NISO STS (Standards Tag Suite)
Technical Working Group Minutes
for STS Draft Version 0.1

For NISO STS Technical Working Group
January 2017

January 11, 2016, 10:00 am - 11:00 pm EST
1 Introduction............................................................................................................................................................1
   1.1 Attendees for January 11, 2017........................................................................................................................1
   1.2 Administrative Business.......................................................................................................................................1
   1.3 NISO STS Technical Working Group Call........................................................................................................1
2 Action Items..................................................................................................................................................................2
3 Comments Requiring Technical Recommendations..............................................................................................2
   3.1 Resolved: ASTM-15: <elocation-id>..................................................................................................................3
          3.1.1 Resolved: ASTM-15-a: Example for <elocation-id>..................................................................................3
          3.1.2 Resolved: ASTM-15-b: URI inside <elocation-id>......................................................................................3
   3.2 Resolved: ASTM-26: Expanding where <proj-id> may be used.........................................................................3
   3.3 Resolved: #00667: Footnote type values...........................................................................................................3
   3.4 Resolved: #00669/#00685: Citation to a Standard (<std>) and its model.........................................................3
          3.4.1 Resolved: #00669-a: What level to place <std>............................................................................................3
          3.4.2 Resolved: #00669-b: External linking elements in <std>..............................................................................4
   3.5 Resolved: #00668/#00685/ASTM-19: <glyph-data>/<private-char>.................................................................4
          3.5.1 Resolved: #00668, ASTM-19: <glyph-data>/<private-char> as rare.........................................................4
          3.5.2 Resolved: #00685-a: <private-char> inside <std>..................................................................................4
          3.5.3 Resolved: #00685-b: <private-char> everywhere text is allowed.........................................................4
          3.5.4 Resolved: #00685-c: Model of <std> versus <mixed-citation>/<element-citation>............................4
   3.6 Resolved: From #00685: Amend model of <std-ref>........................................................................................4
   3.7 Unfinished: #00682: <adoption> with no interior document ........................................................................5
          3.7.1 Resolved: #00682-a: <adoption> with no interior document.................................................................5
          3.7.2 Unfinished: #00682-b: Is <std-xref> as shown above adequate.............................................................5
   3.8 Resolved: #00689: Rationalize Table of Contents models ..............................................................................6
   3.9 Tabled: #00690: Keywords versus subjects ..................................................................................................6
   3.10 Resolved: #00694: Grouping lists ..................................................................................................................9
          3.10.1 Denied: #00694-a Structure to group <ref-list>s....................................................................................9
          3.10.2 Denied: #00694-b Construct paragraph-interrupted <list>s and <def-list>s.................................9
1 Introduction

These are the minutes for the Technical Working Group meeting for the NISO activity to create a standard tag set for Standards. Details on this work item are available to the public at:

http://www.niso.org/workrooms/sts/

These minutes record decisions made and action items assigned during the NISO STS Technical Working Group call on January 11, 2017.

1.1 Attendees for January 11, 2017

- Dreyfuss, Bob (ASTM: observer)
- Galichet, Laurent (ISO)
- Gilson, Howard (ASTM)
- Gooskens, Frans (NEN)
- Hollowell, Bob (ASME)
- Imseke, Gerrit (le-tex Publishing Services)
- Juillerat, Serge (ISO)
- Lagace, Nettie (NISO)
- Lapeyre, Debbie (Mulberry Technologies)
- Markantonatos, Nikos (Atypon)
- Rawson, Ken (IEEE)
- Rosenblum, Bruce (Co-chair, Inera)
- Saari, Antti (SFS: invited expert)
- Salcedo, Ivan (BSI: invited expert)
- Usdin, Tommie (Mulberry Technologies)
- Wheeler, Robert (Co-chair, ASME)
- Winchell, David (XSB)

1.2 Administrative Business

- Bruce Rosenblum welcomed our invited experts, who may be called on during the meeting to explain or qualify their Comment/Requests.

1.3 NISO STS Technical Working Group Call

The NISO STS Technical Working Group will meet by conference call on Wednesday January 18, 2017, at 10:00 am US EST.

