

Using ONIX 2.1 and 3.0 to describe e-books and other digital content

EDItEUR FAQ on e-books and ONIX

EDItEUR published a paper compiled by David Martin on the use of ONIX 3.0 to describe e-books and other electronic and digital products in 2009¹. However there is still widespread variation in practice among publishers in the way that they handle e-books, both in their internal management systems and in the data they provide to their supply chain partners. This document – originally prepared for the BIC Digital Rights Group, but adopted by both BIC and BISG – compiles many of the most frequently discussed issues and provides non-technical answers.

Some answers are expanded upon in EDItEUR's *ONIX 3.0 Implementation and Best Practice Guide* ², but the current document also covers options available in ONIX 2.1. Note however that ONIX 2.1 is considered a 'legacy format' and following a three year transition period, support for it was reduced from January 2015.

1. If some of my products are e-books, can I include them in an ONIX data feed?

Yes. Both ONIX 2.1 and ONIX 3.0 allow you to describe various electronic products – e-books, downloadable audio, educational software and so on. However, support in ONIX 2.1 is relatively limited, and if e-products are an important part of your product range, then you are strongly advised to upgrade your capability to ONIX 3.0.

ONIX data feeds containing e-book records can be supplied to a variety of organisations, and leveraging the standard should be more efficient than supplying unique spreadsheets per trading partner. Major e-book vendors (for example OverDrive, Kobo, Amazon, Google Play) or services providers (*eg* Ingram CoreSource, Libre Digital, Draft2Digital, ePubDirect) accept and use ONIX metadata. And data aggregators (*eg* Nielsen, Bowker) will need in-depth data about all e-books if they are to provide comprehensive e-book metadata, sales statistics or market intelligence.

2. Can I use ONIX to describe mobile 'apps' that contain book content?

Yes. In ONIX 3.0, you should use code ED from List 150 in <ProductForm> to specify a downloadable app, and a code from List 175 in <ProductFormDetail> to specify which platform it's for (code E134 for iOS or E135 for Android). Further details of the operating system or hardware requirements can be included within <ProductFormFeature>.

¹ http://www.editeur.org/files/ONIX%203/ONIX_Books_Digital_Products_3.0.pdf

² <http://www.editeur.org/93/Release-3.0-Downloads/#Best%20practice>

In ONIX 2.1, you should use code DG from List 7 in ONIX's <ProductForm> data element if the app content is 'book-like', and use <ProductFormFeature> to specify the operating system and other hardware requirements for the app. Use the <EpubType> data element (with, for example, code 040) to specify that it is a mobile app.

It's generally the case that such apps are also vendor-specific – see question 6.

3. Can I highlight that a product is a 'digital-exclusive'?

Yes. <EditionType> code 'DGO' from Codelist 21 indicates the product is (or was) a digital exclusive *at the time of publication*. More precisely, DGO means 'digital original', as a product that was exclusive at the time of publication may not remain so throughout its life – non-digital products with the same content may be launched later.

4. How can I describe an 'enhanced' e-book, so it can be differentiated from the more basic version?

You should do two things. First, in ONIX's <EditionType> data element (<EditionTypeCode> in ONIX 2.1), use the code ENH from Codelist 21. This identifies the product as an enhanced edition. Note that this implies the existence of a basic version too – don't use ENH where the version that includes video, audio and so on is the *only* version.

Second, you can use <PrimaryContentType> (ONIX 3.0 only) and <ProductContentType> (ONIX 2.1 and 3.0) to list the various types of content in the product – text, video, audio and so on. The basic version will likely contain just text plus possibly some illustrations, whereas the enhanced version may contain video, audio, animations, and so on. List 81 contains a comprehensive list of content types.

5. Can I list the DRM and usage constraints that apply to the product in ONIX?

In ONIX 3.0, yes. ONIX 3.0 contains a set of data elements that allow you to specify certain usage restrictions. The <EpubUsageConstraint>³ group of elements allow you to specify whether printing, copy/paste, text-to-speech, sharing, lending and so on are allowed, and if so, what restrictions apply to them (for example, how many devices can an e-book be shared across, or how many pages can be printed per month).

