DevMtg 2018-11-07

Developers Meeting on Weds, November 7, 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

Agenda

Quick Reminders

Friendly reminders of upcoming meetings, discussions etc

- DSpace 7 Working Group (2016-2023): Next meeting on Thursday, Nov 8 at 15:00 UTC.
- DSpace 7 Entities Working Group (2018-19): Meeting notes from yesterday at 2018-11-06 DSpace 7 Entities WG Meeting
- DSpace Developer Show and Tell Meetings: On hold until interesting topics arise.

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 (2016-2023))
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 7.x dependency update PRs
   a. https://github.com/DSpace/DSpace/pull/2244 (Apache Commons Configuration v1 v2 and Commons Lang v3)
   b. https://github.com/DSpace/DSpace/pull/2241 (Log4j 1.2 2.x) - Solr upgrade is seemingly dependent on Log4j2
   c. https://github.com/DSpace/DSpace/pull/2058 (Solr v7).
4. DSpace 6x: Merge Legacy and Current Statistics: https://github.com/DSpace/DSpace/pull/1810
5. Brainstorms / ideas (*Any quick updates to report?*)
   a. (*On Hold, pending Steering/Leadership approval*) Follow-up on "DSpace Top GitHub Contributors" site (Tim Donohue): https://tdonohue.github.io/top-contributors/
   b. Follow-up on Curation Task Reporting (PR 2180)
   c. Bulk Operations Support Enhancements (from Mark H. Wood)
   d. Curation System Needs (from Mark H. Wood)
      i. PR 2181 implements per-run task parameters. Ready for review.
      ii. PR 2180 improves reporting. Ready for review.
6. 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

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 [2:00 PM]
@here: It's DSpace DevMtg time. Our agenda for today is at: https://wiki.duraspace.org/display/DSPACE
/DevMtg+2018-11-07

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

Mark Wood [2:01 PM]
is here.

Terry Brady [2:01 PM]
here

James Creel [2:01 PM]
hi!
Totally would have missed this were it not for the "here" callout. Cursed Daylight Saving Time.

Tim Donohue [2:02 PM]
Hi all. Yes, daylight savings ending means this meeting is an hour earlier (for anyone leaving DST)

Mark Wood [2:02 PM]
Duraspace has a public Google Calendar with meetings on it, if that's useful.

Tim Donohue [2:03 PM]
Yep, the DuraSpace public events calendar is listed here: https://wiki.duraspace.org/display/DSPACE
/Developer+Meetings (Look at the bullet *under* the bolded one)
In any case, let's go ahead and get started today
In the "what's going on in other meetings front"....there was an Entities meeting yesterday. It's video &
notes are at: https://wiki.duraspace.org/display/DSPACE/2018-11-06+DSpace+7+Entities+WG+Meeting
Currently, the team is working on reviewing & providing feedback on the prototype code. If anyone else would
like to chip in, this is being done in GitHub (in two large PRs). We always could use more "eyes" and reviewers
The plan is to merge the early code into a `configurable_entities` shared branch...clean it up, finalize & then
submit a final PR to `master` (for both the UI and REST API)
So, keep an eye out for much more in coming weeks/month
The DSpace 7 team has their next meeting tomorrow at 15UTC (10am EST).
I'll also note, there was a DSpace Leadership Meeting today. We talked quite a bit on whether to release
anything (e.g. an "early preview") in 2018 regarding DSpace 7. The reality is that we don't have all features
done, so releasing a preview release would mean releasing something not "feature complete"
That group is less and less excited about a non-"feature complete" release...so, it's unlikely at this time
that we'll have anything in 2018 (other than updated demo sites, of course). But, a subgroup is meeting to
discuss DSpace 7 release plans/scheduling in more detail next week.
So, that's just a brief update there...not anything else official to note at this time.
I think that's it for my giant wall of text/updates. Will pause briefly for any questions/comments though

Mark Wood [2:11 PM]
I think I agree: if it's not complete, we don't want people to remember it as if it is (which *will* happen).

Tim Donohue [2:12 PM]
Oh, and by the way, Leadership Group meeting notes are now all public. They've already been posted here:
https://wiki.duraspace.org/display/DSPACE/2018-11-07+DSpace+Leadership+Group+Meeting

So, feel free to give those a read, if you want.

Mark Wood [2:12 PM]
Thanks for keeping us updated.

Tim Donohue [2:13 PM]
On the DSpace 6.4 side of things.. Again, while I'm sure we'll eventually have a 6.4, I'm also currently
assuming it's not happening until after 7.0 (unless something changes/becomes really high priority). So, no
real updates there.

Terry Brady [2:14 PM]
I presumed our push on the stats issue meant that there might be a need for a 6.4 release
Do we have any other significant features waiting for a release?

Tim Donohue [2:15 PM]
@terrywbrady: I'm glad we pushed that, but to be honest, I'm not going to have mindshare for a 6.4 anytime
soon, even though I think that's important once 6.4 comes around

Mark Wood [2:16 PM]
There is a nasty knot of interdependent dependency upgrades and linked fixes that is pulling e.g. the
statistics schema fixes forward.

