# ONIX International Steering Committee Minutes

*Wednesday 13th April 2016, 1:00–3:00 BST (UTC+1), Dark Room, Olympia, Hammersmith, London*

**Attendees:**

Luc Audrain (Hachette Livre)
Beat Barblan (ProQuest (for BISG))
Emiel van Bockel (CB)
Francis Cave (EDItEUR)
Laurent Dervieu (Electre) Chair
Bente Franck-Satervoll (Bokbasen)
Alex Ingram (EDItEUR)
Anna Lionetti (AIE / mEDRA)
Karina Luke (BIC)

Martin Lüning (MVB / VLB)
Renek Mendrun (ISBN / BN Poland)
Gary Pengelly (ProQuest / APA)
Jesús Peraita (FGEE)
Christer Perslöv (Bokinfo)
Lisbeth Hakansson Petré (Bokinfo)
Marie Bilde Rasmussen (for Danish PA)
Suzanne Rozario (EDItEUR)
Howard Willows (Nielsen Book)

1. **Welcome and introductions**

   LD opened the meeting, welcomed the attendees and invited each attendee to introduce themselves.

2. **Minutes of ISC Meeting held 14th October 2015, Frankfurt Book Fair, and matters arising**

   LD proposed to approve the minutes without detailed review to ensure delegates could review the main ONIX proposals today in detail. Any matters arising from the minutes would be raised within the main agenda items. GB reminded the group that draft versions of the minutes were circulated soon after Frankfurt Book Fair as usual, and had been available in draft form on the EDItEUR website for five months, but corrections can be made any time up to formal acceptance of the minutes. A similar process would occur for today's minutes. **LD called for and received formal approval of the minutes**.

3. **Report of current ONIX development work**

   LD next introduced the activity report submitted by GB, which covers the last six months work.

   As much of the activity report is covered in Item 4 on ONIX 3.0.3, GB said he would skip over those sections.

   GB reported that the sunset of ONIX 2.1 continues to raise very few issues. At the end of 2015, the period of 'twilight' support came to an end. During the 'twilight' period there was no further development of ONIX 2.1, but until the end of 2015, codelists used with 2.1 continued to be updated. Of course shared codelists continue to be updated through 2016 too, as this is a natural consequence of ongoing work on ONIX 3.0 – but from the beginning of 2016, codelists used *exclusively* with ONIX 2.1 are frozen. The only updates now considered would be (unlikely) updates for legal issues. This particularly impacts List 7 (Product Form), List 10 (E-publication Type ) and List 78 (Product Form Detail). List 7 is particularly key, as if there were a new type of binding or audio product it would not be expressible in ONIX 2.1.

   GB suggested the next stage after twilight might be 'complete darkness'. At present, the codelist files distributed for ONIX include the existing frozen codelists that pertain only to ONIX 2.1, as well as the shared and 3.0-only lists. Thus ONIX 2.1 users continue to benefit from additions made to ONIX 3.0. At some stage in the future this could cease, and there should be a final issue of the codelists that would be usable with 2.1. GB said he proposed that Issue 36 should be this final issue, with the removal of 2.1-only codelists starting from the beginning of next year, January 2017. Existing 2.1 codelists would not cease existing (Issue 36 codelists would remain available in the archive), but 2.1 users would not get any new

issues of the codelists. GB asked the group for its comments, and whether members wanted to object to the removal of the 2.1 codelists in 2017, accept or amend the timescale this proposal.

CP asked whether there was a next step beyond 'complete darkness' – could anything further be done, for example to stop ONIX 2.1 working? GB said there wasn't really much more that could be done. HW noted new adopters often still use 2.1, and GB agreed this remained an issue. LD sought clarification as to whether the change was really to remove rather than to obsolete these codelists. GB said this really is about removal, as obsolescence was the strategy for the sunset and twilight periods. ML asked about the issues from people just using values from the code lists that are intended for 3.0 with 2.1. GB said that the issue would be in validation – the messages would not validate – but he was aware that many ONIX users did not make use of validation at all. GB suggested discussion should continue over email over the coming month. GB suggested he could at least warn people of the potential for this change, but in the meantime receive responses from national groups. There was general support for this.

