# ONIX for Books release 3.0.5

**This document summarises the small number of additions and minor changes made in ONIX 3.0 revision 5.** Initial proposals based on requests arising over the period since the release of 3.0.4 in late 2017 were made to the ONIX International Steering Committee. A subset of these requests were ratified directly at the Steering Committee's meeting during the Frankfurt Book Fair in October 2018, and these form revision 5. The remainder of the proposals will be considered further alongside other potential additions for a future revision 6.

## Background

Since 2010, ONIX 3.0 has maintained a roughly biennial cadence of minor revisions. The previous revision was in late 2017, so this proposal is 'early', but the key proposal below is intended to meet a metadata requirement that is evident *now*, rather than being somewhat forward-looking as most revisions are.

For clarity, these enhancements are entirely backwards-compatible. While there have been one or two deprecations, nothing in ONIX 3.0 from 2010 has been removed, nothing mandatory has been added, and correctly-constructed data from that period remains fully compliant with the current 3.0.5 version of the standard.

Full use of this release of ONIX 3.0.5 requires the use of Codelists Issue 43 or later, which is available from the EDItEUR website and via the online multilingual ONIX codelist browser. [1]

## 1. Audio chapter timecodes

There is a growing requirement for audiobook publishers to provide 'tables of contents' with their digital products. In this case, instead of a page number for each chapter, the requirement is a *timecode* – the hours, minutes and seconds point at which a chapter starts within the audio.

Now, for printed products, tables of contents can be delivered in two ways. By far the most common is as *unstructured* data – essentially just text in either <TextContent> or <SupportingResource>, in Block 2 of ONIX 3.0. If using <TextContent>, embedded markup can be used to structure the table of contents as an (X)HTML list. Tables of contents are known to be used by some major online stores: they display the TOC to potential customers and index the text of the TOC for search purposes.

---

[1] https://www.editeur.org/14/code-lists and https://ns.editeur.org/onix
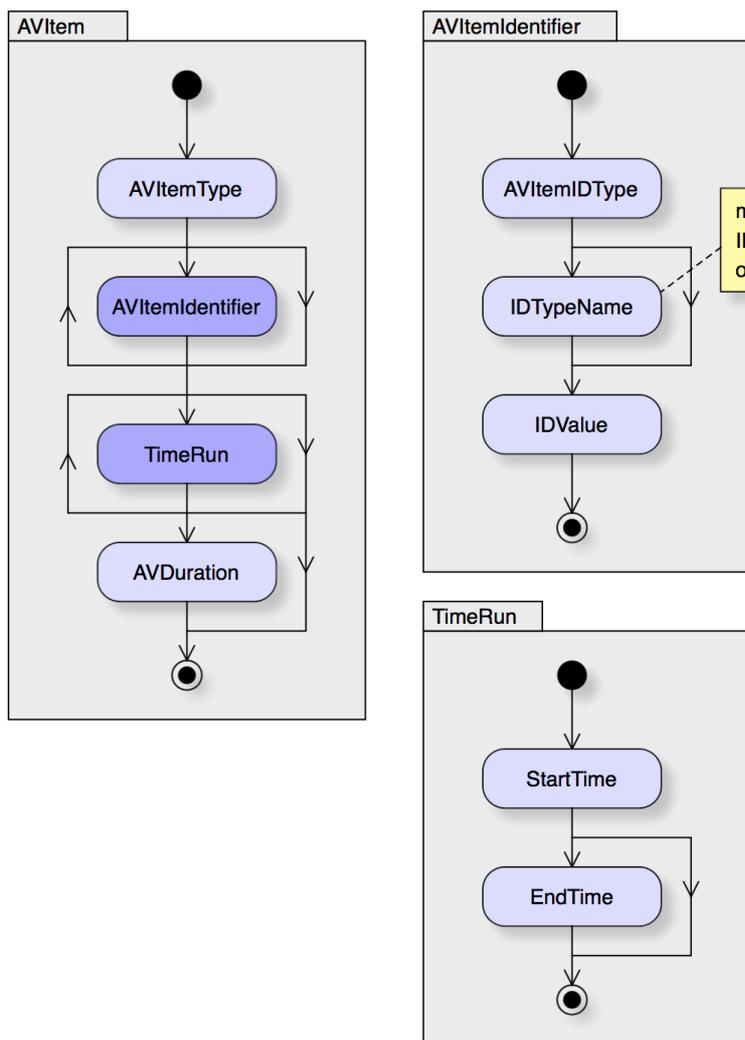
The second method – rarely used except for scholarly texts – is to use Block 3 of ONIX 3.0, with multiple repeats of the <ContentItem> composite. Within <ContentItem>, the <TextItem> composite holds page numbers for each chapter of the book, and other parts of <ContentItem> specify contributors and other marketing material specific to that chapter. However, <TextItem> is not suitable for audio material.

