

2019-07-02 DSpace 7 Entities WG Meeting

Date & Location

02 Jul 2019 at 15:00 UTC (11:00am EST)

Location:

Join from PC, Mac, Linux, iOS or Android: <https://duraspace.zoom.us/my/dspace> (Meeting ID: 502 527 3040)

Additional connection info available on [DSpace Meeting Room](#) page.

Attendees

Actual attendee list will be updated after meeting.

- [Tim Donohue](#) - DuraSpace (On holiday July 3-7. Returning on July 8)
- [Pascal-Nicolas Becker](#) - The Library Code
- [Alexander Sulfrian](#) - Freie Universität Berlin
- [Lieven Droogmans](#) - Atmire
- [Ben Bosman](#) - Atmire
- [Paulo Lopes](#) - FCT|FCCN
- [Mark H. Wood](#) - IUPUI
- [Paulo Graça](#) - FCT|FCCN
- [Jose Carvalho](#) - University of Minho
- [Heather Greer Klein](#) - DuraSpace
- [Dimitris Pierrakos](#) - ARC/OpenAIRE
- [Oliver Goldschmidt](#) (On holiday weeks of June 24 & July 1)

Agenda

| # | Time | Item | Information | Who |
|---|--------|---|--|-----|
| 1 | 15mins | Metadata strategy for Entities | <p>Some interesting discussions on https://github.com/DSpace/DSpace/pull/2443. Two (seemingly different) strategies towards Entities metadata seem to have arisen:</p> <ol style="list-style-type: none">1. Moving as many Entities metadata concepts/fields as possible to schema.org fields2. Or, only using schema.org fields when a concept/field doesn't exist in Dublin Core / DCTerms <p>Some questions from the PR:</p> <ul style="list-style-type: none">• Do Person Entity names belong in "dc.title"? Or is that not descriptive enough, and we should look to use schema.org's Thing schema ("thing.name" field)• Do Person Entity ORCIDs belong in "dc.identifier"? Or again, should we look towards schema.org and use "thing.identifier" or "thing.identifier.orcid" | Tim |
| 2 | 20mins | Revisiting Priorities Listing (see full list below) | <p>We've started #1 and #2. Priorities #3 and #4 are worth digging into a bit more (background reading at https://docs.google.com/document/d/1X0XsppZY0tPtBmq7yXwmu7FbMAfLxxOCONbw0_r17jY/edit#heading=h.wjaqg235p53r):</p> <ul style="list-style-type: none">• Additional data for relations (essentially "metadata" or labels on relations) - Related to many other features / use cases.• Author name variants - Not currently implemented <p>As a reminder, these priorities came from the "Next Steps" listing of the Configurable Entities Documentation: https://docs.google.com/document/d/1X0XsppZY0tPtBmq7yXwmu7FbMAfLxxOCONbw0_r17jY/edit#heading=h.hsdan1o4sj6</p> | All |
| 3 | 10mins | Other Topics? | <p>Clarification on labels of relationships: https://www.dropbox.com/s/m28kgyarazolf9/entities%20labels%20direction.pdf?dl=1</p> <p>Any updates on Submission Integration work to share yet?</p> <p>Any updates on OpenAIRE v4 work to share yet?</p> | All |
| 4 | 15mins | Wrap-up and Assigning tasks | <ul style="list-style-type: none">• Assign tasks / PRs to review (as needed)• Next Meeting?<ul style="list-style-type: none">◦ Tim has a conflict on Tues, July 9. Should we cancel and meet again on Tues, July 16? | Tim |

Current Work

Tickets to Resolve

- All JIRA tickets tagged with "Configurable Entities": <https://jira.duraspace.org/issues/?jql=project%20%3D%20DS%20AND%20component%20%3D%20%22Configurable%20Entities%22>
 - Critical Tickets: <https://jira.duraspace.org/browse/DS-4244?jql=project%20%3D%20DS%20AND%20priority%20%3D%20Critical%20AND%20component%20%3D%20%22Configurable%20Entities%22>

- Major Tickets: <https://jira.duraspace.org/browse/DS-4241?jql=project%20%3D%20DS%20AND%20priority%20%3D%20Major%20AND%20component%20%3D%20%22Configurable%20Entities%22>

PRs Needing Review

- (REST Contract) (Entities) Support for defining relationship lookups in the submission forms <https://github.com/DSpace/Rest7Contract/pull/64> (Alexander Sulfrian - feedback provided, Paulo Graça - feedback provided)
- (REST) (Entities) DS-4223 Metadata Schemas for configurable entities <https://github.com/DSpace/DSpace/pull/2443> (Paulo Graça ✓, ✓ Mark H. Wood)
- (REST) (Entities) DS-4244 Add configurable entities unit tests <https://github.com/DSpace/DSpace/pull/2446> (Tim Donohue - requested inline comments)
- (REST) (Entities) Relationship search by label: <https://github.com/DSpace/DSpace/pull/2450> (Tim Donohue - ✓ but requested inline comments, Alexander Sulfrian - blocked URL but requested a missing javadoc, Paulo Graça ✓)
- (REST) (Entities) [DS-4084] - Relationships should link to Items via REST API: <https://github.com/DSpace/DSpace/pull/2456> (Alexander Sulfrian blocked URL, Paulo Graça ✓)
- (NEW) (REST) (Entities) ITs for deleting relationships <https://github.com/DSpace/DSpace/pull/2457> (Alexander Sulfrian blocked URL, ...)
- (Angular) (Entities) Deleting relationships: <https://github.com/DSpace/dspace-angular/pull/402> (Paulo Graça - changes requested, Tim Donohue)
- (Angular) (Entities) DS-4223 Metadata Schemas for configurable entities <https://github.com/DSpace/dspace-angular/pull/420> (Paulo Graça ✓, Tim Donohue)
- (NEW) (Angular) (Entities) One-sided relationship filtering and refactoring <https://github.com/DSpace/dspace-angular/pull/430> (Ben Bosman, ...)

