EDITEUR

ONIX for Books

Product Information Message

Application Note: Metadata for translated texts in ONIX

 doi:10.4400/eqah

Literature in translation is a growth area for the global book industry, and expansion of the international rights trade has resulted in many more translated non-fiction and children's titles on the market too. So how does ONIX handle translations and the metadata that describes them?

Language the book is in

ONIX metadata should be very clear on the language the book is printed in, using the <Language> composite. The same composite can also be used to specify the language a translated work was *originally* written in. The example is a work written originally in Japanese, but available in English translation:

```
<Language>
    <LanguageRole>01</LanguageRole>          <!-- book is in -->
    <LanguageCode>eng</LanguageCode>         <!-- English -->
</Language>
<Language>
    <LanguageRole>02</LanguageRole>          <!-- translated from original -->
    <LanguageCode>jpn</LanguageCode>         <!-- in Japanese -->
</Language>
```

Translators

Translators should usually be credited as contributors (and should *certainly* be credited in the ONIX if their contribution is noted on the book itself):

```
<Contributor>
    <SequenceNumber>2</SequenceNumber>
    <ContributorRole>B06</ContributorRole>              <!-- translated by -->
    <FromLanguage>jpn</FromLanguage>
    <ToLanguage>eng</ToLanguage>
    <NameIdentifier>
        <NameIDType>16</NameIdentifier>                 <!-- ISNI -->
        <IDValue>0000000109612546</IDValue>
    </NameIdentifier>
    <PersonName>Michael Emmerich</PersonName>
    <!-- other forms of the name can be included here -->
    <BiographicalNote>Michael Emmerich is the translator of more than a dozen
        works by Japanese writers such as Yasunari Kawabata and Banana
        Yoshimoto, as well as an editor of books for students of
        Japanese.</BiographicalNote>
</Contributor>
```

Subjects and Audiences

In principle, where the work in its original language has a set of subject and audience codes (and with *Thema,* any subject qualifiers), those same subject and audience codes should apply to the translated work too. Each audience code, subject code and qualifier should be reviewed, but in almost all cases they should remain the same. Very occasionally, it may be appropriate to remove one or two subject codes, or to modify <Audience>, particularly if the intent is to position the translation differently in the

market from the original – for example a book originally written for adults may be aimed at teens when translated, or *vice versa*.

Additional subject codes may also be appropriate – for example *Thema's* FYT (Fiction in translation) or YFZT (Children's and Teenage fiction: stories in translation):

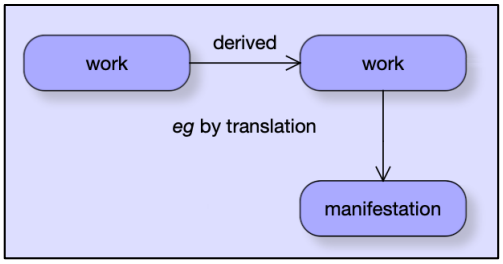
```
<Subject>
    <SubjectSchemeIdentifier>93</SubjectSchemeIdentifier>
    <SubjectSchemeVersion>1.6</SubjectSchemeName>
    <SubjectCode>FYT</SubjectCode>            <!-- fiction in translation -->
</Subject>
```

Note that there is no 'non-fiction in translation'. The FYT and YFZT codes are a convenience for resellers who display translated fiction separately from fiction in their local language.

Works

In ONIX, each product is a manifestation of a work. Several products – a hardback, a paperback or an e-book, all containing the same content – are different manifestations of the same work. When a work is translated, this creates a new 'derived work', so the translated product can be linked to *two* works, the derived work (of which it is a manifestation) and the original work (which was translated to create the derived work):

```
<RelatedWork>
    <WorkRelationCode>01</WorkRelationCode>       <!-- translated product is a
                                        manifestation of this 'parent' work -->
    <WorkIdentifier>
        <WorkIDType>15<WorkIDType>   <!-- work is identified using the ISBN of
                                          the first-published version as a proxy
                                          for a 'real' work ID -->
        <IDValue>9780001234567</IDValue>
    </WorkIdentifier>
</RelatedWork>
<RelatedWork>
    <WorkRelationCode>02</WorkRelationCode>     <!-- this links the translated
                product to the original (untranslated) 'grandparent' work -->
    <WorkIdentifier>
        <WorkIDType>01<WorkIDType>               <!-- work is identified using
                                                  a proprietary identifier -->
        <IDTypeName>JapanesePub KK Work ID</IDTypeName>
        <IDValue>5fb1c5f7-9604-4722-b450-75f893c9189f</IDValue>  <!-- UUID -->
    </WorkIdentifier>
</RelatedWork>
```