Furthermore, with the <EpubTechnicalProtection> element, you can specify whether or not these restrictions are enforced by DRM software, or whether the product is 'watermarked' uniquely for each purchaser.

ONIX 2.1 cannot describe usage constraints in a structured manner, although a free text description of important constraints may be carried in <EpubTypeNote>.

6. Does ONIX cover the territorial limitations on e-book sales, and can it cover the fact that many e-book products are vendor-specific?

ONIX 2.1 and ONIX 3.0 both allow detailed descriptions of sales rights – whether they are territorially-based or vendor-specific.

³ Although there are several ONIX data elements named <Epub...>, they are not specific to the IDPF's EPUB e-book format – 'Epub' indicates 'e-publication' and the data elements apply to all e-book formats.

The issue that arises with ONIX 2.1 is that the *interpretation* of what is often incomplete territorial rights data varies. BIC and BISG have previously advocated significantly different views in their published best practice, which sometimes leads to inconsistent interpretation by global retailers. ONIX 3.0 closes this interpretation gap, and the latest advice from BISG that *full* statements of sales rights are recommended in ONIX 2.1 also greatly reduces the scope for misinterpretation. (See *BISG's Best Practices for Product Metadata*, 2013 revision or later⁴).

It's also true that publisher's internal system limitations often make it difficult to include comprehensive territorial rights information in an ONIX product record. Rights systems – where they exist at all – are often not closely integrated with product metadata systems. This is not a problem with ONIX, but requires publisher investment in improved rights management and systems integration before the ONIX data can be simple to manage, comprehensive and reliable. Where data held in rights management systems is incomplete or uncertain, it is possible in ONIX 3.0 to state explicitly that rights are unknown or unclear using a catch-all '<ROWSalesRightsType>' data element.

7. Can I set different prices in different countries?

Yes, both ONIX 2.1 and ONIX 3.0 allow prices to be described in multiple currencies, and for those prices to be linked to particular territories. Prices can also be based on different trading terms (wholesale, agency), can be fixed or recommended (dependent on the legal regime), can be expressed in a coded fashion (for tiered prices, in ONIX 3.0 only), can be qualified according to the type of customer (corporate, consumer, member *etc*), and can of course be expressed including or excluding value-added or sales tax according to the norms in a particular geographical market.

The more common issue with multi-currency prices is that a few e-book vendors are unable to trade in multiple currencies or handle different prices in different territories, even though they trade internationally via the internet. This is gradually improving, so that international customers can increasingly pay for e-books in their own currency.

8. If I sell an e-book through several different vendors, how many ISBNs do I need?

ONIX allows the use of many different types of identifier, and the specification does not in itself mandate the use of ISBNs. However EDItEUR's view – and the way that ONIX is expected to be used – is in line with the views of the International ISBN Agency, the maintenance agency for the ISBN standard.

The IIA has published a set of guidelines on the use of ISBNs for identification of e-books⁵. The view taken in these guidelines – which are published in the form of a set of FAQs – is that each product needs separate identification, and where a single product is available through more than one outlet, a standard identifier such as the ISBN should be used. 'Product' should be defined by the similarities and differences seen by an end consumer – so for example, differences of content, file format, usage restrictions and DRM are all potentially significant. While a publisher may not wish to differentiate these products (and thus might wish to identify them all with the same identifier), other organisations in the supply chain *do* need to differentiate and often need a standard

⁴ <https://www.bisg.org/product-metadata-best-practices>

⁵ *Guidelines for the assignment of ISBNs to e-books*, available from <http://www.isbn-international.org/faqs/>

identifier like an ISBN to do so. Publishers should consider the needs of the whole supply chain when allocating identifiers, not simply their own internal preferences or requirements.

This view – that each different format of e-book requires a different ISBN – is also broadly supported by major trade organisations such as BISG and BIC. So a publisher selling the ‘same’ e-book as a PDF, as an EPUB and as a Kindle version should not use the same ISBN for all three. *There is no such thing as a single ‘e-ISBN’ that can be applied to all three versions.*

There are clearly some challenges inherent in this view – publishers may not always be fully aware of the downstream conversions that are applied to their e-book master files, conversions that may result in the creation of additional products. The widespread need to collocate the products that result from those conversions heightens the need for a reliable standard work identifier such as the ISTC. IT systems and management processes may not scale cost-effectively to cope with the larger number of metadata records. However, the alternative – the inability to securely identify particular products – is much more fundamental and serious problem that trumps all these issues.

