

E-books, Metadata and the Value of “Books in Print” databases

Brian Green, 

Vilnius, 30 November 2012

www.arrow-net.eu

ARROW Plus is a Best Practice
Network selected under the ICT
Policy Support Programme (ICT
PSP)



Agenda

- EDItEUR and ARROW Plus
- Why is metadata especially important for e-books?
- What is “Books in Print” and what is its value for and beyond e-publishing?
- Developments in Lithuania

What is EDItEUR

- The International Book Trade standards organisation
- Develops, supports and promotes identifier, metadata and transactional standards for the book, e-book and serials supply chains (e.g. ONIX for communication of information about books)
- Based in London, but with a global membership of publishers, distributors, wholesalers, retailers, serials subscription agents, libraries, system vendors, rights organisations and trade bodies
- Manages the International ISBN and ISTC Agencies
- Not-for-profit membership organisation

What is ARROW?

- Accessible **R**egistries of **R**ights Information and **O**rphan **W**orks towards Europeana
- A Best Practice Network under the ICT Policy Support Programme (ICT PSP) of the European Union
- Has developed a system that supports a “diligent search” process for libraries and others involved in projects for the digitisation of books in their collections.
- The ARROW system streamlines the process of identifying authors, publishers and other rights holders of a particular work, including establishing whether the work is in or out of copyright, commercially available, or orphan
- System already running in France, German, Spain and UK
- ARROW Plus extends the project to new countries, mainly in Eastern and Central Europe

The ARROW project and system

- Data providers to the project included:
 - ▣ the cultural sector represented by national libraries (national bibliographies)
 - ▣ the organisations which create and maintain “books in print” databases across Europe
 - ▣ the collective management organisations which maintain a network for the collective management of textual reproduction rights on behalf of authors and publishers (“reproduction rights organisations” or RROs).
- One of the crucial questions that the ARROW system asks is whether a title is still commercially available in print or available as an ebook. This requires a “Books in Print” database
- ARROW Plus aims to help establish sustainable Books in Print services in partner countries that do not have yet them: Hungary, Latvia, Lithuania (LLA), Poland and Portugal

Why is Metadata important?

- For e-books (and printed books sold online), metadata is the way that potential buyers will find your books. The equivalent of browsing in a bookshop
- Both platform-based retailers (e.g. Kindle, Kobo, Google, Apple) and platform-independent ebook “etailers” need rich information about e-books
- In addition to the data fields required for printed books, (e.g. title, author, publisher, ISBN, cover image, subject category, extent, description, author biography, price, availability etc.), e-books require some extra fields and information (e.g. ebook format/platform, user restrictions/DRM, extent (may be file size or number of pages in print equivalent), related print editions)
- N.B. e-books must have separate ISBNs from printed books and each separate format should have its own ISBN



International ISBN Agency

[Home](#)[News](#)[ISBN FAQ](#)[National Agencies](#)[Ranges](#)[International Agency](#)[ISBN Directory](#)[ISBN-13 Converter](#)[Agency Login](#)

5. Guidelines for the assignment of ISBNs to e-books

1. If I publish in two or more different file formats (e.g. epub, PDF) should I assign separate ISBNs?
2. Should different versions of the e-book that use the same file format (e.g. epub) have different ISBNs?
3. Assuming the same content, what are the features that distinguish different e-book products and determine whether separate ISBNs are required?
4. Is there a standard way of describing different product forms and DRM features?
5. If I provide a single master file to a conversion service and am not controlling the different combinations of file format and DRM provided by that service to retailers, should I assign an ISBN to that master file?
6. If my e-books are being supplied by a retailer that is the sole provider of e-books in a proprietary format that can only be bought its own website (e.g. Amazon Kindle) and that retailer does not require ISBNs, should I assign ISBNs to those versions?
7. I provide an e-book conversion service to publishers but they are not providing separate ISBNs for each version that I generate. What should I do?
8. E-book devices offer different features such as type size, text to speech, bookmarking, colour etc. Even if my content, file format and DRM are the same, the user experience varies according to the device used. How does this affect assignment of ISBNs?
9. Should an enhanced e-book that includes audio, video or other additional content have a different ISBN from the standard e-book?
10. Can e-book "apps" (e.g. applications for iPhone, Android etc.) have ISBNs?
11. I am publishing two versions of an e-book, one without DRM and one with 'social' DRM that does not enforce any restrictions on the user (e.g. watermarking). Do I need two separate ISBNs?
12. How can all the different versions of an e-book be linked together?
13. How do I identify individual chapters or other parts of a book that I plan to make available separately?

