

Data Recipient Best Practice

Pathway to ONIX

Prepared by Stephen Long May 2025 Version 1.0

Copyright © 2025 Book Industry Communication Ltd.

Disclaimer:

Please note: The information provided in this document is intended for guidance purposes only. Those involved in the creation, collection, management or distribution of product metadata are strongly advised to seek guidance on compliance with the business policies of their respective organisations.



Index

Executive Summary	4
Introduction	4
Purpose	5
Business Need	6
Current Situation	6
Supply Chain Map	8
The Solution	9
Resources Needed	9
Summary	10
Glossary of Key Terms	11



Executive Summary

ONIX - or ONline Information eXchange - is the international and most widely used standard for capturing and communicating physical and digital book product information electronically. ONIX messages are passed between organisations in the supply chain as a way of sharing information about these products. The format is XML, which put simply means that the messages are structured in such a way to be both human and machine readable. The standard is governed by a strong set of principles that are used to construct the messages. The intention is to automate the exchange of what is increasingly rich and time sensitive product information.

A short and non-technical pre-recorded video briefing on ONIX is available from EDItEUR (see doi: 10.4400/anjr).

The agreement between the sender and receiver to exchange ONIX messages will generally reflect a daily or weekly schedule, or sometimes 'as needed'. Once a data feed has been established with an initial message containing the full set of product records, subsequent messages in the feed normally contain only product records for new products and updates (more accurately, replacement data) for existing products.

The purpose of this document is to provide an overview of ONIX to non-ONIX users wanting to adopt the standard for managing and communicating their product information. The document provides useful background not only on what ONIX is, but also the steps required to transition from a bespoke data exchange to an internationally recognised standard.

1). Introduction

BIC's Data Recipient Best Practice Project has a mandate to agree, document and promote best practice for the collection and management of product information by the various recipients of that metadata in the UK and on the Island of Ireland. By best practice we mean those work practices recognised as being best in class and proven to deliver optimum results when it comes to the collection, handling and management of product metadata.

Ordinarily, any given area of the book supply chain is governed by multiple standards and the best practice routines that control those standards' use. B2B ordering is a good example, where there are multiple national and international EDI standards to choose from, not to mention newer XML standards and numerous bespoke formats too.

Product metadata is different. There is one international standard – ONIX. Since its inception ONIX has evolved to reflect the changing needs of the global supply chain. Unless otherwise stated, when referring to ONIX, we are specifically referring to ONIX 3.1 (the current release published in March 2023 and most recently updated in October 2024). The associated code lists (also called controlled vocabularies) are updated on a quarterly basis. Information on the current ONIX release and code lists can be found at <u>www.editeur.org</u>. EDItEUR's documentation, and the use of the ONIX standard itself, are both free of charge. Please note that release 2.1 and all earlier releases of ONIX are now considered obsolete, and ONIX 3.0 will not have any further updates.



2). Purpose

Product metadata can be thought of as all the data used to describe and trade products through the supply chain. This encompasses both simple, structured and factual information like titles, author names, distribution arrangements and prices, as well as more complex descriptive data, classifications of various types and even parts of the book itself (a table of contents can be seen as valuable descriptive metadata).

Both business intermediaries and end consumers have become used to having access to increasingly rich and detailed product information, allowing them to make informed buying decisions and self-serve without recourse to publisher sales reps (in the case of bookstores) or needing to hold and browse a physical copy (in the case of a consumer).

Today's technology means that many titles no longer go out of print. Short print runs and print on demand solutions mean that it can be cost effective to print small numbers of physical copies. Of course, the advent of digital means that even printing is no longer a prerequisite for getting books into readers' hands. As new products are published, the long tail of publications that remains available grows. Each title must fight harder to be found and chosen.

As the need for ever richer metadata increases and the number of product records grows exponentially, so does the overhead for creating and managing metadata records. These records are dynamic. During the lifecycle of the product, various attributes of the product will change (think price and availability). There is a requirement to regularly update products to maintain accuracy. New publications will come to market just as others may go out of print or be permanently withdrawn from sale.

Product record completeness and accuracy enhances discoverability – that's one key aim of metadata, but not the only one. Having made the product discoverable, it also needs to be efficiently tradable. Orders generate revenue and sales, and ONIX aims to reduce business-to-business customer support costs and alleviate issues within the supply chain itself by ensuring the necessary commercial information is readily available.

