

April 2010

# Standards Column -- eBook Identification: Disrupting the Tried and True ISBN System

Todd Carpenter  
NISO, [tcarpenter@niso.org](mailto:tcarpenter@niso.org)

Follow this and additional works at: <https://docs.lib.purdue.edu/atg>



Part of the [Library and Information Science Commons](#)

---

### Recommended Citation

Carpenter, Todd (2010) "Standards Column -- eBook Identification: Disrupting the Tried and True ISBN System," *Against the Grain*: Vol. 22: Iss. 2, Article 43.  
DOI: <https://doi.org/10.7771/2380-176X.5532>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact [epubs@purdue.edu](mailto:epubs@purdue.edu) for additional information.

## I Hear the Train A Comin' from page 84

community and research centers. An endowment, as well as support from scholarly societies or funding agencies, may also be part of our mature portfolio.

*Is there any role for commercial entities in the long-term sustainability model? If so, what might that look like?*

Possibly. We've had some preliminary discussions with a few of the discipline-relevant scholarly societies about how they might be involved. We hope to continue the conversation with them over the course of the next year or two. Participation by the societies or commercial publishing entities must, of course, be informed by terms and conditions that are acceptable to both the arXiv stakeholders and the outside agencies.

*For other disciplines in which the culture of mass distribution of open-access content by authors is not as strong, are arXiv's recent financial developments encouraging or a cautionary tale?*

Both. The mass distribution of content is costly. There are a variety of ways to recover those costs, and it's both prudent and shrewd to devise a model that will ensure that the resource will realize revenue from a variety of differ-



**ScholarNext**

Consulting Services at the Intersection of  
Technology, Content, & Academia

- ✓ Web 2.0 Functionality
- ✓ Strategic Planning
- ✓ Business Development
- ✓ Policy Analysis
- ✓ Publishers
- ✓ Information Providers
- ✓ Technology Companies
- ✓ Libraries

(510) 295-7504      Greg Tananbaum      greg@scholarnext.com

[www.scholarnext.com](http://www.scholarnext.com)

ent sources. It's critical that the stewards of scholarly resources understand scale and scope economies, think strategically instead of tactically, and be willing to shift directions when their winds of fortune change. 🌱

## Standards Column from page 86

provides the linkage of different manifestations of the same content. For example, the first ISTC was assigned to **Charles Darwin's** *On the Origin of Species*. This text has been reprinted by numerous publishers in various forms, both print and electronic, since it was first published in 1859. A search of the **WorldCat** catalog found 281 different editions with **Charles Darwin** as the listed author, with the latest edition appearing in 2009. While few books have this many editions, in today's digital environment, if each version is unambiguously identified (presumably with unique ISBNs), it's very likely that every single title published today could have at least a dozen different versions and an equivalent number of ISBNs. The ISTC will be the tool that links these different versions together, both for discovery and collection management. On the publisher side, the ISTC also assists with rights management. If the ISBN is used consistently for disambiguation, the ISTC can provide the umbrella framework for all of the multiple product references and could help to

streamline metadata creation and improve its quality. The ISTC standard was published in early 2009 and is currently being used by a few publishers but needs to gain wider adoption to be broadly useful. Additionally, best practices need to be adopted regarding the sharing and consistent application of metadata between the ISBN and the ISTC.

The **International ISBN Agency** recently released a white paper (<http://isbn-international.org/news/view/29>) describing the problems of ISBN assignment to eBooks, emphasizing the need for consistent practices for assignment of ISBNs, and recommending some solutions. From the perspective of a centralized and globally adopted system, ensuring the consistency of the system's application is paramount. This applies not only to different formats, but also to sub-items, such as chapters or other fragments of works, when these items become part of the supply chain. The paper also suggests that there might be a need to identify the generic publication file produced by the publisher that is then used for further re-formatting and distribution. The music industry has such a system in place for the original track, which can then be placed on a variety of albums or even sold separately.

This original track — defined as “an abstract entity representing a bundle of one or more Digital Resources compiled by an Issuer for the purpose of electronic distribution to individual consumers, directly or through intermediaries” — is identified using a Global Release Identifier, or GRID. The GRID standard ([http://www.ifpi.org/content/section\\_resources/grid-standard.html](http://www.ifpi.org/content/section_resources/grid-standard.html)) clearly states that “the Release is not itself a Product.” The **International ISBN Agency** is now discussing with various stakeholders whether a similar system of using generic eBook release identifiers is both feasible and of value.

