ONIX International Steering Committee Minutes

Wednesday 15ᵗʰ March 2017, 1:30–15:00 GMT (UTC+0),
The Dark Room, Olympia, London

Attendees:

Luc Audrain (Hachette Livre)
Veronique Backert (Dilicom)
Graham Bell (EDItEUR)
Marie Bilde Rasmussen (Pruneau)
Maria Börman (Bokinfo)
Pierre Boudigues (Cercle de la Librairie)
Francis Cave (EDItEUR consultant)
Michel Cervellin (BTLF)
Fride Fosseng (Bokbasen, chair)
Alain Fournier (BTLF)
Bente Francke-Sætervoll (Bokbasen)
Noah Genner (BNC)
Alex Haffner (MVB)
Anna Lionetti (AIE)

Karina Luke (BIC)
Martin Lüning (VLB)
Renek Mendruń (BN Poland)
Brian O'Leary (BISG)
Dominique Parisis (CLIL)
Jesús Peraita (FGEE / DILVE)
Christer Perslöv (Bokinfo)
Simonetta Pillon (Informazioni Editoriali)
Vincent Poulvelarie (Dilicom)
Suzanne Rosario (EDItEUR)
Sophie Salmon (CLIL)
Chris Saynor (EDItEUR)
Henan Sun (Ingenta)
Howard Willows (Nielsen)

Summary of decisions in this meeting:

- Minutes of previous meeting approved
- Ratification of the proposals for Issue 37 of the codelists, and agreement not to port documentation of any clarifications back to earlier issues

1. Welcome and introductions

FF call the meeting to order and asked the attendees to introduce themselves. She welcomed all attendees to the committee.

2. Minutes of ISC Meeting held 19ᵗʰ October 2016, Frankfurt Book Fair, and matters arising

FF introduced the draft minutes of the previous meeting, which had been circulated shortly after the meeting and then made available on the EDItEUR, and asked whether there were any comments. **There were no comments and the minutes were approved**.

3. Current ONIX development work

FF asked GB to introduce his activity report.

He first highlighted the end of the process for phased withdrawal of support for ONIX 2.1 (which was initially agreed by the committee in October 2011 and which began with the announcement of planned sunset in January 2012. The main sunset event – withdrawal of development and support for the 2.1 schema – occurred at the beginning of 2015, while post-sunset 'twilight' support for new entries in the codelists continued until the beginning of 2017, as requested by (*inter alia*) BISG and agreed by the ISC. Codelist Issue 36, released in January 3017 is fully compatible with both 2.1 and 3.0.

GB highlighted that 2.1 remains fully compatible, provided you do not update beyond issue 36 of the codelists, and documentation remains available in the archive section of the EDItEUR website. Issue 37 and beyond will not support 2.1. He also noted that the EDItEUR website still receives around 250K http

requests *per month* for the ONIX 2.1 DTD files, as ONIX users continuing to use 2.1 have not updated their validations, and he asked National Group attendees to request that national group members check that they are not making large numbers of automated requests to the EDItEUR website, or that they follow the advice provided to install local copies of the 2.1 DTDs and XSDs (see http://www.editeur.org/15/Archived-Previous-Releases/#2.1%20After%20sunset).

Progress on adoption of ONIX 3.0 has continued, and GB highlighted three stories, in Germany, UK and Spain. In Germany, VLB began to accept ONIX 3.0 in September 2016, and historically it has been a successful driver of change. In the UK, Nielsen reported that two-thirds of *new* feeds are now ONIX 3.0, compared with just a little less than half in October 2016. In response to a question from BFS, HW clarified that this was *data in*, and that Nielsen's new system to be launched later in 2017 will provide 3.0 *data out*. Thirdly, in Spain, JP had reported in January 2017 that DILVE converted its own internal data structure from 2.1-compatible to 3.0-compatible. This affected over 900K records.

Codelists issues 35 and 36 were approved and published. One key new code enabled carrying a chunk of JSON-LD data within the ONIX data. This was, GB explained, intended to allow publishers to provide 'ready-made' schema.org data to retailers, who can include it on their 'buy' page for SEO purposes. Two caveats – first, the syntax is tricky and second there are security issues if retailers do not inspect the JSON-LD to confirm it contains only data (it is possible to embed executable JavaScript too, which represents a security risk). Fuller documentation of the schema.org snippets is expected over the summer as output from a BISG working group (which GB admitted, had been significantly delayed by pressure of EDItEUR work).