9. How can I link a particular e-book to other e-book formats with the same content, or to print books with the same content?

The standards-based method is to use an ISTC⁶ – an ISO standard identifier like the ISBN that identifies a ‘work’. The ISTC would be shared by all versions of that work, whatever format they are in. – hardback, paperback, and all varieties of e-book. In ONIX 2.1, <WorkIdentifier> can specify the ISTC of the work manifested in the product. In ONIX 3.0, the <RelatedWork> data element does the same, and can also specify relationships to other works (for example, specifying the ‘parent’ work from which a translation has been made, or linking a work to its ‘child’ abridged version).

But ISTC is relatively new, and only a few publishers have yet embraced its use, so ONIX also allows you to provide any other work identifier. If you use an internal identifier that serves to collocate all versions of a particular title in your own ‘behind the scenes’ systems, then that identifier would – in principle – be useful to other organizations too.

By providing either an ISTC or a proprietary work identifier, it allows other organizations to create ‘clusters’ of products that contain the same content, and clustering like this should be much more reliable than any collocation based purely on metadata attributes such as title or author name (and much more under the control of the publisher).

Within ONIX data, explicit links between an e-book and other e-book formats with the same content should also be listed in <RelatedProduct>. This is doubly important when a work identifier is not supplied.

10. Can I control when an e-book goes on sale?

Yes, in exactly the same way that you can set strict on-sale dates (*aka* sales embargo dates, national laydown dates) for physical books. The only difference is that because it is common in the e-book supply chain for master files to be distributed to several digital

⁶ <http://istc-international.org>

distributors or conversion houses, it is likely that some conversions would be ready for retail before others – so the need to set a strict on-sale-date more often arises from the need to ‘level the playing field’ and synchronise retail releases through different e-book vendors, rather than from the need to synchronise release of the book with (or more often delay until after) a film premiere or a newspaper serialisation. In ONIX 2.1, you should use the <OnSaleDate> element, and in 3.0 you should use a <PublishingDate> composite that carries an ‘embargo date’.

In ONIX 3.0 only, you can also supply an embargo date for pre-orders (orders placed in advance by consumers, for delivery at the time of retail release).

Of course, using ONIX does not guarantee that data recipients abide by these rules, so embargo dates are often backed by legal agreements (affidavits), industry-wide codes of practice, or formal service level agreements (SLA) between individual senders and recipients. Clearly if you’re providing ONIX data, you should satisfy yourself that the data recipient is willing and able to abide by the date limits in the data you send.

11. My e-book retail release has to be timed to the exact hour, not just a particular day. Can I specify that?

In ONIX 3.0, yes. In ONIX 2.1, a sales embargo can only carry a date.

12. Can I control when a preview chapter becomes viewable online?

A preview excerpt (or other preview content) or a link to a preview excerpt can be provided within an ONIX 2.1 or 3.0 product record, and this preview can be linked to an explicit ‘Valid from’ date.

Best practice guidelines for ONIX 3.0 also say that when no valid from date is supplied, any strict on-sale date / sales embargo date also applies to preview excerpts. But an explicit valid from date overrides any on-sale date, so previews can be valid prior to sale of an embargoed product if required.

It should be noted that this only applies to preview samples provided within the context of the ONIX data itself. If samples are provided by other means, or are abstracted from master files by vendors, then there is little that can be said within ONIX to control their availability beyond the best practice guidance that previews should not be viewable prior to strict-on-sale dates.

13. My e-book vendor wants details of the equivalent print book attached to the data about the e-book. Can this be accommodated in ONIX?

The general approach that’s built in to ONIX is that if you want details of another product, then you should look in the Product record for that other product. Thus ONIX provides a method via <RelatedProduct> to provide an identifier for that other product, and the identifier can be used to link to the other Product record.