Tim Donohue [2:16 PM]
We do have a variety of PRs waiting for reviews/testing and flagged as "possible for 6.4". But, I'm fully concentrated on 7.0 and haven't had time to even really look or review them
PRs current flagged as possible for 6.4: https://github.com/DSpace/DSpace/milestone/27 (There likely are even more out there too)
At this moment, none of these PRs seem so "high priority" as to pull me from 7.0. If one did rise to that level, I'd have to talk to Steering/Leadership about how we'd want to handle it. Currently, they've made it clear that 7.0 is the highest priority

Terry Brady [2:18 PM]
That list is helpful to see. If something comes up, this will be a good reference point.

Mark Wood [2:18 PM]
That doesn't meant that interested others can't give some cycles to these.

Terry Brady [2:19 PM]
Btw, I am going to start work on an upgrade of our repo to 6x. I paused it a couple times in the past due to 6x bugs. I think we will have better luck this time.

Tim Donohue [2:19 PM]
Yep. And, I'm continuing to tag all PRs that come in....so, that list is being added to / enhanced. I just haven't had time to really look at these PRs
And, I agree with @mwood. If anyone *here* is affected by these PRs, testing/reviews right now are more than welcome. No need to wait until later, especially since some of these might also be worth porting to 7.x (in full or part)
Any other questions/comments on 6.x? Essentially, please feel free to review/test PRs (knowing that that effort will bear fruit in the future). But, also be aware that 6.4 may not be out until sometime in first half of 2019
Ok, moving along then to topic #3... helping to update dependencies for 7.0
There's 3 PRs listed here that give major updates to some key dependencies. I'd really like to get the first two in soon (so they can be further tested as we continue development)
Namely...
Apache Commons Config upgrade (already at +1 votes): https://github.com/DSpace/DSpace/pull/2244
Log4j v2 upgrade (also at +1 votes): https://github.com/DSpace/DSpace/pull/2241
The latter is of the most importance (as log4j v1 is EOL & new dependencies don't even necessary work well with it).
But the former also keeps us "up to date", and helps our Spring integration
Is there anyone here who could help @mwood and I move these two upgrades along? We basically need a second review on each

Mark Wood [2:27 PM]
Log4j v2 has some nice things for developers: (1) parameterized message texts like SLF4J; (2) pass a lambda to provide parameter value, *better than* SLF4J; (3) 0-argument getLogger() figures out the category without help.

Terry Brady [2:27 PM]
I put these in my review queue

Tim Donohue [2:28 PM]
I'll also note that our Solr upgrade (the third upgrade in this list) is waiting on log4j v2. New Solr doesn't seem to "like" log4j v1 (which makes sense, as it's been EOL for some time).
So, if you want us to also get off the EOL version of Solr, then we first must get off the EOL version of log4j :wink:

Terry Brady [2:29 PM]
I added that one too

Mark Wood [2:29 PM]
Note: that PR only upgrades the Solr *client*. We still need to move to a recent server.

Tim Donohue [2:29 PM]
@terrywbrady: thanks. Mark & I have reviewed each other's PRs. So, hopefully it'll be easy/quick to give a second +1 on each
Yes, the Solr PR in question currently only upgrades the client (i.e. step #1): https://github.com/DSpace/DSpace/pull/2058
It will need to be followed by a complimentary PR to also upgrade the Solr server. But, we're trying to take this in two steps to avoid a *massive* PR

Andrea Bollini [2:31 PM]
I recommend to keep SOLR client & server aligned. Update the server side could be not trivial, so before to merge the client we should be sure to be able to update also the server side

one thing to note for instance is that recent solr is only provided as standalone service (jetty embedded)

Tim Donohue [2:32 PM]
@bollini: They will be aligned before 7.x. The Server upgrade though is more complex as we need to provide
scripts to upgrade existing data, etc.

Mark Wood [2:32 PM]
It shouldn't be that hard. We install absolutely stock Solr plus two tiny classes that affect nothing in it.
IIRC one of those classes is gone now, and the other is a mere convenience.

Tim Donohue [2:33 PM]
And, yes, the Solr server is now installed standalone. We will no longer be able to "bundle" Solr within
DSpace. That concept no longer is recommended for Solr (edited)
But, we *have* to get off of an EOL version of Solr. There's no other path forward here, especially since we
cannot even resolve security issues being reported on Solr (as we're on an unsupported version). So, the
reality here is that we are simply moving DSpace in the "required" direction to be compatible with modern Solr

Mark Wood [2:35 PM]
To get the server side ready, I will pitch in where I can.

Tim Donohue [2:36 PM]
In any case, the earlier/quicker we can merge the log4j v2 upgrade, the quicker we can upgrade Solr. The
latest Solr really isn't working well (dependency conflicts/mismatches all over) with log4j v1.

