NCBI Bookshelf. A service of the National Library of Medicine, National Institutes of Health.

Journal Article Tag Suite Conference (JATS-Con) Proceedings 2012 [Internet]. Bethesda (MD): National Center for Biotechnology Information (US); 2012.

Implementing XML for Japanese-language scholarly articles

Soichi Tokizane.

Aichi University

Scholarly publishers and typesetting service providers in Japan, as well as the J-STAGE e-journal platform got together to devise a way to implement XML for Japanese-language scholarly articles, and worked with JATS working group to develop multi-language extention of JATS. It is now implemented in the new J-STAGE platform that was launched in May, 2012. Characteristics of Japanese writings, characteristics of Japanese publication, how JATS solved problems of coding Japanese-language articles, and how typesetting companies create JATS XML for J-STAGE are discussed.

SGML/XML for Japanese-language scholarly articles

Several experiments were made to publish Japanese-language articles in SGML in early 1990s, and in XML in early 2000s. These are discussed in detail by the author elsewhere [1].

When the scholarly journal platform for Japanese society publishers, J-STAGE, an e-journal publication platform, was launched by the Japan Science and Technology Agency (JST) in 1999, it was anticipated that the system should realize full text HTML publication. J-STAGE developed a J-STAGE SGML DTD based on JICST-DTD, and encouraged publishers to use it. J-STAGE also investigated the possibility of XML publishing in 2000, and developed DTD for XML, too [2]. But because of the lack of convenient tools, only three J-STAGE journals has been publishing their articles in HTML now. Most of 800 journals on J-STAGE publish articles in PDF.

After many years since its first launch of J-STAGE, JST planned to replace it with a new version in Spring 2012, and decided to implement XML as the content base for the new platform[3], because:

- XML enables flexible presentation of journal articles as demonstrated by many western journal publishers.
- XML allows publishers to distribute their contents globally, for example, via PubMed Central.
- XML enables semantic enrichment of journal contents such as by semantic tagging.

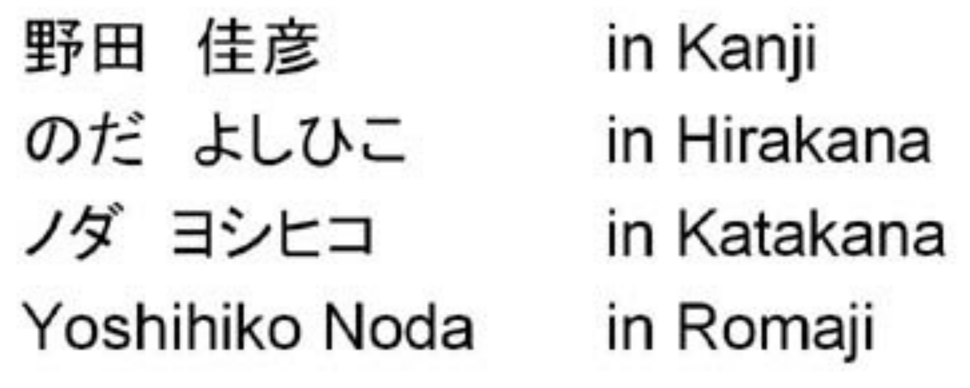
JST decided to adopt JATS 0.4 for this purpose. The detail of this implementation will be discussed later in this paper.

Characteristics of Japanese Writing

Scripts

Japanese writing consists of four kinds of scripts, i. e. Kanji, or Chinese Characters (somewhat modified from traditional Chinese), Hirakana, Katakana, and Romaji or Romanized Japanese. Kanji is used mainly to express personal, geographical, material names, object and abstract names, and most technical terms. Instead, Hirakana and Katakana are phonetic, and used mainly as suffixes of verbs and adjectives, and also to express pronunciation (see Ruby).

A name, Yoshihiko Noda, may be expressed in those scripts as in Figure 1.



野田　佳彦	in Kanji
のだ　よしひこ	in Hirakana
ノダ　ヨシヒコ	in Katakana
Yoshihiko Noda	in Romaji

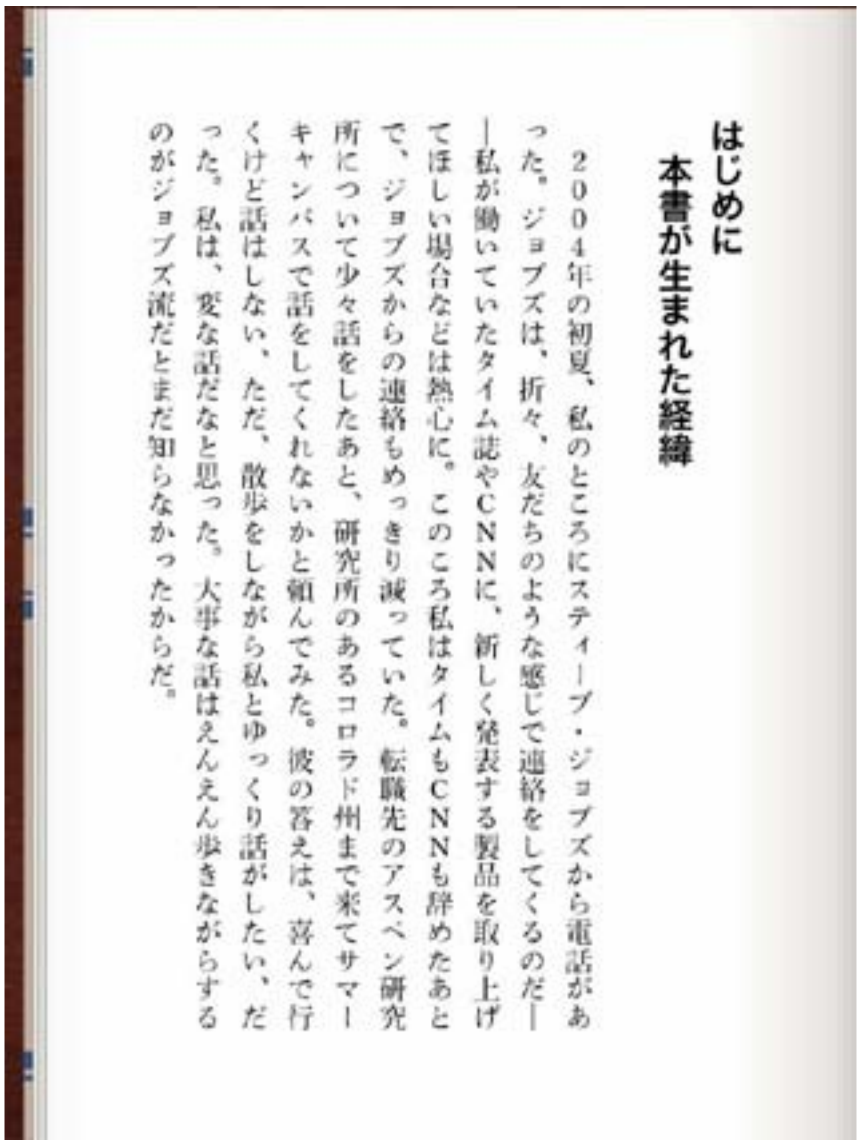
Fig. 1　Script variations of Japanese personal name.

Name order

Please note that, like in many Asian nations, a person's name is written family name first, first name next, in Japanese context. In XML, this is specified by name-style="eastern".

Vertical, right-to-left, and resulting font changes

Traditionally, Japanese writing was vertical (top-to-down) and right-to-left, as used in China until recently (China today uses horizontal writing altogether). It is still very popular for newspapers, magazines, books as well as scholarly publications in humanity and social science area (Figure 2). On the other hand, horizontal writing is used exclusively in business writings. It is also a common practice in science and technology publishing, including most scholarly journals.



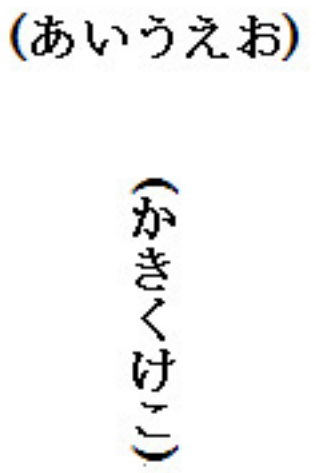
はじめに

本書が生まれた経緯

２００４年の初夏、私のところにスティーブ・ジョブズから電話があった。ジョブズは、折々、友だちのような感じで連絡をしてくるのだ——私が働いていたタイム誌やＣＮＮに、新しく発表する製品を取り上げてほしい場合などは熱心に。このころ私はタイムもＣＮＮも辞めたあとで、ジョブズからの連絡もめっきり減っていた。転職先のアスペン研究所について少々話をしたあと、研究所のあるコロラド州まで来てサマーキャンパスで話をしてくれないかと頼んでみた。彼の答えは、喜んで行くけど話はしない、ただ、散歩をしながら私とゆっくり話がしたい、だった。私は、変な話だなと思った。大事な話はえんえん歩きながらするのがジョブズ流だとまだ知らなかったからだ。

Fig. 2　Example of vertical writing for a Japanese novel.

