NISO STS – An Update

JATS-Con 2017
26 April 2017

Bruce Rosenblum, CEO, Inera, Inc.
Robert Wheeler, Director of Publishing Technologies, ASME
[co-chairs of NISO working group on Standard Tag Suite (STS)]
Lesley West, Director of Product Development, ASTM International
NISO STS and the
Big Hairy Audacious Grail
(NISO STS - An Update)
(with apologies to Monty Python)
Aka...

• A “standard for Standards”
• What has happened since we were last together? Much. Therein lies today’s presentation/update from the NISO STS working group.
• Plus, The True Story of an XML-Based Standards Package.
Last Year at JATS-Con...

• NISO STS Project Overview and Update
  – History of ISO STS
  – ISO STS → NISO STS
  – Project Overview
  – Working group structure
  – Key decisions to-date
Monthly steering group calls
Monthly technical group calls
Sub-group calls as necessary
Committee draft summer 2016
At vote no later than April 2017

(... Shhhhh, we’re not quite there yet.)
Today’s Question

• Question: How many publishing professionals does it take to make a standard?

• Answer: The more there are, the longer it takes

• Corollary: But, if there are not enough, it becomes a custom implementation, not a useable standard.
Quick Review

• NLM DTD begat JATS (2008)
  – NLM project moved to NISO
  – ANSI/NISO Z39.96

• JATS begat ISO STS (2011)
  – Designed for ISO XML requirements
  – Adopted by National Standards Bodies (NSBs)

• ISO STS begat NISO STS (2015)
  – ANSI/NISO Z39.102
Why JATS Foundation for STS?

- Standards & journal articles share structures
  - Sections
  - Tables
  - Figures
  - Equations
  - Bibliography
- Well-developed and well-tested model
- Easily modifiable for Standards use
- Strong third party support
Based on JATS version 0.4

Added:
- New top-level element <standard>
- Standards-specific metadata elements
  - iso-meta, reg-meta, nat-meta
- TBX (ISO 30042:2008) terms and definitions model
- Elements to cite other standards

Posted DTD, Documentation
- http://www.iso.org/schema/isosts/
Why Create NISO STS?

• ISO STS has been:
  – Successful for ISO and NSBs
  – Too limited for other standards bodies

• NISO STS will provide:
  – A stable standard for most standards publishers
  – Guidance to tool and conversion vendors
  – A common format for sharing
    • Metadata
    • Full text
  – Common XML model across pubs (standards, journals)

• Lower barrier to entry for XML publication
NISO STS Goals

- Expand for SDO and other use beyond NSBs
- Align with JATS 1.1 Blue and future versions
- Support additional structures
  - Tables of Contents
  - Indexes
  - CALS tables
- Maintain backwards compatibility for existing users of ISO STS
Why Set Suite?

• Separate models will be available with MathML2 and MathML3 for compatibility
• Separate models will be available with XHTML tables, and with XHTML and CALS tables
  — “Interchange” — XHTML-only
  — “Extended” — XHTML and CALS
• Four models total:

<table>
<thead>
<tr>
<th>XHTML+MathML2</th>
<th>XHTML+CALS+MathML2</th>
</tr>
</thead>
<tbody>
<tr>
<td>XHTML+MathML3</td>
<td>XHTML+CALS+MathML3</td>
</tr>
</tbody>
</table>
NISO STS Development Plan

• Multi-phase project
• Phase 1
  – “Modest” extensions for normative SDO standards
• Future phases may address
  – Non-normative documents
    • consider BITS
  – Dual-language standards
  – Historical standards
Timeline Since April 2016

- 6 months of Technical WG calls
- 1 month WG comment period
  - Boy, did we ever get comments
  - > 80 comments
- 3 more months of Technical WG calls
  (plus a Doodle Poll for “non-controversial” trivial changes resolvable without discussion)
Key Extensions Since April 2016

• New metadata models
• Alternate terms and definitions model
• DOI support
• Adoptions Model
New Metadata Models

Modeling Goals

- Make as clean as possible
- Keep backwards compatible with ISO STS
- Enable encoding of all of the information identified in the STS Metadata Subcommittee Requirements Report (final draft version: 2016-05-27)
- Do not require redundant information in XML documents
New Metadata Models (2)

