DevMtg 2019-07-31

Developers Meeting on Weds, July 31, 2019

Meeting on Hiatus starting August 2019

🚫 *This meeting is on temporary hiatus and will no longer be held on a regular (weekly) basis.* This meeting will resume as soon as one (or more) of the following occurs:

- The DSpace 7 Working Group (2016-2023) finishes its work & ramps down
- Or, the DSpace 6.4 release effort ramps up.
- Or, an interim chair for this meeting is located (as Tim is unavailable to chair this meeting until DSpace 7 finishes up)

In the meantime, ad hoc discussions or meetings are encouraged to occur via the #dev channel in Slack. Developers wanting status updates (e.g. on DSpace 7), or looking for feedback/advice on ongoing work are encouraged to ask for support on that channel, and we would be glad to help out.

Anyone is also welcome & encourages to attend (or "lurk") on DSpace 7 Working Group (2016-2023) or DSpace 7 Entities Working Group (2018-19) meetings (both occur weekly). You need not have a microphone or camera to attend, and can take part or ask questions via text chat in Zoom.

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)
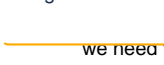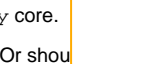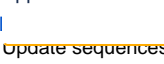
Agenda

Quick Reminders

Friendly reminders of upcoming meetings, discussions etc

- DSpace 7 Working Group (2016-2023): Next meeting is Thurs, Aug 1 at 14:00 UTC.
- DSpace 7 Entities Working Group (2018-19): Next meeting is Tues, Aug 6 at 15:00 UTC.
- DSpace Developer Show and Tell Meetings: On hold until interesting topics arise.
- 2019 DSpace North American User Group Meeting is coming up Sept 23-24.

Discussion Topics

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

1. Quick Updates from other meetings
   a. DSpace 7 Status Updates for this week (from DSpace 7 Working Group (2016-2023) or DSpace 7 Entities Working Group (2018-19))
   b. DSpace 6.x Status Updates for this week

       i. 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.
2. **This general Developers Mtg will go on hiatus, until one (or more) of the following occurs**
   a. DSpace 7 Working Group finishes work / ramps down
   b. Or, DSpace 6.4 effort ramps up
   c. Or, an interim chair for this meeting is found.
   d. (In the meantime, Tim will provide status updates on DSpace 7 efforts via Slack and/or email lists. Also, anyone is welcome to drop in on DSpace 7 Working Group (2016-2023) or DSpace 7 Entities Working Group (2018-19) just to "lurk" or listen in on efforts.)
3. Ongoing Work
   a. Upgrading Solr Server for DSpace

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

       s only happen for **major** releases? Should it be configurable? Can we find a more precise trigger? When **do** we need to reindex?

       ii.   ty core.   Unable to locate Jira server for this macro. It may be due to Application Link configuration.   Or shou

   b. DSpace   ls (Terrence W Brady)

       i. Update sequences on initialization

           1. https://github.com/DSpace/DSpace/pull/2362 - update sequences port
           2. https://github.com/DSpace/DSpace/pull/2361 - update sequences port

       ii. DSpace Launcher Dashboard - Deploy a PR on AWS for Testing
           1. There is a 2 minute video that illustrates this proposal.
       iii. Docker for production - next steps
4. 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. Brainstorms / ideas
    a. (*On Hold, pending Steering/Leadership approval*) Follow-up on "DSpace Top GitHub Contributors" site (Tim Donohue): https://tdonohue.github.io/top-contributors/
    b. Bulk Operations Support Enhancements (from Mark H. Wood)
    c. Curation System Needs (from Terrence W Brady)
2. Management of database connections for DSpace going forward (7.0 and beyond). What behavior is ideal? Also see notes at DSpace Database Access
    a.

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

    b. ~~~~~~nection 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.
    c. 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
    d. Bulk operations, such as loading batches of items or doing mass updates, have another issue: transaction size and lifetime. Operating on 1 000 000 items in a single transaction can cause enormous cache bloat, or even exhaust the heap.
        i. Bulk loading should be broken down by committing a modestly-sized batch and opening a new transaction at frequent intervals. (A consequence of this design is that the operation must leave enough information to restart it without re-adding work already committed, should the operation fail or be prematurely terminated by the user. The SAF importer is a good example.)
        ii. Mass updates need two different transaction lifetimes: a query which generates the list of objects on which to operate, which lasts throughout the update; and the update queries, which should be committed frequently as above. This requires *two* transactions, so that the updates can be committed without ending the long-running query that tells us what to update.

