OAI-PMH Design Meeting

Date: Thursday March 23, 2023 @ 10am Eastern

Where: https://lyrasis.zoom.us/j/85992335527

Attendance:

Ugnius Lukosius

Antanas Streimikis

Andrius

Oliver Schoner

Gytis Vievesis

Vicky Philips

Jared Whiklo

Joshua Westgard

Resources:

https://wiki.lyrasis.org/display/FF/OAI-PMH+Use+Cases

Prior Work

- Fedora Ontology exists https://fedora.info/definitions/v4/2015/03/26/oai-pmh
- Python/Django Fedora 6 implementation https://github.com/jnphilipp/django_oai_pmh
- Islandora implementation https://github.com/Islandora/islandora_oai
- PROAI https://proai.sourceforge.net/ and https://github.com/fcrepo3/proai
  - Similar E-Theses project which used OAI-PMH to harvest NLW theses from Fedora 3 w/ ProAI: https://ethos.bl.uk/
- Wiki on proai implementation https://biowikifarm.net/meta/OAI-PMH_service_for_K2N_Fedora_Commons_based_repository

Use Cases

NLW

- Updating catalog with digital objects
- Sent collections via OAI to external bodies
- Want to share their collections with others

Berlin State

- Possible use case for receiving notifications from other repositories (imaged use cases)
  - Updates to relevant changes in other repositories getting notified about this
- OAI not a dealbreaker for them - they're just curious
- Considerations:
  - Need to consider amount of traffic w/ large metadata and how to handle it

UMD

- Fedora 4
- Participating in DPLA - aggregate metadata from other repos via OAI-PMH
  - Going to build it themselves via a grant
  - Looked at (https://pro.europeana.eu/data/repox)
  - Maybe look to University of Alberta had written something for Fedora 4 but they stopped using it and there were no maintainers
- Use PCDM for modelling

Lithuania (https://en.ktu.edu/) (https://www.elaba.lt/elaba-portal/en/kontaktai/is-prieziuros-darbo-grupe)

- Fedora 3.8 w/ ProAI
- Want to migrate to Fedora 6 in next year
- National repository with shared discoverability/access(Open Aire?? For metadata)

Considerations:

- OAI spec is quite thorough
- Intention to build as a separate service that lives near Fedora, but not built it
- Speed of search

- What constitutes a "set"?
    - Need to figure out how to allow users to set their own parameters on defining "set"
- Deletions?
    - Maybe something with camel to resolve?
    - Does the spec dictate whether or not the URI can be reused?
    - Need to have a way to mark records as having been deleted
    - Delete then purge between harvest would create discrepancies because the tombstone would then be gone and the URI would be available again
- Different metadata types (ie. DC or MODS)
- Permission levels
    - Restricted content could still possibly be harvested by others with the same access levels might be a plausible use case
    - Would be nice to have support for different permission levels within the OAI endpoint to allow to set access permissions
    - Use case for things being discoverable but not accessible
- Need to be able to define the criteria for which the metadata is mapped

Chat notes:

Deleted is a MUST in the spec and has 3 possible implementations http://www.openarchives.org/OAI/openarchivesprotocol.html#DeletedRecords

There's no mention of authz/authn in the spec so that would be implemented in the HTTP layer I expect.