The above relationships can be diagrammed like this, where the ONIX record describes the manifestation – the translated product – at the bottom. The product is linked upwards to its parent work (the translated work, top right) using Work relation code 01, and 'up then across' to its grandparent (the original language work, top left) using relation code 02.

> [In the terms used in the notes attached to codes 01 and 02, the ONIX Product record describes the manifestation – Product A. A is a manifestation of Work X (code 01), and is also a manifestation of a work which is itself derived from work W (code 02).]

In fact, the simple relation code 02 does not specify the *nature* of the 'translation' link between the two works, only that the derivation link exists. But a different code can be used if there's a need to be more precise – the 'up and across' relation can use code 29 instead of 02 – and 29 specifies the derivation is by translation.

Unfortunately, there is no global standard work identifier – no equivalent of the ISBN for works – so publishers must use proprietary identifiers such as a UUID (in the second example above) or use an ISBN of a particular manifestation as a proxy in place of a 'real' work identifier (as in the first example above – this is sometimes termed a 'head ISBN' or a 'title ISBN' by publishers). It is a good practice for publishers to exchange these IDs when they conclude rights deals (for example when buying or selling translation rights) in order that the network of derived works can be 'joined up' internationally.

Translated titles and transliterated author names

Of course, the title of the translated product will almost always be different from the title of the original – it will be in a different language and, potentially, in a different script. But in fact, the original title can also be included in the ONIX too, and may occasionally improve discoverability and clarity of cataloging. Author names are generally not 'translated', but are often transliterated into the script used by the translated language – so author 吉本 ばなな (it is her pen name) becomes Banana Yoshimoto (more correctly, it becomes Yoshimoto Banana, but her name is usually reversed to fit Western name conventions where the family name [Yoshimoto] is last). In ONIX, this might be expressed as [1] [2] [3]:

```
<Contributor>
    <SequenceNumber>1</SequenceNumber>
    <ContributorRole>A01</ContributorRole>                <!-- written by -->
    <NameType>01</NameType>                                <!-- pseudonym -->
    <NameIdentifier>
        <NameIDType>16</NameIDType>                              <!-- ISNI -->
        <IDValue>0000000121446472</IDValue>
    </NameIdentifier>
    <PersonName>Banana Yoshimoto</PersonName>
    <PersonNameInverted>Yoshimoto, Banana</PersonNameInverted>
```

[1] The <NameType> codes in the example are somewhat problematic, as the first name is both a pseudonym *and* has been Romanised (transliterated), the Japanese version is also a pseudonym, and the author has previously also expressed her Japanese pseudonym as よしもと ばなな. The Name type code is only mandatory within <AlternativeName> so it may be clearer if the first instance of <NameType> were omitted. And as she has ISNIs for both her pen name and real name, the second <AlternativeName> could include a second ISNI too.

[2] Note the contrasting treatment in <PersonName> and <PersonNameInverted> of her Westernised name and the two Japanese names – the Westernised name is given name then family name in <PersonName>, and inverts to family name then given name in <PersonNameInverted>. Japanese names are presented as family name then given name in both <PersonName> and <PersonNameInverted>. The same typically also applies to Chinese, Korean, Hungarian and some other naming traditions.