Some characters have to be changed in vertical writing vs. horizontal writing. For example, parentheses have to be horizontal in vertical writing (Figure 3). This is normally handled by typesetting programs. This feature is supported by MS Word.



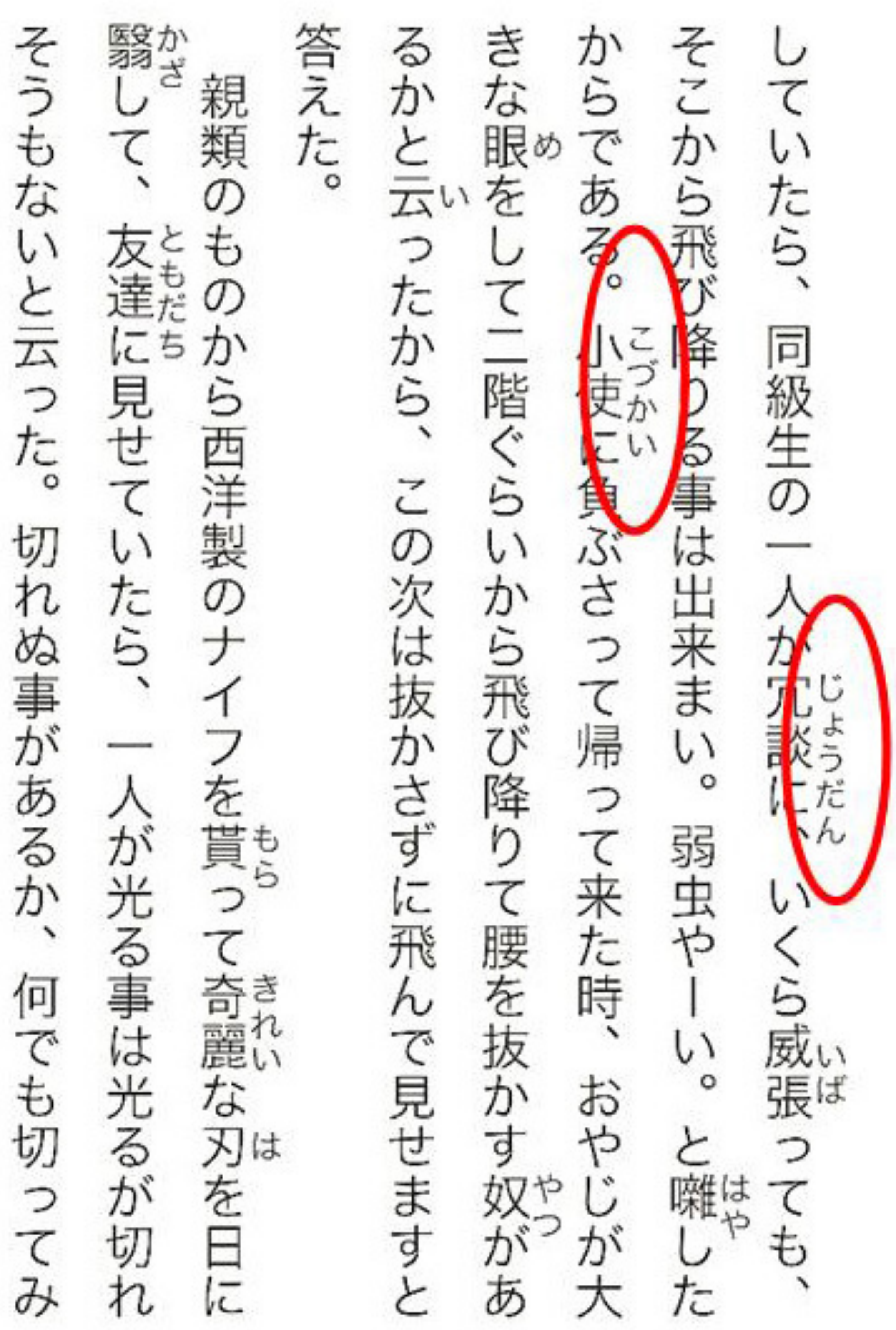
（あいうえお）

（かきくけこ）

Fig. 3　Parentheses for horizontal and vertical writings.

Ruby

As Kanji (Chinese character) writing allows multiple readings (pronunciations), especially for personal and geographical names, and thus are difficult to read sometimes. To help read correctly, we often associate such Kanji characters with Kana (phonetic) characters, usually right-side of a word (for vertical writing) or above a word (for horizontal writing) in smaller fonts (Figure 4). Ruby is also used to teach children how to read Kanji. MS Word supports this feature.



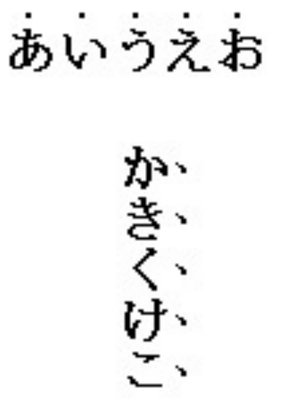
していたら、同級生の一人が冗談（じょうだん）に、いくら威張（いば）っても、そこから飛び降りる事は出来まい。弱虫やーい。と囃（はや）したからである。小使（こづかい）に負ぶさって帰って来た時、おやじが大きな眼（め）をして二階ぐらいから飛び降りて腰を抜かす奴（やつ）があるかと云（い）ったから、この次は抜かさずに飛んで見せますと答えた。

　親類のものから西洋製のナイフを貰（もら）って奇麗（きれい）な刃（は）を日に翳（かざ）して、友達（ともだち）に見せていたら、一人が光る事は光るが切れそうもないと云った。切れぬ事があるか、何でも切ってみ

Fig. 4　Examples of Ruby

Emphasis

In Western articles, underlines, italics and bold fonts are used to emphasize a word. In Japanese writing, we typically use emphasis in dots and other characters. They are placed similar to Rubies, i. e. right-side of a character, or above a character. MS Word support this feature.



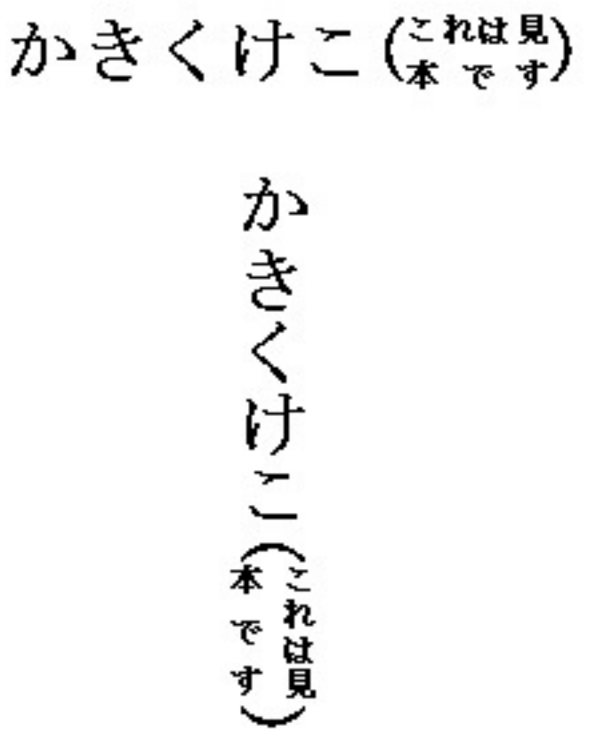
あいうえお

かきくけこ

Fig. 5　Examples of Emphasis

Warichu

Warichu is a short note inserted within a sentence in two lines, typically with parentheses. This is often used in humanity scholarly publications, and supported by MS Word.



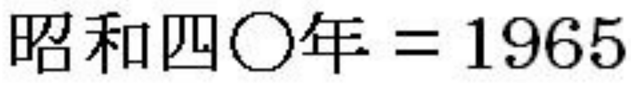
かきくけこ（これは見本です）

かきくけこ（これは見本です）

Fig. 6 Examples of Warichu

non-Gregorian years

Traditionally, Japanese used its own year designations as Chinese did. They were changed very often, and seldom last more than 20 years. Recently, such designations reflect Emperor's era, and may be called, "Emperor year". This is the Japanese formal year designation used by the Government. In scholarly articles, more and more authors use Gregorian years, they still appear in citations of old references.



