Reading 2.0

Tim O’Reilly

O’Reilly Media, Inc.

www.oreilly.com

University of Michigan
March 10, 2006
What Job Does a Book Do?
Internet Public Library
Online public library features directories of online texts, newspapers, magazines, reference materials with special sections for youth.
www.ipl.org/ - 30k - Mar 5, 2006 - Cached - Similar pages
Science Fair - Newspapers - POTUS - Books
More results from www.ipl.org »

The Library of Congress
www.loc.gov/index.html - 10k - Cached - Similar pages

LibrarySpot.com: Encyclopedias, maps, online libraries, quotations ...
Find the best library and reference resources at LibrarySpot.com, including top dictionaries, encyclopedias, newspapers, maps, quotations and much more.
www.libraryspot.com/ - 29k - Mar 5, 2006 - Cached - Similar pages

The WWW Virtual Library
A guide to selected resources across all the main academic areas. The original web directory of Tim Berners-Lee.
vlib.org/ - 6k - Mar 5, 2006 - Cached - Similar pages

THE BRITISH LIBRARY - The world’s knowledge
This is the home page of the British Library website. Use this site to search our catalogues, order items for research, view exhibitions and link to ...
www.bl.uk/ - 24k - Cached - Similar pages

Awesome Library - K-12 Education Directory
Awesome Library organizes 30000 carefully reviewed K-12 education resources, including the top 5 percent for teachers, students, parents, and librarians.
www.awesomelibrary.org/ - 18k - Cached - Similar pages
Library

From Wikipedia, the free encyclopedia

**Alternative meanings:** Library (computer science), Library (electronics), Library (biology)

In the traditional sense of the word, a library is a collection of books and periodicals. It can refer to an individual's private collection, but more often it is a large collection that is funded and maintained by a city or institution. This collection is often used by people who choose not to, or cannot afford to, purchase an extensive collection themselves. However, with the collection or invention of media other than books for storing information, many libraries are now also repositories and access points for maps, prints or other artwork, microfilm, microfiche, audio tapes, CDs, LPs, video tapes and DVDs, and provide public facilities to access CD-ROM databases and the Internet.

Thus, modern libraries are increasingly being redefined as places to get unrestricted access to information in many formats and from many sources. More recently, libraries are understood as extending beyond the physical walls of a building, providing assistance in navigating and analyzing tremendous amounts of knowledge with a variety of digital tools.

**Contents**

1. Etymology of the word
2. History
3. Types of libraries
4. Description
5. Library use
6. Library management
7. Funding problems
8. Some famous libraries
9. See also

A modern-style library in Chambéry
Britannica vs. Wikipedia

Daily Reach (per million)

Britannica.com

Wikipedia.org

0 5,000 10,000 15,000 20,000 25,000 30,000 35,000 40,000

Apr Jul Oct 2006

© 2006 Alexa

2006 Mar 7

O'REILLY®
6.1. THE BASICS OF INFORMATION GRAPHICS

"Information graphics" simply means data presented visually, with the goal of imparting knowledge to the user. I'm including views in that description because they are inherently visual, even though they're constructed primarily from text instead of polygons. Other familiar static information graphics includes maps, flowcharts, bar plots, and diagrams of real-world objects.

But we're dealing with computers, not paper. You can make almost any good static design better with interactivity. Interactive information graphics allow users to hide and show information as she needs it, and they put the user in the "driver's seat" as she chooses how to view and interact with the data.

Even the mere act of manipulating and rearranging the data in an interactive graphic has value—the user becomes a participant in the discovery process, not just a passive observer. This can be invaluable. The user may not produce the world's best-designed infographics, but she understands the data like never before. This may be invaluable. The user may not produce the world's best-designed infographics, but she understands the data like never before. The user may not produce the world's best-designed infographics, but she understands the data like never before.

Ultimately, the user's goal in using information graphics is to learn something. But the designer needs to understand what it is that the users need to learn.
design aspects are most important to you in a given interface.

