

Proceedings of the New Zealand Institutional Repository Community Day 2015

Held Friday 4 September 2015 at University of Canterbury

The following notes were made by attendees (listed at the end of the document) and published under Creative Commons Attribution 4.0 licence.



Report back from Open Repositories 2015

Interesting things:

- DSpace 6.0 is to be released this year (mostly preparation for 7), DSpace 7.0 in 2016 (hopefully to be a single UI, not either jsui or xmlui) - see <https://wiki.duraspace.org/display/DSPACE/RoadMap> for more information.
- Portland Common Data Model: collection (has collection) has object (has object) has file
- Lots of OA 'compliance' tools being developed - interesting though not relevant to funder demands in NZ - Yet!
 - JISC Monitor (UK wide)
 - Lantern by Cottage Labs and JISC
 - Symplectic Elements Open Access Monitor
- IRUS-UK aggregated IRR stats
 - works with trackers a bit like Google Analytics - we don't have to run complicated reports, worry if we're gathering same data as others, interpreting data same way as others, etc...
 - checks for duplicate DOI in same repository :-)
 - JISC interested in international collaborations - unclear if they wanted a pay-to-host deal or to give code so we can do it ourselves

Problems:

- DSpace 5 has a bug around author order - fix for 5.4, 6.0
<https://github.com/DSpace/DSpace/commit/4a6663c2f44638b62f9dae52bef791517f87a34a>

Suggestion:

- If you're upgrading to 5, you may want to stick with it for a while and then look at 7 very carefully, maybe even wait for 8. This is because 6 will change so much architecture that it may impact integrations/customisations using java code (for example custom curation tasks)

Resources:

- Kim's OR2015 slides
<https://docs.google.com/presentation/d/1rYrVSVLymrIUkB5QhHs7r9cp2h3pcsyFsopt3wU8ibo/edit?usp=sharing> - includes more links
 - Andrea's OR2015 slides
https://drive.google.com/file/d/0B_EFnbkiziSaWHVCQ3MzanQ4dG8/view?usp=sharing
 - NZ IR mailing list / google group:
<https://groups.google.com/forum/#!forum/new-zealand-institutional-respositories-community>
-

Straw poll: which institutions are currently on which DSpace versions?

Institution	Currently on	Plans to upgrade?
Lincoln Uni	3.1	5.3 sometime 2016
U of Auckland	3.2 (IR), 5.3 ("superindex")	IR->5.3 this year
Massey	4.1	5.x this year
Otago	5.1 +	5.3 + this year
Victoria	1.8	5 or Islandora
University of Waikato	5.2ish	6 next year probably
AUT	5.1	TBA, probably 5.3
UC	1.8	5.3 this year (next week!)
Open Polytechnic	4.2	Upgraded end '14, no plans.

OA and conferences

Background:

- Deborah Fitchett (DF) has been thinking about how important conferences are going on all the time and papers are usually hosted by the conference site for a while (eg: a year) but then overwritten by later conferences or lost entirely as the conference site is removed.
- There is an idea of a global repository for conference papers but how to get there??

- DF contacted TEC about their dataset of citations submitted for PBRF and has been experimenting with the conference output data to try extracting key data elements. This dataset was once going to be open source but researchers involved refused permission.
- DF interested in a study to google these conferences and find out how many have papers still available / open access 4,5,7 years after the conference.

What are libraries doing with conference literature now?

- UC - since mandatory deposit are receiving huge amounts of conference paper submissions - many very minimalist - for those going into the repository most have a powerpoint rather than a true paper - better than an abstract though
- Researchers submitting “evidence” for PBRF purposes but it’s not good enough for Open Access purposes
- Resistance from academics who consider the journal paper the finished output and don’t want conference paper public
- Rights vary - sometimes author, sometimes society have rights
- OCS problem - conference keeps moving while you might want a static location for proceedings. Every conference is a separate event, often with a separate committee.
- *[I don’t think I mentioned this on the day, but here is NLNZ perspective on conference proceedings - AJ]* Legal Deposit has noticed a big drop off in published proceedings - documentation is usually schedule with abstracts, and then it’s up to individual researchers/speakers to make their content available somewhere if they wish to do so. Also, slight rise in proceedings being in the form of video, often on Youtube. There are legal issues with collecting this material (Youtube not subject to NZ legislation). If conferences move around the globe, then only falls in scope for LD in the years it is hosted in NZ, so the record would be incomplete (LD based on place of publication, not nationality of authors)

