Library Workshop Series 1: Organizing Information

(Helen) Hong Xu
Feb. 22, 2017

Outline

- Data and Information
- Information explosion and information overload
- Organizing Information
- Organizing information in libraries
What Is “Data”?

- Data are facts and statistics collected together for reference or analysis.

- Data are raw, unorganized facts that need to be processed. Data can be something simple and seemingly random and useless until it is organized.

What is Information

- Information are processed, organized data, which are structured or presented in a given context.

- They are meaningful or useful.
Differences between Data & Information

- Computers need data. Humans need information.
- Data is a building block. Information gives meaning and context.

---

Data Organization

<table>
<thead>
<tr>
<th>Hierarchy of data</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Database</td>
<td>(Project database)</td>
</tr>
<tr>
<td>File</td>
<td>(Personnel file)</td>
</tr>
<tr>
<td>Records</td>
<td>(Record containing SSN, first and last name, hire date)</td>
</tr>
<tr>
<td>Fields</td>
<td>(Last name field)</td>
</tr>
<tr>
<td>Characters (bytes)</td>
<td>(Letter F in ASCII)</td>
</tr>
</tbody>
</table>

Examples of Data vs Information

Helen Xu, 8 Duke Ave. Kunshan, JS 215316
051236657207

Helen Xu
8 Duke Ave.
Kunshan Jiangsu 215316
(0512) 3665-7207

Information Explosion

210 billion email messages per day

>560 apps downloaded every second

60 hours of video uploaded every minute
1/4 million full-length Hollywood releases every week (3 billion views per day)

We store 92% of this new information

140 billion photos 108/ Month

2/3 of the world’s mobile data traffic will be video by 2015

By 2015, 1 million video minutes (equivalent to 874 days) will traverse the Internet every second

LAST YEAR’S MOBILE DATA TRAFFIC WAS 3X THE SIZE OF THE ENTIRE GLOBAL INTERNET IN 2009

...and the pace is accelerating

Source: Cisco IBSG, 2006–2011; Cisco VNI, June 2011; TechnoMetrix, Digital Group; IDC; The Economist, Apple; InformationWeek
Information Overload

- There are 1.2 zettabytes of data today with an estimated 35 zettabytes by 2020; People see more than 34 billion bits of information per day – an equivalent of 2 books a day online (https://blog.kissmetrics.com/information-explosion/)
- "It would take over 200,000 years to read all the Internet, allowing for 30 minutes per document" (Baldwin, D. (2001) Information Overload. London: South Bank University.)
- "In the last 30 years mankind has produced more information than in the previous 5,000." (Information Overload Causes Stress. (1997, March/April). Reuters Magazine. Available: Lexis Nexis Universe [4/28/98].)

Amounts of computer storage

- kilobyte (KB) = 1000 bytes
- megabyte (MB) = 1000 kilobytes
- gigabyte (GB) = 1000 megabytes
- terabyte (TB) = 1000 gigabytes
- petabyte (PB) = 1000 terabytes
- exabyte (EB) = 1000 petabytes
- zettabyte (ZB) = 1000 exabytes
- yottabyte (YB) = 1000 zettabytes
Here are the Facts!

- A search on Google on the term “information explosion” : about 3,800,000 results
- A search on Google on the term “information overload” produces about 3,810,000 results.
- A search on Google on the term “information organization” : about 1,150,000,000 results

(---accessed on 2/21/2017)
Five Racks or “LATCH” in Organizing Information

- Location
- Alphabet
- Time
- Category
- Hierarchy/Relation

How is WWW Organized and Accessed?

- Using markup language HyperText Marke Up Language
- Uniform Resource Locator (URL)
- Web browser
- Hypertext Transfer Protocol (HTTP)
- Access
  - Direct: URL
  - Indirect: search engines
How Many Books Published Each Year? (2013)

- United States: 304,912
- China: 256,000
- Russia: 101,981
- India: 90,000
- Japan: 82,589

(https://en.wikipedia.org/wiki/Books_published_per_country_per_year)

Top Ten Largest Libraries in the World

- Library of Congress: 162 million+ Catalogues size (number of items)
- British Library: 150 million+
- Library and Archives Canada: 54 million+
- New York Public Library: 53.1 million
- Russian State Library: 44.4 million
- ...
- National Diet Library: 35.4 million
- National Library of China: 35.1 million

How Books/Materials Are Organized in a Library

- Subject headings
- Classification systems
  - Hierarchical system
    - Library of Congress Classification
    - Dewey Decimal Classification
    - Chinese Library Classification (中国图书馆分类法)
  - Faceted classification
  - etc.
- Call numbers: manually assigned by librarians with professional training
Subject Headings

