



Product Metadata Best Practices for Data Recipients

Version 1.0
September 2008

Book Industry Study Group, Inc.

BISAC: Book Industry Standards and Communications

BISAC Metadata Committee

Product Metadata Best Practices for Data Recipients was written and compiled for the Book Industry Study Group, Inc. (BISG) by Richard Stark and the BISAC Metadata Committee. All enquiries about this document should be addressed to BISG using the contact details given below.

Copyright © 2008 Book Industry Study Group, Inc. All rights reserved.

BISG develops and maintains standards that enable effective communication along all parts of the industry supply chain. We also create policies and best working practices for the industry. BISG has led the way in setting industry standards and conducting vital industry research on behalf of publishers, booksellers, libraries, and vendors.

Book Industry Study Group, Inc (BISG)

370 Lexington Avenue, Suite 900

New York, New York 10017

USA

Telephone +1 (646) 336 7141

Fax +1 (646) 336 6214

info@bisg.org

www.bisg.org

Preface

This document is the result of several months of work on the part of the BISAC Metadata Committee. The committee members who worked on drafting and reviewing this document represented publishers, wholesalers, booksellers, bibliographic data providers, trade associations, and libraries. The final guidelines for each practice covered in this document were approved by a consensus of Committee members and then submitted to BISG membership and to the BISG Board of Directors for review and approval.

These best practices are voluntary guidelines that, if followed, will improve the accuracy of a vendor's data throughout the supply chain and speed the processing of that data among a vendor's trading partners. Recipients of data are encouraged to use these practices to work collaboratively with the suppliers of their data to improve the quality of the product data we all use.

About the Authors

The BISAC Metadata Committee has been a leader in the development of both national and international product metadata standards and best practices for many years. The Committee has overseen the development and maintenance of the EDI X-12 832 Message standard, and since the Association of American Publishers (AAP) empowered BISG to oversee the ONIX standard in our market, the Metadata Committee has had responsibility for the continued development and maintenance of ONIX for Books in this market.

—Richard Stark, on behalf of the BISAC Metadata Committee

Glossary of Terms

consumer-facing catalog systems: systems used by the general public to search for and purchase books and related products. Retailers' web sites (e.g., Amazon.com, BookSense.com, bn.com, etc.) are one example of such systems. Online public-access catalogs (OPACs) used by library patrons are another example. Systems that are intended for use by book-industry professionals are NOT included in this definition.

data aggregator: an organization that collects data from a variety of publishers, vendors, and other data suppliers in order to build a comprehensive catalog of product information. Data aggregators typically license their catalogs to libraries, retailers, and other data recipients.

data contact: the staff person at a data supplier, data recipient, or other organization in the data supply chain who is designated by their organization as its representative on product data matters.

data owner: the holder of U.S. sales and/or distribution rights for the ISBN (or other EAN) indicated in the product record.

If the aforementioned company(s) does not wish to have responsibility for a product's data, it may designate another party as the product's "data owner." Examples of commonly designated data owners are:

- a. The U.S. division of a multinational publishing company
In cases where a book or other product is published in the U.K., for example, and imported by the U.S. division of the U.K. publisher, the U.S. division may be designated as the data owner in the U.S. market.
- b. The vendor of record who sells the publisher's books into the U.S. retail market
This class of data owners includes both distribution companies and sales agents. Distribution companies may be other publishers or companies whose sole business is distribution. A wholesaler of books may serve as the vendor of record for a publisher and therefore the wholesaler could be the data owner for that publisher's books.

- c. The service provider who is contracted to provide product data services to the publisher
Publishers and distributors may designate a data service company to serve as the data “owner” for their books. Examples of such companies are Firebrand Technologies, NetRead, Ingram Book Company, Baker & Taylor, and R. R. Bowker.

data recipient: an organization that receives and loads product data into its catalog and/or other database systems. Libraries, retailers, wholesalers, and distributors are all examples of data recipients.

data supplier: an organization that sends product data to data recipients. Publishers, distributors, wholesalers, publishing systems vendors, and data aggregators are all examples of data suppliers.

distributor: a company that provides warehousing, shipping, and/or other supply-chain services to a publisher or other product manufacturer. For the purposes of Product Metadata Best Practices, distributors often serve as data owners or data suppliers for their clients’ products.

process a data file: The term *process* shall, in the context of Best Practices for Product Data Recipients, mean that a data file has been loaded into all applicable *internal* database systems of the data recipient.

It does NOT necessarily mean that not-yet-released products are made public on consumer-facing catalog systems. Practices on publishing data on public catalogs vary from one data recipient to another, however, updates to product records that are already viewable by data recipients’ customers should be published on consumer-facing catalog systems as soon as is practical and no later than nine calendar days of the receipt of the file.

