

ONIX International Steering Committee Minutes

*Wednesday 15th April 2015, 1:30–3:00 BST (UTC+1),
Conference Room One, Olympia Conference Centre, London*

Attendees:

Luc Audrain (Hachette Livres)	Alex Ingram (EDITEUR)
Alaina-Marie Bassett (BIC)	Anna Lionetti (AIE)
Liv Berg (Bokbasen)	Karina Luke (BIC)
Marie Bilde Rasmussen (Forlæggerforeningen)	Giulia Marangoni (AIE)
Graham Bell (EDITEUR)	Renek Mendrun (Nat Library of Poland)
Maria Börman (Bokrondellen)	Diane Ouellet (BTLF)
Francis Cave (EDITEUR)	Jesús Peraita (DILVE)
Piera Costantini (Informazioni Editoriali)	Christer Perslöv (Bokrondellen)
Laurent Dervieu (Electre) [Chair]	Len Vlahos (BISG)
Fride Fosseng (Bokbasen)	Michael Vogelbacher (MVB)
Alain Fournier (BTLF)	Howard Willows (Nielsen)
Bente Franck-Sætervoll (Bokbasen)	Jon Windus (Nielsen)
Noah Genner (BNC)	Nick Woods (EDITEUR)

1. Welcome and introductions

Graham Bell called the meeting to order and handed over to the Chair, Laurent Dervieu. LD welcomed everyone to the new Olympia venue, and asked the participants each to introduce themselves. LD noted that this was the first meeting since the sunset of ONIX 2.1.

2. Minutes of ISC Meeting held 8th October, Frankfurt Book Fair, and matters arising

LD asked if there were any comments on the minutes of the previous meeting. GB noted that most of the topics will be covered in today's meeting anyway, and noted that the minutes, as draft, have been on the website for a few months already. LD called for and received 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. The main event of the last six months was the sunset of ONIX 2.1. We are now (finally!) beyond the key point in the sunset process. In early Jan, GB removed the key XML tools that provided online validation functionality, and removed the ONIX 2.1 documentation to an archive page on the EDITEUR website. ONIX 2.1 material can still be downloaded and the format can still be used by publishers and retailers. But as intended, sunset has set a strong message that 2.1 will be more difficult to use and that future development will be in 3.0.

GB noted EDItEUR prepared for a large volume of calls and e-mails raising support issues in the first couple of weeks following sunset. But there were many fewer issues than expected. Possibly this is because of the paper EDItEUR published in November describing what you need to do to install and use a local copy of the various XML tools (DTD and XSD files) – this document was very widely downloaded and the fix was implemented by many ONIX users. The paper did a lot of good.

Curiously, there was some evidence that people who use long tags were better prepared than those using short tags. The only likely explanation is that long tags used more Europe and Asia than in North America, but this is not at all a strong rule. Website statistics quoted in the report show that the 2.1 DTD files were downloaded around a million times a month during 2014, but the number of (failed) attempts to download has gone rapidly down following sunset. However, post-sunset there was a huge increase in (failed) attempts to download the 2.1 XSD – there were 1000 times more *failed* attempts to download the XSD in March than there used to be *successes* in 2014 – and almost all these attempts were in any case *misconfigured* (that is, they would not have succeeded even before sunset). It suggests a small number of organisations have botched the installation of local XSD files. GB is attempting to use the website log files to work out who, and to help them correct their installations.

[Post-meeting update: one of the organisations was the German National Library, which is correcting its errors.]

The drop in downloads saves EDItEUR money – it used to cost several hundred € a month, but even with half a million error messages being sent, the bandwidth costs have greatly decreased.

Other aspects of sunset are changes of tone, in particular relating to messages on the ONIX_implement mailing list, and in any training or other communications, emphasising that 2.1 is a ‘legacy’ format in the last stage of its lifecycle.

[Note that while various codelists that are unique to 2.1 are still being updated, this ‘twilight’ support will end as agreed at the end of 2015.]

Codelist revisions Issues 27 and 28 were released since the last meeting of the OISC in Frankfurt. GBs report goes through the main changes, and GB highlighted a handful of the most important. There is a new trade category code for ‘shorts’ – not an ideal solution but the best available – and there is a new ‘digital exclusivity’ statement. There had been some misunderstanding or misuse of the ‘digital original’ edition type code, and clarifying this highlighted the need for a ‘digital exclusive’ code for *genuine* exclusive cases. It was added in the marketing section, a good solution. Also issue 28 added an ISO region code for Kosovo (which is widely internationally recognised but does not have its own country code). This was required by organisations with Kosovan distribution arrangements distinct from arrangements in Serbia. GB raised this partly with reference to Crimea. There has been a request from the Russian National Group for a Crimean region code (as part of Russia). This is not yet widely internationally recognised, and there is no ISO region code that could be used (ISO 3166-2 recognises Crimea as part of Ukraine).

