EDITEUR

ONIX for Books
Product Information Format
Introduction to ONIX 3.0

April 2009, revised April 2019

EDItEUR is the international group which coordinates the development and promotion of standards for electronic commerce in the book and serials sectors. A full licence for the use of EDItEUR standards including ONIX for Books is available (DOI: 10.4400/nwgj).

EDItEUR
United House
North Road
London N7 9DP
UK

Tel +44 20 7503 6418
Fax +44 20 7503 6418
Email info@editeur.org
https://www.editeur.org/

Contents

Corrections

17 July 2009   Added a *release* attribute to the table in Section 2.1.

                 Added a reference to checking the *release* attribute in Section 3.5.

                 Added a reference to the *release* attribute in the table in the Appendix.

28 Mar 2012   Added paragraphs on block updates.

                 Added paragraph on codelists.

                 Added footnote referencing sunset of ONIX 2.x at end of 2014.

                 Minor updates to align with revision 3.0.1.

20 May 2014   Clarified meaning of 'sunset' process.

                 Added reference to *Thema* and keywords.

                 Minor updates to align with revision 3.0.2.

24 Jan 2015   Added references to ONIX Acknowledgement message.

25 Apr 2016   Minor updates to align with revision 3.0.3.

26 Oct 2017   Updates to align with revision 3.0.4.

26 Oct 2018   Minor updates to align with revision 3.0.5

26 Apr 2019   Minor updates to align with revision 3.0.6

1. What is ONIX 3.0?

ONIX 3.0 is the current version of the *ONIX for Books* Product Information Format for book publishers, data aggregators, retailers and other players in the book supply chain. It was first released in 2009, and was the first major update to the format since 2001. It is also the first release in which digital products have been treated as a 'core' element in ONIX coverage.

1.1 ONIX standards

ONIX for Books was the first, and is the most widely-adopted, member of EDItEUR's ONIX family of standards. Other ONIX standards include a family of *ONIX for Subscription Products* standards (originally termed *ONIX for Serials*) and *ONIX for Publication Licenses*, as well as more specialised formats for metadata associated with the registration of identifiers (particularly DOIs) and for the communication of rights and repertoire data between RROs (Reproduction Rights Organizations).

All ONIX standards are XML-based, and all are intended to support computer-to-computer communication between parties involved in creating, distributing, retailing, licensing or otherwise making available intellectual property in published form, whether physical or digital. The standards aim to streamline e-commerce and reduce costs in the global book, e-book and serials trade.

More information about other ONIX formats can be found on the [EDItEUR website](#).

1.2 ONIX for Books

ONIX grew from an initiative by the Association of American Publishers in 1999 to develop a 'fast-track' product information standard which would enable publishers to deliver rich product metadata in a consistent format, in particular to Internet retailers for use on their websites. The AAP's guidelines – published in January 2000 – led directly to EDItEUR's ONIX standard. EDItEUR published version 1.0 of *ONIX for Books* in May 2000.

Release 2.0 was published in 2001, and a backwards-compatible Release 2.1 in 2003. From 2003 until 2009, the format remained very stable, with only corrections and minor revisions to add functionality which had been requested by specific national interest groups.

ONIX for Books was initially adopted in the US and the UK, with Germany, Canada and other countries following soon after. By 2009, it had been taken up in some fifteen countries, whose representatives participate in the maintenance and development of the format through an International Steering Committee and through the activity of national ONIX interest groups. Since then, the number of participating countries has continued to grow, and at the time of writing (2019) there are national groups in North and South America, Europe and the Asia-Pacific region.

Although the format is invariably referred to as *ONIX* or *ONIX for Books*, it has always covered other media and other products produced by book publishers and distributed through a book supply chain.

1.3 The business benefits of ONIX

For publishers, experience has shown that *ONIX for Books* brings two important business benefits. As a communications format, it makes it possible to deliver rich product information into the supply chain in a standard form, to wholesalers and distributors, to larger retailers, to data aggregators, and to affiliate companies. And by providing a template for the content and structure of a product record, ONIX has helped to stimulate the introduction of better internal information systems, capable of bringing together all the 'metadata' needed for the

description and promotion of new and backlist titles. The same core data can also be used to produce advance information sheets, catalogues and other promotional material.

For 'downstream' supply chain partners, *ONIX for Books* means that they can speed up the loading of up-to-date product information into customer-facing systems, with less need for manual intervention and much lower risk of error. It ensures that highly dynamic information can more easily be kept accurate and up to date, reducing customer service issues at each link in the supply chain.

This more efficient and effective data supply chain is a major benefit – but so too is the uplift in sales, particularly sales through online stores, that can be linked directly to the richer, more accurate and more timely data available to potential purchasers.

1.4 The need for a 3.0 release

There have been two main drivers for the new 3.0 release: the need to improve the handling of digital products; and the recognition that the price of maintaining backwards-compatibility has been the increasing number of 'deprecated' elements which have had to be retained in the standard – and supported by ONIX receivers – even though they are no longer recommended for use.

ONIX 3.0 makes a clean sweep of 'deprecated' elements, as well as others that have been made redundant as a result of other changes in the new release. It also enables digital products to be handled more comprehensively and more consistently than before. At the same time, the opportunity has been taken to introduce important improvements in other areas, although there are many data element groups where little or no change has been considered necessary.