JP asked if the schema for 3.0.3 incorporates all the codelists. GB said yes, and currently that is in the form of a shared codelist file used by both schemas and the technical operation proposed after issue 36 would be to split them so that 3.0 could continue to be updated without adding codes to 2.1.

GB noted that there had been nothing particularly contentious added in Issues 31 and 32, and fewer than a dozen codes had been added to key 2.1-only lists during the twilight period.

LA asked about any idea of the relative usage of ONIX 3.0 and 2.1. GB noted that in varied greatly by country: in the UK it might be 20% ONIX 3.0, but close to 100% in Japan. It was still low in the US, while Sweden and Norway were almost entirely 3.0.

GB recapped the ONIX training programme, noting that he had recently completed a three-city tour of New York, Boston, Nashville in March, and training in London continues as usual. GB noted this was very positive for EDItEUR and the national groups involved (BIC in the UK, BISG in the US), and built great engagement among the trainees. GB invited national groups to suggest further locations and support for training elsewhere.

GB reported that the BISG Schema.org working group had taken longer than expected to complete its work, as members were initially unsure of the purpose of schema.org itself. But the group had now wrapped up the initial phases of the work and was moving to publication of its document as joint BISG / EDItEUR publication. It was perhaps more of a 'tutorial' – rather than a mapping as was originally intended – but publication was expected early summer. LD was keen on the possibilities of a translation of the schema.org document, to promote better work on online discovery. GB noted he had no objections to a translation, but would need to agree it with BISG as it is a co-owned document. LD also asked if it was possible to extend to FaceBook's OpenGraph, but GB felt this was out of scope.

LA noted the move of EPUB to become more 'browser-friendly', and the need within that standard to express more metadata within the book itself. Would schema.org meet this need?. GB said that schema.org suffers from being somewhat oversimplified, and while it might meet some needs, it would be best handled by the canonical URIs for ONIX concepts, which was part of the plan for the next few months.

MBR asked if there was any strong evidence for the effectiveness of schema.org markup? GB noted that there is heavy weighting on schema.org tags by Google's search algorithms, and there is anecdotal evidence for it improving rankings significantly. LD noted the test they had experienced of using schema.org: often the difficulty was that web pages with commercial offers can be lowered in the search rankings, countering inclusion of schema.org tags that raise the ranking.

GB noted the effort by Lasse Lehtinen to provide access to ONIX codelists in an online browser. This was a very useful addition, and has garnered some good comments. However, GB does not believe that this effort removes the need for work by EDItEUR to put the codelists online, in part so that canonical URIs can

be provided for every ONIX codelist entry. GB noted that there was a rising interest in this, and (as LA had pointed out) it would be of broad use to EPUB 3 and linked data communities.

Publication of ONIX 3.0.3 and a codelist browser are the two main work items up to Frankfurt Book Fair, with one other item GB would describe later.

**4.   Proposals for ONIX for Books Version 3.0.3**

LD noted the discussion of ONIX 3.0.3 had begun at the Frankfurt meeting of the ISC and should now hopefully come to a resolution. GB agreed that experience showed that the ISC is not a great venue for in-depth technical discussion and on the instructions of the Frankfurt meeting a Technical Working Group had been convened. He introduced the Report and recommendations from the Technical Working Group.

GB said the working group had met weekly from December to March with members from seven countries. GB thanked them for their work, and said he was very happy with the clear recommendation they had provided. The table of the summary recommendations provides detail of 20 proposals that had been discussed and 2 that had been added to them.

GB said he wanted first to explain the 'rejected' proposals. Most of these were rejected by the working group because although each clearly had some value, not *enough* value was seen to be worthy of addition *at present*. Every addition has to justify its place through improving discoverability, smoothing out business processes or perhaps by simplifying the metadata itself, and not all the proposals met these criteria. But business requirements change, and as the business evolves, the value of particular metadata also changes, so each rejected proposal might be reconsidered for the *next* revision – and as evidence of the importance of this, some of the accepted proposals for 3.0.3 had been rejected when considered for 3.0.2. The following proposals were recommended:

Proposal 1 – was not contentious, and simply widened scope of the underused <Conference> element.
Proposal 2 – adding <Territory> for targeting collateral was also straightforward.
Proposal 3 – was not accepted, but the related addition of packaging into <ProductPart> was recommended.
Proposal 7 – was partially recommended (adding <Prize> to <Contributor>).
Proposal 8 – also partially recommended (it has already been implemented in a related DOI standard).
Proposal 10 – recommended, to improve citations in academic publishing.
Proposal 11 – *seemingly* contentious, but in fact not problematic at all. The recommended <Gender> tag is about a *persona,* not a person, and its *gender*, not its sexuality.
Proposal 13 – structured star ratings for reviews was recommended.
Proposal 14 – unpriced items are awkward at present, and the recommendation simplifies this issue.
Proposal 15 – recommended additions to name as subject were relatively uncontentious, and the TWG also recommended further additions under proposal 21
Proposal 17 – recommended to meet German tax requirements in a simple way (note that since the TWG delivered its report, the German group has agreed to this proposal).
Proposal 18 – a straightforward addition, but it is against a particular *order line* ,not an order.
Proposal 19 – the need for textual detail with returns instructions is fairly clear.
Proposal 22 – there had been some interesting discussion about anonymous authors and discoverability that led to recommendations about the restructuring of <UnnamedPersons>.

GB also commented on a couple of the rejected proposals:

Proposal 16 – many requests for something explicit about cover images where they are provisional or have some form of expiry date. The TWG discussed this at length, but noted that *all* ONIX metadata is subject to replacement. And specifically for cover images and other supporting resources, there is already an element <ContentDate> that permits expiry of a resource. So there is already a mechanism to accomplish what was proposed: the difficulty is not the lack of a mechanis, but the lack of support

for it by certain data recipients. GB said he had agreed to emphasise the importance of the existing method in documentation.

Proposal 20 – The need for a combination of *language* and *country* attributes to be able to describe Spanish and Mexican Spanish metadata was somewhat reduced by the recommendation of Proposal 2, which enables marking metadata with a territory. This would imply specific usage, e.g. a French description with a Canada territory would cover French Canada. JP said this was fine, but the exception is in keywords (not covered by Proposal 2). And when there is a limit on the number of keywords there is a problem in handling language issues like 'lift' *vs*. 'elevator'. GB agreed to consider this. GB also noted an increase in the keywords length limit, which was originally 50 characters (!), became 250 and should now become 500 characters. JP challenged this by saying if a site only takes a few keywords there remains a problem. GB agreed this was an issue, and believed the fields at in question were keywords and <BiographicalNote>. Maybe this is a 3.0.4 feature?

GB then returned to the most controversial recommendations:

Proposal 5 –There had been a wide range of opinion and views within the working group, but the group had eventually come to a unanimous agreement. As recommended now it is limited and more conservative than the proposal that was discussed briefly at the Frankfurt ISC. There is much reduced scope for what may be termed 'misuse' or misinterpretation. As it now stands, the recommendation would not mention DRM. Focus would be on limitations on what the purchaser can then do, without any effect on the final reading experience.

Proposal 6 – GB said the working group agreed that a *separate* more specialist working group would need to cover this. This would be a separate and entirely optional block in ONIX that would only used in some messages, so it can be published as an addendum to ONIX 3.0.3.

GB proposed first just ratifying the set of recommendations presented in the Technical Working Group report as it stood, but said he expected discussion on that first. FC noted 1–20 had all originally been proposed by a national group: have they accepted the recommendations? GB said yes, because almost all most were represented on the TWG itself. But he admitted that proposal 5 was complex (it was not unanimously supported by the national group that had proposed it…).

GB noted that the way Proposal 6 is dealt with may well be as an ONIX 3.0.4. CP asked if that then meant the ISC and EDItEUR could add more in an ONIX 3.0.4? GB felt that is possible, but in that case timing would need to be considered. A whole Block 7 is certainly easier to add out of sequence, but further changes in the rest of the message may be a challenge if it is more than a single simple extra proposal. If it were a further half-dozen that would be a concern. Future revisions are coming anyway: 'every two years' is not set in stone, but as we are encouraging ONIX 3 adoption, having rapid iterations to the standard would pose an issue ("why should I adopt 3.0 if you are just going to keep changing it?").

