Accessing library materials via Google and Other Web Sites

*Janifer Gatenby, Strategic Research, OCLC PICA*
Modern users

- Use Google, Yahoo, MSN ++
- Amazon, eBay, Netflix
- Search & find themselves
- Register & order
  - Internationally
  - Request home delivery
Google and Yahoo! timeline

Aug ’03: Google contract signed

Oct ’03: Google harvests 150,000 records

Jan.’04: Yahoo contract signed

May ’04: 725,000 inbound links logged (SSO)

May ’04: Yahoo harvests records

Jun ’04: Yahoo completes indexing of 2 million WC records

Sep ’03: Google given go-ahead to harvest records

Dec.’03: Records begin to appear in Google; 800 inbound-links logged (search-site-originating [SSO])

Mar.’04: 109,000 inbound links logged (SSO)

Aug ’06 Countries with loading contracts to Google

Australia, China, *Czech Republic, *Denmark, Ireland, Israel, Hungary, Lithuania, *Netherlands, Taiwan, *United Kingdom, *United States

* Significant contributions via WorldCat
How Data is Contributed

- XML record
  - Subset at work level
  - ~75% holdings

- Inktomi Data Interchange Format (IDIF)

- Title, author, publisher, identifiers, electronic location, subject, contents, country holding, holdings count

- Google
  - 4.4 million (3m clusters)

- Google Scholar
  - More: whatever of 67 million maps to its dbase

- Google Books
  - 100% mapping TO

- Yahoo
  - 3.5 (3m clusters)

- MSN
  - 4.5 million growing to 10 million
  - Physical sciences & biomedical
Open WorldCat

Limitations on loading imposed by harvesters – Google only interested in titles with many holdings

Exposure
Common Problems

- Coverage
- Ranking
  - Conflicting business models
  - Page hits, referrals + $ revenues
  - 2 billion pages 2004; 12 billion 2007
- Matching
- Work level data needed
- Not enough influence
All 86+ million, representing 1,000+ million holdings, 40,000+ libraries

WorldCat.org

In de Bibliotheek? powered by WorldCat

![Girl with a pearl earring](image)

**Girl with a pearl earring**
door [Tracy Chevalier](#)

Soort: Engels | Boek | Fictie
Uitgever: New York | Dutton | G1099

Onderwerp: Vermeer, Johannes | Nederlandse | Social life and customs | 17th century | Fijnschilderij

Vnl locatiegegevens incl. postzegel, postcode, provincie en land

Bibliotheek 1-98 van 2474 worden getoond.

<table>
<thead>
<tr>
<th>Bibliotheek</th>
<th>Details</th>
<th>Onderwerpen</th>
<th>Edities</th>
<th>Raciale</th>
</tr>
</thead>
</table>

Beperk de resultaten op medium: Alle media tonen

1. **Openbare Bibliotheek Den Haag**
   - Den Haag, 2511 BT Netherlands
   - Boek
   - Services: Go naar de publicatiebladzijde
   - Stel een vraag aan een bibliothecaris: 9 miljoen

2. **Erasmus Universiteit Rotterdam**
   - Rotterdam, 3000 CA Netherlands
   - Boek
   - Services: Go naar de publicatiebladzijde
   - Stel een vraag aan een bibliothecaris: 16 miljoen

3. **Stadsbibliotheek Haarlem**
   - Haarlem, 2011 XP Netherlands
   - Boek
   - Services: Go naar de publicatiebladzijde
   - Stel een vraag aan een bibliothecaris: 17 miljoen

4. **Bibliotheek Rotterdam**
   - Rotterdam, 3011 BV Netherlands
   - Boek
   - Services: Go naar de publicatiebladzijde
   - Stel een vraag aan een bibliothecaris: 17 miljoen

---

OCLC PICA

Exposure
Positive Signs: Statistics

- **March 2007**
  - Total Referrals: 14,118,777
    - Av. Per day: 455,444
  - Links to library services: 834,866
    - % Conversion to lib = 10.5%

- **03 2006**
  - ↑ 96%
  - ↑ 363%
Referrals to WorldCat.org pages

Referrals

Month over month

2005
2006

July
August
September
October
November
December

In millions

Exposure
What we measure:

- Impressions
- Click-throughs
- Conversions
### WorldCat.org Tracked Statistics at-a-glance

<table>
<thead>
<tr>
<th><strong>Impressions:</strong></th>
<th><strong>Click-through:</strong></th>
<th><strong>Conversion:</strong></th>
<th><strong>Conversion rate:</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>300M to 1B/year</td>
<td>10M/ month</td>
<td>550k/month</td>
<td>Overall: 5.5% = “best in class” for web site</td>
</tr>
<tr>
<td></td>
<td>120M per year</td>
<td>6.6M/year</td>
<td>5.9% in January 07</td>
</tr>
</tbody>
</table>

It’s worth noting that our conversion rate is well above Google’s “typical” range of 1-3%.
Worldcat.org Referrals March / April 2007

- Library sites: 31%
- Education sites: 9%
- Yahoo: 26%
- Google books: 14%
- blogs, CNN, AltV+: 8%
- Google: 9%
- Google Scholar: 3%
- Referral sources: <50% of traffic
Web Page Rankings

Rankings:

1. Video search engine
2. Encyclopedia
3. Online shopping site
4. Image hosting site
5. Social networking site

Alexa Placement:

- Rank: 4
- Alexa Value: 19,652 (11,000)
Some Observations

• Need for maximum exposure
  – Sites in addition to Google
  – Google & Yahoo swap places as top referrers

• Users value library information
  – Conversions way above web average

• Alternative means of exposure are effective
Worldcat Websites

People Who Like Worldcat Sites

- mazaazam
  30 / M
  New Jersey USA

- castagna
  30 / M
  England UK

- judell
  M
  New Hampshire USA

- marijar
  F
  Croatia

Featured Worldcat Sites

- Try a search [OCLC - Open WorldCat program]
  Reviewed by Metamorphosis on Jan 7, 2006 10:23pm
  A tool to find books at a library near you, cool.

