2-1-2003

Enhancing Bibliographical Records by Adding Amazon URL Links - Lingnan University Library's Experience

Patrick Lo
Joe Chow

Follow this and additional works at: https://scholarsarchive.byu.edu/jeal

BYU ScholarsArchive Citation
Available at: https://scholarsarchive.byu.edu/jeal/vol2003/iss129/8

This Article is brought to you for free and open access by the All Journals at BYU ScholarsArchive. It has been accepted for inclusion in Journal of East Asian Libraries by an authorized editor of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.
ENHANCING BIBLIOGRAPHICAL RECORDS
BY ADDING AMAZON URL LINKS –
LINGNAN UNIVERSITY LIBRARY’S EXPERIENCE

Patrick Lo and Joe Chow
Lingnan University Library (Hong Kong)

Profile & History of Amazon.com

Amazon.com, a Seattle-based company introduced to the world as the first well-known Web retailer for books and other educational products, was first established in July 1995 with a mission to use the Internet to transform book buying and shopping into an enjoyable experience.

According to their statistics, Amazon.com’s online site http://www.amazon.com is the leading online shopping site for millions of people from more than 220 countries. In addition to books, they also have the largest selection of products, including free electronic greeting cards, auctioned items, and millions of books, CDs, videos, DVDs, toys and games, electronics, kitchenware, computers and more.

General Statistical Data on Amazon.com (obtained in 1999)¹

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total # of Transactions Handled</strong></td>
<td>23 million</td>
</tr>
<tr>
<td><strong>Total # of Customers</strong></td>
<td>6.2 million</td>
</tr>
<tr>
<td><strong>Total # of Repeated Buyers</strong></td>
<td>58%</td>
</tr>
<tr>
<td><strong>Total # of Items Shipped</strong></td>
<td>7.5 million</td>
</tr>
<tr>
<td>(including books, CDs, and videos)</td>
<td></td>
</tr>
<tr>
<td><strong>Total # of Employees</strong></td>
<td>614 employees</td>
</tr>
<tr>
<td><strong>Sales - growth rate compared to 1998</strong></td>
<td>170%</td>
</tr>
<tr>
<td><strong>Inventory - growth rate compared to 1998</strong></td>
<td>650%</td>
</tr>
</tbody>
</table>

Amazon.com’s Coverage & Selection

Amazon offers a selection of over 3 million titles, including 1.5 million English language titles, which are currently in print, as well as everything listed in “Books Out of Print.” That is more than 150 times the 20,000 to 40,000 titles one might expect to find at a typical neighborhood bookstore.

"As Amazon’s advertisements have pointed out, it would be impossible to duplicate that kind of a selection in any physical building – except for the Pentagon. The upshot is that no matter what type of book you want, you more likely can get it from Amazon than from any “real” bookstore.

What Amazon did differently was to integrate listings for all of the books they could get whether they actually owned them or not into a single catalogue that did not distinguish between what they had and what they would have to order. And with that, Amazon transferred a traditional, low-use, special-order service into Earth’s largest Bookstore.”

Amazon.com actually has less than 10% of the 3 million titles advertised in stock. In other words, Amazon.com relies on wholesalers, publishers, and out-of-print book services to deliver the remaining titles upon individual orders/requests.

**Amazon.com International Websites**

Amazon.com also operates five other international websites, with special focus on different language materials, and they are listed as follows:

<table>
<thead>
<tr>
<th>Countries</th>
<th>Websites</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canada</td>
<td><a href="http://www.amazon.ca">http://www.amazon.ca</a></td>
</tr>
<tr>
<td>United Kingdom</td>
<td><a href="http://www.amazon.co.uk">http://www.amazon.co.uk</a></td>
</tr>
<tr>
<td>Germany</td>
<td><a href="http://www.amazon.de">http://www.amazon.de</a></td>
</tr>
<tr>
<td>France</td>
<td><a href="http://www.amazon.fr">http://www.amazon.fr</a></td>
</tr>
<tr>
<td>Japan</td>
<td><a href="http://www.amazon.co.jp">http://www.amazon.co.jp</a></td>
</tr>
</tbody>
</table>

