

Cataloging for the 21<sup>st</sup> Century – Course 3

# **CONTROLLED VOCABULARY & THESAURUS DESIGN**

Instructor's Manual

Original course design by  
Rich Gazan  
University of Hawaii at Manoa



**Association for Library Collections & Technical Services  
Program for Cooperative Cataloging**

*Available from*  
**LIBRARY OF CONGRESS  
CATALOGER'S LEARNING WORKSHOP**

*THIS PAGE INTENTIONALLY LEFT BLANK  
FOR DOUBLE SIDED COPY*

**Cataloging for the 21<sup>st</sup> Century—Course 3**  
**Controlled Vocabulary & Thesaurus Design**

Table of Contents

**Course Introduction and Background**

<b>Session 1</b>	Types of Controlled Vocabularies
<b>Session 2</b>	Vocabulary Terms & Lists
<b>Session 3</b>	Term Selection/Format & Synonyms
<b>Session 4</b>	Hierarchies
<b>Session 5</b>	Associative Relationships & Thesauri
<b>Session 6</b>	Display & Navigation
<b>Session 7</b>	Planning & Maintenance
<b>Session 8</b>	Resources & Future Directions

**Exercises**

**Sample Front Matter**

**Resources**

**Bibliography**

*THIS PAGE INTENTIONALLY LEFT BLANK  
FOR DOUBLE SIDED COPY*

# Controlled Vocabulary & Thesaurus Design

Course Introduction and Background



## **Suggested activities for getting started (before beginning slide content):**

- Introduce instructor(s) and their professional background.
- Ask each participant to introduce him/herself, his/her institution and position. Very brief, as there will be more sharing opportunities to come.
- Go over with participants the contents of their manuals.
- Point out the table of contents, course packet, examples, exercises and bibliography.
- Discuss the planned timing for the two days, for sessions, exercises, breaks, lunch, etc.

## Possible icebreaker questions:

- In one minute, describe either the best or worst web site you've used? Make sure to explain what about the web site made it deserve this rating.
- Briefly describe your position within your organization.
- In one or two minutes, tell us about something you enjoy collecting. If you can't think of anything you currently collect, tell us about something you would like to collect if you had the time and money.



## *Cataloging for the 21st Century:* The five CE course components

1. MARC/AACR2 cataloging of electronic resources
2. Overview of basic concepts of 21st century bibliographic control, including specific metadata standards and applications
- 3. Thesaurus design principles, building a controlled vocabulary, and examination of selected controlled vocabularies**
4. Digital library design, with a project-based component
5. Philosophy and approach to asset management for the 21st century, including evaluation skills

2

Background for this course:

The third of five courses developed as as part of:

Bibliographic Control of Web Resources: A Library of Congress  
Action Plan

Action Item 5.3: Continuing Education (CE)

Continuing Education Implementation Group (CEIG)

Most organizations have, or are looking at developing, a collection of digital resources, often called a digital repository or digital library. The digital library often contains metadata, or information about objects or groups of objects in the repository. One component of the metadata usually includes (or should include) different sets of controlled terms. The Cataloging for the 21st Century courses support each area of development, from controlled vocabularies, to metadata, and finally the entire digital library.

This course builds off of the basics of bibliographic control, metadata can be made more powerful through development of a controlled vocabulary, to conflate related terms and express hierarchical relationships between terms. This course is then built upon by the digital library design course, where they need to have all of the available organizational tools, including controlled vocabularies, available to design a digital library system.



## *Cataloging for the 21st Century:* CE course series objectives

- To equip catalogers to deal with new types of resources and to recognize their unique characteristics
- To equip catalogers to evaluate competing approaches to and standards for providing access to resources
- To equip catalogers to think creatively and work collaboratively with others inside and outside their home institutions
- To ensure that catalogers have a broad enough understanding of the current environment to be able to make their local efforts compatible and interoperable with other efforts
- To prepare catalogers to be comfortable with ambiguity and being less than perfect
- To enable practicing catalogers to put themselves into the emerging digital information environment and to continue to play a significant role in shaping library services



## Goals for this course

- Understand this course as the third in a series of five courses for Cataloging in the 21st Century and see thesauri within a larger context of vocabulary control
- Understand and apply fundamental concepts of controlled vocabulary and thesaurus design, and why they are important
- Understand and apply diverse types of term relationships to structure descriptive terms
- Understand and apply both basic rules and best practices from existing thesauri to the construction and maintenance of thesauri and controlled vocabularies
- Develop a basis for exercising individual judgment for making thesaurus and controlled vocabulary decisions

4

These are some general overall goals for the 2-day workshop as a whole and what participants should take with them.

**Recommended Activity:** Are there other general goals that participants have? What are they here for? What do they want to take with them? What do they want to get out of this 2-day workshop? Instructors may write student goals on a flip-chart and revisit them at the conclusion of the workshop at the end of Day 2.





## Course overview

- Session 1. Types of Controlled Vocabularies
- Session 2. Vocabulary Terms & Lists
- Session 3. Term Selection/Format & Synonyms
- Session 4. Hierarchies & Taxonomies
- Session 5. Associative Relationships & Thesauri
- Session 6. Display & Navigation
- Session 7. Planning & Maintenance
- Session 8. Resources & Future Directions

5

The course consists of 8 sessions to be covered over two days.

One of the best ways to learn something is to do it, so part of this course is actually building parts of an actual thesaurus.

We will be flexible with the schedule, adjusting it to explore areas this group is most interested in.

However, if the schedule says it's time for a break, or if you need one, make sure you let us know.

Mention the ANSI/NISO standard if you have not already!




## Why Controlled Vocabularies?

- Goal of information retrieval systems:
  - Efficient access to information
- Core problems:
  - Within a domain: different terms with same meaning
  - Between domains: same term with different meanings
- Other problems:
  - Scalability: almost all IR research is based on small test collections or user communities
  - User needs vary with time and subject domain
  - Basic disconnect between indexer and user experience

6

As this is likely to be a cataloging group, you can ask them why subject headings are used rather than keyword search alone. Keyword searches are not sufficient for retrieval as the collection grows larger -- a problem of scalability.

Another scalability problem -- Q.: How many words does the average U.S. high school graduate know (not including proper names, numbers, foreign words, etc.)? 45,000



## Who Controls Vocabularies?

- Information professionals
  - Computer scientists, information scientists, librarians, linguists
- Designers vs. maintainers
  - Which are you?
- What is your own interest in Controlled Vocabularies?

7

We need to stress the idea of a designer vs. maintainers. For the longevity of a thesaurus it is important to have a mixture of these types of people.

Those who develop or maintain controlled vocabularies:

Taxonomist  
Lexicographer  
Ontologist  
Thesaurus Builder  
Information Architect

Those who apply terms from a controlled vocabulary:

Indexer  
Cataloger  
Metadata Specialist

Those who seek information using a controlled vocabulary (searcher, user).

Each has a different view and need for words in retrieving information.

Go around the room and have everyone say what their interest in this course

*THIS PAGE INTENTIONALLY LEFT BLANK  
FOR DOUBLE SIDED COPY*

# Controlled Vocabulary & Thesaurus Design

Types of Controlled Vocabularies





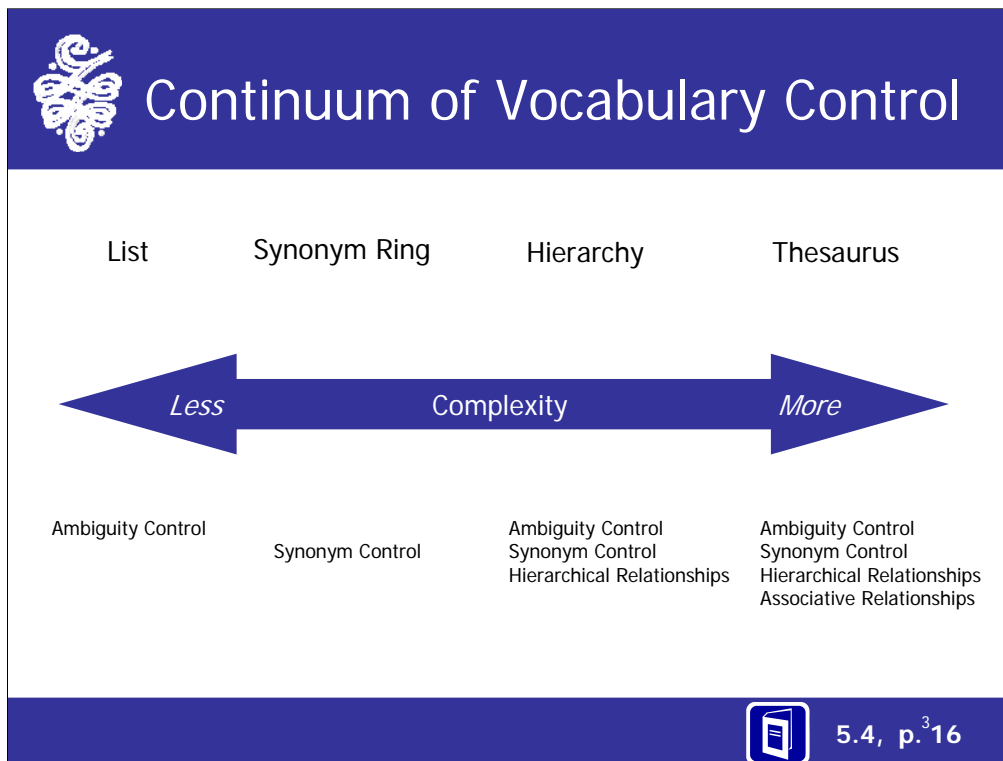
## The Goal

The goal of a controlled vocabulary is to facilitate agreement between the concepts within the site and the vocabulary of the person using it.



1, p. 1<sup>2</sup>

One of the common issues in building and using controlled vocabularies is the disconnect between professionals familiar with the collection and its navigation tools, and users who likely aren't.



From the ANSI/NISO standard

There are different controlled vocabulary solutions to different problems. We will go through the construction of each of these types of controlled vocabulary throughout the workshop.

Note that in some disciplines, thesauri are often referred to as taxonomies or ontologies. In this course we will use the definitions in NISO Z39.19 which we'll cover in more detail later.

Also, work is ongoing on 'semantic networks' which would fall farther to the right than the thesaurus on this continuum.

**List Example 1**

*The Starlite Cafe's Poets Corner*  
Bringing Poets Together, Worldwide  
*For The Love of It!*

**Welcome to The Starlite Cafe**  
Break away from your everyday life and immerse your senses in the splendor of poetry. The Starlite Cafe is your place to relax and enjoy. Browse through our poems. Share some of your own: from love poetry to Goth - all poetry welcome!

- Create your own on-line portfolio.
- Link to your poems from any website.
- Create the perfect look for your poem.
- Improve your writing skills with our writing exercises and poetic challenges.
- Meet new friends!

**Join the Starlite Family**  
Create a profile. Share your work with us. It's fast, it's painless and it's free! Once you've created your profile, you'll have full access to post, modify and delete your poems. Your profile will be linked to all of your new poems. You'll also have full access to our discussion forum and all the custom features of the Cafe.

- ▶ Create a profile
- ▶ Lost your username or password?

**Log In**  
Username:  Password:  **Log In**

**The Poems**

- New Poems
- Latest 100
- To Our Troops
- Love Poems
- Birthdays
- Childrens
- Collaborations
- Congratulations
- Current Events
- Dark
- Dedications
- Fantasy
- Friends
- Grieving & Loss
- Holiday Poems
- Humorous
- I Love You
- Insight

Welcome Raivenne!

4

Example: <http://www.thestarlitecafe.com/>

### The Starlite Cafe's Poets Corner

Discuss the term list shown on the right side of the screen. There is no synonym control, no ambiguity control, and no relationships with other terms shown. Yet, it does what it needs to do -- it controls the terms used to classify the different types of poems in the site's poetry database.



URL: <http://www.inthefirstperson.com/firp/index.aspx>

Site: In the First Person: An index to letters, diaries, oral histories and personal narratives.

Clicking on Historical Events results in the next screen. for an example of a terms list

Click on Search Documents to go to the Advanced Search page. On this page, each facet (field) with a 'Terms' button leads you to a list of terms that the user can select to include as part of the search.

**List Example 2b**

**In the First Person**

General Quick Search:

Home Tables of Contents Search Companion Community Help

Page Repository Collection Document Date Place Historical Event Subject Collection Document Products Community Help

Table of Contents: View Documents by Historical Event

[\[View Collections by Historical Event\]](#)

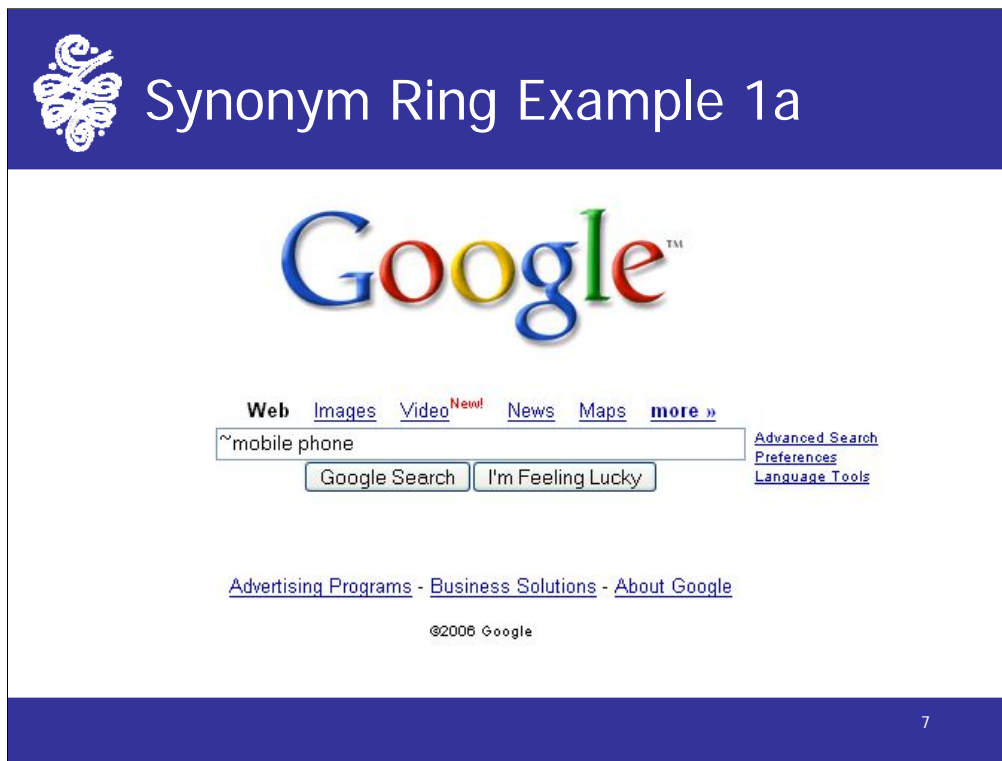
Term	All	Text	Audio	Video
Protestant Reformation, 1500-1650	<u>17</u>	<u>17</u>		
English Protestant Reformation, 1529-1603	<u>3</u>	<u>3</u>		
Abdication of Mary Queen of Scots, 1567	<u>5</u>	<u>5</u>		
Battle of Jarnac, France, March 13, 1569	<u>3</u>	<u>3</u>		
St. Bartholomew's Day Massacre, August 24, 1572	<u>1</u>	<u>1</u>		
Thirty Years War, 1618-1648	<u>1</u>	<u>1</u>		
Irish Rebellion and Cromwell's Invasion of Ireland, 1641-1650	<u>1</u>	<u>1</u>		
English Civil War, 1642-1651	<u>123</u>	<u>123</u>		
Restoration, 1660-1688	<u>6</u>	<u>6</u>		
Jacobite Risings, 1689-1746	<u>3</u>	<u>3</u>		
French and Indian Wars, 1689-1763	<u>41</u>	<u>41</u>		

6

Discuss list of event terms

Terms sorted based on date in term (first year)

Results list includes count of resources with each term assigned



Google search:

~mobile phone

Note: Google has defined the tilde as a term modifier that indicates that the search should include similar terms (i.e., expand the search using a synonym ring if available).

Note:

Extracted from: Leise, Fred, Karl Fast, and Mike Steckel. "Synonym Rings and Authority Files." Boxes and Arrows. 26 Aug. 2003. 21 Aug. 2006. <[http://www.boxesandarrows.com/view/synonym\\_rings\\_and\\_authority\\_files](http://www.boxesandarrows.com/view/synonym_rings_and_authority_files)>

International SEMATECH, the semiconductor research consortium, had a searching problem. Member company employees from all over the world had the ability to upload their own research documents and meeting presentations to the website.

A look at the search logs revealed that people entered search terms that were yielding only a percentage of the documents they were trying to find. The problem was consistency of terminology. A review of the metadata found that those uploading information were equally as likely to call silicon "Si" as they were to spell out the whole name, "silicon." There were many similar examples. Besides chemical symbols, users were both searching and uploading documents with acronyms ("PSM" vs. "Phase Shift Mask") and simple variants in spelling ("low K dielectrics" vs. "low-K dielectrics" vs. "lowk dielectrics").

It was our assumption that when users searched one term, they intended to find the entire set of documents related to that concept. But trying to get such an organization to adopt a style guide for metadata was not viable. The solution was to install a synonym ring into their search engine, Oracle Text.

A synonym ring connects a series of terms together and treats them all as equivalent for search purposes. Synonyms can include different terms for the same topic, acronyms, variant spellings, scientific terms versus popular terms, and terms that are treated as synonymous within the search environment. When a user enters "PSM," for instance, the search term will be sent through the synonym ring to see if there are any equivalent terms. For "PSM" we would find "Phase Shift Mask" as a synonym. The search engine would then retrieve all documents with either "PSM" or "Phase Shift Mask" in their metadata. The searcher would get the complete set of relevant documents as though they had searched both terms (something few people would think to do).

If there is no match in the synonym list, the search is simply sent through the index as usual and any documents with "PSM" are returned. The synonym ring goes into effect only when there is a matching synonym for the term entered into the search box by the user.

But synonyms, in general, quickly become more difficult. Are "medicine" and "drugs" synonyms? Are "fired" and "laid off"? What about "forest" and "woods" or "arid" and "dry"? With these examples, it is more difficult to say for sure. To answer the question about whether two terms are synonyms, you often have to consider the overall content of your site, as well as the site's context and its users.

When creating a synonym ring, or any controlled vocabulary, you will spend a lot of time evaluating near synonyms. What guidelines should you use for making these decisions?



## Synonym Ring Example 1b

[Mobile phone](#) - Wikipedia, the free encyclopedia  
 A **mobile phone** or **cellular (cell) phone** is an electronic telecommunications device. ... The use of a **mobile phone** is prohibited in some rail carriages ...  
[en.wikipedia.org/wiki/Mobile\\_phone](http://en.wikipedia.org/wiki/Mobile_phone) - 74k - [Cached](#) - [Similar pages](#) - [Note this](#)

HowStuffWorks: [Cellular Phones](#)  
 Learn about **wireless** technology, check out our **phone** buying guide, read cell **phone** reviews and **compare prices** and features on **wireless** accessories.  
[www.howstuffworks.com/cell-phone.htm](http://www.howstuffworks.com/cell-phone.htm) - 59k - Aug 20, 2006 - [Cached](#) - [Similar pages](#) - [Note this](#)

Personal: Phones, PDA, Plans - Verizon [Wireless](#)  
 Affordable, reliable **wireless** services for you and your family from Verizon **Wireless**. ... I'm looking for a new **phone** and plan just for me. Start with. ...  
[www.verizonwireless.com/](http://www.verizonwireless.com/) - 26k - [Cached](#) - [Similar pages](#) - [Note this](#)

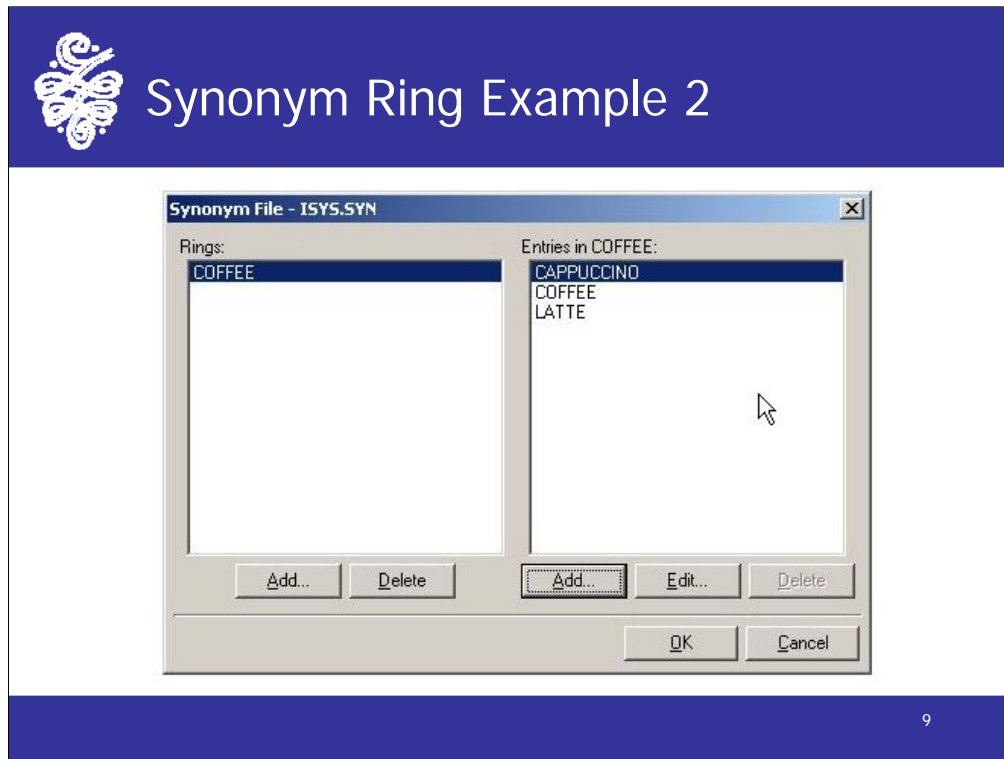
[Compare Cell Phone Prices, Free Cell Phones, Cell Phone plans ...](#)  
 Compare Cell Phones, Research **Cellular** phones, Free cell phones, Compare Cell **Phone** plans, cell **phone** accessories, **wireless** phones, **mobile** phones.  
[www.letstalk.com/](http://www.letstalk.com/) - 74k - [Cached](#) - [Similar pages](#) - [Note this](#)

8

Returns resources containing the terms:

- cell phone
- cellular phone
- GSM phone
- mobile phone
- Nokia phone
- portable phone
- WAP phone
- wireless phone

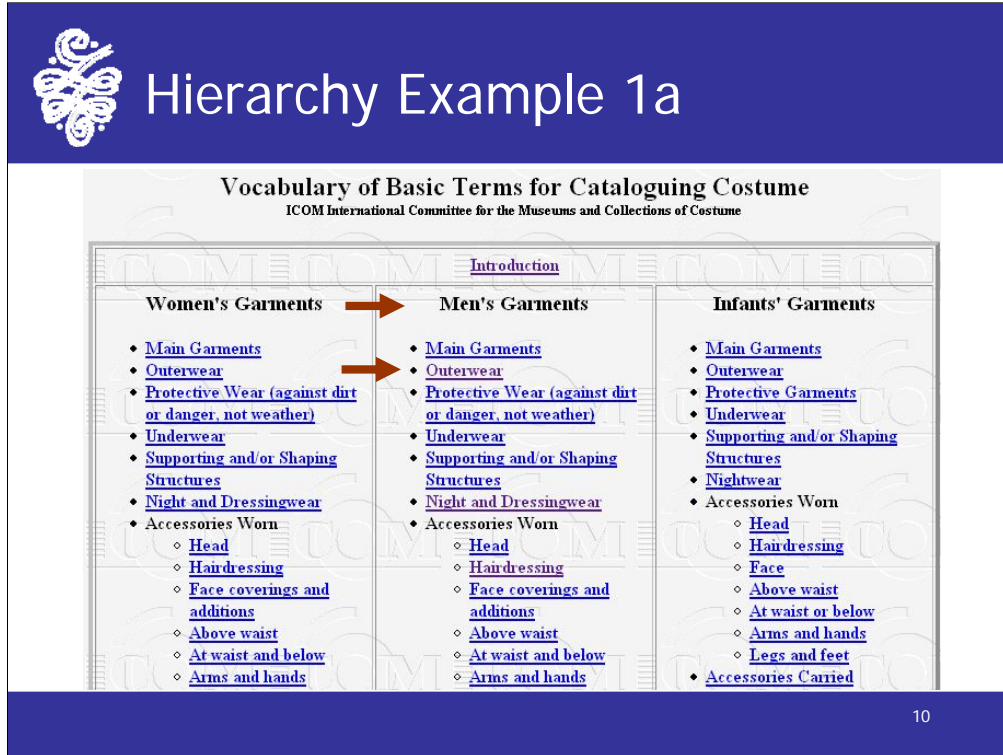
Note: At the time of this search, this synonym ring within Google did not include the term WiFi phone.



Example of a maintenance interface for creating and maintaining synonym rings within the ISYS search software.

Note that "COFFEE" is the name of the ring or record. That is the reason that term is shown in the left column. However, it is NOT a preferred term, as all terms in a ring are co-equal.

Note: Can be viewed on web at (viewed on 21 Aug. 2006):  
<http://www.isys-search.com/support/techtips/synonyms.html>



**Hierarchy Example 1a**

**Vocabulary of Basic Terms for Cataloguing Costume**  
ICOM International Committee for the Museums and Collections of Costume

Introduction

Women's Garments	Men's Garments	Infants' Garments
<ul style="list-style-type: none"> <li>• <a href="#">Main Garments</a></li> <li>• <a href="#">Outerwear</a></li> <li>• <a href="#">Protective Wear (against dirt or danger, not weather)</a></li> <li>• <a href="#">Underwear</a></li> <li>• <a href="#">Supporting and/or Shaping Structures</a></li> <li>• <a href="#">Night and Dressingwear</a></li> <li>• <a href="#">Accessories Worn</a> <ul style="list-style-type: none"> <li>◦ <a href="#">Head</a></li> <li>◦ <a href="#">Hairdressing</a></li> <li>◦ <a href="#">Face coverings and additions</a></li> <li>◦ <a href="#">Above waist</a></li> <li>◦ <a href="#">At waist and below</a></li> <li>◦ <a href="#">Arms and hands</a></li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• <a href="#">Main Garments</a></li> <li>• <a href="#">Outerwear</a></li> <li>• <a href="#">Protective Wear (against dirt or danger, not weather)</a></li> <li>• <a href="#">Underwear</a></li> <li>• <a href="#">Supporting and/or Shaping Structures</a></li> <li>• <a href="#">Night and Dressingwear</a></li> <li>• <a href="#">Accessories Worn</a> <ul style="list-style-type: none"> <li>◦ <a href="#">Head</a></li> <li>◦ <a href="#">Hairdressing</a></li> <li>◦ <a href="#">Face coverings and additions</a></li> <li>◦ <a href="#">Above waist</a></li> <li>◦ <a href="#">At waist and below</a></li> <li>◦ <a href="#">Arms and hands</a></li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• <a href="#">Main Garments</a></li> <li>• <a href="#">Outerwear</a></li> <li>• <a href="#">Protective Garments</a></li> <li>• <a href="#">Underwear</a></li> <li>• <a href="#">Supporting and/or Shaping Structures</a></li> <li>• <a href="#">Nightwear</a></li> <li>• <a href="#">Accessories Worn</a> <ul style="list-style-type: none"> <li>◦ <a href="#">Head</a></li> <li>◦ <a href="#">Hairdressing</a></li> <li>◦ <a href="#">Face</a></li> <li>◦ <a href="#">Above waist</a></li> <li>◦ <a href="#">At waist or below</a></li> <li>◦ <a href="#">Arms and hands</a></li> <li>◦ <a href="#">Legs and feet</a></li> </ul> </li> <li>• <a href="#">Accessories Carried</a></li> </ul>

10

## Vocabulary of Basic Terms for Cataloguing Costume


Shows hierarchy with control of synonyms and ambiguity

Clicking on Men's Garments | Outerwear results in the next screen



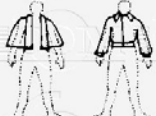
Note:

From site <http://www.mda.org.uk/costume/vbt00e.htm>

Introduction available at [http://www.mda.org.uk/costume/vbt\\_ie.htm](http://www.mda.org.uk/costume/vbt_ie.htm)



## Hierarchy Example 1b

Men's Garments	
<b>Outerwear</b>	
2.1 <i>Unshaped textile, covering upper half of body or more</i> <b>Shawl</b> Shawl Plaid	
2.2 <i>Covering body above and below waist without shaping for arms</i> <b>Cloak</b> <i>with shaping for arms</i> <b>Overcoat</b> Overcoat Mackintosh	
2.3 <i>Covering body above waist without shaping for arms</i> <b>Cape</b> <i>with shaping for arms</i> <b>Jacket</b>	

11

Shows control of synonyms, plus the display of scope notes.

