Improving Journal Article Tag Suite for multilingual articles

Vincent Lizzi
Taylor & Francis Group, Philadelphia, Pennsylvania, USA

Abstract
The scenarios for journal articles that contain more than one language are no longer (and never really were) limited to having an article's title, abstract, and keywords translated to additional languages. Journal Article Tag Suite (JATS) currently has a variety of structures for tagging articles that are in multiple languages or have substantial amounts of content in more than one language. However, these structures are not all coherent and are not up to the tasks of handling some common use cases. A subcommittee of the National Information Standards Organization (NISO) JATS Standing Committee (with participation from members of the Standards Tag Suite (STS) and Book Interchange Tag Suite (BITS) committees and some other invited experts) was formed, in 2021, with the goal of recommending changes to JATS to enable it to usefully encode multilingual articles. The subcommittee has recommended a set of changes that introduce new structures that can be available to JATS users who need them while minimizing the burden JATS users who rarely deal with multilingual content. Most of these changes are backward compatible with earlier versions of JATS. These changes are currently a work in progress and may become available in a future version of JATS. This paper presents a proposal for improving JATS to better support tagging multilingual articles with the hope of garnering feedback and suggestions from the JATS community.

Keywords
Journal Article Tag Suite; Language; Multilingualism

Introduction
The scholarly publishing community is becoming increasingly aware of the need for supporting multilingualism in journal publishing to enable inclusion and to reach global audiences. The National Information Standards Organization (NISO) Journal Article Tag Suite (JATS) Standing Committee and a subcommittee on multilingual content started a project, in 2021, to recommend changes that could be made in a future version of JATS to offer better ways of tagging journal articles that are in multiple languages or have substantial amounts of the content...
in more than one language. The team that undertook this project, which included members of the NISO JATS Standing Committee and other invited experts, gathered and examined examples of journal articles that contain more than one language. The project team began with a “fresh eyes” perspective and tried to produce a way for people who have documents in multiple languages to encode information about the languages used, the relationships among the multiple versions of a portion of the document, and the expected use of these documents or document portions, without imposing a substantial burden on JATS users who do not have multilingual documents. The resulting proposal is being presented in this paper with the hope of sparking discussion and feedback from JATS users who work with multilingual content.

This paper begins by providing a review of the language support that is currently available in JATS (version 1.3 and earlier versions). Next, the proposal is presented in which new attributes are added to existing elements to identify alternative language versions of text, and the concept of language groups is introduced. The proposal also describes a new metadata element and changes to a few existing elements to better support journal articles that contain metadata in more than one language. This paper then provides examples to show how the language attributes could be useful in common scenarios. In its conclusion, this paper requests feedback from the JATS community and describes where to send feedback (to the JATS-List listserv or the author). Suppl. 1 provides an alphabetical list of JATS elements that are currently in scope for the proposed changes.

**Existing Language Support in JATS**

To begin, it is worth giving attention to the mechanisms that are currently available in JATS (including version 1.3 and earlier versions) to support journal articles that contain more than one language.

- Identify the language of text
  - The @xml:lang attribute can be placed on an element to identify the language of content contained within the element. The @xml:lang attribute is available on many elements, but not on all elements.
  - The @hreflang attribute can be placed on linking elements to identify the language of the document that is being linked.
  - @xml:lang and @hreflang hold BCP 47/ISO 639 language codes which can include tags for script, region, variation.
- Contain alternative language versions of text
  - Many elements are allowed to repeat to hold alternate language versions of content with the language identified using the @xml:lang attribute.
  - There are five elements that exist to hold translated versions of certain information: `<trans-title-group>`, `<trans-title>`, `<trans-subtitle>`, `<trans-source>`, `<trans-abstract>`. These elements are documented as being limited to holding translated versions of text for their specific items (although we understand that community usage has been broader than that).
  - Characters of any script can be represented
    - Unicode character set supports scripts for most languages.
    - `<private-char>` and `<inline-graphic>` with image files can be used for glyphs that are not available in Unicode.