昭和四〇年＝1965

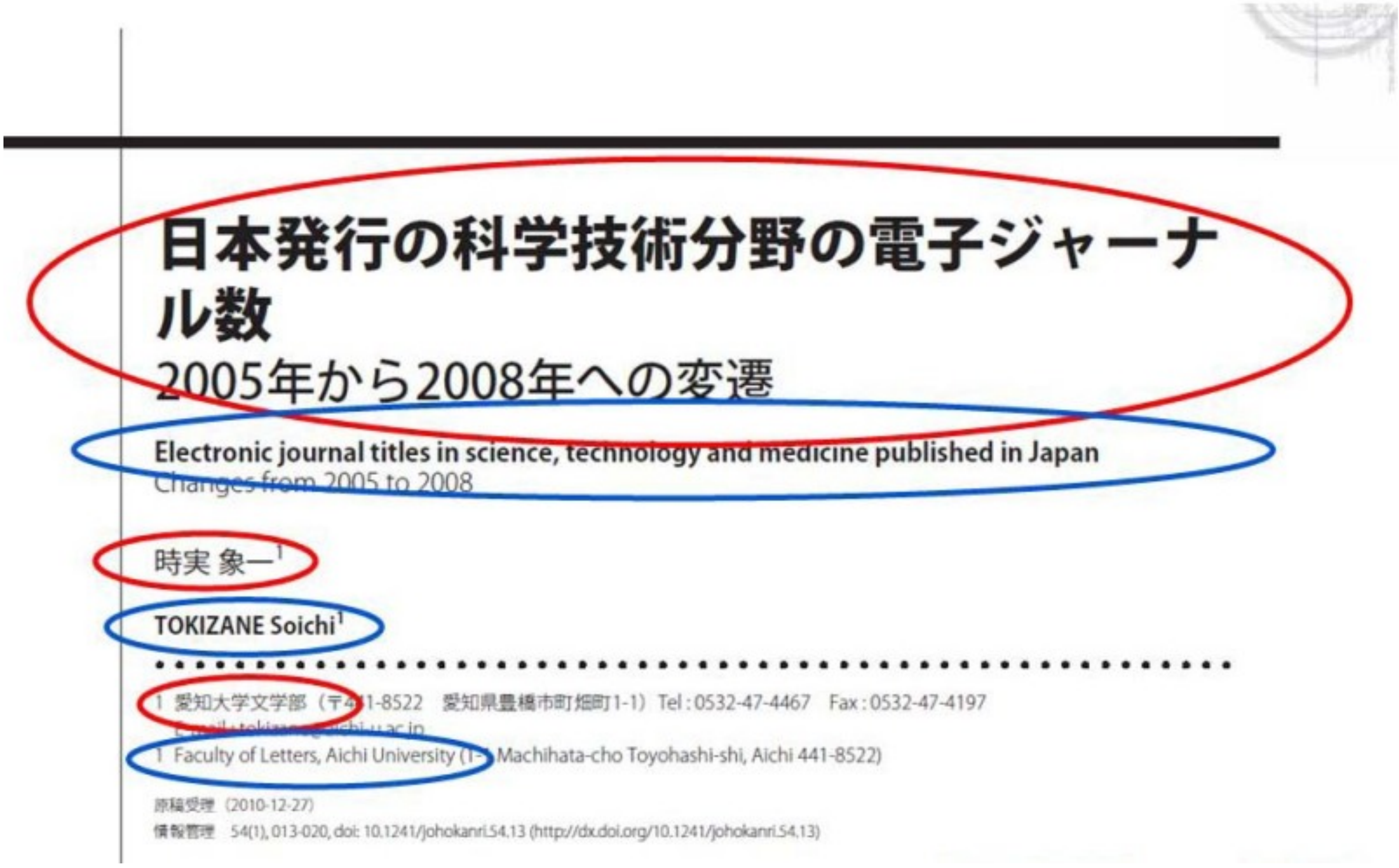
Fig. 7 non-Gregorian year. The left year is read as, "Showa year 40".

Characteristics of Japanese Scholarly Publishing

The author investigated how many scholarly journals were published in Japan in 2005 and 2008 [4]. The year 2008 result showed that there were 3,047 journal titles published in the STM fields in Japan, out of which 2,673 were in Japanese and 374 were in English. The number of ejournals were 774 (29.0%) and 298 (79.7%) respectively.

English-language journals are formatted and published quite the same way as most Western journals. Japanese-language STM journals are, style-wise, quite similar to Western journals, too, except the following differences.

- Article titles, author names, and affiliations are almost always both in Japanese and English (or Romanized), and abstracts and keywords are, too, most of the times (Figure 8).



日本発行の科学技術分野の電子ジャーナル数

2005年から2008年への変遷

Electronic journal titles in science, technology and medicine published in Japan

Changes from 2005 to 2008

時実 象一[1]

TOKIZANE Soichi[1]

1 愛知大学文学部（〒441-8522　愛知県豊橋市町畑町1-1）Tel：0532-47-4467　Fax：0532-47-4197
E-mail：tokizane@aichi-u.ac.jp

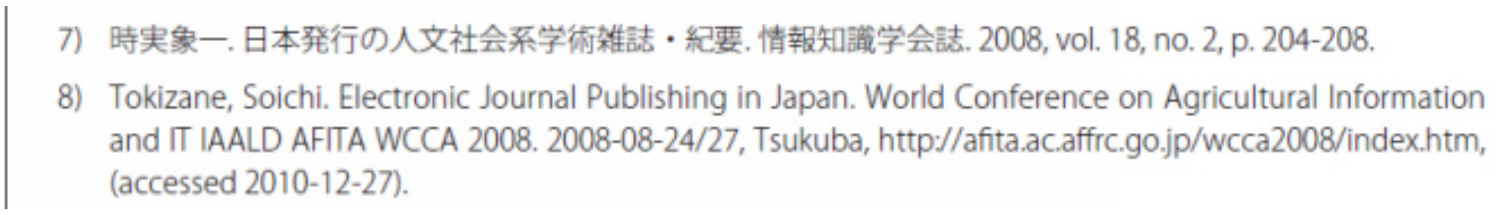
1 Faculty of Letters, Aichi University (1-1 Machihata-cho Toyohashi-shi, Aichi 441-8522)

原稿受理（2010-12-27）

情報管理　54(1), 013-020, doi: 10.1241/johokanri.54.13 (http://dx.doi.org/10.1241/johokanri.54.13)

Fig. 8　Metadata example of a Japanese-language article. Japanese metadata are circled in red, and English ones in blue.

- Cited references are mostly either in Japanese or English (Figure 9), but in a few journals, two language versions are described for a single citation.



7)　時実象一. 日本発行の人文社会系学術雑誌・紀要. 情報知識学会誌. 2008, vol. 18, no. 2, p. 204-208.

8)　Tokizane, Soichi. Electronic Journal Publishing in Japan. World Conference on Agricultural Information and IT IAALD AFITA WCCA 2008. 2008-08-24/27, Tsukuba, http://afita.ac.affrc.go.jp/wcca2008/index.htm, (accessed 2010-12-27).

Fig. 9　Examples of cited references, one in Japanese and one in English (not both).

- Captions of tables and figures are either in Japanese or English, rarely both.
- Emphasis and Warichu are not used.

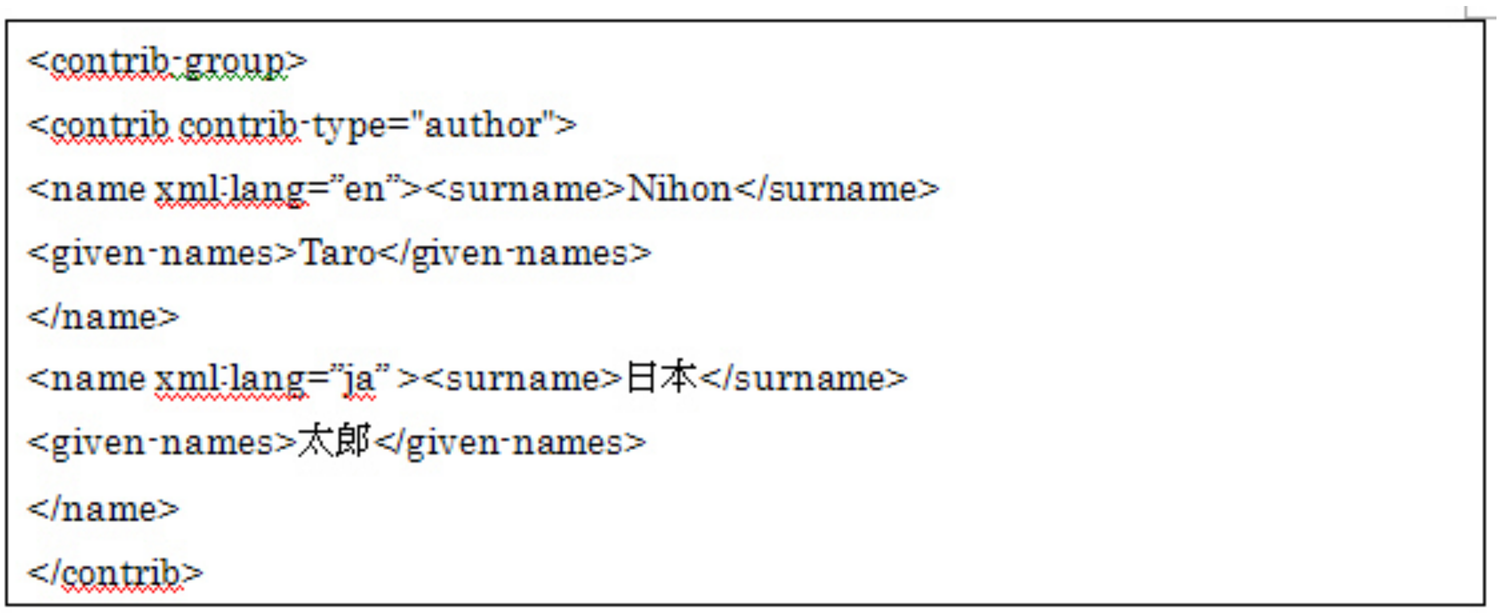
Most humanity, and many social science journals are published quite differently, in that:

- Vertical writing, always for humanities, often for social sciences
- Use of Emphasis and Warichu
- English-language article metadata are not common

Challenges in publishing bi-lingual contents

As discussed above, it is customary and recommended for Japanese-language articles to include both the Japanese and English metadata as well as abstracts. For Japanese-language articles, body texts are certainly in Japanese. Cited references are typically either in Japanese or English, but sometimes both. This situation requires developing a tag set capable of describing such bi-lingual contents.

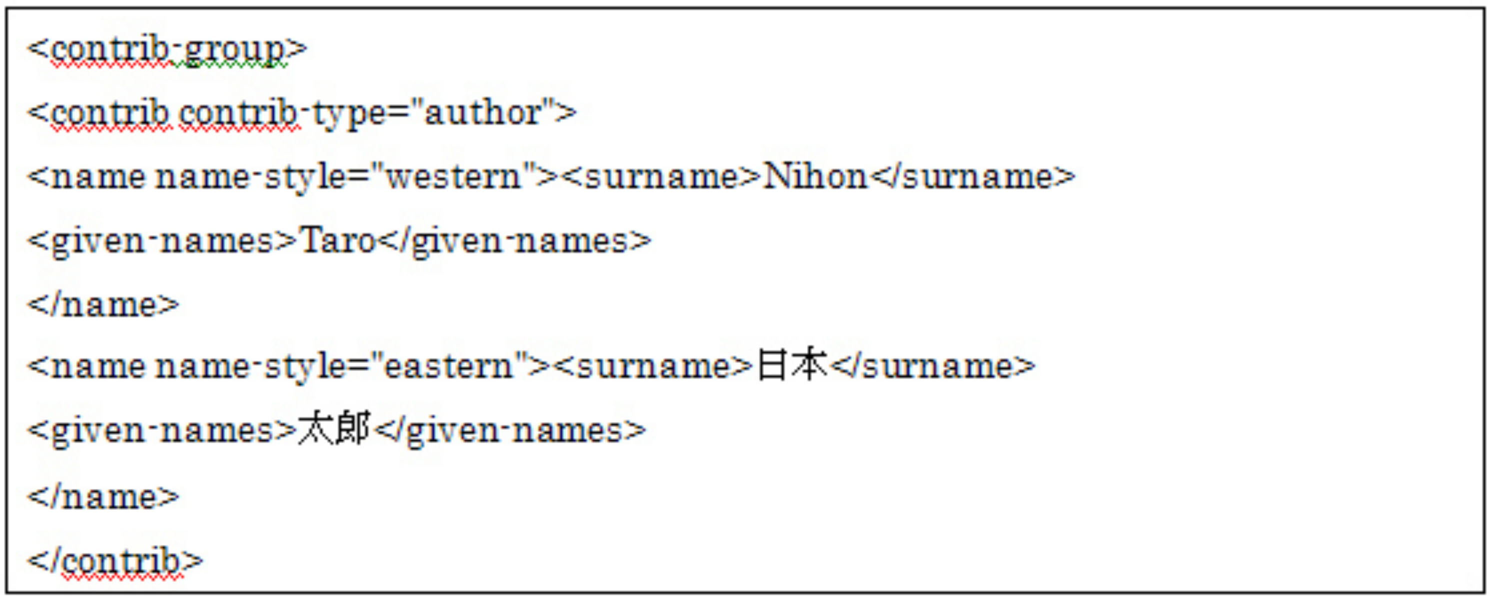
NLM DTD, that has been used widely to exchange scholarly article data, did not support such multi-lingual contents until its version 3.0. For example, as the <name> tag did not allow incorporating the xml:lang attribute, it was not possible to describe an author name in different languages.



```
<contrib-group>
<contrib contrib-type="author">
<name xml:lang=”en”><surname>Nihon</surname>
<given-names>Taro</given-names>
</name>
<name xml:lang=”ja” ><surname>日本</surname>
<given-names>太郎</given-names>
</name>
</contrib>
```

Fig. 10 Incvalid usage of @xml:lang under NLM DTD.

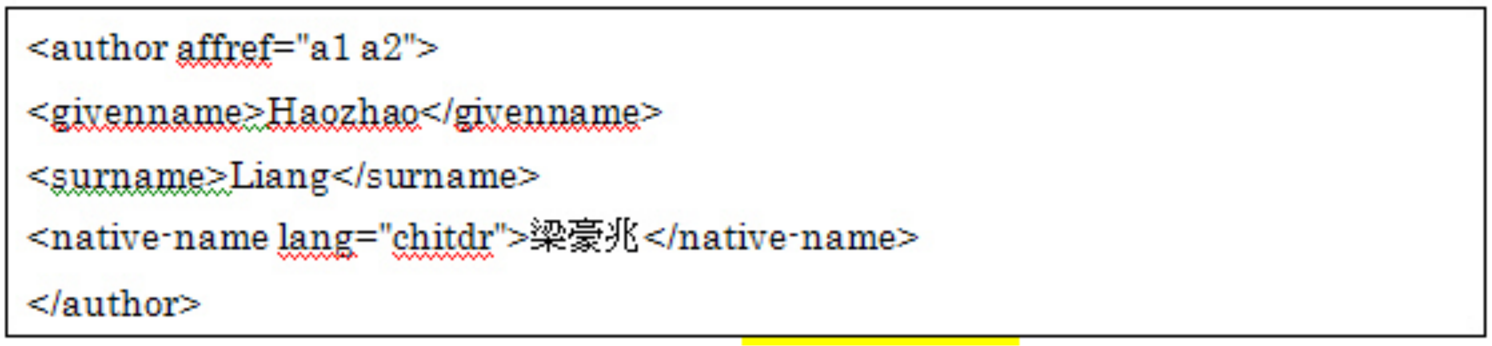
Some used @name-style to indicate that this particular author name is eastern or western, but this is not a right usage of this attribute.



```
<contrib-group>
<contrib contrib-type="author">
<name name-style="western"><surname>Nihon</surname>
<given-names>Taro</given-names>
</name>
<name name-style="eastern"><surname>日本</surname>
<given-names>太郎</given-names>
</name>
</contrib>
```

Fig. 11 Use of @name-style to designate name language.

Another introduced additional tag such as <native-name> to describe non-English description of foreign names.



```
<author affref="a1 a2">
<givenname>Haozhao</givenname>
<surname>Liang</surname>
<native-name lang="chitdr">梁豪兆</native-name>
</author>
```

Fig. 12 Use of <native-name>.

All the above were just bypassing. It was clear that we needed formal extension of NLM DTD to allow describing multi-language contents.

In addition, many elements, such as <kwd-group> and <publisher-name>, were not repeatable, so that it was not possible to describe in two languages. In addition, even for repeatable elements, such as <name> and <aff>, it was not possible to indicate that descriptions in two languages in fact belong to a single identity because of the lack of envelopes.

Scholarly Publishing Japan (SPJ) Working Group

The author was notified by Bruce Rosenblum, Inera, in early 2009 that the working group of NLM DTD was investigating the possibility of expanding the DTD for multi-language contents. The author thought that this is the great opportunity for Japanese publishers to contribute this initiative, and asked volunteers from publishers and typesetting companies in Japan to form a working group to discuss this issue. The working group, Scholarly Publishing Japan, was established in early 2010 and submitted proposals and sample data to the NLM DTD (later JATS) working group. Some of our proposals are as follows.