**Amazon.com Associate Programme**

The Associate Programme of Amazon.com is an online affiliate marketing programme founded in 1996. It was first pioneered by Amazon.com to attract new customers to the site via a new channel. The Amazon.com Associate Programme aims to bring new customers to their website, while helping their Programme Members find another potential financial income source, without

---

impinging on their library services. Because when a customer visits an associate’s website and clicks to make a purchase of the book, the customer is hyperlinked to Amazon.com. If the customer makes a purchase from Amazon.com, Amazon.com will take care of the ordering and associate can earn 5 to 15 percent commission on the sale.

When the Amazon Associate Programme was first established, during the first three months, a total number of 5,000 websites signed up to Amazon.com, ranging from Yahoo!, Netscape, and AT&T Business Network to Explore Madagascar!, the Cigar Journal, and Bexos’s favorite, etc. By the end of 1999, the number of associates exceeded 350,000.3 And according to Amazon.com’s most recent statistics, over 600,000 websites have joined the Amazon.com Associates Programme.

**Enhancing Library’s Bibliographic Records by Introducing Amazon URL Links**

The main advantage of becoming an Amazon Associate Member is that it enables the library to enhance their electronic bibliographic database by providing direct linkage to Amazon.com’s online catalogue. This ultimately leads to much broader and readily available information to the library users, without the burden of inputting and maintaining the information themselves.

In the United States, Tacoma Public Library in Tacoma, Washington (comprising a downtown main library and nine branch libraries) has also become one of the major public libraries to set up partnership with Amazon.com. Users are able to access Amazon.com’s website from their libraries’ online catalogues.

In United Kingdom, Amazon.co.uk has also signed a 3-year agreement with British Library to sponsor the placement of a search facility and banners throughout the British Library’s online public access catalogue, OPAC974. Users following these links will be taken direct to the Amazon.co.uk store. Once a user has carried out a search on the OPAC97, clicking the Amazon link will send the same search to Amazon’s site and bring back the results5. Examples can be found at: http://opac97.bl.uk.

---


4 OPAC97 (Online Public Access Catalogue [1997]) is the version of the British Library’s online public access catalogue, which give access to more than 9 million bibliographic records.

Throughout the years, Lingnan University Library has been exploring various ways to enhance our bibliographical records on the WebPAC with the aim of providing more useful information to our users.

In summer 2002, Lingnan Library became a member of the Amazon Associate programme. Soon afterwards, we grew to be the very first library in the region to offer direct URL links in our WebPAC system to the Amazon book database for English monographs. Our Systems, Reference and Cataloguing staff worked closely together to develop a software for this special project.

By August 2002, more than 76,000 English monograph titles published from 1993 to 2002 have been matched against the holdings of Amazon, with a matching rate of 85%. (see appendix 1)

More than 64,600 of these titles are now added with URL links to Amazon. The following links gives a sample bibliographical record with additional information from Amazon: http://library.ln.edu.hk/search/tgood+to+great (see appendix 2)

Advantages
Since Amazon has the largest electronic book list in the world, linking our bibliographic database to it will significantly enrich our Webpac records. Our users can call up all sorts of value-added information readily available at Amazon.com in the most convenient and speedily fashion, e.g.:

(a) Reviews on books, e.g., full-text reviews, customer reviews, author interviews, recommendations, commentaries, and sometimes full-text excerpts of the books, etc.

(b) Cover arts and text, and book jacket

(c) Table of contents

(d) Request lists of similar books

(e) Users can enter their own comments

(f) Price and availability

(g) Via the online customer book reviews and lists of purchases made, the Library is able to track the users’ preferences. And such data could be used as are most
relevant and valuable indicators for the Library’s collection development

Other Advantages
In addition, users may proceed with the online purchase by entering their own order and shipping details directly from Amazon.com. Services and information are available around the clock and delivery is speedy (shipping within 24 hours) for many items. Location is also not a problem; online shopping could be conducted anytime and anywhere as long as an Internet connection exists.