Possible ways forward:

- Using OJS instead of OCS - treating it as a journal issue
- AUT put papers of conferences at institution into repository even though feel repository should be for institutional research
- Policy change needed? ie that PBRF “evidence” should need an artefact instead of just programme. (But many oral presentations won’t have artefacts.)

Resources:

- OCS: <https://pkp.sfu.ca/ocs/>
- Some conferences held at Otago have used OCS, for example ascilite2014

ORCID and institutions / repositories

- CONZUL interested in looking at national collaboration / use cases / projects? And MBIE?
- ORCID, Digital Science, MBIE having a meeting to discuss benefits this year

- Discussed types of membership, business model / sustainability of org, best ways to integrate at institutional level -> downstream systems, oauth in Elements, authority control in DSpace
 - Encouragement is best, and self-signup is best. Nobody has a stick - except funders and journals! And some of them are already requiring ORCID iDs (eg Wellcome Trust) - might help with promotion
 - Tangents into DOIs, altmetrics - how they fit in (author metadata, ORCID as a “registry of linked identifiers” that make aggregation of things like altmetrics at an author level much easier)
-

Research Data

Environment scan

- Landcare using CKAN
- Canterbury using DSpace
- UoA doing a bit of DSpace, evaluating figshare, CKAN, DSpace
- VUW using DSpace
- Lincoln have proposed Research Data Management Suite including figShare, Data Fabric, Amazon Glacier - going through approval process and continuing to investigate

Issues

- storage/access/preservation vs. discovery/description?
 - tiered storage - figshare, data fabric, glacier
 - Identifiers - Landcare and UoA use EZID (DOI reg agency that registers DataCite metadata as well), others use DataCite direct, most DSpace institutions are using handles
 - DOI trick: <http://dx.doi.org/10182/6673> resolves to <http://hdl.handle.net/10182/6673> :-D
 - versioning - how can we/our id systems allow people to cite versions of datasets or “as at <timestamp>” citations? (as wikipedia or git[hub] history allows)
 - ARK - https://en.wikipedia.org/wiki/Archival_Resource_Key - might be able to do something like this? - can point to ‘fragments’ of an object
 - and, from this..
 - Recent CONZUL working group met to discuss future standards/collaboration?
 - eResearch 2020 / NeSI - visions, strategy for future - early 2015 had a day in Auckland - some notes at <http://deborahfitchett.com/blog/2015/04/towards-2020-research-data/>
 - Do places like DataDryad publish their metadata to eg DataCite / Elements / Primo etc so we can track all institutional research datasets no matter where published?
 - and is there a benefit to a NZ version of data.ac.uk (national data registry)
 - Does figshare have some advantages for institutions (in later implementations) in terms of already holding some publisher copies of data? (maybe!)
 - What about (private) storage for active data? Lincoln seem to be progressing well, everyone else is still implementing/discussing?
-

Symplectic Elements <-> DSpace

Current situation:

- VUW just starting with Elements
 - not linked with repository yet
- Lincoln have just switched on linking between Elements and repository
 - used repository (plus research office system - attempted to merge data) to initially populate Elements
 - some problems with Elements overwriting data in repository
 - crosswalks in place
 - still working out workflows
- Massey have had Elements for a while, have integrated Elements with dark DSpace archive, then harvest from there into public DSpace repository
 - no problems with changes, errors etc since they aren't public
 - but need to get the harvesting right
 - how to get academics to deposit stuff?
- AUT have had Elements since 2010, including DSpace integration
 - dark collection inside public repository
 - customised workflow to process file changes, new items
- UoW have had Elements with DSpace integration for about a year
 - No dark archive, items go into main repository workflow
 - same customised workflow as AUT
 - recently discovered some problems with merging
- UoA have had Elements since 2010, DSpace integration since 2011
 - (largely same issues as others, didn't attend this session but adding this in for the record!)