Note the first bold term in each box is the preferred term, while additional bold terms are synonyms

Centre for Computer-aided Egyptological Research *Global Egyptian Museum*

news egyptian treasures shop egyptological resources virtual exhibition  
support about

Back

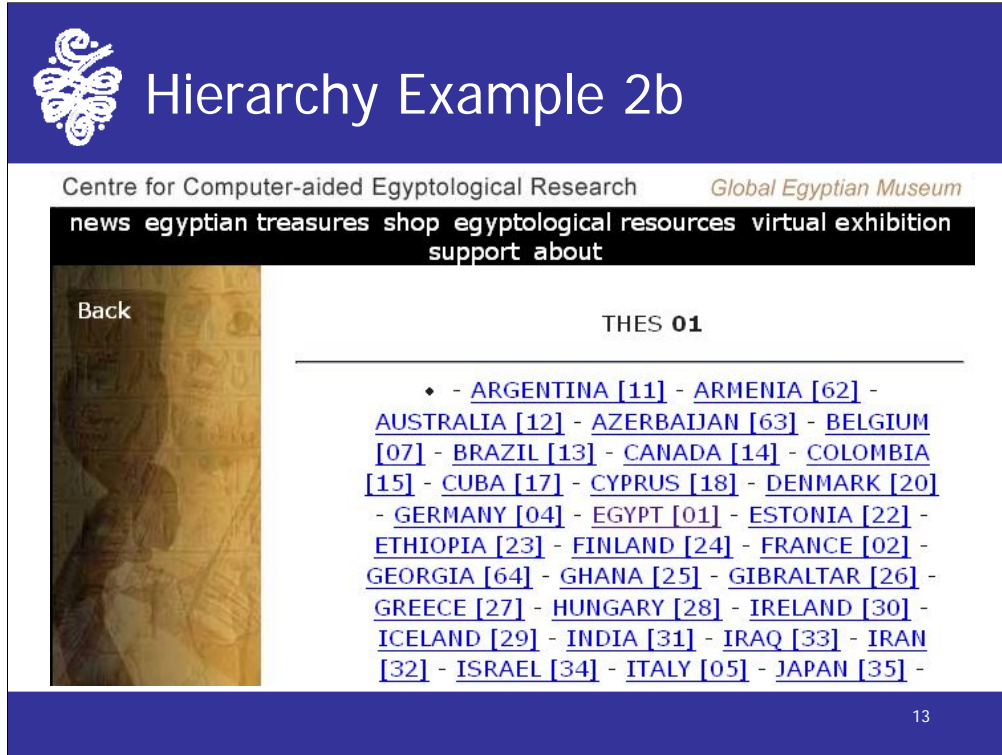
Please select the thesaurus:

- [Thes 01 - Present location](#)
- [Thes 02 - Category](#)
- [Thes 03 - Provenance](#)
- [Thes 04 - Dating](#)
- [Thes 05 - Material](#)
- [Thes 06 - Technique](#)
- [Thes 07 - State of preservation](#)
- [Thes 08 - Description](#)
- [Thes 09 - Language](#)
- [Thes 10 - Writing](#)
- [Thes 11 - Category of text](#)
- [Thes 12 - Text content](#)
- [Thes 13 - Divine names](#)
- [Thes 14 - Royal names](#)
- [Thes 15 - Acquisition](#)

12

Clicking on Present Location results in following screen





Centre for Computer-aided Egyptological Research *Global Egyptian Museum*

news egyptian treasures shop egyptological resources virtual exhibition  
support about


Back

THES 01

- - [ARGENTINA \[11\]](#) - [ARMENIA \[62\]](#) - [AUSTRALIA \[12\]](#) - [AZERBAIJAN \[63\]](#) - [BELGIUM \[07\]](#) - [BRAZIL \[13\]](#) - [CANADA \[14\]](#) - [COLOMBIA \[15\]](#) - [CUBA \[17\]](#) - [CYPRUS \[18\]](#) - [DENMARK \[20\]](#) - [GERMANY \[04\]](#) - [EGYPT \[01\]](#) - [ESTONIA \[22\]](#) - [ETHIOPIA \[23\]](#) - [FINLAND \[24\]](#) - [FRANCE \[02\]](#) - [GEORGIA \[64\]](#) - [GHANA \[25\]](#) - [GIBRALTAR \[26\]](#) - [GREECE \[27\]](#) - [HUNGARY \[28\]](#) - [IRELAND \[30\]](#) - [ICELAND \[29\]](#) - [INDIA \[31\]](#) - [IRAQ \[33\]](#) - [IRAN \[32\]](#) - [ISRAEL \[34\]](#) - [ITALY \[05\]](#) - [JAPAN \[35\]](#) -

13

Clicking on Egypt results in next screen



## Hierarchy Example 2c

01/0126/1	<a href="#">EGYPT [01]</a>
01/0127/2	→ <a href="#">ALEXANDRIA</a>
01/0128/3	→ <a href="#">MUSEUM OF THE FACULTY OF ARTS, UNIVERSITY OF ALEXANDRIA [01/022]</a>
01/1444/3	→ <a href="#">GRAECO-ROMAN MUSEUM [01/026]</a>
01/1445/3	→ <a href="#">JEWELLERY MUSEUM [01/027]</a>
01/0129/2	→ <a href="#">ASWAN</a>
01/0130/3	→ <a href="#">ASWAN MUSEUM [01/023]</a>
01/0131/3	→ <a href="#">NUBIA MUSEUM [01/024]</a>
01/0132/2	→ <a href="#">ASYUT</a>
01/0133/3	→ <a href="#">BENI SUEF MUSEUM [01/025]</a>

14

Shows hierarchy

- 1) Cities Alexandria, Aswan and Asyut (etc.) below Egypt
- 2) Corporate bodies within each city

**Thesaurus Example A**

**NLM NATIONAL LIBRARY OF MEDICINE MEDICAL SUBJECT HEADINGS MeSH**

MeSH Home | Contact NLM | Site Index | Search Our Web Site | NLM Home

Health Information | Library Services | Research Programs | New & Noteworthy | General Information

MeSH Browser (2006 MeSH):  
The files are updated every week on Sunday.  
[Go to 2005 MeSH](#)

Enter term or the beginning of any root fragments: or

Search for these record types:

- Main Headings
- Qualifiers
- Supplementary Concepts
- All of the Above
- Search as MeSH Unique ID
- Search as text words in Annotation & Scope Note

Search in these fields of chemicals:

- Heading Mapped To (HM) (Supplementary List)
- Indexing Information (II) (Supplementary List)
- Pharmacological Action (PA)
- CAS Registry/EC Number (RN)
- Related CAS Registry Number (RR)

15

This is MeSH's front page.



## Thesaurus Example B

### MeSH Descriptor Data

[Return to Entry Page](#)

<b>MeSH Heading</b>	Statistics
<b>Tree Number</b>	<a href="#">E05.318.740</a>
<b>Tree Number</b>	<a href="#">G03.850.520.830</a>
<b>Tree Number</b>	<a href="#">H01.548.832</a>
<b>Tree Number</b>	<a href="#">N05.715.360.750</a>
<b>Annotation</b>	IM SPEC only. SPEC qualif, NIM no qualif, prefer / <a href="#">statist</a> with non-disease terms: Manual <a href="#">19.7+</a> , <a href="#">19.8.68</a> ; <a href="#">statist</a> on dis = / <a href="#">epidemiol</a> ; IM for statistical theory or statistical methods; tables in texts are not necessarily / <a href="#">statist</a> ; do not confuse with <a href="#">MATHEMATICS</a> ; DF: STATIST
<b>Scope Note</b>	The science and art of collecting, summarizing, and analyzing data that are subject to random variation. The term is also applied to the data themselves and to the summarization of the data.
<b>Entry Term</b>	Area Analysis
<b>Entry Term</b>	Correlation Studies
<b>Entry Term</b>	Correlation Study
<b>Entry Term</b>	Correlation of Data

16

Note: MeSH allows poly-hierarchy, as can be seen from the multiple tree numbers.

The rest of the record follows on the next slide...




## Thesaurus Example C

<b>Entry Term</b>	Indirect Estimation Techniques
<b>Entry Term</b>	Multiple Classification Analysis
<b>Entry Term</b>	Service Statistics
<b>Entry Term</b>	Statistical Study
<b>Entry Term</b>	Statistics, Service
<b>Entry Term</b>	Tables and Charts
<b>See Also</b>	<a href="#">Bibliometrics</a>
<b>See Also</b>	<a href="#">Biometry</a>
<b>See Also</b>	<a href="#">Meta-Analysis</a>
<b>See Also</b>	<a href="#">Models, Statistical</a>
<b>See Also</b>	<a href="#">Psychometrics</a>
<b>See Also</b>	<a href="#">Sociometric Techniques</a>
<b>Allowable Qualifiers</b>	<a href="#">CL EC ED ES HI IS LJ MA MT OG ST TD</a>
<b>Entry Version</b>	STATIST
<b>Unique ID</b>	D013223

17

More of the same record...



## Thesaurus Example D

- [Investigative Techniques \[E05\]](#)
  - [Epidemiologic Methods \[E05.318\]](#)
    - [Contact Tracing \[E05.318.270\]](#)
    - [Data Collection \[E05.318.308\] +](#)
    - [Disease Notification \[E05.318.362\]](#)
    - [Epidemiology, Molecular \[E05.318.416\]](#)
    - [Sentinel Surveillance \[E05.318.650\]](#)
    - ▶ [Statistics \[E05.318.740\]](#)
      - [Actuarial Analysis \[E05.318.740.100\]](#)
      - [Analysis of Variance \[E05.318.740.150\] +](#)
      - [Area Under Curve \[E05.318.740.200\]](#)
      - [Cluster Analysis \[E05.318.740.250\] +](#)
      - [Confidence Intervals \[E05.318.740.275\]](#)
      - [Data Interpretation, Statistical \[E05.318.740.300\]](#)
      - [Discriminant Analysis \[E05.318.740.350\]](#)

18

Here is the hierarchical location of Statistics within the MeSH universe.



## Other Means of Access

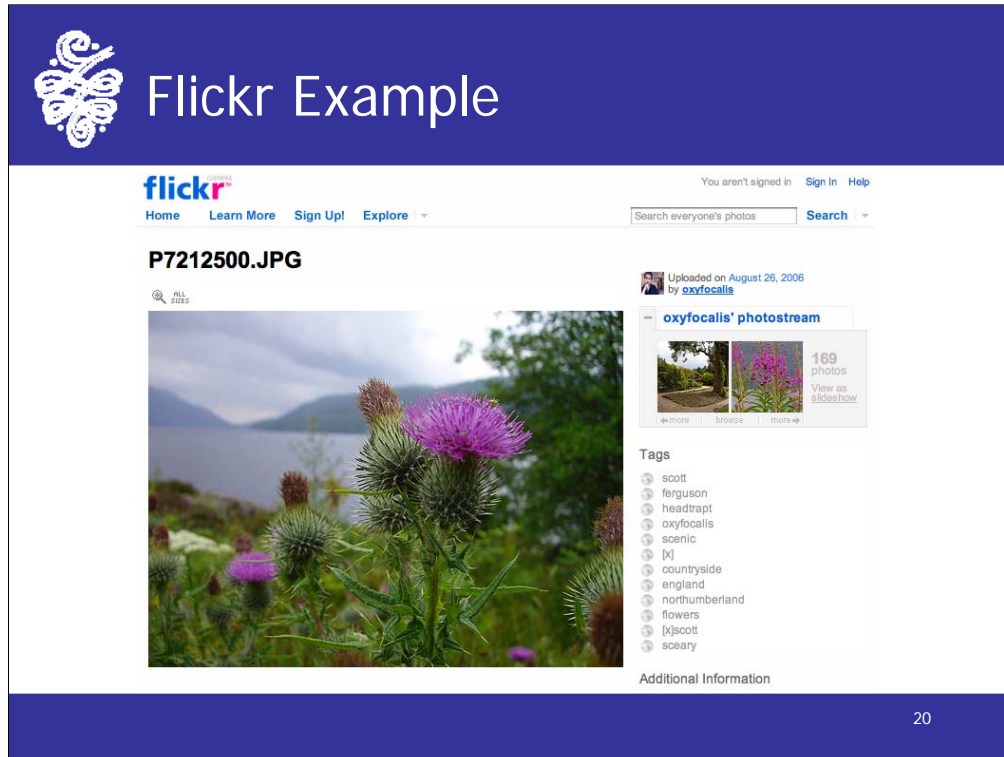
- Folksonomies & tagging communities
  - Technorati - tagging blog entries
  - Digg - tagging webpages
  - Flickr - tagging images
  - Answerbag - folksonomy
- Classification systems
  - Library of Congress Classification Outline
  - Universal Decimal Classification
- Keyword extraction and search

19

Folksonomies are community created and implemented organization schemes, which are typically not controlled.

Tagging communities are equivalent to applying keywords, just with more people doing the applying.

Keyword are not controlled, but work well for access. Seriously hampered by problems of scale.



Tags that have been added by users for retrieval are along the right side. As an information professional, are these the descriptors you would have applied to this image?

We can also use this opportunity to review levels of description -- as a return to cataloging first concepts -- 1) context of the image (photographer, location, picture resolution) 2) Contents of the image (lavendar, green, thistle, mountains, lake, etc) 3) Aboutness (style, themes, etc) of the image (pastoral, vacation, etc). Interesting to see what levels people tag up in Flickr.

That said there are some things that Flickr's community tools that make it very powerful, namely the sheer magnitude of people tagging. They have to be right sometimes, and they require less resources. It is a quantity/quality tradeoff.





**Google Co-op Example**

[Sign in](#)

**Google Co-op**

[Co-op Home](#)

**For Users**

- [Directory](#)
- [FAQ](#)
- [Subscriptions](#)

**For Contributors**

- [Profile](#)
- [Topics](#)
- [Subscribed links](#)
- [Discussion Group](#)

**Topics**

- [Developer Guide](#)
- [FAQ](#)

**Subscribed Links**

- [Developer Guide](#)
- [FAQ](#)

**Other Resources**

- [Resources](#)
- [Glossary](#)

**Health**

Welcome to the Google Co-op page for the health topic. You can help improve search around this topic by contributing your expertise.

The labels for this topic – both those in use and those in development – are listed below. We encourage you to annotate useful and reliable webpages in this topic using the labels below. (Click a label to get its definition.) You can also use the search box below to test out the current version of this topic.

Join in the conversation about this topic at its [discussion group](#).

**Health labels**

<a href="#">Condition info</a>	<a href="#">Drug info</a>	<a href="#">For doctors</a>	<a href="#">Info type</a>
<a href="#">Overview</a>	<a href="#">Drug uses</a>	<a href="#">Research overview</a>	<a href="#">From medical authorities</a>
<a href="#">Symptoms</a>	<a href="#">Side effects</a>	<a href="#">Practice guidelines</a>	<a href="#">Alternative medicine</a>
<a href="#">Tests/diagnosis</a>	<a href="#">Interactions</a>	<a href="#">Patient handouts</a>	<a href="#">For health professionals</a>
<a href="#">Treatment</a>	<a href="#">Warnings/recalls</a>	<a href="#">Continuing education</a>	<a href="#">For patients</a>
<a href="#">Causes/risk factors</a>		<a href="#">Clinical trials</a>	<a href="#">Support groups</a>

Try it out:

21

Google coop is a cooperative labeling initiative, and they define how the labels should be applied by some Coop person. The question is how well this is enforced. 99% of people who will make the effort to use this service are going to be doing it at least pretty well, but there is some question of consistency, and even trained indexers cannot be relied on to 100% consistency.




## Criteria for Selecting a CV Tool

- Institutional requirements
  - Budget, time, manpower constraints
  - Purpose of the CV tool
- Benefits to the user
  - Expected uses for the CV tool
  - Balancing ease of use and functionality



App. B, <sup>22</sup>p. 135

We will revisit these ideas on day 2, but it is good to be thinking about them all along. A CV tool is designed and implemented to meet some need. That need is constrained



## Defining Scope & Users

- Scope of the material
  - What is your core topic or topics?
  - What are your surrounding topics?
  - What is there a need for?
  - What is already out there?
  - What do you have resources for?
- User population
  - Who is intended user?
  - How and when will they use the CV tool?


23

These are some questions to keep in mind, and we will try to get to the point of being able to answer them in our sessions tomorrow.

Design is an iterative process, and requires a lot of trial and error.

You may be able to answer these questions right now, but upon finding materials and through user feedback, you may find you need to take a different approach.

This is perfectly natural, and room for error and backtracking is something that needs to be accounted for in your resource allocation.




## Appropriate CV Applications

- Subject
- Genre
- Physical material
- Personal names
- Corporate/organization names
- Events
- What else?

24

Ask for more examples, discuss name authority lists, ask why they are not thesauri (no explicit relationship between terms, among others)



## CV Tool Discussion

- What is your topic?
- Who is your intended audience?
- What is your institution's need for a CV tool?

25

You can either go around the room for this one or ask for people to volunteer to give preliminary answers to these questions

This discussion is meant to find out what people plan to get out of the program, a networking opportunity, and topics for future examples. Being able to refer back to the challenges students are bringing through the door, will help give them context.


Having a specific purpose and user group in mind during the course will help attendees connect the training with future applications.

*THIS PAGE INTENTIONALLY LEFT BLANK  
FOR DOUBLE SIDED COPY*

# Controlled Vocabulary & Thesaurus Design

Vocabulary Terms & Lists





## Generating Wordstock

- Goal:
  - Find terms that accurately reflect the collection or domain
- Where: sources & warrant
  - People
  - Published materials
  - Data dictionaries
- What: content bearing terms
  - This is tricky...

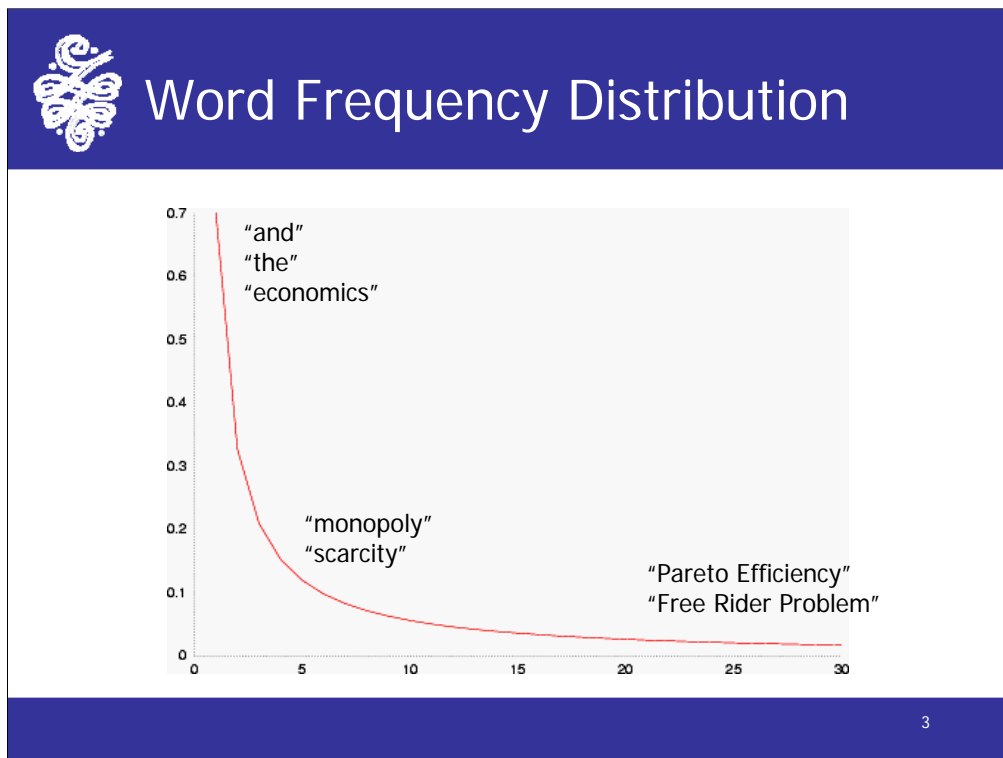
2

People who are experts and literature or published materials are the socially accepted representatives for a domain, and as such are the usual sources for terms.

As a part of the term description the sources of the term are listed to show validity. We call this “literary warrant”.

As it is harder to quote people than published materials, going to an expert for an overview of the domain or a starting point into the literature can be extremely useful. Or if you are the expert, feel free to use your own knowledge of the domain!





The Zipf Distribution or Brown Corpus Distribution (there are other names for it as well) describes the distribution of term frequencies in a given collection of language.

For instance if we were looking at terms in an introductory Economics textbook, we see a large number of uses of common words such as "and" or "the", or even "economics" as this is the subject of the book and we would see many instances of its use. On the other end of the spectrum we will see a large number of uncommon words that are used once, such as "Pareto Efficiency" and "Free Rider Problem". Somewhere in the middle are content bearing terms, such as "monopoly" and "scarcity". These terms are not common enough to be devoid of meaning, and yet common enough to not be too specific.

Content bearing terms will be frequent in a document, but rare over a collection.

Discussion: are content-bearing terms always the most likely to be searched?



## Warrant

Why select particular terms over others:

- *Literary Warrant* - using the terminology found in literature of the topic
- *User Warrant* – using the terminology people familiar with the topic use to describe concepts
- *Organizational Warrant* – based on the needs and priorities of the organization



5.3.5, p. 16 & 6.6.1, p. 30

Sources:

Literary Warrant – articles, books, reports, dictionaries, textbooks, existing vocabularies

User Warrant – user requests, search logs

Organization Warrant – organization manuals, policies

See 6.6.1 on page 30 for selection of neutral terms (e.g., **developing nations** rather than **underdeveloped countries**)

Discussion:

How might terminology differed based on warrant?

Why use one type of warrant over another?



## Exercise 1: Generate Wordstock

- Get into groups of 2-3
- Read the provided materials
- Identify potential content bearing terms
- Write each term on a Post-It

5

Half of your groups should be using top down materials (such as indices or tables of contents), the other half of the groups should be using bottom up materials (such as articles).

Hand out your materials and Post-Its.

Have a list of terms you were able to pull from the articles in preparation.

{5-10 minutes}

The objective of this exercise is purely to extract terms and think about what is content bearing. We will get to synonyms and organization later! (next slide).



## Exercise 1: Discussion

- What terms did you come up with?
- Are they content bearing?
- Are users going to search on these terms?
- Who are the potential users for this collection?
- Were these appropriate sources of wordstock?

6

There are no right or wrong answers to these questions. But you should have some idea in mind of who these materials should be for, such as introductory beer makers or beer drinking enthusiasts or brewers.

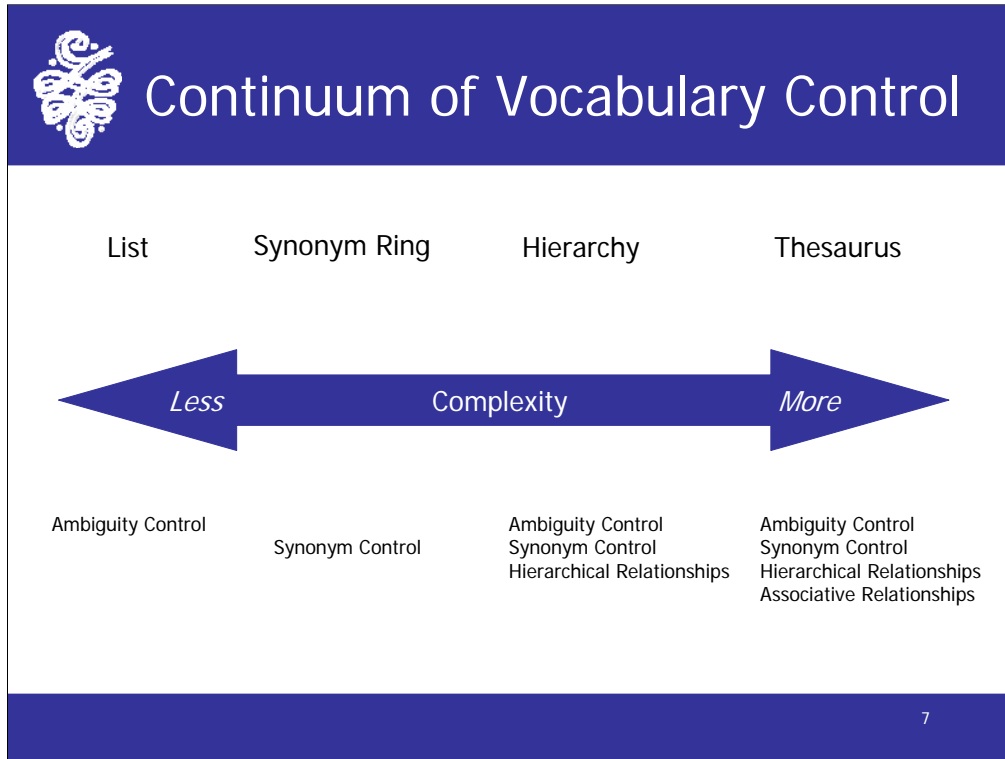
If using visual materials ask if the descriptions of the images hit all three levels of image description:

- 1) Context of the image (photographer, location, date, picture resolution)
- 2) Contents of the image (lavendar, green, thistle, mountains, lake, etc)
- 3) Aboutness (style, themes, etc) of the image (pastoral, vacation, etc).

Hopefully you will be able to identify a potential user for these terms. And hopefully you will find some differences in the terms found and selected as the top down vs. bottom up groups will have different terms.

Have them hold onto their terms as they will be used in the next exercise.

{5-10 minutes}



To return to this slide for a moment...

We are about to launch into the List, so this is a good time to reorient folks to the continuum of vocabulary control.



## The List

- The list literally provides a list of topics represented by your collection
- Terms used have neither hierarchy nor synonyms
- A list generally consists of a small number of terms whose use is intuitive, e.g., format, genre, size, etc.
- List creation
  - Generate wordstock
  - Weed for appropriate terms
  - Organize materials under terms



5.4.1, <sup>8</sup>p. 17

Our first controlled vocabulary tool is the List. And is the most simple. But that does not mean it is not useful!

In practice, two type of lists exist – controlled and uncontrolled. In an uncontrolled list, all terms in a given field (for example) are automatically added to the list. In a controlled list, only the terms in the list can be applied or searched.

The screenshot shows the homepage of the visualcomplexity website. At the top left is a logo of a stylized butterfly. The main heading is "List Example 1". Below the header is a navigation bar with links for Home, About, News, Press, Resources, and Contact. There is also a search bar and a "Search" button. On the right side of the navigation bar, there are links for "UC RSS Feed" and "Support UC: PayPal", along with several small icons representing different services or languages.

The main content area is divided into two sections: "Images" and "Text". The "Images" section displays a grid of 15 network visualizations, each representing a different project or dataset. These visualizations are colorful and complex, showing various types of network structures such as social networks, biological networks, and transportation networks. The "Text" section is currently empty.

On the right side of the main content area, there is a "Filter by:" section. It includes three tabs: "Subject", "Top Years", and "Top Authors". Below these tabs is a list of subjects with their corresponding project counts:

- ALL (353)
- Art (23)
- Biology (32)
- Business Networks (19)
- Computer Systems (19)
- Food Webs (7)
- Internet (26)
- Knowledge Networks (66)
- Multi-Domain Representation (25)
- Others (22)
- Pattern Recognition (11)
- Social Networks (44)
- Transportation Networks (25)
- World Wide Web (34)

At the bottom right corner of the screenshot, there is a small number "9".

Lists are a powerful access tool when used appropriately.

Here are some examples of lists that can be found in the wild...

The visual complexity site hosts projects that are pushing the envelope of how information is presented visually, and I would highly recommend visiting it sometime to be inspired. There is a very short list of topics, by which these projects can be filtered, and they have very nicely listed the frequency of projects in each topic.

<http://www.visualcomplexity.com/vc/>



Etsy materials lists and keyword tags.

Etsy is a site where users can sell their handmade items, and they ask the seller to list the materials used and any keywords that would help purchasers find the wares. Use of the term as a descriptor increases the font size. From this we can see that many of the items sold on Etsy use beads, cotton, glass, paper, and silver as a material. And you can easily see that they have many items that could be described as earrings, jewelry, beads, by color (blue, pink, red, green), or by adjectives like cute.

There is no categorical organization to these terms!

<http://www.etsy.com>





## List Discussion

- What are the potential benefits of a List?
- What are the potential drawbacks of a List?
- When would it be appropriate to use a List for access?

11

This should be a short discussion about when using a list is appropriate. Here are some possible answers to these questions...

Potential benefits:

Easy for users to understand

Good for smaller collections

Relatively inexpensive to create and implement

Potential drawbacks:

Does not allow for variant terms

Not expressive for a big collection - does not scale well! Going back to List Example 1, given the number of items within each filtering term, browsing through 200 items is not a pleasant experience, and a sub category to break up the filter term would be more appropriate.

Like describing every book in a economics collection with the term economics!

*THIS PAGE INTENTIONALLY LEFT BLANK  
FOR DOUBLE SIDED COPY*

# Controlled Vocabulary & Thesaurus Design

Term Selection/Format & Synonyms





## Term Selection

- Defining your scope
  - What is the purpose of this controlled vocabulary?
  - Who is the user of this controlled vocabulary?
- Refining your scope
  - Which terms are central to your collection/topic/mission?
  - Which terms are on the periphery?
- Stick with your scope
  - Focus on the central terms
  - Hang onto the peripheral terms for future consideration

2

We touched on this briefly before in session 2, but here is a more systematic way of thinking about which terms stay and which terms should go. We will start making these decisions during this session's exercise.



## Term Format & Style

- Plural nouns
  - Ex. Cats, Dogs
- Abstract concepts
  - Ex. Happiness
- Abbreviations
  - Use only when the full form is rarely used (e.g., **AIDS, DNA**)
  - When the full form is used, add a cross reference from the abbreviation (and vice versa)
- Capitalization
  - Generally, use capital letters only for proper names, etc.
  - Follow common usage (e.g., **eBay, dBase IV**)



6.3–6.7, p. 23<sup>3</sup>-34

Why would nouns, specifically plural nouns, be the term format of choice?

We are describing things

We are describing multiples

Makes the terminology easier to understand if it is consistent

Cuts down on confusion in the long run

Examples where a single term is insufficient:


stained glass

mass spectrometry

space shuttle


public television

rapid transit



## Compound Terms - Good or Bad?

- Multiple words to express a single concept
- Supported by literary or user warrant  
**Aid to families with dependent children**
- Avoid false hits during retrieval  
Search: **library AND science**  
Desired: **Library science** or **Science library ??**
- When to use compound terms – section 7.5 on page 39
- Generally, avoid multiple concepts in a single expression  
**Children and television** – prefer **Children AND Television**  
**Bird migration** – prefer **Birds AND Migration**  
**Aircraft engines** – prefer **Aircraft AND Engines**


7, p. 36

Terms, however, may consist of one or more words. However, even multiword terms should represent a “single concept”. (6.3.1, page 23). Section 7 on page 36 provides some guidance on when multiword terms should be used. Additional guidance can be found on page 21 in the section on qualifiers.

Examples where a single term is insufficient – words bound together as lexical units:

stained glass  
mass spectrometry  
space shuttle  
public television  
rapid transit

Multiword terms should be used only if the phrase is common in natural language.

Examples:

(phonograph)	phonograph records	<i>rather than</i>	records
	religious tolerance tolerance (religion)	<i>rather than</i>	

There are no hard and fast rules – use natural language and listen to your users and indexers



## Descriptor vs. Subdivision Logic

- Descriptor vocabularies express multiple concepts with separate terms
  - Each concept or aspect warrants a term
  - Intended for post-coordinate searching – user brings concepts together
  - Ex: **Liver AND Blood Vessels AND Diseases**
- Vocabularies like LCSH use subdivisions to refine major concepts with additional aspects
  - Selected aspects are routinely expressed as subdivisions
  - Intended for pre-coordinate searching – vocabulary brings concepts together
  - More complex to create, learn and apply
  - Ex: **Liver—Blood Vessels--Diseases**



7.2.1-2, p. 37<sup>5</sup>



## Terms Can Be Ambiguous

- “Pitch” can represent:
  - a property of sound
  - throwing a baseball
  - a sales talk
  - degree of deviation from the horizontal plane
  - dark heavy viscid substance
  - a high approach shot in golf
  - a card game
  - abrupt up-and-down motion
  - the action of throwing something
  - etc.



5.3.1, p. 13 <sup>6</sup>

Words, especially in English, are often used to represent multiple concepts.

Homonym - same pronunciation and spelling as another word, but a different meaning, e.g., park (action) and park (place).

