

## Minutes of the ONIX for Books International Steering Committee

13:30 – 15:00 Wednesday 9 October 2013, Frankfurt Book Fair, Room Alliance

---

### Present

Luc Audrain	Hachette Livre	Yoichi Kimata	JPOIID
Detlef Bauer	Libri	Hetty Liewerink	CB
Graham Bell	EDItEUR	Karina Luke	BIC (part)
Mark Bide	EDItEUR	Giulia Marangoni	AIE
Maria Börman	Bokrondellen	Haeng-Ung Park	Book Trade Promotion
Francis Cave	EDItEUR	Jesús Peraita	DILVE
Piera Costantini	IE	Christer Perslöv	Bokrondellen
Laurent Dervieu	Electre	Simonetta Pillon	IE
Simon Edwards	BIC (part)	Len Vlahos	BISG
Fride Fosseng	Norske Bokdatabasen	Anthony Watkinson	UCL
Bente Franck-Sætervoll	Norske Bokdatabasen	Nicole van Wijhe	CB
Noah Genner (Chair)	BookNet Canada	Howard Willows	Nielsen
Hirohisa Ishikawa	JPOIID	Nick Woods	EDItEUR
Uwe Janssen	KNV		

### Welcome and introductions

Noah Genner introduced the meeting and welcomed participants. Each delegate introduced themselves.

#### 1. Selection and appointment of new chair for International Steering Committee

NG announced that there was only one nomination for the position of Chair – Laurent Dervieu of Electre – and as a consequence, LD was elected unopposed for a term of three years. Graham Bell thanked Noah for his 3 years' service, and NG explained that by agreement, LD would take over at the beginning of the next meeting.

#### 2. Minutes from ISC Meeting 16 April 2013, London Book Fair

NG asked the Committee to consider the minutes of the last meeting. Nicole van Wijhe raised a question regarding the decision to include both an 'out-of-print' code and a code for 'permanent withdrawal'. GB explained out-of-print means that stock may still be available in the supply chain, and future copies could still be sold. Permanent withdrawal is more often used in the e-book supply chain and means no more copies are allowed to be sold. In the absence of further queries or matters arising, the minutes were formally approved.

#### 3. Report on current ONIX developments

##### 3.1. *Written Report*

GB drew delegates' attention to this report, and apologised that owing to a heavy schedule of travel and training, some things that were expected to be done are not yet completed.

These issues should mostly be finished within the next 3 months.

### 3.2. *Codelist Revisions*

There had been two full releases of the codelists since the last meeting. Issue 21 was approved at the London meeting and released immediately afterwards. Issue 22 was approved by e-mail in the summer. Issue 23 has been prepared for approval and is discussed later.

Two or three key additions were included in issues 21 and 22. First, the beginnings of a set of new codes to describe Open Access books (mostly e-books). Open Access publishing is common in the academic journals world, and is beginning to develop in academic monograph publishing. GB described OA as (broadly) meaning ‘free at the point of use’, but there are many ‘flavours’ of OA, and typically, production of OA books is paid for by *someone* – often the author, an institution or a research council or other funding body.

Despite not having a retail price, they can become part of the supply chain, and EDItEUR has been working with OA publishers to ensure that they can be described properly.

The second significant addition is the description of e-book revisions. E-book vendors may need to keep track of what changed in between minor updates of the same book (correction of literals, technical fixes or new functionality *etc*). A recommended structure now exists (in ONIX 3 only), and example has been included in the *ONIX 3.0 Implementation and Best Practice Guide* (the *Guide*). One particular e-book retailer (Apple) has requested that publishers supply version history like this, and although Apple doesn’t use ONIX, some intermediaries convert ONIX automatically into the Apple proprietary format (Transporter).

### 3.3. *ONIX Training Programme*

GB has delivered several courses over the summer, both in UK and in other countries. He also debuted an ‘advanced’ course as part of BIC’s 2013 training programme, suitable for those with several years knowledge and experience or those who’ve been through standard training. EDItEUR wants to continue to develop this, in partnership with its members, as it’s a great way of increasing ONIX knowledge across the publishing community and improving the consistency of data.

