

Minutes of the ONIX for Books International Steering Committee 13:30 – 15:00 Tuesday 17 April 2012, London Book Fair, Marlborough Room

Present

Luc Audrain	Hachette Livre	Claudia Kong	CNPIEC
Graham Bell	EDItEUR	Karina Luke	BIC
Mark Bide	EDItEUR	Giulia Marangoni	AIE
Maria Börman	Bokrondellen	Paola Mazzucchi	AIE
Pierre Boudigues	Electre	Mona Naustan	Den Norske Bokdatabasen
Francis Cave	EDItEUR	Fawzia Nazir	Publishing Technology
Jean-François Chéatalat	BTLF	Diane Ouellet	BTLF
Piera Costantini	Informazioni Editoriali	Tuula Pelkonen-Tiri	Kirjavalitys
Laurent Dervieu	Electre	Jesús Peraita	DILVE
Tim Devenport	EDItEUR	Christer Perslöv	Bokrondellen
Simon Edwards	BIC	Simonetta Pillon	Informazioni Editoriali
Fride Fosseng	Den Norske Bokdatabasen	Yuji Sanui	Maruzen
Noah Genner (Chair)	BookNet Canada	Skibinskaya Diana	Guild of Book Dealers
Brian Green	EDItEUR	Mathijs Suidman	Centraal Boekhuis
Stella Griffiths	EDItEUR	Len Vlahos	BISG
Michael Hopwood	EDItEUR	Nicole van Wijhe	Centraal Boekhuis
Maarit Huttunen	National Library of Finland	Howard Willows	Nielsen
Uwe Janssen	KNV	Nick Woods	EDItEUR

Welcome and introductions

Graham Bell welcomed the participants to the International Steering Committee meeting. Noah Genner asked the Committee if there were any additions or comments to the agenda. There were none and the agenda was approved.

1. Minutes from ISC Meeting 12 October 2011, Frankfurt Book Fair

Noah asked for the Committee to approve the minutes of the last meeting ISC meeting. The minutes were agreed as an accurate record and were **unanimously approved**.

2. Report on current ONIX developments

2.1. *Written Report*

Graham drew delegates' attention to the separate activity report, then highlighted the key parts as follows:

2.2. *ONIX for Books release 3.0.1*

This release was published in January 2012. The release mostly concerns alterations to enable use in Japanese and the Chinese languages. Graham advised that the schema was effectively a simple replacement to the 3.0 schema. He highlighted the two deprecations <Audience Code> and <Reissue>, and noted that Japanese ONIX users were considering bringing forward their work on ONIX 3.0.

2.3. *Stricter Schemas for ONIX 3.0*

Graham and Francis Cave informed delegates that existing XSD, RNG and DTD schemas do not enforce all the rules laid out in the specification and, as some developers work solely from the schemas, this caused some confusion. A revised version of the schema – which enforces more of the rules present in the specification – is being tested. Graham asked the group whether the current schema should be replaced with this more exacting version.

Francis noted that the rules *were* enforced in the ONIX 2.1 DTD, but relaxation was required for some of the more complex choices in 3.0. Luc Audrain commented that the schema should enforce technically what the specification says as far as possible, and asked if it was possible to document those rules not validated by the schema. Francis noted that theoretically all the known rules that it was not possible express in XSD could indeed be listed using the Schematron schema language, and some work has been done towards this, but this is quite separate from enforcing the ‘structural’ rules in the XSD, RNG and DTD schemas.

Noah asked the Committee if there were objections to the tightening of the schemas. There were none, and the proposal for stricter schemas was **unanimously agreed**.

ACTION: Graham Bell and Francis Cave to work to tighten the schemas. [NB, this has been completed and the more exacting schemas were released 23 April]

2.4. *UML Modelling*

Graham has created what is essentially a ‘picture of ONIX’, currently circulated to a small number of people but which will eventually be made public. The purpose is primarily to assist application developers. Graham noted that these particular diagrams are a low-level, literal translation of ONIX but proposes a phase 2 of creating high level, abstract versions of the diagrams. The current UML diagrams would be most useful for developers. The more abstract diagrams would be a step towards expressing ONIX data using semantic web technologies, and thus mostly of interest to libraries and academics rather than people doing business with ONIX.

Nicole van Wijhe expressed satisfaction with the low-level documents and Graham confirmed that any high-level versions would complement, not replace, them. Luc and

Graham discussed expressing ONIX as RDF: you would need to provide 'landmarks' or guidelines for RDF expressions (as there is no control via a schema). RDF could be important as it provides a method to embed all metadata within the published document. Graham closed by confirming that EDItEUR's rigorous standards would be applied to any RDF (or similar) areas.

2.5. Development status of ONIX for Books release 2.1

Graham referred to documents circulated in November/December 2011 about the so-called 'sunset' date for ONIX 2.1, beyond which this version would move to Legacy status (the third of the four lifecycle stages) and support effort would be reduced. The reason for announcing such a date is to give users certainty and allow planning and budgeting for migration to a newer version. He confirmed that in discussions since the last ONIX ISC meeting the sunset date has been agreed as the end of 2014, but with the caveat (suggested by the United Kingdom ONIX working group) that if an exceptional circumstance, such as a legal issue, were to arise then additional codelist entries would be considered.