1. If I publish in two or more different file formats (e.g. epub, PDF) should I assign separate ISBNs?

Yes. Each different format of an electronic publication that is published and made separately available should be given a separate ISBN.

[↶ back to top](#)

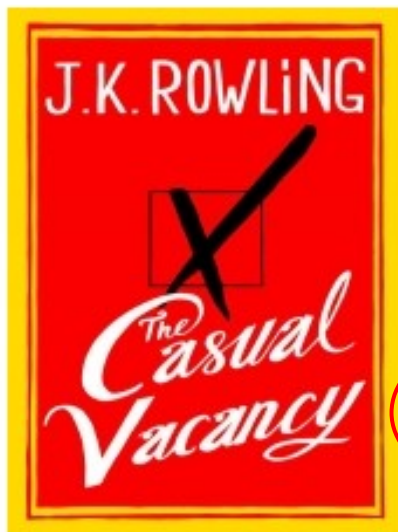
2. Should different versions of the e-book that use the same file format (e.g. epub) have different ISBNs?

If the different versions use the same DRM software (e.g. Adobe ACS4) with substantially the same settings and are interoperable on different devices or software, then a single ISBN should be used. If, however, the same DRM software is used on two versions but with significantly different settings (e.g. one allows printing but the other does not), then each


Communicating metadata

- “*Obscurity is a far greater threat than piracy*” (Tim O’Reilly)
- An ever-increasing number of trading partners will ask for this metadata and each one may require different data fields (e.g. formats, extent, DRM limitations)
- It would be cost-effective to send full metadata to a single trusted hub, if one exists, for transmission to all channels.
- Metadata is normally supplied to retailers and ebook aggregators via web-based forms, in Excel spreadsheets or in ONIX format.
- ONIX 3.0 is the international standard for communicating product information about printed and electronic books and has been designed to include all the data likely to be required (www.editeur.org)

FANATICS MEMBERS EARN **248 POINTS** ON THIS PURCHASE. [LINK YOUR FANATICS ACCOUNT OR](#)



The Casual Vacancy

 [epub]

Author(s): [J.K. Rowling](#)

Published By: Little, Brown Book Group

Date Published: 27 September 2012

Filesize: 780.58 KB

DRM Restrictions:

- No Copy Allowed
- No Print Allowed
- No Lending Allowed
- No Reading Aloud Allowed


Item Code: 9781405519229


Earn 248 Fanatics points [Find out how](#)



Unavailable in your region

Add To Wishlist

Email 

Print 

Digital items require



Get **ADOBE®**
DIGITAL EDITIONS



New to eBooks.com?

[Learn more](#)

Search by title, author, subject or ISBN



Popular Subjects

- Business
- History
- Computers
- Religion
- Health & Fitness
- Science
- Body Mind Spirit

Fiction

- General Fiction
- Children's & Young Adult
- Crime Fiction
- Fantasy
- Literary Fiction
- Romance
- Science Fiction
- Suspense/Thrillers

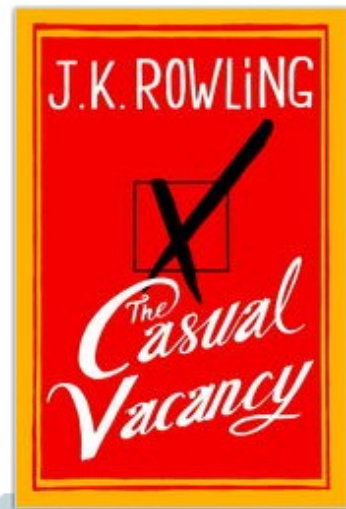
Non-Fiction

- Archaeology
- Architecture
- Art
- Biography & Autobiography
- Business & Economics
- Crafts & Hobbies
- Current Events
- Drama
- Education
- The Environment
- Family & Relationships

The Casual Vacancy

J.K. Rowling

Like 32 people like this.



Add to cart

£ 11.51 (incl. VAT)

[Preview \(read now\)](#)

When Barry Fairbrother dies in his early forties, the town of Pagford is left in shock.

