, 2018	Mark H. Wood	Mark H. Wood	☰	CODE REVIEW NEEDED	7.0
DS-1494	Enable separation of LDAP search credentials from DSpace configuration	↑	Feb 20, 2013	Jul 30, 2018	Mark H. Wood	Brian Freels-Stendel	☰	CODE REVIEW NEEDED	7.0
DS-4410	Send a "welcome letter" to new users	+	Jan 02, 2020	Jan 03, 2020	Mark H. Wood	Mark H. Wood	☰		7.0

Key	Summary	T	Created	Updated	Assignee	Reporter	P	Status
DS-4522	Logical item filtering (with implementation of filtered DOI provider)	+	Jun 08, 2020	Jun 24, 2020	Kim Shepherd	Kim Shepherd	☰	CODE REVIEW NEEDED
DS-4574	Upgrade DBCP dependency	↑	Feb 23, 2021	Jul 06, 2021	Unassigned	Hrafn Malmquist	☰	CODE REVIEW NEEDED
DS-4580	Authority control doesn't update when editing item	⬮	Apr 25, 2021	Apr 25, 2021	Hrafn Malmquist	Hrafn Malmquist	☰	CODE REVIEW NEEDED
DS-3849	REST API items resource returns items in non-deterministic order	⬮	Feb 23, 2018	Mar 28, 2020	Bram Luyten (Atmire)	Andrew Bennet	☰	CODE REVIEW NEEDED

Showing 20 out of 39 issues

2. Newly created tickets this week:

Key	Summary	T	Created	Assignee	Reporter	P	Status
DS-4588	Incorrect spider detector test	⬮	Jul 22, 2021	Unassigned	Alan Orth	☰	CODE REVIEW NEEDED
DS-4587	DSpace's built-in list of spider/robot user agents is old	↑	Jul 22, 2021	Unassigned	Alan Orth	☰	CODE REVIEW NEEDED
DS-4586	Item versioning: enable permanent handle reference to the most recent version of each items	↑	Jul 21, 2021	Unassigned	Jozsef Marton	☰	RECEIVED
DS-4585	Incorrect escaping of citation_ meta tags when metadata contains quote characters	⬮	Jul 20, 2021	Unassigned	Alan Orth	☰	CODE REVIEW NEEDED

4 issues

3. Old, unresolved tickets with activity this week:

Key	Summary	T	Created	Updated	Assignee	Reporter	P	Status
DS-1980	"Private" Items should include a "noindex" meta tag in the HTML	⬮	Apr 21, 2014	Jul 23, 2021	Unassigned	Tim Donohue	☰	CODE REVIEW NEEDED

1 issue

4. Tickets resolved this week:

Key	Summary	T	Created	Assignee	Reporter	P	Status	Resolution
DS-4579	Re-enable partial downloads/Range requests	↑	Apr 07, 2021	Hrafn Malmquist	Hrafn Malmquist	☰	CLOSED	Fixed

