

DCAT Meeting Notes June 3, 2014

Date

- Tuesday, June 3, 2014

Time

- 10:00am Eastern/14:00 UTC

Dial-in

We will use the international conference call dial-in. Please follow directions below.

- U.S.A/Canada toll free: 866-740-1260, participant code: 2257295
- International toll free: <http://www.readytalk.com/intl>
 - Use the above link and input 2257295 and the country you are calling from to get your country's toll-free dial in #
 - Once on the call, enter participant code 2257295

Attendees

- [Bram Luyten \(@mire\)](#) - @mire
- [Elin Stangeland](#) - University of Oslo
- [Sarah Potvin](#) - Texas A&M University
- [Valorie Hollister](#) - DuraSpace
- Sue Kunda

Discussion Items

| Time | Item | Who |
|--------|--------------------------------|-----|
| 40 min | 1) Metadata Panel at OR14 | All |
| 20 min | 2) DCAT Charge | All |

Discussion Notes

Panel Structure

1) Introduction (5 min): VAL

- Goal of discussion: discover what problems people care about - what needs need to be addressed by a DSpace metadata improvement project
- Summary of the metadata improvement project history
 - project spawned from a discussion during the committers mtg at OR10 in Austin at OR11 – committers requested that DCAT help to align current DSpace Dublin Core with current std
 - DCAT worked on the specific project over the last few years, outlining a detailed proposal for the community, gathering feedback, thinking through some of the implementation issues
 - seems like a fairly arbitrary/narrow task – in looking at how DC stds have evolved in recent years it is clear that there are many more complex metadata needs in the community
 - part of the goal for the panel discussion is to begin to broaden discussion w/community
 - even if DSpace was updated with the current version of DC, metadata stds in DSpace would not perform the way the community needs it to - many local workarounds, customizations and some real practical challenges
 - as we look at the future roadmap of DSpace, want to discuss reqmts for metadata
 - most basically and importantly is the challenge around flexibility - must have flexibility around:
 - 1) ingest - if material coming into DSpace is in different std, you need to map so fields import correctly
 - 2) exports / harvesting - metadata needs to be flexible enough to enable crosswalks for exports, harvesting and other purposes
 - if funders have asked for clean ingests / harvesting - more plug and play / easy to use
 - getting data in and out is very important for CRIS - ORCHID - ecosystem is much more complex - need to coordinate metadata on national level
 - 3) local needs - needs to be easy to manage your metadata and content locally (i.e., not req'd to go into XML to make changes)

- identify current challenges
 - we don't just need to standardize to current Dublin Core, need to add new metadata standards
 - clean ingests / or exporting
 - more easy to use - plug and play additions/changes
 - creating separate namespace should be easy - for CRIS, etc. friendly metadata
 - no hard coded fields (like date) - originally DSpace was built to be metadata agnostic
 - creating custom fields - requires you to login to XML in order to do so or sometimes a Java programmer to do and rebuild

2) Brief Panelist Self-Intro (5 min or less): ELIN, BRAM, PEDRO, SARAH

- each panelist introduces themselves and provides a brief summary of where they are from and what their interests are in metadata

3) Panel Discussion

- In advance of panel discussion, panelists will develop a list of the key issues/challenges that DSpace needs to address to meet current (and maybe future) reqmts/needs
- In advance of panel discussion, each panelist will select a specific issue/challenge that they would like to provide a brief specific example about to help spark discussion
- Moderator will limit time for each discussion so that each panelist gets an opportunity to speak
- List of challenges - what are the biggest metadata issues that need to be addressed on DSpace roadmap? what are the things you 'hack' or have work arounds (addressing an unmet need) for locally? what are the implications of workarounds? what should be taken in consideration with future DSpace - so that you don't lose your local customization? what mistakes have you made with DSpace metadata? what would you prefer to do differently? what/how far have you stretched the metadata in DSpace?
 - - adapting metadata to meet external reqmts (OpenAIRE, etc.)
 - exporting metadata to diff schemas DC to METS
 - linked data

Val to create a Google Doc for the panel to use to brainstorm: https://docs.google.com/a/fedora-commons.org/document/d/1x1xt5dq3NiYPVcdYPPKj57caz_YSH4bgJk7IBFCBU/edit

Notes on Metadata Panel from April 30 discussion

[copy of proposal](#)

1) A few questions that could fuel our agenda:

- have about 1 hour for the panel: https://www.conftool.com/or2014/index.php?page=browseSessions&form_date=2014-06-13
- need 4 or 5 panelists - shot for 5
 - Sarah Potvin
 - Elin
 - Bram
 - someone representing large harvesting initiative: OpenAir, SHARE, someone from Fedora linked data or EPrints, RIOXX, Europeana, Metadata For All (Bram)
 - Elin to get list of attendees to see if there is anyone who could be invited to participate
 - Stephanie will check to see if there is anyone from Dublin Core community Jykha Jakula
 - session moderator: Val

2) Structure of the panel

- 10-15 min: introduction highlight why we are having panel, what we want to get out of it
 - validate ideas/plans - not missing a reqmt from the community or reqmt by external standards
- 5 min or less each: each panelist present backgrounds and views
- Discussion: a few topic prompts to guide discussion
 - how does this plan sync up or create conflicts with your current work?
 - linked data - what benefits are there in linked data?
 - how to serve common use cases - as everyone repositories evolve? how have people hacked DSpace to get it to do what they want - make sure current plan addresses some of those needs
 - linking articles and journals in DSpace, ability to define relationships
 - CRIS hierarchical metadata needs to be flattened in order to bring into DSpace
 - what are the priorities and more imminent problems to tackle

3) Meet at OR beforehand to talk through panel

4) Other ideas from Bram

I think it would be nice if we may be able to present a few opposing opinions on the three topics, and see how the audience reacts. e.g.

A new look at supporting and safeguarding standards

Someone who's highly in favor of doing a full migration to dcterms vs someone who presents the idea of a mixed approach where we keep some free text fields like dc.title.

Or someone who's in favor of adding custom/institution specific fields vs someone in favor of reusing external schema's like RIOXX or UKETD

Metadata for all: Maybe there's someone who wants to make a case for treating other dspace objects (coll, comms, bitstreams) differently when it comes to metadata?

Future proofing

Opposing/different opinions on what the priorities should be?

What do you think? would this make sense? How can we ensure that we "dumb it down" enough so that it appeals and affects the audience the most.

Action Items

- Meet at OR beforehand to talk through panel
-