Pagford is, seemingly, an English idyll, with a cobbled market square and an ancient abbey, but what lies behind the pretty facade is a town at war.

Rich at war with poor, teenagers at war with their parents, wives at war with their husbands, teachers at war with their pupils... Pagford is not what it first seems.

And the empty seat left by Barry on the parish council soon becomes the catalyst for the biggest war the town has yet seen. Who will [more](#)

Little, Brown Book Group; September 2012

480 pages; ISBN 9781405519229

[Download in EPUB](#)

More Fiction

- Fifty Shades of Grey**
E L James
2012 £ 3.49
528 pages
- Reflected in You**
Sylvia Day
2012 £ 3.99
400 pages
- Fifty Shades Darker**
E L James
2012 £ 3.49
384 pages
- Don't Blink**
James Patterson
2010 £ 5.49

Subject categories

Fiction > Genre Fiction

[News](#) [Wiki](#) [Blogs](#) [Tweets](#) [YouTube](#) [beta](#)



Jointly with Book Industry Study Group, New York, and
Book Industry Communication, London



4.2 Usage Constraint composite

A new composite element in ONIX 3.0, <EpubUsageConstraint> (P.3.17 to P.3.20) enables a usage type to be specified as *permitted unlimited*, *permitted subject to limit*, or *prohibited*. A quantitative limit can also be specified when applicable.

Code List 145 is used for usage types, with the following initial set of values:

01	Preview	Preview before purchase
02	Print	Print paper copy of extract
03	Copy / paste	Make digital copy of extract
04	Share / use on multiple devices	Use product on multiple concurrent devices
05	Text to speech	'Read aloud' with text to speech functionality

Code List 146 is used for usage status, with these values:

01	Permitted unlimited	
02	Permitted subject to limit	Limit should be specified in <EpubUsageLimit>
03	Prohibited	

What is Books in Print?

- **What it is NOT**
- NOT only about printed books
- NOT a National Bibliography
 - ▣ which includes books in and out of print (no discrimination)
 - ▣ has no updated price and availability data
- NOT a Wholesaler/Distributor/Internet bookseller database
 - ▣ which are limited to books carried by wholesaler/distributor/bookseller
 - ▣ often has only limited information and contact details for publishers and distributors
 - ▣ availability refers to wholesaler or bookseller rather than the publisher

Books in Print

- A trusted book trade hub for product information
- Provides aggregated information to the entire book supply chain
- Listing of all books available
- ... or soon to be published
- Including e-books
- Contact details for publishers / distributors
- Comprehensive in coverage
- Includes descriptive/marketing information
- **Provides updated information on current price, availability and source (i.e. distributor)**

What are the benefits of BiP to the book trade?

- Provides all the information that potential purchasers need to discover, make a purchase decision and obtain the book
- Accurate and up-to-date data about every book available for sale
- Encourages backlist sales and facilitates customer orders
- Promotes export sales (books and licences)
- Enables electronic ordering routing and other valued-added services (sales data collection, anti-piracy systems etc)
- Facilitates online bookstores (including ebook aggregators)
- Enables licensing of in-commerce books for digitisation
- Provides a “hub” for book information
- Saves time and money
- Helps sell more books