Possible ways forward:

- customised workflow (AUT, UoW) appears to work reasonably well but is hacky
- could probably be done using DSpace customisable workflow -- but complex, not well documented, possibly not compatible with Repository Tools for DSpace
- review workflow for changes is a bit of an overlooked use case in all repositories

Resources:

- <http://hdl.handle.net/10289/9339> (OR15 poster on AUT, UoW workflows)
- Symplectic user group Australasia
<https://groups.google.com/forum/#!forum/symplectic-user-group-australasia>
- NZ Symplectic users mailing list
<http://list.waikato.ac.nz/mailman/listinfo/nz-symplectic-users>

Open access resources

- Australasia Open Access Support Group <http://aoasg.org.au/> If you're new to this area, check it out.

- Danny Kingsley article, possibly useful for countering academic arguments against OA
Open access: whipping boy for problems in scholarly communication
<http://aisel.aisnet.org/cgi/viewcontent.cgi?article=3879&context=cais>
response to rebuttal
<http://aisel.aisnet.org/cgi/viewcontent.cgi?article=3885&context=cais>
-

Pre-lunch vote:

“That these notes be considered the ‘proceedings’ of today’s event and deposited into an institutional repository with a Creative Commons Attribution licence.”

--Passed.

Indigenous knowledge notice - presentation by Karaitiana Taiuru

Atua relevant to internet communications:

- Papatūānuku - fibre optics going through the ground
- Ranginui - communications through microwaves through the air
- Tangaroa - cables through ocean
- Tāwhirimātea - signals that pass through the air

No NZ or international laws recognise indigenous property rights. WAI262 is complex and doesn’t help current authors.

Traditional Knowledge Labels <http://www.localcontexts.org/> are more centered around Native American groups.

Calling something a ‘licence’ brings up legal implications - that’s why talking about indigenous knowledge “notice”.

Digital Colonialism - <http://www.taiuru.maori.nz/#article/4068>

- Don’t digitise without consultation - and make sure you consult the right people - eg local iwi; family/iwi of a book’s author. Ensure informed consent by explaining ramifications of how digitised material might be used. Images - Māori carvings and tattoos all have unique meanings
- Data sovereignty vs the USA PATRIOT Act, proprietary formats. Make sure you know where it is being formatted. There is even an issue with proprietary software / platform / server
- Working on a draft of Indigenous Data Sovereignty. Something might be out-of-copyright (eg Elsdon Best) but still belongs to iwi. If you do digitise something, be super wary of putting it on eg Facebook!
- Old books (Books have an atua) - often informants gave permission for tapu information to be put into a book, but had no idea of the internet and info flowing freely via the veins of Papatūānuku - ancestor’s photo might flow through enemy lands, or through cemeteries.

- Concerns of tangi being turned into Virtual Reality from a video recording.

Presentation pdf:

- <http://www.taiuru.maori.nz/wp-content/uploads/NZIRtaiuruk.pdf>

Links to related presentations

- <http://www.taiuru.maori.nz/wp-content/uploads/Lincon-University-Digital-Colonialism1.pdf>
 - <http://www.taiuru.maori.nz/wp-content/uploads/The-Internet-infrastructure-and-technologies-from-an-Indigenous-perspective-comparing-Maori-traditions-and-genealogies.pdf>
 - <http://www.taiuru.maori.nz/wp-content/uploads/Impacts-and-Considerations-for-Indigenous-Populations-using-Open-Source-TaiuruK.pdf>
 - [Via twitter](#): Also see Anahera Morehu – Housing the knowledge of tangata whenua video <http://www.ikebarberlearningcentre.ubc.ca/morehu/>
-

Te Reo and DSpace

Continuing a discussion from the mailing list on getting a translation done.

