ONIX for Books
International Steering Committee Minutes

Wednesday 11th April 2018, 13:30–15:00 GMT (UTC+0),
The Dark Room, Olympia, London

Attendees:


Summary of decisions in this meeting:

- Minutes of the previous meeting were approved
- Approval for Issue 41 of the codelists

1. Welcome and introductions

FF welcomed attendees to the meeting, and the attendees introduced themselves.

2. Minutes of ISC Meeting held 11th Oct 2017, Frankfurt Book Fair, and matters arising

FF introduced the draft minutes of the meeting, which and been previously distributed. There were no comments and the minutes were approved. GB explained for new attendees how the minutes are treated in the period between now and the next meeting of the committee.

3. Report on current ONIX development work

GB introduced his activity report, which had been distributed previously. First, he thanked Marie Bilde and Brian O’Leary for hosting and sponsoring the informal meeting at The Famous Three Kings the evening; other attendees agreed it was an enjoyable event and also thanked them. He then highlighted some of the main points:

a. The previous meeting at Frankfurt agreed a series of proposed additions to ONIX 3.0, and these were documented and published on the EDItEUR website as ONIX 3.0.4 at the end of October.
This included updates to the DTD, XSD and RNG schemas, and to the documentation in PDF and HTML.

b. GB noted that ONIX 2.1 was in the ‘midnight’ part of its life (following ‘twilight’ and ‘moonlight’ support, as the level of support was reduced after the end of 2014. However, he noted that a few data senders had begun to add codes from ONIX 3.0 in ONIX 2.1 messages (eg using codes from list 150 instead of list 7 in <ProductForm>). Such messages are not valid, and may cause technical problems at the recipient end – but DTD validation does not spot this error. This is one of the main reasons for using XSD validation. There have been confused calls to EDItEUR asking ‘what does this code mean?’, when the code does not actually exist (or exists only in a different codelist). Other than this, there have been no issues related to the (diminishing) use of 2.1.

c. There is some good news around the migration to ONIX 3.0, particularly in the US. Previously, this group requested BISG produce a plan to promote ONIX 3.0, in return offering to consider a one-off update to the ONIX 2.1 codelists (currently frozen at Issue 36), if such an update became necessary. BOL reported that readiness for ONIX 3.0 was surveyed across the BISG membership and beyond, with over 100 responses, and BISG has put in place a realistic plan for a series of three ‘metadata summits’ for both data senders and recipients, to discuss and promote better metadata practices. ONIX 3.0 will be a part of this, though not the whole. GB also reported that in the UK, for the first time, more than half of the organisations in receipt of BIC’s accreditation badges for quality and timeliness of metadata are providing ONIX 3.0, and that the revised BIC scheme (due by the end of 2018) will require 3.0 to achieve the highest level of award. Even at its lowest level, it requires the use of ONIX code values in Excel. KL noted that the revised scheme is almost final, but that assessment is still being worked on.

d. One large UK publisher (CUP) had been told by Amazon (.com) that its bespoke data feed to Amazon must be replaced by ONIX 3.0 ‘before the late summer peak’ (ie by September 2018).

e. Release of ONIX 3.0.4 was released alongside issue 39 of the codelists, and Issue 40 was released in mid-January 2018. These incorporated around two dozen new codes, and an equal number of clarifications to existing codes. One particular highlight was incorporation of the definition of ‘publication date’ from the ONIX 3.0 Implementation and Best Practice document into the notes for the codelists, which makes the definition much more visible. Codelists issue 41 will be considered later in the meeting.

f. GB noted the increasing range of ‘application notes’ added to the EDItEUR website. These are 4–6 page of explanation, Q&A or notes on very specific parts of ONIX. They are targeted at ONIX 3.0, but some are useful for 2.1 as well. One of the most recent covers open access monographs.

g. GB drew attention to the ONIX codelist browser (launched immediately prior to the 2017 Frankfurt Fair). He asked that national groups consider contributing translations (and partial translations) so EDItEUR can present them on the browser, with the aim that the browser becomes a ‘hub’ for multiple languages. He also noted that the browser (and its Thema equivalent) has now moved to https://ns.editeur.org/onix for better security and privacy.