Helps sell more books

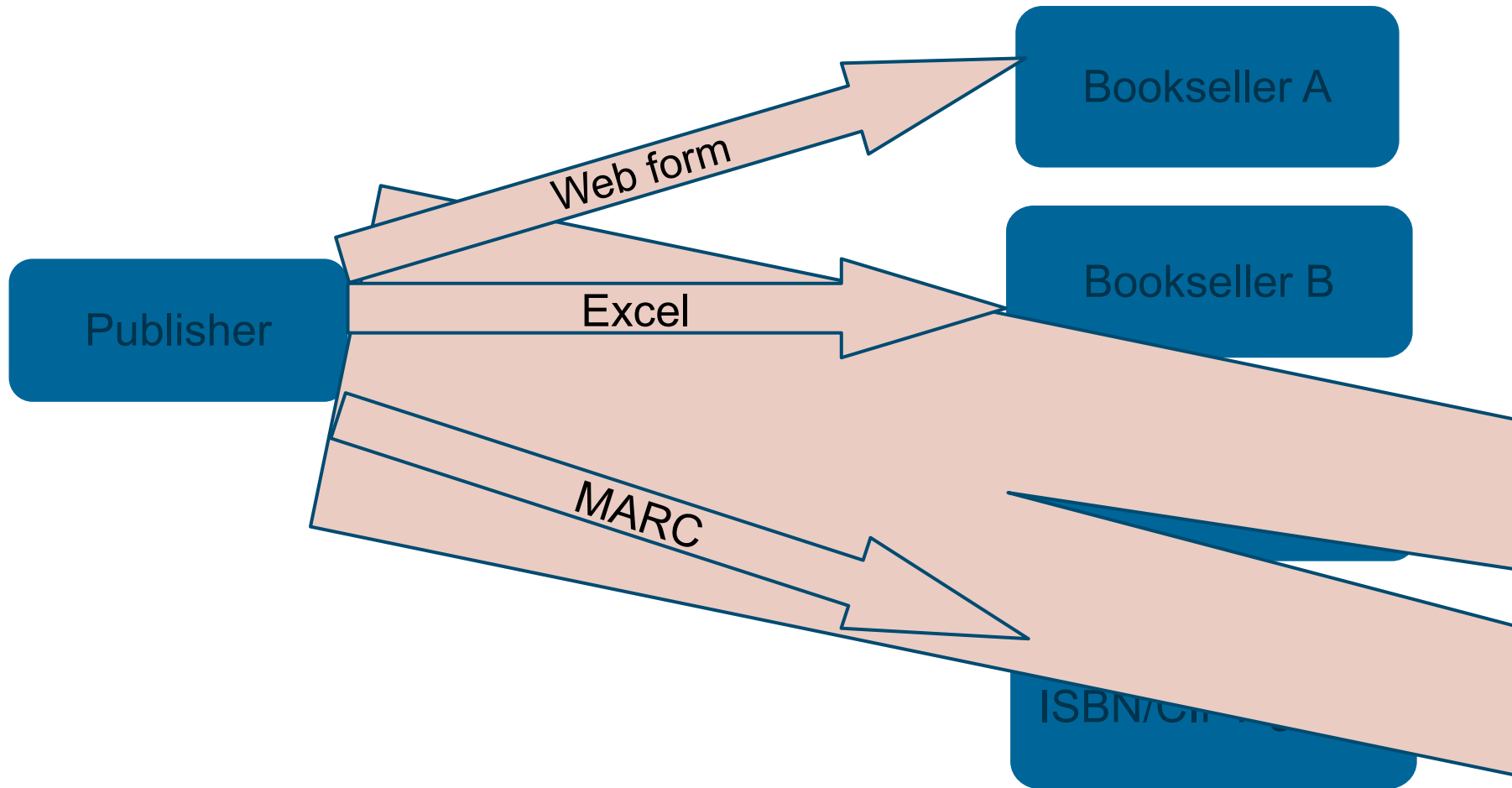
- ***“Those titles in Nielsen’s top-selling 85,000 with complete data records sold 70% more copies on average than those with incomplete metadata.”***

Nielsen (UK Books in Print)

A hub for book information

- Currently publishers must provide product data in many different formats to wholesalers, booksellers, ebook platforms, National Library etc.
- This is time consuming, wasteful, costly and can lead to errors and inconsistencies

