

**A Proposed NISO Work Item:
Development of Standards to Support Bibliographic Data Exchange**

**Proposal for Consideration by the NISO Voting Membership
Approval Ballot Period: December 18-January 23, 2014**

**The following proposed work item is submitted by:
The NISO Content and Collections Management Topic Committee**

**Proposal Last Modified: December 18, 2014
Approved by the NISO [Content and Collections Management Topic Committee](#),
November 14, 2014**

GOAL:

Creation of ANSI/NISO standards covering vocabularies: policies on use and reuse, recommendations on documentation, and requirements on preservation.

INTRODUCTION AND STATEMENT OF PROBLEM:

A number of potential work items were identified during the course of the NISO Bibliographic Roadmap Project¹. To some extent, the identified areas echo those noted in the DCMI Vocabulary Management Special Session Report². Standards organizations can lead the way, and if careful to include representatives of the important players, can create useful tools as a roadmap to the future. Their work is in facilitation of simplifying the path, of finding ways to make information sharing easier. NISO's Content & Collection Management Topic Committee (CCM) is the appropriate body to drive the further practical exploration of existing vocabularies, linked data tools, and methods for effective future exchange of bibliographic data.

This document proposes establishing three Working Groups to consider different areas for further exploration: Vocabulary Use & Reuse; Vocabulary Documentation; and Vocabulary Preservation. CCM recommends a working group structure be established whereby a Steering Committee will oversee the three working groups. The Steering Committee would include the chairs of the working groups and would ensure coordination of effort across the three areas of focus.

DESCRIPTION OF POTENTIAL AREAS OF WORK:

1. Potential Area of Work: Policies Supporting Vocabulary Use and Reuse

One barrier to exchange and interoperability is the lack of policies in several areas relevant to use and re-use of vocabularies by outside organizations (i.e. not the owner or maintainer of the vocabulary). At a minimum, recommended best practices should be established.

This lack of best practice policies includes several categories:

- Rules of conduct and good practice when using, modifying for local purposes, extending

¹ NISO Bibliographic Roadmap Project Summary Report:

http://www.niso.org/apps/group_public/document.php?document_id=13327&wg_abbrev=ccm

² DCMI Vocabulary Management Special Session Report:

http://wiki.dublincore.org/index.php/DC-2011_Vocabulary_Special_Session/Meeting_Report

and aligning a vocabulary owned by another entity. Questions that need to be addressed in this area include both the technical (How do we extend a vocabulary correctly? Is it necessary to use the same license as the originating vocabulary?) and the social (Is an entity extending a vocabulary obliged to notify or ask permission to do so?)

- License and rights requirements and expectations, including expectations around IP rights for language versions, whether done by the owner or others.
- Commercial usage, including the question of what the absence of a license implies. What is the value proposition for a vocabulary owner in allowing reuse, extension, or any other adaptive activity around their vocabulary?
- Proliferation of versions (extensions and modifications) and guidelines around when changes are such that the result is in fact a new vocabulary with “ownership” defaulting to the organization responsible for the modifications.

2. Potential Area of Work: Documentation

Documentation best practice is an area where collaboration with the Linked Open Vocabularies (LOV) project might be very helpful. That group has already put out recommendations for documentation for vocabularies as a whole (see <http://lov.okfn.org/dataset/lov/about/> under the “LOV suggest” section on the bottom of the page for the link to the recommendations).

Documentation about vocabularies is vitally important, particularly for potential users in the stage of evaluation, but also for the community already using the vocabulary. Documentation should include the descriptive information suggested by LOV as well as policy documentation.

Documentation could usefully cover governance of the vocabulary, and information about maintenance and sustainability.

A standard could be established which specifies the minimum information about the vocabulary (i.e. the properties that are most important to users) to be documented in order to meet the minimum needs of the community.

3. Potential Area of Work: Preservation of RDF vocabularies

This problem area includes the issue of orphan vocabularies, an area that DCMI and LOV have also highlighted³. This situation comes up when organizations abandon vocabularies for lack of funding, or when a vocabulary loses funding or cannot make the transition between print and digital. There are community questions implied: under what circumstances can the vocabulary be taken over (‘adopted’) by another entity? Is permission required, and how are the domain and URI issues resolved?

Some entities lack the technical expertise to transition from print or to maintain a digital vocabulary properly. Can some options be explained and a process defined to ‘protect’ existing vocabularies from disappearing? As part of the examination of this problem the minimum requirements for effective preservation need to be defined: structure, documentation, relationships with other vocabularies, or all of the above?

³ “Long-term Preservation and Governance of RDF Vocabularies”, 2013 DCMI International Conference on Dublin Core and Metadata Applications, Lisbon, Portugal.
<http://dcevents.dublincore.org/IntConf/index/pages/view/vocPres>

At its heart this problem is one of outdated business models and lack of agreed remedies when valued vocabularies are abandoned by their entity. This rarely happens with large commercial vocabularies, but is more prevalent in the domains where smaller, special purpose vocabularies have been the norm. The growth of interest in distributed datasets provides yet another reason for attention to these issues, as many datasets include references to external vocabularies.

As part of this task, some effort should be spent determining whether there are pathways possible to manage these vocabularies in a community supported repository, perhaps with a governance model to manage short and long term issues for the vocabulary community.

WORK TIMELINE

Approval of Proposal	January 2015
Appointment of Working Group chairs and First Meeting of Steering Committee	February 2015
Approval of Initial Work Plan and Appointment of Working Groups	February 2015
Completion of Initial Draft Standard(s)	November 2015
Public Comment Period	December 2015
Completion of Final Draft(s)	June 2016