With the release of ONIX 3.0, all ONIX 1.x releases are formally withdrawn, and EDItEUR will no longer support them. ONIX 2.x will continue to be supported until the Steering Committee determines that the take-up of ONIX 3.0 has reached a point where it is appropriate to cease further support. All ONIX users are encouraged to adopt ONIX 3.0 as early as is practicable.[1]

1.5 The advantages of ONIX 3.0

These are some of the advantages of ONIX 3.0 over previous *ONIX for Books* releases:

- Redundant and 'deprecated' elements have been eliminated [2]. Developments introduced in ONIX 2.1 that led to greater use of flexible 'composite' elements instead of dedicated data elements have been continued.

- Digital products can be more fully and consistently described, and the groundwork has been laid for further development in this area, as new product formats and content packages evolve. In particular, the pricing of digital products and licences can be much

---

[1] In 2011, the ONIX International Steering Committee – representing all national ONIX interest groups decided that support for ONIX 2.x should be withdrawn at the end of 2014. This 'sunsetting' of 2.x was announced in January 2012, giving three years notice of the sunset date. The withdrawal of support for versions 2.0 / 2.1 does not prevent supply chain partners continuing to use the 2.1 message format, but online resources such as documentation, DTD and XSD schemas and technical advice have been archived. 'Twilight' support for codelists (controlled vocabularies) used *solely* with ONIX 2.x continues after sunset, but these are 'frozen' as of the end of 2015. Other codelists *shared* with ONIX 3.x are frozen at the end of 2016. As a result, codelists Issue 36 is the last that is compatible with version 2.1.

[2] However, the ongoing development of ONIX 3.0 means that there are (in 3.0.1 and subsequent minor updates) a very small number of newly-deprecated elements.

more complex, and ONIX 3.0 includes new data structures within the *Price* composite to cope with rentals, bundling, licence variations *etc.*

- The handling of series, sets and multiple-item products – an acknowledged problem in earlier releases – has been greatly improved. As part of this improvement, there is a new extended title composite, which enables title detail to be expressed more accurately.

- Publishers and others are using a much greater variety of 'marketing collateral' – often web-based resources – to support the promotion and sale of physical and digital products. In ONIX 3.0, new and more flexible element groups have replaced the previous *Other Text* and *Media File* composites, to allow new types of collateral material to be described and linked.

- For ONIX users working in complex international markets, for example though not limited to the international English-language market, supply-related elements in ONIX 3.0 have been regrouped to allow the status of a product in different markets to be more clearly and accurately described. Geographical information on sales rights, distribution rights and price validity is expressed in a consistent manner throughout ONIX 3.0 (where 2.x is inconsistent and inflexible).

- Textual data such as author biographies or descriptive text can be supplied in multiple languages in parallel (*eg* text in both Korean and English for a book in English sold through the Korean book trade). Dual-language metadata can have significant benefits regarding online discoverability in second language markets, where search engines favour the native language but the book's metadata is naturally in the language of the book.

- With the introduction of the [ISTC (International Standard Text Code)](), products can be related to a parent 'work', and to any 'source' and 'derived' works themselves related to the parent, to help identify groups of different manifestations of the same text, or of texts derived from a common source.[3]

- ONIX 3.0 Product records are 'blocked' in a new way which will permit updates to be sent without complete record replacement, and without the need for a separate 'Supply Update' message type as in ONIX 2.1.

- ONIX 3.0 is subject to ongoing development work. New but entirely backward-compatible minor revisions are released approximately every two years, to add new functionality, improve discoverability and meet new business requirements.[4] Earlier versions of ONIX are no longer being developed.

- The ONIX 3.0 schema definition is available in the ISO RELAX NG schema language, as well as in DTD and W3C XSD schema languages. Implementers are strongly encouraged to use the latest versions of the XSD or RNG schemas for XML validation purposes, rather than relying on the DTD. After validation, recipients may then ignore certain newly-introduced features if their system is not yet fully updated, provided that the features ignored to not modify the interpretation of active features.

- In addition, a 'strict' XSD (using XSD 1.1 and embedded Schematron) is also available, and it extends the range of content-based business rules that can be validated. It also provides warnings where deprecated data elements and codes are used. Implementers are encouraged to use the latest version of the 'strict' XSD particularly during development work, and for onboarding of new data partners.

---

[3] The ISTC standard has subsequently been withdrawn, but other (mostly proprietary) work identifiers may be used in its place

[4] The latest version is ONIX 3.0.6, released in April 2019.

- A new Acknowledgement Message has been defined (January 2015) to allow recipients of an ONIX message to confirm receipt or provide details back to the original sender of any issues encountered during processing of a message.

- For ONIX 3.0, the normative *ONIX for Books: Product Information Format Specification* ('the *Specification'*, available as a PDF or in HTML) is accompanied by an extensive *Implementation and Best Practice Guide* ('the *Guide'*) that is intended to provide guidance on the correct interpretation of the *Specification*, extensive examples and advice, and which is intended to reduce the need for purely national best practice advice. While national guides will still have their role, much of their content can simply refer to the 'global' guidance supplied by EDItEUR. This should reduce the tendency to differentiate national 'dialects' of ONIX, and increase global interoperability.