Amazon.com also offers a range of highly personalized services based on data collected from an individual’s use of their website, which exceeds the customer services of many independent bookstores, e.g.:

1. Collecting customers’ e-mail addresses and generating updates automatically to inform their customers the status of their orders.

2. Sending e-mails to notify their customers when new books of interests to them have been published.

3. Informing users the availability of the next book in the series that customers are reading.

4. Based on one’s buying habits, Amazon can also recommend books to customers based on what others with similar preferences are buying.

5. The majority of books and other items are offered at discounted prices.

6. Moreover, Amazon can help the customers to select gifts, providing optional gift-wrapping service, with personal notes enclosed, and ship them anywhere in the world.

Software Developed for Matching URL Links

In July 2002, Lingnan Library’s Systems, Reference and Cataloguing staff worked closely together and developed unique software especially for tackling this project.
This unique Software is designed to perform mechanical matching of our bibliographic records (in MARC format) against the Amazon.com online catalogue/website. MARC Tag 020, i.e., the ISBN number, is used as the ultimate matching point. After matching titles are identified, the Software will automatically insert the relevant URL links into MARC field 856 of the bibliographic records. The benefits of this Software are evident, as the time and manpower originally required for manual checking can be eliminated. Not to mention the fact that it is a great way to avoid errors caused by human input.

**Software Requirements**

Basic requirements for the locally developed Software are as follows:

1. Pentium III 1GHz or above
2. 256MB RAMs or above
3. Windows 98/ME/2000/XP
4. Visual Basic 6.0

**Procedures for Locating & Matching URL Links in Amazon.com**

Pre-processing & Record Selection

1. Run a review file to select bibliographic records in MARC format that need to be enhanced.

2. Use the FTP function to export the MARC Records out of the library system, INNOPAC.

3. Convert the MARC Records to Text based records via the use of MARCBREAKER.

4. Define the program parameters as:
   - Input file name
   - Output file name
   - Error report file name
   - Wording & description for MARC field 856 subfield $z
   - URL header
   - URL tail
   - Error detection string

5. Input Material:
   - Text based MARC records

6. Actual Processing/Matching
   (a) The programme reads the MARC record line by line and identifies the MARC Tag 020 (ISBN number).
(b) The programme will then automatically validate the ISBN number by using "Checksum" error checking.

(c) If digital error are found in the ISBN numbers, they will be identified in the "Error Report".

(d) If ISBN numbers are proved to be valid, it will generate an online link query for Amazon.com online catalogue by using the predefined URL Header + ISBN + URL Tail. (see appendix 3)

(e) If the title is not stocked by Amazon.com, it will be recorded in the "Error report".

(f) If matching titles are found, the relevant Amazon.com URL links will be inserted to subfield $u$ of MARC field 856, and the public note "Check amazon.com for reviews and other information" will be inserted to subfield $z$.

(g) All records upgraded with 856 will be gathered together in the output file.

(h) At the completion of the matching process, the programme will submit the following reports/files:

- 1st File – featuring upgraded MARC Records (with added Amazon.com URL links in field 856)
- 2nd File – an Error Report, documenting records with no matching title in Amazon.com catalogue.

7. Final/Post-Processing

(a) Text-based file will be converted back to ISO2709/MARC records by using MARCMARKER.

(b) Import upgraded MARC records back to library system via FTP.

(c) MARC records upgraded with Amazon.com URL links can now overlay the original ones.

Program Screenshot

The following is a screenshot taken from the Software programme.
### Input File
- input.txt

### Output File
- output.txt

### Error Log
- Errorlog.csv

### Add Field When
- 020

### URL Head

### URL Tail
- /ingnanuniver-20 target=_blank

### Subfield 2
- Check amazon.com for reviews and other information

### Error Detection String
- &lt;title&gt;Amazon.com Error Page&lt;/title&gt;

---

**Progress Report:**

- 36 Records Read
- 38 ISBN Read
- 32,856 Links Added
- 6 ISBN Error Found

**Start at:** 2002/5/25 AM 03:30:53
Flowchart for Software Programme

START

Read inputed MARC Records line by line and find MARC Tag 020 (ISBN)

Validates the ISBN by "Checksum" error checking

ISBN valid format & check digit?