Additional reading

Safari has identified sections in other books that relate directly to this selection using Self-Organizing Maps (SOM), a type of neural network algorithm. SOM enables us to deliver related sections with higher quality results than traditional query-based approaches allow.

<table>
<thead>
<tr>
<th>Section Title</th>
<th>Book Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. THE BASICS OF INFORMATION GRAPHICS</td>
<td>Designing Interfaces</td>
</tr>
<tr>
<td></td>
<td>By Jennifer Tidwell</td>
</tr>
<tr>
<td></td>
<td>By Jennifer Niederst</td>
</tr>
<tr>
<td>3. Graphics</td>
<td>Web Design in a Nutshell</td>
</tr>
<tr>
<td></td>
<td>By Jennifer Niederst</td>
</tr>
<tr>
<td>4. Graphics</td>
<td>PHP Hacks</td>
</tr>
<tr>
<td></td>
<td>By Jack Herrington</td>
</tr>
<tr>
<td></td>
<td>By Neil Randall, Dennis Jones</td>
</tr>
<tr>
<td></td>
<td>By Dennis Jones, Neil Randall</td>
</tr>
<tr>
<td></td>
<td>By Steve Johnson - Perspection, Inc.</td>
</tr>
<tr>
<td>8. The AWT Graphics Model</td>
<td>Java GUI Development: The Authoritative Solution</td>
</tr>
<tr>
<td></td>
<td>By Vartan Piroumian</td>
</tr>
<tr>
<td></td>
<td>By Molly E. Holzschlag</td>
</tr>
<tr>
<td>10. Slicing Graphics for Placement in a Table</td>
<td>Special Edition Using XHTML™</td>
</tr>
<tr>
<td></td>
<td>By Molly E. Holzschlag</td>
</tr>
</tbody>
</table>
Create, Publish, and Share Customized Computer Science and Information Technology Course Materials.

SafariU is a revolutionary way to teach exactly what you want to teach, how you want to teach it. With SafariU you can give your students more in-depth and focused content, far greater convenience, and unprecedented value. And do it all from your own computer, whenever you're ready.

It's your course. Why not make your teaching tools as dynamic, leading-edge, and unique as possible?

CREATE YOUR CUSTOM TEXTBOOK

Select just the right content from more than 2200 books and 5000 articles from leading technology publishers and experts. Mix in your own materials. Arrange everything just as you want, and have your custom textbooks professionally printed—all at no cost to you.

CREATE YOUR ONLINE SYLLABUS

Go beyond the book with an up-to-date online syllabus. Offer students an electronic bookshelf of up to ten complete books used to create your custom text. And feel free to create links to additional articles, resources, and websites.
What Job Does a Library Do?
Until the Viola Browser Archive is restored, please see a mirror of the website at UC Berkeley

The Global Network Navigator

Features

1. Gingerbread House going online -- Punsel Sugartooth
   What will the future hold for the Gingerbread House as it enters the electronic age? Will kids around the world rejoice at the new-found connectivity to the established bakery? (more...)

2. The black art of debugging -- Mont Todd
   Mont shows us how to avoid classic mistakes in setting debugger breakpoints, and other tricks and techniques.

Columns

Departments

The Whole Internet Resource

No Search

Message:

URL: file://pobble/..../work/viola/docs/tests/tostHPano.html
Archive.org vs. the Library of Congress

Daily Reach (per million)
loc.gov Archive.org

© 2006 Alexa
Why the Google Library Project Matters
iTunes

- Tens of billions of songs ripped from CD
- 1 Billion songs sold
- Software above the level of a single device
- Database back end
- Web services-enabled (CDDB)
- Songs, not Albums
- User controlled playlists
- Local sharing (loose DRM)
## Similar Artists