2. An overview of the format

To many people, *ONIX for Books* looks dauntingly large and complex. However, it is only complex because the products which it has to describe are complex, and because the format has to meet the differing needs of various players in the supply chain, in diverse geographical markets. It is highly modular, and can be broken down into smaller 'Blocks' and 'Groups' that are defined by their function. The aim in this section is to give a broad overview of the format, to provide some signposts to those sections which are central to all applications, and to identify those which are more specialised, and therefore less often used.

References to Groups and Blocks below refer to numbered sections in the *Specification*, where the elements of the ONIX message are defined in detail.

2.1 The message

An *ONIX for Books* message must have a header which identifies the sender, and optionally, the addressee(s), and carries a datestamp.

`<?xml version="1.0"?>`	
`<ONIXMessage release="3.0">`	
`    <Header>……………</Header>`	Message header
`    <Product>……………</Product>`	Product record
`    <Product>……………</Product>`	Product record
`    <Product>……………</Product>`	Product record
`</ONIXMessage>`	

The body of the message consists of an unlimited number of product records, each relating to a single product. Generally, this can be taken to mean a tradable product (*ie* a product for retail sale); but there are exceptions. A product record can also be used to describe an item which is sold only as part of a set; or a set of items which are only sold separately; or a piece of promotional material which is offered to retailers but which is not itself for sale; or a trade pack intended to be broken up by a retailer for sale as individual items.

2.2 The product record

The product record begins with a few elements of record metadata (or 'housekeeping'), among which a record identifier and a coded notification type are mandatory. A product identifier is also required, typically in the form of a GTIN-13 (Global Trade Identification Number) – usually an ISBN-13, unless the record refers to a non-book item whose GTIN comes from a different source. Alternative forms of product identifier can also be sent.

`<Product>`	
Record metadata	Group P.1
Product numbers	Group P.2
`    <DescriptiveDetail> ………… </DescriptiveDetail>`	Block 1
`    <CollateralDetail> …………… </CollateralDetail>`	Block 2
`    <ContentDetail> ……………… </ContentDetail>`	Block 3
`    <PublishingDetail> ………… </PublishingDetail>`	Block 4
`    <RelatedMaterial> …………… </RelatedMaterial>`	Block 5
`    <ProductSupply> ……………… </ProductSupply>`	Block 6
`</Product>`	

The rest of the record is made up of six 'blocks'. Each of the blocks 1 to 6 can occur once only (though Block 6 consists entirely of a repeatable set of groups, and within a single Block 6, this set can repeat once for each different market for which supply detail is being sent).

Each block consists of one or more numbered 'groups'. Data element groups are defined primarily to gather together multiple data elements that are related functionally, and for convenience in presentation of the detailed record specification. They do not necessarily have any structural significance. Groups consists of a mix of single XML data elements and 'composites' that are themselves groups of data elements and other nested composites. These data elements and composites express the semantic structure of the data. Individual data elements and composites may be required or optional, and may be repeatable. The order in which the elements and composites occur within the Group is critical (and this applies equally to the order of groups within a block, and to the order of blocks within a product record).

In a full ONIX product record, at least Blocks 1, 4 and 6 are expected. In other words, a full record should at least describe the product, identify the publisher, and provide supply detail for one or more markets. In practice, it will be very unusual for a full ONIX record to be sent without Block 2 and Block 5, since communicating rich collateral content ('marketing material') and expressing a range of links to other, similar products, is fundamental to most *ONIX for Books* exchanges. Block 3 is less common, but enables rich chapter-level metadata to be carried in the record for the whole-book product.

A typical ONIX product record contains all the data about a particular product. The data recipient is expected to process the record and either create a new record (if the particular product is new, or at least new in the sequence of messages sent to the recipient), or update the entire record it already holds for that product. New and updated product records can be freely mixed in a message. If a block, composite or data element is missing from an updated record, it implies deletion of previously-held data. In contrast, if an entire product record is missing, it means only that no update is required to that record, and does *not* imply deletion of an entire previously-received record.

In a block-level update record – a new feature of ONIX 3.0 distinguished by a notification type indicating a 'partial update' – any block within a product record may be omitted. This signifies that previously-received data *in that block* needs no update. Blocks that are supplied in an update record should be updated as normal by ONIX data recipients. In theory, block-level updates and full records may be freely mixed in a single message, but senders and recipients may agree to gather each into separate messages.

It is also possible to send an 'empty' message giving positive notification that there are no updates at all, which may be necessary in circumstances where there is a strict timetable for the exchange of messages (*ie* where a missed message may generate queries).

2.3 Descriptive detail

Block 1 of the product record covers eleven data element groups, P.3 to P.13. Together, these carry most of the detailed description of the form and content of the product.

`<DescriptiveDetail>`	Block 1
Product form	Group P.3
Product parts (for multiple-item products)	Group P.4
Collection	Group P.5
Product title	Group P.6
Authorship	Group P.7
Event detail (for conferences, exhibitions, etc)	Group P.8
Edition	Group P.9

Language	Group P.10
Extents	Group P.11
Illustrations and ancillary content	Group P.11
Subject	Group P.12
Audience	Group P.13
`</DescriptiveDetail>`	

Group P.3, *product form*, is mandatory within Block 1. For a single-item product, it defines the form (*eg* hardback book, MP3 downloadable file, EPUB e-book), and, where relevant, the physical dimensions and packaging. For e-publications it can specify the file format, licensing terms and any DRM-related usage constraints.