No

Generate a link to query Amazon.com online

No

Available at Amazon.com?

No

ISBN# with error will be reported to Error Report

URL & Public note will be added to MARC Tag 856 $u and $z

All upgraded records with 856 link will be gathered in output file

END
Overall Statistics of Lingnan-Amazon Project

Statistical data on overall matching rate, time required for URL-link matching and time required for reloading the bibliographical records are listed as follows:

Overall Statistical Data for Amazon Project\(^6\)

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of bibliographic records matched against Amazon.com</td>
<td>76,096 records</td>
</tr>
<tr>
<td>Year Range/Coverage</td>
<td>All English language monographs published between 1993-2002</td>
</tr>
<tr>
<td>No. of bibliographic records FOUND matching URL links in Amazon.com</td>
<td>64,521 records</td>
</tr>
<tr>
<td>Overall matching rate</td>
<td>85%</td>
</tr>
<tr>
<td>Publication year that has the highest matching rate</td>
<td>2001 = 89%</td>
</tr>
<tr>
<td>Publication year that has the lowest matching rate</td>
<td>1997 = 81%</td>
</tr>
<tr>
<td>Subject area that has the highest matching rate</td>
<td>Philosophy &amp; related disciplines = 95%</td>
</tr>
<tr>
<td>Subject area that has the lowest matching rate</td>
<td>Social Sciences = 79%</td>
</tr>
<tr>
<td>Hours required for matching Bib. records against Amazon.com</td>
<td>20 hours</td>
</tr>
<tr>
<td>Hours required for reloading Bib. records</td>
<td>12 hours</td>
</tr>
</tbody>
</table>

Traffic Summary\(^7\)

(Period Covered)
1\(^{st}\) July – 1\(^{st}\) October 2002

<table>
<thead>
<tr>
<th>Unique Visitors(^8)</th>
<th>Clicks</th>
<th>Item Ordered</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,149</td>
<td>1,843</td>
<td>3</td>
</tr>
</tbody>
</table>

Experiments on Other International Amazon.com Websites

Lingnan Library also experimented with our foreign-language materials by matching them against the following international Amazon.com websites:

\(^6\) Data obtained in August 2002.

\(^7\) Information about the items your visitors have ordered and the linking methods that work best from your Web site. The numbers are based on items ordered but not necessarily shipped.

\(^8\) The number of visitors who have clicked through to Amazon.com from your Associates links. Repeat click-throughs are not counted as unique visits.
1. Germany -- http://www.amazon.de
3. Japan -- http://www.amazon.co.jp

But the overall matching rate was only 15%, which was relatively low compared to the original Amazon.com site. The low matching rates in other international Amazon.com sites were directly related to the fact that Lingnan Library has a very small collection of foreign-language materials, and many of them are of a language-learning nature. When the Lingnan Library's foreign-language collection expands, the matching rate in the above international Amazon.com websites is expected to increase.

Conclusions & Future Developments

The major advantage of the Amazon.com Project is that it saves our Library immeasurable cost, time and manpower in providing, as well as maintaining the added information in our bibliographic catalogue for our users, since after all, it is a highly labour intensive process.

In addition, information from Amazon.com is readily available, and no special training is required for our users since the user-interface is highly user-friendly and supports typical searching of its catalogue by author, title, and subject. Not to mention that the Amazon.com search engine is highly tolerant of faults, as it offers “nearest match” items if an exact match is not found.