### 3.4. *Best Practice Guides*

Whenever EDItEUR receives an enquiry about ONIX usage, around two thirds can be answered by reference to the *Guide*, and many of the remaining third trigger an addition to the *Guide*! We continue to add to the guidelines, and if your last download of the *Guide* was a year or so ago, download again – they do change. One positive note – GB hasn’t had to *remove or significantly change* any advice in the guidelines, only add, so they are proving relatively robust.

GB noted the recent release of the revised BISG/BookNet Canada Best Practice Guidelines, which are a huge update on the 2005 version. This makes extensive reference to – and builds upon – the EDItEUR *Guide*.

### 3.5. *Schema and Schematron development*

GB noted that there have been several very small updates to the DTD, XSD and RNG files for ONIX 3.0. These do not represent changes to ONIX 3 – they merely align the technical tools (the schema files) more closely with the written specification. NvW asked whether the changed versions are officially released? GB confirmed that if you download any of the schema files, you get the latest version. NvW suggested that CB has noticed instances where there are slight differences. GB emphasised that specification hasn't changed, but the schema – the way that the spec is checked – has been updated, so some *invalid* files that *appeared to be valid* under the older schema are now correctly shown to be *invalid*.

NvW asked about the timing of the releases of updated schema files. GB responded that updated versions are always released alongside codelist updates.

GB noted one schema-related correction that was not included in activity report – a minor change to the XSD for ONIX 2.1. This was done to disallow something that was valid in ONIX 1 but should not have been allowed in 2.1. The corrected XSD now rejects this as it should. Francis Cave pointed out that the XSD for 2.1 was never *officially* released. Like the 2.1 DTD, it is hosted on the EDItEUR website, but the XSD is preferred by a few implementers as it performs a more rigorous check on ONIX 2.1 files than is possible with the DTD.

[Note that the equivalent change has *not* been made in the official 2.1 DTD, as at this stage it might do more harm than good. However, a corrected DTD *for internal use only* is available on request from EDItEUR, should an organisation wish to ensure it is not creating or receiving incorrect message files.]

### 3.6. *Acknowledgement message*

This work was outlined last meeting, but has not progressed significantly over the last few months. However, a number of parties are beginning to pilot the message. There is still some work to do, particularly on the associated codelists, but this should be completed in next few weeks, to allow the pilots to go ahead. GB suggested that by the end of 2013 we should have something that could become version 1.0 of the Acknowledgement Message, and re-emphasised that such a message will be optional.

### 3.7. *Linked Data and Thema*

GB noted that a successful first meeting of the Thema International Steering Committee occurred earlier on 9th Oct. This meeting approved the publication of Version 1.0 of the Thema subject classification scheme.

Development of the database originally intended to be an online resource listing all ONIX codelists has been repurposed as a multilingual database of Thema subject categories (see <http://editeur.dyndns.org/thema>). This is working well, in seven languages so far including Russian and Arabic, and it will also provide the URIs necessary for use of Thema in Linked Data applications.

GB will now roll back and build the database of codelists, and expects this work to be done by the end of 2013.

### *3.8. Adoption of IP Policy*

GB explained that in response to issues raised during the transfer of the intellectual property embodied in Thema from its previous owners (BIC and Nielsen), EDItEUR has decided to adopt a formal policy on IP. The draft policy is very simple, and seeks to protect those who contribute to the development of EDItEUR standards while ensuring that those standards remain freely available to every organisation that wishes to implement them. It should not pose problems for any EDItEUR member, ONIX user, or contributor to EDItEUR standards. The draft is due for ratification at the EDItEUR Board Meeting (to be held 10<sup>th</sup> October), and will subsequently be made available via the website.

### *3.9. Work priorities for the next period*

The priorities for the next period are the continued development of the online tools for Thema and the ONIX codelists, continued improvements in the ONIX 3.0 Schematron, and other items to be discussed later in agenda.

NG thanked GB for the report.

## **4. Proposals for ONIX for Books Codelists Issue 23**

An initial document describing proposed additions to the ONIX codelists (Issue 23) was circulated in September, and an updated document was distributed prior to the meeting. GB highlighted the updated items.

The first addition is to List 9: product classification type. It covers the customs and tax status of a product when it is moved from one country to another, usually via one of the regional variants of the World Customs Organisation's 'Harmonised System'. The addition covers the use of the Mercosul/Mercosur classification of goods used in South America.