- NLM DTD should be as much as structural
- Support both multiple languages as well as multiple scripts using IETF RFC 5646, for example xml:lang="ja-Kana"
- Allow describing @xml:lang for <name>
- Introduce <subbody> to allow multiple language body texts
- Devise ways to describe cited references in multiple languages, for example, <compound-element-citation>
- Allow describing <journal-meta> data such as <journal-title>, <journal-subtitle>, and <abbrev-journal-title> in multiple languages
- Allow specifying <xref> for individual <name> elements so that the institution name in the same script may be linked directly

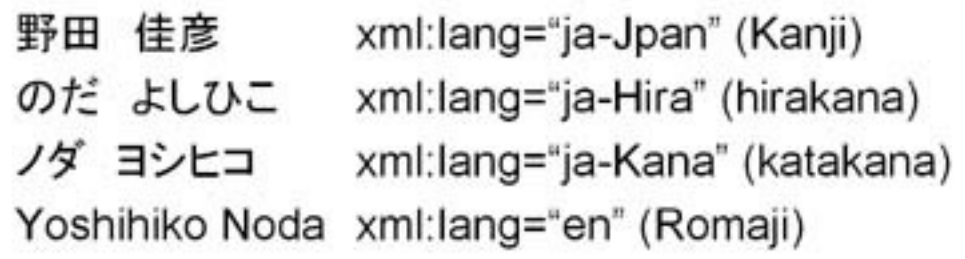
The activities of SPJ were fully discussed by the author in another paper [1].

NLM DTD 3.1 and JATS 0.4

NLM DTD working group reviewed SPJ's input, and developed NLM DTD version3.1 in September 2010. It became NISO JATS 0.4 in March 2011. The multi-lingual features of this version was fully described by Lapeyre and Usdin [5]. The outline is as follows.

- xml:lang is now usable for almost all the elements
- xml:lang is inherited down the XML document tree
- Both language and script may be recorded
- Most elements may be repeatable to describe multiple language expressions
- Wrapping tags are introduced to concatenate repeating elements for a single data, such as <name-alternatives> or <aff-alternatives>

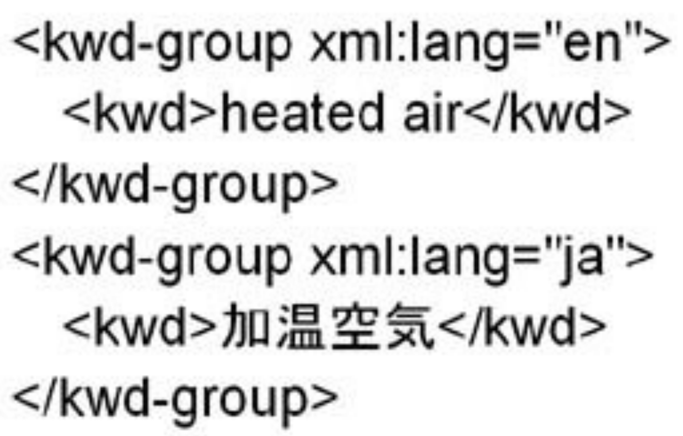
With the introduction of scripts, we can describe the previous name alternatives as follows.



野田　佳彦	xml:lang="ja-Jpan" (Kanji)
のだ　よしひこ	xml:lang="ja-Hira" (hirakana)
ノダ　ヨシヒコ	xml:lang="ja-Kana" (katakana)
Yoshihiko Noda	xml:lang="en" (Romaji)

Fig. 13　@xml:lang for Japanese scripts.

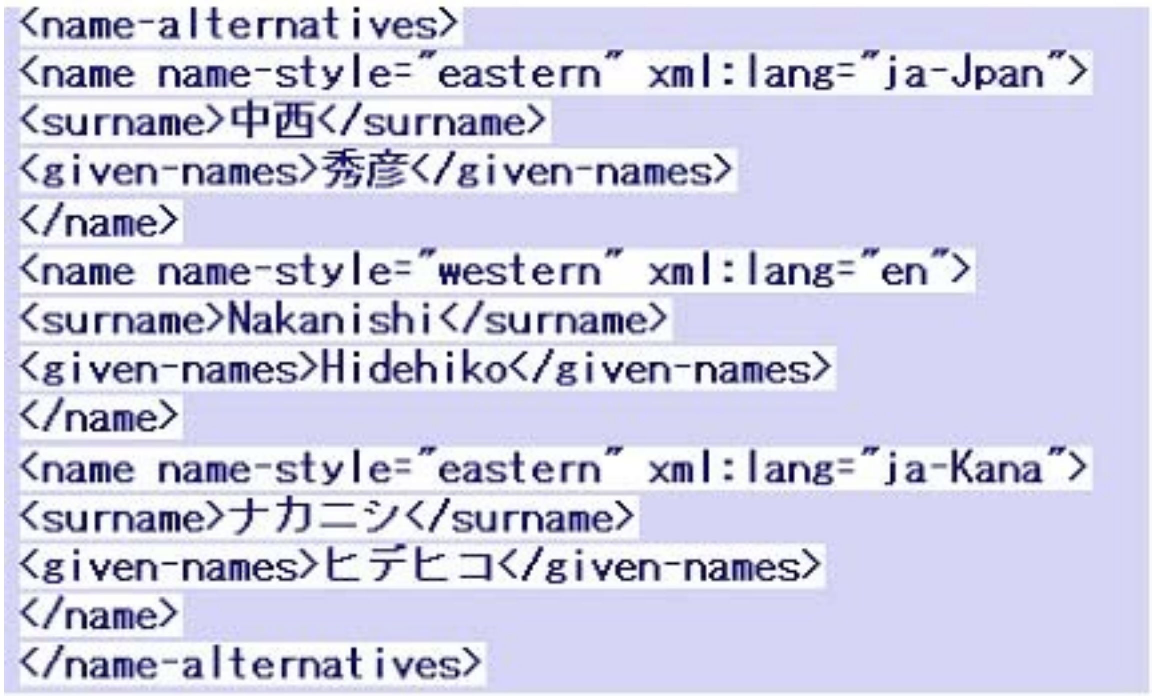
Repeatable elements allow describing the same keyword in different languages as in Figure 14.



```
<kwd-group xml:lang="en">
  <kwd>heated air</kwd>
</kwd-group>
<kwd-group xml:lang="ja">
  <kwd>加温空気</kwd>
</kwd-group>
```

Fig. 14　Use example of repeatable <kwd> element.

An example of JATS 0.4 description of a Japanese author and affiliation is shown in Figure 15.

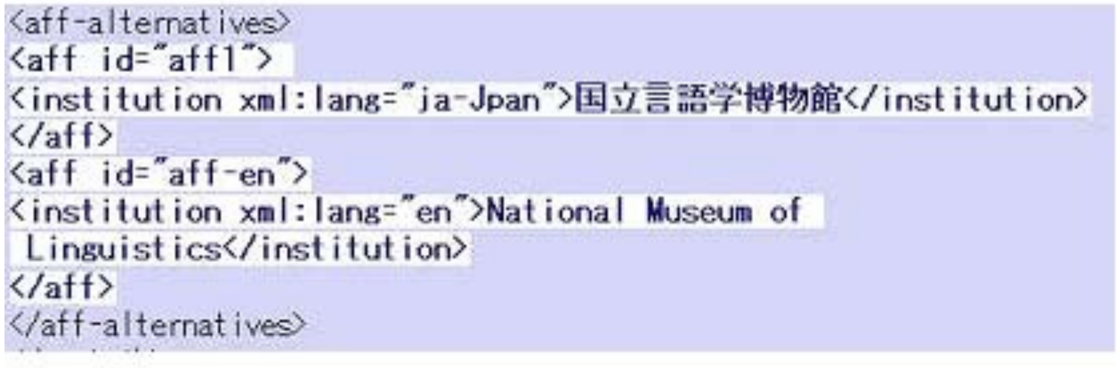


```
<name-alternatives>
<name name-style=”eastern” xml:lang=”ja-Jpan”>
<surname>中西</surname>
<given-names>秀彦</given-names>
</name>
<name name-style=”western” xml:lang=”en”>
<surname>Nakanishi</surname>
<given-names>Hidehiko</given-names>
</name>
<name name-style=”eastern” xml:lang=”ja-Kana”>
<surname>ナカニシ</surname>
<given-names>ヒデヒコ</given-names>
</name>
</name-alternatives>
```

Fig. 15　Description of a Japanese author name in JATS 0.4.

Here the author can be expressed as "中西 秀彦" in Japanese Kanji (ja-Jpan) script, as "ナカニシ ヒデヒコ" in Japanese Kana (ja-Kana) or phonetic script, and as "Hidehiko Nakanishi" in western alphabets.

