

Hierarchical Gaps and Subject Authority Control Processing : an Assessment

Daniel CannCasciato
Head of Cataloging

MaryWise
Catalog Librarian

Central Washington University Library
Ellensburg WA 98926-7548

Introduction:

Subject authority control procedures vary widely from library to library. Some processes are manual, they have varying staffing levels, and many are out-sourced and automated. There are two main goals behind the application of these procedures. The first is to standardize the terminology which is used in bibliographic records. This facilitates patron access to the collections and is significant to the fulfillment of the second of Cutter's objectives—to allow the patron to find what a library has by subject. The secondary effect of a subject authority control procedure is to help guide a patron from general, broad terminology to more specific, narrower headings which he or she might need. This need for records from the upper hierarchy is not unique to subject headings. Name headings for subordinate bodies exhibit the same requirement, although in what is generally a more contained environment. Name headings do not need to fit within any hierarchy other than their own, while subject headings are part of an overall schema. The name authority record (NAR) for a subordinate body, (e.g., *International Business Machines Corporation. Federal Systems Division*), would require that the NAR for the parent body (*International Business Machines Corporation*) also be included in the catalog. For topical subject headings, however, this can become significantly more complex.

The complexity of hierarchical references can be demonstrated in a variety of ways. For example, to guide a patron to the species *Venturia inaequalis* (the fungal cause of Apple Scab) requires a long chain of linkages in terminology. A gap in that linkage compromises the effectiveness of the guidance we provide for our patrons. In Central Washington University Library's catalog the use of the term *Venturia inaequalis* on a bibliographic record required the addition of five subject authority records in addition to the one for *Venturia inaequalis* to provide a full hierarchical guide. The full hierarchical path beginning at *Fungi* is shown in Figure 1. Items in bold are those for which the subject terminology is not on a bibliographic record. The authority records for those terms were added to the catalog to create a full reference structure for a patron to follow. Thus, a patron could enter a search at any point in the hierarchy and be led to the lone item we have that might be of interest.

Figure 1

Fungi
Ascomycetes
Loculoascomycetes
Pleosporales
Venturiaceae
Venturia
Venturia inaequalis

A simpler example is that of the species *Chinese mitten crab (Eriocheir sinensis)*. Our catalog lacked the authority record for the immediately higher term, the genus, *Eriocheir*. Without the addition of that intervening terminology, a patron would not get from the family, *Grapsidae*, to the progressively narrower terms.

In this study we examined nine months of reports of new topical or geographic subject headings in our catalog. We downloaded new authority records when needed. We then examined those newly downloaded records to determine if they needed supporting authority records based on the broader terms from those new headings.

Literature Review

Very little is written on the actual subject control process. Most of the literature available discusses the fact that authority control is beneficial to patrons. “The true purpose of authority control should be to help the user move effortlessly from his/her terminology (natural language) to the terms in use (controlled vocabulary) of the system to locate all materials (objects) that are relevant ...” (Micco, 1996: 1). A problem in academic libraries is subject authority control. Dalrymple and Younger (1991) report that there must be informed feedback for users when performing subject searches.

Some literature relates to specific online catalog systems. Krieger (1990) explains that searching by subject in the Dynix system is somewhat cumbersome because although the first search screen brings up a see or use for reference, related terms, broad terms, and narrower terms must be found by performing a related term command search. In a comparison between two systems, one with authority control and one without, the authors discovered that users performed subject searches more often than title searches (Wilkes, Nelson, 1995). The authors point out that controlled vocabulary is a problem because if users did not know the correct LCSH term, they would often not get results from their search in the system without authority control. The authors state that if users had consulted the print version of LCSH (kept near the terminals) they would have been able to determine broader terms, narrower terms, related terms, “use fors,” etc. Since many users do not understand the idea of controlled vocabulary, it is more useful to have these references built in to the online catalog system. Chan and Vizine-Goetz (1998) discuss the feasibility of automatically generating a subject validation file from OCLC. They do not, however, discuss hierarchical issues.

Micco (1996) explains "... the user should be able to use a hierarchical classification to enter the system at the desired level of specificity in the topic of their choice with the option of broadening or narrowing a search that is not successful." (2). She continues by stating that currently (as of 1996) that option is not available and a hierarchical system needs to be implemented for that to occur.