GB also highlighted the new collection type ‘Collection editoriale’. Luc Audrain, LD and Chris Saynor helped EDItEUR understand some peculiarities of French publishing, where *collections editoriale* can have hundreds or thousands of books, and where they carry some features of an imprint (or *marque*).

GB reported the introduction in January of version 1.0 of the ONIX Acknowledgement message – an optional structured reply to an ONIX message. The initial draft from 2013 was piloted by LA at Hachette Livre for data exchanges with the French National Library. LA reported that the pilot concerned the legal deposit system of national library: Hachette pushed ONIX files to the Library, and wanted to automate the whole process, so if the ONIX files were not processed properly, the Acknowledgement message noted what was incorrect. The pilot worked well, the data structure remained unchanged throughout piloting, and has been updated with the minor change in the final 1.0 version of the Acknowledgement. GB emphasised the Acknowledgement is not a simple reply message – it can simply say ‘I have your file’, or ‘I will process it tomorrow’, but can also say ‘I processed it and there’s an error in record 127’, and can include full error messages, requests for further information and other suggestions for improvement.

GB also noted that in some countries, ISBNs allocated one-by-one by ISBN Agencies. Publishers can use an ONIX for ISBN message (a simple subset of ONIX 3.0) to request an ISBN. The Acknowledgement *could* then be used to inform the publisher of the ISBN that has been assigned by the registration agency. It could similarly pass back details of any *proprietary* identifiers assigned by ONIX data recipients.

It is worth making the point that the Acknowledgement was designed for ONIX for Books, but it will also work for ONIX for Serials. EDItEUR is looking forward to seeing what happens with this message.

LA replied that his concern during the pilot was solely as a simple acknowledgement, but GB raises an interesting opportunity of the reply adding new info. In the case of the French National Library. This could include a URL that points to the CIP record, or any sort of internal proprietary identifier assigned by the ‘recipient’ of the main ONIX data.

Next GB reported on progress with ONIX training. This has been successful over several years, and this year the normal New York programme was extended to Boston and Montreal in March (in collaboration with BISG and BTLF). Some in-house training with a US-based EDItEUR member is planned in May, and the BIC training programme will begin in late April. While GB speaks only English, EDItEUR is looking at other languages, including an arrangement with CLIL in France to deliver the ONIX training in French, where EDItEUR will provide content and branding, and the training will be delivered by a bilingual ONIX trainer. GB emphasised he is happy to discuss training with anyone who wants to partner with EDItEUR. Len Vlahos attested to the quality and authority of the training delivered as part of the BISG programme.

The *ONIX 3.0 Implementation and Best Practice Guide* continues to grow. Every codelist issue results in small additions, with other additions being prompted by queries from implementors.

GB noted that where organisations and National Groups develop national best practice and Certification programmes, they should think about the implications of the sunset of 2.1. When would ONIX 3.0 become *preferred*, and when would it become *mandatory* to achieve the highest level of certification.

GB reported on work done in collaboration with BISG on schema.org. Everyone has used Google, but few know that the little infobox attached to search results comes from schema.org data embedded inside HTML web pages. This includes a limited way of providing info about books. A BISG working group is looking at how ONIX-like data should be represented in schema.org and hence in the structured data behind a webpage (which may itself be highly unstructured).

Schema.org is not a competitor to ONIX, as some have suggested; they have very distinct use cases.

Schema.org itself started by describing how to represent, say, an ISBN. Another contribution from OCLC and librarians allows a slightly richer, library-based view of the data. But our data is more sophisticated and contains elements that schema.org cannot represent. GB emphasised that we need to avoid publishers having to maintain schema.org and ONIX data separately – it needs to be mappable. The Working Group had first identified which bits of ONIX need to be mapped to schema.org (c 75 data elements), and would begin mapping of those to schema.org. Finally it will communicate to the industry on how important it is to use schema.org to improved online discoverability, and the preferred markup. GB noted to the group that there are ways of getting involved – primarily through the working group of BISG, but EDItEUR and BISG are willing to take input from others where they have expertise. LV and GB agreed that having international input to the schema.org process will ensure it carries greater weight.