Similarly, an institution may be express as "国立言語博物館" in Kanji, and "National Museum of Linguistics" in western alphabets as in Figure 16.



```
<aff-alternatives>
<aff id=”aff1”>
<institution xml:lang=”ja-Jpan”>国立言語学博物館</institution>
</aff>
<aff id=”aff-en”>
<institution xml:lang=”en”>National Museum of
 Linguistics</institution>
</aff>
</aff-alternatives>
```

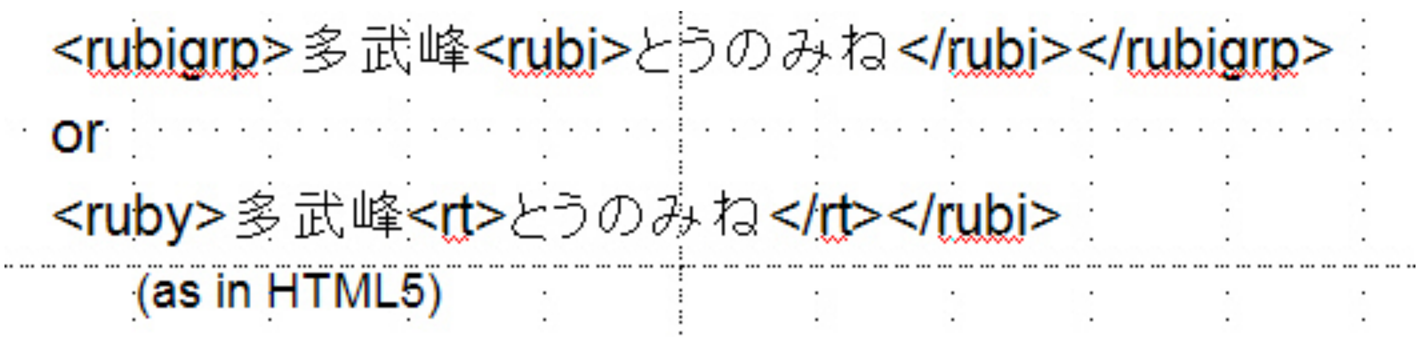
Fig. 16　Description of a Japanese instituion name in JATS 0.4.

As JATS 0.4 invited suggestions until the end of September 2011, we submitted additional comments as follows.

- Add <collab-alternatives> to describe collaborative authors in multiple languages. An example is "日本脳卒

中協会" in ja-Jpan and "Japan Stroke Association" in English.

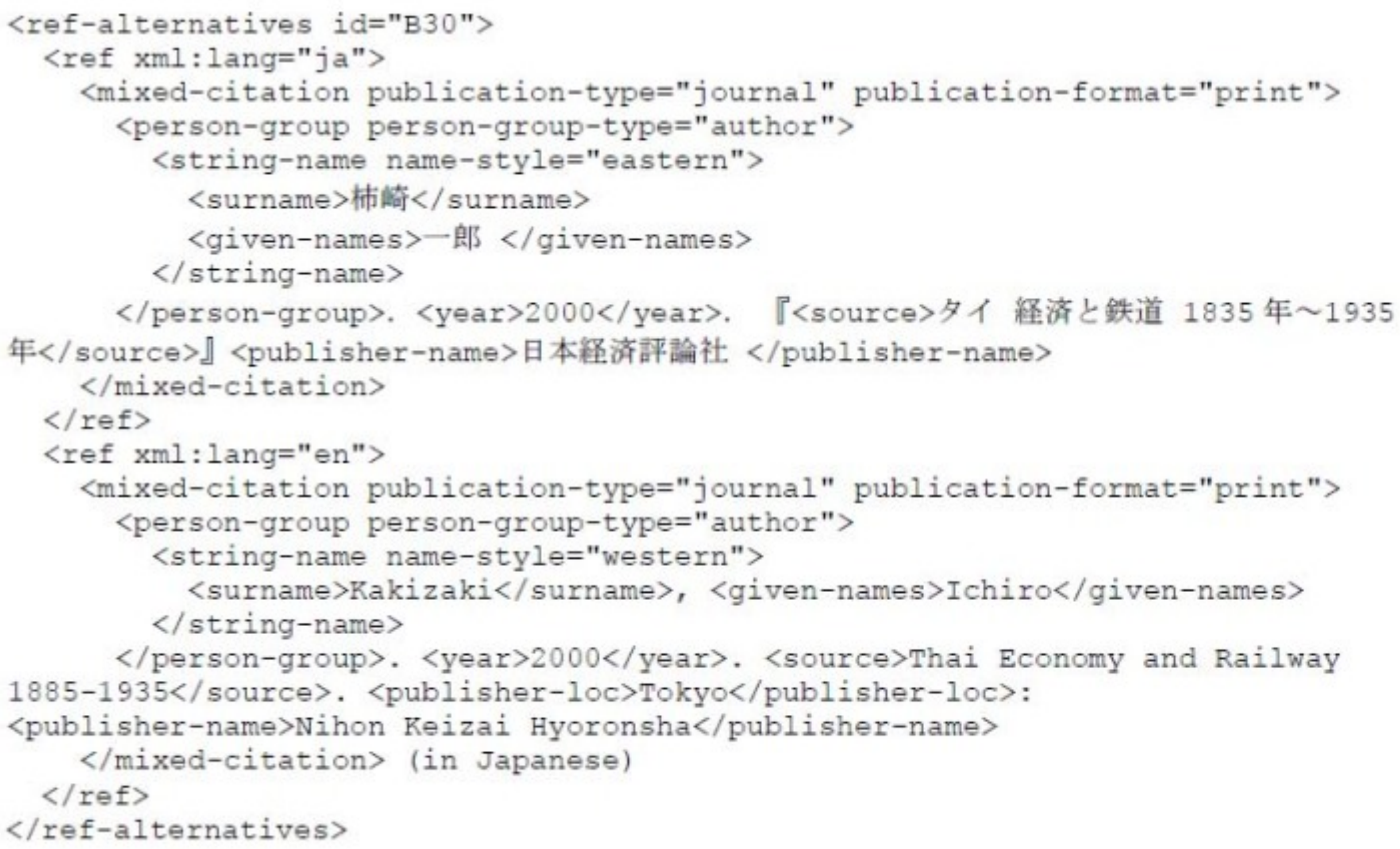
- Allow coding ruby. A suggested coding example is as in Figure 17.



```
<rubigrp>多武峰<rubi>とうのみね</rubi></rubigrp>
or
<ruby>多武峰<rt>とうのみね</rt></rubi>
    (as in HTML5)
```

Fig. 17　Suggested Rubi coding examples.

- Add <ref-alternatives>. There are cases where an English-language citation structure does not match with the Japanese-language counterpart, and thus coding elements individually in two language may be very confusing. A suggested example is in Figure 18.

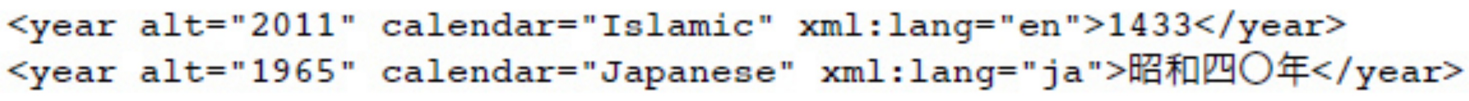


```
<ref-alternatives id="B30">
  <ref xml:lang="ja">
    <mixed-citation publication-type="journal" publication-format="print">
      <person-group person-group-type="author">
        <string-name name-style="eastern">
          <surname>柿崎</surname>
          <given-names>一郎 </given-names>
        </string-name>
      </person-group>. <year>2000</year>. 『<source>タイ 経済と鉄道 1835 年～1935
年</source>』<publisher-name>日本経済評論社 </publisher-name>
    </mixed-citation>
  </ref>
  <ref xml:lang="en">
    <mixed-citation publication-type="journal" publication-format="print">
      <person-group person-group-type="author">
        <string-name name-style="western">
          <surname>Kakizaki</surname>, <given-names>Ichiro</given-names>
        </string-name>
      </person-group>. <year>2000</year>. <source>Thai Economy and Railway
1885-1935</source>. <publisher-loc>Tokyo</publisher-loc>:
<publisher-name>Nihon Keizai Hyoronsha</publisher-name>
    </mixed-citation> (in Japanese)
  </ref>
</ref-alternatives>
```

Fig. 18　Suggested <ref-alternatives> example.

- Add "arXiv" to pubid-type

- Allow coding non-Gregorian years. Examples are as follows.



```
<year alt="2011" calendar="Islamic" xml:lang="en">1433</year>
<year alt="1965" calendar="Japanese" xml:lang="ja">昭和四〇年</year>
```

Fig. 19　Suggested non-Gregorian year coding.

The value "arXiv" for @pubid-type and @calendar for describing non-Gregorian years was included in the version 1.0 published on August 22, 2012. The need for <ref-alternatives> shall be fulfilled otherwise.

As more and more humanities articles get digitized, the needs for extention of JATS increase. Such issues are:

- Ruby (as discussed above)
- Emphasis
- Warichu

Developing a guideline for J-STAGE

Encouraged by the release of draft JATS 0.4 in March, 2011, JST began developing a guideline for J-STAGE based on the JATS 0.4 Journal Publishing Tag Sets. The outline of the guideline is as follows 6.