US and Canada Toll free number and Passcodes are available at

http://www.niso.org/apps/org/workgroup/sts-technical/

under “Upcoming events”.

Global numbers can be found via http://bit.ly/1KCHbsT

For security reasons, the Passcode will be required to join the conference.
2 Action Items

For All

• **Doodle Poll Vote.** Make sure that you have received the doodle pool for voting on the non-controversial items. Please respond no later than Friday January 13, so we can close out non-controversial items and place anything requiring discussion on the next agenda.

• **Subjects and Keywords.** If anyone has clear topic or subject examples, please share by email with the Working Group. If anyone has an in-house or standards-based definition of topics or subjects, please share that as well. If anyone feels up to producing a tightly worded definition of either keyword or subject, please share.

For Bob Dreyfuss

• Provide some examples of topics or classifications of standards, that might be appropriately tagged as subjects and not keywords.

For Mulberry Technologies

• Produce a list of elements that are marked as rare or ‘for JATS only’ in the current NISO STS Tag Library. Email this list to all members of the Technical Working Group and invited experts. The Technical Working Group will consider this list and make the final decisions concerning rarity and deprecation. Any member can veto a suggestion of “rare”. This may be an email vote.

For Ivan Salcedo

• Get someone from his organization to cover the January 18th meeting in his stead.

• BSI makes a distinction between keywords and subjects, but does not put them into the XML content at this time. Ivan will report by email the distinction as BSI sees it, and whether either subjects or keywords would be useful to BSI inside the XML standards document.

For Robert Wheeler

• Provide some examples of ASME topics or classifications of standards, that might be appropriately tagged as subjects and not keywords.

3 Comments Requiring Technical Recommendations

Many of the Comments on the NISO STS draft, as submitted through the NISO website comment form, were trivial changes or typos and could be fixed without discussion. Some Comments could be resolved by the editor (Mulberry) or by the editor in consultation with the original requestor. Other Comments require discussion by the NISO STS Technical Working Group plus invited experts. Those Comments requiring technical discussion are described in this section.

Comments have been combined, divided into multiple sub-comments, rearranged, and/or edited to facilitate discussion and voting. Background material and recommendations have been added to some Comments.
3.1 Resolved: ASTM-15: <elocation-id>

3.1.1 Resolved: ASTM-15-a: Example for <elocation-id>

Documentation Recommendation

- Edit the Remarks section of <elocation-id> to remove mention of DOI and external link.
- Remove any existing tagged examples and replace them with the second example from the JATS Publishing (Blue) <elocation-id> page.

Suggestion

The NISO STS Technical Working Group will suggest this change to the JATS Standing Committee. An <elocation-id> is not an external link, not a DOI; <elocation-id> is meant as an alternative to page numbers.

3.1.2 Resolved: ASTM-15-b: URI inside <elocation-id>

Recommendation

No change will be made to the NISO STS Tag Set or to the Tag Library, since all mentions of DOI and external link will be removed from the discussion of <elocation-id>.

3.2 Resolved: ASTM-26: Expanding where <proj-id> may be used

Recommendation

- Broaden the definition of <project-id> to include any organization’s internal project identifier (not just ISO and ISO-related standards organizations).
- Add <proj-id> to <std-meta> and leave it also in the three ISO-related metadata elements (<iso-meta>, <reg-meta>, <nat-meta>).
- Do not add <proj-id> to <std-doc-meta>.

3.3 Resolved: #00667: Footnote type values

Recommendation

- Make the @fn-type values types unrestricted (CDATA).

Documentation Recommendation

- Since the use of @fn-type in standards documents is currently uncommon, there will be no standards-specific suggested value list at this time.

3.4 Resolved: #00669/#00685: Citation to a Standard (<std>) and its model

3.4.1 Resolved: #00669-a: What level to place <std>

Documentation Recommendation

The element <std> as currently modeled in NISO STS is a parallel structure (peer) to <mixed-citation> and <element-citation> (directly inside <ref>). Since the model for <std> is
mixed content, it is reasonable to think of <std> as a specialized version of <mixed-content>, that has been designed to be specific to the tagging of citations to standards. So standards documents are cited using <std> and books, journals, etc. use the regular JATS citation elements.