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	type	created	updated	assignee	reporter	priority	status	fixversions

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

2. Newly created tickets this week:

key	summary	type	created	assignee	reporter	priority	status

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

3. Old, unresolved tickets with activity this week:

key	summary	type	created	updated	assignee	reporter	priority	status

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

4. Tickets resolved this week:

key	summary	type	created	assignee	reporter	priority	status	resolution

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

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

key	summary	type	created	updated	assignee	reporter	priority

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

Meeting Notes

Meeting Transcript

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

```
Tim Donohue 3:00 PM
@here: It's time for our general DSpace DevMtg for this week. Last minute agenda was just posted to
https://wiki.duraspace.org/display/DSPACE/DevMtg+2019-07-31

Mark Wood 3:00 PM
Hi
Tim Donohue 3:00 PM
Let's do a quick roll call to see who is joining today
Terry Brady 3:00 PM
hello
Tim Donohue 3:01 PM
Looks like just the usual crew... welcome Mark & Terry
So, today, I'm going to jump right into topic #2 on the agenda...as I'd like to allow for plenty of time for
questions/concerns.
After a lot of consideration, I've decided that I need to step away from this meeting on a weekly basis...I've
got too many meetings I'm chairing, and too much on the plate in getting DSpace 7 out the door
Therefore, this meeting will go on indefinite hiatus
more specifically, until either DSpace 7 wraps up, DSpace 6.4 ramps up, or someone else decides they want to
chair this meeting (quite possibly without my regular attendance, until DSpace 7 starts to wrap up)
Terry Brady 3:04 PM
There is a chicken and egg bit.  If this meeting doesn't happen, will there be a place to plug people in for
production issues?
Then again, attendance has been really low for some time.
Tim Donohue 3:05 PM
That's the basics, but I welcome discussion and/or questions.  I'm still more than willing to have ad hoc
meetings to discuss topics that may not be getting enough attention in DSpace 7 WG meetings (or Entities
meetings).  But, I'd like those to be ad hoc
```

Mark Wood 3:05 PM
Understandable.
Tim Donohue 3:06 PM
@terrywbrady: I totally get it. But, as you noted, this meeting is sparsely attended...after a lot of
realization that I have too many meetings (on a weekly basis), this is the least attended and least active
meeting on that list :disappointed:
Terry Brady 3:06 PM
I like the dependencies that you have listed. I think those make sense.
Tim Donohue 3:06 PM
Doesn't meant I won't miss the meeting...and I do hope to restart it sometime in the near future :
slightly_smiling_face:
Terry Brady 3:07 PM
If this meeting is resurrected, I recommend moving away from the chat-based format.
Tim Donohue 3:08 PM
@terrywbrady: there would also be the opportunity for that as well. Honestly, once the DSpace 7 WG starts to
"ramp down", the big question in my mind will also be whether to simply "morph" that into the new DevMtg, or do
we go back to chat-based. It will be a question I will ask
Mark Wood 3:08 PM
Ew, not another video thing. How do you 'grep' a pile of videos?
Tim Donohue 3:09 PM
I don't want to make the decisions on the future of this meeting now. :slightly_smiling_face: But, I'd
encourage everyone to think about how to "recreate" an engaging developer meeting (once the time comes to do so)
In the meantime, honestly I'll do my best to still give regular DSpace 7 updates in Slack (#dev) and/or via
mailing lists. I want to ensure people are aware of what is going on and feel looped in. Also, if topics come
up that need a meeting, we can always borrow this meeting's old timeslot
Mark Wood 3:10 PM
I suspect that a general meeting will become more engaging once everybody is not thinking "I have too many
meetings too, and DS7 deadlines looming...."
:+1:
1

