

FORCE11 - DCIP Executive Meeting
December 2, 2015 08:00 AM Pacific Time

Invitees: Maryann Martone, Tim Clark, Carole Goble, Jeffrey Grethe, Jo McEntyre, Joan Starr, Simon Hodson, Martin Fenner

In Attendance: Martin, Maryann, Joan, Stephanie, Jo, Cassie

Regrets: Simon, Tim, Carole, Jeff

AGENDA

1. Finish Review of [Slides](#)
 - a. Repository expectations (slides 6)
 - i. Support minimal metadata required for citation
 1. Martin – citation styles don't support this; not easy to support versions
 2. PDB does not currently support versions but you can't argue that they are not reliably archiving data. Additionally, not all datasets have authors or version
 3. The good, better, and best levels are more of a matrix instead of linear progression
 - a. Should we drop the language of good, better and best and go by use cases instead?
 - i. Just as there are different ways of citing a book, there are different ways of citing data
 1. DataCite provides three ways to cite data
 - ii. Support minimal metadata required for citation
 1. We should list examples of this to support something more formal
 2. Will provide examples of what data citation looks like – must be clear that this is not a citation style
 - ii. Provide your existing identifiers in manner that is persistently resolvable, and participate in the PID solution provided by the project team
 1. This won't include machine readability (this is in the best category)
 - iii. Best: assign PIDs that are machine readable, globally unique, and widely used by a community to deposited datasets
 1. Citation in the body of the text vs. full citation in the reference list
 - a. The local accession number, is that what you include in the text? This doesn't speak to its actionability (globally or locally)
 - b. John Kunze showed that with minimal effort on the part of the repositories, he can set up a rewrite rule that would

create an actionable form of a PDB number so that people can click on it and see the dataset

- i. We have a resolver like identifiers.org
 - ii. DCIP could contract with CDL – not use EZID to maintain anything (there’s no exchange of metadata, etc.), it’s simply a redirection of metadata
 - iii. No one tracks it, it’s just a rewrite rule
 1. CDL is a member of DataCite
 - iv. The point is to be able to resolve identifiers in some way
 1. It doesn’t matter how identifiers are resolved as long as they are resolved – may not be a long term solution but for this pilot it would be best for it to simply be resolvable
 - a. We should provide examples of what is persistently resolvable (acceptable resolutions, not identifiers themselves)
 - i. Provide your existing identifiers in a manner that is persistently resolvable – change to the better category
 - ii. If repositories do not have a PID solution, we can provide one
- b. Repository Next Level (slide 7)
- i. This is only for repositories that join at the best level
 - ii. Use PIDs that resolve to landing page then to data (via content negotiation, for machines)
 1. Enforcing this will be difficult
 - a. We are not enforcing it. This is what we discuss in machine actionability article as part of what “good citizen” repositories would do. We are not making them do anything, we’re just showing them
 - b. This list is mostly for our own use, it may be too intimidating to show to repositories
 2. HTML meta tags – makes a machine readable landing page
 - iii. Provide a persistence policy statement (refers to machine actionability article)
 1. PANGEA could be an option for this
 - iv. Next step
 1. Pull together some examples of each of these statements to ensure that we’re all on the same page
- c. Exec team (slide 9)

- i. Repositories: determine basic level of readiness; agree on goals; help identify repositories participation in pilot
 - 1. Jeff said he would be the lead on this?
- d. Open issues (slide 10)
 - i. Goals need to be augmented by pulling together example materials
 - ii. Decision on short term PID strategy for repositories (N2t.net)
 - 1. **Action:** post [John Kunze's slides](#) from a previous DCIP meeting
 - iii. Role of the reference management vendors
 - 1. Include HTML tags
 - iv. Identifying the list of repositories
 - 1. List of repositories includes who is participating and who responded to the JATS poll
 - 2. Maryann wants to include some neuro-imaging repositories and Martin will try to include Dryad
 - 3. Anyone that has additional people to invite, add to the list
 - a. We should make a cleaner slide deck for sharing with people – should include examples and remove the “scary” list
 - i. **Action:** create a pitch version of this slide deck
 - b. Maryann and Joan are presenting this slide deck at the bioCADDIE All Hands Meeting
 - c. We should see which repositories endorsed the DDCIP
 - v. Set up a meeting date in January/February
 - 1. **Action:** get a date set for this

< Meeting adjourned >

- 2. Results of JATS Poll
- 3. Determine Next Steps

DCIP Action Items:

https://docs.google.com/spreadsheets/d/1LfkcAq34qvC5Su_4cM9dH-8pCYi6wGNyWSw58ao7mlc/edit?usp=sharing