Historically, the investment required to implement ONIX was deemed too costly for many smaller organisations, particularly where:

- a. The product inventory was smaller, or niche
- b. Where data recipients hadn't expressed a need for ONIX files
- c. Where there were bespoke solutions already in place for the creation, management and exchange of product records
- d. Where responsibilities were outsourced

That's no longer the case. Several third parties and title management systems vendors offer product information management software or cloud-based solutions, and smaller organisations can take advantage of the ONIX 'ecosystem'. An in-depth knowledge of ONIX is no longer required if using a cloud-based solution. Once the product metadata is input, it is the responsibility of the solution provider to validate the data, transform it into ONIX files and communicate that information to a pre-agreed set of data recipients on behalf of the owner/ creator of the metadata.

Nowadays ONIX represents the optimum way of communicating standardised and highly structured metadata from one party to another within the supply chain.



The Value of ONIX

- a. It reduces the costs of data entry and data processing through to automated communication and data sharing within an organisation and across the supply chain
- b. In theory, one comprehensive feed can be repurposed for multiple data recipients, though in practice some limited customisation of data feeds for particular recipients is sometimes required
- c. It is the basis for excellent metadata quality and consistency, promoting sales growth
- d. It offers data interoperability across diverse IT systems, enabling growth of an ecosystem of off-the-shelf software
- e. A truly international collaboration, it is suitable for use across geographies, languages, scripts, currencies, and for all types of books, whether digital or physical, for audio, and for many other book-related product types
- f. It's common understanding of metadata concepts uses an accepted industry-wide lexicon that's understood at the points of creation and of use
- g. It offers the opportunity to develop robust business processes for the creation and maintenance of metadata
- h. It is a consistent and stable data model that reduces change management issues, promotes staff mobility and future-proofs investment in training
- i. It provides semantic compatibility with other standards (for example MARC or BIBFRAME) that enables the exchange of data across the broader sector
- j. The standard is created by expert practitioners and both informs and is informed by national and international consensus and metadata best practices
- k. Robust governance, managed by EDItEUR, provides continuity and protects the investment in metadata and systems
- l. The BIC UK National ONIX Group alongside similar groups in other countries informs the future development of the specification and code list releases to reflect stakeholder needs

3). Business Need

If ONIX is the best way of communicating ever more richer and growing volumes of metadata, what is stopping more organisations from adopting the standard?

Key challenges include:

- a. Criteria for determining that a switch to ONIX makes sense
- b. Understanding the steps and associated costs that any transition involves
- c. Knowing when and where to start
- d. Appreciating the value of ONIX by seeing it in action
- e. Understanding how ONIX differs from other solutions for communicating metadata
- f. Assumption that everyone in the book industry *must* be conversant with ONIX
- g. Awareness of and access to available resources
- h. Knowing what to do when things go wrong
- i. Managing changes to long-established ways of working
- j. Recognising that the volume of new publications requires some automation
- k. Risks of comparing pre and post ONIX worlds where ultimately different processes, ways of working and timescales are involved
- l. Realising that metadata requires continual curation

4). Current Situation

At the moment, a further inhibitor to ONIX adoption is the range of alternatives available:



- a. Retailer, platform or vendor specific arrangements such as Amazon Seller Central. Users share their product information according to the requirements of that retailer, platform or vendor. Where the user only has one or two well established trading relationships, the manual entry or upload of data in a specific format may be achievable.
- b. The bureau service. Offered by several third parties including data aggregators, wholesalers, solutions vendors and title management providers. Users subscribe to a service and share their data in a non-ONIX format for transformation and onward communication by the bureau (that onward communication may however be in ONIX). Data can be manually input via an online portal, shared via printed catalogues and advance information sheets or uploaded via a pre-agreed and formatted template (CSV is common).
- c. Purely manual solutions, reliant on the sharing of catalogues or advance information sheets via postal services, email or fax with key trading partners. Whilst requiring no automation, and only limited resources, such an approach comes with all the risks associated with relying on humans to share, prioritise and input data in a timely and accurate manner.
- d. A multi solution approach involves all the above to a greater or lesser degree, dependent on the type, range and specific needs of trading partners.

