Cataloguing in a library, shelving in a bookshop? Different, but with so much in common.

Both rely critically on metadata. In the case of a library, ensuring patrons can find the resources they need is all about the catalogue, full of in-depth, authoritative metadata about the physical and digital holdings. In a store, physical shelving is often driven by subject, by publication date (those 'new books' shelves), and so on. Online, a bookstore depends on metadata to drive discoverability and purchase. And whether it is in a library, or in a store, standards underpin the exchange of metadata across the information and publishing sector. But although on the surface, the data needs might be thought to be very similar, the data used in libraries and data bookstores looks entirely different, the supply chains – both for the books and the data – are quite different, and the demands on store and library staff are different too.

Focusing on the data itself illustrates some of the contrasts. Library metadata standards like MARC or RDA are well-established, but aren't used widely in the publishing world, other than by a few academic publishers or suppliers whose main customers are libraries. Publishers in North America, across Europe and increasingly in the Asia-Pacific region predominantly use a standard called ‘ONIX’ to distribute data about their book products to bookstores. Libraries use Dewey or UDC for subject classification, whereas publishers and booksellers use a range of subject schemes – ‘BIC’, ‘BISAC’ or a host of other country-specific schemes, and increasingly a global scheme called ‘Thema’.

So who in the commercial publishing and bookselling world creates ONIX or BISAC, CLIL or Thema? These are ‘trade standards’, the prosaic ‘plumbing’ of the data supply chain on which the industry depends, defined and promoted by trade organisations for the benefit of the industry as a whole.

EDItEUR (https://www.editeur.org/) is one such organisation. London-based and with just three and a half staff, it is – like most similar organisations – a not-for-profit, and is best known today for creating, supporting and promoting ONIX for Books, an XML-based message format used widely for communicating metadata about books and book-like products between publishers and booksellers.

Why is ONIX different from MARC, when both broadly describe the same books? Part of the answer lies in the contrast between the librarian’s ‘resource’ and the publisher’s ‘product’. The latter has an overtly commercial outlook, and part of the aim is to maximise sales of those products. ONIX as a standard recognises this – so rather than remaining a purely bibliographic description of the book and its content, ONIX also encompasses a range of marketing, rights and commercial information that would be of little importance to the librarian, but which are vital in reducing ‘friction’ in the book, e-book and audio supply chains. Another contrast is seen in the dynamic nature of the commercial metadata. Publishers begin to distribute information about a new book several months in advance of publication, and continue to update, supplement and even revise that data, not only in the months leading up to the publication date but throughout the lifecycle of the book.

So ONIX has a range of commercial stakeholders, from publishing companies of all sizes to book retailers, data aggregators and logistics organisations. EDItEUR’s membership is built from these stakeholders. Their commitment is a sign that they support an open, standards-based data supply chain. Through direct contact with EDItEUR, members have a voice in the future direction of standards development. EDItEUR relies almost entirely on membership revenue. And yet EDItEUR makes the standards specifications available free of charge to all – because the best outcome for members is that **everyone else uses the same standards too**.

---

**Catalogue and Index**
The development and governance process for ONIX ensures benefits flow broadly and without favouring particular commercial organisations: EDItEUR encourages its members around the globe to form ‘national groups’ representing the whole data supply chain in that country, and ultimately, it is these national groups rather than EDItEUR that have the final say on developments to the standard.

That development and governance model, emphasising wide participation in national user groups, EDItEUR membership to pay the bills, and free availability of the standard for all, is also used for Thema, EDItEUR’s global subject category scheme. Widely used library-focused schemes such as Dewey or UDC are both too complex and too limiting for commercial use, and the book business has developed numerous nationally-focussed subject schemes. BISAC is one such, developed and promoted by the New York-based Book Industry Study Group, and Thèmes CLIL is its quite different French equivalent, but there are many tens of other national schemes creating a babel of competing schemes in the international book trade. Thema is an attempt to create a commercially-relevant subject scheme suitable for global use, and it has grown fast: it’s just five years old, but has already supplanted established national schemes in many European countries. In the UK, Thema is best viewed as ‘the latest update of the previous BIC subject scheme’, and as a development from the BIC scheme is easy to adopt for UK publishers and retailers.