Based on overall listening habits

<table>
<thead>
<tr>
<th></th>
<th>Artist</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Devendra Banhart</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>Joanna Newsom</td>
<td>88</td>
</tr>
<tr>
<td>3</td>
<td>Six Organs of Admittance</td>
<td>79</td>
</tr>
<tr>
<td>4</td>
<td>Vashti Bunyan</td>
<td>78</td>
</tr>
<tr>
<td>5</td>
<td>Josephine Foster</td>
<td>77</td>
</tr>
<tr>
<td>6</td>
<td>Animal Collective</td>
<td>72</td>
</tr>
<tr>
<td>7</td>
<td>Little Wings</td>
<td>71</td>
</tr>
<tr>
<td>8</td>
<td>Jana Hunter</td>
<td>67</td>
</tr>
<tr>
<td>9</td>
<td>Castanets</td>
<td>64</td>
</tr>
<tr>
<td>10</td>
<td>white magic</td>
<td>62</td>
</tr>
<tr>
<td>11</td>
<td>Karen Dalton</td>
<td>60</td>
</tr>
<tr>
<td>12</td>
<td>viking moses</td>
<td>59</td>
</tr>
<tr>
<td>13</td>
<td>Akron/Family</td>
<td>57</td>
</tr>
<tr>
<td>14</td>
<td>M. Ward</td>
<td>56</td>
</tr>
<tr>
<td>15</td>
<td>Michael Palin</td>
<td>55</td>
</tr>
<tr>
<td>16</td>
<td>Scout Niblett</td>
<td>51</td>
</tr>
</tbody>
</table>
The Internet BookMobile
The Orphaned Works Problem

- 4% In Print
  - 1.2 million unique titles sold in 2004*

- ~75% or more The Twilight Zone
  - may be in copyright, but not for sale
  - rights may have reverted to author
  - may be in the public domain

- ~20%** Public Domain

32 million unique titles in all US Libraries **

*Source: Nielsen BookScan
**Source: Online Computer Library Center

http://www.dlib.org/dlib/september05/lavoie/09lavoie.html
Fewer Than 4% of Published Books Are Still Commercially Exploited

- Number of Titles in Twilight Zone
- Number of Titles Commercially Exploited
- Number of Titles In Public Domain
The Near-Term Opportunity
The Long Tail

**ANATOMY OF THE LONG TAIL**

Demand for products *not* available in traditional bricks and mortar stores is potentially as big as for those that are [because of limits on shelf space].

-- Chris Anderson, via email
6.1. THE BASICS OF INFORMATION GRAPHICS

"Information graphics" simply means data presented visually, with the goal of imparting knowledge to the user. I'm including views in that description because they are inherently visual, even though they're constructed primarily from text instead of polygons. Other familiar static information graphics includes maps, flowcharts, bar plots, and diagrams of real-world objects.

But we're dealing with computers, not paper. You can make almost any good static design better with interactivity. Interactive users can choose to show or hide the interface as they need it, and they can put the user in the "driver's seat" as she chooses how to view and interact with the data.

Even the mere act of manipulating and rearranging the data in an interactive graphic has value—the user becomes a participant in the discovery process, not just a passive observer. This can be invaluable. The user may not produce the world's best-designed graphs, but the process of manipulating that plot or table puts her face-to-face with data that she may never have noticed otherwise.

Ultimately, the user's goal in using information graphics is to learn something. But the designer needs to understand what that learning is about.
Does Online Search Drive Discovery?

• Compare sales of physical books versus e-books
  – Sales of physical books, as reported by Bookscan
  – E-Books – sales of same books, but in e-book format from Safari

• Questions
  – Is there a difference on the long tail?
  – If so, how much does it account for?
Physical Books Overview

• Measured by Neilsen Bookscan Top 10,000 Computer Books report
• POS (Point of Sales) data, which represents approximately 70% of US market
  – dominated by Barnes and Noble, Borders and Amazon.
• For first 6 months of 2005, there were:
  – 17,754 distinct SKUs (ISBN’s)
  – A book must have had sales of at least one unit in the period to be included
  – A book must have made its first appearance in before 2005 to be counted
Safari Overview

• Safari is an e-Book service
  – Library of O’Reilly and Pearson titles, which have a combined market share of 46% based on unit sales
  – Subscription based -- subscribers can select a fixed number of books from a large library of books.
• Of the 17,754 physical ISBNs sold in the first 6 months of 2005
  – 2,171 were also available through Safari
  – In this study, a view on Safari occurs when someone accesses a page in a book on their bookshelf.
Methodology Overview