Heteronyms (aka, heterophones) - words spelled the same but have different pronunciations and meanings, e.g., desert (abandon) and desert (arid region).

Homograph - same spelling as another word, but a different meaning (encompasses both homonyms & Heteronyms)

Homophone - same pronunciation as another word, but a different meaning, e.g., carat (unit of mass for gems & a board game), caret (the upside down v symbol), carrot (vegetable) and karat (unit of purity for gold & a rock band in Germany & a Pacific Island fruit).





## Term Disambiguation

- Term qualifiers can be used to disambiguate between two or more terms with different meanings
  - Mercury (planet)
  - Mercury (metal)
- Try to use only when necessary



6.2, p. 21

Qualifiers can be used either when a term is inherently ambiguous or often unclear to the site's users, e.g.,


Fields (database)

or when similar terms exist within the thesaurus, e.g.,

Coatings (process)


Coatings (material)

In addition to using qualifiers to make an ambiguous term clearer, qualifiers have an additional use which we'll cover in Session 4.



## Resolving Synonyms

- Determine which is the preferred term
  - What do your sources say?
    - Which is more frequently used?
    - Which is more likely to be prefaced with “also known as”?
  - Talk to your experts
  - Ask the user community
    - Through user testing
    - Checking search logs
- Save the others as synonyms

 6.6, p.<sup>8</sup> 30

The more frequently uses term should be the preferred term, and the aka's are the synonyms.

Involving your users in the development of your CV tool is covered more thoroughly in Session 7.

Look at 5.3.2 on page 13 for example of Synonymy



## Term Record Components

- Preferred or Main Entry Term
  - SN - Scope note
    - A definition of when to the term applies
    - Anything from a few words to a paragraph
  - UF - Use for
    - Indicates the synonymous terms for which the preferred term should be substituted
- Non-preferred or Non-entry term
  - USE - Use
    - Indicates the preferred term to be used instead of the synonym term



4.1 – 4.2, p. 3-10

Preferred term is also know as an entry term or main entry term, and the non-preferred term can be know as the non-entry term.



## Example Scope Notes

- Cartography
  - SN Science or art of making maps
- Printmaking
  - SN The process of making a picture or design from an impression of an engraved metal plate, wooden block, silkscreen stencil, lithographic stone, photographic negative, etc.
- Patternmaking
  - SN (Note: Do not use for the study of numerical patterns -- see the Identifier "Number Sequences" for that concept)

10

These are from ERIC...

The scope note can be used to describe when the term should be used and when another term should be used, as can be seen in the last example, which is under the broader term of "Occupations" and refers to sewing patterns.

Before we expect them to write their own, it is only fair to show them some scope notes.



## Exercise 2: Terms & Synonyms

Using your terms from the previous exercise...

- Format your terms
  - As plural nouns, abstract concepts, compound terms, etc.
  - Add qualifiers where necessary
- Resolve any synonyms into
  - Preferred terms
  - And their non-preferred terms
- Start term records for each term
  - Preferred terms
    - Add a SN and UF if there are synonyms
  - Non-preferred term
    - Add the preferred term to USE

11

Using the post-it's the groups started in exercise 1, they should select and format their terms. Then they should resolve any synonyms.

If people do not have enough terms to do anything with, have the half with top-down materials trade with the bottom-up groups, so that they can cull terms from both sets of materials. The materials should be sufficiently term rich though.

{10-20 minutes}



## Exercise 2: Discussion

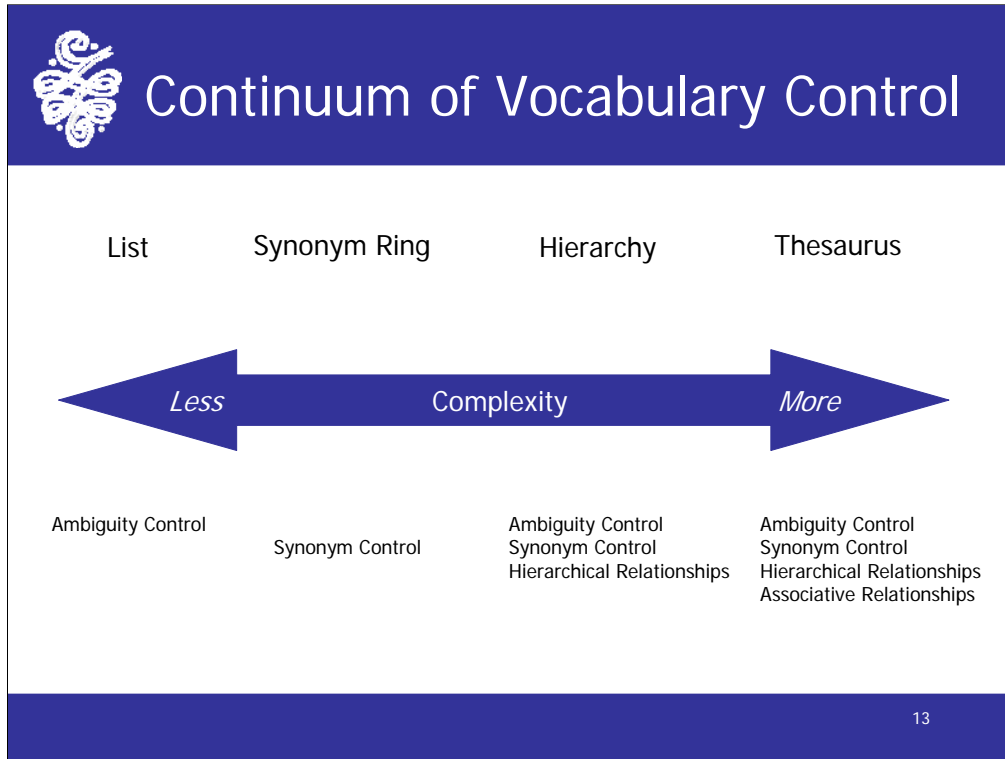
- How did you format your terms?
- What terms did you select as preferred?
- Was everyone in your group in agreement about the terms?

12

Hopefully there are some differences in the wordstock generated so you can examine the top level terms that come from top-down, and the more common words that will come from the bottom level.


Hopefully working in groups will have brought up the idea of linguistic compromise. Thesaurus development is always on the balance between access and being true to the domain, which are sometimes at odds. Additionally it should be noted that language will not be uniform across the same domain. The last question is meant to get to the meat of this problem. They should hold onto their terms, as they are going to use them for the next exercise.

{10-20}




To return to this slide for a moment...

We are about to launch into the Synonym Ring, so this is a good time to reorient folks to the continuum of vocabulary control.



## The Synonym Ring

- Synonym rings are for the controlling of synonymous terms
- Predominantly used for query expansion


 5.4.2, <sup>14</sup>p. 18

Synonym rings are a set of terms that are to be treated as equivalent for the purposes of retrieval.

Synonym rings are not used during input.

Example: Batman, Dark Knight, Caped Crusader





## Synonym Ring Example 1

View results from: [Dictionary](#) | [Thesaurus](#) | [Encyclopedia](#) | [the Web](#)

*Roget's New Millennium™ Thesaurus - Cite This Source* **new!**

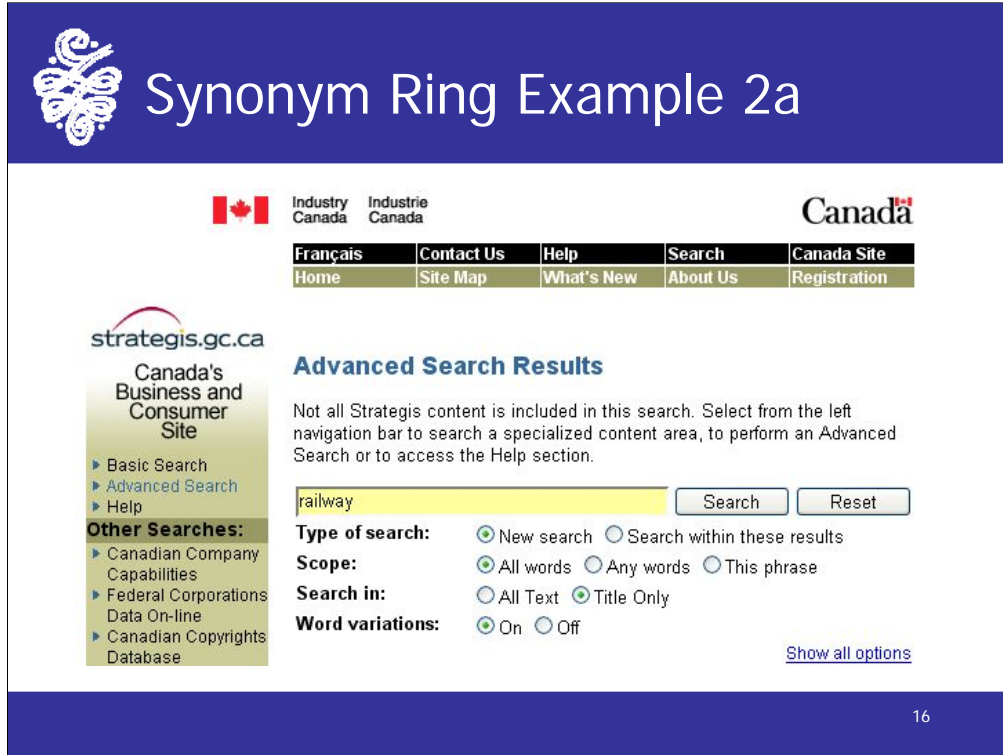
**Main Entry:** information  
**Part of Speech:** *noun*  
**Definition:** facts  
**Synonyms:** advice, ammo\*, break\*, clue, confidence, counsel, cue, data, dirt\*,  
dope\*, dossier, earful\*, enlightenment, erudition, hash, hookup,  
hot poop, illumination, info\*, inside story\*, instruction,  
intelligence, knowledge, leak, learning, lore, lowdown\*, material,  
message, network, news, notice, notification, orientation, pipeline,  
poop, propaganda, report, science, scoop, score, skinny, the  
latest, tidings, tip, tip-off, what's what\*, whole story\*, wisdom,  
word\*

*Source:* Roget's New Millennium™ Thesaurus, First Edition (v 1.3.1)  
Copyright © 2006 by Lexico Publishing Group, LLC. All rights reserved.  
\* = informal or slang

15


From Roget's Thesaurus.

<http://thesaurus.reference.com/>



The screenshot displays the 'Advanced Search Results' page on the Strategis.gc.ca website. The page features a blue header with the title 'Synonym Ring Example 2a' and a logo. Below the header, there are navigation links for 'Français', 'Contact Us', 'Help', 'Search', and 'Canada Site'. The main content area includes a search bar with the term 'railway' entered, and several search options: 'Type of search' (New search, Search within these results), 'Scope' (All words, Any words, This phrase), 'Search in' (All Text, Title Only), and 'Word variations' (On, Off). A 'Show all options' link is also present.

The search engine in the Canadian business database 'Strategis' uses an internal thesaurus to provide automatic synonym expansion. Searching on the term 'railway' [next screen]



## Synonym Ring Example 2b

- 11. Consultation on the Issuance of a Spectrum Licence to the Railway Association of Canada**  
Consultation on the Issuance of a Spectrum Licence to the Railway Association of Canada PDF, 34.8Kb Access to documents Spectrum, Information Technologies and Telecommunications EXPAND engtechnology;telecommunications;licensingConsultation on the Issuance of a Spectrum Licence to the Railway Association of Canada760  
**Strategis Site:** Spectrum Management and Telecommunications  
**Author:** Spectrum, Information Technologies and Telecommunications EXPAND  
2 Kb 2003-09-26
- 12. Annexes Rail and Urban Guided Transit Equipment Sector Competitiveness Framework**  
Similarly, the acquisition of rail equipment manufacturers in Europe and North America has given Bombardier a complete transit equipment line and better access to these markets. Bombardier is one of the few manufacturers worldwide offering a complete range of mass transit and commuter rail cars, such as rubber-tired and steel-wheeled subway cars, light rail vehicles, self-propelled gallery commuter cars, commuter train coaches, bi-level transcontinental train coaches, and monorail and high...

17

Automatically includes related terms in the search such as 'rail' and 'railroad'



## Synonym Ring Discussion

- What are the potential benefits of a Synonym Ring?
- What are the potential drawbacks?
- When would it be appropriate to use a Synonym Ring for access?

18

This is a similar discussion to the List Discussion (ss 2).

Potential benefits:

Ease of use, good for a small collection, remedies that pesky synonym problem, the benefits of query splitting

Potential drawbacks:


Still not descriptive enough for larger collections

{5-10}

# Controlled Vocabulary & Thesaurus Design


Hierarchies





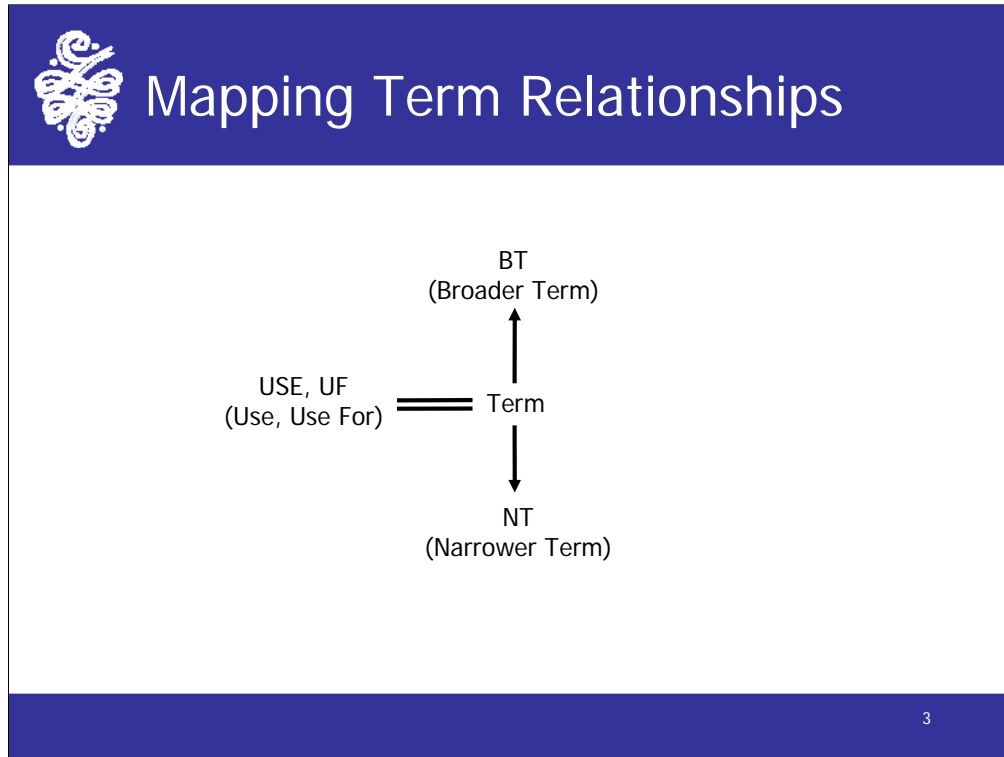
## Term Relationships

- Term equivalency
  - Groups equivalent terms
  - Establishes USE/UF relationship
    - USE Preferred term
    - UF Synonym (non-preferred term)
- Term hierarchy
  - Groups like terms under a broader term
  - Establishes BT/NT relationship
    - BT Broader term
    - NT Narrower term

 8, p. 42


So far we have worked on creating Lists and Synonym Rings, both of which are flat. We are going to move onto the next dimension of term organization -- the Hierarchy!

As examples use some of the preferred terms and synonyms from their exercise 2 work.

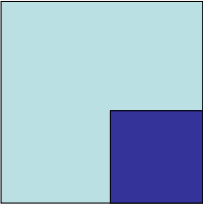
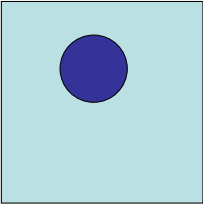
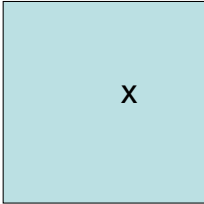



Here is the same set of relationships... in diagram form.

Since we have a solid grasp of the equivalency relationship we will move on to the hierarchy relationship.



## Hierarchical Relationships

Whole/Part BTP	Generic BTG	Instance BTI
		
BTP Heads NT Skulls	BTG Bones NT Skulls	BTI Movies NT The Skulls


 8.3, p.<sup>4</sup>46

There are 3 BT/NT relationships: whole/part, generic, and instance.

While the last one is a stretch, it is very difficult to find a term that will fit under all three of these relationships, I apologize!

Another you can write on the board is BTG Animals NT Cats, BTI Musicals NT Cats.





## Polyhierarchies

- Where the same term fits under more than one broader term

```

graph TD
    NT_Teeth[NT Teeth] --> BTP_Heads[BTP Heads]
    NT_Skulls[NT Skulls] --> BTP_Heads
    NT_Brains[NT Brains] --> BTP_Heads
    NT_Vertebrae[NT Vertebrae] --> BTG_Bones[BTG Bones]
    NT_Femurs[NT Femurs] --> BTG_Bones
    NT_Brains <--> NT_Vertebrae
  
```

8.3.4, <sup>5</sup>p. 49

We have the same narrower term fitting under two different broader terms. It is possible to incorporate a polyhierarchy into your hierarchy, but it is better to try to avoid them if at all possible, just because they are messy.

Some thesaurus construction software packages support the creation of polyhierarchies, but you would need to decide which was the primary location for the term. It is complicated but doable.

In order for a term to be successful in a polyhierarchy the narrower terms beneath the term must fit into all of the hierarchies. This term “Jade” is better suited to the use of qualifiers.



# AAT's "graffiti" Polyhierarchy


**Facet/Hierarchy Code:** V.VC

**Hierarchical Position:**

- ▲▲▲ Objects Facet
- ▲▲▲ .... Visual and Verbal Communication
- ▲▲▲ ..... Visual Works
- ▲▲▲ ..... <visual works>
- ▲▲▲ ..... <visual works by location or context>
- ▲▲▲ ..... graffiti

**Additional Parents:**

- ▲▲▲ Objects Facet
- ▲▲▲ .... Visual and Verbal Communication
- ▲▲▲ ..... Information Forms
- ▲▲▲ ..... <information forms>
- ▲▲▲ ..... <document genres>
- ▲▲▲ ..... <document genres by form>
- ▲▲▲ ..... inscriptions
- ▲▲▲ ..... graffiti



## Polyhierarchy or Qualifiers

- BT4 Materials
- BT3 Inorganic materials
- BT2 Rock
- BT1 Metamorphic rock
- Jade
  - NT Jadeite
  - NT Old jade
- Jade (materials)

- BT4 Colors
- BT3 Green
- BT2 Bluish green
- BT1 Light bluish green
- Jade
- Jade (colors)

7

One reason they are messy is because of the problems with listing the lineage of a term. Normally they narrow as they go to the top, rather than expanding.

This is a constructed example from AAT, for the term Jade, which can be a building/sculptural material or a color, among other things. Image generating a term record for Jade. Which set of broader terms should have precedence? What about the set of narrower terms that will fall under Jade. Do they fit if placed in both the Materials and Colors hierarchies?

By adding qualifiers to these terms, they can have two separate meanings, while using the same term. This sidesteps the problems of integration that we saw on the previous slide, because it is actually referring to two distinct concepts that have the same name.

The qualifier is part of the term.

This is also helpful for disambiguating at the search level. When a user searches for “jade” the returned results would include these two terms, and the user could very quickly decide which jade they were looking for, the material or the color. Getting the user to their results faster.

As you can see it also makes plain similar terms that could be assumed to be the same. For instance I was unaware that “jades” were a type of sculpture until looking up “jade” in AAT.



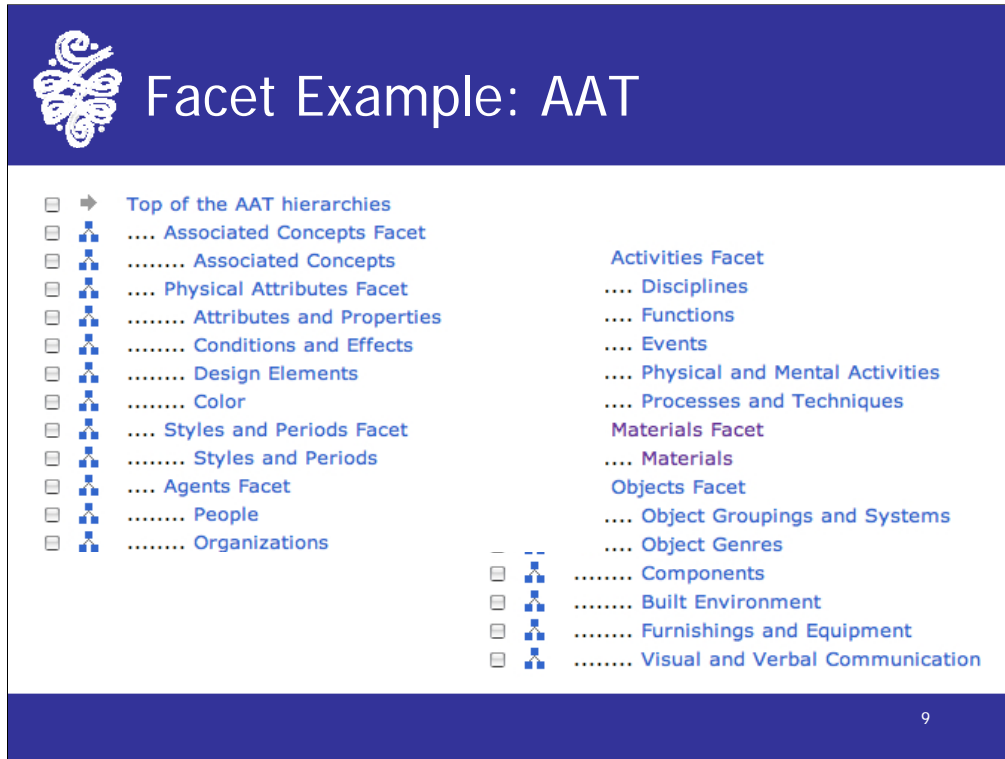
## Facet Analysis & Facets

- The division of terms into categories
- Using one principle of division at a time
- To produce groups which are
  - Homogeneous
  - Mutually exclusive
- The facets comprise a single thesaurus



5.3.4, p. 14 & 9.3.4.6, p. 72

The perfect example of this being AAT.




**Facet Example: AAT**

- Top of the AAT hierarchies
- ... Associated Concepts Facet
  - ..... Associated Concepts
  - ..... Physical Attributes Facet
  - ..... Attributes and Properties
  - ..... Conditions and Effects
  - ..... Design Elements
  - ..... Color
  - ..... Styles and Periods Facet
  - ..... Styles and Periods
  - ..... Agents Facet
  - ..... People
  - ..... Organizations
- Activities Facet
  - ..... Disciplines
  - ..... Functions
  - ..... Events
  - ..... Physical and Mental Activities
  - ..... Processes and Techniques
- Materials Facet
  - ..... Materials
- Objects Facet
  - ..... Object Groupings and Systems
  - ..... Object Genres
- Components
  - ..... Built Environment
  - ..... Furnishings and Equipment
  - ..... Visual and Verbal Communication

9

We look at AAT again some more in the future, but right now I just want to focus on the facets they have chosen.

Activities, Materials, Agents, Objects, Styles and periods, and Concepts. And to give you an idea of what they each contain the first level beneath is a Hierarchy name. Each hierarchy is then composed of terms that fit beneath them. There is some associative relationships across the facets.




## Grouping Terms

- Use of terms to groups subordinate terms


**flowers**

- annuals**
  - marigolds**
  - petunias**
  - zinnias**
- perennials**
  - dianthus**
  - coreopsis**
  - peonies**



5.3.4, p. 14 & 9.3.4.10, p. 72

In hierarchies, there are two methods of grouping terms. The first method groups similar terms under higher-level terms.



## Node Labels

- “Dummy” terms
- Can’t be assigned to objects
- Marked by italics or brackets
- In facets, highest terms are often node labels


**Visual Works**

*visual works by form*

- dioramas**
- diptychs**
- trptychs**

*visual works by function*

- cartoons**
- clip art**



8.3.5, 7.7, 8.4.3, 9.2.2.3

The second method groups similar terms under node labels.

### NODE LABELS

- a “dummy” term that is NOT assigned to documents
- sometimes a phrase
- indicates the logical basis on which a class has been divided

Start at 8.3.5 on page 51, then look at the examples on

7.7 page 41

8.4.3 page 56

9.2.2.3 page 61



## CV Scope (again)

What terms should be included in a thesaurus:

- Only terms needed to describe objects currently in the collection?
- All concepts within the domain (even if there are currently no documents in the collection on a topic)?

Discussion

- What are the benefits of each approach?
- What are the weaknesses of each approach?
- How might you decide which approach to use?



8.3.5, 7.7, 8.4.3, 9.1.2.3

This is the break between having a functional classification system and a full blown thesaurus. One is practical and the other is theoretical.

Classification

Benefits - practical, functional, less expensive to produce

Weaknesses - only describes your collection, not interoperable, etc

Thesaurus

Benefits - can become a more universal representation

Weaknesses - collaboration and agreement is expensive

How?

Review what are your institutional needs, what are your user needs, what is your budget, etc.





## Exercise 3: Hierarchy!

In your groups...

- Take the terms from Exercise 2
- Group like terms
- Develop or find Broader terms to describe the groupings

13

They should be able to put those post-its to use in physically grouping like terms.

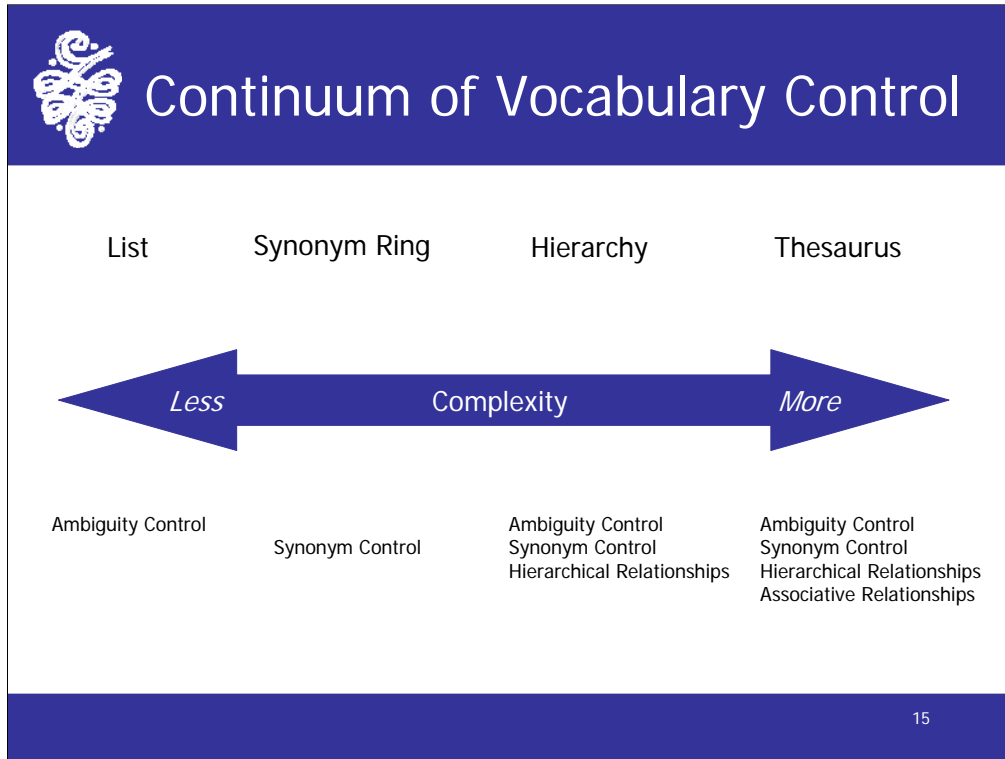


## Exercise 3: Discussion

- What hierarchies did you come up with?
- Did anyone attempt Polyhierarchies or Qualifiers?
- How did this work?
  - Was it harder than you expected?
  - Was it easier than you expected?


14

It is important that by this point they have all of the tools from the previous exercises honed, so that you can launch into thesauri with both feet on day 2.




To return to this slide for a moment...

We are about to launch into the Taxonomies and Ontologies, so this is a good time to reorient folks to the continuum of vocab control.



## Hierarchies

- Taxonomies
  - Hierarchical organizations of terms
  - Controlled for synonyms
  - Prevalent in business and web topic organization
  - Intended to reflect real world organization
- Ontologies
  - Hierarchical organization of terms
  - Not necessarily controlled for synonyms
  - Prevalent in computer science and academia
  - Intended to reflect disciplinary domain
  - Computer readable

 5.4.3, <sup>16</sup>p. 18

These are unfortunately popular buzz words at the moment, and tend to be used interchangeably. And Taxonomy tends to be used for anything that looks like a controlled vocabulary.

Taxonomies are normally a form of classification system, i.e., each concept is limited to a single location in the hierarchy

I have a nice anecdote for this or for institutional support and long term planning: Jillian worked on a Thesaurus for Nonprofit Management, and the power that be decided that Thesaurus did not work for them, since the buzz word of the day was Taxonomy. There was some major staff turn over at the hosting institution, and new leadership could not understand why we were working on Roget's. Without even asking Jillian they renamed it the Taxonomy for Nonprofit Management. While this may sound bitter, it makes sense for the institution and the business domain, it just would have been nice for them to ask.



The screenshot displays the BUSINESS.COM website interface. At the top left is a logo consisting of a stylized butterfly or flower. The main header area is dark blue with the text "Hierarchy Example 1a" in white. Below this is the BUSINESS.COM logo and the tagline "The Search Engine for Managing and Growing Your Business". A search bar is prominently featured with the text "Business Begins Here™" and a "Search" button. The main content area is divided into two columns of links under the headings "Popular Business Topics" and "Popular Products and Services".

**Popular Business Topics**


- [Internet and Online](#)
- [Advertising and Marketing](#)
- [Accounting](#)
- [Transportation and Logistics](#)
- [Starting a Small Business](#)
- [Human Resources](#)
- [Real Estate & Construction](#)
- [Telecommunications](#)
- [Financial Services](#)
- [Management](#)

**Popular Products and Services**

- [Ecommerce](#)
- [Email](#)
- [Computer Networking](#)
- [IT Services](#)
- [Knowledge Management](#)
- [Business Collaboration](#)
- [Incorporation Services](#)
- [Small Business Loans](#)
- [Businesses for Sale](#)
- [Franchise Opportunities](#)
- [Budgeting and Forecasting](#)
- [Payroll Software](#)
- [Time and Attendance](#)
- [Employee Evaluation](#)
- [Hiring and Retention](#)
- [Customer Service](#)
- [Commercial Real Estate](#)
- [Office Furniture](#)
- [Business Toll Free](#)
- [Business Cards](#)

17

As with many websites, the hierarchy is the main means of allowing the user to drill deeper into what the website offers, while providing some search mechanism for the user who knows what they are looking for.