Clack (1990) addresses very briefly the need for hierarchical records in the catalog even when the term has not been used as a subject heading on a bibliographic record. No further detail is given on this topic. Ludy (1985) also briefly describes the benefit of having the authority record for the broader term, even though that term does not appear in a bibliographic record. Michael Gorman (2002) refers to "proceeding from the general to the specific and following the syndetic structures of bibliographic control" as a characteristic of a good librarian (11).

The general concept of subject authority control, and authority control in general, is well covered. Its benefits are noted and recommended. Overall, though, in none of the literature we reviewed have we found a discussion of the frequency of the occurrence of hierarchical gaps in the authority control process—and thus in the catalog—nor of the possible impacts of those gaps.

Methodology

For this study we have reviewed nine months of subject printouts from an Innovative Interfaces, Inc. integrated library system. We looked for new topical and geographic headings only. We omitted names and titles used as subjects. We also omitted most music headings (those covering arrangements, instrumentation, etc.) from this study, as their processing is handled by a separate workflow.

This review resulted in the identification of 331 new topical headings to Cattrax, the Central Washington University Library catalog. For those new headings, we downloaded the appropriate authority records from OCLC . We subsequently searched each of the broader terms for those topical headings in the catalog, identifying those new broader headings that were not represented by an authority record of their own. Those terms that required an authority record were then searched in OCLC and the records downloaded to the catalog. The process was repeated until there were no further authority records needing downloading into the catalog. We searched both explicit and implicit broader terms.

An explicit broader term is one identified on the subject authority record itself as MARC tag 550 (or 551) with a \$w g. It is also known as a broader term, or BT. Implicit broader terms are those of the type for the heading *Pacific Gulf Yupik women*. This topical heading requires the subject authority record for *Pacific Gulf Yupik Eskimos*, because it is only on that authority that references from variant terminology reside. Without the authority record for this implicit broader term, references from the variants *Alutiiq Eskimos*, *South Alaska Eskimos*, *Sugpiak Eskimos*, *Suqpiaq Eskimos*, etc., would not be available for our patrons. Similarly, for river valley headings and watersheds the river itself was searched.

Broader terms created for the specific purpose of filling a reference hierarchy, while searched in OCLC, were not downloaded to the local catalog if references were not needed. In general, these types of headings take the form [Topic] \$x [Sub-topic] [Topic] \$z [Geographic subdivision]. For example, the new heading of *Aguarico River Valley (Ecuador)* had the explicit broader term of *Rivers \$z Ecuador*. The explicit broader term was searched in the utility, but as no broader terms existed on the authority record for that formulated heading, the authority record itself for *Rivers \$z Ecuador* was not downloaded into the local catalog. Depending on the limitations of the catalog software being used by a library, this practice would vary.

Findings

Many of the broader terms for the new headings were already in the database. Of the 331 new headings first downloaded, 42 subsequent new authority records (13.5%) were added. Of those 42, seven required new authority records. Of those seven, five new records were added. Of those five, two were added; and finally, of those two, one new authority record was added to the database. In all, 60 supplemental authority records were added to the database, which was an extra 17% beyond the original new headings list of 331. Without those gaps in the hierarchy being filled, it would have been impossible for a user to get from *Life scientists* to *Women marine mammalogists*, for example.

Contrary to expectations, the new headings were not overwhelmingly scientific in nature (see Appendix). They did include scientific terms, but also included terminology related to agriculture, religion, technology, relationships, geographic locations, culture, and literature.

Conclusion

Based on this study, there are significant numbers of new subjects whose authority records need additional authority records downloaded to fill in the reference hierarchy. There was no predictive subject area in which needed broader term headings are easily identified. Some local library systems' report capabilities might be able to identify subject headings with broader terms not represented in the catalog by bibliographic records. Many will not be able to do so. The Central Washington University Library catalog system would not create a report of headings for which the broader term is not also used on a bibliographic record or not supported by an authority record above it in the hierarchy. It will report if the heading (MARC 1XX field) from an authority record is not found also on a bibliographic record.

Subject authority control procedures have been sporadic at Central Washington University Library. Currency in maintenance of this process could have an impact on the percentage of items needing the additional broader term authority records.