Tim Donohue 3:12 PM
So, "going on hiatus" does mean I'll take this meeting off the DSpace calendar (starting in August, i.e. now).
I'll post a notice on the wiki page (https://wiki.duraspace.org/display/DSPACE/Developer+Meetings) about the
hiatus, and the terms I stated about when this meeting will restart
I'll also note that ad hoc meetings (or general developer advice) can be asked for at any time on #dev Slack
:+1:
1

(and a reminder that all are welcome in the various working group meetings as well, obviously...even just to
lurk)
Any other thoughts/questions/concerns here (including from anyone lurking and reading along)?
Mark Wood 3:15 PM
Updates on #dev will at least keep ongoing development in front of everybody's eyes. I don't want to see
people wander off, and wind up with 1000 forks of DSpace with no "center" like some projects I could name.
:+1:
1

Tim Donohue 3:16 PM
@mwood: completely agreed. I don't think that'll happen :slightly_smiling_face:
And quite frankly...please keep me honest here. If you (or anyone) want an update on DSpace 7 or a specific
feature / concept please ask on #dev and ping me by name
Ian Little 3:17 PM
I've found this meeting more useful than video, as someone randomly jumping in.
Tim Donohue 3:17 PM
I'm always willing to clarify or answer questions. I might not even realize there's confusion out there...but,
I'm always glad to help clarify or add more info, or provide advice, etc
@Ian Little: welcome, and thanks for the feedback :slightly_smiling_face:
FWIW, as a side note, I have started summarizing any Zoom meetings we have (for working groups) in "Notes" on
the meeting pages themselves. So, that's another area to get updates...it's just a summary though, not all the
detailed discussion
Mark Wood 3:19 PM
Thank you.
Tim Donohue 3:20 PM
Ok, that's all I had to say, and I'm not seeing anyone typing...so, I'll leave it at that. Again though,
always glad to answer questions or give advice via Slack or email (though I'll admit to having a preference to
being asked questions more publicly when it makes sense to do so...as it lets others also see the answer and
ask followups if they have any)
So, jumping back up to topic #1 on the agenda...updates from other meetings. I'll admit, not a ton of "big
things" happening right now. But, both DSpace 7 WG and DSpace 7 Entities WG are active, and the meetings are
well attended overall.

As noted, I've been taking notes / summarizing all discussions on the meeting pages. So, for example, the notes for last week's DSpace 7 WG are here: https://wiki.duraspace.org/display/DSPACE/2019-07-25+DSpace+7+Working+Group+Meeting#id-2019-07-25DSpace7WorkingGroupMeeting-Notes

Mark Wood 3:23 PM
TBH all the planning is going on in WG meetings right now. Presumably many of us will be back here to figure out what DS8 will be. :slightly_smiling_face:

Tim Donohue 3:23 PM
Yes, definitely :slightly_smiling_face:
Notes from last week's DSpace 7 Entities WG are also available here: https://wiki.duraspace.org/display/DSPACE/2019-07-30+DSpace+7+Entities+WG+Meeting (Just realized I need to make a direct link to the "Notes" section, it's at the bottom of the page though)
So, I have no major updates on DSpace 7.... work is ongoing, and there's a lot of activity. At the same time, there's quite a bit left to do. In the coming weeks, we'll be working to better estimate remaining work (some scoping of that has already started, but other areas are being discussed and scoped in coming weeks)
Any other questions I can answer though on DSpace 7 or the working group meetings in general?

Mark Wood 3:26 PM
None here.

Tim Donohue 3:27 PM
Ok, that's it for those general updates then
Jumping down to topics under "Ongoing Work" (topic #3)
And, as I sidenote, I'd still like to hear updates here on the subtopics in this section...I'd recommend they move to the DSpace 7 WG (where applicable) or just general Slack updates

Mark Wood 3:28 PM
OK

Tim Donohue 3:28 PM
First up though, Solr Upgrade for DSpace 7. Any updates this week, @mwood? https://wiki.duraspace.org/display/DSPACE/Upgrading+Solr+Server+for+DSpace

Mark Wood 3:29 PM
I've pulled together a test instance with several years of stat. data. And I just realized that I need to patch in the tool for updating the stat. core with UUIDs, which is not released yet.
The statistics pages look weird when there are no stat.s at all.
Loading these data was a good preliminary check of the upgrade process. I am more nearly certain that we won't need my new authority dumper/loader; the statistics tool will do that too, if we document it more fully.