Group P.4, *product parts*, is required for any multiple-item product, to specify the form and number of the items it contains. If the individual items have product identifiers in their own right (and whether or not they are individually commercially available), these are listed and can provide a link to a detailed description sent in a separate ONIX record.

Group P.5, *collection*, carries attributes of a bibliographic collection (a series or set) to which a product belongs, when these are required as part of the description of the product. This typically means a collective title, possibly associated with a volume or part number, and, much more rarely, an ISSN for a continuing series. However, for some collections, the collective title is an integral part of the product title. In this case, it is carried as an element of the product title (Group P.6), and should not be repeated in Group P.5.

Group P.6, *product title*, is mandatory within Block 1, for all products. The structure of the *Title Detail* composite allows collective titles and part numbers to be included as part of the product title when so required.

Group P.7, *authorship*, is required in Block 1 unless a product has no named authorship (and in such a case, a positive assertion of no authorship can optionally be provided).

Group P.8, *event detail*, is included for academic and professional publications which carry the proceedings of a conference, or for other publications linked to artistic, sporting or other major public events. (This replaces the conference information in earlier versions of ONIX.)

Group P.9, *edition*, contains three elements which are important for all *ONIX for Books* users: *edition type*, *edition number*, and *edition statement*. It also carries an extensive set of elements which are specific to Bible publishing (as well as placeholder elements for other religious texts).

Group P.10, *language*, allows the language of the product text to be specified, as well as (for example) the original language of a translated work. In some ONIX applications, the language is explicitly stated only when it is different from a default language, which can be specified in the message header (though the use of defaults set in the header is discouraged). For languages which use different scripts, it is also possible to specify the script.

Group P.11, *extents*, together with *illustrations and ancillary content*, is optional. However, for most printed products and for digital products which are delivered in a non-reflowable format, the number of pages would normally be expected. Other types of extent – for example, the duration of a sound recording, or the file size for a digital download – can also be included. *Ancillary content* covers both illustrations and other matter such as an index or bibliography whose inclusion in a product is considered significant.

Group P.12, *subject*, allows an unlimited number of subject categories to be assigned to a product. Category codes or headings may be taken from any of some ninety or more supported category schemes (for example, the new *Thema* global subject category scheme, *BISAC*, *CLIL*, *Dewey*, *Library of Congress*), and may also carry unstructured lists of

keywords. The same group also allows a personal, corporate or fictional name to be specified as part of the subject of a product (for example, a biography of *Franklin D Roosevelt*), using exactly the same name format as in Group P.6. Subject detail is optional in terms of the *ONIX for Books* schema, but it is expected in most applications.

Group P.13, *audience*, enables the intended audience for a product to be specified in a variety of ways, by coding, age range, educational level *etc*.

2.4 Collateral detail

Block 2 of the product record covers four data element groups, P.14 to P.17. Together, these carry information relating to various kinds of marketing 'collateral', either carried in the ONIX record or referenced elsewhere, by web links or otherwise.

There are important distinctions between Groups P.14, P.15 and P.16. Group P.14, *text content*, is for text which is carried as part of the ONIX record, and which by implication is available to be quoted by the receiver of the record. Group P.16 is for content which is not carried in the ONIX record, but which is offered for linking or download by a publisher (or other supply chain party) to support sales, and is available for use by the receiver, subject to any stated terms. Group P.15 is for content which belongs to a third party and is cited for look-up or linking only, and which remains subject to third-party copyright.

`<CollateralDetail>`	Block 2
Text content	Group P.14
Cited content	Group P.15
Supporting resources	Group P.16
Prizes	Group P.17
`</CollateralDetail>`	

In Groups P.14 to P.17, collateral can be targeted at different geographical territories or markets, so if a product is sold in both North America and Europe, with the same ISBN but with differing covers, each cover can be specified in Group P.16 and linked to the relevant territory.

Group P.14, *text content*, is optional within Block 2, as are the other three data element groups. However, it is common practice in many ONIX applications to send both a short description and a longer description of the product. Text can be delivered in XHTML, as HTML, or without formatting (though note that use of XHTML or HTML is required for reliable delivery of *multi-paragraph* text content).

Group P.15, *cited content*, is used for references to published bestseller lists and media mentions (for example, a TV feature or the full text of a review held online) as well as any other referenced material, whether textual or otherwise.

Group P.16, *supporting resources*, covers such things as cover images, sample content, audio and video clips. The *supporting resource* composite has been structured in a very generalised way so that new types of resource and new attributes can be specified as required by adding code values. It can cover resources in any medium: text, image, audio, video, games *etc*.

Group P.17, *prizes*, details any prizes or awards that have been won by the product or the work in question. Prizes awarded to contributors (on the basis of their body of work, rather than for a specific work) can be listed in Group P.7.

2.5 Content detail

Block 3 of the product record is a single data element group, originally introduced into *ONIX for Books* to allow detailed tables of contents to be carried in a fully structured form. It also allows authorship, marketing collateral and other data to be supplied on a 'per chapter' basis. However, in most applications it has been found sufficient, and preferable, to send simple tables of contents as a text block (with XHTML or HTML markup) in the *Text Content* composite Group P.14).

`<ContentDetail>`	Block 3
Content item detail	Group P.18
`</ContentDetail>`	

Consequently most ONIX users will not need to use Block 3; and users should check with their ONIX exchange partners before adopting it.