Although the digital distribution of content has been gaining momentum for more than a decade and a half, the eBook is still in its relative infancy. But eBook production is growing faster than its predecessors, such as the e-journal. As a result, the problems in the supply chain are getting out of control. We need to act quickly to develop a consensus solution. If you're a stakeholder in this arena, the **International ISBN Agency** (<http://isbn-international.org/>) and the **International ISTC Agency** (<http://www.istc-international.org/>) are good places to start. 🌱

## Rumors from page 63

back to the Palmer House and ALAP (**ALA Players**) was born. Every ALA winter and summer conference there is a dinner dance put together by **Fred Goodman** at LLSI usually on Monday night to commemorate the event. And, it's a hoot to go to the ALAP homepage about the history of this whole event. It's complete with a photo gallery, and memories of the 20th anniversary gala in New Orleans in 1998 [homepage.mac.com/gailmcgovern/PhotoAlbum23.html](http://homepage.mac.com/gailmcgovern/PhotoAlbum23.html).

Speaking of Iceland — The **Steering Com-**

**mittee of Iceland Consortia** for electronic subscriptions has selected **ProQuest Central** to provide electronic resources nationally for the next three years. On behalf of the Ministry of Education, Science and Culture, the **Steering Committee** is responsible for the evaluation and procurement of resources which can then be accessed by all Icelandic citizens from any computer within the country. The project is funded by the Ministry along with more than 200 libraries, educational institutions and private firms. [www.proquest.com/](http://www.proquest.com/)

Yours truly was just interviewed by **Jack McHugh** and the interview is posted on his Website. [www.johnbmchugh.com/](http://www.johnbmchugh.com/). Click on

expert interviews and mine is the fourth one down. Would love comments.

Was interested to see a report on eBook devices from the floor of the **Consumer Electronics Show (CES)** in Las Vegas. Try this link for a look at some of the devices. Ho-ha! [aptaracorp.com/context/Aptara\\_Context\\_Feb\\_2010.pdf#page=3?elq=369ac855817c4b66b93d7b10007dfcc](http://aptaracorp.com/context/Aptara_Context_Feb_2010.pdf#page=3?elq=369ac855817c4b66b93d7b10007dfcc)

**Last but not least! Baby Georgia Helen Walser is here! Born March 7, 2010 at 6:49pm. 7 lbs 10 oz; 19.5 inches long. This is Katina's and Bruce's second grandbaby! And Georgia Helen skyped her father in Iraq in the first 15 mins of her life! She is obviously a genius! ;-)** 🌱

# Standards Column — eBook Identification: Disrupting the Tried and True ISBN System

by **Todd Carpenter** (Managing Director, NISO, One North Charles Street, Suite 1905, Baltimore, MD 21201; Phone: 301-654-2512; Fax: 410-685-5278) <tcarpenter@niso.org> [www.niso.org](http://www.niso.org)

The International Standard Book Number, most commonly known by its acronym ISBN, has been among the most successful and widely-adopted publishing standards since it was launched in the UK during the 1960s. There are more than 140 national ISBN bodies, and the ISBN has been used to identify more than six million items of varying sorts around the world.

However, like all things impacted by the expansion of digital information, the ISBN system is increasingly being strained by the proliferation of electronic books. The problem would seem on its face to be a relatively simple one: how and when do you uniquely identify an eBook? But because of the transformative nature of digital information, the production cycle, supply chain, and final outputs can be quite different for eBooks than in traditional print publishing.

In the print world, there is a fairly straightforward distinction between versions of the same content. We knew, for example the difference between a hard and soft cover, a large-print, and a Braille edition. The publisher has planned in advance for these editions and is in control of the production, including the assignment of ISBNs. However, with eBooks, the publisher often creates a generic file, perhaps in EPUB format, which is then provided to various distributors. The distributors may make further changes to the file before they offer a final product for sale. For example, Amazon can convert the EPUB file to the Kindle format and add digital rights management (DRM) code. The question is whether this transformation constitutes a new format, at what point

does that transformation require a new identifier, and who is responsible for the identifier assignment?

The ISBN standard (ISO 2108:2005) clearly states:

*“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.”*

However, there remain some important questions, even debate, about the implementation of this in the real world.

Let’s start from the most basic principle that we identify things in the supply chain to facilitate commerce and the management of the product — be that physical or digital — through every step of production, distribution, and sale. Such identification is also important to the end users who often use an ISBN to be sure they have purchased the correct edition. If you are visually impaired, for example, the difference between a standard-sized print and a large-print or Braille edition could mean the difference between being able to read the book or not.