Proposed Approach

• Make the metadata model enabling (not enforcing)
• Make virtually everything in the metadata at all levels optional and much of it repeatable
• Create two new structures in <front>
  – Contain metadata that is shared by all of the organizations described in the front matter
  – Describe any organizations that do not fit neatly into the iso, nat, or reg metadata
New Metadata Aspects (1)

- More granular tagging of and more title levels
- Support multiple co-producing standards organizations
- Grouping of IDs (multiple designators, like dated and undated) and coupling (for example designator and DOI or series and DOI)
- Focus use on STS-specific release-date rather than the JATS <pub-date> (which created compatibility issues)
- New attributes on release-date to support standards-specific lifecycles events
- Add regional metadata (<reg-meta>) elements (<release-version-id> and <wi-number>) to std-meta to ease use of <std-meta> in <adoption> and/or eventual migration to <std-meta>
New Metadata Aspects (2)

Benefit/borrow from JATS
- Add JATS <isbn>, <issn>, <issn-l>
- Add JATS <abstract>
- Add JATS <kwd-group>
- Add JATS <subj-group>
- Add JATS <counts>
- Add JATS <self-uri>
- Add JATS <ali:free_to_read> and <ali:license_ref>
Alternate Terms and Definitions Model

- New model, as an alternative or supplement to TBX:
  - Looser and non-enforcing
  - Allows recording term semantics in display order
  - Based on existing <term-display> element with additions
- TBX is a concept-oriented encoding of terminological data
- <term-display> uses natural language to describe terms, and may OR MAY NOT incorporate semantic term element. Semantic tagging is encouraged, but not enforced
Adoptions Model

<adoption>
  <std-meta/>
  <!-- national-specific forward, intro, normative references here -->
  <adoption>
    <std-meta/>
    <!-- regional-specific forward, intro, normative references here -->
    <standard>
      <!-- international standard -->
      <iso-meta/>
      <body/>
      <back/>
    </standard>
    <adoption-back> <!-- CEN-specific ANNEXES here --></adoption-back>
  </adoption>
  <adoption-back> <!-- NSB-specific ANNEXES here --></adoption-back>
</adoption>
DOI Support

- Metadata
- In-text citations
- Normative reference list and bibliography

<std>A17.1/CSA B44
  <std-id-group std-relationship-type="std-as-published">10.1115/ASME A17.1/CSA B44</std-id-group></std>
JATS/BITS/STS Alignment

• JATS & BITS: informally aligned
• JATS & STS: 4 years without alignment
  – <version> resolved
  – <pub-date> complex resolution
  – <std> complex resolution
  – <title-group> BITS “fixed” for 2.0

• Future challenge: JATS/BITS/STS Alignment
  – Messages between groups
  – Working group member overlap
Resolving Complex Conflicts

• `<pub-date>`
  – `<pub-date>`
    • element-only content in JATS
    • character data (#PCDATA) content in ISO STS
  – `<pub-date>` redefined as the original publication date of the standard
  – Leave `<pub-date>` in place for backwards compatibility only
  – Recommend `<release-date>`, not `<pub-date>`, for publication dates
  – Recommend `<meta-date>` to describe all other life-cycle dates (e.g. votes)

• `<std>`
  – Acknowledgement that definition differs between JATS/BITS and STS because the need in STS is fundamentally different
WG Coordination: Term Attributes

• Close JATS/STS WG coordination
• STS WG proposed new <term> attributes
  – @term-status (‘preferred’, ‘deprecated’, ‘allowed’)
  – @term-type (‘variant’, ‘borrowed’. abbreviated’)
• Proposal submitted to JATS for 1.2
• “The JATS Standing Committee requests that the
definition, suggested values, and examples of the
@term-status and @term-type elements be
coordinated with, and probably taken from the
NISO STS documentation”
While discussing CRediT, JATS Standing Committee devised four new attributes to connect to a vocabulary to a taxonomy:

- @vocab: vocabulary name (taxonomy, ontology, database)
- @vocab-identifier: pointer to vocabulary (URI or DOI)
- @vocab-term: canonical form of free form prose content
- @vocab-term-identifier: (likely a URI) pointer (likely URL directly to term in the vocabulary)

Proposed for <kwd-group> & <subj-group> in JATS 1.2
Passed to STS WG for discussion and incorporated into STS

Very useful addition to both JATS and STS
Excellent example of cross-model synergy!
Some Interesting Standards-Specific Additions

- Forms often floating objects in standards
  - Form Attributes
    - is-form: yes/no
    - form-type: blank, filled-in-example, placeholder-explanations, unspecified
  - Added for tables, figures, boxes, preformat

- Unique SDO elements
  - <accrediting-organization>, <authorization>
Project Challenges

- “ACE” Documents
- Backwards compatibility
- “Reading order” preferences
- “Rare” elements
- Unique standards elements
- Keywords versus Subjects
- People *really* cared
Amendments, Corrigenda, and Errata (“ACE” documents) are typically stand-alone documents that contain additions, changes, and error corrections to a standard.