## Hierarchy Example 1b

**Management**

- ◆ [Benchmarking & Best Practices](#)
- ◆ [Business Ethics](#)
- ◆ [Change Management](#)
- ◆ [Consulting Services](#)
- ◆ [Education & Training](#)
- ◆ [Management Theory](#)
- ◆ [Operations Management](#)
- ◆ [Reference](#)
- ◆ [Strategic Planning](#)

[More »](#)


**Media & Entertainment**

- ◆ [Amusement & Family Entertainment](#)
- ◆ [Consultants](#)
- ◆ [Education & Training](#)
- ◆ [Employment](#)
- ◆ [Film](#)
- ◆ [Music](#)
- ◆ [Photography](#)
- ◆ [Publishing](#)
- ◆ [Television](#)

[More »](#)

18

Each major topic opens up to reveal sub-topics, which presumably have their own sub-topics, and so on and so forth.



# Hierarchy Example 1c

**Media and Entertainment**  
 Information media companies and the entertainment industry, including resources for filmed entertainment, live entertainment, broadcasting and journalism.

[Directory Listings](#)
[Company Profiles](#)
[News](#)
[Web Listings](#)

**Media and Entertainment Popular Categories:**

- [Amusement & Family Entertainment](#)
- [Photography Jobs](#)
- [Home Entertainment](#)
- [Lesson Plans](#)
- [Maps and Atlases](#)
- [Music](#)
- [Publishing](#)
- [Radio](#)
- [Talent](#)
- [Television](#)
- [Visual Art & Design](#)

**More Media and Entertainment Categories:**

- [Consultants](#)
- [Education & Training](#)
- [Employment](#)
- [Events](#)
- [Film](#)
- [Journalism](#)
- [Music Industry Consultants](#)
- [News](#)
- [Online Entertainment](#)
- [Organizations](#)
- [Photography](#)
- [Printing for the Publishing Industry](#)
- [Reference](#)
- [Software](#)
- [Sports](#)
- [Streaming Media](#)
- [Technology](#)
- [Theater](#)


Sponsored Links

19

The screenshot displays the AllMusic website interface. At the top left is the AllMusic logo, a stylized butterfly. To its right is the text "Hierarchy Example 2a". Below the logo is the "allmusic" brand name, with "allmusic", "allmovie", and "allgame" listed underneath. A search bar is positioned to the right of the logo, containing a text input field, a "GO" button, and a dropdown menu labeled "Artist/Group" with an "advanced search" button. To the right of the search bar is a logo for "DATA PROVIDED BY AMG All Media Guide". Below these elements is a horizontal navigation menu with categories: ROCK, JAZZ, R&B, RAP, COUNTRY, BLUES, WORLD, ELECTRONICA, CLASSICAL, and MORE... The main content area is divided into several sections. On the left is an advertisement for Bob Dylan's new album "Modern Times" with a "PRE-ORDER" badge and a "Ticketmaster Pre-Sale Offer" link. Below the ad is the iTunes logo. In the center is the "album of the day" section, which currently shows "[no cover]". To the right is a promotional banner for "allmovie" with the text "Check out the new look!". Below the "album of the day" section is a "the whole note" section, which includes the title "AMG Now, Vol. 3" by "AMG Staff" and a short introductory paragraph. The "amg" logo is visible at the bottom right of this section. The page number "20" is located in the bottom right corner of the screenshot.

We see our major categories listed across the top, just under the omnipresent search box...





## Hierarchy Example 2b

**Jazz Styles**

<p><b>Big Band/Swing</b></p> <ul style="list-style-type: none"> <li>▪ Big Band</li> <li>▪ Progressive Big Band</li> <li>▪ Dance Bands</li> <li>▪ Swing</li> <li>▪ Ballroom Dance</li> <li>▪ Modern Big Band</li> <li>▪ Progressive Jazz</li> <li>▪ Jive</li> <li>▪ Retro Swing</li> <li>▪ Orchestral Jazz</li> <li>▪ Sweet Bands</li> <li>▪ Experimental Big Band</li> <li>▪ Continental Jazz</li> <li>▪ British Dance Bands</li> <li>▪ Society Dance Band</li> </ul> <p><b>Bop</b></p> <ul style="list-style-type: none"> <li>▪ Bop</li> <li>▪ Vocalese</li> </ul> <p><b>Cool</b></p> <ul style="list-style-type: none"> <li>▪ Cool</li> <li>▪ West Coast Jazz</li> </ul>	<p><b>Free Jazz</b></p> <ul style="list-style-type: none"> <li>▪ Modern Creative</li> <li>▪ Modern Free</li> <li>▪ Early Creative</li> <li>▪ All-Base</li> <li>▪ Free Jazz</li> <li>▪ Third Stream</li> <li>▪ Free Funk</li> <li>▪ Avant-Garde Jazz</li> </ul> <p><b>Fusion</b></p> <ul style="list-style-type: none"> <li>▪ Jazz-Pop</li> <li>▪ Contemporary Jazz</li> <li>▪ Crossover Jazz</li> <li>▪ Free Funk</li> <li>▪ Jazz-Rock</li> <li>▪ Smooth Jazz</li> </ul> <p><b>Hard Bop</b></p> <ul style="list-style-type: none"> <li>▪ Neo-Bop</li> <li>▪ Post-Bop</li> <li>▪ Hard Bop</li> <li>▪ Modal Music</li> <li>▪ Standards</li> </ul>	<p><b>Latin Jazz/World Fusion</b></p> <ul style="list-style-type: none"> <li>▪ World Fusion</li> <li>▪ Latin Jazz</li> <li>▪ Afro-Cuban Jazz</li> <li>▪ Brazilian Jazz</li> <li>▪ Big Band Latino</li> <li>▪ Cuban Jazz</li> <li>▪ African Jazz</li> </ul> <p><b>New Orleans/Classic Jazz</b></p> <ul style="list-style-type: none"> <li>▪ Boogie-Woogie</li> <li>▪ Ragtime</li> <li>▪ Dixieland</li> <li>▪ Stride</li> <li>▪ Classic Jazz</li> <li>▪ New Orleans Jazz</li> <li>▪ Mainstream Jazz</li> <li>▪ Trad Jazz</li> <li>▪ New Orleans Brass Bands</li> <li>▪ Novelty Ragtime</li> <li>▪ Dixieland Revival</li> </ul> <p><b>Soul Jazz/Groove</b></p> <ul style="list-style-type: none"> <li>▪ Jazz-Funk</li> <li>▪ Soul-Jazz</li> <li>▪ Hard Bop</li> </ul>
--	---	--

21

Some domains lend themselves well to taxonomic organization, such as music and life, for that matter.

The Linnean system for describing organisms is very well developed, and you may remember, breaking organisms into different levels of hierarchy, from Kingdom, Phylum, Class, Order, Family, Genus, Species. There are other more descriptive sub-levels as well, for instance Superclass, Class, Subclass, Infraclass, Parvclass, which are specific types of hierarchical relationships. Akin to our whole:part, instance, etc. hierarchical relationships.

**Hierarchy Example 3a**

Living Taxonomy

navigation

- Math Page
- Community portal - About
- Recent changes
- FAQ
- Help

search

Go Search

toolbox

- What links here
- Related changes
- Special pages

article | discussion | view source | history

Create an account or log in

## Main Page

**Welcome to the Living Taxonomy Project (LTP)**

The Living Taxonomy Project is a collaborative effort aimed at creating a global set of open source, standards-based taxonomies for education. The purpose of these taxonomies will be to provide a free cataloging structure for the collection and sharing of education materials around the world. We will be appending new taxonomies on a regular basis and invite our users to add edit these taxonomies as well as suggest or create new ones. A list of upcoming taxonomy releases can be found on our [What's Next](#) page, and suggestions for new taxonomies can be made on our [Suggestions](#) page.

**Taxonomy Categories**

[Humanities and Social Sciences](#) | [Business and Economics](#) | [Sciences, Engineering and Math](#) | [Education](#)

The [Sandbox](#) is a sample taxonomy, just for trying out editing the taxonomy structures.

### Articles

**Building Learning Communities with Taxonomies and Learning Objects**

It had been a hard day in the e-learning trenches, and online English composition faculty member Jill Deleuze was looking for some place to share her pain. She logged into the College of St. Justine English Department page, and perused the discussion boards and links. It was a remarkably jejune experience.... *Want to read more - check out the article*

### Announcements - Updates - Links

**Announcements**


- The Living Taxonomy project now has an associated blog. <http://livingtaxonomy.blogspot.com/> This space will be used to update the changes to the LTP as well as announce new developments and partnerships. We will also use this space for discussion, to link to other interesting projects and articles, and as a space to work through new ideas and thoughts. We plan to post to the

22


The Living Taxonomy is a wiki-based community for teachers, seemingly teachers of advanced courses for high school and introductory courses for college or community/junior college courses. The taxonomy is an attempt to organize topics for courses.

This examples digs into the taxonomy in Humanities and Social Sciences...

They have a very interesting way of getting around having associative relationships, which you will see in a few slides...



# Hierarchy Example 3b



**LIVING TAXONOMY**

navigation

- [Main Page](#)
- [Community portal - About](#)
- [Recent changes](#)
- [FAQ](#)
- [Help](#)

search

toolbox

- [What links here](#)
- [Related changes](#)
- [Special pages](#)


[Create an account or log in](#)

article
discussion
edit
history

## Humanities and Social Sciences

This area covers taxonomies for the Humanities and Social Sciences, including the arts, languages, and traditional humanities disciplines. From here you can view existing taxonomies - see the list below - edit these taxonomies, and add new taxonomies.

Use the discuss tab above for general discussion about taxonomies in the Humanities and Social Science. You can discuss taxonomies that you would like to see created, general questions about taxonomies, how taxonomies can be used in these academic areas, or other general discussion. In addition, each taxonomy page also has a discussion tab to discuss that specific taxonomy structure.



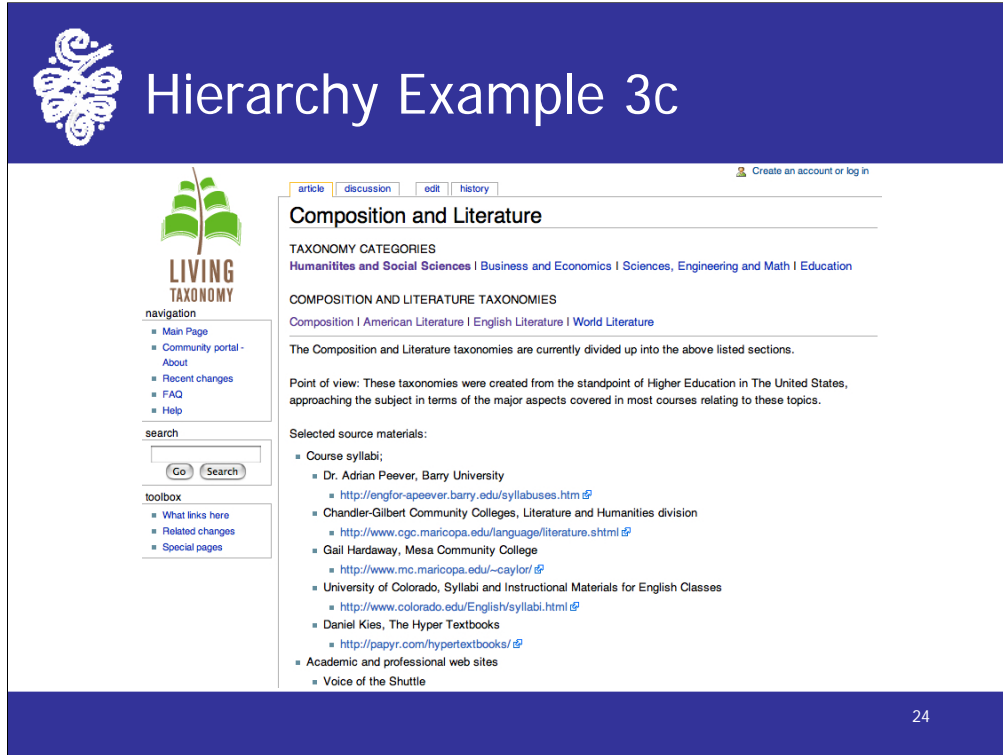
Gustav Klimt *Stoclet Frieze*

The Humanities and Social Sciences section currently holds taxonomies in these areas.

[Composition and Literature](#) | [History](#) | [Music](#) | [Sociology](#) | [Spanish](#) | [Added July 2005 Film Studies](#)  
[Added August 2005 Anthropology](#)  
[Added September 2005 Library and Information Studies](#)  
[Added September 2005 French](#)

23

Digging into Composition and Lit...

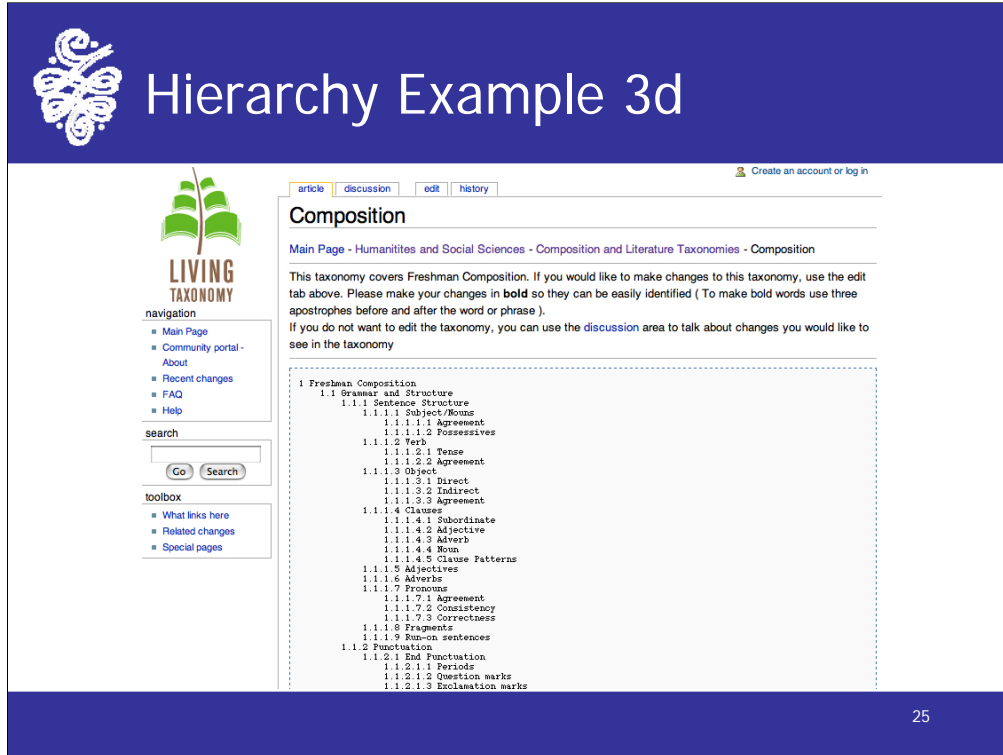


The screenshot shows a web page titled "Hierarchy Example 3c" with a blue header. Below the header is a navigation bar with tabs for "article", "discussion", "edit", and "history". The main content area is titled "Composition and Literature" and includes a "LIVING TAXONOMY" logo on the left. The page content is organized into several sections:

- TAXONOMY CATEGORIES:** Humanities and Social Sciences | Business and Economics | Sciences, Engineering and Math | Education
- COMPOSITION AND LITERATURE TAXONOMIES:** Composition | American Literature | English Literature | World Literature
- Text:** "The Composition and Literature taxonomies are currently divided up into the above listed sections." and "Point of view: These taxonomies were created from the standpoint of Higher Education in The United States, approaching the subject in terms of the major aspects covered in most courses relating to these topics."
- Selected source materials:**
  - Course syllabi:
    - Dr. Adrian Peever, Barry University (<http://englor-apeever.barry.edu/syllabuses.htm>)
    - Chandler-Gilbert Community Colleges, Literature and Humanities division (<http://www.cgc.maricopa.edu/language/literature.shtml>)
    - Gail Hardaway, Mesa Community College (<http://www.mc.maricopa.edu/~caylor/>)
    - University of Colorado, Syllabi and Instructional Materials for English Classes (<http://www.colorado.edu/English/syllabi.html>)
    - Daniel Kies, The Hyper Textbooks (<http://papyr.com/hypertextbooks/>)
  - Academic and professional web sites:
    - Voice of the Shuttle

The page also features a search bar, a navigation menu, and a toolbox on the left side.

At this level they give the literary warrant for what follows.  
We will dig into Composition and the Literature...



The screenshot shows a web page titled "Hierarchy Example 3d" with a blue header. Below the header is a navigation menu with options: article, discussion, edit, history. The main content area is titled "Composition" and includes a sub-header "Main Page - Humanities and Social Sciences - Composition and Literature Taxonomies - Composition". The text below the sub-header reads: "This taxonomy covers Freshman Composition. If you would like to make changes to this taxonomy, use the edit tab above. Please make your changes in **bold** so they can be easily identified ( To make bold words use three apostrophes before and after the word or phrase ). If you do not want to edit the taxonomy, you can use the [discussion](#) area to talk about changes you would like to see in the taxonomy". Below this text is a list of hierarchical topics for "Freshman Composition":

- 1 Freshman Composition
  - 1.1 Grammar and Structure
    - 1.1.1 Sentence Structure
      - 1.1.1.1 Subject/Nouns
        - 1.1.1.1.1 Agreement
        - 1.1.1.1.2 Possessives
      - 1.1.1.2 Verb
        - 1.1.1.2.1 Tense
        - 1.1.1.2.2 Agreement
      - 1.1.1.3 Object
        - 1.1.1.3.1 Direct
        - 1.1.1.3.2 Indirect
          - 1.1.1.3.3 Agreement
      - 1.1.1.4 Clauses
        - 1.1.1.4.1 Subordinate
        - 1.1.1.4.2 Adjective
        - 1.1.1.4.3 Adverb
        - 1.1.1.4.4 Noun
        - 1.1.1.4.5 Clause Patterns
      - 1.1.1.5 Adjectives
        - 1.1.1.6 Adverbs
      - 1.1.1.7 Pronouns
        - 1.1.1.7.1 Agreement
        - 1.1.1.7.2 Consistency
        - 1.1.1.7.3 Correctness
      - 1.1.1.8 Fragments
      - 1.1.1.9 Run-on sentences
    - 1.1.2 Punctuation
      - 1.1.2.1 End Punctuation
        - 1.1.2.1.1 Periods
        - 1.1.2.1.2 Question marks
        - 1.1.2.1.3 Exclamation marks

Note the scope note for the outline -- freshman composition.

It goes on to enumerate the various topics that should be covered as part of learning composition from the basics to the more advanced concepts.

This is a very specific organization of the information, but would it be helpful to any user? Probably not. It seems to be very well suited to the teacher of composition though.

**Hierarchy Example 3e**

**LIVING TAXONOMY**

navigation

- Main Page
- Community portal - About
- Recent changes
- FAQ
- Help

search

Go Search

toolbox

- What links here
- Related changes
- Special pages

article | discussion | edit | history

Create an account or log in

## American Literature

Main Page - Humanities and Social Sciences - Composition and Literature Taxonomies - American Literature

This taxonomy covers American Literature. If you would like to make changes to this taxonomy, use the edit tab above. Please make your changes in **bold** so they can be easily identified ( To make bold words use three apostrophes before and after the word or phrase ).

If you do not want to edit the taxonomy, you can use the **discussion** area to talk about changes you would like to see in the taxonomy.

### 1.2 American Literature

- 1.2.1 Native American Literary and Oral Traditions
  - 1.2.1.1 Navajo
  - 1.2.1.2 Iroquois
  - 1.2.1.3 Tlinglit
  - 1.2.1.4 Tsimshian
  - 1.2.1.5 Zuni
  - 1.2.1.6 Arapaho
  - 1.2.1.7 Sioux
- 1.2.2 Exploration and Early Colonization
  - 1.2.2.1 Outside Influences
    - 1.2.2.1.1 History
      - 1.2.2.1.1.1 English vs. Spanish Colonization
    - 1.2.2.1.2 Cultural
      - 1.2.2.1.2.1 Survival in the Wilderness
      - 1.2.2.1.2.2 Growth of American Identity
    - 1.2.2.1.3 Religion
      - 1.2.2.1.3.1 Exclusiveness - the Right Way to God
      - 1.2.2.1.3.2 Interaction between Religious Groups
  - 1.2.2.2 Literary Genres
    - 1.2.2.2.1 Explorations
    - 1.2.2.2.2 Journals
    - 1.2.2.2.3 Letters
  - 1.2.2.3 Authors
    - 1.2.2.3.1 William Bradford
    - 1.2.2.3.2 Anne Bradstreet
    - 1.2.2.3.3 Christopher Columbus
    - 1.2.2.3.4 Samuel De Champlain
    - 1.2.2.3.5 Thomas Harriot
    - 1.2.2.3.6 Thomas Horton
    - 1.2.2.3.7 John Smith

26

Now we have jumped to the American Literature “taxonomy”, again there is a good scope note.

What gets a little crazy here are the listings of authors under a given subject area -- Exploration and Early Colonialization.

Is this scalable? As information professionals, does this seem like a good way to include authors as a part of their information structure?

It gives me the heeby-jeebies, but only because I want to put all of the authors for that subject under that subject in their system. But because their scope is for introductory, freshman classes, it is likely that for a given topic only a few, good examples of a genre would be included.

Next slide is further down in the same taxonomy...



## Hierarchy Example 3f

```

1.2.4.2.6 John Woolman
1.2.5 Age of Reason and Revolution
1.2.5.1 Outside Influences
1.2.5.1.1 History
1.2.5.1.1.1 Declaration of Independence
1.2.5.1.1.2 Revolutionary War
1.2.5.1.2 Culture
1.2.5.1.2.1 The Enlightenment
1.2.5.1.2.2 Rationalism
1.2.5.1.2.3 Rise of Nationalism
1.2.5.2 Literary Genres
1.2.5.2.1 Autobiography
1.2.5.2.2 Biography
1.2.5.2.3 Histories
1.2.5.2.4 Journals
1.2.5.2.5 Letters
1.2.5.2.6 Political Treatises
1.2.5.2.7 The Federalist
1.2.5.3 Authors
1.2.5.3.1 John Adams and Abigail Adams Letters
1.2.5.3.2 Joel Barlow
1.2.5.3.3 Charles Brockden Brown
1.2.5.3.4 Jonathan Edwards
1.2.5.3.5 Olaudah Equiano
1.2.5.3.6 Benjamin Franklin
1.2.5.3.7 Philip Freneau
1.2.5.3.8 Thomas Jefferson
1.2.5.3.9 Sarah Wentworth Morton
1.2.5.3.10 Thomas Paine
1.2.5.3.11 Annis Boudinot Stockton
1.2.5.3.12 Royall Tyler
1.2.5.3.13 Phillis Wheatley
1.2.6 The Nineteenth Century to 1865
1.2.6.1 Outside Influences
1.2.6.1.1 History
1.2.6.1.1.1 Democracy
1.2.6.1.1.2 Social Reform
1.2.6.1.1.3 Agricultural Economy
1.2.6.1.2 Culture
1.2.6.1.2.1 Humanitarianism
1.2.6.1.2.2 Romanticism
1.2.6.1.2.3 Traditionalism
1.2.6.1.2.4 Transcendentalism
1.2.6.2 Literary Genres
1.2.6.2.1 Autobiography
1.2.6.2.2 Biography
1.2.6.2.3 Histories
1.2.6.2.4 Journals
1.2.6.2.5 Letters
1.2.6.2.6 Oratory
1.2.6.2.7 Political Treatises
1.2.6.2.8 Romance of Nature
1.2.6.3 Authors

```

27

These author lists are never very long -- the one in 1.2.5.2 is only 15 authors long.

This is a very interesting application of the taxonomy to material.



## Hierarchy Discussion

- What are the potential benefits?
- What are the potential drawbacks?

28

Discussion should cover the idea that the information is becoming more structured and subtle. There are also benefits to revealing layers of terms instead of a giant bucket of terms at the same time. It is intended that hierarchies should be able to make explicit, implicit relationships. Thus the user can better understand the domain just by reading the hierarchy (i.e., taxonomy or ontology).

Potential drawbacks are the added complexity might be overkill. It also does not leave room for anything but these broader/narrower and equivalency relationships. And they are that much more time intensive.





## Wrapping Up Day 1

- The role of Controlled Vocabulary tools
- The spectrum of Controlled Vocabulary tools
- Term selection and Lists
- Term refinement and Synonym Rings
- Organizing terms with hierarchy
- Any questions?

29

We have covered a lot today, and there is still much more to cover tomorrow, but we would like to make sure that everything is sinking in so that we can build off of it tomorrow. So this is your opportunity to clear up any confusion, and if you have some confusion I guarantee that you are not the only one, so please do not be shy about asking questions at this point and we will do our best to answer them, or at least make note of them so that we can address them in the future...

Has everything been clear?

Is there anything that needs clarification?

Speak now or hold your peace until morning...


If there is anytime left at this point you might want to find out if anyone can see using lists, synonym rings, taxonomies or ontologies for their CV needs?

*THIS PAGE INTENTIONALLY LEFT BLANK  
FOR DOUBLE SIDED COPY*

# Controlled Vocabulary & Thesaurus Design

Associative Relationships & Thesauri





## Today's Goals

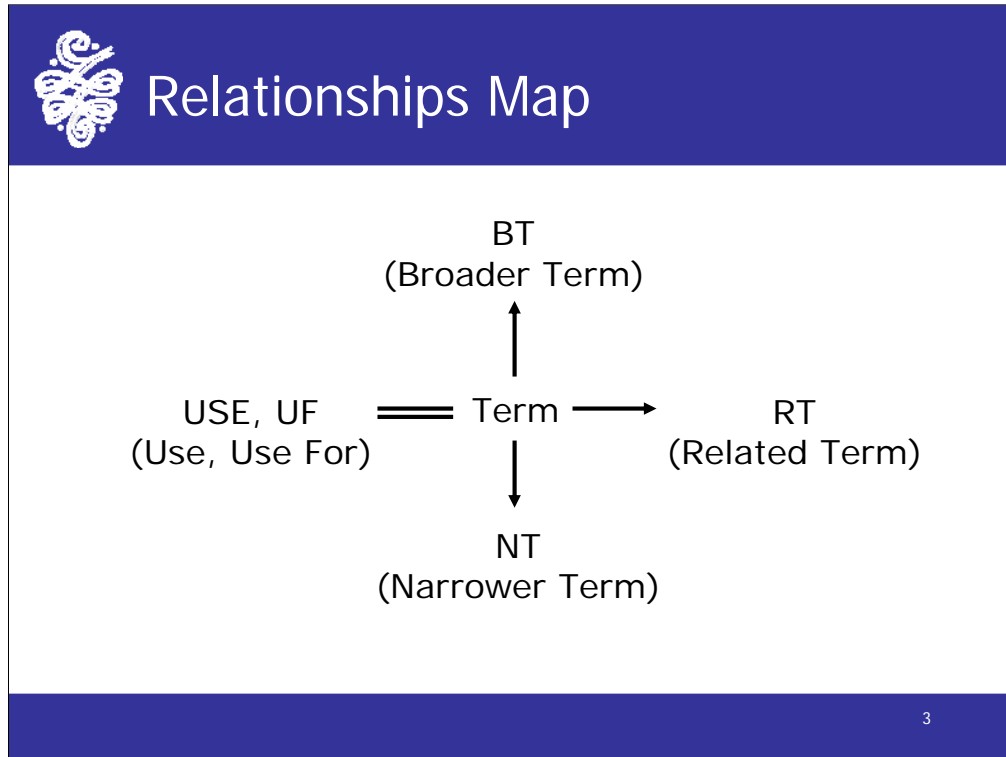
- Questions about yesterday's sessions
- Associative relationships and Thesauri
- CV display and navigation
- CV planning and maintenance
- CV resources and future directions

2

Feel free to mention some of the things that went on yesterday, which have set the stage for the work today.

If people have unanswered questions from yesterday, that will be addressed today, now is a good time to bring them up again. And if you have answers to questions from yesterday that can be taken care of before you start, now is the time.

Now is also the time to see if anyone has any questions from yesterday's material that were not asked at the end of the day yesterday.





## Associative Relationship

- “the most difficult [relationship type] to define” (Z39.19)
- Semantic or conceptual relationship
- Used on the grounds that their inclusion may suggest additional descriptors for indexing or retrieval.
- Overlapping sibling terms (8.4.1.1)
- Terms in different hierarchies (8.4.2)



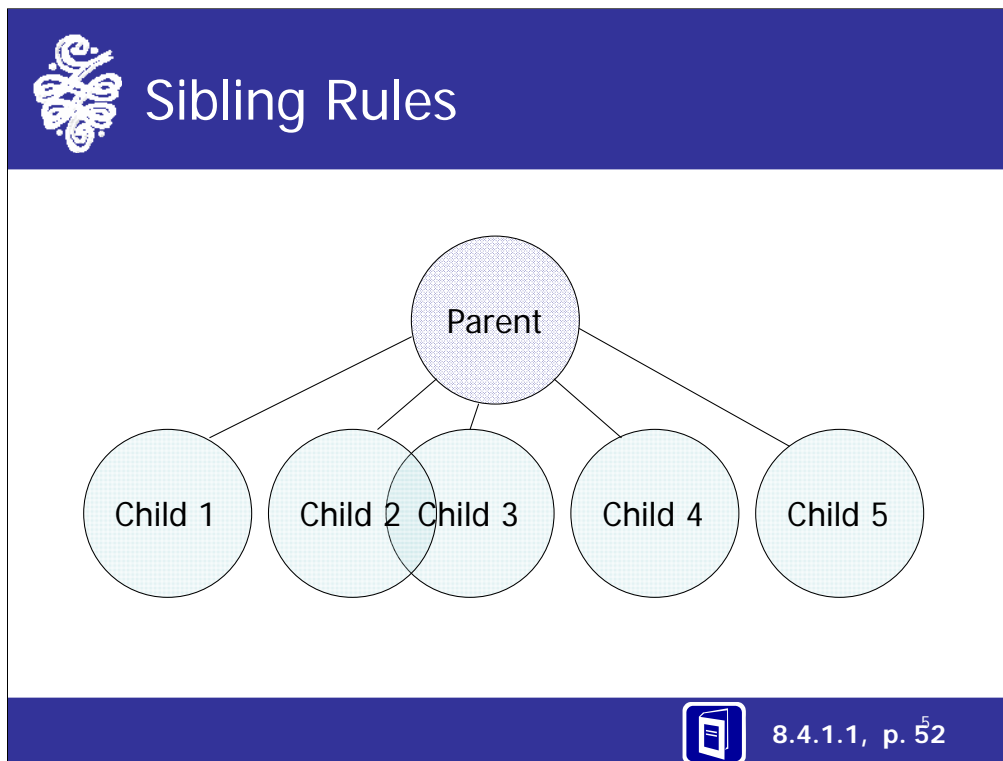
8.4, p.<sup>4</sup>51

**Optional Activity/Discussion:** Can participants think of other types of relationships among resources?

How is a dog related to a house? (lives at, owned by, protects, inside/outside...)

