

Newsletter November 2011

STOP PRESS: ICEDIS North American Meeting in Charleston!

Coinciding almost exactly with the publication of this issue of the Newsletter, ICEDIS is holding its first North American meeting in many years at the Charleston Conference on Wednesday 2 November. Whether or not you have already registered for this event, you would be most welcome to attend the meeting (11:00 to 13:00 at the Charleston Place, 205 Meeting Street) and find out more about our work. Further details can be found in our Forthcoming Events page or you can contact tim@editeur.org. Feedback on the Charleston session will figure in the next issue of the Newsletter.

If word of this meeting reaches you too late, you may care to meet with Mark Bide mark@editeur.org or Tim Devenport, both of whom will be attending the Charleston Conference later in the week.

Frankfurt Book Fair 2011

The 2011 Frankfurt Book Fair has just taken place, with EDItEUR represented between Monday 10 October and Friday 14 October. Our time in Frankfurt was very successful, holding several meetings (details of the individual meetings can be found in the relevant sections of this newsletter) and speaking with members of the publishing community. The highlight of the week was our successful debut hosting EDItEUR's Supply Chain Track within the multi-track Tools Of Change Conference on Tuesday 11 October. Tools of Change attracted over 500 registered visitors and the EDItEUR's Supply Chain Track alone saw an estimated footfall of over 200 people throughout the day. Details of the Supply Chain Track sessions can be found at <http://tocfrankfurt.com/2011-program/editeur-supply-chain-track> whilst some background information on all of the Tools of Change presenters is available at <http://tocfrankfurt.com/2011-program/2011-speakers>.

ONIX for Books

Codelists Issue 15

Issue 15 of the ONIX for Books codelists was approved by the ONIX International Steering Committee (ISC) during its meeting in Frankfurt, and has subsequently been published on the EDItEUR website. The various XML schemas have also been updated to incorporate the new codes.

This Issue of the codelists adds an important new list, 196, that allows description of the 'accessibility' features of e-books which make them more suitable for print-impaired readers (blind, partially-sighted, dyslexic readers, or those with physical problems that affect their use of physical books).

ONIX 2.1 and ONIX 3.0

According to EDItEUR member Maruzen, use of ONIX in Japan has grown rapidly since its introduction at the beginning of 2011. Since then over 200 publishers and 70 distributors, wholesalers and retailers have implemented ONIX 2.1 rev.04. At Frankfurt, the ISC agreed that the various additions to ONIX 2.1 for Japanese use should be carried forward into ONIX 3.0.

This minor update to ONIX 3.0 will also ensure the standard is fully suitable for Chinese implementations (Codelists Issue 15 already includes a number of new codes requested by the Chinese ONIX National Group). And the opportunity will be taken to deprecate a couple of ONIX data elements that are discouraged in the *Implementation and Best Practice Guide* (<http://www.editeur.org/93/Release-3.0-Downloads/#Best%20practice>), <Reissue> and <AudienceCode>. This minor update is planned for release at the beginning of 2012, at the same time as Codelists Issue 16.

ONIX_Implement mailing list

Anyone working with ONIX for Books should subscribe to this list (<http://www.editeur.org/16/Maintenance-and-support/#Implementation%20listserv>). It's the best way to pose quick questions about the details of ONIX, and answers on the list, from EDItEUR or from other implementers, are usually quick, authoritative and – unlike the example below – short. There are only a handful of messages per week, so you won't get inundated with irrelevant e-mails. Here's a particularly key exchange from October:

Q. What is the best practice for creating an ONIX record for a Product that has a print version (eg ProductForm of 'BB') and several e-publications (ie, ProductForm of 'DG') with multiple EpubTypes (eg PDF, HTML, Open Ebook), all sharing the same EAN? Is this one Product record, ProductForm of 'BB' with three RelatedProduct composites with EpubType (PR.23.27) using a RelationCode of 13: 'Epublication based on (print product)? Or two Product records, one having a ProductForm of 'BB', the other with ProductForm of 'DG' with one EpubType specified at (PR.4.1) with two RelatedProduct composites? Or something else entirely?