This study has shown, at least preliminarily, that in subject authority control procedures there is a need to search not only the subject authority record for the new heading itself, but also for those headings that support hierarchical searching and the syndetic structure.

Although it may seem time consuming to check the new heading for hierarchical gaps, doing so will provide the links a user needs to find material on a subject he or she is researching. It may save time in the long run, by providing online guidance to users who do not understand and will not learn about hierarchical relationships in subjects.

One of the primary goals of the catalog is to facilitate the use of the collection by the patron. A fully employed hierarchy must be available to the patron, to assist in the retrieval of pertinent materials.

Bibliography

Chan, Lois Mai and Vizine-Goetz, Diane. (1998) "Toward a Computer-generated Subject Validation File: Feasibility and Usefulness." *Library Resources & Technical Services* (42:1: January 1998): 45-60.

Clack, Doris Hargrett. (1990). *Authority Control: Principles, Applications, and Instructions*. Chicago: American Library Association. 174.

Dalrymple, Prudence W. and Younger, Jennifer A. (1991) "From Authority Control to Informed Retrieval: Framing the Expanded Domain of Subject Access." *College & Research Libraries* (52:2: March 1991):139-149.

Gorman, Michael. (2002) "Why Teach Cataloging and Classification?" *Cataloging and Classification Quarterly* (34:1&2):1-13.

Krieger, Michael T. (1990) "Subject Authority Control on the Dynix System." *Cataloging & Classification Quarterly* (12:1):103-115.

Ludy, Lorene E. (1985) "OSU Libraries' Use of Library of Congress Subject Authority File." *Information Technology and Libraries* (4:1): March 1985:155-160.

Micco, Mary. (1996) "Subject Authority Control in the World of Internet." *LIBRES: Library and Information Science Research*, (6:3)September 1996: 7 pages.

Wilkes, Adeline and Nelson, Antoinette.(1995) "Subject Searching in Two Online Catalogs: Authority Control vs. Non-authority Control." *Cataloging & Classification Quarterly* (20:4)57-79.

Appendix. Downloaded Subject Headings

Terms in **bold face** required broader term (BT) subject authority records be downloaded.
Diacritics are omitted.

First Downloaded Subject Authority Records

Abkhazia (Georgia)

Abu Hureyra, Tall (Syria)

Account books

African American women executives

African American women political activists

Agricultural consultants

Aguarico River (Ecuador)

Aguarico River Valley (Ecuador)

Al-Aqsa Intifada, 2000

Alcohol ignition interlock devices

Alternate takes (Sound recordings)

Alternative rock music

Amended tax returns

American beech

Amish children

Amish Country (Pa.)

Anger in literature

Anti-rape movement

Applied sociology

Archives and national socialism

AskA services

Asparagus-rust

Athapascan mythology

Audited financial statements

Aufeis

Aviano Air Base (Italy)

Aviation medical examiners

Billiard players

Binary principle (Linguistics)

Biographical television program

Black vine weevil

Blacks in motion pictures

Blacks in the motion picture industry

Blood in literature

Boxing films

Brassieres

Broomcorn millet

Butterfly gardening

Butterfly gardens

C# (Computer program language)

Cacao growers

Cascade-Siskiyou National Monument (Or.)

Cenchrus ciliaris

Children of Nazis

Children's poetry, Hispanic American (Spanish)

Christmas in motion pictures

Chukotskii avtonomnyi okrug (Russia)

Chumash beadwork

Coal Miners' Strike, Harlan County, Ky., 1973

Computational intelligence

Computer fonts

Computer sex

Condolence Ceremony (Iroquois rite)

Construction and demolition debris

Construction, Military

Contextualism (Philosophy)

Coordinate measuring machines

Cottonseed products as feed

Craftsman Farms (Parsippany, N.J.)

Criminal syndicalism

Cruise missile defenses

Crusader art

Curves, Quartic

Cyanobacterial toxins

Dampness in basements

Death in popular culture

Decorative arts, Islamic

Decorative arts, Renaissance

Depression in women

Digital signatures

Disorderly conduct

DNA, Fossil

Door-to-door selling

Dots-and-boxes (Game)

Dracula films

Duwamish Indians

DVD-Video discs

Ecological assessment (Biology)

Ecstasy (Drug)

Eightmile River Wild and Scenic Rover (Conn.)