h. GB introduced his work ‘strict’ schema, using XSD 1.1 technology. This strict schema can check rules that for example, validate the check digit of an ISBN, check that e-books do not have physical dimensions, or spot internal contradictions in the sales rights. He thanked those who had volunteered to test the strict XSD with their own ONIX files, and as a result of their testing the XSD is now robust enough for a ‘beta’ release. He stressed it is a useful tool now, though it remains a work in progress. GB noted that XSD works with Saxon, Xerces, Raptor parsers (and thus in oXygen and XML Spy) but does not work with the standard .NET parser or with libxml (which rules out a lot of free XML tools). FC asked whether the XSD only picks up errors, or whether it also provides warnings about best practice. GB first stressed that the strict XSD only checks rules that are already clearly part of the specification – it does not change the nature of ONIX 3.0, nor add any new requirements. But the strict XSD does include optional ‘embedded Schematron’ rules, which do check for deprecated codes and provide warnings. In theory, we could continue to add
warnings about key best practice points. LA asked how detailed such rules might become, and whether work on the strict schema will inform future versions of the main ONIX specification. GB said that the rules can be complex if necessary, and can incorporate logic such as ‘if the subject scheme is Thema, BISAC or BIC then there may be one (and only one) subject code flagged as the ‘main subject’, but if the subject scheme is Dewey, there may be only one subject code and no ‘main subject’ flag. But judgement will need to be used to ensure that rules are not added without sufficient justification – there’s no limit to the number of rules. SB asked whether the ruleset was extensible. GB explained that this was possible, but the architecture of XSD does not make it simple to extend – new rules need to be added within the body of the XSD, not in some separate file. The strict XSD is now available on the EDItEUR website (and the old Schematron schema has been removed).

4. **Proposals for Codelists issue 41**

GB introduced the proposals for Issue 41, which are separately documented, and explained the need for each of the new codes.

BFS queried the need for ‘display order’ (in collection sequence), and JP explained that this was for use when the suggested display order was different from the published order and the narrative or recommended reading order. GB gave an example – should an *hors de sèrie* book be shelved before or after the main sequence of books in the series? – but agreed to refer it back to the proposer before finalising the codelists.

He noted that while there were a very small number of new additions to the lists, but also a range of clarifications. These clarifications – for example, the clearer differentiation between multi-item and multi-component – are also being carried through to other EDItEUR documentation.

GB asked for any objections to ratification of the codelists (subject to the query about display order), and there were none. **Issue 41 was approved.**

5. **Expectations and timetable for further ONIX developments**

GB updated the committee on EDItEUR’s expectation for near-term work on ONIX:

a. He expected continued work on the strict XSD 1.1, because those using it say it is useful
b. EDItEUR will likely convene a working group on ‘production requirements’ and POD, and will also work on promotional events for a potential 3.0.5 update. He called in particular for national groups to nominate production experts to work on the production requirements working group

   i. BOL asked whether a potential production requirements block was preferable to working with XBits. GB noted that he had deliberately not looked at XBits in detail, but adoption was very limited even in the US, and most potential users were already in receipt of ONIX data – and for them, and one-stop solution is attractive. LA noted the new French clic.edit association, which is also working on standard messaging between publisher and production suppliers. GB noted the potential for clic.edit to cherry-pick parts of ONIX, and reminded LA that the licensing of ONIX should be carefully considered

6. **Updates on key migrations and adoptions from national groups**

FF asked for updates from each of the national groups.

NG reported good progress across Canadian publishers in producing ONIX 3.0 (more than 100 publishers on 3.0), but noted there was still no significant Canadian physical retailer able to accept it – though there
are Canadian e-book retailers that use 3.0 extensively. He also mentioned great interest in metadata issues, particularly keywords, at the recent TechForum conference.

CP reminded the group that the Swedish supply chain has been using ONIX 3.0 for nearly six years now (since 2012).

JP reported a decent increase in Spanish publishers using ONIX 3.0, including Random House, but among retailers there is no movement.

AF said that in the French Canadian market, almost everyone uses ONIX, but the main problem was to convince suppliers (i.e., data recipients) to move to 3.0, but he expected movement soon. AF added (later) that BTLF had hired a full-time resource to advise and train publishers and suppliers on a one-to-one basis.

HW described slow but steady progress towards ONIX 3.0 in the UK. Work on accreditation was reported earlier, but he reported that more than 50% of the new data feeds to Nielsen are now 3.0, which was not the case even a few months ago.

SB reported that in the Netherlands most work at CB was to ensure that publishers used the right fields for the right information, but progress was positive. CB is also discussing the adoption of Thema across the Dutch book trade.