- Characters

  Characters are in UTF-8. Entity references such as, &amp;, &lt;, &gt;, &apos;, &quot;, and characters of ISO8879(SGML), MathML characters, and JATS specific characters such as &gcaron;, &Hmacr;, &euro;, and &franc;. may be used.

- Font attributes

  <bold>, <italic>, <monospace>, <roman>, <sans-serif>, <sc>, <overline>, <strike>, <sub>, <sup> and <underline> may be used.

- XML declaration and DOCTYPE

  version="1.0" encoding="UTF-8"

  <!DOCTYPE article PUBLIC "-//NLM//DTD JATS (Z39.96) Journal Publishing DTD v0.4 20110131//EN" "http://www.jstage.jst.go.jp/dtds/JATS-journalpublishing0.dtd">

- Journal meta

  <journal-id> and <issn> must exist.

  Characters

- Article meta

  In <article-id>, "publisher-id" of @pub-id-type should be JOI, or JST Object Identifier, not publisher's own id.

  <contrib-group> must exist.

  <name-alternatives> and <aff-alternatives> are required.

  <xref> id has to be in <aff-alternatives>.

- Mathematical formulas

  Text, graphic, MathML and Tex/LaTeX are supported.

- Figures and Tables

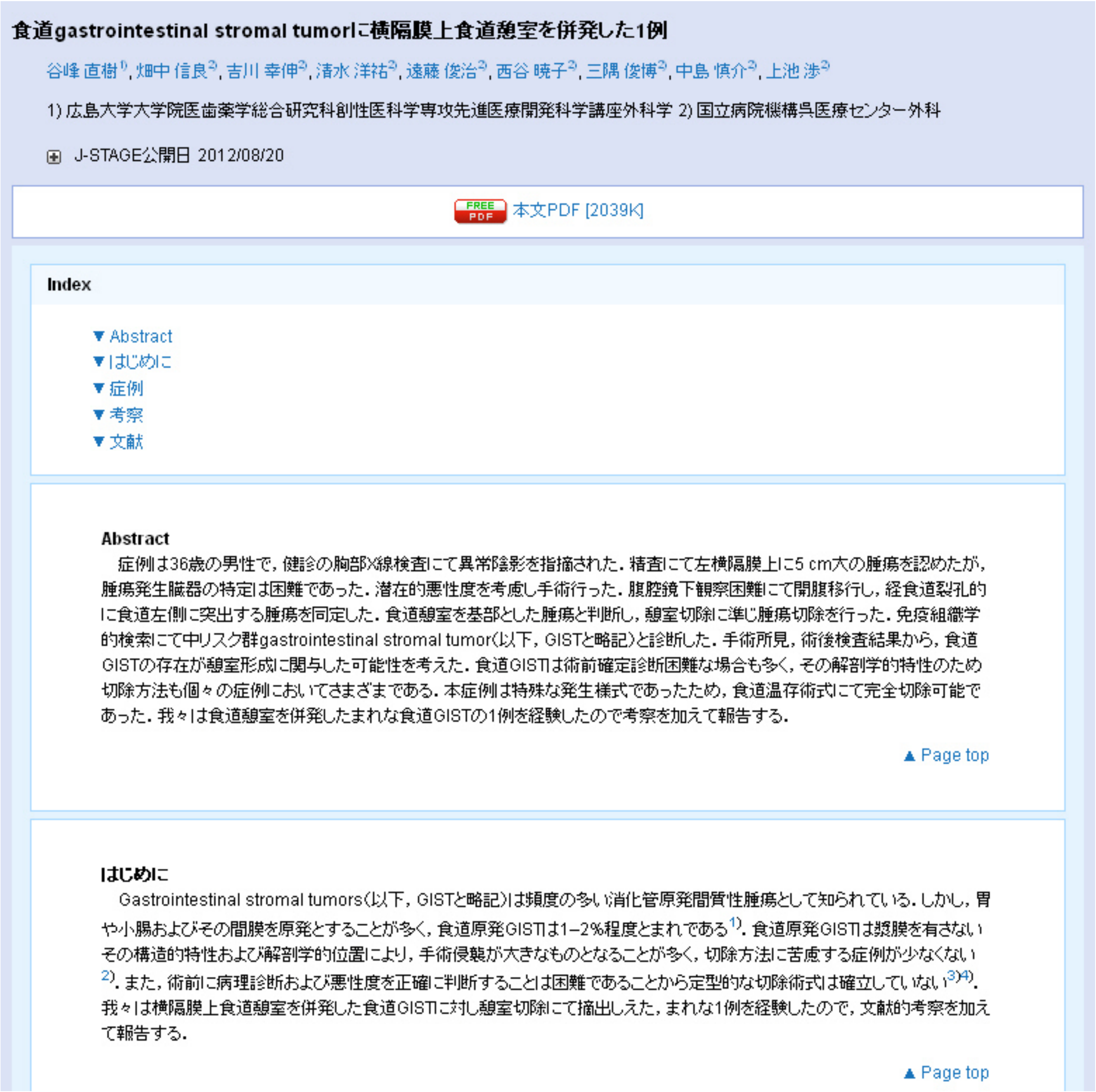
  They should appear where they are mentioned in the body text. Does not support OASIS CALS in the first version.

- References

  <mixed-citation> is supported, not <element-citation>. @xml:lang should be described in <ref> rather than in <mixed-citation>. Use English punctuations, such as ",", rather than Japanese punctuations, such as "，".

Launch of the new J-STAGE and implementation of XML

The new J-STAGE was launched in May, 2012 7. It hosts 1,658 journal titles (including title changes and merging) with 2,387,426 articles as of September 6, 2012. The first English-language XML-produced articles became online on July 13, 2012, for the journal, "Genes & Genetic Systems" 8, and the first Japanese-language ones (Figure 20 and 21) on July 18 for "Nippon Shokaki Geka Gakkai Zasshi", or the "Japanese Journal of Gastroenterological Surgery" 9.



食道gastrointestinal stromal tumorに横隔膜上食道憩室を併発した1例

谷峰 直樹[1], 畑中 信良[2], 吉川 幸伸[2], 清水 洋祐[2], 遠藤 俊治[2], 西谷 暁子[2], 三隅 俊博[2], 中島 慎介[2], 上池 渉[2]

1) 広島大学大学院医歯薬学総合研究科創性医科学専攻先進医療開発科学講座外科学 2) 国立病院機構呉医療センター外科

⊞ J-STAGE公開日 2012/08/20

本文PDF [2039K]

Index

Abstract

症例は36歳の男性で，健診の胸部X線検査にて異常陰影を指摘された．精査にて左横隔膜上に5 cm大の腫瘍を認めたが，腫瘍発生臓器の特定は困難であった．潜在的悪性度を考慮し手術行った．腹腔鏡下観察困難にて開腹移行し，経食道裂孔的に食道左側に突出する腫瘍を同定した．食道憩室を基部とした腫瘍と判断し，憩室切除に準じ腫瘍切除を行った．免疫組織学的検索にて中リスク群gastrointestinal stromal tumor(以下，GISTと略記)と診断した．手術所見，術後検査結果から，食道GISTの存在が憩室形成に関与した可能性を考えた．食道GISTは術前確定診断困難な場合も多く，その解剖学的特性のため切除方法も個々の症例においてさまざまである．本症例は特殊な発生様式であったため，食道温存術式にて完全切除可能であった．我々は食道憩室を併発したまれな食道GISTの1例を経験したので考察を加えて報告する．