Best Practices for Recipients of Product Data

1. Data recipients should have a clearly defined system for contacting their organization regarding product data

Recipients may provide a list of responsible parties and their contact information or some other means of contacting them that facilitates effective communication on product data. Whenever it is feasible, a single primary contact person (who can communicate with other concerned parties at their company) should be named, with back-up contacts noted.

2. Data recipients should acknowledge that files have been received

Recipients of product data should acknowledge the receipt of incoming data files within two business days of a file's receipt.

This acknowledgment may be active (such as an e-mail message) or passive (such as a web service that allows a supplier to query the recipient's FTP logs).

The day when this acknowledgment is sent is when the seven-day period described in Section no. 3 begins.

3. Data recipients should establish a service level for file processing

Recipients of product data should process incoming weekly delta files, in the order they were sent, from data suppliers within seven calendar days of the file's receipt.

Daily delta files should be processed within 48 hours of their receipt, in the order they were sent.

Recipients should process a full catalog file as often as they request such a file or at least once each year. Recipients of product data should process incoming full catalog files from data suppliers within fourteen calendar days of the file's receipt.

A confirmation that the product records in a file have been processed (i.e., loaded into the recipient's database systems) or an indication of what prevented timely processing should be provided to data suppliers within three business days after a file has been processed.

4. Data recipients should provide data suppliers with a clear statement of their practices for processing and displaying data

This statement should include an indication of when the data will be processed, as well as when it will initially display on consumer-facing public catalog systems (e.g., 90 days prior to a book's publication date). This statement should be publicly available to all data suppliers from whom a recipient receives product data. This statement should include some description of pre-release sales practices (e.g., the recipient will begin accepting customer orders on new book products 60 days prior to a book's publication date).

This statement should detail how a recipient treats products whose release dates are postponed. Recipients should indicate under what conditions postponed products would stop being displayed on consumer-facing public catalog systems and under what conditions such products would begin displaying again on said catalog systems. Recipients should also detail how they treat products that become temporarily unavailable. An indication should be made of how long a product is unavailable before the recipient cancels customer orders.

5. Data recipients should use delta files on a regular basis instead of full weekly files

Data recipients should inform their data suppliers how often they wish to receive a full file (i.e., a file containing ALL of a data supplier's records for products that have been active during the past year). Recipients should process a full file from each supplier at least once each year, although some data recipients may wish to receive full files quarterly or semi-annually.

Data recipients should inform their trading partners how often they wish to receive update files. For some data recipients this might be a weekly update file from all major trading partners, while for other recipients there may be no need for anything more frequent than a semi-annual update file.

6. Recipients should, upon request of the data owner, identify the source of data from suppliers other than the data owner

Data recipients (including data aggregators) are free to load data from any source they wish to use. Data owners, however, should be informed by data recipients, upon request, of the source of a given piece of relevant product data to the extent that the recipients are legally able to do so under their agreements with other data suppliers.

Data recipients should store information on the source of their current product data for at least one year after the data is received.

Data recipients should reply to such requests within seven calendar days.

7. Data recipients (at the request of a data supplier) should allow data owners' data feeds to overwrite data that was manually entered by the recipient

As a general practice, data recipients are encouraged to establish policies that allow updates from a data owner to overwrite data the recipients have manually updated in their systems. In cases in which recipients have “locked” data they have manually entered in order to prevent said data from being updated by subsequent data feeds, it is incumbent on the data recipient to work with the data supplier to allow newer, mutually agreed-upon data to update older data, even if that older data was entered to correct information previously sent by the data supplier.

Data suppliers must, therefore, commit to a practice of including any manual updates they have requested since their last update file in their next batch update file. Data recipients should provide an explanation for any case in which they do not allow a data supplier's update files overwrite the recipient's existing data. Data recipients should provide such explanations within seven calendar days.

8. Data recipients should provide facilities for “emergency” updates

Data recipients should provide data suppliers with a clear statement of the procedures that data suppliers should follow to make “emergency” updates to product data in the recipient's systems. Any such updates should be made within one business day of a data supplier's request and an acknowledgment that the update was (or was not) made should be provided to the data supplier within one business day.

Data recipients should provide data suppliers with a clear statement of their practices for emergency updates. A given data recipient's practice may simply be to supply a list of key contact persons and the contact information for each of those persons.

Such updates will often require the manual update of data and therefore it is incumbent on data recipients to allow subsequent batch updates to the data in question (see above).

