2016-03-02 Research Data Repository Interoperability BoF Meeting Minutes

Ideas Behind the BoF

- Started at P6 in Paris to discuss a standardized way to communicate between different repository platforms
  - Idea presented during Repo Platforms session. Positively received
- Held a phone conference in October
  - Collected additional technologies and candidates for interoperability
  - Standards may include generic APIs, import/export formats, etc.
- Use cases
  - Deposit/retrieve digital objects into/from research data repos
  - Migration/replication of digital objects from/to different platforms
  - Retrieval of information regarding the repo itself or its contents
- Value proposition
  - Remove barriers
  - Easier collaboration
  - Creating commonalities
- Related work (list on slide). Others:
  - IIIF
  - Shared Canvas
- Deliverables
  - 1. Interoperability primer
  - 2. Interface specification draft
  - 3. Interface specification
- Next steps
  - Identify gaps and additional contributors
  - Finalize case statement
- More information
  - https://goo.gl/yjDTn6 (RDA Group Page)
  - https://goo.gl/8WJ5oJ (Case Statement)

Questions

- What is the problem tried to be solved, what is the focus?
  - Machine-machine communication in focus
  - Ability to 'rescue' a Digital Object in case of a technology change
- People who tried to do this in the past and problems they faced
  - e.g. mapping metadata, terms not in a schema, etc.
- Survey repos to get state of the art

- - - We already have a lot of interoperability components but they are not widely adopted
    - A new API does not solve the problem
  - FAIR port
    - Implements repo based on linked data, FAIR
    - FAIR point
      - Any data source should have an interface
      - Must be a simple protocol (e.g. HTML/HTTP)
- Find a consensus around what kinds of functions a repository must support
- Protocol requires repos to actually implement it
- Use case: Pull different objects from different repos to work on them simultaneously
- SWORD is not performant, need 2 endpoints
- Functional requirement specification
  - Similar to practical policy group
- Data centres have more complicated data structures
  - Database and file based worlds are very different
- Use web technologies rather than layering things on top of (web based) technologies
- Some repos are very old and underfunded
  - In many domains there is no activity to bring together repos
  - Old technologies like FTP
- An RDA recommendation would resonate with funding agencies to implement the standard
- SHARE API
  - Pulls metadata for search
  - Uses custom harvester for every single one
- Should the WG just have the right conversation, or must we focus on a solution?
  - Comment: Strive for a consensus and not for a solution.

IEEE Big Data Conference

- Big conference, 800+ attendees
- This topic might be feasible as a workshop
- Need volunteers to be additional co-chair (1) and program committee members (approx. 5)
- Proposal will be submitted by WG chairs till May
- Call for papers for workshop approx. beginning at June (announced via WG mailing list)

Case Statement

- Must submit within 2 months

Scope

- Here is the potential scope and here is what we are doing
  - e.g. aggregation use case