



Standardizing the discovery of digital book content: a work in progress

Michael Healy, Book Industry Study Group

Working to create a more informed, efficient, and empowered book industry

Goals

- To enable digital book content to be discovered, browsed, searched and distributed in a standardized way
- To allow publishers to manage the quality and availability of the content
- To achieve these goals by defining a set of HTTP transactions between a publisher's digital archive and the websites of syndication partners



Background

- AAP Digital Issues Working Group
- Project Brief and Preliminary Requirements were published in June 2007
- Work transferred to BISG in July 2007
- Establishment of Digital Standards Committee
- Wider industry participation: booksellers, digital repositories, search engines etc.

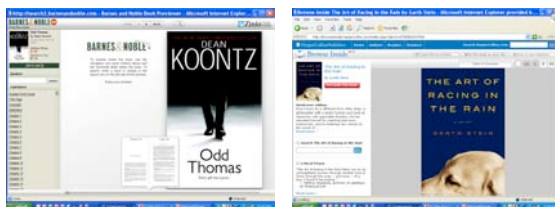


Participating Organizations

- AAP
- ABA
- Barnes & Noble.com
- BookNet Canada
- Bridgeport National Bindery
- CBA International
- Copyright Clearance Center
- Digital Library Federation
- FYI Solutions
- Google
- GS1 US
- Hachette Book Group USA
- Harcourt
- HarperCollins
- IDPF
- Ingram Digital Group
- John Wiley & Sons
- Kirtas
- LibreDigital
- LJNDawson
- McGraw-Hill
- Metacomet Systems
- Microsoft
- NACS
- O'Reilly Media
- Oxford University Press
- Pearson
- Penguin Group (USA)
- Publishing Technology
- Quality Solutions
- R.R. Bowker
- Random House
- Rosetta Solutions
- Sensible Solutions
- Simon & Schuster
- Smerillo Associates
- Workman Publishing



The requirement



....but these implementations referenced two different web service specs – that's not scalable!



Use Cases Covered by the Specification

| | Use Case Reference | Transaction Code | |
|-----------|--------------------|------------------|---|
| SEARCH | UC-01 | <TR01> | Get Number of Phrase Matches in Archive |
| | UC-02 | <TR02> | Search Archive for Phrase |
| | UC-03 | <TR03> | Search Book for Phrase |
| ENUMERATE | UC-04 | <TR04> | Enumerate Archive |
| | UC-04.5 | <TR04.5> | Enumerate Book Info |
| | UC-05 | <TR05> | Enumerate Book Page Images |
| | UC-06 | <TR06> | Enumerate ONIX Info |
| | UC-07 | <TR07> | Enumerate Book Text |
| | UC-08 | <TR08> | Browse by Context Pages |
| RETRIEVE | UC-09 | <TR09> | Browse by Sample Pages |
| | UC-11 | * | Get Full-Page Media |
| | UC-12 | * | Get Thumbnail Media |
| | UC-14 | * | Get Highlighted Page Media |



Project Participants

- **Enumerate Group:** David Mandelbaum (Barnes & Noble) - Chair
- Paul Gore (Hachette)
- Brian Hogue (Ingram Digital)
- Steve Kotrch (Simon & Schuster)
- Ricky Leung (ABA)
- David Maralack (Microsoft)
- Rich Rothstein (Harpercollins)
- **Get Group:** Dave Haase (Random House) – Chair
- Neil DeYoung (Hachette)
- Brian Hogue (Ingram Digital)
- Steve Kotrch (Simon & Schuster)
- Chris Odom (LibreDigital)
- Rich Rothstein (Harpercollins)
- **Search Group:** Brian Hogue (Ingram Digital) – Chair
- Matt Bennett (Hachette)
- Suraj Gaurav (Microsoft)
- Stephanie Halpern (R.R. Bowker)
- Ricky Leung (ABA)



Current status

- Initial BISG Board approval, May 2008
- Technical specification is being finalized
- Prototyping phase – June-August 2008
- Participation of publishers, digital repositories and syndication partners
- Final specification review
- Publication – September 2008



Questions?

Michael Healy
michael@bisg.org