• Rank the 17,754 books by total unit sales and divide into blocks of 500 based on range
  – Range 1-500 = top 500 books by unit sales
  – Range 501-1000 = next 50 for next block
• Find the Safari views for the books in each block
• Are there any differences?
Total Units Sold vs. Safari Views

Physical Rank Range

Physical Units
Safari Hits
Notes

• Unit totals follow power law (Long tail)
• Views appear to follow power law, but begin to climb back up at the far end of the long tail
  – These are total views, though, so what happens when you look at weekly average views per title for each block?
• Outliers removed to avoid skewing data
Averages

- Avg. Physical Units per Week
- Avg. Safari Hits per Week

Physical Rank Range

- 1,500
- 1,001-1,500
- 2,001-2,500
- 3,001-3,500
- 4,001-4,500
- 5,001-5,500
- 6,001-6,500
- 7,001-7,500
- 8,001-8,500
- 9,001-9,500
- 10,001-10,500
- 11,001-11,500
- 12,001-12,500
- 13,001-13,500
- 14,001-14,500
- 15,001-15,500
- 16,001-16,500
- 17,001-17,500
How Big is the Long Tail?

• Need to know where the long tail starts, a point we refer to as C
  – all books with rank less than C are short head, all books with a rank greater than C are long tail
• No direct data source to know which books are available online only. A best guess estimation:
  – Create two groups of titles, those that appear in the in the BookScan 3k list (better selling titles) and those that only appear in the BookScan 10k list (potential long tail).
  – Find the rank where more titles only appear in the BookScan 10k list and use that as C*
Cutoff Estimate

- $C \sim \text{a rank of 4,500}$
Conclusion: Shift in Demand Curve

- Additional demand from the long tail effect
Rough Estimates

- Using 4,500 as a cutoff, we can compute rough estimates for long tail effect using proportions of unit sales and views

<table>
<thead>
<tr>
<th></th>
<th>Short Head</th>
<th>Long Tail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical Books</td>
<td>94%</td>
<td>6%</td>
</tr>
<tr>
<td>Safari</td>
<td>77%</td>
<td>23%</td>
</tr>
</tbody>
</table>

- Long Tail is important part of Safari
  - 23% of Safari views
  - 6% of physical book sales
Executive Summary

• Our findings confirm that there is demand for technical books that are generally unavailable offline.
  – In fact, among the top books in our online service are books that either out of print, or older editions of titles that are no longer carried by traditional bookstores.

• The long tail effect online is 2 to 3 times that of physical books
Google Print vs. Bookscan
Deciles: Page Views Per Day - Correlate w/ BookScan Sales

- We group the 685 titles into Daily Page View Deciles, e.g.,:
  - Decile 1 contains the 10% of titles with the most page views / day
  - Decile 2 contains the next 10% of titles based on page views / day
- We calculate the percentage of BookScan unit sales attributed to each Decile
- Titles in Top page view Decile account for 26% of BookScan unit sales; Titles in the Bottom page view Decile account for 3% of BookScan unit sales.
  - Titles in the Top 2 Deciles account for 41% of all Bookscan unit sales
Year of Publication

- Older titles receive a disproportionate share of page views. 2004/2005 titles represent 67% of unit sales but only about 40% of page views. 2000 or earlier titles represent 12% of page views, but only 4% of unit sales.
Series Analysis

- 7 major series account for 43% of all titles, 52% of all page views in GBS, and 57% of our Bookscan units.
- Reference-oriented series have a disproportionate share of GBS views relative to their Bookscan units.

![Q4-2005: Titles Available in Google Print](chart.png)
Series Type

- 4 Major Series Types, accounting for 64% of all titles, 72% of all page views, and 81% of all Bookscan units.
  - "reference" titles have a higher share of page views than Bookscan units.