A. First, you have to view the hardcover and the e-book(s) as at least two separate Products. After all, they are likely to carry separate prices, they have different fulfillment processes, and they clearly have different product forms (BB and DG in ONIX 2.1, BB and ED in ONIX 3.0). And that means they have two or more separate <Product> records. In ONIX, you simply cannot have a single 'product' with multiple different product forms – different product forms by definition means different Products.

And an ISBN always identifies a 'product' too. ISBNs should never be assigned to 'a group of products'. Most people take this for granted in the physical world – when faced with hardcover and softcover books with the same content, they give them two different ISBNs. Even two distinct softcovers, one rack-sized and one trade paperback, get two different ISBNs. And so it should be with e-books – if you have a hardcover and an e-book, they should clearly never share an ISBN.

[NB you mention EANs, and of course ISBNs are just a particular small subset of EANs that are reserved for identifying book and book-like products. ISBNs are EANs that begin with 978... or 979... And these days, EANs are more properly called GTIN-13s.]

So the ISBN is used to distinguish between different products which have the same content. It also, of course, distinguishes between similar products that have different content – for example the hardcover versions of the first and second editions of a particular title should always have different ISBNs. And in ONIX,

you'll have something like this:

```
<Header>
...
</Header>
<!-- product A, hardcover -->
<Product>
    Product ID A (ISBN #1)
    Product form A (BB)
</Product>
<!-- product B, e-book -->
<Product>
    Product ID A (ISBN #2)
    Product form B (DG)
</Product>
```

Second, there is an understandable wish to bring together (or to 'collocate') different products that contain the same content. You want some way of identifying all the different products, and you do this with a shared 'work ID'. Now the standards-based way of doing this is the ISTC (see <http://www.istc-international.org>). Think of this (very loosely) as an 'identifier for content'. If two different products have the same content, they have the same ISTC (and this is true even if the products come from different publishers). But the ISTC is relatively new, and few publishers have implemented it yet. For those that haven't, you can use a 'proprietary work ID'. This would be some ID that is shared between the hardcover, the softcover and all the e-book versions that contain the same content. It may be an ID that comes from your internal IT systems, and be pretty meaningless to everyone else. Ideally, the first and second editions would not share the same work ID, but for a proprietary identifier that's somewhat flexible.

In ONIX 2.1, you can include work IDs in PR.7.17. For a proprietary work ID, you also need a name (such as 'MyCompanyWorkID') in PR.7.16.

Some publishers choose to use the ISBN of the first published product (eg the hardcover) as the work ID. So the second product gets a new ISBN as its product ID, but shares the first ISBN as its work ID. Publishers sometimes call this the 'head ISBN' or the 'title ISBN' – and it is, in effect, a proprietary work ID. And so in ONIX you get something like this...

```
<Header>
...
</Header>
<!-- product A, hardcover -->
<Product>
    Product ID A (ISBN #1)
    Product form A (BB)
    Work ID A (ISBN #1)
</Product>
<!-- product B, e-book -->
<Product>
    Product ID B (ISBN #2)
    Product form B (DG)
    Work ID B (ISBN #1)
</Product>
```

Now in terms of best practice, if you look at Codelist 16 which is used to specify different types of work ID you'll see that you can use an ISBN as a work ID in this way (code 15). My advice would be – don't. If you use an ISBN as a work ID, it is too easy for someone else to misunderstand the context and start treating it as a product ID. So if you're using an ISBN as a work ID, label it as a proprietary work ID, and prefix it with a couple of letters or something so it doesn't look quite so much like an ISBN any more...

```
<WorkIdentifier>
```

```

<WorkIDType>01</WorkIDType> <!-- proprietary -->
<IDTypeName>MyCompanyWorkID</IDTypeName>
<IDValue>MyCo9781590307465</IDValue>
</WorkIdentifier>

```

You mentioned also the <RelatedProduct> composite. Within the Product record for the hardcover, you might list the e-book ISBN as an 'alternative format', and in the Product record for the e-book, you might list the hardcover ISBN with the relation 'E-publication based on...'.

Third, there is the question of several different e-book products sharing an ISBN. There is – it has to be said – some disagreement within the book trade about how to handle some of the details of this, but there is reasonable consensus about the particular case you cite. If you as a publisher are creating multiple e-books with the same content, and they are in quite different file formats (you list PDF, HTML, OEB) then they are different products. They should have different <Product> records and – if they are given ISBNs as product IDs – they should each have different ISBNs (different from each other as well as different from the hardcover).