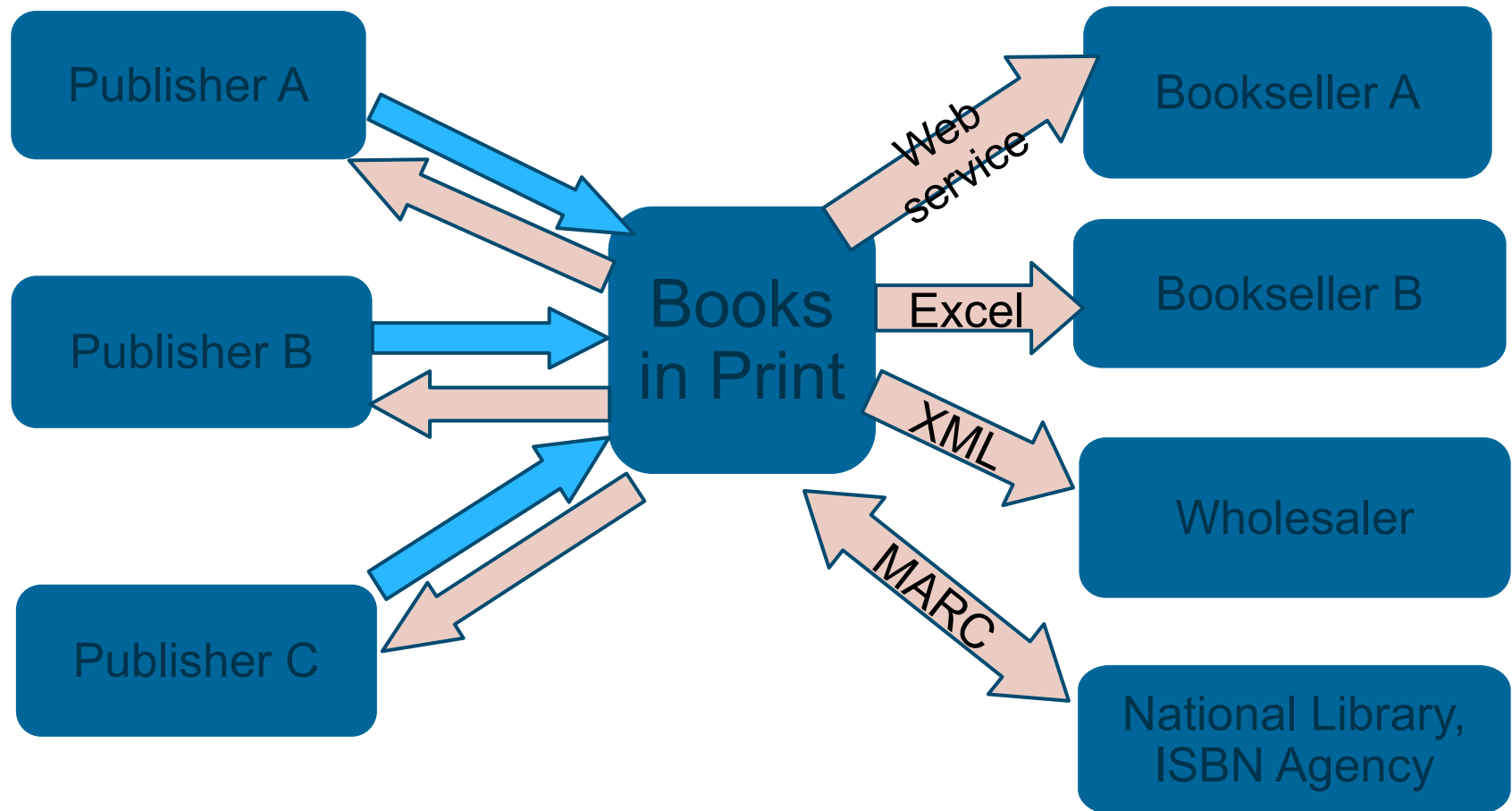
A hub for book information



A hub for book information

- A Books in Print service can provide a “hub” to which publishers can send their product information in agreed formats
- The Books in Print hub can check the data and convert it into the formats required by wholesalers, distributors, booksellers, libraries etc.
- This checked and formatted data can also be made available back to the publisher to provide a well-formed product database for producing catalogues and other promotional material

A hub for book information



Key elements of a Books in Print

- Comprehensive coverage of titles
 - ▣ Books in Print should include all books currently available
- Depth (richness) of coverage
 - ▣ Maximum information about each individual title including cover image, description, current price, availability (in or out of print) etc.
- Accuracy
 - ▣ Information in Books in Print must be accurate and trusted by the trade
 - ▣ Publishers must understand the need to check, add and update the information about their books
- Timeliness
 - ▣ BiP needs not only to be comprehensive, rich and accurate; it must also to be up-to-date – both in terms of new books and updated dynamic data (e.g. price and availability).

Which European countries have BiP services?

- Belgium
- Denmark
- Finland
- France (includes French-speaking Belgium & Switzerland)
- Germany (includes Austria and German-speaking Switzerland)
- Greece
- Italy (includes Italian-speaking Switzerland)
- Netherlands
- Norway
- Romania
- Spain
- Sweden
- UK (includes Ireland)

When was BiP established?