The `@xml:lang` attribute is defined in the W3C Recommendation *Extensible Markup Language (XML) 1.0 (Fifth Edition)* [1], which is the standard that defines XML itself, to “identify the natural or formal language in which the content is written.” The `@xml:lang` attribute is defined to hold a language identifier that conforms to the standard IETF BCP 47 *Tags for the Identification of Languages* [2], which is based on several standards including ISO 639 *Language Codes*. The `@xml:lang` attribute is scoped to apply to the elements’ textual content, the elements’ attribute content, and all the elements’ descendants until another `xml:lang` attribute is encountered. The W3C has published an informative document *Language tags in HTML and XML* [3] that explains how the `@xml:lang` attribute can be used to identify language, script, region, and variations. The JATS Tag Library also contains informative documentation about the `@xml:lang` attribute [4].

The scripts for most writing systems in the world can be used in JATS documents thanks to the Unicode Character Set (ISO/IEC 10646) [5]. Unicode allows the scripts for most languages, both modern and ancient, to be represented in JATS documents. Glyphs that are not available in Unicode can be represented in JATS documents using the `<private-char>` and `<inline-graphic>` elements along with an image file.

**Proposal**

This proposal recommends creating a new set of attributes for marking up multilingual content. These new attributes, which are referred to collectively as the lang-* attributes, should be added as a group to block-level elements, text-containing metadata elements, and most elements that currently have the `xml:lang` attribute. That is, an element that allows any of the lang-* attributes should allow all of them so that all the lang-* attributes will be available on elements where users might want them. All the lang-* attributes will be:

- optional, as we expect most documents will not use any of them.
• provided in the DTD modules as a parameter entity, so users who do not want them can easily remove them in their local subsets.

A key concept is that of a language group. A language group consists of all the alternate language versions of some content. This may include versions that are considered parallel, or an original and translation(s) or other nonoriginal versions. In some documents it is expected/required that all of the language versions of content be displayed, in other documents it is expected that a display system will select appropriate version(s) to display. The members of a language group are not wrapped in a structural element but they are associated with each other. The lang-* attributes provide the mechanism for associating the members of a language group.

The use of attributes, rather than elements, to hold language information builds on the existing @xml:lang and @hreflang attribute mechanism for marking up the language of content, and has the advantages that it:
• avoids cluttering the tag set with body elements that most users will not need.
• puts structures that are provided in multiple languages in the same place in the document as the same structure would be if there were only one version of it.

The language attributes will contain metadata to:
• Identify language using existing attributes @xml:lang and @hreflang with standard 2-letter or 3-letter language codes, and optionally include identifiers for region, dialect, and ancient languages.
• Associate alternative language versions of the same passage using a @lang-group attribute.
• Label each alternative language version (original, translation, etc.) using a @lang-variant attribute.
• Indicate the source of an alternative language version of a passage (author, editor, translator, machine, etc.) using a @lang-source attribute.
• Mark text that should not be translated (for example by automated services like Google Translate) using a @lang-translate attribute.
• Indicate intent for how alternative language versions of a passage should be displayed (primary, secondary, etc.) using a @lang-focus attribute.

Example:
<p> <italic id="phrase001" lang-group="phrase001" xml:lang="la" lang-variant="original" lang-translate="no"> carpe diem </italic> ( <styled-content lang-group="phrase001" xml:lang="en" lang-variant="translation"> seize the day </styled-content> ) </p>

The lang-* attributes are described in more detail in the following sections.

There is a distinction between article metadata that is present in multiple languages and article content that is present in multiple languages. Article metadata typically consists of an article's title, authors, keywords, and other information that can be contained within the <front> element. Article content typically consists of an article's text, figures, tables, and other content that can be contained within the <body> and <back> elements. The nature of article metadata allows alternative language versions of metadata items to appear next to each other, while the nature of article content is such that alternative language versions of content items may need to occur in separate locations throughout a document.

• In situations where article metadata is present in more than one language the element that holds the metadata item can repeat with a different @xml:lang attribute value for each alternative language version. For example, the kwd-group element can repeat to hold alternative language versions of keywords thus forming a language group for keywords.
• In situations where article content is present in more than one language the element that holds the content item can repeat with a different @xml:lang attribute value for each alternative language version. The alternative language versions of content items may be in separate locations throughout a document. The @lang-group attribute provides a mechanism to associate alternative language versions of a content item.