Andrea Bollini [2:36 PM]
ok, thanks to clarify. It will be nice to move to a recent solr version

Terry Brady [2:37 PM]
Once the server is separated out, we will need to update our docker stuff to reflect the additional service.

Mark Wood [2:38 PM]
Good point. If there is a ticket for the server side, it would be good to note that there.

Tim Donohue [2:38 PM]
@terrywbrady: yes, makes sense. We'll also need to update our DSpace Installation process anyhow (as Solr will
be a dependency to install separate). But, that's a good point
Ok, so I don't have anything else to note here. We could just really use more eyes on the open PRs, as I think
all of these are important for DSpace 7. The quicker we merge them, the better. That way I also get these off
my plate (as I'm trying to ensure DSpace 7 doesn't get released with *any* EOL dependencies or known security
vulnerabilities in dependencies)
Are there any other important dependency or backend updates we want to talk about for DSpace 7? (Obviously
keeping in mind work on REST API and Angular will be talked about in tomorrow's mtg)
Ok, not hearing any
Next on the agenda was to get some final approval for Solr Statistics Schema update PR here: https://github.com
/DSpace/DSpace/pull/1810
And, actually that's already at +2 :slightly_smiling_face:
So, I suspect this can just be merged... The update script code looks good to me. And, I know @mwood and
@terrywbrady have been testing it on live data

Terry Brady [2:45 PM]
And a 6x port is ready to go: https://github.com/DSpace/DSpace/pull/2256
I suppose the title of both should be updated to remove the schema change reference

Tim Donohue [2:46 PM]
@terrywbrady: I'm assuming the two PRs have been kept "in sync"? If so, we likely can just do a quick sanity
check of the 6.x port, once the `master` version is merged
Also, as I noted in the 1810 PR, we will need to create some documentation here for 6.4 and 7.0

Terry Brady [2:47 PM]
Yes. The 6x was created after the latest change for 7x.
I will create documentation after the merge is done.

Tim Donohue [2:49 PM]
Sounds good. Assuming no objections, I can go ahead and merge #1810, and then do a "sanity check" comparison
of #2256 (and merge it, assuming everything looks the same). I can do that after this meeting
We'll keep the JIRA ticket open until the docs are done
Is there anything else to discuss on this topic?
not hearing anything...moving along
Ok, so we have <10mins left. The last topic on our agenda is the ongoing brainstorms/ideas that we've just
been carrying over from meeting to meeting (in case of updates)
Rather than going through all these, in the essence of time, are there any updates to add? Anything requiring

more eyes/reviews?

Terry Brady [2:52 PM]
what is up with your contributor report

Mark Wood [2:52 PM]
Nothing to add.

Tim Donohue [2:53 PM]
@terrywbrady: I've had zero time to even move it forward to be honest. I'm swamped right now...and Steering
/Leadership is swamped with DSpace 7 decisions. So, while I'd like to see it "official", I'm not sure when
it'll happen
It still "exists", but I also haven't found time to update it with Oct data yet: https://tdonohue.github.io/top-
contributors/
tdonohue.github.io
DSpace's Top GitHub Contributors
Ranking DSpace GitHub contributors, since 2018

Terry Brady [2:54 PM]
Since you have a solution, it is a shame that it requires approval to use. I think it is a nice way to
recognize contributions from folks.
It could be a meaningful form of recognition for new contributors to the project.

Tim Donohue [2:55 PM]
It doesn't really *require* approval, but I'd rather it get some sort of approval/look prior to moving it under
the "DSpace" name.
The only reason I want that...is that Fedora developers specifically didn't like the idea. I don't know all
the ins & outs of "why". But, I'd hate to release it and have folks say: "why are we doing this?"

Terry Brady [2:56 PM]
For what my approval is worth... I think it is a nice solution. :slightly_smiling_face:

Tim Donohue [2:56 PM]
There *is* an upcoming Outreach subgroup of Leadership & Steering (forming now) that I plan to bring this to.
However, they haven't formed yet

Terry Brady [2:57 PM]
That sounds like a good audience.

Tim Donohue [2:57 PM]
In the meantime, it exists & it's public. So, I'll try to keep it up to date (will find a few minutes to get
the Oct stats in soon): https://tdonohue.github.io/top-contributors/

Terry Brady [2:58 PM]
I need to join another meeting. Have a good week!

Mark Wood [2:59 PM]
Should dspace-api enable deletion of copy request records? Was this omission intentional? IIRC @bollini was
connected with this code. Do you recall?

Tim Donohue [2:59 PM]
Bye. Ok, as we only have a few minutes here, I'm assuming we likely don't have time for any other topics. But
is there anything to add to next week's agenda?
@mwood: i don't think that omission was intentional. So, I'm in favor of adding it in (especially since it
could be personal info that requires deletion per GDPR or country-specific regulations)
As we're at the top of the hour, and discussion has slowed, I say we close out the meeting. If there are
topics to add for next week, pass them my way and I'll get them on our agenda.
Thanks all & have a good rest of the week

Mark Wood [3:02 PM]
Thanks, all.