The element <std> is only available inside <mixed-citation> and <element-citation> to be compatible with JATS and backward compatible with ISO STS, and this usage should be documented as deprecated in NISO STS.

3.4.2 Resolved: #00669-b: External linking elements in <std>

Recommendation

• Add the elements <ext-link> and <uri> to the content of <std>.

3.5 Resolved: #00668/#00685/ASTM-19: <glyph-data>/<private-char>

3.5.1 Resolved: #00668, ASTM-19: <glyph-data>/<private-char> as rare

Documentation Recommendation

Remove the prose in the Tag Library that indicates <glyph> and <private-char> are rare in standards documents.

3.5.2 Resolved: #00685-a: <private-char> inside <std>

Recommendation

Add the element <private-char> to the content of <std>.

3.5.3 Resolved: #00685-b: <private-char> everywhere text is allowed

Recommendation

<private-char> will be allowed anywhere that can contain any kind of formatted text, in the same way the emphasis elements can.

3.5.4 Resolved: #00685-c: Model of <std> versus <mixed-citation>/<element-citation>

Recommendation

• The element <private-char> will be allowed inside <std> [#00685-a].
• The linking elements <ext-link> and <uri> will be allowed inside in <std>.
• The element <break> is only useful for forcing linebreaks to replicate a look-and-feel and will not be allowed inside <std>.
• None of the other over 90 elements allowed inside <mixed-citation> and <element-citation> will be allowed inside <std>.

3.6 Resolved: From #00685: Amend model of <std-ref>

Recommendation

Add no additional elements to the content of <std-ref>, since the use of any elements inside <std-ref> is deprecated in the documentation. The element <std-ref> is intended to be a string.
Documentation Recommendation

In the documentation for `<std-ref>`, keep the best practice note:

**Best Practice Content:** Although many elements are allowed inside a `<std-ref>`, best practice is to make the content just text. For best practice, the elements (such as `<year>`) that are allowed inside `<std-ref>` should be tagged as part of the `<std-ident>` element, leaving `<std-ref>` as a simple string.

3.7 Unfinished: #00682: `<adoption>` with no interior document

3.7.1 Resolved: #00682-a: `<adoption>` with no interior document

**Recommendation**

Allow `<std-xref>` (in lieu of `<standard>` or `<adoption>`) as the content of an `<adoption>`, in order to hold a reference to an adopted standard like this:

```xml
<adoption>
  <adoption-front>...metadata, notes, sections ...</adoption-front>
  <std-xref type="adopts">
    <std-id-group>
      <std-id std-id-type="dated" std-id-link-type="urn" originator="ISO">urn:iso:std:iso:13849-1:ed-1:en</std-id>
    </std-id-group>
    <std-ref>ISO 13849-1:2006</std-ref>
  </std-xref>
  <back>...</back>
</adoption>
```

With an `<adoption>` document, the adopting organization may want to or be able to distribute the original standard’s XML.

**Documentation Recommendation**

- The documentation will point out that this allows a standards organization to reference a document, such as a PDF, inside an `<adoption>` rather than including the tagged XML.

3.7.2 Unfinished: #00682-b: Is `<std-xref>` as shown above adequate

This issue was raised by Robert Wheeler toward the end of the meeting and was not resolved. Is the current `<std-xref>` model with the current `<std-xref>` attributes adequate for the kind of inclusion/link shown in #00682-a? There are no generalized linking attributes or external link elements involved in `<std-ref>`, but they may not be necessary.

**Background**

- `<std-xref>` Cross-Reference to a Standard is described in the Tag Library as: “Used within one of the standards metadata elements to identify, and possibly provide a link to, another standard. This element is used to name a related standard, for example, a standard that has been revised, amended, or replaced by the current standard.”