- Steering Committee defined three possible styles of ACE documents: Structural Style, Format Style, and Embedded Style.
- Steering Committee decided, following a review of our timeline to that point and projecting out when we could get to a version 1, and considering that the only current use cases were “format style,” that “Format Style” was what would be considered in scope for Phase I NISO STS.
- We defined a new element `<editing-instructions>` to hold the instructions (not part of the standards document) such as “Replace Section 2.3 with the section below.”
Backwards Compatibility

• Some things we wanted to do, but could not to retain backwards compatibility
  – Remove ISO STS elements we “think” no one uses
  – TBX – did not add <p> in <tbx:definition>
  – Single definition of <std> in STS and JATS
  – Single definition of <pub-date> in STS and JATS
  – Change <content-language> to @content-language
  – Deprecate iso-meta, reg-meta, nat-meta for std-meta
“Reading Order” Preferences

• Traditionalist XML view: reading order is the domain of a transform or style sheet
• But even JATS has added “string” elements
• Some WG members desired more support for “reading order” tagging
• NISO STS (cf. ISO STS) is looser to support reading order
• New element allow rendering order
  – <term-display>
  – <adoption> model
“Rare” Elements

- Many JATS elements appeared to have no obvious applicability in standards
- “This element was inherited from JATS and may be very rare in standards.”
- Examples:
  - <glyph-data>
  - <private-char>
  - <access-date>
  - <author-comment>
  - <state>
  - <media>
- Technical Working Group was unable to agree which elements might be “rare” in standards documents
- All such comments removed, except from deprecated elements.
Keywords Versus Subjects

Extensive discussion: keywords versus subjects

• Keywords will be defined as specific terms or concepts found within the text of a standards document, or implied (broader, narrower, etc.) from terms found within the text. Even though keywords may be found in the text of the standards document, the source of the keywords could be a taxonomy, a thesaurus (https://www.iso.org/obp/ui/#iso:std:iso:25964:-1:ed-1:v1:en:term:2.62), a controlled vocabulary (https://www.iso.org/obp/ui/#iso:std:iso:25964:-1:ed-1:v1:en:term:2.12), or an uncontrolled vocabulary.

• Subjects will be defined as overarching categories, classifications, topics, or themes of a standards document, as a higher-level descriptor of the content. In JATS, subjects are used to organize articles in a Table of Contents or similar. In standards documents, subjects have been used to organize standards into series. The source of the subjects can be a classification scheme (https://www.iso.org/obp/ui/#iso:std:iso:25964:-1:ed-1:v1:en:term:2.6), such as a taxonomy.
Unique Standards Elements

• The saga of <ics>
• Initial consensus: just like <subj-group>
  – Best “theoretical” way to tag ICS codes
• Post-call discomfort; the decision:
  – Overlooked easy and consistent interchange
  – Was not consistent with existing DTDs for standards
  – Often appear in print (uncommon for “subjects”)
• Working group revisited and reversed
• Added <isc-wrap> and <isc-desc>
People *Really* Cared

- Huge interest in working group
- Great attendance at call... after call... after call
- Almost Talmudic wrestling over some issues
- Invited experts in comment resolution
People *Really* Cared (1)

Were there any challenging aspect(s) of the project to date that took you by surprise?

- “How technical the steering group was.”
- “The complexity of the metadata aspects associated with the development of the NISO STS schema took us by surprise. It was only when delving more deeply into discussions with other participants in the NISO standardization process that we realized how diverse were the views and requirements other organizations have with regard to standards metadata and how these requirements/views determine the way they use or intend to use the schema.”
- “When I first read the proposed discussion items for NISO STS, I didn’t expect that more than 5 or 6 calls would be needed.”
People Really Cared (2)

Were there any pleasant surprises?

