The Development, Use, and Modification of Book Product Metadata

A Report from the United States and Canada
BISG Mission

To create a more informed, empowered, efficient book industry.
Friction
The Impact of Bad Metadata
Research Project

- Contracted with Magellan Media
- 30 companies interviewed in U.S. and Canada
- Survey of BISG and BNC members
- Test data tracked through supply chain
Project Background

- BISG *Best Practice* guides
  - For *Data Senders* (2005)
  - For *Data Recipients* (2010)

- Other organizations have also studied the topic
  - “Streamlining book metadata workflow”, NISO/OCLC 2009
  - “The link between metadata and sales”, Nielsen 2012
Metadata Flow: The Ideal

- Large Publishers
- Medium-size Publishers
- Smaller Publishers
- Metadata management services (e.g., Firebrand) – not always used
- Distributors and aggregators (e.g. wholesalers, Bowker, etc.)
- Retailers (e.g., B&N, Amazon, indie stores, etc.)
Metadata Flow: The Reality

Metadata management services (e.g., Firebrand) – not always used

Large Publishers

Medium-size Publishers

Smaller Publishers

Distributors and aggregators (e.g., wholesalers, Bowker, etc.)

Content converters

Retailers (e.g., B&N, Amazon, indie stores, etc.)

Other Consumer Facing Companies (e.g., social sites, etc.)
High Level Findings

- Publisher concerns with modified, added data; risk of bad data
- Recipients report continued weakness in supplied metadata
- Separate feeds for physical and digital products
- Metadata is added to improve discovery and purchase
- In the U.S., ONIX 3.0 is off to a slow start
High Level Findings

- Publisher concerns with modified, added data; risk of bad data
- Recipients report continued weakness in supplied metadata
- Separate feeds for physical and digital products
- Metadata is added to improve discovery and purchase
- In the U.S., ONIX 3.0 is off to a slow start
Overlapping formats (suppliers)

- Among publishers, transmission formats vary
- 56% send multiple versions of a single product feed; another 23% “not sure”
- 85% maintain separate digital feeds
Overlapping formats (suppliers)

- Among publishers, transmission formats vary
- 56% send multiple versions of a single product feed; another 23% “not sure”
- 85% maintain separate digital feeds
Metadata quality (suppliers)

- 43% use a metadata management system to validate
- 36% “check manually”; 21% don’t check/not sure
- A third either don’t track metadata quality/not sure
- 85% get feedback; 63% “follow up on every issue”
- A third of publishers see a need for other “critical” fields
Metadata quality (suppliers)

- 43% use a metadata management system to validate
- 36% “check manually”; 21% don’t check/not sure
- A third either don’t track metadata quality/not sure
- 85% get feedback; 63% “follow up on every issue”
- A third of publishers see a need for other “critical” fields
Metadata quality (suppliers)

- 43% use a metadata management system to validate
- 36% “check manually”; 21% don’t check/not sure
- A third either don’t track metadata quality/not sure
- 85% get feedback; 63% “follow up on every issue”
- A third of publishers see a need for other “critical” fields
Metadata quality (suppliers)

- 43% use a metadata management system to validate
- 36% “check manually”; 21% don’t check/not sure
- A third either don’t track metadata quality/not sure
- 85% get feedback; 63% “follow up on every issue”
- A third of publishers see a need for other “critical” fields
Downstream metadata handling

- A quarter of publishers seek “substantial improvement” in recipient practices for processing and updating data
- 60% of publishers want “substantial improvement” in reporting on data taken from other sources
- Only 5% of publishers said their data is “never” altered
- 47% said they “don’t know” where it is altered
A quarter of publishers seek “substantial improvement” in recipient practices for processing and updating data.

60% of publishers want “substantial improvement” in reporting on data taken from other sources.

Only 5% of publishers said their data is “never” altered.

47% said they “don’t know” where it is altered.
Downstream metadata handling

- A quarter of publishers seek “substantial improvement” in recipient practices for processing and updating data.
- 60% of publishers want “substantial improvement” in reporting on data taken from other sources.
- Only 5% of publishers said their data is “never” altered.
- 47% said they “don’t know” where it is altered.
Process opportunities

- Confirm a shared vocabulary (e.g., page count, rights)
- Create feedback loops (improve the data supplied)
- Clearly articulate when updates occur, what gets updated and what doesn’t get updated
- Improve discussion about what is changed, added, deleted
Process opportunities

- Confirm a shared vocabulary (e.g., page count, rights)
- Create feedback loops (improve the data supplied)
- Clearly articulate when updates occur, what gets updated and what doesn’t get updated
- Improve discussion about what is changed, added, deleted
“Future-proofing” metadata

- Automate data workflows and compress cycle times
- Prepare for more frequent updates (especially price)
- Harmonize supply-chain uses of ONIX; move to 3.0
- Separate structure and content from display
“Future-proofing” metadata

- Automate data workflows and compress cycle times
- Prepare for more frequent updates (especially price)
- Harmonize supply-chain uses of ONIX; move to 3.0
- Separate structure and content from display
Product Data Certification

- Panel of industry experts certifies sender data
- Retooling program to be scalable
- Pilot planned for Q1 2013
Executive Summary and Full Report Available at BISG.org

Coordinating Partner: BookNet Canada

Silver Sponsors:
Disney Publishing
Hachette Book Group
Random House
John Wiley & Sons

Supporting Sponsors:
Readerlink Distribution Services
Sensible Solutions
Sourcebooks

Friends:
Safari Books Online
Sally Dedecker Enterprises
Executive Summary and Full Report Available at BISG.org

Coordinating Partner: BookNet Canada

Silver Sponsors:
Disney Publishing
Hachette Book Group
Random House
John Wiley & Sons

Supporting Sponsors:
Readerlink Distribution Services
Sensible Solutions
Sourcebooks

Friends:
Safari Books Online
Sally Dedecker Enterprises

50% off FULL REPORT for Frankfurt Supply Chain Attendees!
Executive Summary and Full Report Available at BISG.org

Coordinating Partner: 
BookNet Canada

Silver Sponsors: 
Disney Publishing
Hachette Book Group
Random House
John Wiley & Sons

Supporting Sponsors: 
Readerlink Distribution
Sensible Solutions
Sourcebooks

Friends: 
Safari Books Online
Sally Dedecker Enterprises

Email Nadine@bisg.org to take advantage.

“Frankfurt Metadata”
Contact

Len Vlahos
Executive Director
Book Industry Study Group
len@bisg.org
@lenhouse