Jesús Peraita asked who is responsible for making sure the data is correct? Publishers are seeing their metadata changed as it passes through the ‘metadata supply chain’. GB agreed, there is an issue of ‘updateability’. In ONIX, the metadata is considered ‘provisional’ – it is always subject to being updated by the next ONIX message, but this is not how the library sector or schema.org see the data. Where ONIX data is seen as dynamic, library data is not. This will need to be emphasised to schema.org users when documentation is compiled. GB noted that this issue neatly encapsulates why schema.org is not a competitor. It has a different purpose and outlook.

Noah Genner suggested that this brings up other mappings. BNC has maps of ONIX to other schemes, including mapping between ONIX and MARC. This is well documented by OCLC and linked from the EDItEUR website, but if other mappings exist we would be happy to add them.

4. Proposal for ONIX for Books Codelists Issue 29

LD moved on to the next agenda item: approval of additions to the ONIX codelists. GB noted this was issue 29, and the proposals were circulated to the national groups 5 weeks ago. As a result of feedback, five or so new codes had been added to the proposals, and OISC delegates had copies of the latest.

The primary additions were around extension of use of ONIX in the comic book area. The proposals included an extra identifier (List 5) and extra contributor codes (List 17) – for example as Alex Ingram explained, in many comic production teams, the illustrator of a comic book only pencils in, another contributor inks over the illustration, another is responsible for colours and another adds the letters. Each is an important and distinct role. GB said this also accounted for a minor clarification to List 7. A pamphlet doesn’t have separate cover, no proper binding, and can include small-extent unbound or stapled comics. This code was originally defined in German as ‘geheftet’, and Michael Vogelbacher confirmed that this term would cover comics. GB suggested a ‘soft’ unspecified limit of 36 pages.

List 27 now includes placeholder for the BISG educational taxonomy. While this has not yet been made available, LV agreed he is happy to add this in.

The addition to List 33 comes from Russian group, but will be useful in other areas. It allows statements of official approval (*eg* for a textbook by a Ministry of Education). LD stated there is a similar process in France, and he was happy to add the new code.

List 59 allows different price qualifiers for different groups of customers. Issue 29 includes new, separate codes for library, education and corporate customers. This means that code 6 has a different label (though an unchanged *meaning*). Howard Willows queried whether best practice would be to use new codes 10, 11, 12 rather than 06. GB replied that this depends on whether you use code 06 to mean ‘everyone except consumers’, If so then continue to use 06. But if you use code 06 to mean specifically ‘library only’ (AI noted that this appears to be common) or ‘corporate only’, then use 10, 11, 12. LV suggested that further additional granularity for different educational customers (*eg* K12, HE) may be required for future. Francis Cave asked whether, with these more specific price type qualifiers, there was a need to add a code for ‘everything’ else? GB replied that there were already two codes that could be used: an ‘unqualified price’ and ‘consumer price’. However, he agreed we need to think about this, as this area of <Price> is getting far more use than before. Are there customer groups for whom the existing codes don’t work? LA said that ONIX definitely needs to improve this aspect of pricing (possibly in 3.0.3) and we need to re-consider the various price qualifiers. There is no proper hierarchy. GB agreed. JP saw some danger of duplication between price qualifiers and sales restrictions. GB and JP discussed the difference between restrictions and qualifiers.

The proposals for Issue 29 include two suggestions relating to ‘subscription services’ such as Oyster or Scribd (where the service purchases copies outright). Some publishers want to make it clear that this book is not for sale to subscription services (at all, or at least until a particular date).

Bente Franck-Sætervoll noted the Norwegian group had made a request to cover another option where the book is available to subscription services under a revenue share. GB agreed he had looked at this suggestion, but did not feel he understand revenue-share well enough yet. BFS replied that what the Norwegian group needed was a positive way of saying ‘Available, but under a revenue share model’, but they did not need any indication of what that revenue share was. This indication need not be a price qualifier. LV asked whether the absence of a not for sale code indicate that it IS for sale? BFS replied that it needed to be positive. GB suggested that a positive indicator that it could be sold to subscription could a price type or price qualifier. But as FC pointed out that becomes would not work for revenue-share. If it was available to subscription services under a revenue share model, it could have an unpriced item type. Fride Fosseng asked whether you can you define both in one record? GB affirmed it was possible, but a bit inefficient, as there was quite a high-level distinction between markets where the product was unpriced and markets where the product had a price (or multiple prices).

