The metadata value chain

Jonathan Nowell – President, Nielsen Book
Metadata value ... the systems

1858 → 2012
What we do ...

- **BookScan**
  - Measuring & analysing book sales

- **Registration Agency**
  - Identifying publishers & books

- **BookNet**
  - E-trading

- **BookData**
  - Search and discover books

Book Supply Chain
Agenda

• The value of metadata
• Lines of communication
• Aligning data needs
• The benefits of a central metadata aggregator
• Summary
The value of metadata

• Discovery

• Persuasion

• Transaction

• Fulfilment

• Measurement…

… all of which contribute to effective trading and maximised sales
Lines of communication
The matrix of data relationships rapidly becomes complex

Data suppliers

Data receivers
A central data aggregator dramatically reduces the number of data relationships.
More data players can lead to exponential growth in the number of data relationships

<table>
<thead>
<tr>
<th>Number of data senders</th>
<th>Number of data receivers</th>
<th>Data relationships in an everyone-to-everyone scenario</th>
<th>Data relationships via a central aggregator</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>4</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>32</td>
<td>12</td>
</tr>
<tr>
<td>8</td>
<td>12</td>
<td>96</td>
<td>20</td>
</tr>
<tr>
<td>20</td>
<td>30</td>
<td>600</td>
<td>50</td>
</tr>
<tr>
<td>50</td>
<td>100</td>
<td>5000</td>
<td>150</td>
</tr>
</tbody>
</table>
The difference between multiplying or adding complexity

• Introducing more data suppliers to an unmediated matrix can **MULTIPLY** the data relationships

• Introducing more data suppliers to a system mediated by a central data aggregator merely **ADDS** data relationships
Data senders to Nielsen Book

• 320 sending data via electronic feeds (ONIX, EDI and bespoke files)

• 12,800 registered users of our web interface (Pubweb)

• Over 50,000 records added via manual entry in 2012 so far
Data relationships via Nielsen Book

• Approximate number of data senders (publishers, marketing agents, distributors etc) =

   6,720

• Approximate number of data receivers (retailers, wholesalers, libraries) =

   3,500

• Without the mediation of a central aggregator this environment has the potential to generate a matrix of:

   23.5 million data exchange relationships
Data traffic

- Our web interface (Pubweb) users send over **5,000** new records, and make approximately **10,000** edits per month – with new registrations to use the interface at around **200** per month

- In total we have created almost **900,000** new records on our database so far this year (excluding records provided by Ingram)

- Each month we process an average of **2.8million** price and availability updates, with January seeing a peak of **3.5million** updates
Aligning data needs

**Data senders**

- Differing data elements included, both in scope and order/structure within the data

- Differing conventions for presenting data elements:
  - i.e. author name: Rowling, J.K., J K Rowling, ROWLING J. K.

- Differing methods of sending data:
  - i.e. ONIX, non-standard files, web portal (pubweb), paper records

**Data receivers**

- Differing requirements in scope and order/structure within the data

- Differing requirements for how data elements are presented

- Differing requirements for receiving data:
  - i.e. ONIX, non-standard files, web services
Aligning data needs

Data sender

Different data receivers require different as well as multiple data feeds

Data receivers
Aligning data needs

Data sender

Central data aggregator

Data transformation, additional fields derived, bespoke feeds generated

Data receivers
Benefits of a central bibliographic aggregator

• Reduces the number of data relationships a data sender/receiver needs to have (potentially to one)

• Reduces the number of formats of data exchange a data sender/receiver needs to have (potentially to one)

• The aggregator can have flexibility of both input and output, to meet the variety of sender and receiver needs

• Helps to maximise publisher reach

• As well as having big resource implications for larger publishers/retailers these factors also ensure that smaller players can communicate with the full market
Benefits of a central bibliographic aggregator

- Standardisation and harmonisation of data

- Consistency
  - improves discovery and comparison
  - also allows greater quality assurance (by identifying inconsistencies/errors)

- Measuring and validating data standards
  - BIC accreditation
  - ONIX compliance

- Other services
  - Order routing (BookNet)
  - Sales monitoring (BookScan)
Summary

• The book trade represents a complex ecosystem of data senders and receivers

• This leads to a huge number and variety of data relationships, with a range of different requirements

• The action of a central metadata aggregator helps to:
  – mediate the complexity of the system
  – manage the matrix of data relationships
  – add value to the data through standardisation and harmonisation
  – support ancillary services such as order routing and sales measurement
Thank you