2.6 Publishing detail

Block 4 of the product record covers important information about the imprint and publisher, and about publishing status and sales rights [5].

`<PublishingDetail>`	Block 4
Imprint and publisher	Group P.19
'Global' publishing status and copyright	Group P.20
Sales rights and restrictions	Group P.21
`</PublishingDetail>`	

Group P.19, *imprint and publisher*, is mandatory within Block 4 to the extent that either imprint or publisher detail (or both) must be sent. Group P.19 may also include contact information for staff (usually at the publisher) responsible for dealing with queries related to production, marketing, publicity, accessibility *etc*.

Group P.20, *'global' publishing status and copyright*, is optional but strongly recommended when the information is available. For products which are distributed in international markets, it is sometimes the case that an ONIX sender is only aware of the status and 'local publication date' of a product within their own territory, which need not be the same as the 'global' status or publication date. In this case, it is sufficient to use Group P.25 to send the local market publishing status and pub date. However, provision of both global and market-specific status and dates in P.20 and P.25 is preferred where both are available. Group P.20 may be used alone where there is only a single pub date or a single market, or where there is a requirement to specify the global date of publication (*ie* the earliest publication in *any* market).

Group P.21, *sales rights and restrictions*, is optional but expected in most applications. It specifies both territorial and non-territorial sales rights and restrictions which apply to the product. These should not be confused with distribution rights and restrictions which may apply to a supplier of the product, and which are specified in Block 6.

2.7 Related material

Block 5 of the product record has two optional data element groups which respectively carry pointers to related works and related products.

---

[5] ONIX does not directly carry information about 'publishing rights'. It describes *products*, and how, where and by whom those products can be traded.

`<RelatedMaterial>`	Block 5
Related works	Group P.22
Related products	Group P.23
`</RelatedMaterial>`	

Group P.22, *related works*, makes it possible to group products which embody, or are derived from, a single work, identified by an ISTC or a similar work identifier.

Group P.23, *related products*, may identify a variety of different types of relationship between products, for example an alternative format, an earlier or successor edition, or a set of which a product forms part.

In both P.22 and P.23, the information sent in the ONIX record specifies simply a relation type and an identifier of the related product or work. The identifier must be followed up elsewhere to obtain more detail.

2.8 Product supply

Block 6 of the product record covers supply and availability detail for an implicitly or explicitly defined market. The complete block is repeatable for each different market for which such detail is sent.

`<ProductSupply>`	Block 6
Market	Group P.24
Market publishing status	Group P.25
Supply detail: availability and price within market	Group P.26
`</ProductSupply>`	

Group P.24, *market*, defines a market territory and specifies any additional sales restrictions within the territory. It is mandatory within Block 6 except in ONIX applications which are expressly limited to a single market, where the parties exchanging ONIX feeds have agreed that the market can be taken as implicit.

Group P.25, *market publishing status*, identifies a publisher's representative (if any) within a market, and specifies the publishing status and publication date of a product within a specified market. For products which are distributed in international markets, particularly the English-language market, this group is mandatory within Block 6. It may be omitted in ONIX applications which are expressly limited to a single market, where the parties have agreed to use Group P.20.

Group P.26, *supply detail*, is mandatory within Block 6 (although Block 6 itself is optional within a product record). It is a very substantial data element group covering details of distribution sources, availability, prices, *etc* within a market. For many applications, only a few elements are required: the identity of a supply source, an availability code, a price type and a price amount, and optionally also a geographical territory where that price is valid. However, to meet requirements which have been expressed by user groups in different countries, Group P.26 also covers such things as discounts, returns conditions, stock levels and applicable taxes. It allows communication of price categories in lieu of actual amounts, and of prices that are conditional upon other factors (particular qualifying customer types, prior ownership of other products and so on). It also enables communication of different prices for perpetual licences (*eg* an e-book *purchase*) and time-limited licenses (an e-book *rental*), or e-books where there are different limitations on usage depending on the price paid.

2.9 ONIX codelists

An important part of the ONIX for Books framework is the set of 'codelists', a series of controlled vocabularies that are used with particular XML data elements. Codelists can be revised without necessitating a new release of the main XML message specification, and are the main means of ensuring ONIX can remain relevant as new business requirements arise.

Codelists are typically revised at least two (and usually four) time per year. Some codelists are unique to ONIX 3.0, and some are shared with previous versions of ONIX.

The first version of the codelists intended for use with ONIX 3.0 was Issue 9, but later minor revisions require more up to date version of the codelists. The earliest version of the codelists that can be used with the latest revisions of ONIX 3.0 (*ie* with 3.0.5 or 3.0.6) is Issue 43.

2.10 Changes from Release 2.1

The notes in the preceding section are written without referring to previous releases, so that they apply equally to new and experienced users. If you are already using ONIX 2.1, and are planning a transition to 3.0, please see additional notes on changes in the Appendix.

3. Implementation

Most of this section is aimed at prospective senders of *ONIX for Books* messages – primarily publishers – although some may be equally relevant to recipients of ONIX data feeds. The aim is not to offer detailed implementation advice, but to record a few points that have come up repeatedly from experience with previous releases.

