2019-07-16 DSpace 7 Entities WG Meeting

Date & Location

16 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)***

Attendees

Actual attendee list will be updated after meeting.

- Tim Donohue - DuraSpace
- 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 - Hamburg University of Technology (TUHH) - cannot make it on July 16

Agenda

#	Time	Item	Information	Who
1	5mins	Agenda	Any additional topics to today's agenda	All
2	10mins	Revisiting Contract for Relationship lookups	Based discussion last week on https://github.com/DSpace/Rest7Contract/pull/64, it should be determined whether any of the suggestions should be changed. Please check: https://github.com/DSpace/Rest7Contract/pull/65.	All
3	20mins	New Task #14	Bring examples to support the need of having the #14 task - https://docs.google.com/presentation/d/1aD_JbbmfHOl_t-IoouqY0xaPNUyd0JtWTuLJJjeygp0	RCAAP and All
4	20mins	Overview of Tasks #5 and #7	Reintroduce discussion of these two tasks, so that team can begin to think more about what implementation may look like: • #5 : Configuration of batch import (via CSV) for Entities • #7 : Deleting objects with Relations (How or should deletion propagate between closely related objects, e.g. delete entire Journal)	All
5	5mins	Wrap-up and Assigning tasks	• Assign tasks / PRs to review (as needed) • Next Meeting is Tues, July 23.	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

1. (REST Contract) (Entities) Support for defining relationship lookups in the submission forms https://github.com/DSpace/Rest7Contract/pull/64 (blocked URLAlexander Sulfrian, Paulo Graça - feedback provided)
2. **(UPDATED)** (Angular) (Entities) Deleting relationships: https://github.com/DSpace/dspace-angular/pull/402 (Paulo Graça , Tim Donohue)
3. (Angular) (Entities) One-sided relationship filtering and refactoring https://github.com/DSpace/dspace-angular/pull/430 (blocked URLBen Bosman, Paulo Graça)

PRs Merged this week!

1. ✅ (REST) (Entities) DS-4244 Add configurable entities unit tests https://github.com/DSpace/DSpace/pull/2446
2. ✅ (REST) (Entities) ITs for deleting relationships https://github.com/DSpace/DSpace/pull/2457
3. ✅ (REST) (Entities) DS-4223 Metadata Schemas for configurable entities https://github.com/DSpace/DSpace/pull/2443
4. ✅ (Angular) (Entities) DS-4223 Metadata Schemas for configurable entities https://github.com/DSpace/dspace-angular/pull/420

Task List

This list was roughly prioritized 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. *NOTE: Keep in mind, just because an item is listed here does NOT guarantee it will be completed for DSpace 7. Some of these tasks may need to be delayed for a future release.*

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.
14. **(NEW)** The ability of pick the proper affiliation of a Person for a specific context. DSpace should address this use case to allow the user to describe something like in this document http://repositorium.sdum.uminho.pt/bitstream/1822/46268/1/1-s2.0-S1877050917302788.pdf regarding the authors and affiliations. You have different persons, each can belong to an institution at the time of that publication. The affiliation shouldn't be changed afterwards. And the user should be able to pick the proper one if an Author has more than one.

Notes

- Revisiting Contract for Relationship lookups (PR: https://github.com/DSpace/Rest7Contract/pull/64)
    - Decision: no strong opinion here on one being "more correct" than the other
    - We will move forward with PR#64 as-is (except Ben will remove the "relationship: true" flag noted as not necessary)
    - As we move to implement REST API and/or Angular UI interface, if we find this approach is complicating implementation then we will look towards alternative PR#65: https://github.com/DSpace/Rest7Contract/pull/65.
- Task #14 above
    - All agree this is a real need & use case. However, it's not currently met by the existing implementation
    - We can only provide a one-to-one link. So publication can be linked to an Org and a Person. Or we can provide a date range for a Person to Org relationship. However, we cannot yet say "This Person was affiliated with this Org when they wrote this Publication"
    - While this is a real use case, it might be difficult to achieve in DSpace 7 without a much larger change to implementation (which we don't have time for).
    - We will keep this task on the list, but will note that it's more likely to be achieved in DSpace 8 (i.e. post DSpace 7)
- Overview of Task #5 (didn't get to Task #7 today)
    - Atmire proposed a way to define "temporary" identifiers in CSV spreadsheets, in order to allow one CSV to create two new Entities & link them together (i.e. the UUID for those new entities will be unknown in the spreadsheet)
    - Tim noted this overlaps some with the idea of the "+" placeholder for creating new Items. Should we use "+" as part of that temporary identifier? (e.g. "+1", "+2", etc.)?
        - Batch Metadata Editing#AddingMetadata-OnlyItems
    - It's important that these identifiers only exist in the CSV spreadsheet. They should NOT be stored in metadata in system as they are not guaranteed unique.

Action Items

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