

What's Happening in JATS, BITS, and STS in 2022

Journal Article Tag Suite
Book Interchange Tag Suite
Standards Tag Suite

Debbie Lapeyre

Mulberry Technologies, Inc.
1714 Tweed Street
Rockville, MD 20851-1425
dalapeyre@mulberrytech.com



What's Happening in JATS, BITS, and STS in 2022

About This Talk and Slides

STATUS UPDATE: JATS (last year)

STATUS UPDATE: JATS 2022

Changes Already Recommended for JATS 1.4

Current Discussions for JATS 1.4

In Draft: Multilingual Publishing

Ref: How Multi-lingual Metadata Works

Ref: How Multi-lingual Texts are Structured

Ref: How Multi-lingual Mechanism Works

Ref: Optional Multi-lingual Attributes Describe

Ref: Repeatable element `<content-language>`

Discuss: Change to Collaboration (`<collab>`)

Discuss: Citing Standards (inside mixed and element citations)

Discuss: External and Supplemental Objects

Already Recommended Attributes for JATS 1.4

Aside: About Custom Values for Set Value Lists

Ref: More Attributes for Discussion

STATUS UPDATE: NISO STS (Standards Tag Suite)

Aside: Next NISO STS will be Version 1.2

Last year I said: NISO STS 1.2 was *Nearing Completion*

NISO STS Content Status

NISO STS 'Standard' Status

After ANSI Approval

Ref: What's New in NISO STS 1.2?

Ref: No Major TBX/`<term-display>` Changes

Ref: Not Included: Marking Up Requirements in Standards

Ref: New and Cool in STS Elements

Ref: Legend added to Figures, Tables, Equations, etc.

Ref: New Metadata *About the File* (`<processing-meta>`)

Ref: Reference List: Citing Episodic Material

Ref: Other Element Modifications

Ref: What's New and Cool in Attributes

Ref: Custom Values for Set Value Lists

Ref: Processing Metadata Modeling Attributes

Ref: `@dtd-version` Attribute Changes

Ref: New `@hreflang` Attribute

Ref: New `@style-detail` Attribute

STATUS UPDATE: BITS (JATS for Books)

BITS 2.1 New and Cool for Elements (overview)

Legend added to Figures, Tables, Equations, etc.

BITS-only Element `<content-version>`

Citing Episodic Material

New Metadata About the File (<processing-meta>)

Processing Metadata Modeling Choices Attributes

Ref: Other Element Modifications

Ref: Other BITS Modified Elements

BITS 2.1 New and Cool for Attributes (overview)

@dtd-version Attribute Changes

New @hreflang Attribute

Ref: New and Changed BITS Metadata Attributes

Make Your voices Heard!

About This Talk and Slides

- I have 20 minutes plus questions
- I could talk about JATS or BITS or STS for a week
- This talk asks for “and” not “or”
- Some of you are interested in the techie details (some are not)
- These slides will be the XUG website for a time

Therefore, warning: there are detailed slides I don't plan to talk. These are for your reference.

STATUS UPDATE: JATS (last year)

- Journal Article Tag Suite (1.3) ANSI/NISO Z39.96-2021
 - Approved: 6/10/2021 (its model date/year)
 - Standards Published: 6/18/2021
 - New DTDs, RNGs, XSDs
 - Redesigned Tag Library with better documentation

STATUS UPDATE: JATS 2022

JATS Standing Committee working on

- JATS 1.4 (next regular release)
- JATS 2.0 (non-backwards-compatible version)
- Mined JATS 2.0 Recommendations
 - for things that *could* go into JATS 1.4
 - Discussing those suggestions now for recommending or refusing

Changes Already Recommended for JATS 1.4

- Make `<permissions>` repeatable in more places, so copyrights and years can be grouped
- Change MathML 3 modules to latest version (2012-04-10)
- Change ALI (Access & License Indicators) to latest version (2021)