The type of business may mean that one or all the above options for communicating metadata are perfectly acceptable and reasonable. However, there are potential risks to consider:

- e. The resources required to input or upload data (human effort, financial cost, time).
- f. The potential for delays where there is no automation or appreciation of the importance of the data.
- g. Any process involving manual input to text files or spreadsheets, or manual upload may come with little or no validation leading to further potential delays and the possibility of errors in product records.
- h. Dependent on the bureau service, the subscription may be free or paid for. Free subscriptions usually come with limits on the volume or depth and richness of the data. Data may also be put in a queue of product information from other non-paying subscribers, to be processed as and when resource and time allow (first come, first served). There may be no named contact at the bureau to speak to if there are problems with the product information.
- i. The involvement of any third party creates a dependency, with limited visibility of the third party's processes once the product information has been received.
- j. Each data recipient will have specific criteria to manage what data is accepted or acceptable. The involvement of an intermediary may mean that the criteria is confidential (between the third party and the recipient), poorly documented or completely opaque to the data sender. Where there are data delays, errors or rejections, the cause may not be immediately apparent.
- k. Every time product metadata is subject to manual intervention or intermediate steps between its creation and collection by the data recipient, there is a risk of delay, inaccuracy or error creeping in.

Overleaf is an illustration of how metadata moves around the book supply chain. Many organisation types not only create and curate metadata, but also pass it along the supply chain, not necessarily to the ultimate recipient, but to further intermediaries.





5). The Solution

In terms of best practice, there is only one current and recommended standard, namely ONIX 3 (which encompasses both 3.0 and the newer 3.1).

Robust governance of ONIX, managed by EDItEUR with a range of user groups (including BIC's UK ONIX National Group), provides continuity, protects the investment in metadata and reflects evolving business needs in the book industry.

- a. The current version of ONIX has a logical structure. It is modular, extensible and predictable, meaning that there is usually only one clear way of doing something. This makes it simpler for developers to work with.
- b. It works with different languages, different scripts, different currencies etc, and is suitable for global use.
- c. The standard is well documented, via both the *Specification* and EDItEUR's *Global Implementation and Best Practice Guide.* These are supported by a range of application notes and other collateral. This makes the requirements to ensure global interoperability clearer.
- d. There is a wide range of XML tools available different XML schema options including the 'strict schema' for validation purposes.
- e. 'Block Updates' offer a more efficient way of updating product information.
- f. ONIX 3.0 and later treat digital products as a core part of the product (think streaming products and digital licensing).
- g. It offers a logical and universal approach to the treatment of sets and series (*i.e.* 'collections').
- h. It features a single consistent method of limiting sales rights and distribution rights, the applicability of prices and the geographical targeting of collateral material.
- i. It offers prioritisation of items of marketing collateral (*e.g.* which of the five reviews is the most important).
- j. International distribution arrangements are captured in a logical and flexible way, reflecting real-world international distribution agreements.
- k. Product pricing is more flexible and covers a range of contemporary business models including rentals, bundling, *etc.*
- l. It allows for more information about works and derivation of works.
- m. It carries parallel multilingual metadata (e.g. an author biography in multiple languages within the same product record).
- n. It can include information about promotional events (e.g. book signings, festival appearances), and manufacturing detail for POD, e-book and digital audio products. This can help eliminate 'sidecar files' on those workflows data that could be included in a single structured metadata file that would otherwise need to be in some associated file.
- o. ONIX continues to evolve in a controlled manner to meet new business requirements as they arise.

In simple terms, ONIX facilitates the creation, management and communication of product metadata globally. As it is truly an international standard, it caters for the needs of multiple markets to reflect local language, currency, classification, sales and distribution needs.

In theory, one feed can be repurposed to fulfil the requirements of multiple data recipients (both intermediaries and those further along the supply chain), without a significant overhead. It removes the need for bespoke, trading partner specific solutions that have a complexity and overhead all their own.

6). Resources Needed

Given that ONIX is already in widespread use globally and is almost ubiquitous among large and midsize publishers in the UK, US, Canada and many countries across western Europe, the optimum way in which senders and receivers should exchange product information does not warrant a project to



determine the best way forward. ONIX is the way forward.

In effect, each organisation interested in the adoption of ONIX needs to have a clear understanding of what's involved, the resources available and an appreciation of its potential impact on their respective supply chains. With this information it can establish its strategic importance.

a). What's Required

From a technical viewpoint, knowledge of the XML format is required – but only if an organisation plans to develop its own data management application. ONIX is specified in an XML schema.