- The element `<std-ref>` has a model of:
  The following, in order:
  - One or more of any of:
    - `<std-ident>` Standard Identification Block
    - `<std-ref>` Standard Reference Designation
    - `<std-id-group>` Standard ID Group
    - `<release-date>` Release-Date, zero or more
    - `<meta-date>` Other Metadata Date, zero or more
• The `<std-xref>` element `@type` attribute may take any values. However, the following values are considered to be best practice:

<table>
<thead>
<tr>
<th><code>@type</code> attribute</th>
<th>Cross-reference description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>adopted_from</code></td>
<td>The cross-reference is to a standard from which the current standard is derived.</td>
</tr>
<tr>
<td><code>amends</code></td>
<td>The cross-reference is to the standard which the current standard amends. Thus the current document is an amendment to this referenced document.</td>
</tr>
<tr>
<td><code>consolidates</code></td>
<td>The cross-reference is to the standards which the current standard consolidates.</td>
</tr>
<tr>
<td><code>corrects</code></td>
<td>The cross-reference is to the standard which the current standard corrects.</td>
</tr>
<tr>
<td><code>replaces</code></td>
<td>The cross-reference is to the standard which the current standard replaces.</td>
</tr>
<tr>
<td><code>revision_of</code></td>
<td>The cross-reference is to the standard the current standard revises.</td>
</tr>
<tr>
<td><code>supersedes</code></td>
<td>The cross-reference is to the standard the current standard supersedes.</td>
</tr>
</tbody>
</table>

3.8 Resolved: #00689: Rationalize Table of Contents models

 Recommendation

Mulberry will make a rational subset of the BITS ToC models, minimally removing the ability to have `<contrib>` on every section and removing elements such as `<statement>` and `<verse-group>` from the ToC.

3.9 Tabled: #00690: Keywords versus subjects

Tabled Discussion

The NISO STS Draft 0.1 includes both `<subj-group>`s and `<kwd-group>`s, but does not make a clear distinction between them, and why/when you would use one versus the other. Both can utilize controlled vocabularies.

The Working Group suspects that both subjects and keywords may exist, but we cannot decide this matter without better examples from which to build clear and unambiguous definitions.

Unless we can decide what `<subjects>` are, and why we need them, Mulberry recommends that we remove `<subj-group>` (and its contained elements `<subject>`, `<compound-subject>`, and `<compound-subject-part>`) from the NISO STS vocabulary. For the sake of clarity and interchange, we suggest that all such material be tagged as keywords.

First Bone of Contention: Description

• Keywords (`kwd-group`) *may be* specific terms or concepts found within the text of a standards document, or implied (broader, narrower) from terms found within this text. Although found in the text of the document, the source of the keywords could be a taxonomy.

• Subjects (`<subj-group>`) *may be* the overarching categories/classifications or themes of a standards document, as a higher-level descriptor of the content. (Some organizations call these ‘topics’ or ‘classifications’. ) In JATS, subjects are used to organize articles in a Table of Contents or similar. In standards documents, subjects may be used to organize standards into series. The source of the subjects could be a taxonomy.

Second Bone of Contention: Examples

The Tag Library also needs a few clarifying examples:

• The index terms provided by INSPEC from ASME and IEEE documents have been tagged as keywords:
• Keywords provided by Working Groups (while rare) are definitely keywords:

  <kwd-group kwd-group-type="committee"> (from ASTM)
  <kwd>general delivery requirement</kwd>
  <kwd>pressure containing parts</kwd>
  <kwd>pressure vessel steels</kwd>
  <kwd>steel plates</kwd>
  <kwd>steel plates for pressure vessel applications</kwd>
</kwd-group>

• There are other kinds of keywords:

  <kwd-group kwd-group-type="controlled"> (from ASTM)
  <kwd>Boilers And Pressure Vessels</kwd>
  <kwd>Plates</kwd>
  <kwd>Pressure Vessels</kwd>
  <kwd>Pressure-Containing Parts</kwd>
  <kwd>Steel Plates</kwd>
</kwd-group>

• ASME “topics” are probably be subjects. May we see an example?

• The Tag Library says that IPC codes are keywords. Is it correct?