There is almost no end to the types of relationships that resources may have with one another, and this is worthwhile to bring out, because they are often inherent in a catalog, and in online catalogs often take the form of active hyperlinks to indexes. For example: subject relationships (resource about the same topic, person, place, or thing), author relationships (resource created or otherwise affected by the same person or corporate body), etc.

Discuss semantic vs. user focus – both have their roles



A potentially confusing problem...

Siblings are already related through the parent, but if their scope overlaps or is in some way entangled with another sibling, then they can be RTs as well.


Example:

The Chrysler 300, Dodge Charger and Dodge Magnum are all cars made by Chrysler.

However, they also are all built on the same frame, the Chrysler LX platform.

Should their relationship be included using RT's? (no correct answer – what do the users need/expect)

Example in Z39.19 using vehicles, ships, boats



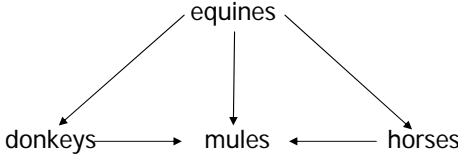
## Derivational Relationships

**donkeys**  
 BT equines  
 RT mules

**equines**  
 NT donkeys  
 NT horses  
 NT mules


**horses**  
 BT equines  
 RT mules

**mules**  
 BT equines  
 RT donkeys  
 RT horses



```

    graph TD
      equines --> donkeys
      equines --> mules
      equines --> horses
      donkeys --> mules
      horses --> mules
    
```


8.4.1.3, p. 52

Look for instances of a sibling that is derived from two or more siblings.  
 These relationships usually require RTs.





## Hierarchy & Associativity


- Hierarchy depth rules
  - Try to be as specific (deep) as possible
  - Try to explicate the relationship
- Inverse relationship between
  - Depth of hierarchy
  - Number of associative relationships
  - Eg. A flat thesaurus will have many more associative relationships



11.1.7, p. 94


The depth rules might need to be explained, but it is also something that students will be better at understanding once they have tried to develop some of their own associative relationships and come across the problem of where to establish the relationship.

Try to explicate the relationship = try to put it into a sentence



## Additional Associate Relationships

- Relationships between terms belonging to *different hierarchies*
  - Process - Agent
  - Process – Counteragent
  - Action – Property
  - Action – Product
  - Action – Target
  - Cause – Effect
  - Concept or Object – Property
  - Concept or Object – Origins
  - Concept or Object – Units or Mechanisms of Measurement
  - Raw Material – Product
  - Discipline or Field = Object or Practitioner
  - False Common Parent
  - Antonyms



8.4.1, p. 42 & 8.4.2<sup>8</sup>, p. 53

These types of relationships often are listed as RTs.

Review some of the examples shown in Examples 119 – 130.

Point out the summary table in 8.4.1 on page 42.



## Administrative Information

- Includes
  - Scope note: defining the scope of use for the term
  - Creator: the person who entered the term
  - Creation date: date of entry
  - Language: term language
  - Modified by: last modifier of the term record
  - Modified date: term last modified on date
- Keeps track of your decisions for you and others!



11.1.4, 11.3 & 9.3.3

It really is best practices to keep on top of this information. Some of it can be collected automatically by the thesaurus software, such as creation and modification information. Other info is clearly your responsibility though, such as the scope note and language, especially if working on a multilingual vocabulary.



## Exercise 4: Associate

Given your groups and terms from yesterday...

- Establish as many associative relations as possible
- Revise your scope notes, hierarchy, etc when necessary

10

Using the term records from the previous example, and using string to attach them with dental floss and tape to cross the hierarchies.

{10-20}

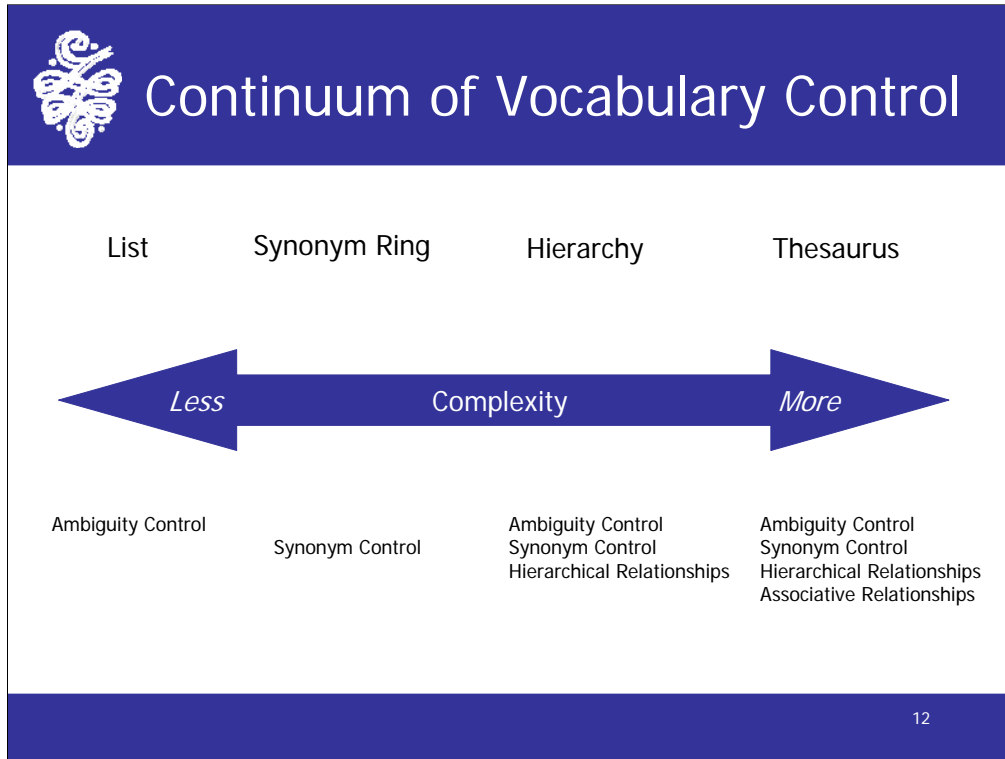


## Exercise 4: Discussion

- What associations did you establish?
- Questions?
- Problems?

11

This is one of the most frustrating activities, and hopefully it will produce some interesting discussion about the interplay between hierarchy and association.



To return to this slide for a moment...

We are about to launch into the Thesaurus, so this is a good time to reorient folks to the continuum of vocab control.



## Thesauri

- A tool for vocabulary control of a specific subject domain
- It contains:
  - Preferred terms and non-preferred terms
  - The semantic relations between terms
  - Rules for use and other administrative information
- It presupposes:
  - A particular collection of documents
  - A particular group of users



5.4.4,<sup>13</sup> p. 18

The screenshot shows the website **thesaurus.maths.org**. At the top, there is a navigation menu with links for Home, Preferences, Help, UserGuides, Credits, Download, and Forum. Below this is a search bar with the text "Find word or phrase" and a dropdown menu set to "containing". The search bar is set to "English" and has a "Go" button. The main content area is titled "Connecting Mathematics" and contains the following text:

This site contains brief explanations of mathematical terms and ideas in English, Danish, Finnish, Hungarian, Lithuanian, Polish, Slovak. We have also made a start on Spanish. Go to the preferences page to choose your language(s) and level of explanation.

Mathematical ideas are linked to others using 5 types of link:

- Broader**  
A link to a broader - more general or more encompassing - idea.
- Narrower**  
A link to a narrower - more specific - term.
- References**  
A link to a term used in explaining this idea.
- ReferencedBy**  
Links to ideas which build on this one
- See Also**  
Just that - a link to something relevant to this term.

You can either navigate the site by following the links on each page, or, if you have the Sun Java plugin, you can use our graph browser. This allows you to see ideas that

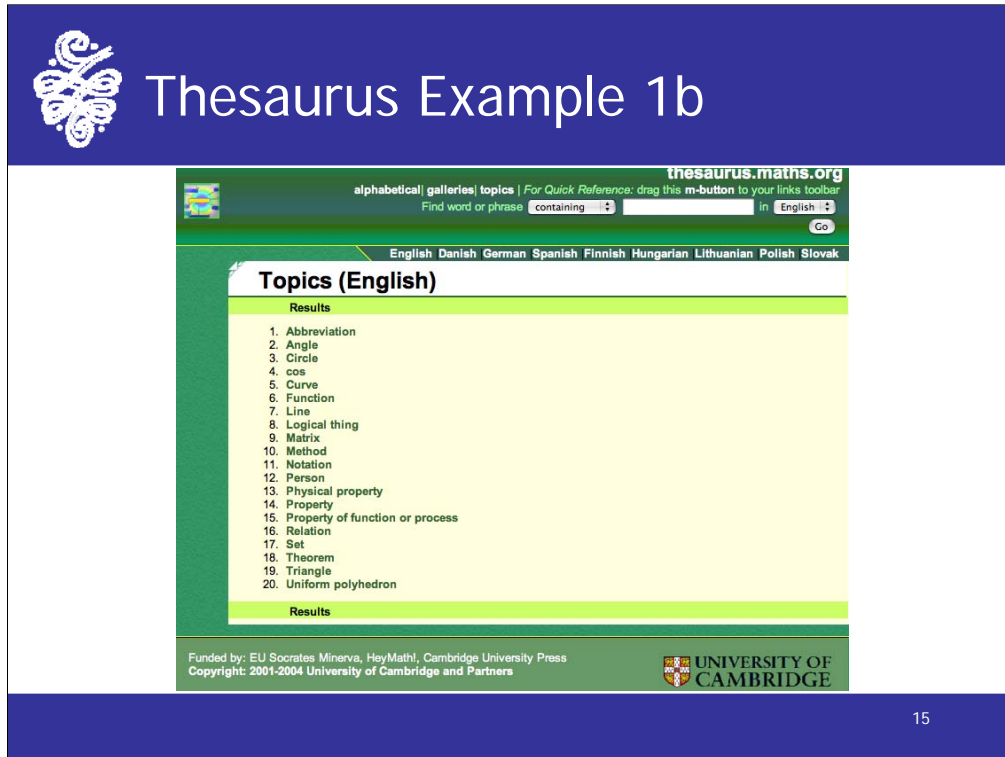
You will find thumbnails linked to all the illustrated entries in these media galleries:

- Image Gallery**  
Various images. No plugin required.
- Flash Gallery**  
Flash animations. Needs the Flash Player plugin.
- Cinderella Gallery**  
Dynamic Geometry diagrams. Needs Sun Java plugin.
- Live3D Gallery**  
3D viewer. Needs Sun Java plugin.
- Java3D Gallery**  
3D viewer. Needs Sun Java plugin with java3D extension.
- Shockwave Gallery**  
3D viewer. Needs the Shockwave player
- VRML Gallery**  
Virtual Reality 3D viewer. Needs the Cortona VRML plugin.

The page number "14" is visible in the bottom right corner of the screenshot.

<http://thesaurus.maths.org/>





The screenshot shows the website **thesaurus.maths.org**. At the top left is a logo of a stylized butterfly. The main title is "Thesaurus Example 1b". The website interface includes a search bar with the text "Find word or phrase" and a dropdown menu set to "containing". The language is set to "English". Below the search bar, there are language options: English, Danish, German, Spanish, Finnish, Hungarian, Lithuanian, Polish, and Slovak. The main content area is titled "Topics (English)" and lists 20 results:

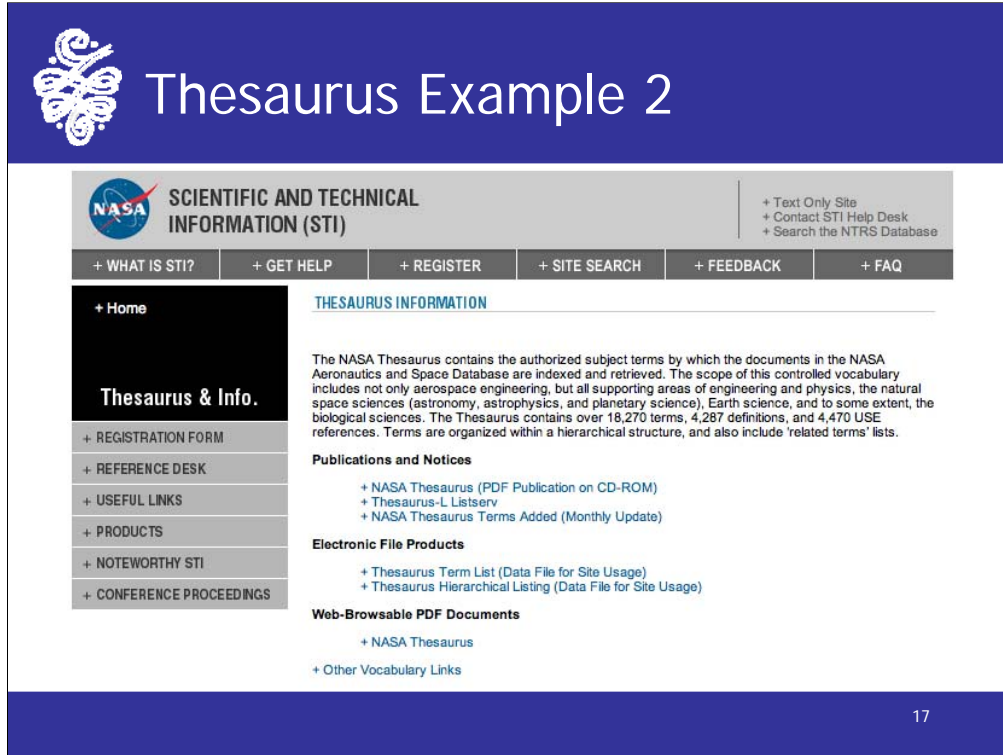
1. Abbreviation
2. Angle
3. Circle
4. cos
5. Curve
6. Function
7. Line
8. Logical thing
9. Matrix
10. Method
11. Notation
12. Person
13. Physical property
14. Property
15. Property of function or process
16. Relation
17. Set
18. Theorem
19. Triangle
20. Uniform polyhedron

At the bottom of the page, there is a footer with the text: "Funded by: EU Socrates Minerva, HeyMath!, Cambridge University Press" and "Copyright: 2001-2004 University of Cambridge and Partners". To the right of the footer is the logo of the "UNIVERSITY OF CAMBRIDGE".

The screenshot shows the website **thesaurus.maths.org** with a search for "Antiderivative". The page features a navigation bar with links like Home, Preferences, Help, UserGuides, Credits, Download, and Forum. Below the search bar, there are language options: English, Danish, German, Spanish, Finnish, Hungarian, Lithuanian, Polish, and Slovak. The main content area is titled "Antiderivative (English)" and includes a search bar with "Antiderivative" entered. Two definitions are provided, each with a level indicator circled in red: "Definition (advanced level)" and "Definition (undergraduate level)". A "Relations" section on the right, also circled in red, lists "broader: (en) Integral" and "references: (en) Derivative". A "Show graph Requires Java" button is visible. The footer contains funding information and the University of Cambridge logo.

16

It should be noted that this thesaurus essentially has two sets of scope notes, one for beginners to mathematics and more advanced mathematicians.



The screenshot shows the NASA Scientific and Technical Information (STI) website's Thesaurus page. At the top left is the NASA logo and the text "SCIENTIFIC AND TECHNICAL INFORMATION (STI)". To the right of the logo are three links: "+ Text Only Site", "+ Contact STI Help Desk", and "+ Search the NTRS Database". Below this is a navigation bar with six items: "+ WHAT IS STI?", "+ GET HELP", "+ REGISTER", "+ SITE SEARCH", "+ FEEDBACK", and "+ FAQ".

On the left side, there is a vertical menu with a "Home" link at the top. Below it is a "Thesaurus & Info." section with a list of links: "+ REGISTRATION FORM", "+ REFERENCE DESK", "+ USEFUL LINKS", "+ PRODUCTS", "+ NOTEWORTHY STI", and "+ CONFERENCE PROCEEDINGS".

The main content area is titled "THESAURUS INFORMATION". It contains a paragraph describing the NASA Thesaurus, followed by three sections: "Publications and Notices" with links to a PDF publication, Listserv, and monthly updates; "Electronic File Products" with links to term lists and hierarchical listings; and "Web-Browsable PDF Documents" with links to the thesaurus and other vocabulary links.

The page number "17" is located in the bottom right corner of the screenshot.

<http://www.sti.nasa.gov/thesfrm1.htm>

**Thesaurus Example 3a**

**The Cook's Thesaurus**

The *Cook's Thesaurus* is a cooking encyclopedia that covers thousands of ingredients and kitchen tools. Entries include pictures, descriptions, synonyms, pronunciations, and suggested substitutions.

**Categories**

- Vegetables** - includes roots, tubers, leafy greens, squash, and sea vegetables.
- Fruits** - includes berries, citrus fruit, melons, tropical fruit, and tomatoes.
- Dairy** - includes milk, cream, cheese, eggs, and cultured milk products
- Flavorings** - includes sweeteners, herbs, spices, chocolate, and extracts.
- Liquids** - includes alcoholic beverages, stocks, juices, and vinegar.

**Ingredient of the Month**

Braised red cabbage goes well with apples, onions, nuts, and red wine, and it makes an excellent bed for pork, duck, and salmon. Choose firm, shiny heads that are as dark as possible. Don't overcook red cabbage or you'll sacrifice flavor and vitamins.

18

Instead of associative relationships we have substitution relationships!

<http://www.foodsubs.com/>



## Thesaurus Example 3b

### Vegetables Category

---



Vegetables is a catch-all category that includes many of the edible parts of a plant, like stems, roots, flowers, and leaves. We don't usually consider the fruits of a plant to be vegetables, except for fruits that aren't very sweet. Tomatoes, squash, peppers, eggplants, and beans, for example, are all fruits, but we usually refer to them as vegetables.

- [Roots](#)
- [Tubers](#), including [potatoes](#) and [sweet potatoes & yams](#).
- [Stalk Vegetables](#)
- [Onions](#), including [green onions & leeks](#) and [dry onions & shallots](#).
- [Garlic](#)
- [Ginger & Other Rhizomes](#)
- [Cabbages](#)



## Thesaurus Example 3c

### Sweet Potatoes & Yams

batata

batata dulce

batiste



**boniato** = tropical sweet potato = Cuban sweet potato = white sweet potato = white-fleshed sweet potato = **batiste** = **batata** = **batata dulce** = **camote** *Notes:* Boniatos aren't as sweet and moist as other sweet potatoes, but many people prefer their fluffier consistency and more delicate flavor. Store them at room temperature and use them soon after your purchase them, since they tend to spoil quickly. *Substitutes:* sweet potato (This is sweeter and moister than a boniato.) OR yam (moister) OR russet potato


camote

Cuban sweet potato


cush-cush See tropical yam.

greater yam See tropical yam.

igname *Pronunciation:* EE-nyahm



# Thesaurus Example 4



ICPSR INTER-UNIVERSITY CONSORTIUM FOR POLITICAL AND SOCIAL RESEARCH

Need Help?

About ICPSR Data Access & Analysis Data Use Tutorial Depositing Data Education & Instruction Research at ICPSR Special Topic Archives

MyData: Login/Account Info - Data Cart - Logout ICPSR Home - Contact Us - OR Site - Search

## Thesaurus

The ICPSR Thesaurus is a controlled vocabulary system composed of three separate lists: Subject Thesaurus, Personal Names Authority List, and Geographic Names Thesaurus. Brief descriptions and links to each are provided below.

- > [Subject Thesaurus](#) is an alphabetical listing of social science subject terms. The scope of this thesaurus is multidisciplinary and is intended to reflect the subject range of the ICPSR archive. Social science disciplines represented include: political science, sociology, history, economics, education, criminal justice, gerontology, demography, public health, law, and international relations.
- > [Personal Names Authority List](#) provides controlled name forms and spellings. Unlike the Subject and the Geographic lists, it is non-hierarchical (there are no broader/narrower, or related term designations). This list is designed to provide name form consistency and, through the use of a "description" (DSC) note field, enable users to distinguish between same or similar names and to determine whether the person/name is appropriate to their indexing or searching needs.
- > [Geographic Names Thesaurus](#) provides guidance on the preferred and/or current names of geographic and geopolitical entities. It is hierarchically arranged using broader/narrower notation, allowing one to navigate the relationships among region, continent, country, and in some cases, cities. Scope notes briefly detailing the history of country name changes are included along with instructions and restrictions for their use and application.

21



# Thesaurus Example 5a



## Chocolate Thesaurus

IS 463 Indexing and Thesaurus Construction, Fall 2004

[Home >>](#)

**Introduction**

[Purpose](#)

[Construction](#)

[Guide to Application](#)

[Updating & Enhancing](#)

**Strengths & Weaknesses**

[Analysis](#)

**Appendices**

[Assignment](#)

[Credits & Resources](#)


**Thesaurus**

[Alphabetical](#)


[Hierarchical](#)







# Thesaurus Example 5b



## Chocolate Thesaurus

IS 463 Indexing and Thesaurus Construction, Fall 2004


A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

<ul style="list-style-type: none"> <li>Home</li> <li><b>Introduction</b></li> <li>Purpose</li> <li>Construction</li> <li>Guide to Application</li> <li>Updating &amp; Enhancing</li> <li><b>Strengths &amp; Weaknesses</b></li> <li>Analysis</li> <li><b>Appendices</b></li> <li>Assignment</li> <li>Credits &amp; Resources</li> <li><b>Thesaurus</b></li> </ul>	<p><b>C</b></p> <p><b>CACAO BEANS</b>  <i>cacao butter</i> (use: COCOA BUTTER)  <i>Cacao de Zaan</i> (use: NETHERLANDS, THE)  <i>cacao fruits</i> (use: CACAO PODS)  <i>cacao nibs</i> (use: CACAO BEANS)  <b>CACAO PODS</b>  <i>cacao powder</i> (use: COCOA POWDER)  <i>cacao seeds</i> (use: CACAO BEANS)  <b>CACAO TREES</b>  <i>Cadbury</i> (use: UNITED KINGDOM)  <i>Cafe-Tasse</i> (use: BELGIUM)  <i>Caffarel</i> (use: ITALY)  <b>CAKES</b>  <i>Callebaut</i> (use: BELGIUM)  <i>Camille Bloch</i> (use: SWITZERLAND)  <b>CANADA</b>  <i>candies</i> (use: CONFECTIONS)  <b>CARAMEL</b>  <b>CAROB</b>  <b>CASHEWS</b></p>	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 15%;">Current Term</td> <td>CACAO BEANS</td> </tr> <tr> <td><b>UF</b></td> <td>cacao nibs</td> </tr> <tr> <td></td> <td>cacao seeds</td> </tr> <tr> <td></td> <td>cocoa nibs</td> </tr> <tr> <td></td> <td>cocoa seeds</td> </tr> <tr> <td><b>RT</b></td> <td>CACAO PODS</td> </tr> <tr> <td></td> <td>CACAO TREES</td> </tr> <tr> <td></td> <td>CHOCOLATE LIQUOR</td> </tr> <tr> <td></td> <td>CLEANING</td> </tr> <tr> <td></td> <td>COCOA BUTTER</td> </tr> <tr> <td></td> <td>COCOA POWDER</td> </tr> <tr> <td></td> <td>CRIOLLO</td> </tr> </table>	Current Term	CACAO BEANS	<b>UF</b>	cacao nibs		cacao seeds		cocoa nibs		cocoa seeds	<b>RT</b>	CACAO PODS		CACAO TREES		CHOCOLATE LIQUOR		CLEANING		COCOA BUTTER		COCOA POWDER		CRIOLLO
Current Term	CACAO BEANS																									
<b>UF</b>	cacao nibs																									
	cacao seeds																									
	cocoa nibs																									
	cocoa seeds																									
<b>RT</b>	CACAO PODS																									
	CACAO TREES																									
	CHOCOLATE LIQUOR																									
	CLEANING																									
	COCOA BUTTER																									
	COCOA POWDER																									
	CRIOLLO																									


23

Point out use of:

- ALL CAPS (or bold)
- use of an alphabetical listing
- UF & RT relationships listed



# Thesaurus Example 5c



## Chocolate Thesaurus

IS 463 Indexing and Thesaurus Construction, Fall 2004

---

- Home
- Introduction**
- Purpose
- Construction
- Guide to Application
- Updating & Enhancing
- Strengths & Weaknesses**
- Analysis
- Appendices**
- Assignment
- Credits & Resources
- Thesaurus**
- Alphabetical
- Hierarchical >>

**HIERARCHICAL THESAURUS**

- Regions
- Natural Sources
- Types
- Chocolate Goods
- Processes
- Associated Ingredients

24

## Use of Facets



## Thesaurus Discussion

- What are the potential benefits of a thesaurus?
- What are the potential drawbacks?

25

The thesaurus is a very informative document. **It is a blueprint for understanding the nuances of a domain.**

They are rather **expensive** to develop! And many users will not know what to do with them, unless they are hooked up in the background of their finding activities.


*THIS PAGE INTENTIONALLY LEFT BLANK  
FOR DOUBLE SIDED COPY*

# Controlled Vocabulary & Thesaurus Design

Display & Navigation




This is a big session! Be warned!



## General Text Guidelines

- Font and capitalization
  - Bold or all caps for preferred terms
  - Italics or lower case for non-preferred terms
- Indentation conventions
- Order of relationships
  - (e.g. BT, then NT, then RT), or by facet
- Be consistent!

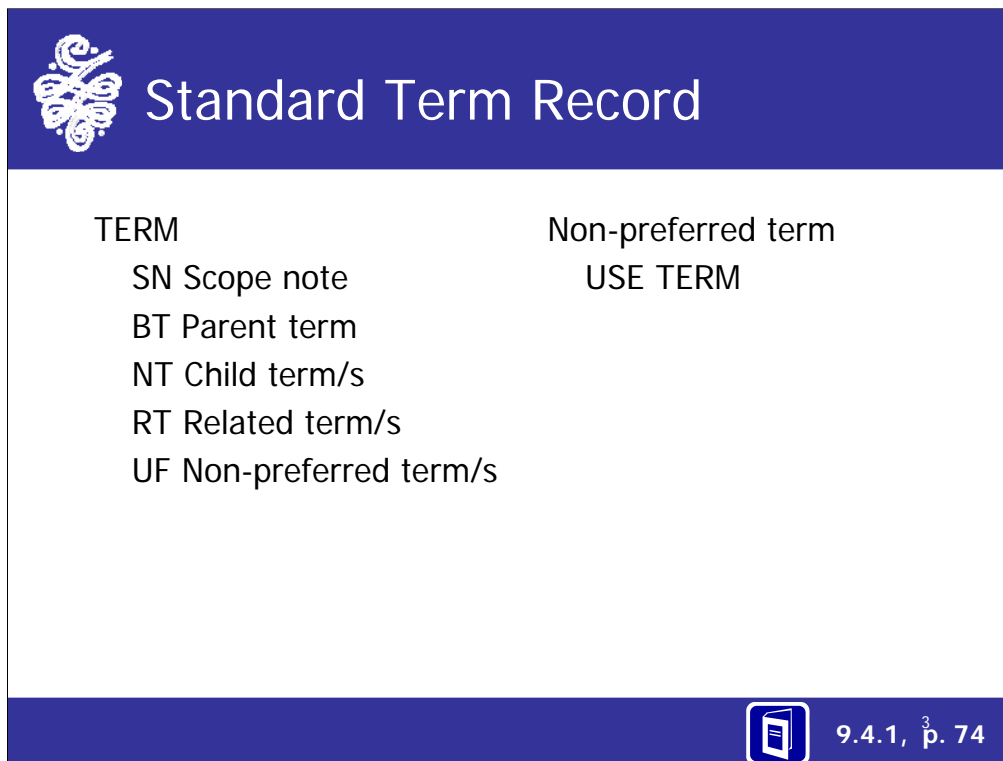
 9.2.3 & 9.2.4, <sup>2</sup>p. 61

Hierarchy can be implied through text style and indents.

The Term is the title for a term record, and should be the most visible.

How you format your text will tell the user about the content and gives you as the designer more tools to delineate between the preferred term and the attendant information for that term, or between the relationship qualifier and the term being related.

Keep in mind, these are general guideline, really you can format your text how you like. If you feel uncomfortable with setting up your own system of style and format, you can always borrow from someone else. Probably the most important is to be consistent! Once your user gets a feel for what the various styles mean it is not nice to start switching them around!



The slide features a dark blue header with a white decorative logo on the left and the title "Standard Term Record" in white text. Below the header is a white rectangular area containing two columns of text. The left column is headed "TERM" and lists five types of relationships: SN Scope note, BT Parent term, NT Child term/s, RT Related term/s, and UF Non-preferred term/s. The right column is headed "Non-preferred term" and contains the text "USE TERM". At the bottom right of the slide, there is a small icon of a book and the text "9.4.1, <sup>3</sup>p. 74".

This is the standard for a term record and a non-preferred term, for printed term records. Some controlled vocabularies still maintain this as standard, but you do what is necessary.