- A controlled vocabulary or thesaurus of subject headings applied to every item within collections
- Word(s), phrase, name of a region/person
- Formal/standard English
- Inverted form
- Subdivision

Library of Congress Subject Headings (LCSH)

- Developed, maintained, and revised by LC since 1898 (38th edition in 2016)
- 24,156 personal name headings (family, corporate, meeting, conference)
- 484 uniform titles
- 241,047 topical subject headings
- 60,839 geographic subject headings
- 5,000 new headings added each year
Library Classification Systems

- A library classification is a system by which library resources are arranged according to subject.
- Use a notational system that represents the order of topics in the classification and allows items to be stored in that order.
- Group related materials together, typically arranged in a hierarchical tree structure: Class-division-subdivision.
Library of Congress Classification

- Developed by the Library of Congress.
- Used by most research and academic libraries in the U.S. and several other countries.
- Invented by Herbert Putnam in 1897
- Driven by LC’s practical needs

LCC Classes

<table>
<thead>
<tr>
<th>Letter</th>
<th>Subject area (21)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>General Works</td>
</tr>
<tr>
<td>B</td>
<td>Philosophy, Psychology, and Religion</td>
</tr>
<tr>
<td>C</td>
<td>Auxiliary Sciences of History</td>
</tr>
<tr>
<td>D</td>
<td>World History and History of Europe, Asia, Africa, Australia, New Zealand, Etc.</td>
</tr>
<tr>
<td>E</td>
<td>History of Americas</td>
</tr>
<tr>
<td>F</td>
<td>History of America</td>
</tr>
<tr>
<td>G</td>
<td>Geography, Anthropology, and Recreation</td>
</tr>
<tr>
<td>H</td>
<td>Social Sciences</td>
</tr>
<tr>
<td>I</td>
<td>Political Science</td>
</tr>
<tr>
<td>J</td>
<td>Law</td>
</tr>
<tr>
<td>L</td>
<td>Education</td>
</tr>
<tr>
<td>M</td>
<td>Music and Books on Music</td>
</tr>
<tr>
<td>N</td>
<td>Fine Arts</td>
</tr>
<tr>
<td>P</td>
<td>Language and Literature</td>
</tr>
<tr>
<td>Q</td>
<td>Science</td>
</tr>
<tr>
<td>R</td>
<td>Medicine</td>
</tr>
<tr>
<td>S</td>
<td>Agriculture</td>
</tr>
<tr>
<td>T</td>
<td>Technology</td>
</tr>
<tr>
<td>U</td>
<td>Military Science</td>
</tr>
<tr>
<td>V</td>
<td>Naval Science</td>
</tr>
<tr>
<td>Z</td>
<td>Bibliography, Library Science, and Information Resources (General)</td>
</tr>
</tbody>
</table>
Dewey Decimal Classification

- The DDC is published by OCLC Online Computer Library Center, Inc. OCLC owns all copyright rights in the Dewey Decimal Classification, and licenses the system for a variety of uses.
- DDC assigns each book a number based on its subject matter.
- Organized all categories into three levels.
- 10 main classes, 100 divisions, and 1000 sections.

Ten Main Classes in DDC

000 – General works, Computer science and Information
100 – Philosophy and psychology
200 – Religion
300 – Social sciences
400 – Language
500 – Science
600 – Technology
700 – Arts & recreation
800 – Literature
900 – History & geography
Understanding Call Numbers

- A call number represents an item’s subject matter and indicates the location on the shelf where the item can be found.
- Call number = Classification number + Cutter number(s) (+ Publication date) (+ Copy) e.g.
  - CT1828.Y55 A3 2009
  - E183.8 .C5 W3416 2013 c.2
Call Numbers on Book Spines

Interpreting Call Number

Call number =

E183.8
.CS W3416
2013
2013 c.2

Class number +
Cutter number(s) +
Date
Copy
How to Read a Call Number

Read the first line in alphabetical order:
A, B, BF, C, D, L, LA, LB, LC, M, ML...

183.8

Read the second line as a whole number:
1, 2, 3, 45, 100, 101, 1000, 2000, 2430...

.C5

The third line is a combination of a letter and numbers. Read the letter alphabetically. Read the number as a decimal, eg:
.C5 = .5 .C724 = .724

Some call numbers have more than one combination letter-number line.

2013

The last line is the year the book was published. Read in chronological order:
1985, 1991, 1992...

c.2

That means we have at least 2 copies for this book.

Shelving Books According to Call Number
Cutter Number

- The Cutter number, or Cutter, is a combination of letters and numbers that follows the class number and is preceded by a decimal point.
- Designed by Charles Ammi Cutter (1837-1903)
- Arranging books by author within a given class
Q & A

Thanks!