Implementation questions are routinely discussed on the *ONIX_Implement* listserv, and notices of *ONIX for Books* developments are sent out to the listserv as well as to national groups. If you have not signed up to the list, you can do so here.

And there is a detailed *ONIX 3.0 Implementation and Best Practice Guide* that accompanies the *Specification*.

3.1 Implementation options

The options for publishers who want to start sending *ONIX for Books* are of three kinds: develop or commission bespoke software; buy in a third-party system for product data management; or contract to use a web-based service which supports online data entry and delivery of ONIX output to designated receivers. The availability and practicality of each of these options will vary from country to country and publisher to publisher: some ONIX national groups [6] may be able to provide contacts with technical suppliers.

3.2 A communications format

ONIX for Books is a format for communicating information from one computer system to another within the book and e-book supply chain. It is not in itself designed as a database format. However, it is obvious that communication can only work if the systems at each end are substantially 'ONIX compliant' and support data element structures, data semantics and encoding which are no less exact – and ideally no less detailed – than those used in ONIX. It is important to verify at the earliest possible stage that any product data management system you plan to use is 'ONIX compliant' in this sense.

---

[6] EDItEUR aims to provide links to as many national groups as possible from the *ONIX for Books* pages at https://www.editeur.org/.

Most ONIX data originates with publishers and is eventually consumed by retailers and retail platforms. It may pass direct from publisher to retailer, or may pass through various other organisations such as bibliographic data aggregators. This should be viewed as a compled 'metadata supply chain', analogous but subtly different from the supply chain for the products themselves.

As well as communicating externally between organisations in the supply chain, ONIX is also used as an internal 'data bus' to communicate between applications within a single organisation.

3.3 Data entry

When much of the content of a product record is free text which cannot be automatically validated, it is all too easy to deliver something which is correct in format terms but wrong in data content. For example, title fields in some book trade systems are traditionally used to carry added data, such as an edition number or the binding type, which differentiates one product from another. It is a key principle of ONIX that fields are not used in this way, and that only one piece of information is used in each field – thus there are separate fields for the text of the main title, the binding type and the edition number. The internal system can only create correct ONIX if it too maintains the data in this structured and granular way. So not only the internal database structure, but also the disciplines followed by data entry staff, need to be consistent and 'ONIX compliant'.

3.4 Determining what to include

Nobody uses *all* the elements of the *ONIX for Books* format – though, where an element is used, it should be used in a very controlled and consistent way. As well as the EDItEUR-supplied *ONIX 3.0 Implementation and Best Practice Guide*, in many countries where it has been widely adopted national groups or trade organisations publish their own guidelines for implementation and 'good practice'. They often use the EDItEUR guide as the basis of the national guide. If you are buying a third-party system or service, you need to check that it complies with both the EDItEUR guidelines and any additional national guidelines that are in use in your particular market(s). Some receivers may choose not to use certain elements, or may ask for one option rather than another. However, no recipient should reject a valid message because it includes optional content which they have chosen not to use.

3.5 Validating ONIX messages

Developers can use a variety of widely-available XML tools to validate an ONIX message against the DTD, XSD or RNG schema. With the DTD, only the basic XML tag structure is validated. With the XSD or RNG schemas, both the tag structure and the code values are validated. An ONIX message which fails at either of these levels is invalid, and it is reasonable to expect it to be rejected by a receiver.

However, there are also ONIX business rules which (for example) specify conditional requirements. EDItEUR has developed a 'strict' XSD to check many of these requirements, However, the tool requires XSD 1.1 and (optionally) Schematron, and these requirements limit the XML software that can be used for valiation purposes (Saxon and Xerces XML parsers are known to work, libxml is known not to work). A data recipient may also reject data that does not meet the requirements of this strict schema. A few national groups and independent application developers offer validation tools which check some of these added criteria, and many off-the-shelf third-party software applications implement similar rule-based validation or scorecarding.

Several national groups operate accreditation schemes which involve a more rigorous test process, typically taking into account not just the basic validity of ONIX messages but also their timeliness and the accuracy of the metadata content.

Note that for XML validation with XSD or RNG, most XML tools will require the XML namespace attribute to be added to the top level message tag if it is not already present (*eg* <ONIXMessage release="3.0" xmlns="http://ns.editeur.org/onix/3.0/reference">).

3.6 Order of elements in the *Product* record

With previous ONIX releases, there has been evidence that some implementers have not understood that elements in the *Product* record must be delivered in the sequence defined by the schema. A message in which elements occur out of sequence will not validate.

3.7 Empty elements

In XML it is possible to deliver an empty element in either of two forms: <Tag></Tag> or <Tag/>. *Neither* of these two forms should ever occur in an *ONIX for Books* message, except in a few instances where an element (*eg* the <MainSubject/> flag in Group P.12) is specifically defined as always empty, in which case we strongly recommend that the second form <Tag/> should always be used. [7]

If there is no data for an element that is not defined to be empty, the element should be omitted completely. Such 'illegal' empty elements are detected by validation against the XSD or RNG schemas, but not by validation against the DTD. Consequently there is a risk that a message could pass DTD validation even though it includes mandatory elements for which there is no data. We therefore strongly recommend that the XSD or RNG schema be used for validation where system constraints allow it.

