Indexing and Abstracting

Introduction

Kuang-hua Chen
Department of Library and Information Science
National Taiwan University
khchen@ntu.edu.tw

Purpose

- Indexing and abstracting are practical arts.
- What is important is the information in those records founded, and the validity and completeness of that information still relies on the expertise of human indexers and abstractors.
- The purpose of this course is to present the fundamentals of indexing and abstracting as a foundation for entry-level professional practice.
The Need for Indexes

- The function of an index is to give users systematic and effective shortcuts to the information they need.
- Indexes are needed for any information collection, except the very smallest.
- An index has two general purposes
  - Minimize the time and effort
  - Maximize the searching success

Dual Spaces

- Document space
  - Original documents
- Surrogate space
  - Indexes
  - Abstracts
- Indexing and abstracting are to ensure surrogate space similar to document space
Indexing

Document space → Controlled terms → Free terms → Surrogate space

Abstracting

Document space → Controlled terms → Surrogate space
Nature of Indexes

- Come in many forms and formats
  - Printed
  - Electronic
  - Both
- Serve many purposes
  - Name indexes
  - Subject indexes
  - Map indexes
  - Artifacts indexes
  - Address indexes

Nature of Indexes (continued)

- Surrogates of different materials
  - Books
  - Periodicals
  - Images
  - Database
  - Numerical data
- Compiled by
  - Human being
  - Computers
  - Both
Makers of Indexes

- Being a good, fast reader is also a real plus for an indexer
- The good indexer becomes the surrogate of the reader
- Makers of indexes are also censors
  - Unrecognized bias
  - Prejudice
  - personal agendas

Joseph Goebbels

Source: http://www.us-israel.org/jsource/Holocaust/goebbels.html
Related Books

- 1878 - Henry Wheately
  *What Is an Index*

- 1975 and 1976 – Alan G. Brown
  Volume 1 –
  *Introduction to Subject Indexing: A Programmed Text. Subject Analysis and Practical Classification.*
  Volume 2 –
  *Introduction to Subject Indexing: A Programmed Text. UDC and Chain Procedure in Subject Cataloguing.*

Related Books (continued)

- 1975 – Harold Broko and Charles Bernier
  *Abstracting Concepts and Methods*

- 1978 - Harold Broko and Charles Bernier
  *Indexing Concepts and Methods*

- 1979 - G. Norman Knight
  *A Guide to the Indexing of Books and Periodicals*

  - Eugene Garfield
    *Citation Indexing-Its Theory and Application in Science, Technology and Humanities*
Related Books (continued)

- 1979 – Jennifer E. Rowley
  *Abstracting and Indexing*

- 1983 – Hilda Feinberg
  *Indexing Specialized Formats and Subjects*

- 1986 – Timothy C. Craven
  *String Indexing*

- 1991 – F. W. Lancaster
  *Indexing and Abstracting in Theory and Practice*

Indexing -- Historical Background

- Bible
- Alphabetic order
  - Witty(1976)
    - Although we have in Graeco-Roman times the use of alphabetic order of certain non-fiction works, we do not...have anything like an alphabetic index to a work before the Middle Ages.
  - 12th century:
    - social institutions
    - intellectual discourse
  - 14th century: annotation
    - Manuscripts
    - Library catalogs bibliographies
Indexing -- Historical Background (continued)

- Blank space: do-it-yourself index
- Early indexes
  - Concordances indexes
  - Topical subject indexes
- 18th century: indexes to books were prepared primarily by their authors
- 19th century: compile indexes to cover entire fields
- 20th century: subject index consisted of a list of terms with no subdivisions or with subdivision in a simple hierarchical structure of two or three levels (main heading and modifier or main heading, subheading, and modifier)

Journal Indexing -- Historical Background

- Scholarly journals
  - with the rapid growth of journals, subject access become necessary
- Late 19th century - taxonomy
- 1850s, W. F. Poole
  - publication of an index across many journals
- 1882 – ALA
  - Index to Periodical Literature
- 1892 – Paul Otlet, Henri La Fontaine
  - International Institute of Bibliography
  - Title – word indexing
Journal Indexing -- Historical Background (continued)

- 1900 - H. W. Wilson
  - *Readers' Guide to Periodical Literature*
- “Information explosion” during and following World War II
- 1950s –
  - Computers had entered the indexing and abstracting arena and efforts were begun to evaluate indexing using quantitative methods.
  - Hans Peter Luhn - KWIC indexes

Journal Indexing -- Historical Background (continued)

- Cranfield Project
  - 2 stages: 1957, 1963
  - C.W. Cleverdon
  - Major values of the Cranfield Project was the development of techniques for the quantitative evaluation of indexing
- 20th century –
  - Indexing has outstripped the ideas of previous ages. We have progressed from indexes to individual works, through indexes to several volumes, to cooperative indexes on an international scale
Abstracting - Historical Background

- 1665 – *Journal des Scavans*
- 1700s – other abstracts began to appear
  - 1703: *Monatsextracte*
- 19th century – abstract journal proliferate

How to Measure Information

- A specific event $x_i$
  - $-\log(p(x_i))$
- A random variable $X$ represents possible events
  - Entropy of $X$ is $- \sum p(x_i) \log(p(x_i))$