For EDItEUR, the international dimension to standards is critical: our industry is increasingly international in scope, our members are spread internationally, and publishing and bookselling issues are often common across many countries. The standards themselves reflect this:

```
<Contributor>
  <SequenceNumber>1</SequenceNumber>
  <ContributorRole>A01</ContributorRole> <!-- author -->
  <NameIdentifier>
    <NameIDType>16</NameIDType>  <!-- ISNI -->
    <IDValue>0000000083888172</IDValue>
  </NameIdentifier>
  <PersonName>서영은</PersonName>
  <KeyNames>서</KeyNames> <!-- Seo -->
  <NamesAfterKey>영은</NamesAfterKey> <!-- Yeong-eun -->
  <BiographicalNote textformat="05">서영은 강원도 강릉 출생이며, 강원도 강릉사범학교를 졸업하였으며 건국대학교 영어영문학과를 졸업하였다.</BiographicalNote>
</Contributor>
```

The example extract of metadata describes a Korean writer, Seo Yeong-eun (서영은) – and you’ll note that the XML-based format of ONIX works as well in Hangul, or Arabic or Cyrillic, as it does with Latin. It can include text in many languages, in parallel if necessary. The standard specification has been adopted as a national standard in China (GB/T 30330), and is also the preferred format for delivery from publishers to Amazon, Flipkart (in India), BOL (in The Netherlands), or to Waterstone’s online stores. But the stakeholders are not all huge organisations – EDItEUR’s smallest members are two-person companies, who see the standard as levelling the playing field.
Equally, *Thema* is intended for global use. The whole multilingual scheme can be viewed at ns.editeur.org/thema. It’s not intended as a replacement for Dewey or UDC and lacks the detail and rigour demanded by academic libraries – but it might be a good fit for the public library service. A book like Colson Whitehead’s *The Underground Railroad* might have the *Thema* categorisation:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>FV</td>
<td>Historical fiction</td>
</tr>
<tr>
<td>FBA</td>
<td>Modern and contemporary fiction</td>
</tr>
<tr>
<td>NHTS</td>
<td>Slavery and abolition of slavery</td>
</tr>
<tr>
<td>1KBB-US-SE</td>
<td>US South Atlantic states</td>
</tr>
<tr>
<td>3MN6</td>
<td>1820s</td>
</tr>
<tr>
<td>5PB-US-C</td>
<td>Relating to African American people</td>
</tr>
</tbody>
</table>

This provides multiple ways of accessing the book, whether via a faceted search or a more traditional ‘browse by subject’. The codes (e.g. ‘FBA’) are independent of language, whereas the literals (‘Modern and contemporary fiction’) are available in around 20 languages so far with others on the way. MKE means the subject is dentistry, but also *zahnheilkunde*, *fogászat*, طب الأسنان and 牙科.

So if these standards have an overtly commercial tenor, where do they intersect with library practices? ONIX metadata records are usually distributed widely across the book trade several months prior to publication, so that ONIX data is often used to create an initial skeleton MARC record. The British Library’s CIP programme (operated by BDS Ltd.) works in this way, and the US Library of Congress is actively soliciting more ONIX data from publishers. The MARC data may later be enhanced (or even *corrected*, for publishers are less bound by cataloguing rules than are librarians) via a book-in-hand process, but its genesis frequently lies in ONIX. ONIX of course uses the same range of identifiers and authorities as library catalogues, so interoperability is good. Each product described has an ISBN, contributors can be listed with their ISNI or ORCID identifiers, and there is a publicly-available mapping procedure developed by OCLC to convert from ONIX to MARC (MARC to ONIX is less successful, because the commercial data that provides the value in ONIX would be missing).

In addition to developing and supporting its own standards like ONIX and *Thema*, EDItEUR also provides management support to other identifier standard organisations, in particular the ISNI International Agency (www.isni.org) and the International DOI Foundation (https://www.doi.org/) which manage the ISNI and DOI systems on behalf of the ISO. Until recently it provided similar support to the International ISBN Agency. All this, combined with EDItEUR’s global outlook and spread of members and contacts, gives EDItEUR an unparalleled insight into the use of identifier and metadata standards across the industry, and this expertise can be called upon by EDItEUR members when necessary.
Graham Bell, Executive Director, EDItEUR (ISNI: 000000427566266)

Graham is Executive Director of EDItEUR, responsible for the overall development of EDItEUR’s standards and the management services it provides on behalf of other standards agencies. He joined EDItEUR as its Chief Data Architect in 2010, focussed on the continuing development and application of ONIX for Books, and on other EDItEUR standards for both the book and serials sectors. Graham previously worked in IT for HarperCollins Publishers in the UK, and as an Editor in the magazine industry with Redwood Publishing and BBC Magazines.

EDItEUR in summary

- a global trade standards body for the book, e-book and audiobook sector
- not for profit, membership-supported
- responsible for ONIX, Thema, EDItX and EDIFACT standards
- provides management services for ISNI, DOI agencies

For more details see https://www.editeur.org or e-mail info@editeur.org