AH also characterised progress in Germany as slow but steady, and looked forward to the upcoming meeting of the Product Metadata IG, where there is activity to promote 3.0. So far, only a few German publishers are supplying ONIX 3.0 to VLB (accounting for only 12,000 products), but there are a number of large publishers on the point of switching to 3.0 and the first data recipient for 3.0 files from VLB is going live. The more flexible marketing information is key to adoption in Germany. The IG is also producing best practices in German. The treatment of multi-item and multi-component products is a particular sticking point.

BFS said the Norwegian market is in the same position as Sweden – it has been working with ONIX 3.0 (only) for several years. The focus is now very much on the quality of the data, and Bokbasen has launched a new incentive scheme to encourage publishers to use more of the capabilities of ONIX. Bokbasen also runs regular ONIX training sessions for publishers and retailers.

VB reported that 90% of the book distributors in France use ONIX 3.0 (in this case, 3.0.0) now, and Dilicom is investigating whether it can update to a later subversion (e.g., 3.0.4). GB noted that 3.0.4 is fully back-compatible with 3.0.0, so any valid 3.0.0 is also (automatically) valid 3.0.4. VB explained that her issue was where data was supplied to Dilicom as 3.0.4 and needed to be sent to a recipient which could only accept 3.0.0. For this, the only solution is to drop any elements that cannot be expressed in 3.0.0. [One exception – sales restrictions may require work.] There is also ongoing work on creating French best practices. DP noted that some of this work within Commission FEL in cooperation with BIC. KL noted this took the form of some shared workshops.

OF noted that in Flanders, data suppliers were using a mix of ONIX 2.1 and 3.0 and Meta4Books was encouraging those on 2.1 to switch. Meta4Books itself is mostly using 3.0, and plans to phase out 2.1 completely. As with Norway and Sweden, the emphasis is on quality, richness and timeliness.

RC reported that MetaBooks was still less than a year old, so the emphasis was on collecting metadata – any metadata – rather than worrying about which version of the standard. However, those publishers using ONIX were exclusively using 3.0. MetaBooks is focusing on awareness and education at present.

BOL added to his earlier update. He noted that BISG has no requests for continued 2.1 support at this time. At BISG, there has been some discussion of the value of a central repository of metadata akin to that provided by BookNet Canada, but there is a realistic view that this would be a huge task to create, given the fragmented nature of the US market. He asked also whether other national groups could contact him if they had particular success stories they could share that grew from the adoption of 3.0.
MB reported that the Danish user group was now a year old, and had been using 3.0 particularly for digital books. Awareness is growing among traditional publishers that their physical and digital products are using different versions of the standard. The group is working on establishing a common national subset or baseline metadata requirement that everyone can support. GB suggested that BIC shares its work on PDEA with the Danish group, and reminded MB of the importance of recipients not rejecting anything that is outside the subset (but still valid ONIX 3.0). She agreed, and noted that the emphasis was on ensuring publishers delivered a richer data set than they were previously used to.

KL outlined the BIC ONIX 3.0 marketing plan that includes training, workshops and events provided in collaboration with EDItEUR, BIC’s social media presence, documentation such as BIC Bites and other elements. BIC will also be surveying the readiness of the UK market for ONIX 3.0, and will be launching an ambitious Metadata Map project to assess the whole range of metadata exchanges, issues and opportunities across the book trade.

LA reported that Hachette was working internationally (US, UK, France) to harmonise the ONIX 3.0 across the three markets. Hachette aims to produce very similar ONIX from three quite different IT systems. (GB remembered that he was tasked with doing something similar for HarperCollins to standardise the ONIX 2.1 across UK, US and AU, but left the company before making much progress.)

SP said that in Italy, IE-Online had been talking with Amazon, but there was no progress on 3.0 yet.

7. Any Other Business and Next Meeting

FF asked for any other business. GB noted that BOL has become Chair of the Board of Directors of EDItEUR, replacing JP. GB made a point of thanking JP for five years of service to the EDItEUR exec committee, as deputy then as Chair, and there was a round of applause from committee members.

The next meeting will be during the Frankfurt Book Fair, on October 10th.

GB reminded attendees that the previous day, the 9th, the International Supply Chain Seminar will be in the offices of MVB, and not within the Fair as previously.

Graham Bell
EDItEUR
11th April 2018