The International ISBN Agency published an FAQ at the end of 2010 that covers exactly this question – see <http://www.isbn-international.org/faqs/view/17>. And the Book Industry Study Group in the USA is in the late stages of preparing a policy document on the application of ISBNs to e-books. While I cannot speak for BISG, I believe they are likely to broadly endorse this approach too.

[NB it need not be the publisher that assigns all these ISBNs to different e-book file formats. If a publisher creates a single 'master file' and a third-party conversion service turns that master file into the various different products that go on sale, then either the publisher or the conversion service can take responsibility for the ISBNs and the rest of the metadata.]

So in ONIX 2.1, you get...

```

<Header>
...
</Header>
<!-- product A, hardcover -->
<Product>
  Product ID A (ISBN #1)
  Product form A (BB)
  Work ID A (ISBN #1)
  Related Product Ai (alternative format ISBN #2)
  Related Product Aii (alternative format ISBN #3)
</Product>
<!-- product B, PDF e-book -->
<Product>
  Product ID B (ISBN #2)
  Product form B (DG)
  E-publication type B (002)
  Work ID B (ISBN #1)
  Related Product Bi (e-publication based on ISBN #1)
  Related Product Bii (alternative format ISBN #3)
</Product>
<!-- product C, HTML e-book -->
<Product>
  Product ID C (ISBN #3)
  Product form C (DG)
  E-publication type C (001)
  Work ID C (ISBN #1)
  Related Product Ci (e-publication based on ISBN #1)

```

Related Product Cii (alternative format ISBN #2)
</Product>

Of course, this does make your ONIX a bit bigger – you have more products, and more product records. A large part of each product record is the same, because ONIX has a pretty flat structure (it's 'denormalised'). But this doesn't mean that your internal IT systems need to have the same flat structure. A good design of internal IT system would be more hierarchical (more normalised), and would allow you to manage much of the data for each product at 'work' level, so if you need to change the spelling of an author's name, you only do it once. The data is then flattened for transmission as an ONIX message (and the hierarchy might then be rebuilt by the ONIX recipient, to match their internal IT system).

[Bonus question – why is ONIX flattened like this? Two reasons. First, because not everyone has the same hierarchical structure, and second because often, only the publisher needs to manage most of this data. Highly normalised structures are good for data management, but if all you want to do is use the data (without editing it or managing it in any way) then a partly or wholly denormalised structure is simpler and usually has better performance.]

ONIX for Books news in brief

- The minutes of the ONIX for Books International Steering Committee meeting in Frankfurt are available from <http://www.editeur.org/16/Maintenance-and-support/#Steering%20committee>
- The ONIX 3 documentation has been available in an 'enhanced' HTML version for some months. If you prefer to use PDF, that version too now incorporates the various additions made to the HTML version – more example data, a full sample message, and improved XML technical notes. You can get whichever version you prefer from <http://www.editeur.org/93/Release-3.0-Downloads/> along with the new issue of the codelists.
- The latest ONIX 3.0 documentation changes the recommended 'XML namespace' for ONIX 3.0 data elements, from <http://www.editeur.org/onix/3.0/reference> to <http://ns.editeur.org/onix/3.0/reference>. The change is intended to highlight that, although namespace URIs look like web URLs, they're different, and they generally don't resolve to real web pages. Most ONIX implementations don't use namespaces at all, but if you do, old and new namespaces should be treated as synonymous, and you should migrate to the new namespace as soon as possible.
- Don't forget that if you've implemented ONIX in your business, you should add your details to the Users and Services directory (<http://www.editeur.org/111/Users-and-services-directory/>).

EDItX

EDItX is a family of XML transaction message formats, developed by EDItEUR. Each format is tailored specifically to an individual book trade application, while using common elements across the whole set of messages. XML is recognised as the syntax of choice for future trading communication. In particular, it lends itself to the immediacy of delivery and response which are possible when web services are used.