PRs Merged this week!

1. (first PR merged goes here)

Priorities

This list was voted on in the meeting on [May 23, 2019](#) (just before OR2019). The prioritization below may change, but it gives a high level overview of what still needs to be done.

1. **(Lieven, Ben, Tim, Fernando, Jose, Mark, Oliver, Paulo)** Submission integration (creating Entities & relations using the Item submission process) - Mockups already created by Paulo previously.
2. **(Lieven, Ben, Tim, Jose, Oliver, Paulo)** Which metadata fields should be used for each Entity type. (DS-4223).
3. **(Lieven, Ben, Tim, Mark)** Additional data for relations (essentially "metadata" or labels on relations) - Related to many other features / use cases.
4. **(Oliver, Paulo)** Author name variants - Not currently implemented
5. **(Jose)** Configuration of batch import (via CSV) for Entities - Already a CSV import available, but can only link entities in CSV to existing entities (in the system). Need to decide how to represent relations in CSV.
6. **(Mark)** Permissions on Relations (who has permissions to add/modify/delete relations) - Currently, if you have Edit permissions on the Entity, then you can edit/delete any relationships to/from that Entity.
7. **(Fernando)** Deleting objects with Relations (How or should deletion propagate between closely related objects, e.g. delete entire Journal) - Currently, deleting a relation just decouples the two Entities. E.g. if you delete a Person entity, that Person may no longer be listed on any Publications it is linked to (may want to copy info over after deletion).
 - a. Relates to GDPR
8. **(Alexander)** AIP Backup & Restore (of Entities)*
9. Dynamic display of Relations - determine automatically how a list of entities displayed on an Item page (list vs search). Currently hardcoded based on entity type (in item page template). Want to make it configurable/dynamic.
10. SWORD integration (submission of Entities via SWORD) - Uses same format as AIP. Once AIP is implemented, SWORD should be easy.
11. OpenAIRE v4 implementation using Entities* - Brought up in Steering. Possibly just an OAI-PMH configuration which maps Entity metadata fields to OpenAIRE v4
12. ORCID integration with Entities (for Person Entities).
13. Best Practices around Entities in Collections. We've suggested in the Preview Release to structure Collections based on Entity Type (Person Collection, Projects Collection, etc). We should better document and formalize these best practices. Can we hide these Collections which only serve to store Entity Type.

Notes

- Metadata strategies for Entities
 - Decision to move forward with PR as-is: <https://github.com/DSpace/DSpace/pull/2443>.
 - Tim will create a JIRA ticket to describe the two metadata options (rely more on DC, or move more towards schema.org). Will ask DCAT (and others) to review and provide feedback.
 - List pros/cons to each approach. Note adding schema.org support makes entities metadata easier, but may add complexities to Browse/Search/Submit configurations (we could properly configure out-of-box) and other features that relate to metadata (e.g. Batch metadata editing, OAI-PMH, etc)
- Discussion of Additional data per relations.
 - Two approaches suggested in the Google Doc: https://docs.google.com/document/d/1X0XsppZY0tPtbmq7yXwmu7FbMAflxxOCONbw0_r17jY/edit#heading=h.lknzuix8dggk
 - Simple approach, extending relations to mostly support Author name variants
 - Complex approach, come up with a solution to also support temporal relationships (e.g. author with changing affiliations, Journals moving between publishers, Journals changing names...over time)
 - Atmire recommends simple approach, as it'll be much quicker to get to for 7.x. But, this means no support for Temporal relationships (other than possibly recommending versioning Entities that have changed relationships over time, e.g. version 1 of a Journal may have one name, while version 2 has a different name).
 - Agreement among team
 - Decision:
 - Version 7 will concentrate on Author variants (simple label on relations)

- Temporal relations will be left for Version 8. More use cases may help us design this better.
 - May even want to consider whether these are better met by somehow storing relationships as triples (as RDF has solved some of these issues via usage of triples)
- OpenAIRE v4 updates (from Jose): https://docs.google.com/presentation/d/1wRwkP6z-Eep7yaexf32B_FSAqUwtPhQkqujnpT55vql/edit
- **Next meeting July 9.** Paulo Graca will lead this meeting, as Tim has an all day meeting conflict.

Action Items

Any assigned actions will appear here, along with details of who they are assigned to.