**IMPORTANT! What follows are the display options for controlled vocabularies. We are going to discuss a type of display and then look at some examples for each type. At the end we are going to discuss things that we liked and didn't like, that you could see your users taking advantage of or balking at, so please make a note of displays you liked or any other thoughts about the displays. So let us think critically about each display you see, and please do not hesitate to ask me to stop on a slide. We have screen shots of each slide, but if you are really interested in a specific display, we could try to see it in it's native habitat (web).**

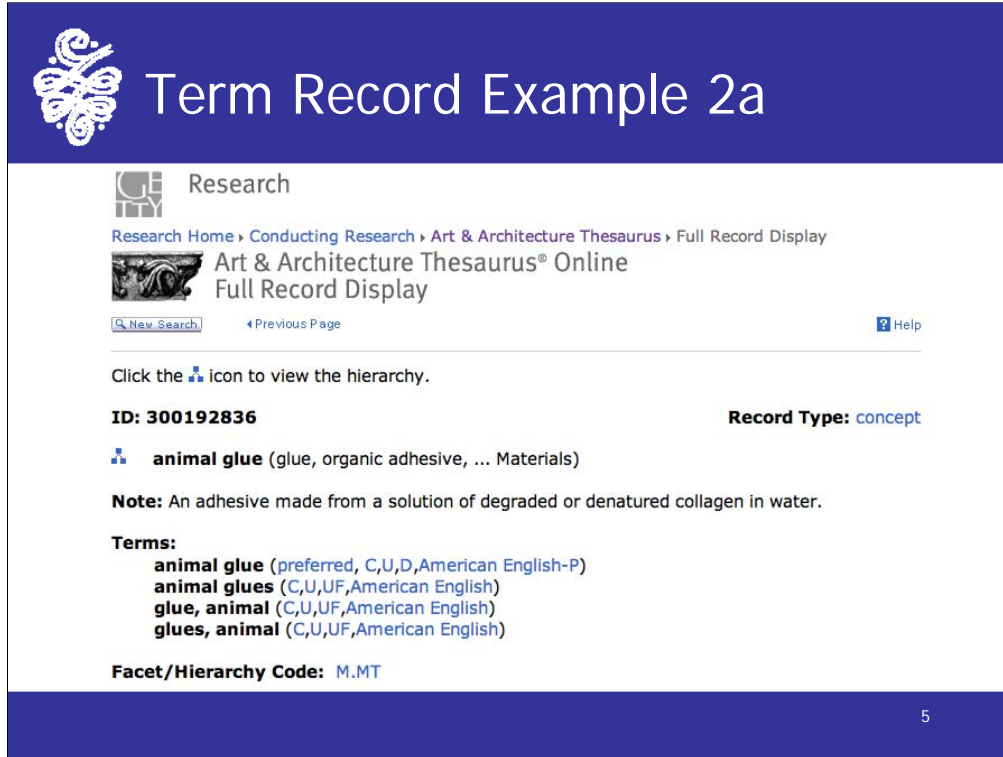
The screenshot shows the ERIC website interface. At the top left is the ERIC logo, a stylized butterfly. To its right is the text "ERIC Education Resources Information Center". Further right are links for "About ERIC", "ERIC News", "Contact ERIC", "Online Submission", and "Help". Below these are navigation tabs for "Home", "ERIC Search", "Thesaurus", and "My ERIC". The main content area is titled "Thesaurus Search Results" and features a search bar with the text "Start an ERIC Search using Ecology as a search criteria". Below the search bar, the term "Ecology" is displayed in a blue box. The record details are as follows:

- Record Type:** Main
- Scope Note:** Study of the interrelationships between organisms and their environment
- Category:** Science and Technology
- Broader Terms:** [Biological Sciences](#);
- Narrower Terms:** n/a
- Related Terms:** [Adjustment \(to Environment\)](#); [Area Studies](#); [Biodiversity](#); [Biology](#); [Botany](#); [Climate](#); [Conservation \(Environment\)](#); [Ecological Factors](#); [Energy Conservation](#); [Environment](#); [Environmental Education](#); [Environmental Standards](#); [Ethology](#); [Evolution](#); [Human Geography](#); [Indigenous Knowledge](#); [Marine Biology](#); [Mycology](#); [Quality of Life](#); [Radiation Biology](#); [Recycling](#); [Scientific Research](#); [Social Biology](#); [Sustainable Development](#); [Wastes](#); [Water](#); [Water Pollution](#); [Water Quality](#); [Weather](#); [Zoology](#);
- Used For:** [Bogs \(2004\)](#); [Ecosystems](#); [Estuaries \(2004\)](#); [Fens \(2004\)](#); [Habitats \(2004\)](#); [Marshes \(2004\)](#); [Rain Forest Preserves \(2004\)](#); [Rainforests \(2004\)](#); [Salt Marshes \(2004\)](#); [Swamps \(2004\)](#); [Temperate Rainforests \(2004\)](#); [Tropical Rainforests \(2004\)](#); [Wetlands \(2004\)](#);
- Use Term:** n/a
- Use And:** n/a
- Add Date:** 07/01/1966

The page number "4" is visible in the bottom right corner of the screenshot.


A term record from Eric.





The screenshot shows the 'Full Record Display' for the term 'animal glue' in the Art & Architecture Thesaurus Online. The page has a dark blue header with a white logo and the title 'Term Record Example 2a'. Below the header, the 'Research' logo is visible, followed by a breadcrumb trail: 'Research Home > Conducting Research > Art & Architecture Thesaurus > Full Record Display'. The main title is 'Art & Architecture Thesaurus® Online Full Record Display'. Navigation links include 'New Search', 'Previous Page', and 'Help'. A note instructs the user to click a hierarchy icon to view the hierarchy. The record details include: ID: 300192836, Record Type: concept, and the term 'animal glue' with its definition '(glue, organic adhesive, ... Materials)'. A 'Note' states: 'An adhesive made from a solution of degraded or denatured collagen in water.' The 'Terms' section lists: 'animal glue (preferred, C,U,D,American English-P)', 'animal glues (C,U,UF,American English)', 'glue, animal (C,U,UF,American English)', and 'glues, animal (C,U,UF,American English)'. The 'Facet/Hierarchy Code' is 'M.MT'. A page number '5' is located in the bottom right corner of the screenshot.

Here is the AAT version of a term record, or at least the first half. Including the term, the SN, the synonyms, what type of term it is, in this case a “concept”. AAT also has guide terms, hierarchy names, and facet terms. The guide terms are not terms at all but a way of making the hierarchy explicit without creating terms.



## Term Record Example 2b

**Hierarchical Position:**

- ⊞ Materials Facet
- ⊞ ... Materials
- ⊞ ..... materials
- ⊞ ..... <materials by function>
- ⊞ ..... adhesive
- ⊞ ..... <adhesive by composition or origin>
- ⊞ ..... organic adhesive
- ⊞ ..... glue
- ⊞ ..... animal glue

**Related concepts:**

related to .... **gelatin**  
 ..... (<processed animal material>, animal material, ... Materials) [300011812]

**Sources and Contributors:**

animal glue..... [VP Preferred]  
 ..... Mills and White, *Organic Chemistry* (1987) 75

animal glues..... [VP]  
 ..... Rose and von Endt, eds., *Protein Chemistry* (1984) 42  
 ..... Wehlte, *Materials and Techniques of Painting* (1982) 210

glue, animal..... [VP]  
 ..... Getty Vocabulary Program

glues, animal..... [VP]  
 ..... Getty Vocabulary Program

**Subject:** ..... [VP]

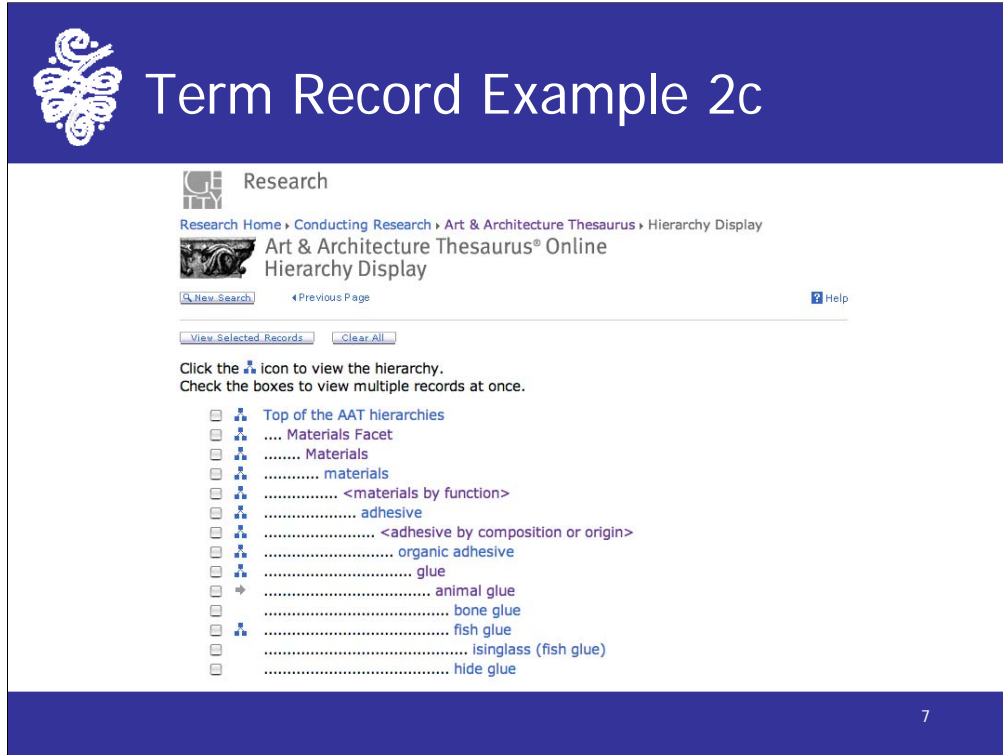
**Note:** ..... [VP]


6

Here is the second half, showing all of the broader terms above it, until the top level of the hierarchy, being the “Materials Facet”, we have and RT “Gelatin”, our SRCs.

The guide terms are in carrots <materials by function>, the terms are lower case, the Hierarchy names start with a capital letter, and the facet term ends in facet. That is their convention and they stick to it.


If you actually want to see the narrower/child terms for the current term, you need to go through the hierarchy symbol in the hierarchical position.



 Term Record Example 2c


Research













Research Home > Conducting Research > Art & Architecture Thesaurus > Hierarchy Display

 Art & Architecture Thesaurus® Online  
Hierarchy Display

[New Search](#) [Previous Page](#) [Help](#)

[View Selected Records](#) [Clear All](#)


Click the  icon to view the hierarchy.  
Check the boxes to view multiple records at once.

-  Top of the AAT hierarchies
-  .... Materials Facet
-  ..... Materials
-  ..... materials
-  ..... <materials by function>
-  ..... adhesive
-  ..... <adhesive by composition or origin>
-  ..... organic adhesive
-  ..... glue
-  ..... animal glue
-  ..... bone glue
-  ..... fish glue
- ..... isinglass (fish glue)
- ..... hide glue

7


This brings us to the total hierarchy and the narrower terms.

This is their way of dealing with polyhierarchical terms. They can put multiple parent hierarchies on a term record, and then only by clicking through the hierarchy you want can you see the narrower terms that are actually part of that hierarchy. This makes their site less user friendly, but it is entirely necessary given the structure they have set up.



## Displays: Alphabetical

- Similar to a back-of-the-book index
- Alphabetical list of all preferred terms
- Can also list non-preferred terms, with USE pointers to preferred terms



9.3.1, <sup>8</sup>p. 64

The alphabetical display is similar to an index in how it is accessed.

How would help the user? How would it hinder them?

Trying to elicit: As long as the user knows the term they are looking for, this is the quickest method of access, but if they do not know the term or even an equivalent term they could be out of luck.

Most web-based alphabetical displays are a hybrid of alphabetical and hierarchical, because of the functionality that hyperlinking affords.

**IIPA Full Text** HOME | ABC

[\[need help?\]](#)

**Search articles**

Limit to full-text articles only [More information](#)  Peer-reviewed articles only  Exclude re

**Keyword**

**Article Title**  [> select from a list](#)

**Author**  [> select from a list](#)

**Subject Categories\***  
*e.g. film*  [> select from a list](#)

**Subject Terms\***  
*e.g. production companies*  [> select from a list](#)

**Document Type**  [> select from a list](#)

9

The International Index to Performing Arts database allows the user to view the indexed terms by clicking on the 'select from a list' link. [next screen]

**IIPA Full Text**

[need help?] **Go**

- SEARCH ARTICLES
- SEARCH JOURNAL TITLES
- BROWSE JOURNALS
- SELECTED RECORDS
- SEARCH HISTORY
- LIBRARIANS' RESOURCES

**Select Subject Terms**

Subject Terms  **LOOK FOR**

**UP**

- Gherman, Alexei
- Ghery, Frank
- Ghesse Haye Kish, Abolfazl Jalili (1999, Film)
- Ghetonia Ensemble
- Ghetto Cabaret, Joshua Sobol (n/a, Drama)
- Ghetto Superstar: The Man That I Am, Billy Porter (n/a, M...
- Ghetto-Kids, Christian Wagner (2002, Film)
- GhettOriginal Productions
- Ghezlbash, Regis
- Ghezzi, Enrico
- Ghia, Carmen (Character)

**DOWN**

**OK**

10

Which displays a simple alphabetical listing of indexed terms for the user to browse.

**Alphabetical Example 2a**

**Grove Art Online**  
for all aspects of the visual arts world wide

Search

Home Advanced search Browse Explore

Article Abbreviations Contributors **Index** Go to

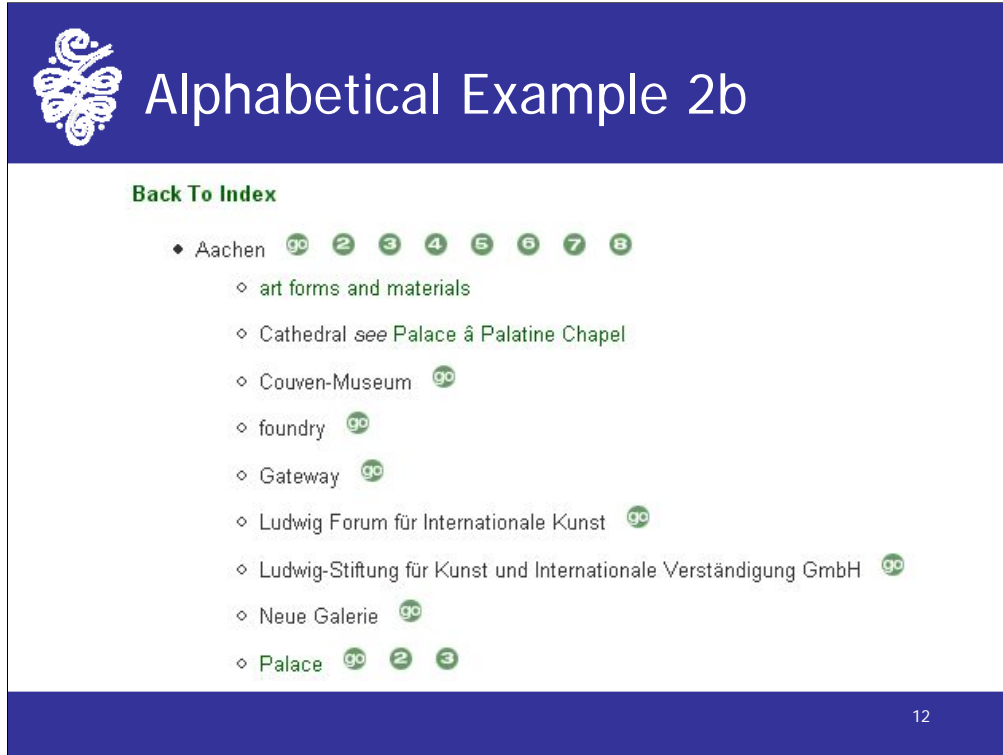
A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z

- A, Painter
- Aa, Dirk van der
- Aa, Jan van der
- Aa, van der (family)
- A.A.A. *see* Allied artists' association; American abstract artists
- A.A.A.-Anderson Galleries
- A.A.A. Gallery
- Aachen
- Aachen, Hans von (1552-1615)
- Aachen, Johann von (fl 1540)
- Aachen Altar, Master of the
- Aachen Gospels *see under* Gospel books → individual manuscripts; Gospel books → individual manuscripts
- Aachen Situla
- AACR *see* Anglo-american cataloging rules

11

Grove Art Online, on the other hand, provides the user with a multilevel alphabetical display.

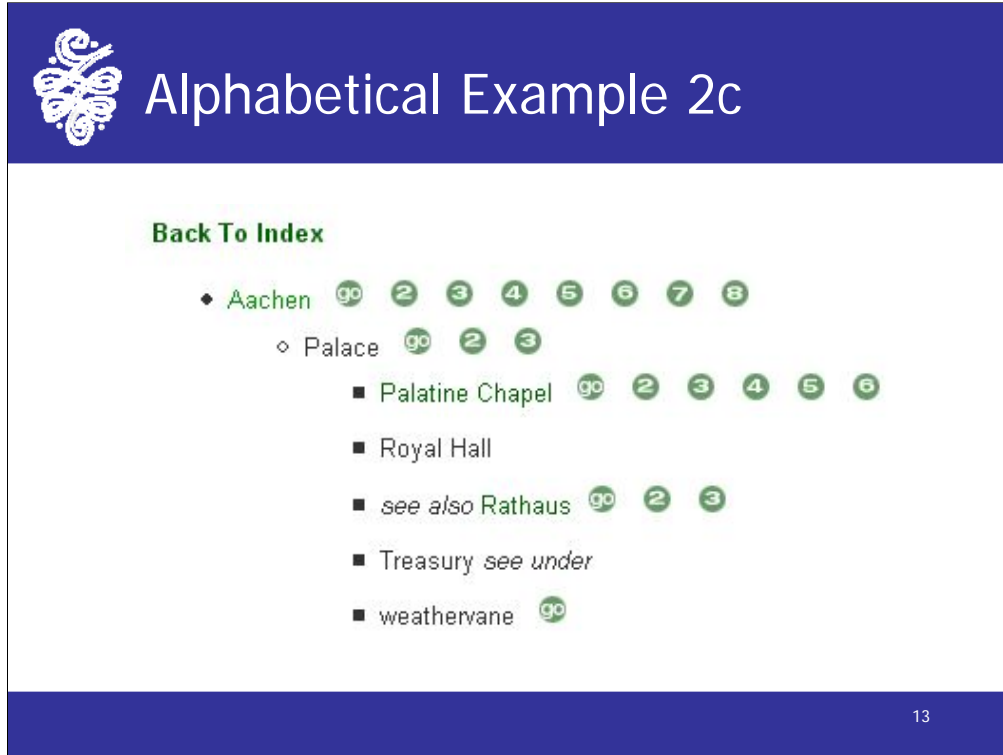
In this alphabetical listing of indexed terms, the GO icon is simply a link to an article indexed under that term. Terms with additional sub-levels of information are indicated both by the plus icon and a green font.



The screenshot shows a web interface with a dark blue header containing a white logo and the title "Alphabetical Example 2b". Below the header, the text "Back To Index" is displayed in green. A main menu item "Aachen" is shown with a green "go" icon and a row of numbered green circles (2, 3, 4, 5, 6, 7, 8). Underneath, a list of sub-items is shown, each with a diamond icon and a green "go" icon: "art forms and materials", "Cathedral see Palace à Palatine Chapel", "Couven-Museum", "foundry", "Gateway", "Ludwig Forum für Internationale Kunst", "Ludwig-Stiftung für Kunst und Internationale Verständigung GmbH", "Neue Galerie", and "Palace". The "Palace" item has a row of numbered green circles (2, 3). The page number "12" is located in the bottom right corner of the interface.

Clicking on Aachen in the initial browse index results in the display above. For example, the article indexed with the term 'Couven-Museum' is also indexed under the term 'Aachen'. The icons with numbers just indicate there are additional articles indexed under the indicated terms. In this screen, you can also see the display of a Used For term for Cathedral. Also, any term in green, again, has an additional level of indexing underneath.






The screenshot shows a digital interface with a dark blue header containing a white logo and the text "Alphabetical Example 2c". Below the header, the text "Back To Index" is displayed in green. The main content area is white and contains a hierarchical list of items:

- ◆ Aachen [go](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#)
  - ◇ Palace [go](#) [2](#) [3](#)
    - Palatine Chapel [go](#) [2](#) [3](#) [4](#) [5](#) [6](#)
    - Royal Hall
    - *see also* Rathaus [go](#) [2](#) [3](#)
    - Treasury *see under*
    - weathervane [go](#)

The page number "13" is located in the bottom right corner of the interface.

Clicking on Palace results in the display above. In this screen, we also see a display of an associative relationship for Palace, *see also* Rathaus.

Note: I assume 'Treasury *see under*' is telling the user that articles that would be indexed under 'Aachen-Palace-Treasury' are indexed under 'Aachen-Palace', but I was unable to confirm this assumption in a cursory look at their help file.




## Displays: Hierarchical

- Display terms within the context of the hierarchy
  - Category lists that can be drilled into
  - Starting at the top and listing all children
- Typically lists only preferred terms, no non-preferred

TESTING

- . Aerospace testing
- . . Aircraft testing
- . Electronic equipment testing
- . . Circuit testing
- . . . Integrated circuit testing
- . . . Printed circuit testing



9.3.4, <sup>14</sup>p. 68

This display literally starts at the top of the hierarchy and drills down.


Where would this help the user? When would it not?

Trying to elicit: when the user is browsing and not sure of a term. Similarly this display is difficult to use if you already know your term, but do not know the context.

Some problems with the hierarchical display are how to show the levels of depth within the hierarchy. If it is very deep or contains many narrower terms, then it is possible for the user to lose sight of the broader term, or levels above. It is a compromise between space and showing more of the broader/narrower relationships.


As the designer, you might find it is hard to decide how to order your broadest terms, or Top Terms, and really the option is up to you. If there is some logical order to the broadest terms, then by all means exploit it, but if not you might want to arrange them alphabetically.

Again the web-based display tends to make even a hierarchy more usable, allowing for the expansion of topics in ways that cut down on the confusion that looking at a printed display of hierarchy would create.



# Hierarchical Example 1a

**UNESCO Thesaurus: hierarchical list**



United Nations Educational, Scientific and Cultural Organization.

### 3. Culture


---

- [3.05 Culture](#)
- [3.10 Cultural policy and planning](#)
- [3.15 Philosophy and ethics](#)
- [3.20 Religion](#)
- [3.25 History](#)
- [3.30 Linguistics](#)
- [3.35 Languages](#)
- [3.40 Literature](#)
- [3.50 Visual arts](#)
- [3.55 Performing arts](#)
- [3.60 Museums](#)
- [3.65 Leisure](#)

15

This is a very traditional display of hierarchy. It comes from the UNESCO thesaurus.

Jump to the Performing arts section 3.55...



## Hierarchical Example 1b

### 3.55 Performing arts

[Back to hierarchical index](#)

---

#### Cinema

Narrower Term

- NT1 Cinema attendance
- NT1 Film festivals
- NT1 Film industry
- UF Film distribution, Film production, Movie industry*
- NT1 Film making
- UF Cinematography*
- NT2 Film scripts
- NT1 History of cinema

#### Dance

Used For

- UF Choreography

Narrower Term

- NT1 Ballet
- UF Classical ballet*
- NT1 Modern dance
- NT1 Traditional dance
- UF Folk dance*

16

Note the multi-level hierarchy..

**Hierarchical Example 2**

**CENTER FOR NONPROFIT MANAGEMENT**

**NONPROFIT RESOURCE LIBRARY**

**TAXONOMY of NONPROFIT MANAGEMENT**  
 Hierarchical  
 Alphabetical  
 About TNM  
 Thesauri  
 How to use  
 Credit  
 Feedback

- Nonprofit Organizations
  - Organization lifecycle
    - Startup
    - Collaboration
    - Merger
    - Dissolution
      - Involuntary dissolution
      - Voluntary dissolution
  - Legal
    - Legal structures
      - Nonprofit corporations
        - Federal corporations
        - Single-parent title-holding corporations
        - Charitable organizations
          - Religious organizations
          - Scientific organizations
          - Educational organizations
          - Literary organizations
          - Cemetery companies
          - Credit unions
          - Social welfare organizations
          - Labor and agriculture organizations
          - Business leagues
          - Social and recreational clubs
          - Fraternal beneficiary societies
          - Volunteer employee beneficiary associations

Expand All Collapse All

Home | About Us | Contact Us

**Using the Hierarchical View of TNM**


You can navigate through the terms by clicking on the term to be expanded or collapsed. Clicking on a term will expand the next level of terms within the upper level term and bring up the record that contains information about the term, such as the scope note (SN), "use for" non-entry terms (UF), broader term hierarchy (BT), narrower terms (NT), related terms (RT), and the term source (SRC).

In order to return to this introduction to the hierarchical display, please click on the **Hierarchical** link to your far left.

SN	Scope Note	NT	Narrower Term
UF	Use For	RT	Related Term
BT	Broader	SRC	Source note

Hierarchical view from TNM.

NASA uses the hierarchical NASA Thesaurus to index literature of interest to NASA employees or related to NASA activities. In the NASA search interface, you can choose to browse the NASA Thesaurus hierarchy by clicking on links which lead through a hierarchical map of the Thesaurus.



## Hierarchical Example 3b

**CATEGORY SEARCH**

Category Home > **NASA Missions and Projects**

Search in: NASA Missions and Projects


Include sub-categories

**+ GO**

**CATEGORIES**

- + **Flight Research** (32357)
- + **Humans in Space** (72509)
- + Robotics Missions (123)
- + **Solar System Exploration** (80488)

19



## Hierarchical Example 3c

---

**CATEGORY SEARCH**

Category Home > NASA Missions and Projects > **Humans in Space**

Search in: Humans in Space


Include sub-categories

---

**CATEGORIES**

<ul style="list-style-type: none"> <li>+ Apollo (5428)</li> <li>+ Apollo-Soyuz (359)</li> <li>+ <b>Astronauts</b> (40605)</li> <li>+ Gemini (1037)</li> <li>+ <b>International Space Station (ISS)</b> (35428)</li> </ul>	<ul style="list-style-type: none"> <li>+ Mercury (573)</li> <li>+ Shuttle-Mir (10520)</li> <li>+ Skylab (2926)</li> <li>+ <b>Space Shuttle</b> (19852)</li> </ul>
---	---





# Hierarchical Example 3d

**CATEGORY SEARCH**

Category Home > NASA Missions and Projects > Humans in Space > **International Space Station (ISS)**

Search in: International Space Station (ISS)

Include sub-categories

**+ GO**

**RESULTS**

**Documents: 1 - 10 of 200 returned**

+ Prev 1 2 3 4 5 + Next

**NASA - NASA Selects International Space Station Managers**

NASA Selects International Space Station Managers

[http://www.nasa.gov/home/hqnews/2004/may/HQ\\_04157\\_ISS\\_Mgrs.html](http://www.nasa.gov/home/hqnews/2004/may/HQ_04157_ISS_Mgrs.html) - 21KB

15 Jun 06

[+ Find Similar](#)

**NASA - NASA Names International Space Station Program Manage ...**

William H. Gerstenmaier Selects Michael Suffredini to Manage ISS Program

15 Jun 06

21

**Hierarchical Example 4a**

NCBI MeSH  
A service of the National Library of Medicine and the National Institutes of Health

My NCBI [Sign In] [Register]

All Databases PubMed Nucleotide Protein Genome Structure DMIM PMC Journals Books

Search MeSH for [ ] Go Clear

Limits Preview/Index History Clipboard Details

About Entrez  
Text Version  
Entrez PubMed  
Overview  
Help | FAQ  
Tutorials  
New/Noteworthy  
E-Utilities

**MeSH** is NLM's controlled vocabulary used for indexing articles for MEDLINE/PubMed. MeSH terminology provides a consistent way to retrieve information that may use different terminology for the same concepts.


- Use the MeSH database to find MeSH Terms and build a search strategy.

MeSH database tutorials:

- Searching with the MeSH Database  
Quick Tour

22

The National Library of Medicine, through its search interface Entrez, makes the MeSH vocabulary available for users to use in developing their search strategy.



## Hierarchical Example 4b

### MeSH Tree Structures

- [Musculoskeletal Diseases \[C05\]](#)
  - [Foot Deformities \[C05.330\]](#)
    - [Flatfoot \[C05.330.448\]](#)
    - [Foot Deformities, Acquired \[C05.330.488\] +](#)
    - [Foot Deformities, Congenital \[C05.330.495\] +](#)
    - ▶ [Hallux Valgus \[C05.330.610\]](#)
    - [Hallux Varus \[C05.330.612\]](#)
    - [Hammer Toe Syndrome \[C05.330.615\]](#)

Here is a small example of a hierarchical term display in MeSH's famous tree structure.



## Displays: Visual

- Some visual organization of the terms via their relationships
  - Tree graphs - family tree
  - Topic maps - clouds of terms
- Difficult to implement well
- Good for visual users!

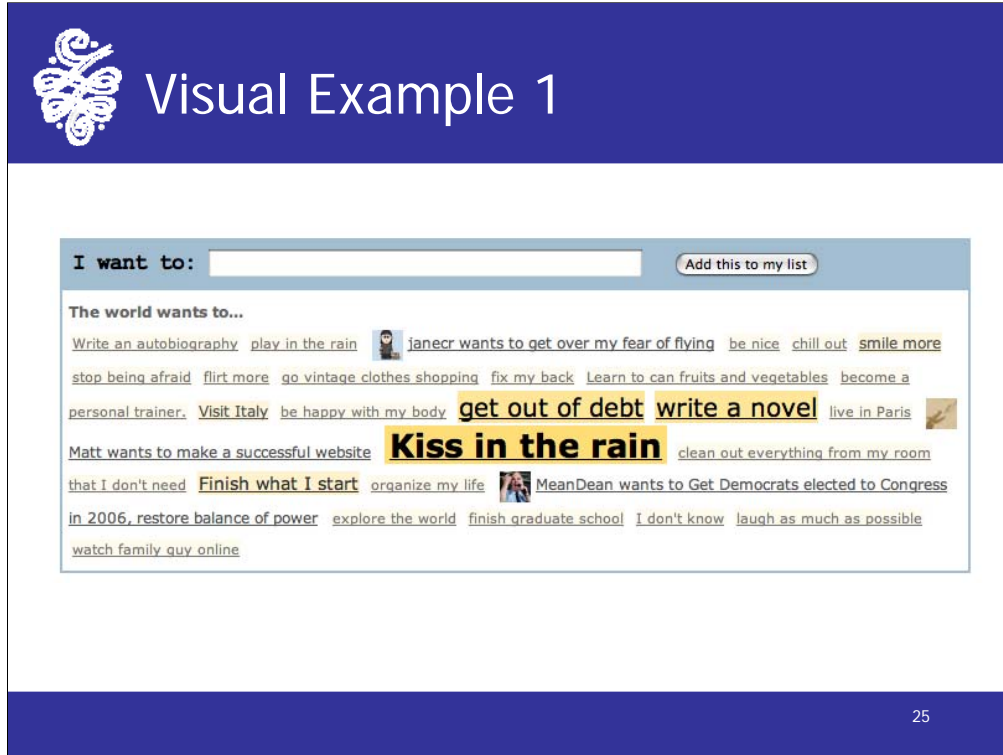


9.3.5, <sup>24</sup>p. 73

People are visual learners, aural learners, and kinetic learners. Some are better at interfacing with tables of data and others are better with graphs. People are all different, and visual thesauri are certainly a fun option for those visual learners out there.

You need to have someone who is very savvy with data visualization to implement them.

We will see how you all like this type of display...



**Visual Example 1**

I want to:  [Add this to my list](#)

The world wants to...