The addition in ONIX 3.0.5, diagrammed left, is a new composite, <AVItem>, which is used *instead* of <TextItem> for audio (and possibly video) products. It closely mirrors the structure of <TextItem> but with elements more suitable for time-based rather than paged media.

- <AVItem> contains a type code similar to that in <TextItem>, where the existing codes are used to indicate whether the item is part of the front matter, the back matter or the main body of the product.
- <AVItemIdentifier> allows each chapter of the audio work to be given an identifier, separate from the identifier for the whole product or whole work.
- <TimeRun> and <Duration> give the start and end timecodes and duration of the chapter, always in the form HHHMMSS.

One practical note: digital audio products are large files – certainly when delivered in a lossless format, but even when delivered in lower quality compressed formats like mp3. For example, 10 hours of stereo CD-quality audio is 6GB. Even a compressed mp3 file can be 1MB per minute of audio (so 600MB for ten hours). This often means that digital audio is delivered in chunks (*ie* several individual files), rather than as a single gigantic data file. Chunks are not necessarily aligned with chapters or other subdivisions of the product. *It is vital to note that the timecodes are relative to the start of the product, and take no account of the chunking of the audio data*.

This is because products are occasionally *re-chunked* – that is, chunk sizes might be adjusted within the supply chain, as files might be transcoded into a different compression scheme, making existing chunks too large or too small. Using times relative to the start of the product greatly reduces the

likelihood of confusion – chunk size is only relevant to timings at the final point of delivery to the consumer.

> Note: some digital audio retailers call these chunks *tracks*, by extension from the tracks on a physical CD, but a typical chunk is much longer in duration than a typical CD track.

## 2. Other minor additions

**3.1.** A minor limitation on the ability to Block update previously-supplied data has become apparent. Some Blocks – specifically Blocks 1, 4 and 6 – contain at least one mandatory element, whereas Blocks 2, 3 and 5 do not. This means that it is not possible to use a Block update to 'reset' all previously supplied marketing collateral (*eg* to replace all Block 2 marketing collateral with nothing, or to remove all previously-specified related products and works, since <CollateralDetail> or <RelatedMaterial> would be empty elements – which previous versions of the ONIX specification do not allow).

> Note: a 'reset' of previously-supplied Block 2, 3 or 5 data *can* be done with a full-record update, simply by omitting the Block.

The documentation for ONIX 3.0.5 has been modified specifically to allow Blocks 2, 3 and 5 (only) to be empty in Partial or 'Block updates' (only). The blocks must not be empty in ordinary 'full record' messages.

**3.2.** In some countries, for example Chile, books are specifically *exempted* from VAT or Sales tax. This status is different from being taxed at zero % (as is the case for VAT on physical books in the UK, for example), and different from prices supplied without Sales tax details (as is the case in North America). To cover this, a new <TaxExempt/> empty element is added as an alternative to <Tax>, for use with 'ex tax' prices where the product is specifically tax-exempt.

**3.3.** A <PalletQuantity> tag has been added, similar to <PackQuantity>, to contain the number of copies of a product on a pallet or skid.

<div align="right">

Graham Bell, Chris Saynor

EDItEUR

26th October 2018

</div>