This proposal also recommends changes to bring more consistency to tagging article metadata that is present in two or more languages. Most structures within article metadata are able to repeat with different @xml:lang attribute values to hold alternate language versions. This proposal recommends making this consistent for all article metadata elements that could have alternative language versions. There are a few metadata elements that are not allowed to repeat in the current version of JATS but which may have alternative language versions so these elements should be allowed to repeat. In the current version of JATS there is a set of five elements that can hold a translated title, subtitle, abstract, and source; these elements are consistent with the way that JATS handles alternative language versions of other metadata so this proposal recommends deprecating these elements in favor of better consistency. Also, since community use of these translation elements is not limited to translations, we recommend using elements that meet both the needs of current users and of producers of multilingual articles.

Language Group attribute @lang-group
The @lang-group (Language Group) attribute is used to associate all members of a language group that present alternate language versions of the same content. The members of a lan-
guage group may appear next to each other, in different places within a document, or in different documents.

The value of the @lang-group attribute should be the same for all the members of a language group to support processing. For example, the members of a language group may be processed using `<xsl:for-each-group select="/" group-by="@lang-group">` to produce a display of the language group.

@lang-group is defined as IDREF, which enables the DTD to assist users in creating links using ID/IDREF. The group name must be a valid @id attribute value in the same document. For example, the @id attribute could be on one member of the language group or on a relevant parent element.

Language Variant attributes @lang-variant and @lang-variant-custom

The @lang-variant (Language Variant) attribute indicates how language of the text is acting on the content, for example, "transliteration," "phonetic," or "spoken". The @lang-variant attribute describes the relationship between alternative language versions of the content. The @lang-variant attribute labels an alternative language version of a passage; a passage can be any piece of content, as small as a word or as big as an entire journal article. The list of values for the @lang-variant attribute should include:

- original – a passage in its original language
- translation – a translation of a passage into another language
- interpretation – a rewording of a passage into another language
- transcription – a representation of spoken language in a written form
- transliteration – a mapping from one system of writing into another
- phonetic – a representation of speech sounds using phonetic symbols
- spoken – a passage spoken aloud
- unknown – the language relationship is unknown. This is not the same as omitting the attribute; this is truly "we do not know."
- custom – any other label or any combination of the above labels, to be used in conjunction with the @lang-variant-custom attribute.

The @lang-variant-custom (Language Variant Custom) attribute holds a text string to describe the language variant, which may be a combination of the values above or any text string, when the value of @lang-variant is "custom". This attribute works in the same way as the @*-custom mechanism works for other fixed lists.

The list of values for the @lang-variant attribute may grow over time. If usage shows that some combinations of values are popular/common those combinations can be added as hyphenated values of the base values (e.g., "transcription-translation").

By design, this value list does not include "equivalent". The members of a language group will be considered to be equivalent if no @lang-variant attribute is present to differentiate by naming one, for example, a "translation". For example, a title and a parallel title (where neither is a translation) are merely two titles with different language attribute values, so they are equal or equivalent titles.

Language Focus Suggestion attributes @lang-focus and @lang-focus-custom

The @lang-focus (Language Focus Suggestion) attribute indicates how members of a language group are related to each other, which may be used as a hint for how multiple language variants might be displayed. The manner in which the members of a language group are displayed is application specific, perhaps driven by reader preference. However, the @lang-focus attribute can be used to provide hints in markup about the authors’ intention for how a language group should be displayed. The list of values for the @lang-focus attribute should include:

- primary – The text has a more central focus than other language variants in the group. In display, such text is typically made more prominent or be the only focus displayed.
- secondary – The text is not the primary textual focus in the language group. In display, such text is typically less prominent.
- undefined – No recommendation is made concerning the relative focus of the language variants, thus all variants are intended to be displayed the same way. In display, all language variants are typically displayed in document order.
- custom – The language relationship is not any of the specific listed values. The @lang-focus-custom attribute should be used to specify the relationship.

The @lang-focus-custom (Language Focus Suggestion Custom) attribute holds a text string to indicate how members of a language group are related when the value of @lang-focus is "custom". This attribute works in the same way as the @*-custom mechanism works for other fixed lists.

Language Source attributes @lang-source and @lang-source-custom

The @lang-source (Language Source) attribute indicates the source of an alternative language version. The list of values for the @lang-source attribute should include:

- author – provided by an author of the article
- editor – provided by an editor of the article
• translator – provided by a human translator
• machine – provided by a machine (for example, automated translation software)
• custom – The source is not any of the specific listed values. The @lang-source-custom attribute should be used to specify the source.

The @lang-source-custom (Language Source Custom) attribute holds a text string to indicate the source when the value of @lang-source is “custom”. This attribute works in the same way as the @*-custom mechanism works for other fixed lists.