But in practice, e-book vendors may not wish to receive Product records for products that they cannot sell, so some *limited* information about a comparable print product can be provided in a record that describes an e-book. In ONIX 2.1, this is limited to physical details about the related product (*eg* its product form). In ONIX 3.0, there is an additional capability to provide price details in <ComparisonProductPrice>. Great care

needs to be taken to ensure that these details about a related product are kept in sync (because some types of change to that related product must trigger an update of the e-book metadata as well).

14. For backlist conversions, where I am creating a new e-book version of a print book that was published some years ago, my e-book vendor wants the pub date of the print book attached to the e-book. Is this right?

No – it’s wrong to conflate the two quite separate dates. Both ONIX 2.1 and 3.0 allow you to provide *both* pub dates. In ONIX 2.1, you use the <PublicationDate> element to carry the publication date of the e-book itself, and <YearFirstPublished> to carry the publication year of the print version.

In ONIX 3.0, these dates are carried in repeats of the <PublishingDate> composite, with different values for <PublishingDateRole>.

15. I am creating a new e-book which is a ‘bundle’ – it contains three e-books in one. How do I specify this?

This would be an ‘omnibus edition’ (use code CMB in <EditionType>). The three titles can be listed as *Sense and Sensibility / Pride and Prejudice / Northanger Abbey* (spaced slashes are conventionally used to separate the titles), or you can give the bundle a distinctive name of its own (*The Jane Austen Omnibus, Part One*) and use <ContentItem> to describe each title individually. If the three are also available as separate novels, you should use <RelatedProduct> to link them together.

16. How can I specify that the e-book is in EPUB 3 format?

For e-book formats where the version is important (for example, EPUB 2 and EPUB 3, where compatibility with reading devices may be critical), you should always specify which version a particular product conforms to. In ONIX 3.0, the version information should be carried in the <ProductFormFeature> composite with <ProductFormFeatureType> = 10. In ONIX 2.1, you should use <EpubTypeVersion>.

17. I’ve gone to some effort to optimise the accessibility of my e-book for readers who are print-impaired – blind, partially-sighted, dyslexic or physically disabled. Can I ‘advertise’ that?

Yes – and you should. Making your e-publications more accessible might increase the number of potential readers by 10 or 15%. The text-to-speech and large print features of e-book reading devices are a boon to print-impaired readers, and it’s important that you don’t (even inadvertently) disable those built-in features. Beyond that, you can use the <ProductFormFeature> composite with <ProductFormFeatureType> = 09 to specify accessibility features you’ve added like alternative textual descriptions that can be read in place of images, detailed navigation improvements, or technical features like the use of MathML.

Right now, few retailers make use of this information, but the more publishers make their books accessible – and advertise that fact – the more retailers will take advantage of that to provide a better service to readers and grow their customer base.

18. Can ONIX cope with e-book rentals?

As the e-book market develops, many e-books will likely be available for both purchase and rental (or in other words, with either a perpetual license or a time-limited license). Some publishers and retailers will simply agree on a fixed proportion of the purchase (perpetual license) price for each rental period, so that to rent an e-book for three days will be a certain proportion of the retail price, and to rent for seven days will be a larger proportion. If the proportions are fixed and apply to all products, then rentals may not need separate descriptions.

However, other publishers and retailers will wish to set and manage purchase and rental prices individually for each product. ONIX 3.0 supports this through the <PriceCondition> composite – a single e-book can have a price without conditions (which is the purchase, or perpetual license, price) plus one or more prices with conditions – each corresponding to a different rental duration.

Note however that this <PriceCondition> method does not associated each rental duration uniquely with a product identifier. So a retailer reporting revenues must either report a single, aggregated figure against the product identifier comprising both sales and rental revenues, or report against some composite key by combining the product identifier and a proprietary sale/rental duration indicator. This is made simpler using the <PriceIdentifier> composite in ONIX 3.0.

For retailers and publishers who require a unique product identifier or proprietary SKU for each rental period – for example because they demand highly granular revenue reporting – then each should have a separate ONIX Product record, and the rental period should be stated using the <EpubUsageConstraint> composite.