At the Frankfurt Book Fair meetings there was growing interest in implementation of the EDItX Sales Report format (intended to be used by a retailer, distributor or other intermediary, or by an agent reporting on their behalf, to report sales to a publisher or content owner.) EDItEUR is keen to promote and support implementation of the Sales Report format, and to that end is looking for ways of improving communication with current and potential EDItX implementers. Implementers are encouraged to use the EDItX_IMPLEMENT email discussion list already set up for this purpose. EDItEUR would also like to know what additional documentation and other resources might be helpful for implementers. Please feel free to e-mail us at info@editeur.org with any suggestions.

ICEDIS developments & ONIX formats for subscription products

Frankfurt Meeting 2011 – Decisions & Priorities

A well-attended ICEDIS Committee Meeting took place on 10 October in Frankfurt, just ahead of the Book Fair. Updates were delivered on various work streams that had been progressed over the summer, notably on price catalog messaging, print claims pilots and a new format for preservation holdings. And to freshen up the format of the discussions, we held our first breakout sessions – on New Orders and on Price/Payment Reconciliation: these were lively, participatory and well received and will probably become a feature of future meetings.

The committee decided upon a number of propositions relating to version upgrades and message branding (see below) and voted on a ranking order for ICEDIS's development and rollout priorities through 2012. Top of the development list came initiatives on:

- New Orders together with Structured Addresses.
- Quotations for customer-based prices.
- Improving the Price/Payment Reconciliation process.

In parallel, work will continue on wider rollout of the price catalog message and the draft print claims standard, as well efforts progressively to extend the usage of ICEDIS messaging to new players.

Further and more detailed information is available in the Minutes of the Frankfurt meeting, available at

<http://www.editeur.org/files/Collaborations/ICEDIS/20111010%20ICEDIS%20Frankfurt%202011%20Minutes.pdf>.

Next in line for development

Exploring the updated priorities in a little more depth, ICEDIS decided on a mix of enhanced support for business models that have grown in importance recently, whilst continuing its work to automate and remove more of the “pain” from traditional, “nuts and bolts” activities.

Customer-based pricing (based for instance on the number of researchers or users in a given institution) has become a very important alternative to list prices, particularly for the purchase of online resources. To accelerate the process of obtaining reliable pricing information, ICEDIS will investigate ways of supporting Quotation requests and responses between agencies and publishers. Some of the background structural work has already been approached via ONIX for Price Catalog (see elsewhere) and it is conceivable that a web service model may be more appropriate for this purpose than traditional file exchanges.

The need to reconcile and settle numerous differences between amounts paid by agents and the prices expected by publishers continues to be an irritant in the supply chain. Worse, it can in some circumstances lead to delays and interruptions to service as well as impacting cash flows over amounts in dispute. A group of ICEDIS members have started to brainstorm ways of improving or eliminating aspects of this process and it is expected that proposals will be made in due course regarding best practices to be followed in this area.

Finally, in focussing efforts on automating the process of placing new orders (and/or orders for new customers), ICEDIS wishes to streamline what has been very largely a manual task to date. Similar issues affect two key aspects of the process and hinder attempts to automate. The phrases “who is this for?” and “exactly which product is being requested?” summarize the main challenges.

For customer identification, ICEDIS believes that a solution lies in a combination of standard organizational identifiers and a well-structured address record to facilitate matching. The lack of any widely accepted product identifier makes unambiguous product identification more difficult, but useful ideas may emerge from the price catalog exchanges already underway. Whichever approach is followed, these initiatives will also represent a major step forward in re-engineering and updating our legacy standards for orders and eventually renewals.

Looking ahead to Glasgow in 2012

ICEDIS is already starting to plan for its next European meeting, which will as usual coincide with the final day of UKSG next spring. UKSG will be held next year in Glasgow, Scotland, 26 – 28 March 2012 and the ICEDIS meeting is scheduled for early afternoon on Wednesday 28 March. More details of time and venue will be circulated nearer the date.

Version updates coming soon

Following decisions by the ICEDIS Committee in Frankfurt, changes will be made shortly to the versioning of several important serials standards. Three of the ONIX for Serials standards will be moved to formal version 1.0 status, without substantive change but recognizing that all three have been extensively piloted over several years and are already in live use. The formats concerned are the Serials Release Notification (SRN), the Serial Product and Subscription format (SPS), and the Coverage Statement – the latter used to convey complex enumeration schemes and resource ranges.