Write an autobiography play in the rain  janecr wants to get over my fear of flying be nice chill out smile more  
 stop being afraid flirt more go vintage clothes shopping fix my back Learn to can fruits and vegetables become a  
 personal trainer. Visit Italy be happy with my body **get out of debt write a novel** live in Paris 

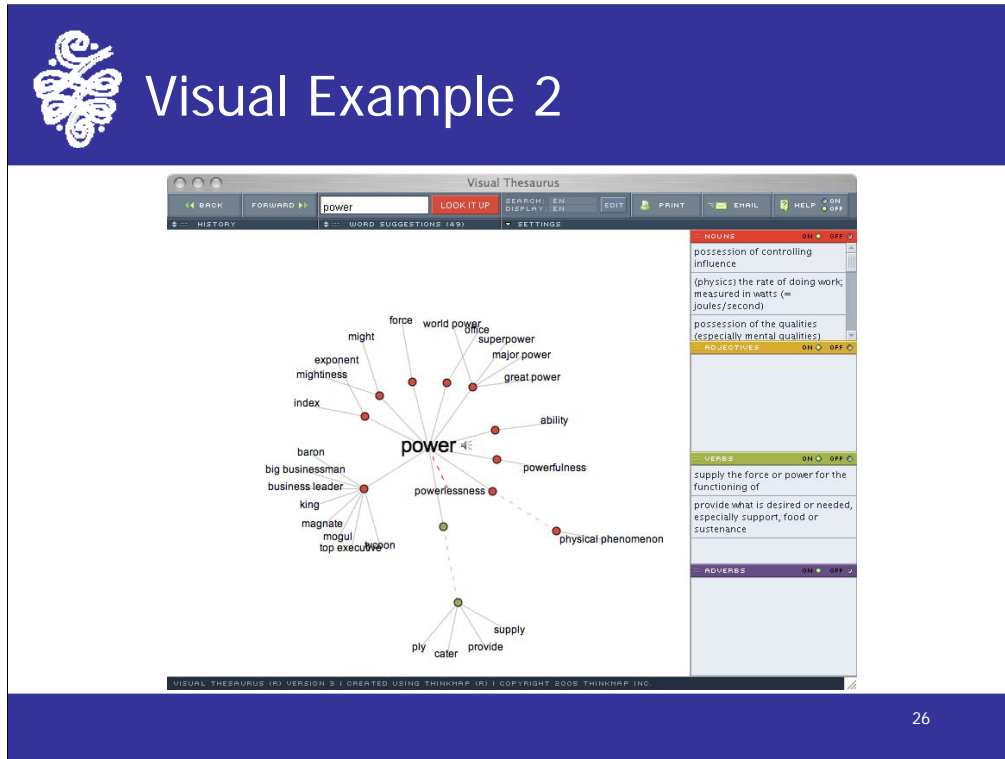
Matt wants to make a successful website **Kiss in the rain** clean out everything from my room  
 that I don't need **Finish what I start** organize my life  MeanDean wants to Get Democrats elected to Congress  
 in 2006, restore balance of power explore the world finish graduate school I don't know laugh as much as possible  
 watch family guy online

25

This is an interesting approach to a List, giving popularity by text size and background color density. Sort of a topic map, but not quite...

Similar to the Etsy example of materials used for crafts being sold.

<http://www.43things.com/>



This is literally the Visual Thesaurus, which is of course a Synonym Ring just like Roget's. But it also has some associated relationships and it gives you a visual notion of proximity. This is a more traditional topic map and tree graph.

<http://www.visualthesaurus.com/>

**Visual Example 3**

**BugGuide** Identification, Images, & Information For Insects, Spiders & Their Kin For the United States & Canada   [Register](#) [Log In](#)

[Home](#) | [Guide](#) | [ID Request](#) | [Recent](#) | [Frass](#) | [Forums](#) | [Donate](#) | [Help](#)

**Clickable Guide** **Taxonomy** **Browse** **Info** **Images** **Links** **Books** **Data**

[Home](#) > [Guide](#) > [Arthropods \(Arthropoda\)](#) > [Hexapods \(Hexapoda\)](#) > [Insects \(Insecta\)](#) > [Winged Insects \(Pterygota\)](#) > [Mantids and Cockroaches \(Dictyoptera\)](#) > [Mantids \(Mantodea\)](#) > [Mantidae](#) > [Tenodera](#)

Photo#1925

Copyright © 2003 [Troy Bartlett](#)

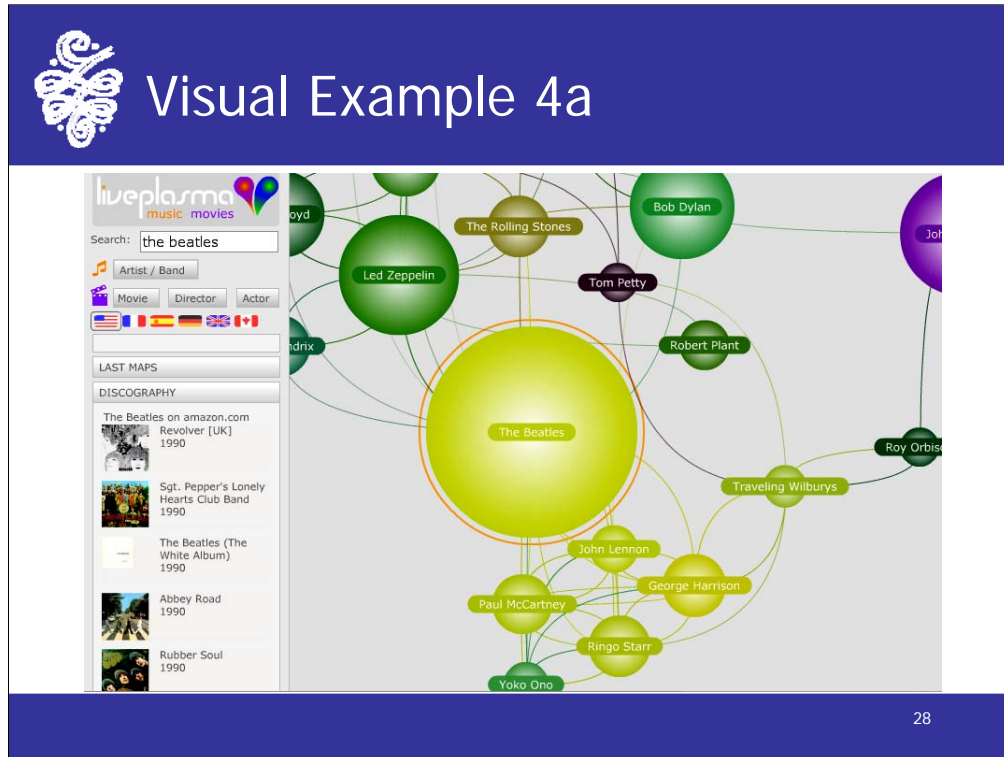
**Mantid - Tenodera**  
Duluth, Georgia, USA

27

The Clickable Guide on the left of this image serves as an entrance to this bug taxonomy, which also includes arachnids and other critters that are not technically bugs. More technically an arthropod taxonomy.

Handy because bugs are typically classified by their shape, like size of thorax, wings, etc.

<http://bugguide.net/>



LivePlasma is a way of interfacing with the statistics gathered by Amazon about user preferences. This is an image of how bands/musicians are related to the Beatles.

Bands can be looked at as a very flat thesaurus. As some bands have individuals that produce music as solo performers (children of band) part of other bands (polyhierarchy). The music styles can be the associative relationships, or in other words, if you bought this you would probably enjoy that, because they are in the same vein.

<http://www.liveplasma.com/>



The image shows a screenshot of the LivePlasma website's movie explorer feature. At the top left, there is a search bar with the text 'roger moore' and a search button. Below the search bar are tabs for 'Artist / Band', 'Movie', 'Director', and 'Actor'. The main content area is a radial map with 'Roger Moore' at the center. Surrounding him are various movie titles, each with a small image and a colored box. The colors of the boxes indicate the genre: yellow for action (e.g., 'The James Bond Collection, Vol. 1', 'The Spy Who Loved Me', 'The Man Who Haunted Himself'), blue for drama (e.g., 'The Saint', 'The Persuaders!'), and green for mystery (e.g., 'The Sea Wolves'). The map is connected to a sidebar on the left with menu options like 'LAST MAPS', 'DESIGN', 'MOVIE INFO', 'LEGEND', 'HELP', and 'MEMBERS ZONE'. The number '29' is visible in the bottom right corner of the screenshot.

This is again LivePlasma, but this is using their movie explorer feature, sorted by actor. So in the center is Roger Moore, and surrounding him are movies available for purchase in which he is an actor. Some interesting points to note, the Bond movies are all linked together, and the Bond collections are linked together, but they are not linked to one another. Distance from the center must be indicative of something. The color of the movie box is indicative of the genre.

If music can be viewed as a very flat thesaurus, then movies would be the polyhierarchical thesaurus to end all polyhierarchical thesauri.

<http://www.liveplasma.com/>



## Displays: Systematic

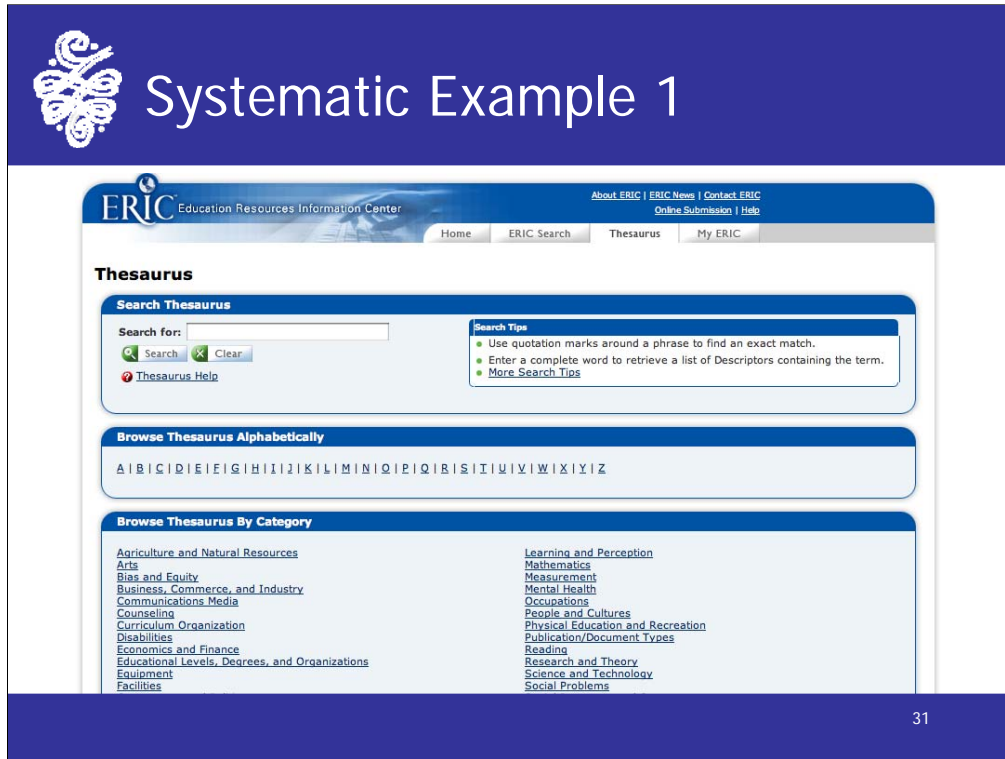
- A systematic view is some combination of views to allow multiple means of access
- The best of all worlds!

30

The real question is -- How can we help both the user who needs to browse for a term and the user who knows the term but not necessarily the context?


Just as library systems have used different access points for catalogued items, giving the user the option of a various displays, gives them the power to access the term they need with the information they have. The alphabetical display is analogous to author, and the hierarchy is analogous to subject.

The Systematic Display is a combination of these two, with an alphabetical index and a subject arranged or “systematic” display.




ERIC Education Resource Information Center thesaurus has 3 means of access, and you are presented them on the main thesaurus page. Search, Alphabetical browse, and Hierarchical drilling. Beautiful!


<http://www.eric.ed.gov/>



# Systematic Example 2a



## UNESCO Thesaurus



United Nations Educational, Scientific and Cultural Organization.

This web site allows you to [browse](#) the *UNESCO Thesaurus*. It also provides information [about](#) the *UNESCO Thesaurus*. To search the *UNESCO Thesaurus*, please go to the separate [search interface](#) which is now available on UNESCO's website.

1. [About the UNESCO Thesaurus](#)
2. [Browsing the UNESCO Thesaurus database](#)
3. [Purchasing the UNESCO Thesaurus](#)
4. [About this web site](#)
5. [Copyright statement](#)

### 1. About the UNESCO Thesaurus

The *UNESCO Thesaurus* is a controlled vocabulary developed by the [United Nations Educational, Scientific and Cultural Organisation](#) which includes subject terms for the following areas of knowledge: education, science, culture, social and human sciences, information and communication, and politics, law and economics. It also includes the names of countries and groupings of countries: political, economic, geographic, ethnic and religious, and linguistic groupings.


The *UNESCO Thesaurus* allows subject terms to be expressed consistently, with increasing specificity, and in relation to other subjects. It can be used to facilitate subject indexing in libraries, archives and similar institutions.

As in other subject thesauri, the terms in the *UNESCO Thesaurus* are linked together by three types of relationships:

- **Hierarchical relationships**, which link terms to other terms expressing more general and more specific concepts - i.e. broader terms and narrower terms. Hierarchically related terms are grouped under general subdivisions (known as "microthesauri"), which in turn are grouped into the areas of knowledge covered by the *Thesaurus*. In the *UNESCO Thesaurus*, broader terms are indicated by the prefix **BT**, narrower terms by the prefix **NT** and microthesauri by the prefix **MT**.
- **Associative relationships**, which link terms to similar terms (related terms) where the relationship between the terms is non-hierarchical. Related terms are indicated by the prefix **RT**.
- **Equivalence relationships**, which link "non-preferred" terms to synonyms or quasi-synonyms which act as "preferred"

32


This is the main page of the UNESCO thesaurus, which is also systematic. It is sort of hard to find either entry points in this page. Take this opportunity to note that people do not scroll down!



## Systematic Example 2b

**ULCC**

### UNESCO Thesaurus: alphabetical list

 [United Nations Educational, Scientific and Cultural Organization.](#)

This page allows you to search the *UNESCO Thesaurus* alphabetically. Each group of 40 terms links to the terms in that group. The list includes preferred terms and non-preferred terms. The following abbreviations are used: BT (broader term), FR (French equivalent), NT (narrower term), MT (microthesaurus heading), RT (related term), SN (scope note), SP (Spanish equivalent), UF (non-preferred term). Terms which are acronyms (e.g. UK) and compound terms which begin with acronyms (e.g. AIDS education) are listed at the start of their alphabetical section, rather than in alphabetical order.

[Return to UNESCO Thesaurus Home Page.](#)

---

[Abandoned children - Accessibility of information](#)

[Accessions - Admission requirements](#)

[Adolescence - Afrikaans](#)

[After school activities - Agricultural prices](#)


[Agricultural production - Aliens](#)

[Allegory - Animal feeding](#)


[Animal genetics - Aptitude](#)

33


<http://www2.ulcc.ac.uk/unesco/>



## Systematic Example 2c



### UNESCO Thesaurus: hierarchical list



United Nations Educational, Scientific and Cultural Organization.

---

This page allows you to search the *UNESCO Thesaurus* hierarchically, in the following way:

- The seven main sections of the *Thesaurus* link to the microthesaurus headings in each section.
- Each microthesaurus heading links to an alphabetical list of the preferred terms which are entered under that microthesaurus.
- Each term includes a link to a display of its broader terms, narrower terms, related terms, scope notes, non-preferred terms, French equivalent and Spanish equivalent.

[Return to UNESCO Thesaurus Home Page.](#)

---

- [1. Education](#)
- [2. Science](#)
- [3. Culture](#)
- [4. Social and human sciences](#)
- [5. Information and communication](#)
- [6. Politics, law and economics](#)
- [7. Countries and country groupings](#)

---

All rights reserved. The copyright of this web site belongs to UNESCO and the University of London Computer Centre. The information provided by this web site may be freely used and copied for educational and other non-commercial purposes, provided that any reproduction of data is accompanied by an acknowledgement of this web site as the source. Under no circumstances may copies be sold without prior written permission from the copyright holders.

34

<http://www2.ulcc.ac.uk/unesco/>

The screenshot displays the 'NONPROFIT RESOURCE LIBRARY' page from the Center for Nonprofit Management. The page features a navigation menu on the left with options: 'TAXONOMY of NONPROFIT MANAGEMENT', 'Hierarchical', 'Alphabetical', 'About TNM', 'Thesauri', 'How to use', 'Credit', and 'Feedback'. The main content area is titled 'NONPROFIT RESOURCE LIBRARY' and lists various categories under 'Nonprofit Organizations', such as 'Organization lifecycle', 'Legal', 'Governance', 'Organizational planning', 'Management', 'Financial management', 'Human resources', 'Operations and facilities', 'Information technology', 'Fund development', 'Evaluation', 'Marketing and public relations', 'Program management', 'Volunteerism', 'Ethics', 'Advocacy', 'Nonprofit sector', and 'Philanthropy'. A 'Using the Hierarchical View of TNM' box explains navigation options: 'SN' (Scope Note), 'NT' (Narrower Term), 'UF' (Use For), 'RT' (Related Term), 'BT' (Broader Term), and 'SRC' (Source note). A legend at the bottom right defines these abbreviations. The page also includes 'Expand All' and 'Collapse All' links, a footer with 'Home | About Us | Contact Us', and a copyright notice for 2004.

### Taxonomy of Nonprofit Management

Opens with the Hierarchical display, with the option of using the Alphabetical display.

**Systematic Example 3b**

**CENTER FOR NONPROFIT MANAGEMENT**

**NONPROFIT RESOURCE LIBRARY**

**TAXONOMY of NONPROFIT MANAGEMENT**  
**Hierarchical**  
**Alphabetical**  
**About TNM**  
 Thesauri  
 How to use  
 Credit  
 Feedback

Nonprofit Organizations  
 ▶ Organization lifecycle  
 ▶ Legal  
 ▶ Governance  
 ▼ Organizational planning  
 ▶ Organization statements  
 ▶ Assessments  
 ▼ Strategic planning  
     SWOT analyses  
     Operational plans  
     Capacity building  
     Succession planning  
 ▶ Management  
 ▶ Financial management  
 ▶ Human resources  
 ▶ Operations and facilities  
 ▶ Information technology  
 ▶ Fund development  
 ▶ Evaluation  
 ▶ Marketing and public relations  
 ▶ Program management  
 ▶ Volunteerism  
 ▶ Ethics  
 ▶ Advocacy  
 ▶ Nonprofit sector  
 ▶ Philanthropy

Expand All      Collapse All

Home | About Us | Contact Us  
 © 2004 Center for Nonprofit Management. All rights reserved.

<b>SN</b>	Refers to an evaluation of the strengths, weaknesses, opportunities and threats of an organization that is used during the strategic planning process.
<b>BT</b>	Strategic planning

<b>SN</b>	Scope Note	<b>NT</b>	Narrower Term
<b>UF</b>	Use For	<b>RT</b>	Related Term
<b>BT</b>	Broader Term	<b>SRC</b>	Source note

Opening the hierarchy up...


<http://www.cnmsocal.org/library/FinalTNM/TNM/index.html>



The Alphabetical view, opens at A, and lists all of the available letters and numbers.


This is a good point to ask if they can see the difference between main entry terms and non-preferred terms in this display...

Here, Preferred terms are in regular text, and the non-preferred terms are in italics. This was a poor mistake on my part, and if I had the chance to do it again, which presumably I will, I will use the all caps vs not all caps means of distinguishing non-preferred terms.



## Displays: Permuted Index

- Separate listing for every word in every term with pointers to the full compound term
  - Twist on the alphabetical display
  - Mainly used for compound terms
- Typically part of a systematic display
  - Serves as an index to the thesaurus



9.3.2, <sup>38</sup>p. 65


This is a very interesting type of thesauri, and the idea of using the alphabetical display as an index to the thesaurus is very handy, which once again sacrifices space for making it more user friendly.

This is basically the print version of a keyword search on the terms.

The screenshot displays the ILLUMINA Thesaurus interface. At the top, there is a logo for CSA ILLUMINA and the University of Virginia Libraries. The interface includes a search bar with the text "Search the Thesaurus" and a "Go" button. Below the search bar, there are several tabs: "Combine Searches", "Alerts", "History", "Command Search", "Thesaurus", and "Indexes". The "Thesaurus" tab is selected. The main content area shows a dropdown menu for "Select Thesaurus" set to "Technology Terms" and a text input field for "Browse Thesaurus for:" containing the word "environmental". There are three radio buttons for "Select Display": "Alphabetical Index", "Hierarchy", and "Rotated Index", with "Rotated Index" selected. Below the search options, there is a section titled "Definitions:" with three entries: "Alphabetical List", "Hierarchy", and "Rotated Index", each with a brief description of what the search results will look like.

39

The Civil engineering abstracts database from Cambridge Scientific Abstracts allows the user to search in and display a permuted or rotated index.

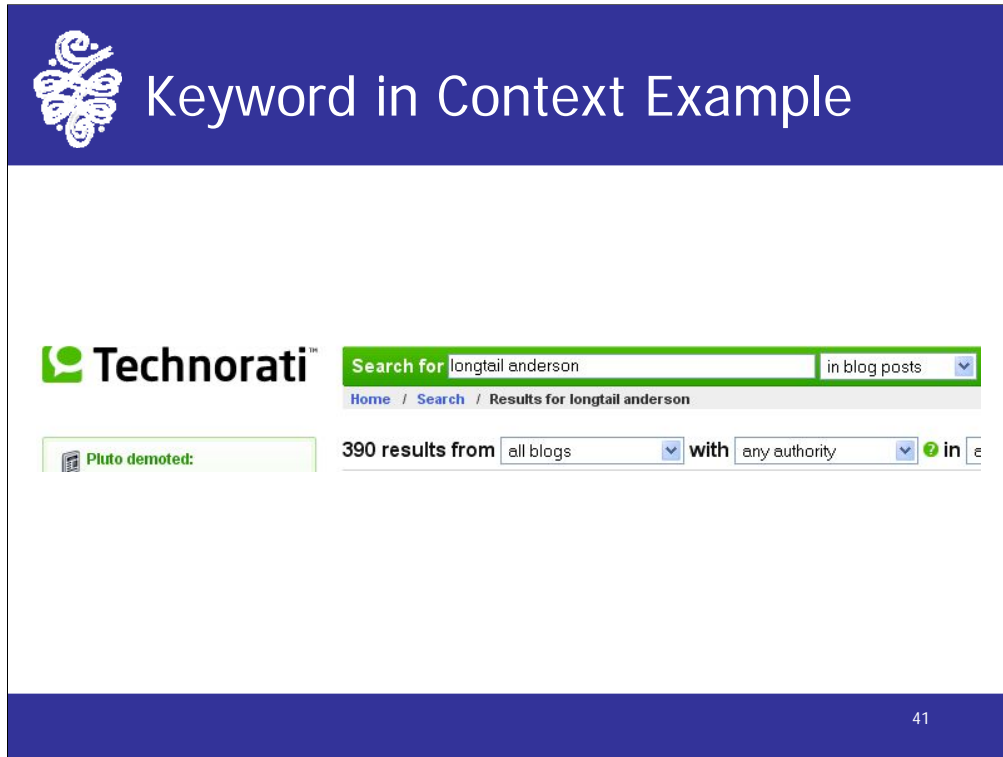


## Permuted Index Example

- [Environmental tests](#)
- [Environmental Trusts](#)
- [Integrated environmental assessment](#)
- [Protest, Environmental](#)


40

Every term with the word “environmental” in it is listed together. This is a nice way to get access to compound terms you might not know the first word of...



The image shows a screenshot of the Technorati search engine interface. At the top left is the Technorati logo, a stylized butterfly. To its right, the text "Keyword in Context Example" is displayed in white on a dark blue background. Below this, the main search area features the Technorati logo on the left. To the right of the logo is a search bar containing the text "longtail anderson". To the right of the search bar is a dropdown menu set to "in blog posts". Below the search bar is a breadcrumb trail: "Home / Search / Results for longtail anderson". To the left of the search results area is a small green box with the text "Pluto demoted:". To the right of this box, the text "390 results from" is followed by a dropdown menu set to "all blogs", the word "with", another dropdown menu set to "any authority", and a small green icon with the text "in e". At the bottom right of the screenshot, the number "41" is displayed.

Keyword in Context has become a very popular feature of many web search engines. This example is from the Technorati website.




## Keyword in Context Example

**Faux Granite Video**  
2 days ago in **SalesTail** Because Niche ain't always Kitsch · [1 blog links here](#)  
Faux Granite Video August 25th, 2006 Long Tail Product of the Day: Faux Granite Video Purchased and downloaded from Blish ... on **Longtail** by Chris **Anderson** The original post describing the **longtail** market dynamics VCs Ask ... ) Long Tail Product of the Day (21) Long Tail Quote of the Day (1) **Longtail** Incarnations  
[» Show details](#)

**links for 2006 08 25**  
2 days ago in **Micro Persuasion** by **steverubel** · [4,228 blogs link here](#)  
Wired News: iWoz Logs Leap From Geek to Icon Interview with Woz about his new book. (tags: woz apple) ... execs and Chris **Anderson**, author of the Long Tail. (tags: **LongTail** 37Signals Yahoo) TechCrunch UK  
[» Show details](#)

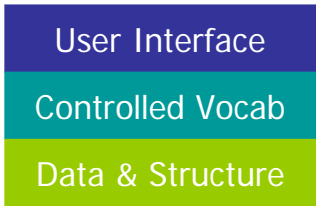
42

Notice the bold instances of the search terms.



## User/Indexer Views

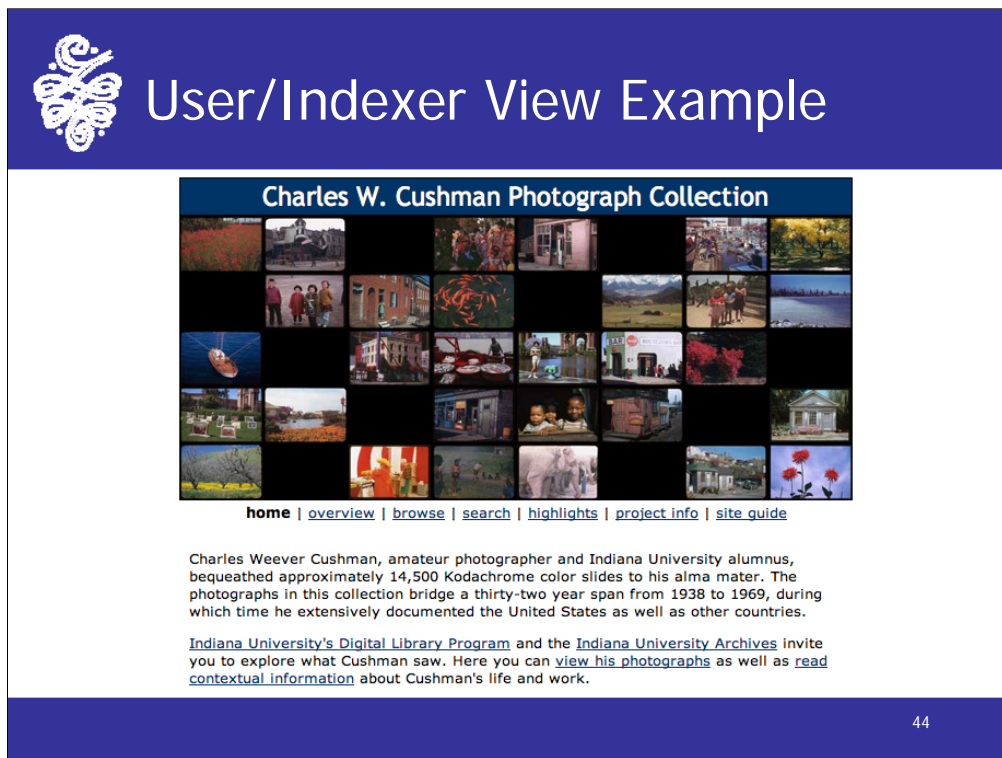
- A CV is layered over the data
- Potentially a user does not need to see the CV guts
- The indexer may need to see more of the admin info than the user
- How can both be accommodated?



43

### Non-integration issues of ILSs -- Discussion

Sometimes you can convert you CV into some sort of authority list within an ILS, and apply terms. That sort of compromises your structure, but it is a reasonable tradeoff.



**Charles W. Cushman Photograph Collection**

home | overview | browse | search | highlights | project info | site guide

Charles Weever Cushman, amateur photographer and Indiana University alumnus, bequeathed approximately 14,500 Kodachrome color slides to his alma mater. The photographs in this collection bridge a thirty-two year span from 1938 to 1969, during which time he extensively documented the United States as well as other countries.

[Indiana University's Digital Library Program](#) and the [Indiana University Archives](#) invite you to explore what Cushman saw. Here you can [view his photographs](#) as well as [read contextual information](#) about Cushman's life and work.

44


Dalmau, Michelle, Randall Floyd, David Jiao, and Jenn Riley. "Integrating Thesaurus Relationships into Search and Browse in an Online Photograph Collection." *Library Hi Tech* 23, no. 3 (2005): 425-452. <http://www.dlib.indiana.edu/~jenlrile/publications/cushmancv/cushmanCV.pdf>

Dalmau, Michelle, and Jenn Riley. "Use of Controlled Vocabulary in Searching and Browsing the Cushman Collection." *Indiana University Digital Library Program Brownbag*. Indiana University, Herman B Wells Library, Bloomington. 07 May 2004.

Cantara, Linda, Michelle Dalmau, and Jenn Riley. "The Charles W. Cushman Collection: Enhancing Visual Resource Discovery Through Descriptive Metadata Based on Subjective Image Analysis." *Web X: A Decade of the World Wide Web. The Joint International Conference of the Association for Computers and the Humanities/Association for Literary and Linguistic Computing (ACH/ALLC)*, University of Georgia, Athens, Georgia, 29 May - 2 June, 2003. Full paper submitted for publication in conference proceedings, forthcoming.

Dalmau, Michelle, and Jenn Riley. "Cushman Exposed! Exploiting Controlled Vocabularies to Enhance Browsing and Searching of an Online Photograph Collection." *Digital Library Federation Spring Forum 2003*. Sheraton New York Hotel & Towers, New York. 15 May 2003. <http://www.dlib.indiana.edu/education/brownbags/spring2004/cushmanCVbb/cushmancvbb.pdf>






# User View

## Charles W. Cushman Photograph Collection

Indiana University Archives / Digital Library Program



[home](#) ■ [overview](#) ■ **[browse](#)** ■ [search](#) ■ [highlights](#) ■ [project info](#) ■ [site guide](#)

### Browse the Cushman Collection

[home](#) >> [browse](#)

[Year](#)  
Browse photographs chronologically by calendar year.

[Location](#)  
Browse photographs by geographic location.

[Subject](#)  
Browse photographs by subject or topic matter.

[Genre](#)  
Browse photographs by a particular style, form, or general content.

[Notebook](#)  
Browse the notebooks in which Cushman documented his photographs.

