DCAT Meeting March 2016

- short link to this page: http://bit.ly/dcat-march

Date & Time

- March 8th **16:00** UTC/GMT - **11:00** EST

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

Agenda

Let's talk about OAI-PMH

For many institutions, the OAI-PMH interface is a key feature of DSpace, sometimes even a central driver for establishing an institutional repository.

This interface enables your repository to provide metadata to Initiatives such as OpenAIRE, RIOXX, SHARE, DART Europe, BASE-SEARCH and OAISTER (now part of WorldCat)

Even though the protocol has been stable for many years, the DSpace implementation has evolved. The most important milestone was a complete rewrite of the DSpace OAI code back in DSpace 3, where also a web UI was added so administrators can easily preview results exposed through this interface: http://demo.dspace.org/oai/request?verb=Identify

Despite the popularity and the massive uptake of this feature, a number of issues remain to be addressed by the DSpace community.

Friedrich Summann will join us in this call, representing the perspective of BASE-SEARCH, a large scale OAI-PMH harvester. He will share a number of issues that are consistently coming up while trying to harvest old and new installations of DSpace.

The goal of the discussion is that we clearly identify and list issues affecting the OAI-PMH implementation of DSpace and that we identify those use cases to make OAI-PMH in DSpace even more user friendly for Repository managers and external harvesters.

Update: Future of the DSpace User Interface

Tim Donohue will bring us up to speed on the latest news concerning the future of DSpace UI.

Open Repositories 2016 Dublin

Who is attending the Open Repositories conference in Dublin (June 13-16th)?

To which topics should we dedicate this opportunity to meet face to face?

Preparing for the call

Let's talk about OAI

Consider following questions in preparation of the call:

- What harvesters is your repository registered with?
- Which issues did you experience in the registration process with these harvesters?

- Which issues do you continue to experience while being harvested?
- Is there anything you wish to configure/customize, but that you are currently unable to, due to a lack of functionality or documentation?
- Are there any open OAI-PMH related issues in JIRA that we still need to resolve?

Open Repositories 2016 Dublin

Are you attending/planning to attend the conference? What would you like to discuss during the week?

Meeting notes

OAI-PMH harvesting

Friedrich Summann joined the meeting to present some of the most common pitfalls he noticed in harvesting numerous Open Access repositories to populate the BASE-SEARCH initiative. Although some 1000 DSpace's are properly configured for harvesting and indexing. For about 200 DSpace repositories base did encounter some issues. Mr. Summann lists the following problems:

- DSpace is often not correctly configured. This problem mostly occurs in African countries, India, China, Colombia and Ecuador. The problem lies in the incorrect configuration of the handle system, which causes OAI-PMH to be unable to harvest the repository. The exact configuration errors do vary. In some cases a default handle URL is shown, in other cases the end user UI and PMH interface show different links, or display the same erroneous link. Around 170 DSpace repositories suffer from this.
- Registered handle numbers are not correct or not working: It is not certain why this issue occurs. It might be caused by repository administrators who configure handles on their own, without registering with handle.net. Around 40 DSpace repositories have this problem.
- Someone tries to repair the situation: When Friedrich notices problems with a certain repository, he contacts the DSpace administrator. This often results in an attempt to repair the situation. In reality this sometimes results in a repository containing a mixture of old, still incorrect data and new, correct data. It is possible this is caused by incorrectly running the update script. Facilitating the correct use of this script through enhanced documentation could be a solution.
- The administrator is unreachable: Related to the previous problem, the repository administrator is often difficult to contact. In many cases the administrator email address is not configured. Often there is also no contact email address to be found in the end user UI. In case there is a contact form provided, submitting an entry to this frequently results into no response.
- The harvesting process crashes: This issue was found with newer versions of DSpace. During the OAI-PMH harvesting process an internal server error causes the harvesting problem to stop.
- OAI-PMH webapp is not deployed: Sometimes there is a working end-user interface, but no OAI-PMH webapp deployed.
- There are no UTF-8 characters in the responses but question marks: This is likely not a DSpace issue, but a Tomcat misconfiguration issue.
- Problem with the LYNCODE interface: This specific issue delivers a "No matches for the query" for ListRecords. It is possible this is caused by cronjobs which are not running (frequently).

Document that describes these issues in detail, including examples: http://bit.ly/BASE-SEARCH-OAI-issues-DSpace

OAI harvesting related questions from the community

OAISTER/Worldcat only supports HTTP, it does not support HTTPS. This problem was overcome on DSpaceDirect by configuring OAI to go through HTTP, while leading all other traffic through HTTPS.

The question rises if it would be possible to hide a collection for OAI-PMH. In case a collection is already access restricted in your own repository the contents will also not be harvested by OAI-PMH. It is possible the collection's name will be harvested, and thus be visible in the harvesting repository or platform. The contents of that collection on the other hand, will not be visible.

Update: Future of the DSpace User Interface

A prototyping challenge was held to prototype new UI technologies, as there already was a consensus the future of the DSpace UI lays neither with JSPUI nor XMLUI. In January all the candidate technologies were demonstrated. Since last week there is a group digging deeper in the different technologies, looking mainly for similarities.

However, the final decision for a technology is yet to be made. This will be discussed during next week's DuraSpace summit. One of the main points of disagreement is the question whether to stick to a server-side approach, or move on to a client-side approach using for example a javascript framework.

The DSpace UI working group has already agreed there is a need for User Experience (UX) expert who can help improve the end user experience. In case you have such a person in-house, who would be able to contribute to the new UI's user experience, feel free to get in touch with the working group.

DCAT meeting on Open Repositories in Dublin

As the conference program is already tight scheduled, we will try to have a short meeting over lunch.

DCAT members are free to suggest topics of interest. Currently an interest in discussing Statistics & Analytics is the only one addressed.

Call Attendees

- Bram Luyten - @mire
- Ignace Deroost - @mire
- Maureen Walsh - Ohio State University
- Tim Donohue - Duraspace
- Friedrich Summann - University of Bielefeld & BASE harvester
- Andrew McLean - Imperial College, London
- Irene Berry - Naval Postgraduate School

- Jonathan Green - University of Nottingham
- Monica Rivero - Rice University
- Felicity Dykas - University of Missouri–Columbia
- Susan Borda - Montana State University
- Pauline Ward - University of Edinburgh
- Sarah Potvin - Texas A&M University
- Daniel Draper - Colorado State University
- Marie-Hélène Vézina - Université de Montréal