Language Translate attribute @lang-translate
The @lang-translate (Language Translate) attribute is used to specify whether an element’s content is to be translated when the content is localized or whether to leave the content unchanged. Text that should not be translated, for example automatically by Google translate, can be tagged with attribute lang-translate = “no” to indicate that the text should not be translated. The @lang-translate attribute can hold a value of either “no” or “yes”.

If a passage of text that should not be translated that is tagged with lang-translate = “no” contains text that can be translated, the setting can be toggled by adding attribute lang-translate = “yes” to the text that can be translated.

The @lang-translate attribute is based on the HTML5 @translate attribute, which is defined in the HTML5 standard [6] and described in “Using HTML’s translate attribute” [7].

Add @xml:lang and lang-* attributes to face markup elements
It is very common for text that is in a different language than the surrounding text to be set apart stylistically, for example in italicized font using the <italic> element. However, the <italic> element does not allow the @xml:lang attribute in JATS 1.3 and earlier versions of JATS. The @xml:lang and lang-* attributes should be available on all face markup elements.

- italic
- roman
- sans-serif
- sc
- strike
- bold
- monospace
- overline
- underline
- rb
- rt
- styled-content

The lang-* attributes will also be added to structural elements. Suppl. 1 to this paper contains a list of all elements that are proposed to receive lang-* attributes.

New <content-language> element
In current versions of JATS (version 1.3 and earlier) there is no way to identify the languages of an article in cases where the main content of the article is provided in two or more languages.

For the majority of journal articles, which have the content of the article in one language and perhaps have metadata and some text passages in other languages, the @xml:lang attribute on the <article> element can identify the primary language of the article. However, for journal articles in which the text of the article is provided in more than one language, the @xml:lang attribute cannot hold multiple values. The @xml:lang attribute can hold a value “mul”, which is defined by ISO 639 to indicate “multiple languages”, but this does not identify what the multiple languages are.

Book Interchange Tag Suite (BITS) [8] and NISO STS [9], which are part of the family of tag sets that are based on JATS, have a <content-language> element that can hold metadata to identify each primary language when a document contains more than one primary language.

The BTIS Tag Library says of <content-language>:
• Part of the metadata of a document used to identify the primary language(s) used in the document. This element should appear once for each primary language. For Best Practice, the content of <content-language> should be the two-letter ISO 639 code for the language that are typically used as values for @xml:lang., for example, “en” for English, “de” for German, or “es” for Spanish.
• The tag set is agnostic on how “primary” is defined, leaving that decision to each producer. However, the intent of this element is to record the principle languages used in a multi-lingual document, not to state that three Latin quotations are intermixed with primarily German content. An abstract in Spanish in an otherwise Portuguese paper would be a single primary language.

The <content-language> element is only useful in situations where the <article> element has attribute @xml:lang = “mul” to identify what the multiple languages are. The <content-language> element should appear once for each primary language used in the text of a multi-lingual document. There is no value in using the <content-language> element in a monolingual document (one where the @xml:lang attribute on <article> element has a value other than “mul”).

The <content-language> element should be added to JATS, and allowed to appear as a repeatable optional element within <article-meta>. Within the content model of the <article-meta> element, the <content-language> element should be
placed after <author-notes> element. This placement is based on the content model of the <book-meta> element in BITS.

For articles that have the main content of the article presented in more than one language it will be recommended to:
1. Use @xml:lang = "mul" on the <article> root element.
2. Include a <content-language> element with a language code for each primary language in the main <article-meta> in <front>.
3. Tag a <sub-article> element with attributes @xml:lang and @lang-group for each alternative language version of the main content. The @lang-variant, @lang-focus, and @lang-source attributes could also be usefully applied to the <sub-article> elements if desired.

**Changes to <article-meta> and <front-stub> for metadata in more than one language**

The content model of the <article-meta> and <front-stub> elements should be updated to:
1. Add <content-language> element as optional and repeatable after <author-notes>.
2. Allow <author-notes> to repeat in order to allow alternative language versions of author notes.
3. Allow <supplement> to repeat and clarify its use in the documentation of this element.
4. Allow <title-group> to repeat in order for article title to be provided in more than one language.

All other article metadata elements are already able to repeat for multiple languages. Article metadata elements are either able to repeat directly within <article-meta> (and <front-stub>) or they can be placed in a grouping element where it is allowed to repeat.

