Ad Hoc, User-Determined Classified Displays Based on Faceted Indexing

James D. Anderson

School of Communication, Information, and Library Studies
Rutgers University, New Brunswick, NJ 08903, USA

SUMMARY

End user specification of classified arrays of text surrogates for browsing electronic text databases or for organizing retrieved records.

CONVENTIONAL INDEXING

Indexing usually refers to extraction or assignment of terms to describe features of texts. These terms name classes to which a text belongs. These terms (class names) are typically displayed or accessed alphanumerically, although they are sometimes complemented with displays of term hierarchies and other conceptual relationships among terms via thesauri or internal cross references. The alphabetic displays of ERIC databases (Figure 1), for example, are supplemented by the Thesaurus of ERIC Descriptors (Figures 2-A and 2-B), while indexes published by the H. W. Wilson Company (Figure 3) and most library catalogs (Figure 4) display conceptual relationships in cross references (normally restricted to standard thesaural relations of broader-than, narrower-than, associative, and equivalent, although all but the last are sometimes collapsed into a single category).

Classification usually refers to display of terms (class names) in some conceptual order based on relations among classes (Figure 5), as opposed to the arbitrary order of the alphabet. Large classified displays are often accompanied by an alphabetical index to particular classes (Figure 6). Libraries, for example, display their books in classified arrays using class notation to represent class names, and catalogs function as alphabetic indexes to these classified displays (LC and DDC class notations in Figure 4). Similarly, the main section of the printed version of the MLA International Bibliography (Modern Language Association) is displayed in classified order (classes are specific literatures / periods) (Figure 7), with alphabetic indexes for authors and subjects (Figure 8).

Homophobia
Teaching about Gay and Lesbian Sexual and Affectional Orientation Using Explicit Films to Reduce Homophobia. *Journal of Humanistic Education and Development;* v28 m1 p18-34 Sep 1989 EJ 402 275

Homosexuality
Teaching about Gay and Lesbian Sexual and Affectional Orientation Using Explicit Films to Reduce Homophobia. *Journal of Humanistic Education and Development;* v28 m1 p18-34 Sep 1989 EJ 402 275

Figure 1. Alphabetic Display in ERIC Current Index to Journals in Education.
HOMOSEXUALITY
SN Sexual attraction and/or intercourse between members of the same sex (note: use a more specific term if possible)
NT Lesbianism
BT Sexuality

Figure 2-A. Narrower-than (NT) / Broader-than (BT) in Alphabetic Descriptor Display in *Thesaurus of ERIC Descriptors*.

:SEXUALITY
HOMOSEXUALITY
.LESBIANISM

Figure 2-B. Two-Way Hierarchical Term Display in *Thesaurus of ERIC Descriptors*.

Homophobia See Attitudes toward homosexuality
Homophones See English language—Homonyms
Homosexuality
See also
Attitudes toward homosexuality
Gay studies
Lesbianism
The campus environment for gay and lesbian life. J.
D’Emilio. il *Academe* 76:16-19 Ja/F ’90

Figure 3. Cross references in Wilson *Education Index*.

HOMOSEXUAL LIBERATION MOVEMENT
see Gay Liberation Movement

HOMOSEXUALITY
see also Bisexuality; Church work with homosexuals; Gay Liberation Movement; Lesbianism

My father & myself. [1st American ed] New York,
Coward-McCann [1969, 1968]

HQ76.A27 1969 301.41’5 69-16083 MARC

Figure 4. Subject Catalog Display with Cross References and Entry with LC (Library of Congress) and DDC (Dewey Decimal Classification) Notation, in *LC Catalog—Books: Subjects*. 
Institutions pertaining to relations of the sexes
Class here interdisciplinary works on sex (formerly 612.6), sexual love, sexual relations
Unwed parenthood relocated to 306.856
For sexual ethics, see 176; problems and controversies concerning various sex relations, 363.4; sex offenses, 364.153; sex customs 392.6; sex hygiene, 613.95; sex techniques, 613.96; sex practices viewed as treatable disorders, 616.858
See Manual at 155.34 vs. 306.7

General institutions
Class here dating behavior
Group sex relocated to 306.77
For marriage, see 306.85

Celibacy

Courtship
See also 392.4 for customs of courtship

Cohabiting
Former heading: Nonmarital relations
Including ménage à trois, free love

Extramarital relations

Illegitimacy
Relocated to 306.874

Homosexual marriage

Prostitution
See also 331.76130674 for prostitution as an occupation, 363.44 for prostitution as a social problem, 364.1535 for prostitution as a crime

By females
Class child prostitution in 306.745

By males
Class child prostitution in 306.745

By children

Sexual orientation
Class practices associated with specific orientations in 306.77

Neutral sexual orientation
Heterosexuality
Bisexuality
Homosexuality
Gay liberation movement relocated to 305.90664

Male homosexuality
Female homosexuality (Lesbianism)

Figure 5. Classified Array in Dewey Decimal Classification (Part of Schedule 306).
Homosexual marriage 306.738
Homosexuality 306.766
ethics 176
religion 291.566
Buddhism 294.356 6
Christianity 241.66
Hinduism 294.548 66
Islam 297.5
Judaism 296.385 66
literature 808.803 53
history and criticism 809.933 53
specific literatures T3B—080 353
history and criticism T3B—080 353
medicine 616.858 34
social problems 363.49
social theology 291.178.357 66
Christianity 261.835 766
Homosexuals 305.906 64
T1—086 64
female 305.489 664
T1—086 643
male 305.389 664
T1—086 642
pastoral theology 259.086 64

Figure 6. Alphabetic Index to Dewey Decimal Classification.

American literature/1900-1999

HEILBRUN, CAROLYN G. (1926- )

Fiction


Figure 7. Classified Display in MLA International Bibliography (Class is American literature/1900-1999).