1 issue

5. Tickets requiring review. This is the JIRA Backlog of "Received" tickets:

Key	Summary	T	Created	Updated	Assignee	Reporter	P
DS-4586	Item versioning: enable permanent handle reference to the most recent version of each items		Jul 21, 2021	Jul 22, 2021	Unassigned	Jozsef Marton	
DS-4584	Cross site scripting		Jun 08, 2021	Jun 09, 2021	Unassigned	Jaime Solorzano	
DS-4583	SQL injection		Jun 08, 2021	Jun 09, 2021	Unassigned	Jaime Solorzano	
DS-4576	Export/import of Solr authority core with SolrImportExport not handling ORCID "label" multi-value fields properly		Feb 25, 2021	Feb 25, 2021	Unassigned	Sean Kalynuk	
DS-4573	Filter Media relies on Bitstream Name rather than UUID, skipping files with identical names		Feb 23, 2021	Feb 23, 2021	Unassigned	Claudia Jürgen	
DS-4572	Deleted Bitstream corresponding extracted files not deleted		Feb 23, 2021	Feb 23, 2021	Unassigned	Claudia Jürgen	
DS-4571	Registration allows to register more than once with the same email		Feb 18, 2021	Feb 18, 2021	Unassigned	Santiago Tettamanti (SEDICI)	
DS-4570	UUID fragments included in Legacy Usage Reports (log file based usage statistics)		Feb 12, 2021	Feb 12, 2021	Unassigned	Anne Lawrence	
DS-4569	Date Display in Browse Indices		Feb 01, 2021	Feb 01, 2021	Unassigned	Claudia Jürgen	
DS-4568	Possible to create multiple epersons with the same email address		Jan 29, 2021	Jan 29, 2021	Unassigned	Anne Lawrence	
DS-4567	Solr statistics error		Jan 26, 2021	Jan 26, 2021	Unassigned	Humberto Blanco	
DS-4566	Export search results feature incorrectly escapes query parameters in search filters		Jan 26, 2021	Jan 26, 2021	Unassigned	Alan Orth	
DS-4565	Unable to update archived Item using SWORDv2 as collection / community administrator		Jan 19, 2021	Jan 19, 2021	Unassigned	Jakub íhák	
DS-4558	RTS - Update bagit-support to 1.1.0		Oct 21, 2020	Oct 21, 2020	Michael Ritter	Michael Ritter	
DS-4556	Replication Task Suite: Support the reporting of individual failures when part of a larger curation action		Oct 14, 2020	Oct 14, 2020	Unassigned	Bill Branan	
DS-4553	XMLUI generates useless links		Sep 29, 2020	Sep 29, 2020	Unassigned	Hrafn Malmquist	
DS-4550	OpenAIRE and dc.relation		Sep 01, 2020	Sep 02, 2020	Unassigned	Gabriela Mircea	
DS-4549	Significant performance issues when displaying items with many (100+) bitstreams		Sep 01, 2020	Sep 01, 2020	Unassigned	Martin Walk	
DS-4548	org.dspace.core.LogManager is poorly named		Aug 31, 2020	Aug 31, 2020	Mark H. Wood	Mark H. Wood	
DS-4547	Google Analytics code should be moved to <head>		Aug 25, 2020	Aug 25, 2020	Unassigned	Bill Branan	

Showing 20 out of 94 issues

## Meeting Notes

### Meeting Transcript (IRC Bot is not working)

- IRC Transcript is available at <http://irclogs.duraspace.org/index.php?date=2018-08-15>

Log from #dev-mtg Slack (All times are CDT)

Tim Donohue [3:01 PM]

@here: it's DSpace DevMtg time. As noted in #dev, the agenda is at: <https://wiki.duraspace.org/display/DSPACE/DevMtg+2018-09-12>

Let's do a quick roll call to see who is able to join us today

Mark Wood [3:02 PM]

Hi

Tim Donohue [3:03 PM]

Hi Mark, welcome. Hmm...we may be a small group today. Well, I'll do a quick few announcements and we'll see who else joins along the way

First, and perhaps *\*most importantly\**, reminder that our next DSpace 7 Community Sprint is coming up Oct 1-12. Signup here: <https://wiki.duraspace.org/display/DSPACE/DSpace+7+Community+Sprints>

DSpace 7 Entities had its second meeting yesterday. Video Recording is at: <https://wiki.duraspace.org/display/DSPACE/2018-09-11+DSpace+7+Entities+WG+Meeting>

That's pretty much it for DSpace 7 updates at this time. (Though next Dspace 7 weekly meeting is obviously tomorrow at 14UTC)

No 6.x updates

So, that's basically the first two agenda items :slightly\_smiling\_face:

We do have an upcoming 5.10 release (code is all merged on `dspace-5\_x` branch, but being tested).

@terrywbrady is the release manager, but isn't able to be here today

More info on that is at <https://wiki.duraspace.org/display/DSPACE/DSpace+Release+5.10+Status>

And if anyone here would like to see 5.10 happen *\*sooner\**, get in touch with Terry to see how we can get this out the door quickly.

I'm guessing it's still just you and me, @mwood. Anything else you want to discuss here?

Mark Wood [3:09 PM]

I added a bit to discussion topic 4b: curation system. There are a couple of PRs to look at, if anyone is interested. Terry has commented on one.

Thanks to a wet weekend when I couldn't do yard work....

Tim Donohue [3:10 PM]

Aha, yes, I see those now. Are these PRs connected in any way (does one depend on other)?

Mark Wood [3:11 PM]

They are mostly independent. One line will have to be fixed in one when the other is merged.