There is no support for rentals in ONIX 2.1.

19. My e-books are also available via ‘subscription’ services such as Scribd and Oyster. Does ONIX handle that?

Most of these services are structured as subscriptions for the consumer, but revenue is accounted for as conventional trade sales when reported from the subscription service to the publisher. Once a subscriber selects and reads a certain proportion of the book, it triggers a full purchase (and that purchased copy is linked irrevocably with that one particular subscriber). So there is no separate ‘subscription price’ – it is just a normal price (the subscription service may also receive a trade discount, just as would a conventional retailer).

EDItEUR is investigating requirements around controlling availability through such subscription channels separately from standard retail channels (for example, allowing a product to be offered via a subscription service some months after initial publication), and it is likely this will be added to ONIX 3.0 soon.

Some other ‘subscription’ services are built on a revenue share model, where revenues are not reported to the publisher as conventional per copy sales. Each competing service offers a particular and often complex revenue share calculation that’s unsuitable for standardisation. These revenue share models are not supported in ONIX at present.

20. How can I indicate my e-book rights are about to expire?

In much the same way as you would indicate an out of print or deletion⁷ date for a conventional book.

In ONIX 3.0, you can indicate a forthcoming deletion date using the <PublishingDate> composite and a <PublishingDateRole> of 13. In ONIX 2.1, you should use the <OutOfPrintDate> element – it's perfectly okay to put in a *future* date. On that date, the <PublishingStatus> should be updated to 'out of print' (code 07), the <SalesRights> adjusted to reflect the loss of rights, and an ONIX update circulated to your supply chain partners. Distributors should update the <ProductAvailability> to 'not available' (for example using code 51, 'publisher indicates OP / deleted').

21. My e-book is 'open access'. Can I indicate that in ONIX, other than just setting the price to zero?

Yes – and in fact, because of the risk of errors in applications handling 'blank' prices, zero prices are not valid in ONIX. There is a special <UnpricedItemType> for free of charge items.

'Open access' can mean a great variety of things, and the exact licence under which the e-book is published is critical. E-books can use the ONIX 3.0 <EpubLicense> composite to link to the exact terms of the licence (for example, to a Creative Commons licence or the publisher's own open licence). The <Publisher> composite can list the organisation that *funded* the publication alongside the actual publisher.

Note that the <EpubLicense> is not unique to open access – it can be used with *any* e-publication, and with proprietary and limited licences. But in addition, all open access products should carry an open access statement in P.14 <TextContent>, a short 'headline' highlighting the open access nature of the book. Presence of this statement acts as a 'flag' to indicate the product is open access, and the statement text can be displayed as a one line summary of the licence terms

22. I already provide e-book metadata in ONIX 2.1 format. What are the advantages if I upgrade to ONIX 3.0?

Support for e-books in ONIX 2.1 is relatively limited, whereas ONIX 3.0 was designed from the outset to include data elements of specific value to e-publications. For example, ONIX 3 can specify the type of technical protection measures (DRM) applied to an e-book, and the various usage limitations that apply, such as ability to cut-and-paste text, print, lend to another reader or use text-to-speech. It provides a mechanism to link to the exact terms under which the e-book is published. And it also allows clearer specification of various added-value content included in 'enhanced' e-books.

ONIX 3 includes facilities for tiered pricing and print-equivalent comparison prices which are intended for use primarily with e-books, and greatly improved facilities for description of complex international sales rights, markets and distribution channels, and multi-currency trading – it is much better match for the complex trading relationships and supply chains prevalent with e-books.

⁷ 'deletion' in the sense of 'deletion from the catalogue'. Of course, copies already sold to consumers will not be affected.

Finally, the block-update mechanism of ONIX 3 can significantly reduce the amount of data transferred between sender and recipient, which can be important where e-book prices are varied very frequently. And the advent of EDItEUR's global implementation and best practice guidance should significantly ease cross-border trading of metadata within the increasingly international e-book market, by removing the country-by-country idiosyncrasies of ONIX 2.1.

Graham Bell
EDItEUR
27 June 2011

revised 25 Jan 2013, 9 Jan 2015