The ONIX 3.0 XSD schema (and the equivalent RNG) can be used to validate the structure and codelist content of an ONIX message. However, it cannot check all data, so many large-scale ONIX users check data content at application level. GB reported that EDItEUR has been exploring XSD version 1.1, which allows further checking of the data (for example, checking of identifier check digits, or various co-occurrence constraints). XSD 1.1 duplicates many of the things that Schematron does, but the experimental Schematron that is available for ONIX 3.0 remains an unusual part of the XML toolset and has not been adopted wisely. GB stressed that the XSD 1.1 will not replace the XSD 1.0, as not all XML validation tools are compatible with 1.1. FC noted that the extra checking in the XSD 1.1 should be documented in such a way that implementers can ensure the checking is compatible with any application-level checks they currently implement. AH noted that consistent plausibility checks at application level or database level would be a benefit.

LA suggested some collaborative contributions to such business rule checking. GB confirmed that EDItEUR would welcome receiving such contributions of XSD 1.1 rules (asserts), but that they must be part of the ONIX *Specification*, not proprietary requirements. Suggested rules could be described in abstract terms, or concretely as XSD 1.1 asserts.

HS asked about end-user validation tools. GB replied that the XSD provides the template to be used within a validation process, but it is likely to be run by more technical users using an XML software library (*eg* xmllib, Saxon) or within an XML editor application (*eg* oXygen, XML Spy). End users could use simple tools like Notepad++. HS also asked whether any representatives had experience dealing with ONIX in an environment where ISBNs are widely misused (*eg* same ISBN on many different books). GB replied that while duplication of ISBN is very poor practice within the supply chain, in ONIX it is less of an issue: although you need a unique identifier for each product, ISBN is not the *only* identifier that can be used.

ONIX training has been a part of EDItEUR's programme for several years, and EDItEUR is keen to extend the scope and volume of the training – each course means a few more people have an improved understanding of the standard and its benefits. Over the previous period, EDItEUR has delivered courses in London (for BIC), New York, Washington, Seattle, San Francisco (for BISG), Frankfurt (for MVB) and in-house at two EDItEUR members. Seattle in particular was an important session as it was requested specifically by Amazon. EDItEUR can now deliver training in French too, via CS.

GB reported that work on the online ONIX codelist browser was essentially complete, and that it would be launched within a couple of months. It is multi-lingual, so various language translations would be added over the summer. At present, there is no 'preferred format' for translations (EDItEUR could work with 'three columns in a spreadsheet'), but a more formal template would be produced in due course. He noted the equivalent *Thema* online browser is very popular, and expected similar popularity for the ONIX.

JP asked whether usage statistics were available. GB replied that the *Thema* log files were not analysed formally, but lists of successful and unsuccessful search terms used had been used in creating an initial set of synonyms added in early 2017. This is likely to be extended, with further synonyms being informed by real search behaviour. (GB noted that this is not likely to be as useful for the ONIX browser.)

In conclusion, GB suggested that members of the Steering Committee should begin thinking about some proposals for ONIX 3.0.4, most likely for release in early 2018. This is in addition to the planned Working Group on a POD/e-book specification Block within ONIX.

4.   Proposals for Codelists Issue 37

A short list of proposals for Issue 37 were circulated a month prior to the Book Fair. NOTE this update of the codelists will not be compatible with ONIX 2.1.

GB summarised the proposed additions to the lists, which were documented separately. The proposals were adopted.

BFS raised the question of whether *clarifications* such as those in List 45 added to the codelists at Issue 37, and argued that clarifications should *not* be ported back to earlier issues – or more specifically to the archived version of issue 36 for use with ONIX 2.1. The point of *not* porting back is to avoid sending a mixed message about ongoing support of ONIX 2.1. AH noted that there would be issues were the meaning of a code modified in any way in ONIX 3.0, but not changed in 2.1, and JP agreed. AF, AL, KL, LA, NG and others supported BFS's view. GB suggested that in practical terms these are only minor clarifications, and are unlikely to cause any real-world issues. AH asked whether there could be some solution with within the online browser, and GB agreed to consider how this could be dealt with.