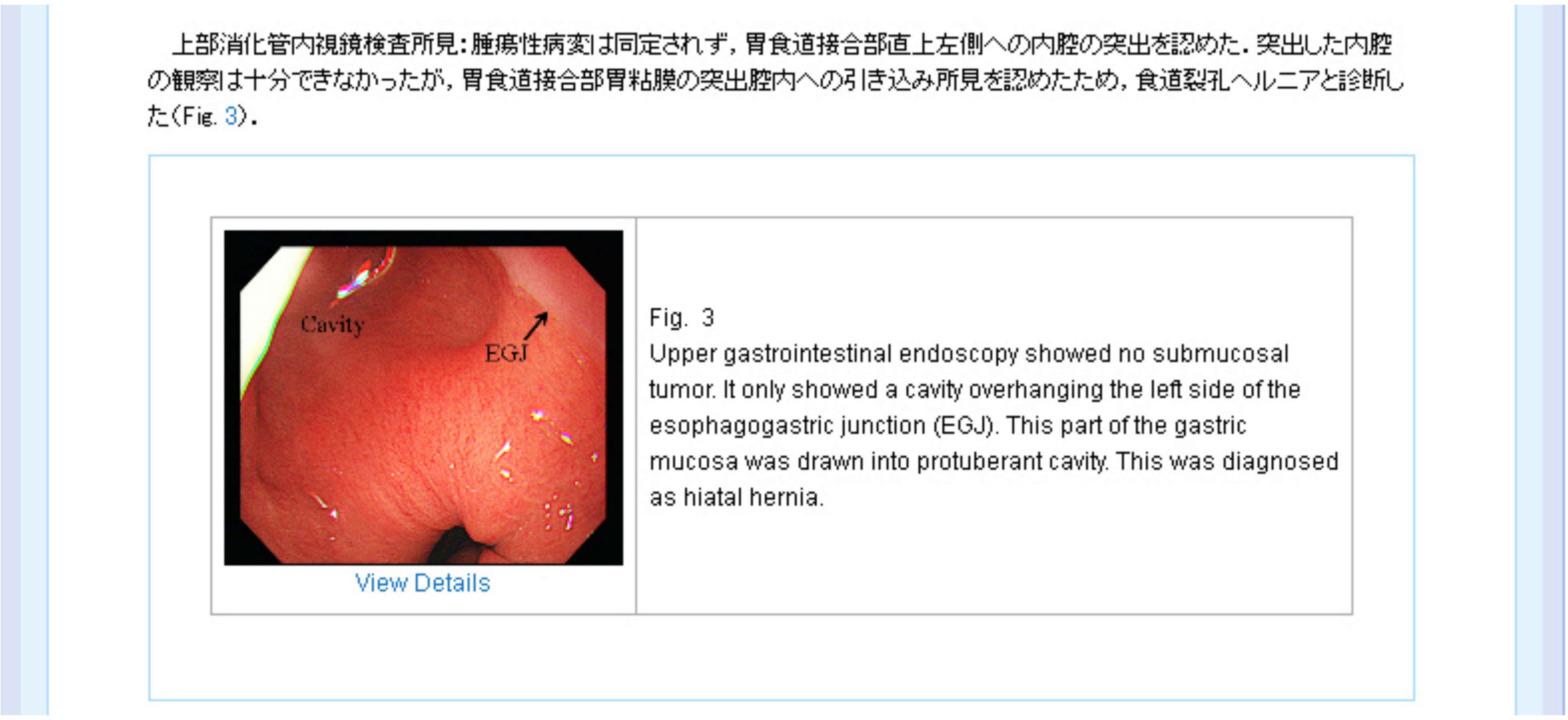
▲ Page top

はじめに

Gastrointestinal stromal tumors(以下，GISTと略記)は頻度の多い消化管原発間質性腫瘍として知られている．しかし，胃や小腸およびその間膜を原発とすることが多く，食道原発GISTは1–2%程度とまれである[1]．食道原発GISTは漿膜を有さないその構造的特性および解剖学的位置により，手術侵襲が大きなものとなることが多く，切除方法に苦慮する症例が少なくない[2]．また，術前に病理診断および悪性度を正確に判断することは困難であることから定型的な切除術式は確立していない[3][4]．我々は横隔膜上食道憩室を併発した食道GISTに対し憩室切除にて摘出しえた，まれな1例を経験したので，文献的考察を加えて報告する．

▲ Page top

Fig. 20 Japanese language articles on J-STAGE produced in XML ("Nippon Shokaki Geka Gakkai Zasshi") (1)



上部消化管内視鏡検査所見：腫瘍性病変は同定されず，胃食道接合部直上左側への内腔の突出を認めた．突出した内腔の観察は十分できなかったが，胃食道接合部胃粘膜の突出腔内への引き込み所見を認めたため，食道裂孔ヘルニアと診断した（Fig. 3）．

View Details

Fig. 3
Upper gastrointestinal endoscopy showed no submucosal tumor. It only showed a cavity overhanging the left side of the esophagogastric junction (EGJ). This part of the gastric mucosa was drawn into protuberant cavity. This was diagnosed as hiatal hernia.

Fig. 21 Japanese language articles on J-STAGE produced in XML ("Nippon Shokaki Geka Gakkai Zasshi") (2)

Production Workflow

Typesetter A

It generates JATS XML from MS Word 2007 XML via its own program. The XML is converted to XSL-FO using an XSLT stylesheet, and then PDF is produced from XSL-FO using the FO rendering engine, both on AH Formatter of Antenna House. Tables are in XHTML. Any corrections resulted from author proofs are made using an XML editor. The typesetting generally goes well including placing figures and tables.

Typesetter B

It converts MS Word files to XML using eXtyles. Manual editing is needed for many Japanese-language elements, for example, personal names, cited references, etc. It also converts TeX to XML using its own program for English-language articles. The XML is fed to 3B2 for typesetting. 3B2 can handle tables in OASIS and mathematical formulas in TeX. 3B2 produces both JATS XML and PDF. Corrections are made on 3B2.

Typesetter C

It pastes texts to FrameMaker to typeset. The final texts exported from FrameMaker are then processed by eXtyles to generate XML, which is then converted to JATS XML using XSLT. FrameMaker will be replaced by Typefi/InDesign in the near future. Then, the Typefi XML (contents XML) will be converted to JATS XML, again using XSLT.

Typesetter D

It uses eXtyles (without text cleaning) to generate XML from MS Word Japanese-language file. It then feeds the XML to InDesign using Typefi to typeset. The InDesign XML is converted to JATS XML using its own stylesheet. Any corrections are made on InDesign.

Problems and Issues of J-STAGE XML Guideline

According to typesetting companies, the Guideline needs further review and fixes as follows.

- <name-alternatives> and <aff-alternatives>

  It looks like a wrong decision to require <name-alternatives> and <aff-alternatives>, even for English-language articles. This practice is not compatible with that of PMC.

- Unsupported <element-citation>

  As PMC's default is <element-citation>, not supporting it cause additional work in production.

- "Publisher-id" should be a publisher's own id, rather than JOI.

Publishers and typesetters hope these will be fixed in the near future.

Establishing Scholarly XML Publishing Association

The people who worked together at SPJ formed an organization called, "Scholarly XML Publishing Association (SXPA)" in June 28, 2012, to promote the use of XML, especially JATS XML, in scholarly publishing in Japan. Soichi Tokizane was elected as the President. It will have a symposium about XML publishing on September 19, 2012. This new organization will succeed SPJ as a contact point of the NISO JATS Working group in Japan.

What are next?

It was very impressive that JATS XML has been very quickly implemented via typesetters in variety of ways. I believe the future of XML scholarly publishing is very bright. Several challenges are still exist, however.

- J-STAGE has to be enhanced to take full advantage of XML. It is still premature.
- J-TAGE XML Guideline should be fixed to eliminate unnecessary restrictions and to become compatible with that of PMC.
- Processing humanity/social science articles is still challenging. Typesetting such articles needs more experiments and practices.
- Encouraging scholarly book publishers to use JATS XML is important. We hope JATS-compatible book tag set will be developed soon.

Acknowledgments

The author to thank Soichi Kubota, Japan Science and Technology Agency (JST), Nakanishi Printing Co., Ltd., Komiyama Printing Co., Ltd., Letterpress Co., Ltd. and J Prepress Co., Ltd. for their information. The author also thanks the member of the Scholarly Publishing Japan (SPJ) working group for their contribution, as well as members of NLM DTD (later JATS) working group for their help in developing proposals for multi-language DTD.

1. Tokizane Soichi. From NLM DTD to JATS: XML for scholarly articles in Japanese. Joho Kanri. 2011; 54(9): 555-67 [Japanese with English abstract].
2. Shirokizawa Yoshiko, Obara Michio, Omi Asako, Shimizu Takahiko. A prototype system for the application of XML to J-STAGE. Joho Kanri. 2001; 44(2): 113-24 [Japanese with English abstract].
3. Sato Ryuichi, Kubota Soichi, Aoyama Kota, Tsuchiya Eri. New J-STAGE system accelerates digitization and distribution of academic journals from Japan. Joho Kanri. 2012; 55(2): 106-14 [Japanese with English abstract].
4. Tokizane Soichi. Electronic journal titles in science, technology and medicine published in Japan : Changes from 2005 to 2008.. Joho Kanri. 2011; 54(1): 13-20 [Japanese with English abstract].
5. Lapeyre Deborah A, Usdin B Tommie. Introduction to Multi-language Documents in NISO JATS. Proceedings of the Journal Article Tag Suite Conference (JATS-Con). 2011. Sep.
6. XML Guideline. Japan Science and Technology Agency; 2012. Available from: https://www.jstage.jst.go.jp/pub/html/AY04S230_ja.html [Japanese].
7. J-STAGE. Japan Science and Technology Agency. Available from: https://www.jstage.jst.go.jp/.
8. Genes & Genetic Systems. Genetics Society of Japan. Available from: https://www.jstage.jst.go.jp/browse/ggs.
9. The Japanese Journal of Gastroenterological Surgery. The Japanese Society of Gastroenterological Surgery. Available from: https://www.jstage.jst.go.jp/browse/jjgs/-char/ja.

Bookshelf ID: NBK100380