Current Discussions for JATS 1.4

- More detail for affiliations and funding:
 - `<institution-department>`
(subunits of an institution)
- Let `<anonymous>` (currently empty) contain content in Publishing (for anonymous Peer review)
- Add `<legend>` from BITS

In Draft: Multilingual Publishing

- Tag documents where
 - entire document is in 2 or more languages
 - substantial portions of content in 2 or more languages
- The goal is
 - enable true multi-lingual documents without requiring changes to monolingual documents
 - read/display only one language, or all of them
 - search in only one language, or all

Ref: How Multi-lingual Metadata Works

- Metadata elements allowed to repeat, differentiated by `@xml:lang`
(article titles, abstracts, funding statements, copyright statements, etc.)
 - `<title-group>`, `<issue-title-group>`, `<journal-title-group>`
 - `<author-notes>`
 - `<permissions>`
- Also: All `<trans-xxx>` elements deprecated

Ref: How Multi-lingual Texts are Structured

- Same text structure repeats in several languages
(sections, paragraphs, figures, etc.)
- Structures need not be co-located
- Single language need not be contiguous
- Therefore, no wrapper element would work

Example: alternate sections repeat the same content first in French and then in English

Ref: How Multi-lingual Mechanism Works

- Language variants of the same content called a “language group”
(the same paragraph in Romanian, Greek, and French)
- Language groups identified by `@lang-group` attribute
 - `@lang-group` holds an IDREF attribute
 - IDREF value is the `@id` of one of the group members
- `@xml:lang` put nearly everywhere

Ref: Optional Multi-lingual Attributes Describe

- `@lang-variant` — relationship between the language variants (original, translation, transliteration, etc.)
- `@lang-source` — where a given variant originated (author, editor, translator, machine-translation, etc.)
- `@lang-focus` (content marked as primary, secondary, etc.)
- `@lang-translate` (content to be translated?, yes or no)
- Attribute on `<processing-meta>` says this document uses language groups

Ref: Repeatable element `<content-language>`

- Borrowed from NISO STS
- Inside `<article-meta>`
- Each element names one primary language of the article
- Element content should be the two-letter ISO 639 code for the language (“en”, “de”, “es”)

(Not really useful for monolingual articles)

Discuss: Change to Collaboration (`<collab>`)

- `<collab>` was supposed to be the *name* of a collaboration
- But people put all sorts of junk into it
- Solution:
Deprecate current `<collab>` and `<collab-alternatives>`
- New elements for old function:
 - `<collab-name>` (name of the collaboration)
 - `<collab-wrap>` (to hold all the stuff now in `<collab>`)

Discuss: Citing Standards (inside mixed and element citations)

Current elements (`<std>` and `<std-organization>`) don't describe how standards are cited in the real world

Proposal:

- Deprecate `<std>` and `<std-organization>`
- Add `<std-ref>` (“Standard Designator”) to citations (from NISO STS)
 - the way standards are cited in standards
 - content just a string, encapsulating (untagged):
 - originating organization
 - standard number
 - year

Discuss: External and Supplemental Objects

- Mark supplemental figures, tables, sections
 - with a “supplemental” attribute
 - Put them directly into the article
- Deprecate `<supplemental-material>`
- Deprecate `<inline-supplemental-material>`

Already Recommended Attributes for JATS 1.4

- Add `@xml:lang` to `<permissions>`
- Add vocabulary attributes to `<see>` and `<see-also>`
- New “interview” suggested value for `@article-type`
- Add value “custom” to `@question-response-type` and `@custom-type` to `<question-response>`

Aside: About Custom Values for Set Value Lists

- When a schema gives an attribute a set list of values

In the DTD: `color (red | blue | green)`
In the XML: `<ribbon color="red">`

- No way to add a value, without changing the schema
- Now, all set value lists have one additional value (“custom”)

In the DTD: `color (red | blue | green | custom)`

- New `@custom-type` attribute holds additional value
- In the document (two attributes):

```
<ribbon color="custom" custom-type="purple">
```