Electron beam curing

Electronic reference services (Libraries)

Embryonic stem cells

ERIC (Information retrieval system)

Eritrean-Ethiopian War, 1998

Erotic prints, Japanese

Etton Site (England)

Europa (Satellite)

Expectation (Psychology) in children

Ex-welfare recipients

Federal aid to housing

Female sex offenders

Feminist criminology

Fengate Site (England)

Flood insurance claims

Fluid-structure interaction

Folk dancing, Romanian

Folk songs, Czech

Food in popular culture

Football team owners

Formation damage (Petroleum engineering)

Fort Reno (Okla.)

Fort Stewart (Ga.)

Four Horsemen of the Apocalypse

Frazil ice

Freak shows

Furniture, Mission

Fuzzy decision making

Galway (Ireland)

Game and game-bird stocking

Gangsters

Gas air conditioning

Genre painting, French

Geophysicists

Gordon, Flash (Fictitious character)

Granddaughters

Graphic calculators

Grave goods

Grierson's Cavalry Raid, 1863

Guantanamo Bay Naval Base (Cuba)

Gypsophila paniculata

Haying equipment

Heating-pipes

Heterosexuality in art

Highway-railroad grade crossing

Hispanic American teenage girls

Hopi magic

Hopi sculpture

Horror in music

Hosokawa family

Humanitarian assistance, Europe

Hunkpapa Indians

Hurricane Mitch, 1998

Identity (Philosophical concept)

Income tax deductions for moving expenses

Indian activists

Indian civic leaders

Indian cordage

Indian women healers

Individualized education program

Information theory in aesthetic

Integrated agricultural systems

Integrated solid waste management

Interfaith families

International District (Seattle, Wash.)

Internet in public administration

Internet pornography

Internet Relay Chat

Internet users

Interstate 5

Irony in art criticism

Jai alai \$x Betting

Jari River (Amapa and Para, Brazil)

Jari River Valley (Amapa and Para, Brazil)

JPEG (Image coding standard)

Jukebox industry

Kepeh, David (Fictitious character)

Kerosene industry

Koryaks

Land mines (International law)

Landslide dams

Lecanium

Lentil industry

Lesbianism in motion pictures

Lettres de cachet

Liability for flood damages

Libraries and distance education

Linaria vulgaris

Log buildings

Loneliness in old age

Lower Rio Grande Valley National Wildlife Refuge (Tex.)

Male elementary school teachers

Mam Indians

Manic-depressive illness in adolescence

Manitowish River (Wis.)

Mantle plumes

Manuscripts, Nahuatl

Marshall (Minn.)

Masculinity in art

Massage therapy

Mathematics on postage stamps

Maturin, Stephen (Fictitious character)

Mechanics' liens

Media consultants

Medicare beneficiaries

Meteorological services

Mexican American labor union members

Microbial pesticides

Microsoft .NET Framework

Milk fever in animals

Minorities in mathematics

Minority women educators

Mixed radioactive wastes

Mobile computing

Molecular diagnosis

Monroe (N.C.)

Motion picture editors

Motion pictures, Brazilian

Motion pictures, British

Motion pictures, French

Multiculturalism in art

Mutton industry

Nanjing (Jiangsu Sheng, China)

Nanking Massacre, Nanjing, Jiangsu, China, 1937

National characteristics, Korean

Navajo calendar

Naxi (Chinese people)

Network File System (Computer network protocol)

New Age consumers

Nisqually River (Wash.)

Non-church-affiliated people

Northern sky (Astronomy)

Oakley Plantation (La.)

Oil as pesticide

Oklahoma Panhandle (Okla.)

One-person shows (Performing arts)

Onion thrips

Online stockbrokers

Operation Allied Force, 1999

Operation Hardtack, 1958

Oregon silverspot butterfly

Overweight children

Pacific Gulf Yupik women

Packets

Palestinian American journalist

Payette River (Idaho)

Pediatric emergency services

Pennsylvania Dutch Country (Pa.)

