

Minutes of JAV conference call 5th January 2005

Present: John Ober, Peter McCracken, Scott Plutchak, Claire Saxby, Catherine Jones, Andrew Wray, Bernie Rous, Cliff Morgan, Evan Owens, Bev Acreman

CM had suggested taking a more simple, practical approach as a starting point. As such CM's "strawman" model, based on 5 top-level stages of an article, formed the basis of the discussion

1. JO agreed that a more minimalist approach is useful but felt strongly that the strawman doesn't address relationships, although CM felt they could be expressed through DublinCORE. It was suggested that we should agree the terms and what they mean first and then return to the question of relationship - CM suggested that we could incorporate a relationship metadata tag, like a "bucket" to include any relevant information (because relationship changes over time). We could provide guidance for describing these relationships.
2. TSP outlined Elsevier/NLM removal/retraction discussions. There seemed to be some disagreement within the community as to whether removing an accepted manuscript version that had been posted online, without replacing this with an explanatory notice, is acceptable. PubMed indexes such versions (e.g. Elsevier Articles in Press, Online Early etc.) and so when these articles are removed by the publisher this leaves dead links. CS said that when OUP publishes an accepted manuscript version in this way through 'Advance Access' this establishes the official date of publication, and this version remains accessible when the final definitive version has been published online (whereas Elsevier appear to remove Articles in Press when the final version is published). This raises the question of how 'publication' is defined - where would an Article in Press sit in CM's strawman? Publication may be defined as being made 'publicly available'; CM suggested that publication should mean that an organisation has taken legal responsibility for the work. BR suggested that because terminology such as 'published' already comes with different interpretations, we may need to reject this and come up with new terminology. Version of Record was proposed as an alternative (It was commented that some use this to distinguish between HTML and PDF versions... Elsevier have used the term 'final definitive/archival version).
3. The 'version of record' will not always be the latest version of an article. We discussed how best to refer to an 'updated version of record'.
4. It was asked how working papers/technical reports (often deposited in IRs) fit into the scheme. These are a final product even if not peer-reviewed, but they may then inspire or be tightly linked to the

definitive version of an article in a journal. This relationship seems important. Do we need a label for these? It was commented that technical reports/conference papers etc. do not always become journal articles and are not necessarily of less intrinsic value, whereas the 5 journal article stages outlined so far have in common that value is added from 1 to 5 - each subsequent version in a sense replaces the previous version. Therefore are technical reports/conference papers etc. a separate category altogether and as such out of scope? AW said he tends to think of these as different articles not journal article versions, but when they do have a relationship with a journal article it is useful for this to be exposed, especially if such a work has the same title/authors as a journal article and can be found on the web. The group could try to provide guidance here.

5. AW highlighted the existence of "Living reviews" which are added to: this he felt reinforced the need for relationships to be spelt out as different "editions" of an article may fork. A new journal bibliographic reference is assigned to each new edition of the review.

6. TSP noted that an author's "final published article" may not be formally published at all, but reside in an IR. The importance of defining 'publication' was again discussed in the context of institutions encouraging faculty to 'publish' their work in an IR rather than a journal, and how to bring these 'publications' into the scope of the project.

7. CM said that the terms agreed for each stage should be accompanied by a comprehensive list of attributes.

8. AW noted that on the ArXiv pre-print server a comments box allows users to enter free-text to explain what a version is and how it related to other versions. It was commented that this community is very familiar with and well disciplined with regards to self archiving. However even in this case free-text comments may be confusing.

Action - all to review terminology for the 5 (or more) top level stages and circulate alternatives for discussion prior to next phone conference. We will then apply terms to the use case to see how they would work in