HOMOSEXUALITY

Used for: Homeroeticism.
See also related term: Homosexual.

Heilbrun, Carolyn G. Characterization; plot. Treatment of academia; HOMOSEXUALITY; androgyny. I:8146.

Figure 8. Subject Index Display for MLA International Bibliography.
Unlike conventional indexing, which can only produce alphabetical (or alphanumeric) displays of terms, faceted indexing includes the basic building blocks of classification, and therefore can provide classified, as well as alphabetical (alphanumeric) displays. Faceted indexing also tends to be more "request-oriented" than conventional indexing. Facets can be used to describe the categories of queries that users are expected to have and, therefore, for which indexers should seek answers. In the words of Soergel [1985, p. 397], "A properly designed facet frame captures the essential conceptual structure of a field and is instrumental in eliciting the concepts to be included in the index language, in assisting in the analysis of a search topic, and in the analysis of an entity in indexing." A list of facets serves as a questionnaire, reminding an indexer what users are expected to seek. One of my favorite examples is Ranganathan's list of facets for a database on diesel engines; indexers were to seek answers for questions about manufacturer or brand, country of manufacture, purpose, environment (i.e., for engines designed for under-water use, humid climates, deserts, etc.), costs, compression ratios, cycle of strokes, size, number and arrangement of cylinders, bore diameter, etc., etc. [Ranganathan, 1965, p. 221-222].

Faceted indexing can be illustrated by the database of the Modern Language Association. Descriptors assigned to documents are tagged by their facets at the time of indexing. For example, descriptors describing texts about literature belong to one of the following facets:

- specific literatures
- performance media
- languages
- periods
- individuals (e.g., authors)
- groups/movements
- genres
- works
- features
- literary techniques
- themes/motifs/figures/characters
- influences (recipients)
- sources
- processes
- types of scholarship
- methodological approaches
- theories
- devices/tools
- disciplines
- scholars
- document types

**FACETS AND HIERARCHIES**

Concept classes may be created either for single or multiple concepts. For example, a multiple-concept class, such as French poetry (combining French from the language and/or nationality facet with poetry from the genre facet) would be created if the single concept classes from which it is derived would be too general, e.g., encompass too many documents, to be useful by themselves. As this example illustrates, creation of compound concept classes depends on combinations of classes from different facets and distinctions among classes within a facet — e.g., combining French and poetry on the one hand, and hierarchical distinctions among French, Romance languages, Indo-European languages, and sibling relations among French, Spanish, and Italian on the other.

Within facets, single concept classes may be displayed alphanumerically on the basis of names assigned to classes, chronologically on the basis of some stated or implied date, or hierarchically based on such relations as genus/species or whole/part.
USER ORDERED CLASSIFICATION DISPLAYS

Facets within domains or compound classes may be arranged in any order, resulting in fundamentally different displays. If a domain is described on the basis of, say, five facets (most domains, such as diesel engines, need many, many more), there are at once 120 different display orders possible, ignoring options within facets (5x4x3x2x1=120). Seven facets, not a large number, would result in 5,040 possible classified display arrangements. Any facet can serve as the primary facet. Each primary facet can be subdivided by each of the remaining facets, and so on until all facets are displayed. Thus, we can display themes treated in literature, subdivided by nationality or language or genre or period; or genres, subdivided by themes or periods or nationality or language, and so on. The following are two such displays:

Nationality, eg., English literature  
  Period, e.g., 20th Century  
  Genre, e.g., Novel  
  Theme, e.g., Homosexuality  
  Writer, e.g., Forster, E. M.  
  Work, e.g., Maurice  
  Approach, e.g., Psychological analysis

Theme, e.g., Homosexuality  
  Genre, e.g., Novel  
  Period, e.g., 20th Century  
  Nationality, eg., English literature  
  Writer, e.g., Forster, E. M.  
  Work, e.g., Maurice  
  Approach, e.g., Psychological analysis

IMPLEMENTATION AND RESEARCH

To take full advantage of faceted indexing, a searcher should be able to request a classified display of surrogate records based on any or all facets used in an indexing scheme. As seen by the MLA example shown earlier, the required data (facet plus descriptor) exist in at least one publicly-available database. However, implementation of this possibility depends also on the capabilities of search or front-end systems, and the necessary procedures have not yet been provided by the major database vendors. Actual implementation is not at all complicated. At Rutgers, we have this capability in our small, simple text database management system used in instructional labs ("Information Organization based on Textual Analysis: IOTA," [Anderson, 1987] — not to be confused with the very different French IOTA information retrieval system! [Chiaramella and Defude, 1987]).

Designers and creators of new databases should seriously consider the possible benefits of faceted indexing, not only in terms of the variety of potential displays available for print or on-screen browsing, searching, and viewing, but also in terms of the contribution of a facet frame to increased client/request orientation.

Vendors who provide search systems for existing databases should consider the advantages of making classified displays available in the context of the growing desire for more varied...
approaches to information, beyond the present Boolean search.

R&D should explore the idea of applying faceted indexing "after the fact" to previously indexed databases. This may be possible if a thesaurus is available and terms are grouped into hierarchies. Each hierarchy could be treated as a facet (indeed, in many cases that's exactly what they are). Every term in the database could be tagged with a hierarchy indicator perhaps at the time of retrieval, and this information could be used to create classified displays.

In any case, if any of this makes sense, we need to investigate whether faceted indexing is easier or harder to do than "conventional" indexing, whether it results in more effective indexing (ie, more effective retrieval), and whether the kinds of classified displays that faceted indexing makes possible are useful in the overall search process, as we work to make flexible navigation, scanning, and browsing real options for online retrieval.

**BRIEF BIBLIOGRAPHY**