[3] The *textscript* attribute can optionally be added to the various name fields if there is any doubt, and names in non-alphabetic (*eg* East Asian) writing systems can also include a *collationkey* attribute for sorting purposes. Typically, for Japanese names, the collation key is phonetic Hiragana, and so it can also be used as a gloss (furigana) when *displaying* the name. Glosses in other textual data fields that allow (X)HTML can be included using <ruby> markup tags, and for fields that *don't* allow (X)HTML, Unicode interlinear annotation delimiters can be used – see section X.13 *Extended character sets and encoding declarations* within the ONIX *Specification*.

```
        <AlternativeName>
            <NameType>01</NameType>                    <!-- also a pseudonym -->
            <PersonName collationkey="よしもと ばなな">吉本 ばなな</PersonName>
                                    <!-- but this time in Japanese script, a mix
                                    of Kanji and Hiragana. Note the collationkey
                                     attribute for phonetic sorting purposes. It
                                  could be added to <PersonNameInverted> too -->
            <PersonNameInverted>吉本 ばなな</PersonNameInverted>
        </AlternativeName>
        <AlternativeName>
            <NameType>04</NameType>                          <!-- 'real' name -->
            <PersonName collationkey="よしもと まほこ">吉本 真秀子</PersonName>
                   <!-- also in Japanese script, with a collationkey attribute -->
            <PersonNameInverted>吉本 真秀子</PersonNameInverted>
        </AlternativeName>
        <BiographicalNote textformat="05"><p>Banana Yoshimoto is the pen name of
            Japanese author Mahoko Yoshimoto (吉本 真秀子), whose work includes the
            novels <cite>Kitchen</cite> and <cite>Goodbye
            Tsugumi</cite>.</p></BiographicalNote>
    </Contributor>
```

Similarly, an alternative <TitleDetail> composite could be added to give the original Japanese title alongside the English title, but in the example case – a book titled *Goodbye Tsugumi* when published in English translation – the Japanese title is actually *TUGUMI* (explicitly in all-caps Latin characters! [4]):

```
    <TitleDetail>
        <TitleType>01</TitleType>          <!-- title on the (translated) book -->
        <TitleElement>
            <TitleElementLevel>01</TitleElementLevel>
            <NoPrefix/>
            <TitleWithoutPrefix>Goodbye Tsugumi</TitleWithoutPrefix>
        </TitleElement>
    </TitleDetail>
    <TitleDetail>
        <TitleType>03</TitleType>.             <!-- title in original language -->
        <TitleElement>
            <TitleElementLevel>01</TitleElementLevel>
            <NoPrefix/>
            <TitleWithoutPrefix textcase="01">TUGUMI</TitleWithoutPrefix>
        </TitleElement>
    </TitleDetail>
```

Publication dates

The various publishing dates of a translated product are entirely conventional: the Publication date is a nominal date on which it becomes available to purchase and obtain (nominal, because many books are delivered to book stores and go on sale immediately, potentially a few days before this 'official' date – but it is this publication date that is used for planning and many administrative purposes). If the book *must not* go on sale even one or two days early, then a Sales embargo date should be set as

[4] Note the *textcase* attribute which signals that the title is *deliberately* all caps, not because of some system limitation (in the rare case of a system limitation, it would be *textcase* code 03). *Language* and *textscript* attributes can optionally be added to the title text to avoid any doubt.

well (this sales embargo date is sometimes called a 'strict on sale date'). A third potential date in the ONIX could be the 'Date of first publication' – the date of publication of the earliest manifestation of the work (see the section above on Works). For a trade product, a paperback, this 'date of first publication can often be the publication date of the equivalent hardback,

But for books in translation, there is another useful date that can be set -- the Date of first publication *in the original language*:

```
<PublishingDate>
    <PublishingDateRole>01</PublishingDateRole>      <!-- publication date -->
    <Date>20030606</Date>                    <!-- of the product itself (a -->
</PublishingDate>                               <!-- paperback in English) -->
<PublishingDate>
    <PublishingDateRole>11</PublishingDateRole> <!-- date of first publication
                 (in this case, publication date of the English hardback) -->
    <Date>20020801</Date>
</PublishingDate>
<PublishingDate>
    <PublishingDateRole>20</PublishingDateRole> <!-- date of first publication
                   in the original language (ie of the Japanese hardback) -->
    <Date>19890801</Date>                          <!-- a dozen years prior to
                                              publication of the transation -->
</PublishingDate>
```

Other metadata

Of course, the other metadata will also change for the translated product, albeit *not because it is a translation* – the translated product will have a different ISBN, may well have a different physical size, different prices, most likely a different publisher and publication date, different sales rights and so on. The description and other collateral will also be different [5].

Graham Bell
EDItEUR
19th March 2026

[5] See also EDItEUR's application note *Multilingual metadata in ONIX* (doi:10.4400/knpd) when inclusion of textual metadata in both the original and the translated language in a single Product record would be useful.