Adding content reporting to DSpace

Andrea Schweer, The University of Waikato ITS, Hamilton, NZ; schweer@waikato.ac.nz
Jenni Barr, AgResearch Ltd, Hamilton, NZ; Jenni.Barr@agresearch.co.nz
Deirdre Congdon, AgResearch Ltd, Hamilton, NZ; Deirdre.Congdon@agresearch.co.nz
Megan Symes, AgResearch Ltd, Hamilton, NZ; Megan.Symes@agresearch.co.nz

Session Type: Poster
Conference Themes: Building the Perfect Repository
Keywords: Content Reporting; DSpace
Audience: Repository managers and their managers; repository developers

Abstract

This poster presents a content reporting add-on to DSpace, developed for AgResearch Ltd by the IRR support team at the University of Waikato's Information Technology Services Division. We outline the motivation for developing this add-on, give a high-level description of its implementation and report initial insights on its reception and uptake.

Background

Open repositories have become core systems in many research institutions. With this increase in visibility comes an increase in what is expected of repositories, potentially redefining the nature of the "perfect repository". A key requirement of the AgResearch repository from initial setup was that the repository – as the location of the institution's research outputs – feed data about these outputs into institutional reporting systems. While the OAIS reference model [1] includes reporting as a part of the functional model's Data Management component, its focus is mainly on managing the contents of the Archive rather than providing reports to an external system. Perhaps as a consequence, content reporting functionality is quite limited in current repository software.

Presentation content

At tertiary education institutions, academic staff are reasonably free to choose their publication venues and information is typically collected by the institution only after successful publication. In contrast, a significant proportion of research at AgResearch results in client reports that are commercially sensitive and not formally published. To protect intellectual property rights, all research outputs produced at AgResearch are mandated to undergo a release approval process before submission, resulting in institutional interaction with each research output at an earlier stage. Until recently, data gathering for reporting as well as archival of outputs were attached to the (paper-based) approval process. It was at times haphazard and involved many manual steps.

To address the need for more systematic archival of and reporting about research outputs, an internal-access repository was set up using DSpace in the second half of 2014. The approval

process was changed to an electronic process that feeds into the repository. The repository now is the authoritative location of all research outputs produced by AgResearch. The ability to extract data from the repository for institutional reporting was considered non-negotiable by stakeholders and essential to gain buy-in.

Consultation with user groups to identify requirements revealed that users wanted the ability to specify various criteria and then receive lists of all matching DSpace items in a format suitable for importing into a software package (e.g. MS Excel) and filtering or aggregating as needed. We added screens to DSpace XMLUI to enable the user to choose from a fixed list of report types and set parameters such as the start and end date for the report and a recipient email address. For some report types, the user can further restrict the items included by choosing from a list of all values of a given field (e.g. to obtain a list of reports prepared for a given client). Our code generates a corresponding query for the metadata indexed in the Solr core for the Discovery module. The Solr response, in CSV format, is post-processed and then sent to the specified email address.

Initial feedback has been positive, with anticipated use cases including quarterly and annual KPI reporting for the whole organisation or a specific subset such as a team or other organisational unit.

We believe a poster is the most appropriate format in which to present our work as the basic concept can be explained quickly (well suited for the "minute madness"); explanation of the work during the poster session can be as high-level or as technical as is suitable for a given visitor; and the interactive poster session enables discussions about implications for the "perfect repository".

Conclusion

Our poster demonstrates the need for content reporting as a part of repository software and shows a relatively lightweight addition to DSpace that meets this need. Our code is available under an open-source licence. However, we hesitate to propose inclusion of the code in the DSpace codebase at this point, both because the functionality goes beyond the traditional feature set of an Open Repository and for technical reasons. We anticipate that our poster will stimulate discussions at Open Repositories about the best way forward.

Acknowledgements

Stefan Mutter, formerly of the University of Waikato's Information Technology Services Division, implemented a major part of the reporting add-on. We thank Stefan for his work.

References

[1] CCSDS (June 2012), Reference Model for an Open Archival Information System (OAIS) – Recommended Practice. CCSDS 650.0-M-2 (Magenta Book).