The proposed changes to List 41 (prize achievements) are clarifications. List 41 is a fairly abstract way of describing how close to winning the book was. It doesn't convey the *exact terminology* used by a particular prize – for example, you may be the 'winner', the 'grand medallist' *etc.* Different wording for different prizes. The exact terminology may be added in a new data element in the future.

In List 49 (region codes), there is already a code for the airside of airports in Britain. This is a tax-free zone (ground side is not.) The same rules apply in Ireland, and a new code is proposed for this purpose.

The fourth addition is to List 71 (sales restrictions). This list allows positive statements (*eg* 'for sale only via this retailer', or 'for sale only to libraries'), and some negative statements (*eg* 'not for sale to libraries'). The addition intends to allow a statement such as 'for sale except via this retailer'.

Finally, GB's updated document includes a correction from FC to the diacritics in a Polish code label.

[This is an illustration of a particular feature of the ONIX codelists. While the codes themselves are always limited to the ASCII characters (and in fact only use the characters 0–9, A–Z, a–z and the hyphen character), whereas labels and notes can incorporate any Unicode character.]

GB asked for and received unanimous approval of the proposals. They will be published as Issue 23 of the codelists by the end of October.

GB noted that at end of updated proposal document, he had added three issues for discussion over the next months, which will most likely come back for inclusion in Issue 24. The US national group (BISG's metadata committee) has requested clarification of 'digital original', and there is still some discussion over the definition even within US group.

Len Vlahos noted that the definition of the DGO (digital original) edition type currently referred to the absence of a print equivalent, but if a digital version is published first, then whether or not there was going to be a print version published later should be inconsequential. The US group is close to, but not at, a consensus. NG noted that this is important for Canada, where BookNet Canada wants to produce bestseller lists for digital products – and thus needs a very clear definition for this purpose.

GB commented that the question revolves around how long or significant a delay there is between digital release and print release. If a digital product is released, and the print equivalent released two days later, is that a digital original? Or is at least some significant delay necessary? Or is it about *expectation* – whether or not a print equivalent is planned at all (no matter what the delay). LV agreed, noting that a digital-only release could be planned but events mean it only predates print release by a handful of days.

HW asked whether the definition could be linked to the original public announcement of the product? But if a digital-only product is announced, then simultaneous release of a print product is announced later, is it still a digital original?

GB suggested each national group considers the exact meaning of this term.

The Dutch national group has requested a method to describe products as 'shorts'. NVW commented that it will shortly be possible for publishers to offer 'shorts' within the e-book fulfilment service run by CB Logistics. This is a broadly accepted category of e-book product, but one that is hard to define. Is it up to each publisher to define what a 'short' is? Some publishers might say up to 100 pages, others may have wordcount limit, but it is unlikely that a broad consensus on a specific definition can be reached. It may simply be a marketing term: the publisher may want to promote the product as a 'short' purely for sales purposes. GB agreed: it may be that the only way to define a 'short' [by analogy with recorded music, sometimes also termed a 'single'] is 'whatever the publisher decides'. However, this should probably be discussed by national groups. LD asked whether a 'short' is always a book? GB replied that it might not always be: in Japan there are 'mooks' (magazine-like books) that might qualify also as 'shorts'. But the term is mostly used for e-books, and in fact there are retailers who only sell 'shorts'.

HW asked why 'short' is proposed as an 'edition type': is it not a separate product form? GB responded that mook is in <ProductFormDetail>, but this may not be right for short. The proposal for List 21 is only tentative, and reflects a lack of clarity on what the 'short' property really is. FC questioned whether if it is a short, could it also be anything else? GB replied that yes, it would almost certainly be an e-book (or some other product form), and equally might also be some other edition type too.

Simonetta Pillon remarked that if it is 'short', it remains short. It is about the *content*, not about form or edition. Luc Audrain asked whether a short is a shortened version of something else, like a preview? GB emphasised that shorts are not previews, they are a complete and usually standalone products.

GB suggested each national group considers the exact meaning of this term.

Finally, GB raised the question of 'inspection copies', also known as 'approval copies'. These are sometimes available to educators and librarians, usually but not always direct from the publisher, often free of charge or on sale or return terms. They are intended to allow consideration of the product for adoption or recommendation to a particular course of study.

LA noted these are called 'specimen copies' in France. KL asked whether these were like proof copies? GB clarified that they were *not* bound proofs or 'advance reading copies' (ARCs), but were fully finished products.

