Introduction & Context

The American Library Association’s Metadata Standards Committee (MSC) contains representation from three of its topical Divisions: the Association for Library Collections & Technical Services, the Library Information Technology Association, and the Reference & User Services Association. The MSC welcomes the opportunity for public comment on the draft NISO Z39.102-201x, STS: Standards Tag Suite.

The MSC has published a set of “Principles for Evaluating Metadata Standards,” http://metaware.buzz/2016/08/04/principles-for-evaluating-metadata-standards/, which we have used to structure our comments. Please note that our principles are tailored more towards structured metadata, rather than markup of full text documents. Nevertheless, we hope that our comments are useful to the NISO STS committee.

1. Metadata standards should be part of a shared data network

As a document markup language defined as an XML DTD, W3C XML Schema, and RelaxNG schemas, it would be difficult to expose data encoded in NISO STS as Linked Data on the Web. For example, the current NISO STS element and attribute definitions do not seem to make a systematic effort to provide places to encode URLs for named entities. Given the Web community’s focus on Linked Data today, the NISO STS standard would benefit from some thought in how to make the standard a bit more Linked Data friendly.

We note that most, if not all, of the NISO STS document markup could be done in the Text Encoding Initiative (TEI) community standard. We recognize that TEI is not nationally or internationally formally standardized and there is, in many communities, a strong need for a more official standard for encoding of standards documentation. However, the TEI does have wide adoption and use among certain communities, and clear statements on the relationship between the two would benefit both current and potential user communities for both languages, and promote linking between them.

Crosswalks for descriptive metadata features of NISO STS to other popular metadata standards, especially in the bibliographic realm, would promote use of STS encoded documents by other communities. We do note that the decision to not require a default namespace in the interests of retaining backwards compatibility does somewhat limit its ability to play nicely with metadata in other namespaces.

2. Metadata standards should be open and reusable

NISO makes its standards open for reading and download, and that is appropriate for this standard. Making the standard open is an important part of encouraging adoption and use.
It is slightly unclear to us what the exact relationship is between the ISO STS, version 1.1 and this NISO STS draft. The NISO standard would be easier to select, adopt, and use if this relationship and the differences between the two were made clearer in the documentation.

As with all NISO standards, the governance and oversight of STS is clear and structured through regular NISO procedures, which promotes confidence in the standard. The detailed documentation in the NISO “workroom” for this standard further makes it accessible.

We note that NISO standards are issued as copyrighted documents as a matter of course, which comes with an implicit statement that all rights are reserved. A more open license for these standards documents themselves would increase their utility to the community. The DTDs, XML Schemas, and RelaxNG schemas would also benefit from explicit and permissive licensing terms.

3. Metadata standards and creation guidelines should benefit user communities

Per the comment in item 6, below, the standard may benefit from use case exploration that would develop approaches to lighter-weight use cases. The general introduction does provide an overview of the defined use case “standards bodies, standards producing organizations, publishers, commercial vendors, and archives can publish and exchange standards documents.” Are there other use cases or other user communities that would benefit and whose needs should be considered? For example, a discipline-based community formalizing its own metadata standard might not even know NISO STS exists, or if it does, would likely find it difficult to implement due to its size and complexity. It might be useful to consider how to make NISO STS easier to implement for use cases such as these.

4. Metadata standards should support creative applications

Modern metadata standards are at their best when they promote computational and other derivative uses of the data and documents encoded in them. The definition of NISO STS in three formats -- DTD, W3C XML Schema, and RelaxNG -- increases flexibility in implementation, allowing integration into multiple types of technical platforms. This formal document structure easily promotes more advanced applications such as text mining. While definition of modern standards in RDF-friendly technologies is desirable to participate in Linked Data communities, we recognize this is difficult due to STS’s nature as a document markup language.

5. Metadata standards should have an active maintenance and governance community
NISO’s open, community-based, and formal governance and revision practices foster active engagement with standards such as this one, and this approach is commendable. The nature of STS as a new standard formalizing long-held practices among a stable implementation community speaks well to its ability to remain relevant over time.

6. Metadata standards should be extensible, embeddable, and interoperable

The incorporation of MathML into NISO STS is an advantage, as it does not re-implement the features of MathML in a new standard, but rather relies on a specialist community to maintain in its area of expertise. We note that NISO STS does make its own definitions of metadata features that other established communities have standards for, including bibliographic information, text formatting, geographic places, rights information, and names.

NISO STS is a very large and complex standard. Presumably implementation of such a standard requires significant expertise and resources. Providing both the Interchange and Existing Tag Sets, while adding to the standard’s flexibility, also increases its complexity. While we note that the “Scope” section in the STS draft for comment indicates that the tag set may be restricted to meet the needs of a given project, no formal mechanism is given for doing this, and as such does not meet our definition of “extensible”. The committee might consider an official “lite” version that is more accessible, to promote wider adoption of the standard.

It appears that some elements have recommended vocabulary and abbreviations (e.g. http://niso-sts.org/TagLibrary/niso-sts-TL-0-2-html/element/doc-type.html) that do not correspond to external standards, but are rather embedded in the DTD/Schema for STS. It may be that defining external vocabularies at this level is impossible for this specific standard but could be a challenge to future stability of the standard. There are multiple ways in which a standard community could engage with this issue. By further defining the vocabulary this specific standard could make contributions that other standards might adopt.

7. Metadata standards should follow the rules of “graceful degradation” and “progressive enhancement”

As a markup language for standards documents, rather than only a metadata format, graceful degradation of the STS standard is difficult. Some features, such as <std-doc-meta> do make it conceivable that some data might be extracted (here, a bibliographic record for the standard), but the committee might consider what other use cases there are for automatically extracting data from STS and ensure the standard is structured in such a way to support this.

The principle of progressive enhancement relies upon a design whereby a format starts at a
relatively simple state and then allows complexity to be added as needed. The design of STS as a pair (Interchange and Extended) of DTDs, W3C XML Schemas, and RelaxNG schemas containing all allowed elements and attributes does not embody this principle. Other implementation strategies, even with these technologies, could be used to make the vocabulary more modular, such that implementers could choose only features in categories useful to them. See, for example, customization options for the TEI as one method by which this might be done.

8. Metadata standards should be documented

In addition to the primary alphabetical listing of STS elements and attributes, at http://niso-sts.org, there is more helpful documentation available about the standard. The hierarchical view of the schema is especially helpful in enabling those new to the standard to more easily learn it. This page also offers helpful examples of elements and attributes used in context. Documentation grouping elements by theme would have benefited us as readers to be able to more quickly understand the features that the standard covers. In addition, full encoded examples would be helpful.

9. Metadata standards should be inclusive and transparent about historical and cultural biases

The <glyph-data> and <glyph-ref> elements provide useful means for expanding the content of an STS encoded document beyond Unicode characters.

We note that the content model for <name>, with its sub-elements for <surname> and <given-names>, may not work for all cultural traditions. <surname> is particularly problematic as a “family” name is not always a “surname”, and even then a “family” name is not a universal construct.

Similarly, the content model for <address> seems to assume Western-style addresses.

The standard includes a few elements that begin with the word “trans-”. We suggest either using the full term “translated” or a different abbreviation, to avoid using a word that carries with it other cultural implications in the English language, as a way of showing sensitivity to the LGBTQ community.

Other issues

@xlink:href is listed as an available attribute on <media> in the PDF but not the tag library http://niso-sts.org/TagLibrary/niso sts-TL-0-2-html/attribute/xlink href.html