Issues to consider:

- Local dialects - eg 'ng' vs 'k' can do find/replace. Vocabulary would be more an issue but this isn't so big an issue for technical vocabulary.
- Macrons probably okay (different than issues some have had with metadata)
- Localisation works in xmlui - has messages.xml file. In v3 which one displays depends on browser settings; in v5 user can switch back and forth.
- How much sustainability issues - additions on each upgrade?
 - Brought this question to the "DSpace v5.3 upgrade" session later which responded:
 - messages.xml will change completely for v.7. DSpace UI working group working on prototype - not sure if internationalisation is on their radar so check that!
 - <https://wiki.duraspace.org/display/DSPACE/DSpace+UI+Working+Group>
 - someone's written a tool that checks if there are new codes not in your existing message <http://dspace-i18n-check.esy.es/web/>
 - More than just messages.xml: also email templates, messages.properties (for jspui - but may be affecting other places too), and each aspect

Path forward:

- Seems to be support to go ahead
- Karaitiana can put the file into an app which makes the translation work easier
- Could do v1 - machine translated; (v2 crowdsourced?); v3 professionally
- Needs proper scoping - Deborah and Amanda C to coordinate on Google Doc
- Make clear with translator that this is an open source project so outcome is to be open source

- Beyond this: Would there be value in a inter-university group (uni-wide not just library) that could be a go-to group for translations/questions?

Resources (urls, people, places):

- Moodle has a Te Reo interface; Lumio; Linux; Microsoft Office; Silverstripe - meaning much vocabulary has already been sorted.
-

NZResearch / KRIS OAI harvesting discussion

Background - in KRIS, work was done to harvest all the university repositories, via OAI-PMH (simple DC), a lot of standardisation was done on metadata particularly around type

It turned into nzresearch, which was still then a separate 'thing' to DigitalNZ, they're a bit more merged now. [need more help on background / harvesting issues!]

- there's an international (COAR) proposal on type vocab (see session below), we should consider those
- also access rights
- (and i want to add: possibility of NZ doing things like Aus do with funding / grant identifiers mandated in exposed metadata)
- who still does ANZSRC codes? do we still need to?
- "custom" / specialist feeds are only going to get more popular (subject specific things like CEISMIC, etc.), we can collaborate to make things easier for harvesters
- peer reviewed indicator?

DSpace XOAI makes it tons easier to crosswalk and standardise our metadata output, exposed sets. Can we get together and work on some standardisation ourselves, then liaise with nzresearch?

Some perceptions/concerns

- Has DigitalNZ 'swallowed' nzresearch a bit: harder to contact and get things changed or done?
- Does nzresearch need more promotion/awareness among the research community? (people who "should" be using it aren't, our experience tells us!)
- Do we still need it at all? What would be done otherwise? (in terms of DigitalNZ / separate interfaces)

nzresearch's talking points:

- this is a high value service - specific audience... but that narrow audience means it's hard to get more resource to put into changes, etc.
- a collection response/submission would be more useful than individual requests

Action - Andrea / Allison to ask DigitalNZ/NZResearch whether they'd have resources to adjust their harvesting if we were all to adjust our oai feeds to include things like rights information, free-to-read, peer review status etc

DSpace - upgrading to 5.3

Not big difference between 3 and 4

5 has xoad improvements

Need to start bugging Symplectic about support for upgrade to v.6 as Graham has left them (but is still a DSpace committer). What if we can get them to release RT4DS open source?

- For v5 there were problems with rt4ds - version dependency issue. Solution would be easy at Symplectic end. In meantime ask Andrea on mailing list.

Problem with geoinformation in usage statistics on upgrade

Might be good to go to 5.4 if it's out in time as it should fix bugs with searching - fix is currently being tested.

COAR proposed types:

Background:

- In NZ KRIS metadata standard document tried to agree on list of types. Ended up with 2, one for DSpace and one for ePrints.
- Then some unis started using PBRF types.
- Then Elements has its own list of types and you can create your own too.
- COAR proposing this list as a comprehensive list of research output types, then would lobby for eg DSpace, Elements, EPrints to change their default set. Note it's a hierarchy and it's linked data with authoritative translations. They're asking for feedback.