HW felt that the meaning of 'rejected' is a bit more nuanced, as GB was describing it as more like 'deferred'. GB accepted that 'rejected' really meant 'rejected at this particular time by the working group' and drew HW's attention to the sentence in the report saying that any rejected proposal could be reconsidered at a future revision. Proposal 2 had been rejected during the work on ONIX 3.0.2, when those behind it had become less keen during that work, whereas now there is a distinct demand for it.

BFS wanted to ask about subtitles being repeatable in cases where books effectively have more than one subtitle, which is not something the working group had picked up. Could it be thought of in later 2016? GB said it could be (maybe that is the one extra addition that might be made at 3.0.4?), but also there could be an alternative within <TextContent> [GB to pick up with the Norwegian group after the meeting].

LA wanted to emphasise the need of the market to cover Proposal 5, because there are models already underway for library lending in a number of countries for which this would be needed. Some countries are already using something that is far from best practice to do so. GB agreed with this. BB said he and the

relevant BISG committee had reviewed and supported the proposals for ONIX 3.0.3 – including Proposal 5 – and that he viewed Proposal 5 a practical response to the market.

FC asked about any concerns from recipients as regards the proposals of ONIX 3.0.3? What if the adoption process is slow? GB agreed that there are some hard elements, including Proposal 5, but as LA had noted some were already handling it often in non-standard ways. Similarly, 14 and 17 are simplifications and standardisations that should make recipients' lives easier. One option could be to postpone some of the proposals that may take time to pick up until ONIX 3.0.4, particularly if it were as soon as late 2016. GB replied he would not be keen to delay, unless there was significant opposition: all that would accomplish is to delay the *start* of the adoption process. And ideally, 3.0.4 should be 95% about a further Block 7.

JP said that some of these issues make it harder for recipients to be sure that a message is valid and correct. In particular, if Proposal 5 goes through, then isn't there an issue about the possibility of incompatible usage constraints being described. Messages could be incomprehensible whilst still being valid. Similarly what if there are issues with related products, if these are not present, is the message valid? Doesn't this risk making more trouble for recipients and higher requirements on cross-validation *between* records? JP was concerned that increased difficulty of validation would lead to *less* validation. What about missing prices for countries and so on? GB agreed that it is very easy to create internal inconsistencies within an ONIX message and the normal validation tools (XSD and RNG) cannot check for them. With the exception of Proposal 5, the recommendations for 3.0.3 do not make this much worse. But Proposal 5 – by enabling more 'dimensions' to describe a price – creates some new cases where this is possible. The ONIX Schematron – the most fully-featured validation tool – can only point out so many of the issues.

LA made reference to the need for a reference implementation of an ONIX parser: is there anything about the way to ingest ONIX that would help answer this issue? There are already six dimensions for expressing price, and Proposal 5 adds a seventh. Already with six, you can supply a range of territories both with and without prices. Perhaps a group concerned about consistency could advise? JP liked the idea of a 'Consistency Group'. GB said an extended implementation guide could cover some of these issues. GB noted that if a 'Readium of ONIX' reference parser were created, it would not help define or explain the *interpretation* of the data – only how to import the data from an ONIX message into a database. GB encouraged those interested (particularly LA, JP) to pick perhaps 2–3 examples of interpretation or consistency difficulties between them to discuss, to form the basis of additions to the implementation guide. FC felt territory would a good candidate for further explanation. (At least in ONIX 3, the <Territory> structure is now consistent, where it was not in ONIX 2.1.)

BB felt that there are some good opportunities in both making metadata granular as it needs to be and simultaneously ensuring that the added granularity does not then lead to a problem of inconsistency. LA felt that there was definitely a variance in the importance of certain elements being consistent or not. There was a degree of impact upon customers and upon sales reporting. FC summarised that there was consensus that some analysis work in this area is justified (at the same time, he noted that inconsistencies in the ONIX were usually a problem with the data suppliers' underlying data). LA, JP and GB were tasked to analyse some examples on consistency to see if this would be a useful process.

**The recommendation of the Technical Working Group for ONIX 3.0.3 was approved in full without dissent.**

GB noted the main HTML documentation, XML tools have already been prepared, and he expected publication of 3.0.3 in May.