- online.wsj.com/article_email/0,,SB110298438601999135-IRjf...
  Reviewed by judell on Dec 14, 2004 4:47am
  Digging back into a separate database, Google’s service allows consumers to type in their ZIP Code and
  by libraries near them.

- WorldCat [OCLC]
  Reviewed by mamabear10 on Feb 25, 6:29am
  This looks associated!
More observations

- **Google Print**
  - 950,765 in March 2007
  - Link from snippets to metadata
  - Copyrighted material searching holdings

- **Hard working with search engines**
## Referrals to Amazon.com Q1 2007

<table>
<thead>
<tr>
<th></th>
<th>Jan-07</th>
<th>Feb-07</th>
<th>Mar-07</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ger</td>
<td>Referrals</td>
<td>0</td>
<td>1546</td>
</tr>
<tr>
<td>Can</td>
<td>Referrals</td>
<td>1493</td>
<td>1311</td>
</tr>
<tr>
<td>UK</td>
<td>Referrals</td>
<td>7291</td>
<td>6525</td>
</tr>
<tr>
<td>US</td>
<td>Referrals</td>
<td>71072</td>
<td>82813</td>
</tr>
<tr>
<td>total</td>
<td></td>
<td>79856</td>
<td>92195</td>
</tr>
<tr>
<td>Ger</td>
<td>Conversions</td>
<td>0</td>
<td>4.72%</td>
</tr>
<tr>
<td>Can</td>
<td>Conversions</td>
<td>3.82%</td>
<td>1.68%</td>
</tr>
<tr>
<td>UK</td>
<td>Conversions</td>
<td>8.30%</td>
<td>9.16%</td>
</tr>
<tr>
<td>US</td>
<td>Conversions</td>
<td>4.84%</td>
<td>4.26%</td>
</tr>
<tr>
<td>av.</td>
<td></td>
<td>5.65%</td>
<td>6.60%</td>
</tr>
<tr>
<td>Ger</td>
<td>Orders</td>
<td>373</td>
<td>104</td>
</tr>
<tr>
<td>Can</td>
<td>Orders</td>
<td>57</td>
<td>22</td>
</tr>
<tr>
<td>UK</td>
<td>Orders</td>
<td>605</td>
<td>598</td>
</tr>
<tr>
<td>US</td>
<td>Orders</td>
<td>3438</td>
<td>3525</td>
</tr>
<tr>
<td>total</td>
<td></td>
<td>4100</td>
<td>4218</td>
</tr>
</tbody>
</table>

---

*Remember clicks to libraries 834,886*
Wanted title found

Pedigrees of Polperro ...
by Frank Hill Perrycoaste

Type: English: Book
Publisher: [Liskeard, Cornish Times] 1925.
OCLC: 26501597

Enter Location Information: [field]

Displaying libraries 1-3 out of 3

1. Corporation of London Libraries
   London, EC2P 2EJ United Kingdom
   - Book
   Services: Library Information

2. London Library
   London, SW1Y 4LG United Kingdom
   - Book
   Services: Library Information

3. University of Oxford
   Oxford, OX1 3LU United Kingdom
   - Book
   Services: Library Information, Ask a librarian

Discovery to Delivery Gap
Get It Button

![Search Results for Moby Dick or the Whale](image.png)

<table>
<thead>
<tr>
<th>Supplier</th>
<th>Author</th>
<th>Format</th>
<th>Supplier</th>
<th>Fee</th>
<th>Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moby Dick or the Whale (Modern Library)</td>
<td>Herman Melville, Rockwell ...</td>
<td>Book</td>
<td>Amazon</td>
<td>$14.93</td>
<td>Purchase</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Book</td>
<td>Barnes &amp; Noble</td>
<td>$21.95</td>
<td>Purchase</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Book</td>
<td>Open WorldCat</td>
<td>Free</td>
<td>2 week loan</td>
</tr>
<tr>
<td>On Time and Being</td>
<td>Martin Heidegger</td>
<td>Book</td>
<td>Amazon</td>
<td>$24.03</td>
<td>Purchase</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Book</td>
<td>Barnes &amp; Noble</td>
<td>$29.85</td>
<td>Purchase</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Book</td>
<td>Open WorldCat</td>
<td>Free</td>
<td>2 week loan</td>
</tr>
</tbody>
</table>

OCLC PICA

Discovery to Delivery
Delivery Components

- online user registration, (free)
- registry of digital masters (free)
- global library registry (free)
- direct home delivery pilot
- NLA also conducting ambitious DD pilot

The delivery gap will not be solved just by making available new technical components. Libraries must make the necessary policy changes to cooperate on an international scale.
The motivation of exposure has been to attract one’s own user population to the resources of one’s library.

• BUT ALSO
  – Your users find things in other libraries on a global scale
  – Other libraries’ users from anywhere in the world find things in yours
In Conclusion

- We cannot expect the user to come to the library web site anymore
- It doesn’t matter if our sites are found by the back door: the more the merrier
- Now is the time to focus on delivery
  - Focus is on discovery at the expense of Delivery

Just how many of our users who are currently using Amazon would acquire from a library instead if it were just as easy
Acknowledgments to:
Chip Nilges and Bill Brembeck