For the purposes of these Best Practices, an "emergency" update should be considered any correction or addition to the data on a product that could, if it is not acted upon, cause legal action to be brought against the data supplier or significant harm to their reputation.

Simple updates to data on books that are not attracting a great deal of media attention are NOT emergencies. Updates to strict on sale dates on high-profile titles are emergencies. A significant change to a product's price is an emergency; an update of one dollar in a product's price as a part of a normal reprint process is NOT an emergency.

9. Postponed publications should be made active again upon the receipt of updated status data from the data supplier

Data recipients should allow titles that have been postponed for normal reasons (i.e., titles that have been delayed due to legal considerations are NOT included in the definition of postponed for "normal reasons") to be updated through a data supplier's regular data feeds. It should not require manual intervention on a data supplier's part to have such books reinstated on a recipient's catalog systems. Data recipients always have the final decision, however, on what products they wish to list in their catalogs.

Data suppliers must, therefore, supply revised availability dates or status updates as soon as they are known.

Titles that have been postponed due to legal considerations will require that data suppliers notify recipients using the emergency update provisions detailed in Section no. 8.

10. Publications that are cancelled should not be displayed on consumer-facing catalog systems

Recipients should remove publications that are cancelled or abandoned from sale and display on consumer-facing catalog systems. Catalog systems designed for use by book industry professionals are specifically excluded from this practice; however, such systems should indicate the correct status of any cancelled publication. Any outstanding consumer orders for such publications should be cancelled.

11. Publications that are withdrawn from sale for legal reasons should not be displayed on consumer-facing catalog systems

Recipients should remove publications that are described by their publishers as “withdrawn from sale for legal reasons” from sale and display on consumer-facing catalog systems. Catalog systems designed for use by book industry professionals are specifically excluded from this practice; however, such systems should indicate the correct status of any product withdrawn from sale. Any outstanding consumer orders for such publications should be cancelled.

Data recipients who wish to continue selling used or second-hand copies of books that have been withdrawn from sale for legal reasons should consult with the publishers of such books regarding the legal ramifications of selling used or second-hand copies of such books.

12. Data recipients should provide mechanisms for suppliers to view or compare data

Recipients should provide some way for data suppliers to query a recipient’s database to automate those comparison processes. This service may be provided by web services, secure web sites, ONIX, or Excel reports pushed to suppliers, or any other secure means for data suppliers to get this information.

13. Data recipients should document what systems are updated from the data sent by data suppliers

Recipients should document how suppliers’ files are used and what internal systems are updated by the data received from the suppliers. For example, if a supplier’s files are used to update public catalog systems (e.g., retail web sites), purchase-order systems, billing systems, returns systems, and warehouse operations systems, a supplier should be given a clear indication of how its data is being used by each of the recipient’s systems.

Re-distributors of product data (such as data aggregators) are NOT required to document how their customers use product data; however, they should (to the extent that they are legally able to do so and insofar as it does not harm their business), upon request, provide a list of the companies to whom they are re-distributing data.

Recipients should indicate if a supplier’s data is being utilized in any of the following types of systems, and they should indicate what data from a supplier is being used in each system:

- **consumer-facing catalog systems and retail e-commerce web sites** (e.g., a recipient might use ONIX product records for such web catalog systems)
- **purchasing systems** (e.g., a recipient might use ONIX product records for purchasing systems but also rely upon EDI purchase order acknowledgments and advanced shipping notices to update certain data points related to product availability)
- **store operations systems** (e.g., a recipient might use ONIX product records to populate a store's catalog system used by booksellers)
- **point-of-sale reporting systems** (e.g., a recipient might use ONIX product records to update sales-reporting tools used by their trading partners)
- **warehouse operations systems** (e.g., a recipient might use ONIX product records to populate systems used to receive and traffic products through the recipients' warehouses)
- **returns systems** (e.g., a recipient might use ONIX product records to populate systems used to generate and process product returns to their suppliers)
- **in-store kiosks** (e.g., a recipient might use ONIX product records to populate consumer-facing catalog systems in retail stores)
- **online public-access catalogs (OPACs)** (e.g., a recipient might use ONIX product records to populate consumer-facing catalog systems in libraries)
- **union catalogs** (e.g., a recipient might use ONIX product records to populate shared catalog systems used by library consortia)

14. Data recipients should have a clear policy for releasing embargoed title data

Recipients should have the ability to prevent embargoed data from displaying on consumer-facing catalog systems and other public web sites before the date and time set by the data supplier (as indicated in the ONIX announcement date field, for example). A recipient's policies and practices regarding embargoed data should be publicly available to all data suppliers from whom a recipient receives product data.