Ref: More Attributes for Discussion

- Add lots more suggested values for `@article-type`
 - data paper, collection, community comment, protocol
 - expression of concern, Jupyter Notebook(?), many more
- Add vocabulary attributes to `<see>` and `<see-also>`
- Add `@sort-key` to `<index-term>`

STATUS UPDATE: NISO STS (Standards Tag Suite)

(JATS family for standards)

- Two flavors
 - Extended (XHTML tables + OASIS CALS tables)
 - Interchange (XHTML tables only)

Status: new NISO STS standard still *ALMOST* ready

Aside: Next NISO STS will be Version 1.2 (rather than NISO STS 1.1)

- Because ISO STS 1.1
 - is already a standard
 - was the last ISO update before NISO STS
- NISO STS 1.1 and ISO STS 1.1 sounded *too much* alike (confusing)

NISO STS 1.0 to be followed by NISO STS 1.2

Last year I said: NISO STS 1.2 was *Nearing Completion*

- Committee voted final changes October 2021
- Mulberry made DTDs, XSD schemas, RNGs, Tag Libraries
- Version 1.1d1 to STS Standing Committee for approval November 2021
- Mid-February 2022, STS SC approved Version 1.1d1
- Version 1.1d1 went out for public comment

And then my cloudy crystal ball said:
If STS 1.2 standardization process starts in February,
we should have a new NISO STS by June

NISO STS Content Status

Nothing has changed in the models or documentation since February 2022!

NISO STS 'Standard' Status

This year I say: NISO STS 1.2 is at Completion

- Survived public comment
- NISO Topic Committee approved it
- NISO members Voting Pool approved it
- NISO will request ANSI approval (mid-October 2022)
- ANSI *may* approve in October (TBD)
- Mulberry nearly done making new DTDs, RNGS, XSDs, Tag Libraries, and standard document

After ANSI Approval

- NISO STS 1.1d1 will become NISO STS 1.2
- DTDs, XSD schemas, RNGs, Tag Libraries changed to reflect new number, publication date
- Minor wording changes in the standard
- Then
 - NISO can publish the standard
 - Release DTDs, XSDs, RNG, Tag Libraries on non-normative site NISO-STS.org

Ref: What's New in NISO STS 1.2?

Slides are here, but I'm not going to talk them

Ref: No Major TBX/term-display Changes (tweaks)

- New element `<term-display-string>`
 - to add text to `<term-display>`
 - `<term-display>` has element-only content
- `<tbx:geographical>` usage now repeatable
- New `@term-type` value “translation”
- New suggested content `<tbx:grammaticalGender>` “masculine-feminine”

Ref: Not Included: Marking Up Requirements in Standards

(NISO STS 1.1 Will NOT Model Requirements)

- By design JATS-based standards (like STS) do not *lead* their user community
- They *reflect* industry best practice
- They rely on broad-level consensus

There is no industry-wide consensus on Requirements

Many ongoing initiatives (bespoke schema, RDF, database, link bases, ontologies, other)

Ref: New and Cool in STS Elements

- New element `<inline-code>`
- Keywords and subjects inside more elements
- `<std-ident>` has new optional component `<subpart-number>` for section numbers inside parts (with `<part-number>`)
- New `<legend>` element
- Redefined `<part-title>` and `<source>`
- New `<processing-meta>` (just like JATS)

Ref: Legend added to Figures, Tables, Equations, etc.

<legend> is

- A list or table of the explanations for symbols, characters, terms, or abbreviations
- Also known as
 - “Key” (“Légende” in French, “Legende” in German)
 - where list
 - variable list
 - symbol chart
- Legends accompany figures, tables, equations to act as a key to the symbols, characters, or abbreviations used in the graphic, figure, table, equation, etc.