ONIX for Price Catalog

Another change in designation was agreed in Frankfurt to more clearly signal the purpose (and lineage!) of the price message recently introduced. The ICEDIS Publishers' Price List is to be renamed as ONIX for Price Catalog (ONIX-PC), without making any substantive changes, and will take its place alongside the growing list of specialized ONIX standards. More significantly, it was agreed that all future subscription price catalog development will be carried out on ONIX-PC, while the existing "Product List Priced" variant of the SPS message will be frozen at version 1.0 and not developed further. We hope that the altered title will make the purpose of this particular offering clearer and more accessible to future adopters.

The Frankfurt meeting reviewed positive progress made over the summer by early adopters of the price message – which essentially comprised a full scale and in-depth pilot, conveying information for the 2012 subscription year. Constructive lessons-learned sessions, both before and during the meeting – have resulted in suggestions for sample files, improved documentation and specialized training, all of which should assist ICEDIS in a significantly wider rollout of the standard early in 2012. Organizations that have already implemented include Wiley-Blackwell, Taylor & Francis, Elsevier, Swets, LM Information Delivery and Lehmanns.

ONIX for Preservation Holdings

Following recent approval from the Preservation working group, a draft of the new ONIX for Preservation Holdings (ONIX-PH) standard will shortly be published on the EDItEUR website. As described in earlier Newsletters, ONIX-PH will convey details of preserved e-resource holdings – in the first instance from preservation agencies such as the British Library and Portico to a new service, The Keepers Registry (<http://thekeepers.org/thekeepers/keepers.asp>), maintained by EDINA as an outcome of the PEPRS project (<http://edina.ac.uk/projects/peprs>.) Initial test files have already been generated by BL and it is hoped that other contributing agencies will progressively adopt ONIX-PH as a means of communication through 2012.

Other Major projects

Enabling Technologies Framework

Since the publication of *Accessible Publishing, Best Practice Guidelines* for Publishers there has been significant interest worldwide and we are pleased to report that 5 translations have been undertaken – French, Spanish, German, Italian & Japanese. With so many positive endorsements the guidelines are being disseminated as widely as possible to all types of publishers, large or small.

The ratification of the new ONIX for Books codelists at The Frankfurt Book Fair has significant impact for the accessibility community with far greater detail being given for the description of accessible products. Details of these updated codes will be included in the next update of the guidelines.

The event *Accessible Publishing with EPUB 3* held in conjunction with IDPF at The Frankfurt Book Fair had a very good attendance and has resulted in further interest in the project and guidelines from a number of major publishers and advocacy organisations. Speakers included Bill McCoy (IDPF), Markus Gylling (IDPF & DAISY), Dave Gunn (RNIB), Cristina Mussinelli (AIE), Graham Bell (EDItEUR) and Sarah Hilderley (EDItEUR) and slides will be made available via the EDItEUR website shortly.

For more information on this and the Enabling Technologies Project please contact Sarah Hilderley sarah@editeur.org

<http://www.editeur.org/109/Enabling-Technologies-Framework/>

Linked Heritage

Michael Hopwood, EDItEUR's Linked Heritage project lead, has continued work on a 'best metadata practice' report, the first deliverable of this EU-funded project. EDItEUR's role in the project is to enable contribution of commercial sector metadata to the Europeana cultural discovery portal (www.europeana.eu). Reviewing this initial report on metadata and identifier standards for books, music, photography and film has helped develop a glossary of technical terms to aid understanding of commercial concepts by cultural heritage sector staff.

Commercial sector engagement with Europeana is made more difficult by the recent announcement that all metadata aggregated by Europeana will be made available under the Creative Commons CC0 licence (effectively a waiver of all legal rights to the data). Michael has joined a "task force" within the Linked Heritage network that aims to develop technical and licensing responses to this decision.

Since commercial metadata clearly has a significant market value (and good metadata costs a significant amount to create and manage), some protection for data contributors and perhaps links to purchase commercial content are essential. It is of particular interest to note a recent (28 October 2011) Europeana Commission press release detailing a Recommendation which urges EU Member States to "[g]et more in-copyright material online, by, for example, creating the legal framework conditions enabling large-scale digitisation and cross-border accessibility of out-of-commerce works."