Q4-2005: Titles Available in Google Print
Key Questions for Authors & Publishers

• Does search help or hurt book sales?
• Are there differences in access depending on the type of book?
• For reference works, where search competes with the book, what opportunities are there for significant revenue?
• How will publishing itself change as it moves online?
Visions of the Future
The Internet As Platform

web2.0
Information Businesses
Software As a Service
Harnessing Collective Intelligence
1. Users Add Value

The key to competitive advantage in internet applications is the extent to which users add their own data to that which you provide.
2. The Perpetual Beta

When devices and programs are connected to the internet, applications are no longer software artifacts, they are ongoing services.
Library
From Wikipedia, the free encyclopedia

Alternative meanings: Library (computer science), Library (electronics), Library (biology)

In the traditional sense of the word, a library is a collection of books and periodicals. It can refer to an individual's private collection, but more often it is a large collection that is funded and maintained by a city or institution. This collection is often used by people who choose not to, or cannot afford to, purchase an extensive collection themselves. However, with the collection or invention of media other than books for storing information, many libraries are now also repositories and access points for maps, prints or other artwork, microfilm, microfiche, audio tapes, CDs, LPs, video tapes and DVDs, and provide public facilities to access CD-ROM databases and the Internet.

Thus, modern libraries are increasingly being redefined as places to get unrestricted access to information in many formats and from many sources. More recently, libraries are understood as extending beyond the physical walls of a building, providing assistance in navigating and analyzing tremendous amounts of knowledge with a variety of digital tools.

Contents [hide]
1 Etymology of the word
2 History
3 Types of libraries
4 Description
5 Library use
6 Library management
7 Funding problems
8 Some famous libraries
9 See also
Rough Cuts

Get behind the scenes to stay ahead of the curve.

As the book is being written...

Get access to digital content

Main message

When you buy a Rough Cuts book:

- Get access to an evolving PDF manuscript
- Download and print it, or read it online
- Submit suggestions to help shape the final book
- Send comments directly to authors and editors
- Be a part of the new program from Safari Books Online

Any questions?

Find out more
3. Software Above the Level of a Single Device

The PC is no longer the only access device for internet applications, and applications that are limited to a single device are less valuable than those that are connected.
4. Data is the Next “Intel Inside”

Applications are increasingly data-driven.

*Therefore*: Owning a unique, hard-to-recreate source of data may lead to an Intel-style single-source competitive advantage.
5. A Platform Beats an Application Every Time

- Excel 1-2-3
- WordPerfect
- Netscape Navigator
Two Types of Platform

• One Ring to Rule Them All

• Small Pieces Loosely Joined
Open a Recent Item

Create New
- HTML
- ColdFusion
- PHP
- ASP VBScrip
- ASP.NET C#
- JavaScript

Create from Samples
- CSS Style Sheets
- Framesets
- Page Designs (CSS)
- Starter Pages
- Page Designs
A Web Services-Based Help System
Event to Content Correlation

---

**Essential SNMP**
by Douglas Mauro, Kevin Schmidt

- **Publisher:** O'Reilly
- **Pub Date:** July 2001
- **ISBN:** 0-596-00020-0
- **Pages:** 326

**Section: Polling & Thresholds**
SNMP gives you the ability to poll your devices regularly, collecting their management information. Furthermore, you can tell the NMS that there are certain thresholds that, if crossed, require some sort of action. **More...**

---

**Network Troubleshooting Tools**
by Joseph D. Sloan

- **Publisher:** O'Reilly
- **Pub Date:** August 2001
- **ISBN:** 0-596-00185-X
- **Pages:** 364

**Section: Overview of SNMP**
SNMP is a management protocol allowing a management program to communicate, configure, or control remote devices that have embedded SNMP agents. **More...**

---

**Description**

- SNMP data collection on interface 10.100.100.55 failed.
- SNMP data collection on interface 10.100.100.55 previously failed and has been restored.
- SNMP data collection on interface 10.100.100.55
“I’m an inventor. I became interested in long term trends because an invention has to make sense in the world in which it is finished, not the world in which it is started.”

-Ray Kurzweil
For More Information

- http://tim.oreilly.com
- http://radar.oreilly.com
- http://safari.oreilly.com
- http://safariku.oreilly.com