  <kwd-group kwd-group-type="ipc codes">
  <compound-kwd>
    <compound-kwd-part content-type="code">B82B1/00</compound-kwd-part>
    <compound-kwd-part content-type="value">Nano structures</compound-kwd-part>
  </compound-kwd>
  <compound-kwd>
    <compound-kwd-part content-type="code">H01L21/02</compound-kwd-part>
    <compound-kwd-part content-type="value">Manufacture or treatment of semiconductor devices or of parts thereof</compound-kwd-part>
  </compound-kwd>
</kwd-group>
• We have assumed that UNSPSC codes and UNS codes are keywords. Is this correct?

<kwdbroup kwdbroup-type="UNSPSC"> (from ASTM)
<compound-kwd>
  <compound-kwd-part content-type="code">30102204</compound-kwd-part>
  <compound-kwd-part content-type="value">Steel Plate</compound-kwd-part>
</compound-kwd>

Third Bone of Contention: ICS Codes (<ics>)

To further complicate matters, the <ics> element is in the NISO STS (inherited from ISO STS) to contain ICS Codes, which might otherwise be considered subjects or keywords.

• The <ics> element holds a code that identifies the subject matter of the standard using code values from the International Classification for Standards. When there is more than one ICS code for a standards document, the element <ics> appears more than once. (Typical ICS codes consist of a sequence of decimal digits and points [full stops]; they may denote fields (two decimal digits), groups (two digits, a point, and three digits), or a sub-group (two digits, a point, three digits, a point, and two digits).

• The issue: ICS Codes are currently modeled as #PCDATA (<ics>13.110</ics>), which cannot hold both a code and its expansion except as text:

<ics>23.020.30 Pressure Vessels, Gas Cylinders</ics>

• These code/values pairs could be modeled explicitly with parts if ICS Codes were tagged as either keywords or subjects. Compound parts for ICS codes have been requested.

<kwdbroup kwdbroup-type="ICS"> (from ASTM)
<compound-kwd>
  <compound-kwd-part content-type="code">23.020.30</compound-kwd-part>
  <compound-kwd-part content-type="value">Pressure Vessels, Gas Cylinders</compound-kwd-part>
</compound-kwd>
<compound-kwd>
  <compound-kwd-part content-type="code">77.140.50</compound-kwd-part>
  <compound-kwd-part content-type="value">Flat Steel Products And Semi-Products</compound-kwd-part>
</compound-kwd>

• Could these ICS codes have been tagged as subjects? Would that be incorrect? Is there a situation in which the ICS codes for a document would be different from ICS codes that are in the subjects of a document?

<ics>23.020.30</ics>

<subj-group originator="Mulberry" subj-group-type="ICS">
<compound-subject>
  <compound-subject-part content-type="code">23.020.30</compound-subject-part>
  <compound-subject-part content-type="text">Pressure Vessels, Gas Cylinders</compound-subject-part>
</compound-subject>
<compound-subject>
  <compound-subject-part content-type="code">77.140.50</compound-subject-part>
  <compound-subject-part content-type="text">Flat Steel Products And Semi-Products</compound-subject-part>
</compound-subject>
</subj-group>
3.10 Resolved: #00694 : Grouping lists

3.10.1 Denied: #00694-a Structure to group <ref-list>s

*Recommendation*

The Working Group decided that the Reference List example provided to support this request could be tagged either as a section (<sec sec-type="bibl">) with three titled <ref-list>s inside it or as a single <ref-list> containing recursive <ref-list>s. With two ways to handle this in the current Tag Set, there seemed to be an insufficient use case for creating a new grouping structure to handle this situation.

3.10.2 Denied: #00694-b Construct paragraph-interrupted <list>s and <def-list>s

*Recommendation*

Although the solution offered in the request was perhaps more elegant than that in the existing NISO STS, there is a solution (which has been used with standards documents) which ISO STS inherited from JATS. Lists separated by paragraphs are called “discontinuous lists”. There is an attribute on <list> and <def-list> called @continued-from (list continuation), which holds the ID of a list which the current list continues. This attribute is used to connect discontiguous lists that constitute one logical list. List item elements (<list-item>) hold their own <label> elements, so they can begin with “8”, or “42”, or whatever continuation is needed.