Ref: New Metadata *About the File*

(<processing-meta>)

- At the beginning of the document
- NOT part of the content, the structure, or the publishing metadata for the document
- Metadata about the XML (at the “file level”)
- Content:
 - <extended-by> names tag set extensions (superset) the XML document claims to follow
 - <restricted-by> names tag set restrictions (subset, secondary schema, etc.) the XML document claims to follow
 - <custom-meta> elements describe other metadata, as name-value pairs
- Attributes describe the modeling choices made

Ref: Reference List: Citing Episodic Material

(Just like JATS for book chapters, podcasts, streaming)

Redefined citation elements `<source>` AND `<part-title>`

- Used to be defined in terms of a “document”
- Now defined in terms of a “resource”

`<part-title>` — Title of a portion of a larger resource (for example, a chapter in a book, module in a course, episode of a podcast) described in a bibliographic reference.

`<source>` — Title of the resource (for example, journal, book, podcast, conference proceedings, radio or TV show) that contains (is the source of) the material being cited in a bibliographic reference or product.

In the case of a resource that has titles at multiple hierarchical levels, this is the title of the largest titled resource (e.g., the TV show name and not the episode title).

* `<chapter-title>` is deprecated

Ref: Other Element Modifications

- Subjects and keywords allowed in more elements
figure, question, appendix group, equation, etc.
- New `<inline-media>`
- `<xref>` allowed in more places, used to point to the data *behind* a figure, media object
- `<disp-formula>` can take a `<captionn>`
- Table cells can contain speeches, quotes, statements, verse

Ref: What's New and Cool in Attributes

- Vocabulary attributes on more elements
- `@assigning-authority` on more elements
- `@stage-code` on `<release-version>` holds numeric ISO International Harmonized Stage Codes
- New `@list-type` value “arabic”
- `@dtd-version` no longer fixed
- New `@hreflang`
- New `@style-detail`

Ref: Custom Values for Set Value Lists

- When a schema gives an attribute a set list of values

```
In the DTD:    color    (red | blue | green)
In the XML:    <ribbon color="red">
```

- No way to add a value, without changing the schema
- Now, all set value lists have one additional value (“custom”)

```
In the DTD:    color    (red | blue | green | custom)
```

- New `@custom-type` attribute holds additional value
- In the document (two attributes):

```
<ribbon color="custom"    custom-type="purple">
```

Ref: Processing Metadata Modeling Attributes

@tagset-family

("sts")

@base-tagset

STS (e.g., "interchange" or "extended")

@table-model

("xhtml", "oasis", "both")

@mathml

(MathML 2.0 or MathML 3.0)

@math-representation

How many ways are mathematical expressions represented in this document?

(e.g., "latex images" or "mathml plaintext images")

@terminology-model

("tbx" or "term-display")

Ref: @dtd-version Attribute Changes

- Used to be a #FIXED attribute, set to the latest version number (1.2, 1.1d1, 2.0, whatever)
- Now a choice from a list
 - STS (iso-0.9 | iso-1.0 | iso-1.1 | 1.0 | 1.1d1 | 1.2)

Ref: New @hreflang Attribute

- Language of the *target* to which a link is pointing
 - not the language of the link
 - nor the document the link is inside
- Allowed everywhere that has external links

Related article in German:

```
<related-article id="ral"
  related-article-type="alt-language"
  hreflang="de"
  ext-link-type="doi"
  xlink:href="https://dx.doi.org/10.5414/ALP33164"/>
```

Ref: New @style-detail Attribute

- used with @list-type="bullet"
- sets the bullet character (for example “dashed”)

```
<list list-type="bullet"
  style-detail="dashed">
```

STATUS UPDATE: BITS (JATS for Books)

(Book Interchange Tag Suite)

- BITS Version 2.1 has been Published
- New DTDs, RNGS, XSDs
- Same redesigned Tag Library style as JATS

BITS 2.1 New and Cool for Elements (overview)

- `<legend>` [*requested for JATS*]
- `<content-version>`
- `<processing-meta>` (similar to JATS)
- Redefined existing `<part-title>` and `<source>`

Legend added to Figures, Tables, Equations, etc.

`<legend>` is

- A list or table of the explanations for symbols, characters, terms, or abbreviations
- Also known as
 - “Key” (“Légende” in French, “Legende” in German)
 - where list, variable list, symbol chart

Legends act as a key to the symbols, characters, or abbreviations used in a figure, table, graphic, equation, etc.

