

Identifiers for digital content

Brian Green
Executive Director
International ISBN Agency



Why do we need identifiers?

- Identifiers allow each tradable item to be uniquely identified throughout the supply/value chain
 - Discovery and differentiation (via product lists and bibliographic databases)
 - Trading (ordering, supplying, pricing etc)
 - Managing digital products in internal systems (publisher, distributor, bookseller, library)
 - Cataloguing in libraries (e.g. differentiating formats)
 - Reporting sales, loans etc. (royalties, lending rights)
- Must be a standard identifier to provide interoperability



The ISO TC46 family of content identifiers

- Text
 - ISBN – books
 - ISSN – serials
 - ISTC – textual works
- Music
 - ISMN – printed music
 - ISRC – master recordings (not end products)
 - ISWC – musical works
- Movies
 - V-ISAN – film and video
 - ISAN – audio visual works



The ISBN system

- ISO 2108:2005 Introduction to the standard
 - *The ISBN system serves as a key element of ordering and inventory systems for publishers, booksellers, libraries and other organizations.*
 - *It is the basis for collecting data on new and forthcoming editions of monographic publications for directories used throughout the book trade.*
 - *The use of ISBN also facilitates rights management and the monitoring of sales data for the publishing industry.*



ISBN and eBooks

- Digital publications are specifically included in the ISBN standard under “types of monographic publications to which an ISBN” viz:
 - *electronic publications either on physical carriers (such as machine-readable tapes, diskettes, CD-ROMs) or on the internet;*



ISBN and different formats

- Under “Rules of assignment”, the standard says:
 - *Different product forms (e.g. hardcover, paperback, Braille, audio-book, video, online electronic publication) shall be assigned separate ISBNs*
 - *Each different format of an electronic publication (e.g. '.lit', '.pdf', '.html', '.pdb') that is published and made separately available shall be given a separate ISBN.*



...but some publishers only assign a single ISBN for all formats

- Some of their arguments:
 - "We only "publish" one generic format (e.g. .epub) and assign an ISBN to that"
 - "We are not responsible for formats, DRM protection or functionality provided by third part intermediaries"
 - "We don't care whether or not different product formats are listed in bibliographic databases. Customers will find our books through their preferred platform"
 - "We don't require sales reporting by format"
 - "We are concerned about bloat of ONIX records and the cost of ISBNs"



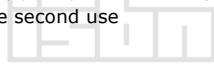
What is .epub?

- IDPF (International Digital Publishing Forum) standard for ebook text files
- Replaces OEB standard
- Two possible uses
 - .epub files rendered natively by reader devices/software as consumer products
 - .epub files supplied by publishers as distribution format and converted by third parties to other formats



What is .epub?

- IDPF (International Digital Publishing Forum) standard for ebook text files
- Replaces OEB standard
- Two possible uses
 - .epub files rendered natively by reader devices/software as consumer **product**
 - .epub files supplied by publishers as **distribution format** and converted by third parties to other formats
- In an ideal world a single file could be rendered on any device / platform (cp. unprotected MP3)
- ...but currently mostly the second use
-



ISBNs and .epub

- If the .epub file is a publicly available product then an ISBN should be assigned to it and product information (e.g. ONIX record) communicated to the trade
- If the .epub file is simply a distribution format between publisher and reseller for conversion into different formats, it should not, strictly speaking, be assigned an ISBN (although an internal ISBN may be used as long as it isn't communicated to the trade)
 - We really need a new identifier (cp ISRC)



Identifying expressions

- We have a product identifier (ISBN)
- We are about to have a work identifier (ISTC)
- The music industry has an intermediate identifier, the International Standard Recording Code (ISRC) to identify the actual recording
- These may appear in various different products
 - e.g. ISRCs are already used for audiobook recordings that may appear as MP3 files, CDs etc
- Do we need a new identifier for generic ebook formats before they become products?



What the trade tells us

- Resellers, bibliographic agencies, libraries and booksellers have told the International ISBN Agency that their systems and procedures require separate ISBNs for each ebook version / format
- Some have said that if publishers don't provide ISBNs they will need to apply their own identifiers
- Some have already started using ISBN-like identifiers, leading to duplication



The International ISBN Agency response

- ISBN agencies may assign registrant prefixes to eBook resellers to enable them to allocate ISBNs to individual eBook formats if, and only if, the publisher has not provided an eBook ISBN for each separate format
- The original publisher will appear in the bibliographic records that the resellers provide to the bibliographic agencies
- If the publisher does provide separate ISBNs for each separate format, then these should always be used in preference to the reseller's ISBNs



What about formats tied to specific devices and platforms?

- Consumer devices, e.g. Kindle, Sony eReader sell content direct and don't currently use ISBNs
- Will they be available from other sources?
- Does publisher want those versions listed
 - in bibliographic databases ?
 - in sales data / bestseller reports ?
- ...if so, YES – otherwise maybe NO
- What about library platforms?
 - Libraries tell us that they need to differentiate using separate ISBNs



What constitutes a separate version, format or product?

- This needs further discussion for digital publications
 - Should two versions of an ebook with the same format (e.g. pdf) but different DRM and user functionality be separately identified?
 - What about two MP3 versions of an audiobook using different compression ratios and therefore different user experience?
- Principle of functional granularity
 - They should be separately identified if they need to be anywhere in the supply chain



Chapters and other fragments

- ISO 2108: 2005 scope statement
 - *This International Standard is applicable to monographic publications (or their individual sections or chapters where these are made separately available) and certain types of related products that are available to the public.*



What about customised products?

- Not if they are produced one-off for a single customer
- Yes, if they are traded in multiples through normal distribution channels (e.g. college course books available through campus booksellers)

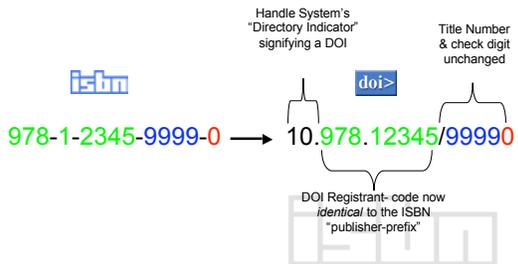


What about DOI?

- DOI is not a supply chain identifier and wouldn't be recognised in most supply chain systems
- However, DOI is a good way to "web-enable" ISBNs
- Syntax has been agreed for ISBNs to be incorporated into DOIs
- Currently being piloted in German *libreka!* full text book search system



"Actionable ISBN" Syntax



Thanks. Looking forward to some lively discussion.

> www.isbn-international.org