GB suggested we could put in a price qualifier for fixed-price sales to subscriptions services. BFS agreed that revenue shares could be deferred to the next Issue. LV summarised: fixed price sales to subscription services could be addressed now in both a positive and a negative manner, and the revenue share model would be addressed later.

GB moved on to the final update relating to describing soundtrack types, mono, stereo *etc*.

Renek Mendrun suggested that List 17 needed a ‘scientific editor’ role. LA asked whether this was covered by the existing role of ‘technical editor’. RM asserted they were different, and in

Poland it is a valued role. GB agreed, and proposed to include in Issue 30 (*ie* not in the current Issue 29). LA and HW agreed, but noted that the difference between the technical and scientific editors need to be explained well in accompanying documentation.

LD called for and received approval of the proposals for Issue 29. The ratified updates would be published towards the end of April.

5. Prioritisation of work on ONIX 3.0.3

GB noted that he felt that enough suggestions were building up for a minor update to ONIX 3.0 – version 3.0.3 – and suggested aiming for early 2016 release. For that timetable to be met, work should begin on 3.0.3 during May–August, so that a full outline can be presented and ratified at the Frankfurt meeting of the Steering Committee. A numbered list of suggestions was included in the activity report. GB suggested that all national groups should consider this list and any other suggestions or requirements they have. However, it is important to understand that not every suggestion made will make it into ONIX 3.0.3. EDItEUR and the Steering Committee will have to strike the balance between richness and complexity.

(One addition that *must* go in is funding information for Open Access publications. GB and the Italian group have already discussed the addition of grant numbers.)

LD asked for longer explanations for each of the suggestions. GB agreed to provide them, and will circulate explanations to the various national groups. He asked that National Groups should discuss and reply to EDItEUR by July, and there was general agreement. This would give us proposals ready for FBF and publication by November/December 2015 or January 2016 at the latest.

There was general agreement on this timetable for 3.0.3.

BFS noted that e-book pricing models for libraries is an area of significant concern. There are so many different business models now. This was echoed by LA, who pointed out it was on the list (suggestion 5), and by NG. But more detail was needed so that each suggestion could be considered properly. GB agreed point 5 was a complex issue – possibly the biggest structural issue for 3.0.3 – and reminded the Steering Committee that at this stage, EDItEUR wants ‘requirements, not solutions’. The national groups should let EDItEUR know what is needed, in detailed terms, but groups need not suggest specifically how those requirements are met (*eg* suggesting specific new tags, codes *etc*).

6. Any Other Business

LD suggested that national groups might provide a brief update on ONIX activities.

LA brought up EDItX. As well as being a major publisher, Hachette is a digital distributor and deals with parties who send EDItX sales reports. With whom should he deal if he has modifications to suggest. GB and FC had already identified EDItX sales reporting as a priority area for work in 2015, and suggestions should be directed to FC.

LV reported that in addition to schema.org, BISG has a working group dedicated to ONIX 3.0 implementation. BISG planned to publish a grid showing which American organisations are creating/

accepting 3.0. Usage of 3.0 in the USA is still slow, but growing, and the most significant recent change was Bowker which could now both receive and send ONIX 3.0.

HW noted the UK ONIX group was grateful for inclusion of item 15 in the list of 3.0.3 proposals, as this had been the source of much debate in the UK.

Finally, GB related some news about the EDItEUR board. LC replaces Paola Mazzucchi as Chair. The board also elected JP as Deputy Chair. However, LV is leaving BISG later in the year, to becoming a bookseller in Denver, Colorado, and it is expected that JP will become Chair after LV leaves.

7. Next Meeting

LD advised the next meeting would be at the Frankfurt Book Fair, and Nick Woods confirmed that it would be on Wednesday 14th October, in Room Facette, Halle 3 Via within the Frankfurt Messe. Exact time to be confirmed, but likely to be 1pm.

BFS asked about having ISC meetings in North America. GB replied that EDItEUR does usually attend BookExpo America, but LV agreed that it's unlikely that more than two or three national groups could attend an ISC meeting if held in New York (or Chicago next year). However, BISG was planning a 'metadata summit', largely about ONIX, on May 8th, which EDItEUR would also attend, and EDItEUR staff do attend many BISG metadata committee meetings.

Graham Bell / Nick Woods
EDItEUR
15th April 2015