BITS-only Element <content-version>

Book “versions” and book “editions” are not the same thing

- <content-version> acts for books and book parts as <article-version> does for articles
- Added to both <book-meta> and <book-part-meta>
- <content-version-alternatives> element holds multiple <content-version>s for both books and book parts, differentiated by their context and their attributes.

Citing Episodic Material

(For book chapters, podcasts, streaming, just like JATS)

Simplified citation elements <source> and <part-title>

- Used to be defined in terms of a “document”
- Now defined in terms of a “resource”

<part-title> — Title of a portion of a larger resource (for example, a chapter in a book, module in a course, episode of a podcast) cited in a reference.

<source> — Title of the resource (for example, journal, book, podcast, conference proceedings) that contains (is the source of) the material cited in a reference.

* <chapter-title> now deprecated

New Metadata About the File (`<processing-meta>`)

- NOT part of the content, the structure, or the publishing metadata for the document
- Positioned at the beginning of the document
- Metadata about the XML (at the “file level”)
 - `<extended-by>` tag set extensions (superset) the XML document claims to follow
 - `<restricted-by>` tag set restrictions (Schematron, secondary schema, etc.) the XML document claims to follow
 - `<custom-meta>` describes other metadata, as name-value pairs
- Attributes describe modeling choices

Processing Metadata Modeling Choices Attributes

`@tagset-family`

FIXED to value “bits”

`@table-model`

(“xhtml”, “oasis”, “both”)

`@mathml`

FIXED to value “3.0”

`@math-representation`

How many ways are mathematical expressions represented in this document? (e.g., “latex images” or “mathml plaintext images”)

Ref: Other Element Modifications

- Subjects and keywords allowed in more elements
 - figure, question
 - appendix group, equation, etc.
- New element<inline-media>
- <xref> allowed in more places, used to point to the data *behind* a figure, media object
- <disp-formula> can take a <caption>
- Table cells can contain speeches, quotes, statements, verse

Ref: Other BITS Modified Elements

- <pub-date> added to <collection-meta>
- <self-uri> added to most book-parts and block elements
 - <sec>, <boxed-text>, <app>, <glossary>,
 - <index>, <ref-list>, <toc>, <abstract>, <notes>, <question>, etc.

BITS 2.1 New and Cool for Attributes (overview)

- New attribute extension mechanism (custom)
- @dtd-version no longer fixed
- @hreflang (name the language of an external target)

@dtd-version Attribute Changes

- Used to be a #FIXED attribute, set to the latest version number (1.1, 2.0, whatever)
- Now a choice from a list of values (1.0 | 2.0 | 2.1)

New @hreflang Attribute

- Language of the *target* to which a link is pointing
 - not the language of the link
 - nor the document the link is inside
- Allowed everywhere that has external links

```
<related-article id="ra1"
  related-article-type="alt-language"
  hreflang="de"
  ext-link-type="doi"
  xlink:href="https://dx.doi.org/10.5414/ALP33164"/>
```

Ref: New and Changed BITS Metadata Attributes

- @assigning-authority given to
 - same identifiers as JATS (and subjects, keywords, and <custom-meta>)
 - but also to <book-part-id>, <book-volume-id>, <collection-id>, <content-version>
- @sort-key added to <index-term>

Make Your voices Heard!

- Technical details on non-normative supporting material:
 - All Tag Sets: <http://dtd.nlm.nih.gov/>
 - STS Tag Suite (all Tag Sets): <https://www.niso-sts.org/>
 - BITS Tag Suite (all Tag Sets) <https://jats.nlm.nih.gov/extensions/bits/>
- How to comment/request:
 - JATS https://groups.niso.org/apps/group_public/add_comment.php?document_id=25759
 - STS https://groups.niso.org/apps/group_public/add_comment.php?document_id=18492
 - BITS: Comments to JATS public email list: jats-list@lists.mulberrytech.com