Examples of metadata elements that repeat within a grouping element are:
1. <article-version> can repeat within <article-version-alternatives>.
2. <article-categories> can contain repeatable <subject-group>, <series-title>, and <series-text> elements.
3. <volume-series> can repeat within <volume-issue-group>.
4. <issue-part> can repeat within <volume-issue-group>.
5. <pub-history> can contain repeatable <event> elements.
6. <history> can contain repeatable <date> elements.
7. <permissions> can contain repeatable <copyright-statement>, <copyright-holder>, <license> elements.
8. <name> for a contributor can repeat within <name-alternatives>.
9. <aff> can repeat within <aff-alternatives>.
10. <institution> in a funding source or in an affiliation can repeat within <institution-wrap>.

There are a few article metadata elements that are not allowed to repeat and are not likely to have alternative language versions, as is the case with the page range and <elocation-id> elements.

**Article title, issue title, and journal title**

The metadata elements that hold article title, issue title, and journal title should all handle alternate language versions in a consistent manner. Each title group element (<title-group>) for article title, <issue-title-group> for issue title, and <journal-title-group> for journal title) should be able to repeat to hold alternative language versions. This approach keeps the corresponding title elements for each language version together (for example, a title group with the French article title with the French article subtitle, and a title group with the English article title with the English article subtitle).

The @xml:lang attribute should be placed on the group element, unless the article has a very rare exception in which a title and corresponding subtitle are in different languages. The lang-* attribute group should be added to the <title-group> element.

Within the group element, the title element should be required once (<article-title> in <title-group>, <issue-title> in <issue-title-group>, and <journal-title> in <journal-title-group>).

The current version of JATS contains inconsistencies in how the metadata elements for article title, issue title, and journal title are modeled. This proposal recommends bringing consistency with these changes:
- For article title, the <title-group> element should be allowed to repeat within <article-meta> so that the article title, subtitle, and alternate titles can be provided in more than one language. The @xml:lang and @lang-* attributes should be added to <title-group>.
- For issue title, the <issue-title-group> element can repeat to hold alternative language versions of an issue title and issue subtitle. The lang-* attributes should be added to the <issue-title-group> element.
- For journal title, the <journal-title-group> element can repeat to hold journal title and subtitle in more than one language. The <journal-title> element, which is optional in current version JATS, should be required once in <journal-title-group>. The lang-* attributes and the @xml:lang attribute should be added to the <journal-title-group> element.

**Retire translation elements**

The current version of JATS (version 1.3) and all previous versions, including JATS’ predecessor NLM DTD, have five
elements that are designated to hold translations of certain metadata information. These elements are: <trans-abstract>, <trans-title-group>, <trans-title>, <trans-subtitle>, and <trans-source>. Each of these elements has a corresponding main element that can take an @xml:lang attribute and repeat for alternative language versions. Providing designated translation elements for a few metadata items is inconsistent, potentially misleading, and not as capable of handling the complexities of real multi-lingual documents.

This proposal recommends documenting the translation elements as deprecated to discourage their use, and then removing the translation elements in a future non-backwards compatible version of JATS.

The five translation elements and their preferred replacement elements are:

1. <trans-abstract> - instead use <abstract>
2. <trans-title-group> - instead:
   o In <title-group> repeat the <title-group>
   o In <journal-title-group> repeat the <journal-title-group>
   o In <issue-title-group> repeat the <issue-title-group>
3. <trans-title> - instead:
   o In citation contexts use <article-title> or <part-title>
   o In <journal-title-group> use <journal-title>
   o In <issue-title-group> use <issue-title>
4. <trans-subtitle> - instead:
   o In <title-group> use <subtitle>
   o In <journal-title-group> use <journal-subtitle>
   o In <issue-title-group> use <issue-subtitle>
5. <trans-source> - instead use <source>

Processing metadata
A new optional attribute @lang-grouping (Language Grouping Use) should be added to the <processing-meta> element. The @lang-grouping attribute flags to users that language grouping features are used in the document and may need to be processed. The values that are allowed in the @lang-grouping attribute are “yes” and “no.”

The @lang-grouping attribute is only used on the <processing-meta> element, so @lang-grouping is the only language attribute that is not part of the lang-* attribute group.