Notes to feedback:

- Need to add honours theses; undergraduate research report
- Need to make sure PBRF types are included usefully - link to TEC report
- Others missing: architectural drawings, 3d schematics
- sound - should include audiobooks, interviews?
- ephemera
- blog, software paper (cf 'data paper' which they have)
- creative works, exhibitions
- "lecture" is there but it's an event not an object. PBRF also allows other types of performances which has the same issue. An object might be photos or videos which would come under those headings.
- "video" shouldn't be subset of image as includes audio
- How does it work for collections? Zip files? Could dc.type be a repeatable field? What about a collection of multiple images?

Our response:

- Need to get CRIs involved

- Do we want a national or individual response? Only have a couple of weeks so individually is probably easier. Deborah to mention to TEC and Ex Libris. Someone else to talk to Symplectic. Andrea making a note to bring it up with DSpace committers.
- Provide feedback at <http://www.coar-repositories.org/activities/repository-interoperability/ig-controlled-vocabularies-for-repository-assets/deliverables/>

Resources:

- http://vocabularies.coar-repositories.org/documentation/resource_types

Session topic:

Current situation or problem(s) described:

- Legal Deposit/Digital Preservation

Possible ways forward or solution(s) suggested:

- Question: how would your institution feel if NLNZ scraped all the Open Access content (that your institution is the publisher of) and made copies openly available directly from NLNZ's own repository?

a) That would be fine, more access is more access

b) Yes sure, but we would want good reporting on how our content is being used

c) We'd really prefer that you send researchers back to our content, please

d) Other _____

Would it impact what goes into your agreements with your depositors/digitisation partners/etc?

- NLNZ to discuss the relationship between LD and indigenous knowledge

Resources (urls, people, places):

- Draft NLNZ collections policy: <http://natlib.govt.nz/about-us/strategy-and-policy/new-collections-policy>. Please send comments to amy.joseph@dia.govt.nz

Wrap-up and Future of NZIR Community Day:

What people got out of today:

- many mentioned continuing discussion re indigenous knowledge, digital colonialism - IP, legal deposit; conflict with open access is challenge to grasp with and balance
- upgrading discussions
- interesting to see what goes on behind the scenes making a database work
- looking forward to contribute to COAR vocab
- lots of topics to follow up on - DAMS
- hearing perspectives from other institutions - easy to get tunnel vision so good to open up
- research data management, hearing what Lincoln's up to
- encouraging academics to submit; controlled vocabulary
- squeezing conversations about identifiers into every session

- expected it to be about finding resolutions but was more about finding common problems
 - lots in notes about people aiming to feedback so important we continue these conversations
- challenges of working with researchers
- meeting people have corresponded with over the years
- being part of discussion again
- need to go back to organisation and talk about digital preservation and what is and isn't being done
- how other universities are working, different challenges, sharing issues and ideas
- realising haven't mentioned indigenous knowledge in copyright sessions currently being run - going back to fix that
- picking brains between sessions
- future of DigitalNZ
- support to progress Te Reo translation of DSpace

Actions

- Te Reo translation scoping
- Metadata we can expose to DigitalNZ
- COAR vocabulary feedback

Future of NZIRC Day

- Agreement that there's value in having an NZ separate event instead of being part of an Australian group - ease of attendance, and see all the interest in local issues like indigenous knowledge, working with DigitalNZ, National Library etc.
 - Kim's been asked to volunteer Auckland as a location for next year. Discussion ensued in general favour.
-

List of attendees

Allison Brown
Amanda Curnow
Amanda He
Amy Joseph
Andrea Schweer
Andrew White
Anton Angelo
Aurelia Arona
Deborah Fitchett
Emma Richardson
Fiona Wilson
Glen Walker
Helen Thomas
Karaitiana Taiuru
Kate Miller
Kerry Gilmour
Kim Shepherd
Leah Pengelly
Mariko Yoshioka
Max Sullivan
Michael Parry
Nekerangi Paul
Peter Kennedy
Roger Dawson
Rudy bin Mahli
Sarah Tritt
Simon Cooke
Steve Donaldson
Sue Allan
Tom Avery
Valerie Anderson