Michael presented some initial findings from his work at a joint Europeana / Federation of European Publishers event at the Frankfurt Book Fair, alongside speakers from EDItEUR members Pearson and MVB / Libreka.

New Members

Federación de Gremios de Editores de España

We are delighted to announce that, following a meeting at Frankfurt between Antonio María Ávila, Eefke Smit (EDItEUR's chair) and Mark Bide, Federación de Gremios de Editores de España (FGEE, the Spanish publishers' association) has announced its intention to become a Charter Member of EDItEUR with effect from January 2012. FGEE's representative on the Board will be Jesús Peraita, well known to the ONIX for Books community.

Kogan Page

EDItEUR welcomes new member Kogan Page (<http://www.koganpage.com/>), a leading independent publisher of business books based in London, Philadelphia and New Delhi.

Forthcoming events

ICEDIS North American Meeting at Charleston

2 November 2011, 11:00 to 13:00

Charleston Place, 205 Meeting Street, Charleston, South Carolina 29401

It's now only a few days until ICEDIS hosts its first North America meeting for many years. This will take place on Wednesday 2 November during the Charleston Conference "Issues in Book and Serial Acquisition" 2011. We invite those with an interest in communicating serials information and in automating serials transactions to come along and find out more. Indeed we welcome representatives of publishers, libraries, subscription agencies, system vendors or other intermediaries.

Alongside an update on current ICEDIS work the meeting will encourage participants to share their ideas and business requirements with existing ICEDIS members and consultants from EDItEUR. The venue, Charleston Place, is conveniently located within a short walk from the Conference at the Francis Marion Hotel.

The meeting is entirely free and there is still time to register. All an interested party has to do is e-mail Tim Devenport at tim@editeur.org. It is open to both existing EDItEUR members and ICEDIS members and other interested parties. This meeting can of course be found on the events page of EDItEUR's website at <http://www.editeur.org/3/Events/All/0>.

XXXI Charleston Conference – "Issues in Book and Serial Acquisition"

2 November 2011 to 5 November 2011

Various venues in Charleston, South Carolina, USA.

The Charleston Conference is an informal annual gathering of librarians, publishers, electronic resource managers, consultants, and vendors of library materials in Charleston to discuss issues of importance to them all.

As part of the Conference Programme, there will be a presentation about EDItEUR's work and the standards it maintains. Both Mark Bide and Tim Devenport will be attending on behalf of EDItEUR; if you would like to arrange to meet at Charleston, please email info@editeur.org.

Digital Publishing Forum: Digital Marketing to and with Libraries

2 November 2011

University College London, Gordon Street, London

Libraries are a key channel for publishers and authors to have direct contact with readers. Amongst other services, they provide a discovery tool for new authors and subjects, venues for discussing books and showcases for backlists and older works. Yet, these benefits fade into obscurity if they are not communicated efficiently, and they will be lost entirely without regard to the digital age. This forum will examine the various digital marketing initiatives which bring together publishers, authors and libraries to promote reading through library services.

This seminar will be chaired by Emma House, Trade and International Director, The Publishers Association. Speakers will include Sandy Mahal, Programme Manager, The Reading Agency, Karen Halliday, Senior Library Marketing Manager, EMEA, Wiley, and Jim Thompson, Development & Quality Manager, Edinburgh Library and Information Services, City of Edinburgh Council.

JISC Collections 2011 Conference and AGM

24 November 2011

National Railway Museum, York.

JISC Collections will be holding a Conference on 24 November 2011 in York. With a focus on the lifecycle of electronic resources, the morning session will include a review of the past year and presentations from the HE and FE perspectives. In the afternoon, JISC will be running parallel sessions on some of their leading projects and initiatives and how these will develop in the next 12 months.

Online Information 2011: Exhibition and Seminar

29 November 2011 to 1 December 2011

National Hall, Olympia and Conference Centre, Hammersmith, London

Online Information is the largest event dedicated to the information industry, providing an annual meeting place for the global information and information management profession.

This event consists of an exhibition with more than 150 international exhibitors plus an extensive educational seminar programme. Full details at www.online-information.co.uk/

Mark Bide, Executive Director

EDItEUR

United House

North Road

London N7 9DP

UK

Tel: +44 (0)20 7503 6418

Mobile: +44 (0)7785 306267

Email: mark@editeur.org