Maintaining full text semantics
The modifications that we propose making to the JATS tag set do not alter the way in which the contents of a journal article may be semantically tagged in JATS. The lang-* attributes provide ways to group together alternative language versions in text that is tagged with the usual semantics of the JATS syntax. The use of two or more languages may often suggest repetition of parts of the text. Most elements (such as paragraphs and face-markup) may naturally repeat when there are alternative language versions of text present. Some elements may be grouped with elements of another kind and in another location to express their language relation (such as display quotes or verses and their translation). Some components might be more challenging. For example, while tables can contain translated content, the fundamental table structure must not be broken; supplying additional cell entries in a row would collide with how tables work and should not be done. The handling of multilingual content may in some cases demand some design insight from the user.

Examples

The following examples show some of the ways in which the language attributes could be useful in common scenarios:

- article metadata (article title, keywords, abstract, etc.) in two or more languages
- substantial portions of content in two or more languages
- the entire document in two or more languages

Metadata items in two or more languages

**Article title**

```
<issue-title-group xml:lang="en" lang-variant="original" lang-source="author">
  <issue-title>EXPOSURE TO COVID-19 RISK REPRESENTATIONS AND STATE DEPRESSIVE SYMPTOMS IN A UNITED KINGDOM SAMPLE: A PRELIMINARY EXPERIMENTAL STUDY</issue-title>
</issue-title-group>
```

**Issue title**

```
<issue-title-group xml:lang="es" lang-variant="translation" lang-source="translator">
  <issue-title>REPRESENTACIONES DE RIESGOS REFERENTES A LA EXPOSICIÓN AL COVID-19 Y SÍNTOMAS DEPRESIVOS ACTUALES EN UNA MUESTRA DEL REINO UNIDO: UN ESTUDIO EXPERIMENTAL PRELIMINAR</issue-title>
</issue-title-group>
```
This study examined the impact of being in lockdown... This study examined the impact of being in lockdown...

Este estudio analizó el impacto del confinamiento... Este estudio analizó el impacto del confinamiento...

Social representations of risk

Socio-demographic characteristics of the sample

Características demográficas de la muestra
Full text in two or more languages

There are two general approaches that publishers follow when publishing journal articles where the full text of the article is provided in two or more languages. The JATS tag set can support either of these approaches.

- Full text in two or more languages published as one article.
- Full text in two or more languages published as separate articles.

The decision of which of these approaches to use can be influenced by a variety of factors that may include timing (for example, whether a translation or alternate language version is ready before or after publication), publishing agreements, and content management needs.

Full text in two or more languages in one article

A journal article that contains the full text of the article in more than one language can be tagged as follows.

- The <article> root element should have attribute xml:lang="mul" (multiple languages).
- Each language should be identified by a <content-language> element containing a 2-letter language code within <article-meta> in <front>.
- Each alternative language version should be tagged in a <sub-article> element with an @xml:lang attribute to identify the language and a @lang-group attribute to connect the alternative language versions.

This approach allows a wide variety of options to choose from in regard to how much of the article’s content is presented with alternate language versions or with only one language. For example: figures and tables might be present with alternative language versions or only once, a reference list might be present only once. Any content item that is presented with alternate language versions can be tagged using the language attributes to identify each one as part of a language group.

Full text in two or more languages in separate articles

When a journal article has alternative language versions published as separate articles, each separate article may be linked together using the <related-article> element.

The <related-article> element uses the @xlink:href attribute to point to another article. The @hreflang attribute can identify the language of the article that is pointed to by the @xlink:href link. The @related-article-type attribute identifies the type of relationship that the related article has to the current article.

The @related-article-type attribute value “alt-language” indicates that the linked article is an alternative language version of the current article.

As greater numbers of researchers and publishers become interested in publishing journal articles that contain multilingual content, and as the technical challenges of doing so are increasingly solved, there may be new opportunities. Language metadata in journal articles could be used in a variety of ways such as: improving accessibility and discovery by providing metadata that identifies the language to screen readers and search engines; linking and navigation within an article; repackaging multi-language content in new forms; and reusing content in the world of publishing in general, and it is likely that there are use cases for multilingual content in existence that have not come to the attention of the project team. Therefore, this paper concludes with a call to the JATS community to provide additional use cases, suggestions, feedback, and discussion regarding the proposal set forth in this paper. Comments can be sent to the JATS-List listserv or to the author of this paper. The details about how to subscribe to the JATS-List listserv are available at https://www.mulberrytech.com/JATS/JATS-List/.
Conflict of Interest

No potential conflict of interest relevant to this article was reported.