Tim Donohue [3:11 PM]

Also, it might be good to note how to test these (tips/hints). I'm assuming it'll be semi-obvious on reviewing code, but it always helps testers to know exactly what to do.

Mark Wood [3:12 PM]

I'll see what I can add, on that. There are automated tests included.

Tim Donohue [3:13 PM]

RE: PR#2181, do we actually have any curation tasks that are updated to *\*use\** properties, or is this just establishing the code to allow (future) tasks to use properties?

Mark Wood [3:13 PM]

I'll do a little manual testing of the workflow-linked task reporting, since that one's really hard to automate.

Tim Donohue [3:13 PM]

<https://github.com/DSpace/DSpace/pull/2181>

Mark Wood [3:14 PM]

Task properties are a nice idea that never really got taken up. I don't recall whether any stock tasks use them. There is a test task included in the patch, which does.

Tim Donohue [3:15 PM]

err..by "properties", I guess I mean "parameters"... Is there a way to test this using existing curation tasks, and simply pass it new parameters (to see behavior change)?

Documentation could clear this up a bit more.

Mark Wood [3:16 PM]

I just did a quick grep through `org.dspace.ctask.general` and didn't find any use of properties, alas.

Yes, I should work over the curation documentation. For one thing, it's *\*all\** under developer stuff in the Official Doco. It should be split into development and use.

Tim Donohue [3:18 PM]

I'm not sure if I have my lingo mixed up here, or if we're talking about two different things. I guess I'm

asking for more docs on what this feature "adds". Does `-p`` option mean you can override configuration settings? Does it only add to configuration settings? Is it only usable if a curation tasks specifically "looks for" params?

I'm just trying to understand this new feature -- it sounds interesting, but there's no info in the PR or ticket on what `-p`` is (edited)

Mark Wood [3:20 PM]

From a task's point of view, there are only task properties. From the user's point of view, there are properties (which are set in the DSpace configuration) and run parameters (which are set by `-p`). Curator effectively merges the two groups.

Good point, I'll explain more thoroughly in the PR.

Tim Donohue [3:20 PM]

Which one "wins" if you have a configuration & run parameter of the same name?

Thanks, yes, I think a PR explanation (even just a few examples or sentences) would make this a lot clearer

Mark Wood [3:21 PM]

Run parameters are consulted first, so they override configured properties.

DSpaceSlackBot (IRC) APP [3:21 PM]

\*karias\* has quit the IRC channel

Mark Wood [3:21 PM]

None of this makes any difference to a task that doesn't ask its Curator for property values.

That's unchanged from before.

Tim Donohue [3:22 PM]

Ok, makes sense. I'm just trying to figure out how to help get this tested...and how to better document it.

So, it sounds like adding an explanation of this feature to the PR would be nice...any testing hints/tips would be nice.

Mark Wood [3:23 PM]

My understanding is that task properties were introduced to make it possible for a given CurationTask class to run differently when invoked under different names. Run parameters just piggyback on this.

I will update the PR.

(There's a mapping from fully qualified class names to short "taskIDs" in `dspace.cfg`, so you can map multiple names to one class.)

Tim Donohue [3:25 PM]

FWIW, there is an example of using Task Properties for "thumbnail settings" in the official documentation:

<https://wiki.duraspace.org/display/DSDOC6x/Curation+System#CurationSystem-TaskProperties>

Not sure if that's helpful in documenting or testing this feature, but it may be

Mark Wood [3:26 PM]

OK, thanks.

Tim Donohue [3:26 PM]

Thanks though...glad to see these moving along. I'll also try to take a closer look at the code here when I get a chance

Mark Wood [3:26 PM]

Appreciated.

There is some discussion in the other PR about just which part of curation controls the reporting.

Tim Donohue [3:27 PM]

I honestly don't have anything else to discuss today...and it seems to be just two of us. Shall we close up the meeting? Or did you have other things you wanted feedback on?

Mark Wood [3:27 PM]

That's all I have right now.

Tim Donohue [3:28 PM]

I hadn't looked at the other PR (2180) either yet...but will try to do so soon. (edited)

Mark Wood [3:28 PM]

Thanks!

Tim Donohue [3:28 PM]

Ok, let's close up the meeting & leave it to a quick ~30 mins. Have a good rest of your day!

Mark Wood [3:29 PM]  
Good day!