Given such a successful experience, Lingnan Library is planning to further explore other alternatives of enhancing our bibliographic records in a similar fashion in the future, e.g.:

1. Further enhancing English books, e.g. Barnes & Noble.

2. To identify other effective ways to introduce Amazon.com URL links to our bibliographic records for audio-visual materials in any languages, e.g., CD, DVDs, CD-ROMs, etc. While ISBN numbers are not used for these materials, it poses a different challenge for our project.

3. To explore other potential websites or web book retailers which specialize in the selling of Chinese and other Oriental-language materials.
Bibliography References

   *Searcher*, vol. 7, issue 3, pp. 34-47.

   issue 13, pp. 44-47.


## Appendix 1

<table>
<thead>
<tr>
<th>Subject Areas</th>
<th>Matching Rate</th>
<th>Pub. Years</th>
<th>Matching Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>001-099 Generalities</td>
<td>94%</td>
<td>1993</td>
<td>84%</td>
</tr>
<tr>
<td>100-199 Philosophy &amp; Related Disciplines</td>
<td>95%</td>
<td>1994</td>
<td>86%</td>
</tr>
<tr>
<td>200-299 Religion</td>
<td>91%</td>
<td>1995</td>
<td>85%</td>
</tr>
<tr>
<td>300-399 Social Sciences</td>
<td>79%</td>
<td>1996</td>
<td>83%</td>
</tr>
<tr>
<td>400-499 Language</td>
<td>85%</td>
<td>1997</td>
<td>81%</td>
</tr>
<tr>
<td>500-599 Pure Sciences</td>
<td>91%</td>
<td>1998</td>
<td>83%</td>
</tr>
<tr>
<td>600-699 Technology &amp; Applied Sciences</td>
<td>87%</td>
<td>1999</td>
<td>87%</td>
</tr>
<tr>
<td>700-799 Arts</td>
<td>86%</td>
<td>2000</td>
<td>88%</td>
</tr>
<tr>
<td>800-899 Literature</td>
<td>92%</td>
<td>2001</td>
<td>89%</td>
</tr>
<tr>
<td>900-999 Geography and History</td>
<td>90%</td>
<td>2002</td>
<td>88%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>85%</strong></td>
<td><strong>Total</strong></td>
<td><strong>85%</strong></td>
</tr>
</tbody>
</table>

Subject Areas – Matching Rate in Amazon Catalogue

![Subject Areas Matching Rate](image-url)
Publication Years – Matching Rate in Amazon Catalogue

Appendix 2a
Sample bibliographical record with additional information from Amazon

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>Good to great: why some companies make the leap, and others don’t / Jim Collins</td>
</tr>
</tbody>
</table>

Click on the following to:
Check amazon.com for reviews and other information

URL links to Amazon.com
Appendix 2b

Sample pages from Amazon.com
Appendix 3

Record added with online link for Amazon.com

<table>
<thead>
<tr>
<th>File</th>
<th>Edit</th>
<th>New</th>
<th>Save</th>
<th>Remove</th>
<th>Print</th>
<th>Copy</th>
<th>Commands</th>
<th>Notes</th>
<th>Help</th>
</tr>
</thead>
<tbody>
<tr>
<td>B13590881 29-01-2002 LAN:eng COU:nyu LOC:cie</td>
<td>^E &gt; END</td>
<td>^F &gt; Show FULL Menu</td>
<td>OVERWRITE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

092 658 COLL 2001
010 2001024818
020 0066620996 (hc)
040 DLC|dLDC
008 010423s2001 nyuac b 001 0 eng cam a
050 00 HD57.7|b.C645 2001
082 00 658|221
100 1 Collins, James C.|q(James Charles),|q(d1958-
245 10 Good to great :|bwhy some companies make the leap, and others don't / |cJim Collins
300 xii, 300 p. :|bill., port. ;|c25 cm
504 Includes bibliographical references and index
650 0 Leadership
650 0 Strategic planning
650 0 Organizational change
650 0 Technological innovations|xManagement
656 40 |uhttp://www.amazon.com/exec/obidos/ASIN/0066620996/lingmanuniver2-20
   target= _blank|zCheck amazon.com for reviews and other information
997 bn