**[GB to pick up subtitle issue with Norwegian group]**
**[GB to constitute a Technical Working Group to consider proposal 6 and a potential ONIX 3.0.4]**
**[LA, JP, GB to consider how consistency can be  improved]**

### 5. Proposal for ONIX for Books Codelists Issue 33

LD introduced the proposals for Issue 33 of the codelists.

GB summarised the additions and refinements to the codelists. Specifically he drew attention to the addition to List 27, which had originally been proposed for Issue 32, then withdrawn before issue 32 was finalised. Here, it was being re-proposed. He also noted that a couple of the additions to List 139 (retail sales outlet would most likely not go ahead because the organisations were not retailers but digital distributors (and he would discuss this with the original proposers). Finally in list 217, he suggested that the GTIN should not be added (at least at this point). BFS questioned this, given that it was simply an option added to a list. GB explained that it was a concern over the nature of the proposed use of a GTIN to identify a *price*, because  GTIN is a *product* identifier. LD agreed there may be a concern from GS1 about GTINs used in the proposed way as a price identifier, and other concerns about it being an expensive and restrictive solution, and JP agreed with this. But he explained there is a need for a clear identifier. BFS was concerned about the risk of collisions if just proprietary identifiers were used (although proprietary identifiers need namespaces to prevent such collisions). GB suggested that a UUID / GUID would work, and the method of minting them removed any practical risk of collisions.

GB noted that some of the proposals established new code lists to support ONIX 3.0.3 – these are only starter lists and further additions may be necessary in the fullness of time. FC added that for 'Flourished around', the note should include '(*floruit*)' to incorporate the common Latin text used in library practice.

JP requested clarification on how these codelists would be published. GB said they would be issued in a combined listing, not with separate lists for 3.0.3. That meant that List 20, where the values are intended for use only in 3.0 would potentially be used in 2.1 (though it would be marked as 'for 3.0 only').

**The proposals for Issue 33 were approved**, subject to removal of GTIN from list 217 and careful consideration of the additions to List 139.

### 6. Any Other Business

LD asked if there was AOB, or any specific updates from national groups.

ML confirmed that the German group (led by MVB) wants the industry in Germany to move to ONIX 3.0 this year, with VLB (books in print system) and others making the push. This was welcomed by the group.

BFS asked about Amazon and ONIX 3.0, as she had been told Amazon were not taking ONIX 3.0 for e-books. CP said he was sending ONIX 3.0 to Amazon. GB said there are issues with retailers being inconsistent across teams and local systems. Amazon definitely *were* taking some ONIX 3.0 into parts of their systems and territories. There was discussion of other retailers, and GB noted system implementers are now much more supportive of ONIX 3.0 and he had now spoken to some new implementers recently who would *only* be supporting ONIX 3.0 (*ie* not accepting ONIX 2.1 at all).

CP asked why the ISTC website is down. GB said that EDItEUR provides a limited service to the ISTC agency which did not include the website, but currently the web site was moving hosts and there was a technical issue during that move. More generally, he noted that ISTC as a standard was somewhat 'on hold', and a proposal for revision had been put forward. That is going through ISO for decision, ahead of beginning work on the revision itself. CP said the issue was 'what to do whilst ISTC is not seeing take-up?' GB suggested 1) list all related products in ONIX records with reference to their identifiers, 2) use a proprietary work identifier to group works, and 3) keep an eye on ISTC revision. BFS asked for clarification on the ISO process, and GB made clear that what is now with ISO is an outline proposal for a potential revision, but that once that is approved by participating member countries within ISO TC46/SC9, then the revision process could begin. The revision itself would take some time, likely two years. [Details of participating ISO member countries and their national standards bodies can be found at http://www.iso.org/iso/iso_technical_committee.html%3Fcommid%3D48836]

**7.  Next Meeting**

LD advised the next ONIX ISC meeting would be during the Frankfurt Book Fair, and AI stated that it would be on **Wednesday 19<sup>th</sup> October (the first day of the Fair) at 1:00pm**, location to be confirmed later.

GB noted EDItEUR's **38<sup>th</sup> International Supply Chain Seminar** would be at 1:00pm on the Tuesday 18<sup>th</sup> October (the day before the Frankfurt fair opens), with arrangements similar to the successful 2015 seminar.

LD thanked the attendees and closed the meeting.

Graham Bell / Alex Ingram
EDItEUR
13<sup>th</sup> April 2016