Funding

The author received no financial support for this article.

Acknowledgments

The members of the JATS Multi-lingual Article subcommittee are: B. Tommie Usdin (Mulberry Technologies, Inc.) (Chair), Antti Saari (Finnish Standards Association SFS), Holly Bodger (Canadian Medical Association Journal [CMAJ]), Hugh Cayless (Duke University), Jan Driesen (Brepols Publishers NV), Laura Randall (National Library of Medicine [NLM]), Martin Latterner (NLM), Mary Seligy (Mary Seligy Independent), Mathieu Pigeon (Érudit), Nettie Lagace (National Information Standards Organization [NISO]), Patricia Feeney (Crossref), Vincent Lizzi (Taylor & Francis Group).

The members of the JATS (Z39.96) Standing Committee are: Jeffrey Beck (NLM) (Chair), B. Tommie Usdin (Mulberry Technologies, Inc.) (Chair), Ardie Bausenbach (Library of Congress), Brooke Begin (Silverchair Information Systems), Bruce Rosenblum (Inera Inc.), Cinthia Vieira (SciELO), Cory Schires (Scholastica), Debbie Lapeyre (Mulberry Technologies, Inc.), Franziska Buehring (De Gruyter), John Meyer (ITHAKA/JSTOR/Portico), Joni Dames (Inera Inc.), Josh Pyle (Atypon), Kathleen Sheedy (American Psychological Association [APA]), , Kennett Rawson (IEEE), Laura Randall (NLM), Mark Doyle (American Physical Society [APS]), Mary McRae (Orbis Technologies), Mathieu Pigeon (Érudit), Michael Parkin (Europe PMC), Nettie Lagace (NISO), Nick Nunes (Nick Nunes Individual), Patricia Feeney (Crossref), Paul Donohoe (Macmillan Science and Education), Sara Groveman (NISO), Soichi Tokizane (Aichi University), Vincent Lizzi (Taylor & Francis Group).

Supplementary Material

Supplementary file is available from: https://doi.org/10.7910/Suppl.1. An alphabetical list of JATS elements

References

1. Bray T, Paoli J, Sperberg-McQueen CM, Maler E, Yergeau F, editors. Extensible Markup Language (XML) 1.0 (Fifth edition) [Internet]. W3C; 2013 [cited 2022 Apr 21]. Available from: https://www.w3.org/TR/xml/
2. Phillips A, Davis M, editors. Tags for identifying languages [Internet]. 2009 [cited 2022 Apr 21]. Available from: https://www.rfc-editor.org/info/rfc5646
3. Ishida R. Language tags in HTML and XML [Internet]. W3C; 2014 [cited 2022 Apr 21]. Available from: https://www.w3.org/International/articles/language-tags/
4. @xml:lang Language in Journal Archiving and Interchange Tag Library NISO JATS version 1.3 (ANSI/NISO Z39.96-2021) [Internet]. Bethesda: National Library of Medicine; 2021 [cited 2022 Apr 21]. Available from: https://jats.nlm.nih.gov/archiving/tag-library/1.3/attribute/xml-lang.html
5. Basic questions [Internet]. Mountain View: Unicode Inc; c1991-2022 [cited 2022 Apr 21]. Available from: https://www.unicode.org/faq/basic_q.html
6. The translate attribute in HTML Living Standard [Internet]. WHATWG community; 2022 [cited 2022 Apr 21]. Available from: https://html.spec.whatwg.org/multipage/dom.html#the-translate-attribute
7. Using HTML’s translate attribute [Internet]. W3C; 2014 [cited 2022 Apr 21]. Available from: https://www.w3.org/International/questions/qa-translate-flag
8. <content-language> (Content Language) in Book Interchange Tag Suite (BITS) version 2.1 Tag Library [Internet]. Bethesda: National Library of Medicine; 2022 [cited 2022 Apr 21]. Available from: https://jats.nlm.nih.gov/extensions/bits/tag-library/2.1/element/content-language.html
9. <content-language> (Content Language) in NISO STS 1.0 Tag Suite Components including TBX-TML [Internet]. Baltimore: National Information Standards Organization; 2017 [cited 2022 Apr 21]. Available from: https://www.niso-sts.org/TagLibrary/niso-sts-TL-1-0-html/element/content-language.html