GB said that we probably want to be able to say, 'yes, there are inspection copies available for this product'. However, do we want *extra* information? Who are they available from? Are they returnable? What are the terms and conditions under which they are available? [And additionally, do we want an analogous method for noting the availability of ARCs?]

Jesús Peraita asked whether the process for obtaining inspection copies is market dependent? GB replied that this is likely to be the case, so we need to be clear what the range of international practices is.

GB asked that the national groups think about these three issues (digital originals, shorts and inspection copies), and feed views back within the timetable for proposals for Issue 24 [ie by the end of November].

## **5. Proposal for development of 3.0.2**

NG introduced this item, and the document circulated prior to the meeting. GB commented that this has been an agenda item for the last two meetings of the ISC, and EDItEUR is getting to the stage where it has concrete proposals. Not yet an XSD file, but close to it. All additions are entirely optional. 3.0.2 will be a minor update, similar in 'size' to ONIX 3.0.1, which was discussed at FBF 2011 and released in Jan 2012. For this update, we aim to release in Jan 2014. In total there are 8 proposed additions so far.

GB described each proposal in detail.

1. 'No prefix' is about clarity. Are we saying there is no prefix, or that we don't know whether there a prefix? In all versions of ONIX up to 3.0.1, there is no way of differentiating. This request came up originally in Italy. Christer Perslöv noted that the worked example in the proposal document is in Swedish, as *all* Swedish books would have the no prefix tag [the Swedish language does not contain a definite article ('The'), and the indefinite article ('A' or 'An') is not used in titles].

2. A tag for associating a contributor with a town or city is the most widely-requested update. Ability to be more specific about where the author comes from is important for marketing purposes, but only country or region is possible at present. The proposal reuses existing <LocationName> tag, which is currently used only for naming stockholding locations.

3. The third addition also concerns stock. In EDItX, one can make detailed declarations about physical stock in a particular location, including number of copies, the precision of that number, and the rate of stock depletion. In ONIX only stock copy numbers can be carried. Some organisations have policies against the release of exact numbers. This proposal brings ONIX into line with what you can already do in EDItX. NvW asked if it's already possible in EDItX, why add it to ONIX? GB replied that some people want to use an ONIX 3.0 Block 6 update for stock reporting, especially if they have implemented ONIX but not EDItX. The important point is that the semantics of the messages must be identical.

4. In all versions of ONIX, you can say 'this book is available in a certain country' and 'this book can be only sold by a certain retailer'. In ONIX 2, it is not clear whether these two statements combine to mean 'only that retailer in that country', or 'only that country, plus

that retailer everywhere else'. In ONIX 3, this has been somewhat clarified (to mean 'only that retailer in that country') by the addition of <SalesRightsType> 07 and 08.

However, in ONIX 3, the sales rights structure in Block 4 is still not entirely clear, as slightly different structure is used in Block 6. The proposal suggests updating Block 4 to bring internal consistency with Block 6 (and would obviate the need to use codes 07 and 08).

5. This is the ability to link 2 products: 'If you have bought product A you get Product B at a lower price'. Over the last few months, several real-world examples have arisen, including Canadian bookseller BitLit offering print and e-book linking ('take a photo of your print book to get e-book discount').

Therefore e-book can have two prices: a normal price, and a lower price contingent upon prior purchase of product A.

6. Proposal 6 is about enabling a change in sales reporting. Multiple prices now exist for the same product (*eg* prices for different customer groups, perhaps consumer, corporate and library prices). Sales reporting is increasingly complex, and needs to say 'I have sold 5 copies at price X, 6 copies at price Y'. A proprietary price identifier would allow more flexibility in sales reporting. This addition would also be added to the EDItX Sales Report message if approved in ONIX 3.0.2.

Finally a few minor items: new tags for a Prize statement and for a link to a licence, plus a suggestion that the deprecation of 'complexity' be removed – commercial measures of text complexity are becoming increasingly important.

If these proposals are approved, we plan to release ONIX 3.0.2 in Jan 2014, alongside Issue 24 of the codelists.

NG asked whether there were any questions? Giulia Marangoni raised an extra point: in Italy – specifically within the LIA project that produces 'accessible' e-books for print-impaired readers – there are some issues with the version of EPUB files (EPUB 2, EPUB 3). The current method within ONIX 3.0 using <ProductFormFeature> is not very easy to use, as it is essentially a textual data element. The ONIX 2.1 <EpubTypeVersion> tag was a better way of checking, as it could more easily be validated. LA asked for clarification. GB agreed with GM that this could be useful, as it could limit the data to a number.