Those with commercial responsibilities should understand the supply chain and how the use of ONIX can simplify, standardise and automate the creation, management and communication of increasingly rich and complex product information. Reducing customisation and the need for ad hoc processes can lessen support overheads whilst also enhancing business relationships.

Not all players in the metadata supply chain use the same terminology, so standardisation is the ideal way to remove potential confusion and misunderstandings.

b). Resources and Expertise Available

The ONIX Specification, an extensive *Implementation and Best Practice Guide*, and a range of XML schemas are available on the EDItEUR website: <u>www.editeur.org</u>. EDItEUR runs an email list and discussion forum via Groups.io (anyone can join – send a blank email to <u>onix+subscribe@groups.io</u>) and it manages the ONIX for Books International Steering Committee. BIC operates the UK ONIX National Group which is open to BIC/ EDItEUR members.

c). Potential Impact

- a. Addresses a current supply chain challenge without the need for manual intervention.
- b. It enhances product discoverability and tradability.
- c. As an actively maintained, global standard, it provides a recognised and authoritative set of guidelines reflecting book industry needs.
- d. It removes the need for bespoke, trading partner specific formats and routines.
- e. It reduces development, support and cost overheads.
- f. It reflects the needs of the book supply chain and its stakeholders:
 - i. The product range is increasingly diverse and growing (physical, digital, book, non-book).
 - ii. Rich, market specific product information drives online sales and frequent changes to price and availability information. Regular and timely updates to information are critical to a well-oiled supply chain.
- g. It complements and extends the use, relevance and applicability of current XML standards, reducing the development and support overhead for today's technical teams.
- h. It comes with a range of XML tools (schemas) that can ensure your ONIX data is properly structured (although these tools cannot ensure that the data is accurate).
- i. It allows benchmarking of an organisation's metadata against industry expectations via BIC's Metadata Excellence Award (MEA) accreditation scheme.

7). Summary

Using an existing business standard that has been developed collaboratively, with the needs of both current digital and physical book supply chains in mind, offers a straightforward way to replace, simplify and automate manual processes.

Extending the use of ONIX will deliver a more efficient and resilient supply chain – one where misunderstandings and queries relating to out-of-date product information are removed, where



manual intervention no longer risks causing delays or errors to the exchange of information and where globally one standard and controlled vocabulary that everyone understands form the basis for doing business.

8). Glossary of Key Terms

A more extensive glossary of terms related to Data Recipient Best Practice is available via the BIC website at www.bic.org.uk

BIBFRAME: A relatively new data format for library bibliographic data, intended to replace MARC.

Block Update: A partial update of a product record in ONIX.

EDI or Electronic Data Interchange: A structured, technical format for the exchange of time, business or financially critical transactions in an automated and standardised form between computers. Transactions are processed in batches rather than in real time.

MARC or Machine Readable Cataloguing: A family of metadata formats used in library cataloguing since the 1960s. MARC21 is the latest version in use in North America and the UK.

Metadata: Strictly, data about other data. More usefully in the context of the book and e-book supply chain, metadata can be thought of as all the data used to describe and trade products through the supply chain. This encompasses both simple, structured and factual information like titles, author names, distribution arrangements and prices, and richer, more complex descriptive data, classifications of various types and even parts of the book itself (a table of contents can be seen also as valuable descriptive metadata). ONIX messages are a method of communicating standardised and highly structured metadata from one party to another within the supply chain.

ONIX or ONline Information eXchange: The international and most widely used standard for capturing and communicating book product information electronically. This includes both physical and digital products.

ONIX Feed: Arrangement and supply of an ongoing series of ONIX messages exchanged between data sender and recipient, generally on a daily or weekly schedule or 'as needed'. Once a data feed has been established with an initial message containing the full set of product records, subsequent messages in the feed normally contain only product records for new products and updates (more accurately, replacement data) for existing products.

ONIX Message: A complete ONIX data file, generally one in a series of messages (an ONIX feed) passed between a data provider and a data recipient. A single message may contain one or many product records, or none.

Schema: Describes the structure (and rules) relating to an XML document (such as an ONIX message). The XML schema is used to ensure that the document meets the necessary requirements for a valid XML document.

XML or Extensible Markup Language: A set of codes or tags that describes text in a digital document. The structured human/ computer readable format is used in the book industry for the latest standards such as ONIX.