The group questioned the value of the proposed code 09 in list 65 and asked EDItEUR to go back to the original proposer to ask for a real-world use case before including it in Issue 37. There was also extensive discussion of the slight modification to code 08 in List 145, concluding that the proposal was acceptable.

5.   Updates on key migrations and adoptions from national groups

AH reported on progress in Germany, including changes of the VLB data model which required conversion of over 6 million records, and noted there would be strong encouragement to all German publishers to update to 3.0 over the coming year – although inevitably VLB would continue to receive 2.1 for at least a couple of years. He described the German guidelines for delivery to VLB, noting that they supplement rather than replace the EDItEUR guidance, and noted the delivery of ONIX 3.0 data to VLB-tix alongside additional non-ONIX event information and digital previews.

VB described Dilicom's challenging project in France to collect ONIX 3.0, with the aim of reaching 100% 3.0 before the end of 2017. (This is alongside an equivalent effort for Thema). MBR reported that in Denmark, ONIX is mostly limited to the e-books supply chain – although it is all 3.0. Publishers and retailers are engaged with a project to improve physical supply chain data, and extending the use of ONIX is being considered, and awareness is growing, but no decision has yet been made.

HW reported that supply of 3.0 in the UK is accelerating. AL described the position in Italy – while IE-Online uses ONIX extensively, few publishers implement the standard. It is hoped that a wider range of services can be supplied to the publishers if they begin to adopt ONIX 3.0. AF reported that nearly 100% of French Canadian publishers send data to BTLF in ONIX 3.0, though wholesalers continue to use 2.1 (with only a gradual move towards 3.0). In Norway, BFS reported that ONIX 3.0 has been in widespread use since

before sunset, and there is no significant remaining use of 2.1. The current project is to persuade libraries to use ONIX to describe licence terms for e-book lending, and FF noted that the new <PriceConstraint> elements are being introduced. Training is an important part of this, to ensure that price and terms are accurately communicated.

JP had already reported the introduction of ONIX 3.0 into DILVE, but he noted the remainder of the supply chain is still using 2.1 or – in many cases – Excel spreadsheets. The market is moving very slowly, with only a handful of publishers using ONIX 3.0, and he suggested there is a disconnect between publisher's staff who handle metadata and the staff who make decisions. In response to a question from FC, he reported that all but a couple of ONIX users in Spain use in-house applications rather than off-the-shelf commercial applications, which may be a cause of the relatively slow speed of migration.

BOL reported that the US remains a mixed market, with much less than half of ONIX-capable publishers using the updated standard. CS had attended an event in New York which aimed to begin addressing the problem and BISG plans a programme to outline the business case for 3.0, to be presented to publishers later in 2017. BOL also noted that Ingram and Barnes and Noble still received over 1000 spreadsheet files from publishers each month, which presents a significant opportunity for streamlining the supply chain (irrespective of the *exact version* of ONIX).

NG echoed some of BOL's comments, and also noted that BNC had shut down its Excel-to-ONIX 2.1 converter in late 2016 in order to concentrate efforts on ONIX 3.0. The BNC data certification process will also be tweaked to reduce financial incentives for use of 2.1. But BNC has few downstream recipients for its 3.0 so far (excepting Kobo).

HS reported that the Chinese Govt had earmarked significant financial resources for development of ONIX capabilities, and 14 publishing houses had been selected as the first group to receive this funding and apply the standard, and 7 technology companies are currently integrating CN-ONIX into their ERP applications. The funding also covers work on ONIX-to-MARC transformation tools to support the library market. HS asked about whether ONIX would ever allow data to flow back from distributors to publishers. GB replied that there is a separate standard, EDItX, for returning sales information to the publisher.

6. **Any Other Business and Next Meeting**

FF asked for any other business. There being none, she advised the next ONIX ISC meeting would be during the Frankfurt Book Fair, October 11th 2017. FF and GB thanked the attendees and closed the meeting.

Graham Bell
EDItEUR
15th March 2017