Similarly, if you purchase a **Kindle** edition of a book from **Amazon** and hope to read that file on your **Nook** device, you will likely be disappointed because of the differences in the two file formats or the DRM encoded in the file. If you wanted high-quality, detailed page-layout in your file, a pre-formatted PDF might be more suited to your expected reading experience than an EPUB file. Finally, even in publicly available scans of files that lack DRM, the quality of the resolution in a scan might be different from one supplier to

another, and one can envision situations where these differences are important, if not critical, to the end user’s experience. If all of these versions have the same ISBN, there could be considerable confusion in exactly which product is which by distributors, by online or bricks-and-mortar book stores, and ultimately by the end user. As reading devices proliferate in the market and eBooks become widely available from multiple sellers in multiple formats, lack of easy identification of the variations will result in both user dissatisfaction and increased costs to suppliers — from added customer support needs, for example.

This is not unlike the problems that the journal supply chain underwent in the mid- to-late 1990s, when the rules for the assignment of ISSN numbers were unevenly applied to electronic versions of print journals. Some publishers would assign the same ISSN to both print and electronic versions, while others would assign a different ISSN to the electronic edition. It wasn’t until the revision of the ISSN standard in 2007 (ISO 3297:2007) that the assignment of unique ISSNs to e-versions was made mandatory, and a system, the Linking ISSN or ISSN-L, was proposed to link all the versions of a title — print and electronic. Digital book delivery is currently suffering from the same loose application of ISBN rules.

One of the difficulties with eBooks and ISBN is that in the print supply chain the publisher creates the different versions and assigns the ISBN. With eBooks, the publisher frequently passes on a generic file and the various formats are produced by someone else further down in the supply chain. These entities, such as distributors and aggregators, have never been in the “business” of assigning and managing ISBNs. Last summer the **International ISBN Center** adopted the policy that if a publisher has not or does not want to assign and manage the separate ISBN numbers for each variant of an eBook, then a distributor can do so. Last fall, some distributors began assigning ISBN numbers for eBook products on behalf of publishers. This, however, introduces the new problem about which entity owns and manages the canonical source of metadata about a book. Many would argue that publishers should be the ultimate source of verifiable data about a book. There are already known issues about the quality of publisher-supplied data, which were highlighted in the report that **NISO** published last summer. (See the white paper, *Streamlining Book Metadata Workflow* available from: [http://www.niso.org/publications/white\\_papers/StreamlineBook-MetadataWorkflowWhitePaper.pdf](http://www.niso.org/publications/white_papers/StreamlineBook-MetadataWorkflowWhitePaper.pdf).) Increasing the number of entities with responsibility for creating and delivering book metadata could make things even worse.

Another aspect to this situation is the use of the International Standard Text Code (ISTC) standard (ISO 21047:2009) and its relation to the ISBN. The ISTC identifies the unique content of a work rather than its various formats and



## ADVERTISERS' INDEX

59	AAAS	75	THE CHARLESTON ADVISOR	88	MIDWEST LIBRARY SERVICE
39	ABC-CLIO	10	THE CHARLESTON REPORT	65	MODERN LANGUAGE ASSOCIATION
17	ACM	15	GOLD SPRING HARBOR LAB PRESS	21	OECD
13	ACS PUBLICATIONS	7	COUTTS INFORMATION SERVICES	37	OSA - THE OPTICAL SOCIETY
55	AMERICAN ECONOMIC ASSOCIATION	61	EASTERN BOOK COMPANY	9	PALGRAVE MACMILLAN
29	AMERICAN INSTITUTE OF PHYSICS	49	EBSCO INFORMATION	25	PROQUEST
43	ANNUAL REVIEWS	77	EMERY-PRATT	69	RITTENHOUSE
41	APPI	71	THE ENDOCRINE SOCIETY	85	SCHOLARNEXT CONSULTING
45	ASBMB	23	IEEE	11	SPIE
33	ASME INTERNATIONAL	2	IGI GLOBAL	73	ACQUISITIONS INSTITUTE AT TIMBERLINE LODGE
5	ATG	27	IGI GLOBAL	35	WILEY
14	BASCH SUBSCRIPTIONS, INC.	79	INFORMS	3	YBP LIBRARY SERVICES
31	BEGELL HOUSE PUBLISHERS	81	McFARLAND		
87	THE BOOK HOUSE INC.	19	M. E. SHARPE		

For Advertising Information Contact: **Toni Nix**, Ads Manager, <justwrite@lowcountry.com>, Phone: 843-835-8604, Fax: 843-835-5892.

*continued on page 85*