- 1858 UK
- 1970 
- 1971 Germany
- 1975 Italy
- 1984 France
- 1985 Norway
- 1996 Denmark
- 1998 Greece
- 2000 Belgium
- 2001 Finland
- 2003 Sweden
- 2004 Netherlands
- 2005 Spain
- 2010 Romania

Ownership of Books in Print

- Belgium Meta4Books, (non-profit industry association) - ISBN
- Denmark The Danish Booksellers Association
- Finland Kirjavälitys Oy, (wholesaler)
- France Cercle de la Librairie, (professional association) - ISBN
- Germany MVB (marketing arm of PA/BA) - ISBN
- Greece The National Book Centre of Greece (Culture Ministry)
- Italy Informazioni Editoriali (bibliographic company)
- Netherlands Centraal Boekhuis (wholesaler) - ISBN
- Norway Publishers, booksellers and wholesalers consortium
- Romania National Book Centre (public/private funds)
- Spain Federation of Spanish Publishers Guilds - ISBN
- Sweden The three major publishing houses in Sweden + the leading wholesaler.
- UK Nielsen Book Services (commercial company) - ISBN

Ownership of Books in Print

- **Belgium** Meta4Books, (non-profit industry association) - ISBN
- **Denmark** The Danish Booksellers Association
- **Finland** Kirjavälitys Oy, (wholesaler)
- **France** Cercle de la Librairie, (professional association) - ISBN
- **Germany** MVB (marketing arm of PA/BA) - ISBN
- **Greece** The National Book Centre of Greece (Culture Ministry)
- **Italy** Informazioni Editoriali (bibliographic company)
- **Netherlands** Centraal Boekhuis (wholesaler) - ISBN
- **Norway** Publishers, booksellers and wholesalers consortium
- **Romania** National Book Centre (public/private funds)
- **Spain** Federation of Spanish Publishers Guilds - ISBN
- **Sweden** The three major publishing houses in Sweden + the leading wholesaler.
- **UK** Nielsen Book Services (commercial company) - ISBN

Collaboration is a good idea

- In all countries where the BiP and ISBN Agency are not co-located, there is close collaboration between the two functions
- National Libraries managing legal deposit schemes are well-equipped to collect and aggregate good quality comprehensive data but not to maintain dynamic data such as price and availability or other market data
- Collaboration between ISBN agencies and trade bodies is the best solution
- Lithuanian Books in Print should be a good example of this

A chicken and egg problem

- Booksellers will only use and benefit from BiP if the information is comprehensive, accurate, up-to-date, rich and well presented
- Publishers will make the effort to provide and maintain their book information only if they know that booksellers are using BiP to find and buy books
- The ARROW Plus BiP model tries to solve that problem by uploading National Bibliography and ISBN/CiP data and inviting publishers to check, enhance and update the information about their books (cover image, description, author biography, price, availability etc.)

Lithuanian Books in Print

- To create “critical mass” initial data will be ingested from the National Library
- Publishers will be asked to update and enhance their own title records online using secure password access (i.e. the publisher can only change their own records)
- In the future, publishers will be expected to enter the initial information in advance of publication and keep it updated
- This may be using online forms or by batch delivery of data for multiple titles in agreed formats (e.g. Excel templates, ONIX)
- This may be connected to ISBN/CiP registration
- It will certainly include ebook information and can be developed to provide added value services such as linking different media versions of the same book, electronic ordering from booksellers to publishers, sales data monitoring etc. etc.

Value of Books in Print

For booksellers:

- Finding sources of books for customer orders
- Discovering information about new books
- Selecting books for stock
- Licensing to build bookseller's own database, saving costs
- Possibilities for electronic ordering direct from database

For publishers:

- Data sent to a single hub for transmission to entire book trade
- Equal visibility for large and small publishers, new and backlist titles
- Possibility of linking different versions of the same work (ebook, paperback, hardback, audiobook etc)

For the whole book trade

- Helps to sell more books. All parties can benefit from better data quality

The word "ARROW" is written in a large, bold, orange, sans-serif font. The letters are slightly overlapping and have a subtle gradient.

<http://www.arrow-net.eu>

ONIX for Books:

<http://www.editeur.org/93/Release-3.0-Downloads>

ISBN FAQs:

<http://isbn-international.org/faqs>

FURTHER INFORMATION

Brian Green

brian@editeur.org

ARROW Plus is a Best Practice Network
selected under the ICT Policy Support
Programme (ICT PSP)