3.8 Reference names and short tags

Each element in *ONIX for Books* has a 'plain language' reference name (for example, <NotificationType>) and a short tag (for example, <a002>). The schema definitions allow *either* reference names or short tags to be used to label the elements in an ONIX XML message. Each reference name tag has an equivalent short tag. However, they cannot be mixed in the same message. Users have a choice between readability and conciseness – using short tags usually reduces message size by around a quarter, though given ever-increasing speed of file transfers over the internet, file size itself is decreasingly important, and in any case, zipping the file is *much* more effective. And the choice of long or short tags has no bearing on the *complexity* of the message – there is no significant saving in processing time, memory requirements and so on for either sender or recipient.

In the very earliest ONIX releases, the initial letters of the short tags indicated an attempted logical grouping of elements. As the format developed, this grouping quickly became impractical, so that there is now no significance to be drawn from the initial letters. (The numeric part of the tag has always been unique by itself.) In Release 3.0, all new elements have been assigned short tags of the form <x*nnn*>, so that they can be immediately recognised as new.

---

[7] The half dozen or so 'defined to be empty' elements must *always* be empty. Other elements must always contain data – and are not valid ONIX if left empty.

Appendix: key changes in ONIX 3.0

This Appendix is for ONIX users who are planning a transition from *ONIX for Books* 2.1 to 3.0. For each numbered data element group in the Release 2.1 Product record, the table below shows the equivalent in 3.0, and notes where there have been significant changes.

The table below does not list *all* changes between the two formats – in particular, it does not list most *additions* to the message format which enable new functionality. It focuses on the key changes that an implementer of ONIX 2.1 would have to consider as the 'first phase' of a migration – and this first phase would then enable 'minimal' ONIX 3.0 to which new functionality can then be added.

The trickiest areas are likely to be – for most organisations migrating from 2.1 to 3.0 – the changes in Groups P.5 and P.6 relating to series and set titles, the provision of good sales rights information in P.20, and the restructuring of P.24, P.25 and P.26. Each of these changes aims to fix deficiencies in the data supplied using ONIX 2.1 that lead to misinterpretation of data or the over-simplification of complex commercial arrangements.

2.1	3.0	Notes
-	-	The <ONIXMessage> tag (the XML 'root element') now has a mandatory *release* attribute and may include an optional XML namespace declaration (*release* was optional in 2.1). The message should *not* contain a DOCTYPE declaration. The use of Unicode and UTF-8 is strongly recommended, particularly for ONIX messages exchanged internationally. Throughout the message, most textual data elements are repeatable in multiple parallel languages.
		<Header>: some restructuring with elements renamed, and redundant elements have been removed.
		<Product>: a new six-part block structure within the product record, intended to support more granular, block-by-block updating, is new to ONIX 3.0. The ONIX Price and Availability Update message is no longer a separate message format – it is simply a block update containing only an identification preamble and Block 6. For regularly-timetabled updates, it is possible to send a message containing no Product records.
PR.1	P.1	*Record reference number, type and source*: no significant changes
PR.2	P.2	*Product identifiers*: the Barcode element has been replaced by a new <Barcode> composite allowing barcode type and position to be specified separately, and redundant elements have been removed. No other changes.
		Block 1
PR.3	P.3, P.4	*Product form*: there are significant changes in the elements and coding used for multiple-item products and for digital products, including DRM, usage constraints and licensing details. For examples and detailed notes, see separate documents on *How to describe sets, series and multiple-item products in ONIX 3* and *How to describe digital products in ONIX 3*. There is little change in the handling of most single-item physical products. Provision has been added to allow a country of manufacture to be specified when this is required for cross-border supply.
PR.4		*Epublication detail*: this data element group has been deleted, since product form description for digital products is now integrated into Group P.3.
PR.5	P.5	*Series*: the *Series* composite in Release 2.1 has been replaced with a new *Collection* composite. This carries collective attributes (of series or sets) when they are not carried in P.6 (*ie* where they are not required as part of the distinctive title of the product). A new *Collection sequence* composite can carry details of ordered collections). For examples and detailed notes, see separate document on *How to describe sets, series and multiple-item products in ONIX 3*.
PR.6		*Set*: this data element group has been deleted. Insofar as sets have collective attributes which are required as part of the ONIX description of individual products within the set, they are now handled in the same way as 'series', in P.5.