2.6. Codelists revisions

Since the Frankfurt ONIX ISC meeting, Codelists Issues 15 and 16 have been released. These latest releases will form the start of a regular, established release timescale for Codelists every 3 months. Codelists Issue 17 is dealt with under agenda item 3.

2.7. Other developments and activities

- Graham advised the Committee that the ISNI standard has recently been published. ISNIs can be used in ONIX and are a good match with the <Contributor> data composite.
- Book Industry Communication is making progress on the international subject categories, and a 'pan-European' category list is being worked on. Howard Willows clarified that IBIC is (and is intended to be) international, not simply European. IBIC will shortly cease its 'open house' period for new organisations to join, though there will be phase 2 recruitment later.
- Graham next spoke about the identification of ebooks. He noted that BISG recently published a policy on this topic which is expected to introduce clarity amongst publishers, particularly in North America, and that the BISG policy nearly exactly matches the guidance from the International ISBN Agency. Len Vlahos spoke about the next step from BISG: they intend to test some of the wording in the policy, after which it will be locked down as final.

- EDItEUR is to concentrate on ONIX training and implementation help. Graham spoke about EDItEUR's two approaches – provide fully commercial training like that organised by BIC, or provide training courses as and when invited by a member or consortium, as happened recently in Sweden. Christer Perslöv confirmed that the Sweden training went very well and helped with the switchover to ONIX v3.0. Graham then extended the offer of help and training to all Committee members.

3. Proposals for ONIX for Books Codelists Issue 17

Graham spoke first about the process of releasing a new issue of the codelists. There is a set time period for suggestions, then afterwards a second period where comments are invited on those suggestions put forward. This is theoretically appropriate but in practice the chronology is being broken: people continue to put forward new suggestions in the second period, which should be reserved for comments. Graham stated that, from Issue 18, the chronology will be more strictly adhered to.

The full proposal for codelists issue 17 is addressed in the accompanying document, but Graham drew the Committee's attention to some key points:

- List 12. Because of differing legal meanings of the word 'textbook' in different territories (and hence different tax codes) it is necessary to add multiple 'textbook' codes. Additionally, in Issue 17 a new code (13) has been added because this will affect the way a book is marketed and sold in the United States.
- List 15. Two new codes have been added to allow for dates to be precise to the minute or even the second. Graham noted that the time listed must by default be specific to the sender's time zone, or otherwise explicitly stated. Piera Costantini and Graham Bell raised the question of whether these new codes should only be added to version 3.0 (to avoid unnecessary work by implementers of 2.1). Len Vlahos and Luc Audrain felt the need for (for example) embargo dates with hour or minute precision is increasing, and that it is too early *not* to also include in ONIX 2.1 and Graham agreed to include these codes in ONIX 2.1.
- List 80. Graham noted that the change here comprised a more thorough explanation of the existing codes – e.g. clamshell – for clarification and requested that working groups contact EDItEUR if there are codes for which the definition remains unclear.
- List 98. This change is another where some terms (in this case, 'Navy' as a synonym for 'dark blue') will not be universally understood across the world.
- List 166. Graham noted that the new code in this list is required to clarify matters in the Japanese supply chain.

Graham then invited delegates to discuss the proposals. Luc Audrain noted an absence of the 'full product' code in list 158 prior to Issue 17, and Francis Cave confirmed this was merely an oversight in the design of ONIX 3.0 and the initial codelists. Noah Genner then asked the Committee for approval, and the publication of Codelists Issue 17 was **unanimously approved.**

Actions; Graham to complete and publish Codelists Issue 17, including the new List 15 codes in both ONIX 2.1 and ONIX 3.0 [NB: this action is completed and the new codelists were released 23 April]

4. Any Other Business

- Graham raised an issue originally highlighted by both Uwe Janssen and the BISG Metadata Committee (the US ONIX National Group): how to identify products that are created by scraping data, usually from the internet. Uwe noted that in Germany this request came from publishers, while Len noted that in the United States that it was retailers (keen to avoid consumers buying what turns out to be a 30-page staple bound volume of text scraped from Wikipedia.) Graham asked the Committee if the ability to identify scraped data products should be included in ONIX. Howard Willows noted that the UK's approach was to identify scraping at a publisher level, or even at a 'source of data' level. Graham closed the discussion by asking the national groups to consider this issue and report back.

Action: ISC National Groups to report back on their views on including (or otherwise) the ability to identify products formed of scraped data. How would this fit into the ONIX message – and who in the supply chain would provide the data?

- Luc Audrain suggested that review text in ONIX could be supplemented by a generalised way of including 'star ratings'. Luc then explained the desire for some publishers to have the ability in ONIX to include what social networks say about their products. Luc suggested that maybe one tag could be added, but noted that the question could be generalised to include any user-generated promotional material. And Graham added that we would need to consider copyright (and possibly privacy) issues if we explored this route.

Action: ISC National Groups to report back on their views on including ratings and user-generated promotional material within ONIX.

5. Next Meeting

The next meeting will be held during the Frankfurt Book Fair on Wednesday 10 October 2012. The exact venue is to be confirmed.