• “I did not think we would necessarily reach a conclusion given the various organizations and standardization processes represented. But, we did and that speaks volumes to all those who participated.”

• “We think it is remarkable that most of the discussions were characterized not so much by trying to bring home one’s own favoured solution for particular problems, but by a wish to understand what the needs of the others were and trying to find solutions acceptable to all.”
What do you feel will be the single most beneficial aspect of the STS project when completed?

• “We need one schema to be used by all standards bodies so we can provide our standards users with consistent content and functionality. That makes me very enthusiastic about this project, I think we are finally getting there.”

• “Well that’s easy: We will all have a standard now that describes how to structure standards content using XML.... A great step forward, too, in making XML standards content attractive for customers.”

• “Customers rarely if ever use just standards from one SDO. This interoperable standard would allow organizations to build a consistent experience for users thus reducing the cost and friction of using Standards for customers.”
Current NISO STS Timeline

- **April 24 to May 24:** Public comment period
- **May 25 to July 4:** Respond to public review/comments and finalize/revise materials
- **July 5 to July 17:** STS Steering and Technical Group Vote
- **July 18 to July 26:** NISO Topic Committee vote
- **July 27:** Submission of BSR-8 to ANSI for inclusion in Standards Action
- **July 27 to August 10:** Notification of ballot and 15-day period to close NISO Member Voting Pool
- **August 11:** Standards Action publication including NISO BSR-8 notice (45-day review period starts)
- **August 11 to September 9:** NISO Member Vote
- **September 9 to September 17:** Respond to any "no" votes from NISO Voting Members
- **September 25:** ANSI 45-day review period ends
- **September 26:** NISO Submission to ANSI (ANSI Approval: 2-4 weeks)
- **October 2017:** Target formal, published standard
Follow NISO STS Developments

• http://www.niso.org/workrooms/sts/

NISO STS Working Group

Work Description

This work will standardize a specific tag set used for standards publishing, and link it officially to JATS (ANSI/NISO Z39.96-2012 JATS: Journal Article Tag Suite), a widely used specification which defines a set of XML elements and attributes for tagging journal articles and describes several article models.

Background

There are currently several DTDs used for tagging standard-type information based on JATS and a number of others that have been independently developed. This variety of DTDs used in standards publishing makes interoperability between organizations difficult and increases any integration costs.

At the end of 2011, the International Organization for Standardization (ISO) revamped its publishing systems and together with Mulberry Technologies, Inc. developed a derivative of JATS to be used for ISO standards publishing: the ISOSTS (ISO Standard Tag Set). This DTD has been in full production since, with little or no changes. The DTD and documentation are openly available. A number of ISO’s members (BSI (British Standards Institution), SIS (Swedish Standards Institute), NBN (Netherlands Standardization Institute), SA (Standards Australia)) and some distributors have adopted the ISOSTS.

Several standards development organizations (SDOs) and distributors in the United States are looking to upgrade their publishing systems. Some are already familiar with JATS and have looked at ISOSTS. However, there is reluctance to adopt ISOSTS as it is not currently an official standard. Concern also exists that if JATS is updated, its updates may not filter into ISOSTS.

It would be beneficial to all stakeholders to move ISOSTS toward standardization and create an official relationship with JATS. The two standards will then remain coupled and the ISOSTS solution will be adopted more readily. This move will result in greater interoperability of standards, which in turn will aid our end users, and improve the future of standards publishing.
Non-Normative Materials

- [www.niso-sts.org](http://www.niso-sts.org)
- All models in interchange and extended suites
  - DTDs, RelaxNG, XSD
  - MathML 2, MathML 3
- Tag library documentation *and* examples
- Coming soon: STS standard as XML document tagged in STS
  - Going meta where no one has gone before
One Last Story, Discovery

- Helping subscribers find paid content via Discovery Services.
- ASTM overhauled how to feed content to DSVs.
- Push vs. Pull
- Aggregation of data via a single source vendor.
- Reportedly, “everyone” wanted _______________.

Conclusions

• NISO STS: longer-than-expected gestation
• Excellent working group input and consensus
• Expected ANSI standard: October 2017
• Benefits of using a standard XML model
  – Production efficiencies
  – New product opportunities
  – Easier interchange with development and distribution partners
Contact Us

• Bruce Rosenblum
  – brosenblum@inera.com

• Robert Wheeler
  – wheelerr@asme.org

• Lesley West
  – lwest@astm.org