2.1	3.0	Notes
PR.7	P.6	*Title*: the *Title* composite in Release 2.1 has been replaced by an expanded *Title detail* composite. This can include collective title elements (of series or sets) when these are required as part of the distinctive title of a product. For examples and detailed notes, see separate document on *How to describe sets, series and multiple-item products in ONIX 3*.
PR.8	P.7	*Authorship*: there has been little or no significant change in the elements typically used for a personal or corporate contributor name. However the *Contributor* composite has been restructured to make it more logical, and the *Name identifier* elements have been redefined. Associations between contributors and places can be defined in a more flexible way, and the gender of a contributor persona can be specified. The handling of *Unnamed persons* has been modified to make it possible for anonymous contributors to have identifiers, alternative names *etc*.
PR.9	P.8	*Conference*: the entire composite has been renamed *Event* to allow more flexible use for events other than conferences (*eg* artistic or sporting events). Some redundant *Conference* elements have been deleted altogether and the remainder have been deprecated.
PR.10	P.9	*Edition*: no significant change.
PR.11	P.10	*Language*: redundant elements have been deleted. A new element allows 'script' (Latin, Cyrillic, Arabic *etc*) as well as language to be specified. Otherwise unchanged.
PR.12	P.11	*Extents and other content*: redundant elements have been deleted, so that *all* extent types, including number of pages, are now handled in a slightly extended *Extent* composite. Minor adjustments in the treatment of illustrations and ancillary content.
PR.13	P.12	*Subject*: redundant elements have been deleted, and both main and subsidiary subjects are now handled as repeats of a single *Subject* composite. The *Person As Subject* composite in Release 2.1 has been slightly restructured and extended, and renamed as *Name As Subject*, consistent with changes to the *Contributor* composite in P.7. From version 3.0.4, its use is broadened to encompass the names of fictional entities as well as of real persons and corporations.
PR.14	P.13	*Audience*: redundant elements have been deleted, otherwise no change.
		Block 2
PR.15, PR.16	P.14, P.15, P.16	*Descriptions and other supporting text / Links to image/audio/video files*: these two data element groups have been replaced by three new groups, significantly extending the capability for handling supporting materials, particularly resources made available on the web. Collateral material can be targeted at specific markets or geographical areas.
PR.17	P.17	*Prizes*: one redundant element has been deleted, otherwise no significant change. Prizes given to authors for their whole body of work (rather than for a specific book) should be specified within *Contributor* in P.7.
		Block 3
PR.18	P.18	*Content items*: changes are mostly limited to those which follow from the revision of some of the composites from other data element groups which are re-used here. In addition, content items may have *Related products* in addition to *Related works*. Both text and audiovisual content items are supported.
		Block 4
PR.19	P.19	*Publisher*: redundant elements have been deleted. The *Imprint* and *Publisher* composites have been revised so that handling of identifiers is consistent with general ONIX practice. Provision for providing promotional contact details on a per-product basis has been added, as has provision of funding details for open access publications.

2.1	3.0	Notes
PR.20	P.20	*Publishing status and dates, and copyright*: individual date elements have been deleted and replaced by a repeatable *Date* composite. For books published in international markets, it is no longer mandatory that there will be a single publishing status and pub date in P.20, if that 'global' status is unknown (a publisher would be aware of the global status, but other ONIX senders may not be). Instead, status and date can be specified for individual markets in P.25. A new element for Latest Reprint Number has been added, for use only in certain countries where this information is understood to have legal significance. Copyright details are more flexible, and may also be used to specify Phonogram rights.
PR.21	P.21	*Territorial rights and other sales restrictions*: the *Not For Sale* composite has been deleted, with adjustments to the *Sales Rights* composite so that all cases can now be handled within this single composite. A new *Rest of World sales rights* data element has been added to define the rights associated with countries not specifically listed (which can mean an explicit statement that they are unknown). The country, region and territory elements have been replaced by a new *Territory* composite which is also used consistently in P.14–16, P.24 and P26, so that wherever a geographical market is specified it is handled in the same way. The underlying logic, however, remains the same. Provision has also been added to allow a non-geographical restriction to be stated with date limits, and interpretation problems relating to retail exclusivity have been resolved by moving the *Sales restriction* composite inside the *Sales rights* composite.
PR.22		*Dimensions*: the Measure composite has been moved to P.3, and redundant elements have been deleted. Consequently this data element group is no longer required.
		Block 5
PR.23	P.22, P.23	*Related products*: a new *Related Work* composite has been added as P.22. The *Related Product* composite has been radically cut back, by general agreement, so that it carries only a relation type and an identifier.
		Block 6
PR.24, PR.25, PR.26	P.24, P.25, P.26	*Supplier, availability and prices / Market representation / Sales promotion information*: these three data element groups have been significantly restructured within a new *Product Supply* block. P.24 now specifies a territorial market together with any non-territorial sales restrictions which are specific to that market. P.25 details the publishing status of a product within a particular market. P.26 details the distribution source, availability, and prices of the product within a particular market. Within the new structure, many of the most frequently used elements remain the same, though there are some detailed changes and additions, notably a new and more flexible *Tax* composite, price conditions for rentals and linked product offers, tiered pricing, and a facility for provision of prices of other, comparable products and for price identifiers. The treatment of free of charge and other unpriced items is restructured, there is new provision for specifying minimum orders, and a new *Price constraint* composite can be used to specify certain contractual limitations linked to prices. Where a product is made available under different licences at different price points, links to the licence can also be included. The recommended treatment of reissues has been changed and the *Reissue* composite deprecated.

It should be noted that from the point of view of a data supplier with a well-designed product data management application that is already 'ONIX compliant', the migration from 2.1 to 3.0 can be viewed as little more than 'developing a new report format' (and one which can re-use large parts of the older report format).

There should be little need to modify or restructure the application's underlying database. The most likely areas where some restructuring might prove necessary are in the storage of titles (in order to accommodate changes in the way collections are defined) and in the management of markets and suppliers. ONIX 3.0 also encourages delivery of comprehensive sales rights information where 2.1 tolerated minimal sales rights, so work may be necessary to improve the underlying data available in the application's database.

However, for most ONIX implementers, the opportunity should be taken to improve compliance with the standard, to improve the quality of data supplied and increase the range of data elements supported.

Some major recipients and data senders may consider implementing the ONIX Acknowledgement message, in order that recipients can provide automated and structured confirmation of receipt or error information back to ONIX message senders.