Per diem allowances

Persian Gulf syndrome

Personal communication service

Philosophy, Igbo

Philosophy, Yoruba

Photographic historians

Pitching (Softball)

Planetary volcanism

Pocket computers

Police films

Police in mass media

Pollinators

Portrait painting, Chinese

Possessiveness in literature

Poverty Point culture

Poverty Point State Park (La.)

Predatory insects

Prince Rupert Harbour (B.C.)

Problem-based learning

Punctuated equilibrium (Evolution)

Puppets in literature

Purdah

Python (Computer program language)

Qualitative reasoning

Racism in popular culture

Rajasthani poetry

Random graphs

Red hake

Red River Parish (La.)

Refugees, Hutu

Refugees, Palestinian Arab

Religious fiction, American

Report cards

Rhabdocline pseudotsugae

Road films

Safety harness (Aeronautics)

Saint John River (Me. and N.B.)

Scanning force microscopy

Schubert varieties

Seahenge Site (England)

Securities commissions

Seismites

Self-injurious behavior

Selkies

Semantics, Historical

Septoria diseases

Shaft, John (Fictitious character)

Sharpeville Massacre, Sharpeville, South Africa, 1960

Sheer fabrics

Shell beads

Shimura varieties

Sinixt Indians

Sitona lineata

Slavery and Judaism

Smart materials

Soho (London, England)

Soil scientists

Somerset Place (N.C.)

South Fork Dam (Pa.)

Southern Paiute Indians

Spaghetti Westerns

Sports films

Spotted alfalfa aphid

Squalene

Stable isotopes

Stagonospora diseases

Steppe ecology

Strawberry-weevil

Stromateidae

Swarm intelligence

Swashbuckler films

Symphyla

Tactile sensors

Tardiness

Taxpayer account numbers

Technology in art

Telephone credit cards

Television and theater

Tort liability of managed care

Townshend (Vt.)

Transparency in government

Trials (Terrorism)

Triangles (Interpersonal relations)

Triticale

Trout industry

Tunneling spectroscopy

Underground ecology

Underpayment of taxes

Urban pollution

Values in adolescence

Verismo (Opera)

Vesara style (Architecture)

Video games and children

Visual programming languages (Computer science)

Volunteer workers in environmental education

Warfare, Prehistoric

Wellhead protection

Westover Air Force Base (Mass.)

Wexford, Inspector (Fictitious character)

Whirligigs

Whitestone Lake (Wash.)

Wind power plants

Wind River (Wash.)

Wind River Watershed (Wash.)

Wireless Internet

Wit and humor in art

Wizard of Oz (Fictitious character)

Women archaeologists

Women coaches (Athletics)

Women marine mammalogists

Women, !Kung

Young adult consumers

2nd Round of Downloaded Subjects

These were all BT terms (explicit or implicit) on the previous list

Alkenes

Architecture, Hindu

Beech

Beneficiaries

Biomolecules, Fossil

Blank-books

Butterfly attracting

Calcium

Cenchrus

Christian children

Computer users

Distributed artificial intelligence

DVDs

Fuzzy mathematics

Gypsophila

Hispanic American poetry (Spanish)

Hispanic American teenagers

Historical lexicology

Humanitarian assistance

Image compression

Legume industry

Linaria

Marine mammalogists

Microsoft .NET

Moving, Household

Music trade

Non-monogamous relationships

Palestinian Americans

Pencil games

Pipe

Portable computers

Psychoses in adolescence

Rajasthani literature

Realism in opera

Referral centers (Information services)

Religious fiction

Rhabdocline

Rye

Sideshowes

Sitona

Tibeto-Burman peoples

Women mammalogists

Women marine biologists

Yi (Chinese people)

Zerene fritillary

3rd Round of Downloaded Subjects

These were all BT terms (explicit or implicit) on the previous list

Caryophyllaceae

Cultural industries

Hemiphacidiaceae

Speyeria

Women aquatic biologists

Women marine scientists

Women zoologists

4th round of downloaded subjects

These were all BT terms (explicit or implicit) on the previous list

Aquatic biologists

Marine scientists

Nymphalidae

Women aquatic scientists

Women biologists

5th Round of Downloaded Subjects

These were all BT terms (explicit or implicit) on the previous list

Aquatic scientists

Women life scientists

6th Round of Downloaded Subjects

This was a BT term (explicit) on the previous list

Life scientists