Tim Donohue 3:31 PM
Good to hear no extra upgrade tools need to be built for that!

Mark Wood 3:31 PM
After I get the UUIDs set and can see some stat.s in the UI, I'll actually run an upgrade to 7 and see what breaks.
That's all for now.

Tim Donohue 3:32 PM
Sounds good. And just to clarify, the UUID issue you are talking about is just ensuring your old (presumably 5.x or below) stats are first updated to 6.x compatibility?

Mark Wood 3:32 PM
(I also need to patch that tool into the production instance from which I got the data, because it won't be showing any stat.s either!)

Tim Donohue 3:33 PM
So, you are trying to get to 6.x stats first...then doing the 7 upgrade?

Mark Wood 3:33 PM
Yes, I got the data from a v5 system that we upgraded to 6.
Yes, I took data from a 6.2 system for testing.

Tim Donohue 3:34 PM
Gotcha. So, this is an area we'll need to likely just add some notes to docs.... Make sure you first update your Stats to 6.x compatibility (turn into UUID based) before moving to 7.x. It sounds like you'd need to do that to go from 5.x (or below) to 7.x

Mark Wood 3:34 PM
It's been fun figuring out how to make a copy of a large-ish repo. without finding space on my office system for 300GB of documents.
Yes, I'll be working over the procedure documentation again after I finally get this testing done.

Tim Donohue 3:36 PM
Makes sense. Thanks for all the work here & for working through the procedure to nail it down. Overall, I'm really glad to hear we should (hopefully) have all the tools we need...just need to finalize the procedure & have others try it out, etc

Terry Brady 3:37 PM
If someone goes from 5x -> 7x, will they be able to upgrade stats records in 7x or will they need to perform that action in 6x?

Mark Wood 3:37 PM
I think it should work, to upgrade to UUIDs in 7x.
It's something to test.
:+1:
1

It's a good question, thanks. I've made a note.
Terry Brady 3:38 PM
As long as the records can get into solr, I think the upgrade tools will work on them.
Mark Wood 3:39 PM
This shows why we need some lurkers in the WG meetings, to stir up the folk who are head-down and racing for the finish line.
Tim Donohue 3:41 PM
@mwood: not sure I understand the full context there, but I'm always glad to have lurkers in WG meetings : slightly_smiling_face:
Mark Wood 3:41 PM
"How do we go from 5.x to 7.x, w.r.t. statistics?" Nobody asked until now.
Tim Donohue 3:42 PM
Ah, right. Well, those questions could/should be asked in DSpace 7 WG meetings too, as we do updates on this...
And honestly questions in WG can be asked via voice or via text chat too (we have had some attendees attend without a microphone)
So, I agree, lurkers welcome & encouraged to ask questions however they see fit
Mark Wood 3:43 PM
Maybe we will get a 6.4 out before 7.0, and then the doco. will make UUID upgrading part of the 6.x process.
But I've made a note to test doing it in 7.x.
:+1:
1

Anyway, I would invite interested folk to disrupt our headlong rush if there's something that's been overlooked.
Terry Brady 3:44 PM
I may have an open PR relating to the stats fix. Let me look at that.
https://github.com/DSpace/DSpace/pull/2292/files
Tim Donohue 3:45 PM
Thanks, at this time I don't know which will come first (6.4 or 7.0), to be totally honest. If I were to guess though, I'd currently guess 7.0 as there's more activity there...but 6.4 is a quicker release to achieve if someone(s) decides to run with it.
Looks like #2292 needs another review/test, and eventually a port to master
Ok, so, it sounds like this topic has wrapped up? A couple good points / next steps here...anything else to add
Mark Wood 3:48 PM
I've added a note to my "review these" list.
:+1:
1