FC asked whether this approach should be extended to other items? GB: Some other tags are already limited to carrying numbers, of course, but we should perhaps review the specification and tighten up the requirements (and the schema) where it would be useful.

GB noted that in this case there is a reasonable argument for reintroduction of a specific <EpubTypeVersion> tag, but in some sense it is a backward step to treat e-books as 'exceptional'. And would you want a code that says '2' or '3' (*ie* be just an integer), 2.0 or 2.1 (*ie* might be a real number), or even 2.0.1 (not a number at all, but at least has a structure)?





GM replied that precision of version numbering is needed [by implication, this means the third option].

LA noted that this was discussed two years ago, and it was decided *not* add another tag then, but to use <ProductFormFeature>. GM replied that the need to distinguish very clearly between minor versions of a file format is crucial for accessibility, and it may also be critical for other market segments.

[ONIX users are reminded that if the file format changes (*eg* from EPUB 2 to EPUB 3), the latter should ideally be treated as a new product, since it will affect the range of customers who can read the file. For small updates that do *not* change the file format – for example the addition of new accessibility features, typo corrections or additions of new content or functionality, the advice is to use <EditionNumber> and <EditionVersionNumber> to distinguish between releases.]

LD said that it's important to introduce a *generic* ability to define this. GB agreed, noting that any suggestion would be limited to e-publications, but not solely to EPUB.

FC indicated the requirement may just be a codelist, as there will always only be a finite number of releases of a particular format. This approach may be plausible, now that codelists are revised very frequently. GB agreed, and suggested that FC and GB look at GM's suggestion over the next few weeks, with a view to incorporating it into either 3.0.2 or Issue 24.

NG received general assent for the proposals for 3.0.2.

## **6. Any Other Business**

NG asked for any other updates from other national groups?

LV reported that BISG's Metadata Committee and BookNet Canada have published an updated Best Practices document three weeks ago. Freely available via the BISG and BookNet Canada websites (*eg* <http://www.booknetcanada.ca/blog/2013/9/20/better-living-through-better-metadata.html>). LV said that BISG and BNC are happy for other groups to use this document and NG also clarified that the document does not aim to replace EDItEUR's *ONIX 3.0 Implementation and Best Practice Guide*, but builds further upon it. LV also said that BISG has formed two new working groups, one looking at the new Common Core US State educational standards initiative, and the other at keyword and tagging practices. Both aim to recommend best practices for ONIX. Finally, he noted that due to the US Government shutdown, BISG's data certification system is offline because parts of the Library of Congress website are unavailable.

Piera Costantini reported that the Italian group has completed translation of all the ONIX codelists into Italian.

LD noted that translation of the codelists and a new best practice spec are also in progress in France (with LA). GB said that these translations will be made available online as part of the new database.

Yoichi Kimata reported that the JPO's Forward Book Information Centre – the book data hub forming a key part of the Japanese book supply chain – migrated from 2.1.3 to ONIX 3 around May this year. The FBIC involves almost 450 publishers and 100 bookstores and data aggregators. GB noted that the Japanese group delivered a presentation on FBIC during the CONTEC conference. It was impressive, given that three years ago, there was no ONIX in Japan at all.

Uwe Janssen reported the planned German national group meeting on Friday. The group is working on best practices for version 3, and will attempt to agree a date for the introduction of ONIX 3.

GB noted that given the emphasis from several groups on best practices, it could be useful to have links from the EDItEUR website to national groups' own websites so that people can easily access the national best practice guidelines, alongside the main EDItEUR *Guide*. Detlef Bauer agreed that this would be a good idea.

JP mentioned that in Spain, DILVE has been using ONIX 3 since September 2010, although publishers don't yet use the whole array of possibilities available.

## **7. Next Meeting**

Nick Woods stated that the next meeting will be during the London Book Fair. Fair dates have changed a little for 2014, and the meeting is expected to be on Wednesday 9<sup>th</sup> April 2014, at 13:30. Note the different day within the week, and earlier week within the year.

NG and GB thanked all meeting participants, and looked forward to LD taking over as chair for the next meeting.

Nick Woods  
EDItEUR  
14/10/2013