Nothing from me just now.
Tim Donohue 3:48 PM
Ok, sounds good then. Let's move along then to other topics
@terrywbrady: Any updates/comments on the DSpace + Docker efforts? I've been reading along on #dspace-docker channel, but not sure if you wanted to summarize, or report updates?
Terry Brady 3:50 PM
I hope to pull together some interested folks to chat about this. So far only @pbecker and @kshepherd have expressed interest.
Tim Donohue 3:51 PM
I'm interested in hearing about the discussion...but not sure I'm needed for the actual meeting (assuming that's the next step)
Terry Brady 3:51 PM
The following section describes some of the potential docker deployment options and some of the parts of the system that might need to change in order to accomplish that goal. https://wiki.duraspace.org/display/DSPACE /DSpace+Docker+and+Cloud+Deployment+Goals#DSpaceDockerandCloudDeploymentGoals-DeploymentOptions
Tim Donohue 3:51 PM
Plus, it sounds like getting a meeting scheduled between the 3 of you (on different continents) will be complex enough :wink:
Terry Brady 3:53 PM
When you start talking about a DSpace 8 roadmap, I recommend spending some time evaluating this info.
Tim Donohue 3:54 PM
Definitely....and some of these topics have already come up elsewhere too...like "Migrate all CLI functionality to REST API calls". That's come up in DSpace 7 discussions already, but I don't think it'll happen entirely in v7
Mark Wood 3:54 PM
Is there a public list of "things we're putting off until DSpace 8"?
Tim Donohue 3:55 PM
Not a fully maintained one. But there is this page. Feel free to add to it: https://wiki.duraspace.org/display /DSPACE/DSpace+Release+8.0+Status
Mark Wood 3:55 PM
Ah, of course, thanks.
Tim Donohue 3:55 PM
Only one item is there right now though
Terry Brady 3:56 PM

Could this be added to that list as "Refactor for Docker/Container Deployment".
Tim Donohue 3:57 PM
Possibly, though it might be worth breaking that down some. There's a lot on your list here.... That said,
putting it on there as a massive list is better than nothing for now.
Terry Brady 3:57 PM
I would also like to propose the migration of the REST reporting tools since they will become deprecated.
Mark Wood 3:58 PM
Hit "e" and add it.
Kim Shepherd:musical_note: 3:58 PM
Sorry folks, I slept in this morning!
Tim Donohue 3:59 PM
Yes, add the REST reporting tools to that list please
Terry Brady 3:59 PM
Is it OK to add to this page? I was not certain if you curated this @tdonohue
Tim Donohue 3:59 PM
And the Docker stuff should be added for more discussion. I was just thinking some of the Docker changes
you've listed here are non-obvious (to me). There's a few that I'd say immediately "yes, let's do that".
Others, we'd need to discuss more
Mark Wood 3:59 PM
Welcome @kshepherd.
Tim Donohue 4:00 PM
No, this list is not curated at this time @terrywbrady. Please add anything to it. As DSpace 8 discussions
ramp up, we'll curate & comb through it more.
Mark Wood 4:00 PM
I reckon that that list will become a source of discussion and may have some items moved to "DSpace release 9.0
status".
:+1:
1

Tim Donohue 4:01 PM
(And I don't think DSpace 8 discussions will ramp up anytime soon, so adding things while we are thinking about
them is a good idea...otherwise we'll forget them once DSpace 8 discussions start)
Mark Wood 4:01 PM
So, we shouldn't be too shy of making reasonable additions.
Tim Donohue 4:02 PM
Correct. If it's something that has come up in a discussion between two Committers or two trusted developers...
add it in. Crazy brainstorms run by someone else first for a sanity check...then possibly add it :wink:
(edited)
Terry Brady 4:03 PM
Great. 2 items added.
Tim Donohue 4:04 PM
Ok, I'm realizing we are over time here. So, it's time to wrap up this meeting. As noted, unfortunately this
will be our last general DevMtg for a while...but I'd encourage attendance (or reading notes) of DSpace 7 WG
meetings & ad hoc discussions (possibly meetings) via Slack.
In the meantime, also as noted, don't be shy about asking me questions. Glad to help or clarify or give ad hoc
updates on #dev or similar.
Thanks all! I hope you have a good rest of the week, and hope to catch you in tomorrow's DSpace 7 WG meeting
or on Slack or GitHub
Mark Wood 4:05 PM
Thanks, 'bye.
Kim Shepherd:musical_note: 4:05 PM
Thanks