45

The screenshot shows a web page with a dark blue header containing a white logo and the text "User View". Below the header is a dark blue banner with the text "Charles W. Cushman Photograph Collection" and "Indiana University Archives / Digital Library Program". To the right of the banner are three small photographs. Below the banner is a navigation menu with links: "home", "overview", "browse", "search", "highlights", "project info", and "site guide". The main content area is titled "Browse by Subject" and includes a breadcrumb trail: "home >> browse >> subject" followed by a question mark icon. Below the title is a horizontal list of letters from A to Z, with "All" centered under the letters M and N. A horizontal line is positioned below the letters. Underneath the line is the text "Select a letter above to browse subject terms." The page has a dark blue footer with the number "46".

**User View**

Charles W. Cushman Photograph Collection  
Indiana University Archives / Digital Library Program


home ■ overview ■ browse ■ search ■ highlights ■ project info ■ site guide

**Browse by Subject** [home](#) >> [browse](#) >> [subject](#) | ?




A B C D E F G H I J K L M N O P Q R S T U V W Y  
Z All

- [Cabanas](#) (2)
- [Cabarets](#) [maps to nightclubs] (5)
- [Cabins](#) (10)
- [Cable cars](#) [maps to cable railroads] (28)
- [Cable railroads](#) (28)
- [Caboozes](#) [maps to railroad cabooses] (4)
- [Cabs](#) [maps to taxicabs] (18)
- [Cabs, Hansom](#) [maps to carriages & coaches] (68)
- [Cactus](#) (169)
- [Cadavers](#) [maps to dead animals] (4)
- [Caddies](#) (1)
- [Cafes](#) (64)
- [Cages](#) (16)
- [Calamities](#) [maps to disasters] (45)
- [California poppies](#) (3)
- [Calla](#) [maps to lilies] (23)
- [Calliopes](#) (1)

47

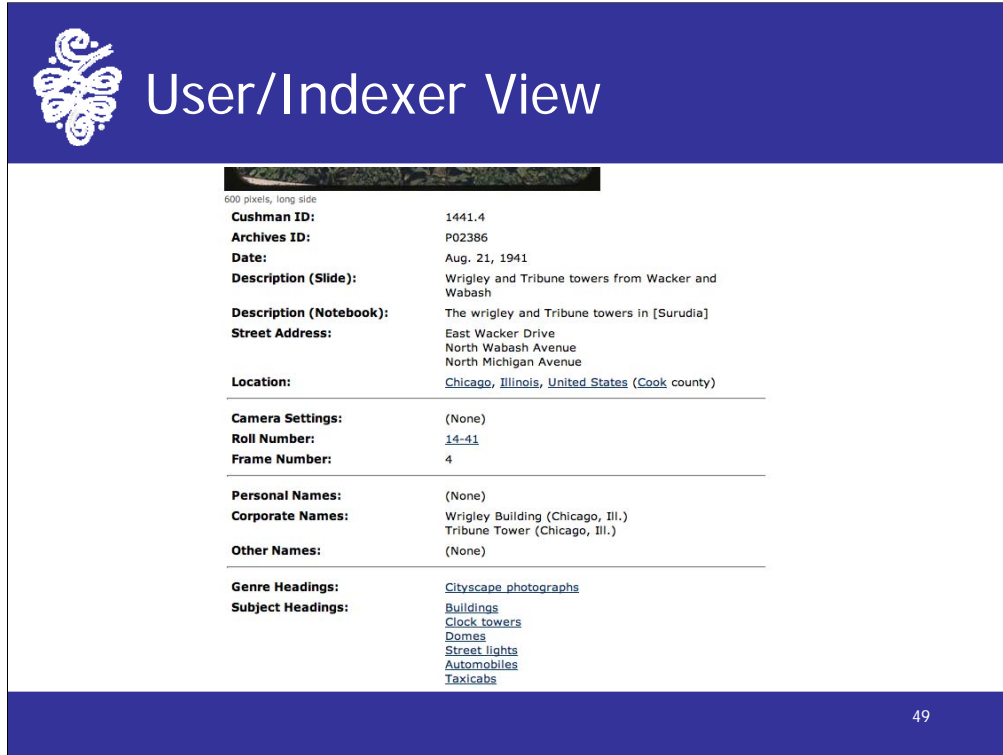



## User View


<p><b>Browse Suggestions</b> Modify your original browse results by selecting options within any of the following available facets: year, location, genre and/or subject.</p> <p><b>Year</b></p> <ul style="list-style-type: none"> <li>1941</li> <li>1948</li> <li>1949</li> <li>1952</li> <li>1954</li> <li>1955</li> <li>1957</li> <li>1961</li> <li>1964</li> </ul> <p><b>Location</b></p> <ul style="list-style-type: none"> <li><a href="#">Austria</a></li> <li><a href="#">United Kingdom</a></li> <li><a href="#">United States</a></li> </ul> <p><b>Genre</b></p> <ul style="list-style-type: none"> <li><a href="#">Cityscape</a></li> <li><a href="#">photographs</a></li> <li><a href="#">Landscape</a></li> <li><a href="#">photographs</a></li> <li><a href="#">Snapshots</a></li> <li><a href="#">Views</a></li> </ul>	1		<p>Cushman ID No.: 1441.4                  Roll No.: <a href="#">14-41</a>                  Date: Aug. 21, 1941                  Location: <a href="#">Chicago, Illinois, United States</a> (<a href="#">Cook county</a>)                  Description: Wrigley and Tribune towers from Wacker and Wabash  <a href="#">view details</a></p>
	2		<p>Cushman ID No.: 748.13                  Roll No.: <a href="#">7-48</a>                  Date: Aug. 23, 1948                  Location: <a href="#">Chicago, Illinois, United States</a> (<a href="#">Cook county</a>)                  Description: Old Chicago street car Chgo RR fair  <a href="#">view details</a></p>
	3		<p>Cushman ID No.: 848.18                  Roll No.: <a href="#">8-48</a>                  Date: Sep. 13, 1948                  Location: <a href="#">Chicago, Illinois, United States</a> (<a href="#">Cook county</a>)                  Description: State St. cable car - Chgo - RR fair and a Gig  <a href="#">view details</a></p>

48

Notice: results interface shows and provides access to broader and narrow terms on left of results screen Click on [view details](#) next to any image to see a larger image with the full metadata records, including the terms assigned from the controlled vocabulary



 **User/Indexer View**

  
500 pixels, long side

**Cushman ID:** 1441.4  
**Archives ID:** P02386  
**Date:** Aug. 21, 1941  
**Description (Slide):** Wrigley and Tribune towers from Wacker and Wabash  
**Description (Notebook):** The wrigley and Tribune towers in [Surudia]  
**Street Address:** East Wacker Drive  
 North Wabash Avenue  
 North Michigan Avenue  
**Location:** [Chicago, Illinois, United States](#) ([Cook county](#))

---

**Camera Settings:** (None)  
**Roll Number:** [14-41](#)  
**Frame Number:** 4

---

**Personal Names:** (None)  
**Corporate Names:** Wrigley Building (Chicago, Ill.)  
 Tribune Tower (Chicago, Ill.)  
**Other Names:** (None)

---

**Genre Headings:** [Cityscape photographs](#)  
**Subject Headings:** [Buildings](#)  
[Clock towers](#)  
[Domes](#)  
[Street lights](#)  
[Automobiles](#)  
[Taxicabs](#)

49

This is the very detailed information about the image that may be useful to the user, but is more likely to be useful to the indexer



**Index View**

**Charles W. Cushman Photograph Collection**  
Indiana University Archives / Digital Library Program

[home](#) ■ [overview](#) ■ [browse](#) ■ [search](#) ■ [highlights](#) ■ [project info](#) ■ [site guide](#)

**Cataloging Information** [home >> project info >> cataloging](#)

Cushman's notebook annotations and corresponding documentation on slide mounts have been exactly transcribed, even when abbreviations or misspellings are included. These descriptions are keyword searchable. Dates associated with the slides are also searchable. In addition, descriptive metadata for the subject content, geographic location, genre and physical characteristics of each image provides enhanced access to the collection. Terms have been selected from the following controlled vocabularies:

[Library of Congress Thesauri for Graphic Materials](#)  
The Library of Congress Thesaurus for Graphic Materials I: Subject Terms (TGM I) and Thesaurus for Graphic Materials II: Genre and Physical Characteristic Terms (TGM II) include topic headings developed by the Library of Congress Prints and Photographs Division over the past fifty years to catalog graphic materials. The thesauri adhere to ANSI/NISO Z39.19-1994, the American National Standards Institute/National Information Standards Organization Guidelines for the Construction, Format and Management of Monolingual Thesauri. Terms are expressed in natural language word order following American English spelling conventions. First published in 1987, both thesauri are now available online.

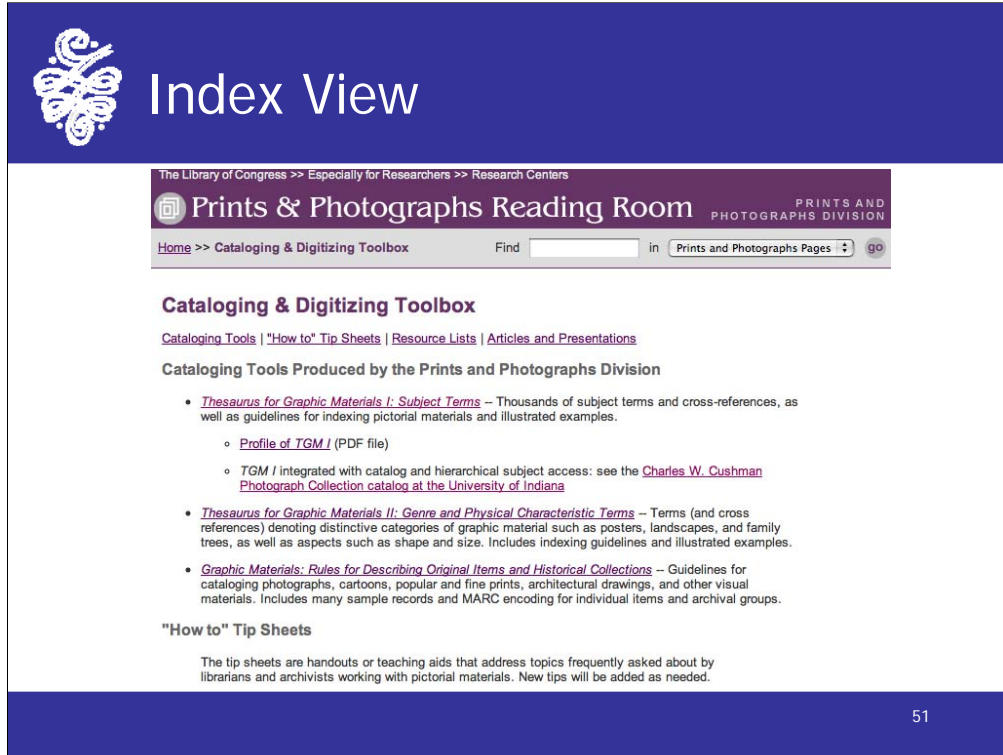
[Subject Terms \(TGM I\)](#)  
TGM I comprises more than 6,300 terms and approximately 5,000 cross-references. The thesaurus includes terms for objects, activities, occupations, events and concepts represented in visual resources. TGM I is updated regularly, with several hundred new terms added annually. Catalogers have selected applicable terms from TGM I to describe prominent subjects and activities represented in each image. As appropriate, terms for concepts (e.g., poverty, happiness, anger) are also documented.

[Genre and Physical Characteristic Terms \(TGM II\)](#)  
TGM II, the companion thesaurus to TGM I, comprises more than 600 terms and more than 450 cross-references for documenting categories of visual resources as well as the processes.

50

From the back end of the project, we can see that they are using a variety of controlled vocabularies to handle their materials...

This introduces a concept that will be covered in more detail in the next session, namely the use of other controlled vocabularies, or pieces there of in your own.



The Library of Congress >> Especially for Researchers >> Research Centers

Prints & Photographs Reading Room PRINTS AND PHOTOGRAPHS DIVISION

Home >> Cataloging & Digitizing Toolbox Find  in Prints and Photographs Pages

## Cataloging & Digitizing Toolbox

[Cataloging Tools](#) | ["How to" Tip Sheets](#) | [Resource Lists](#) | [Articles and Presentations](#)

### Cataloging Tools Produced by the Prints and Photographs Division

- [Thesaurus for Graphic Materials I: Subject Terms](#) – Thousands of subject terms and cross-references, as well as guidelines for indexing pictorial materials and illustrated examples.
  - [Profile of TGM I](#) (PDF file)
  - [TGM I](#) integrated with catalog and hierarchical subject access: see the [Charles W. Cushman Photograph Collection catalog at the University of Indiana](#)
- [Thesaurus for Graphic Materials II: Genre and Physical Characteristic Terms](#) – Terms (and cross references) denoting distinctive categories of graphic material such as posters, landscapes, and family trees, as well as aspects such as shape and size. Includes indexing guidelines and illustrated examples.
- [Graphic Materials: Rules for Describing Original Items and Historical Collections](#) – Guidelines for cataloging photographs, cartoons, popular and fine prints, architectural drawings, and other visual materials. Includes many sample records and MARC encoding for individual items and archival groups.

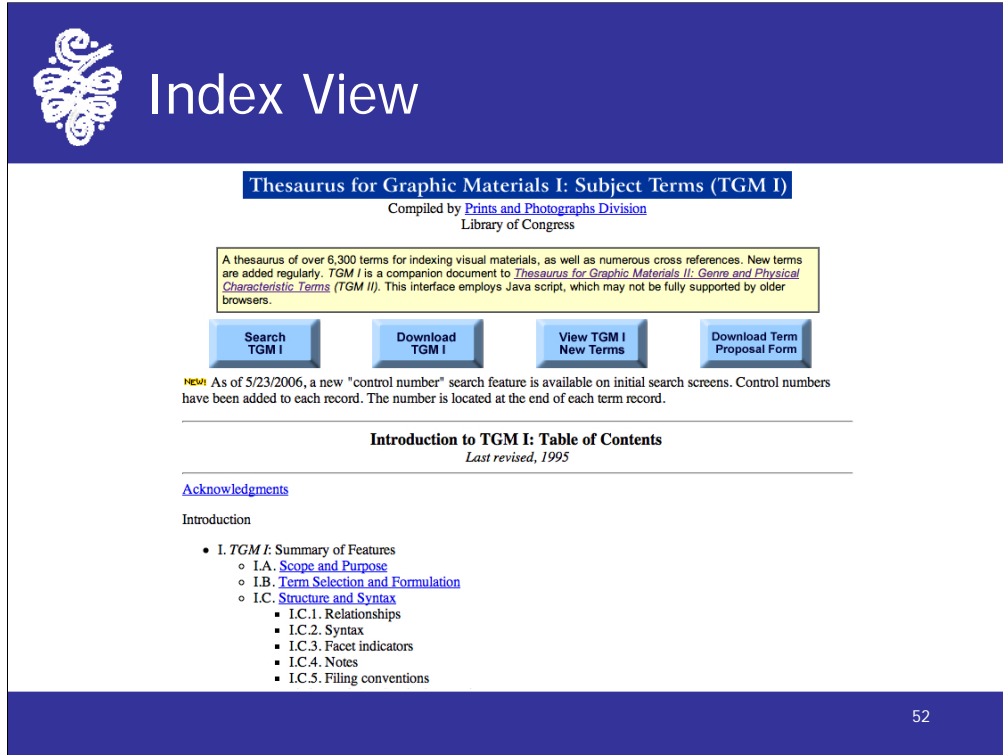
### "How to" Tip Sheets

The tip sheets are handouts or teaching aids that address topics frequently asked about by librarians and archivists working with pictorial materials. New tips will be added as needed.

51

LC resources for Prints and Photographs  
One of the contributing CVs.

Optional.



**Index View**

**Thesaurus for Graphic Materials I: Subject Terms (TGM I)**  
Compiled by [Prints and Photographs Division](#)  
Library of Congress

A thesaurus of over 6,300 terms for indexing visual materials, as well as numerous cross references. New terms are added regularly. *TGM I* is a companion document to *Thesaurus for Graphic Materials II: Genre and Physical Characteristic Terms (TGM II)*. This interface employs Java script, which may not be fully supported by older browsers.

[Search TGM I](#)   [Download TGM I](#)   [View TGM I New Terms](#)   [Download Term Proposal Form](#)

**NEW!** As of 5/23/2006, a new "control number" search feature is available on initial search screens. Control numbers have been added to each record. The number is located at the end of each term record.

**Introduction to TGM I: Table of Contents**  
*Last revised, 1995*

[Acknowledgments](#)

Introduction

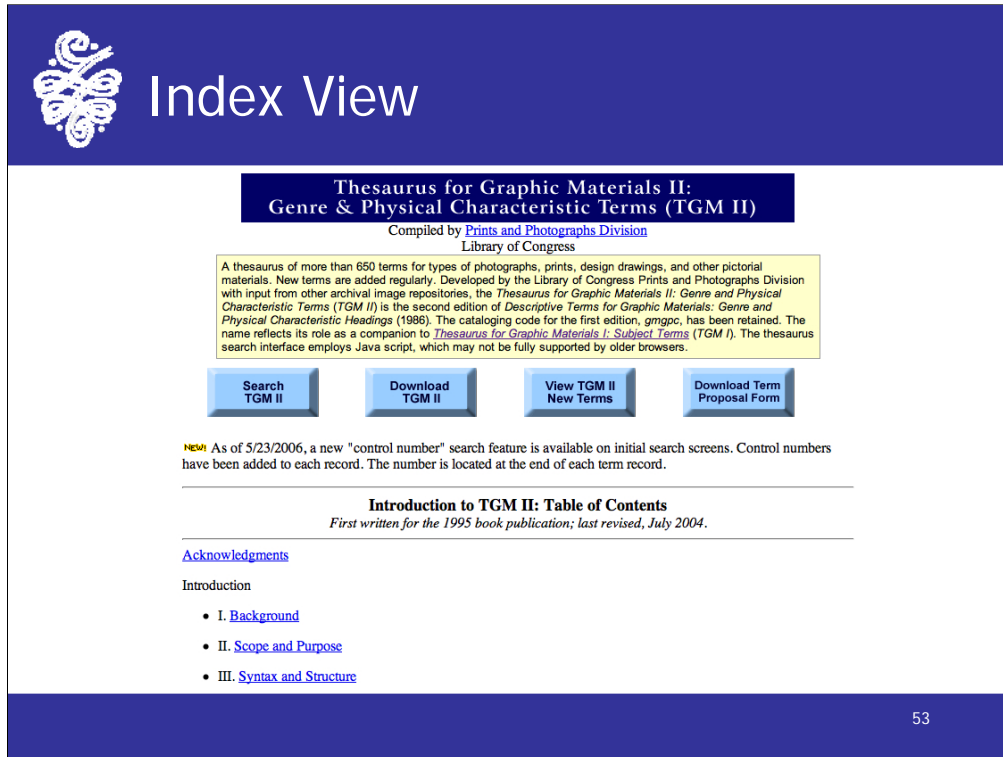
- I. *TGM I*: Summary of Features
  - I.A. [Scope and Purpose](#)
  - I.B. [Term Selection and Formulation](#)
  - I.C. [Structure and Syntax](#)
    - I.C.1. Relationships
    - I.C.2. Syntax
    - I.C.3. Facet indicators
    - I.C.4. Notes
    - I.C.5. Filing conventions

52

Thesaurus for Graphic Materials I: Subject Terms (LC)  
One of the contributing CVs.

Optional.



The image is a screenshot of a web page titled "Index View" for the "Thesaurus for Graphic Materials II: Genre & Physical Characteristic Terms (TGM II)". The page has a dark blue header with a white logo on the left and the title "Index View" in white. Below the header, the main content area is white. At the top of this area, the title "Thesaurus for Graphic Materials II: Genre & Physical Characteristic Terms (TGM II)" is displayed in a dark blue box, followed by "Compiled by [Prints and Photographs Division](#) Library of Congress". A yellow box contains a detailed description of the thesaurus, mentioning it has over 650 terms and is the second edition of a 1986 publication. Below this are four blue buttons: "Search TGM II", "Download TGM II", "View TGM II New Terms", and "Download Term Proposal Form". A "NEW" notice states that as of 5/23/2006, a new "control number" search feature is available. The page includes a section for "Introduction to TGM II: Table of Contents" with a note that it was first written in 1995 and last revised in July 2004. This section lists "Acknowledgments", "Introduction", and a numbered list of three items: "I. Background", "II. Scope and Purpose", and "III. Syntax and Structure". The page number "53" is located in the bottom right corner of the screenshot area.

**Index View**

**Thesaurus for Graphic Materials II:  
Genre & Physical Characteristic Terms (TGM II)**  
Compiled by [Prints and Photographs Division](#)  
Library of Congress

A thesaurus of more than 650 terms for types of photographs, prints, design drawings, and other pictorial materials. New terms are added regularly. Developed by the Library of Congress Prints and Photographs Division with input from other archival image repositories, the *Thesaurus for Graphic Materials II: Genre and Physical Characteristic Terms (TGM II)* is the second edition of *Descriptive Terms for Graphic Materials: Genre and Physical Characteristic Headings* (1986). The cataloging code for the first edition, *gmppc*, has been retained. The name reflects its role as a companion to *Thesaurus for Graphic Materials I: Subject Terms (TGM I)*. The thesaurus search interface employs Java script, which may not be fully supported by older browsers.

[Search TGM II](#)   [Download TGM II](#)   [View TGM II New Terms](#)   [Download Term Proposal Form](#)

**NEW** As of 5/23/2006, a new "control number" search feature is available on initial search screens. Control numbers have been added to each record. The number is located at the end of each term record.

---

**Introduction to TGM II: Table of Contents**  
*First written for the 1995 book publication; last revised, July 2004.*

[Acknowledgments](#)

Introduction

- I. [Background](#)
- II. [Scope and Purpose](#)
- III. [Syntax and Structure](#)

53

Thesaurus for Graphic Materials II: Genre & Physical Characteristic Terms (LC)

One of the contributing CVs.

Optional.

The screenshot displays the 'Index View' of the Getty Thesaurus of Geographic Names (TGN) website. The page has a dark blue header with the Getty logo on the left and the text 'Index View' in white. Below the header is a navigation bar with the text 'The Getty' and several menu items: Home, Visit, Art, Education, Research, and Conservation. Underneath these are sub-links: Conducting Research, Scholarly Activities, and About the Research Institute.

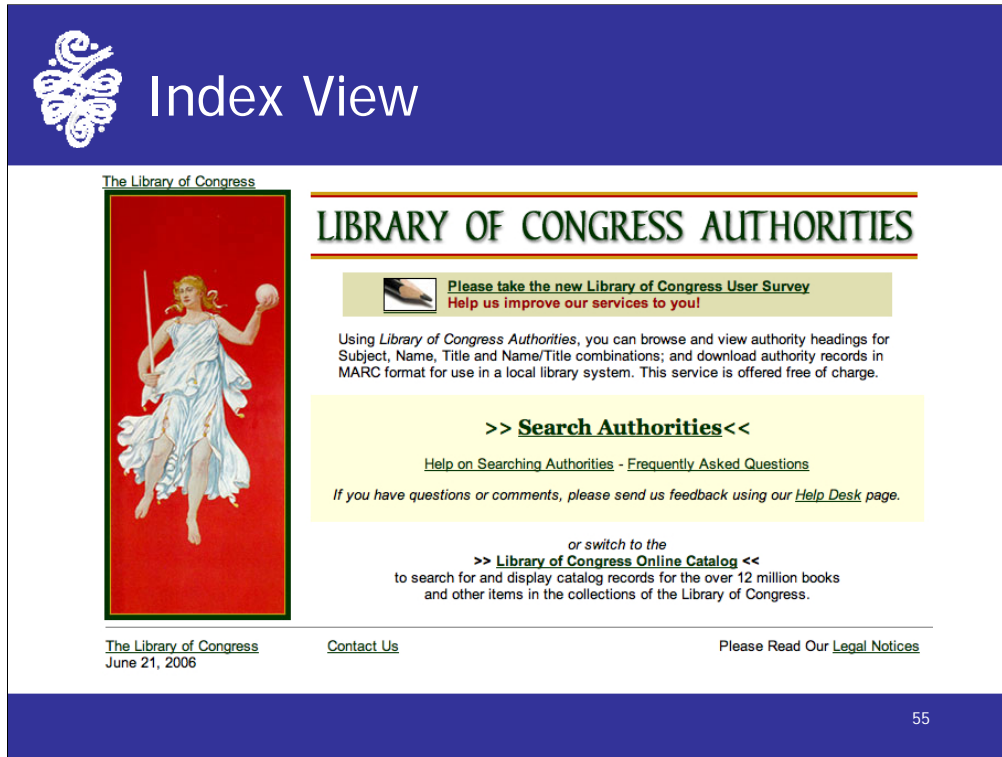
The main content area is white and contains the following elements:

- Navigation:** A vertical list on the left side includes 'Conducting Research', 'Research', 'Scholarly Activities', 'About the Research Institute', and 'Institute'.
- Search Form:** A central box titled 'Search the TGN' contains three input fields: 'Find Name:', 'Place Type:', and 'Nation:'. Each field has a 'Lookup' button below it. To the right of the form are buttons for 'Search', 'Clear', and 'Help'.
- Related Sections:** A sidebar on the right lists several links: 'Learn about the Getty Vocabularies', 'Contribute to the Getty Vocabularies', 'Frequently Asked Questions', 'Getty Vocabularies Download Center', and 'Editorial Guidelines'.
- Search Tips:** A section below the search form provides instructions: 'For the Find Name field, you may use **AND** and **OR** [e.g., 1) san carlos, 2) carlos OR charles, 3) carl\* OR charl\*, 4) san AND carlos, 5) carlos AND (san OR saint), 6) (carlos OR charles) AND (san OR saint)]. Boolean operators must be in all caps (AND and OR).'.

The page number '54' is located in the bottom right corner of the screenshot.

Getty Thesaurus of Geographic Names  
One of the contributing CVs.

Optional.



The screenshot shows the "Index View" of the Library of Congress Authorities website. At the top left is the Library of Congress logo. The main heading is "LIBRARY OF CONGRESS AUTHORITIES". Below this, there is a survey prompt: "Please take the new Library of Congress User Survey Help us improve our services to you!". A central text block explains the service: "Using Library of Congress Authorities, you can browse and view authority headings for Subject, Name, Title and Name/Title combinations; and download authority records in MARC format for use in a local library system. This service is offered free of charge." Below this is a yellow button labeled ">> Search Authorities <<". Further down, there are links for "Help on Searching Authorities - Frequently Asked Questions" and a feedback link: "If you have questions or comments, please send us feedback using our Help Desk page." At the bottom, there is a link to "Library of Congress Online Catalog" with the text: "or switch to the >> Library of Congress Online Catalog << to search for and display catalog records for the over 12 million books and other items in the collections of the Library of Congress." The footer includes "The Library of Congress June 21, 2006", "Contact Us", and "Please Read Our Legal Notices". The page number "55" is in the bottom right corner.

Library of Congress Authorities  
One of the contributing CVs.

Optional.



## Displays: Discussion

- What displays did you like?
- Which would be most helpful to your user?
- What thesauri have you used before and how was your experience?
- Print or web?

56

Try to make this a critical discussion of what works and what does not, as well as keeping in mind the variety of uses and users they are trying to serve. The goal of this is to open them to the possibilities for their own controlled vocabularies.

Please feel free to go back and ask to see CVs from the presentations or that the students have some experiences from their own uses of CVs.

Please bring in print examples if you can find some to let them peruse. AAT's 3 volumes are a great example of how print thesauri were trying to do what technology can do in a much smaller space.



Note: In the 9/06 class, session 7 only required 50 minutes (and session 8 was also short)

NOTE: A new exercise was added to the end of the session on 3/06, and will probably bring the time up to a full hour and half if you give them enough time to draft the front matter for their thesaurus...



## Controlled Vocabulary Review

- What?
  - What Controlled Vocabulary is right for you?
- When?
  - When should the CV be developed and implemented?
- Why?
  - Why is this CV a necessary development project?
- How?
  - How is the CV going to be developed?



## Thesaurus Design Questions

- Is a controlled vocabulary really necessary?
- What is the lowest level of vocabulary that will get the job done?
- Will natural language searching be sufficient?
- Will an interface design improvement alleviate the need for a controlled vocabulary?
- Will there be more than one indexer?
- Is someone available with the time and the skills to develop a thesaurus?
- Will someone be available in the future?

3

Ref: Z39.19 section 11.1 (partial)

Constructing a controlled vocabulary is expensive – it takes a lot of time. Make a thorough search for existing thesauri. Developing your own vocabulary for an area that already has an established, standard vocabulary is reinventing the wheel.

Finding vocabularies that overlap your scope or have an analogous underlying structure will help you get a good start on building your vocabulary [good place for a personal story]. For instance, when researching Nonprofit Management, we came across an overlapping taxonomy for Grantmakers, which helped to narrow our scope and to determine that a thesaurus was actually necessary. Other business thesauri helped us develop an underlying structure.



## Project Justification

- Cost of finding (time, frustration)
- Cost of not finding (bad decisions)
- Cost of training (staff turnover)
- Value of discovery (related information, browsing)
- Language is ambiguous – synonyms, abbreviations, acronyms, misspellings, homonyms, antonyms, etc.

4

Ref: Peter Morville of Semantic Studios (partial)

Specific objectives – more than a vague goal of usability.

“Power users find great value in using a known, granular indexing language that can surface the most relevant items and filter out items of peripheral or no interest. . . With the increasing quantities of content available to users, and with prospects of these quantities growing exponentially, information users cannot jeopardize their productivity by spending lots of time hunting for the information they require to do their work.” *Jan Sykes, Information Management Services*

“Our ability to create information has substantially outpaced our ability to retrieve relevant information.”  
Judi Vernau, SchemaLogic





## Project Justification

- What are the specific objectives of the project?
- Are essential objects hidden in a lot of chaff?
- Are a few good objects sufficient? Or is it necessary to find the best, the one that makes a difference, or everything on a topic?
- Use easily understood terms like *common vocabularies* rather than technical terms like *taxonomies*
- Stories tell it best.



## Project Justification

- “Users of ... intranets frequently express frustration with how much time it takes to find items—both when searching for known items and when browsing to see if items on a particular topic exist in the system. . . Browsing and search functions are much enhanced if the indexing and topic hierarchy, or taxonomy, make sense to the user and are customized to reflect the content of the source documents.” Jan Sykes, Information Management Services, February 2001



## Project Justification

- “Power users find great value in using a known, granular indexing language that can surface the most relevant items and filter out items of peripheral or no interest.” Jan Sykes, Information Management Services, February 2001
- “Keyword search captures only 33% of relevant information.” Chris Wilkie, BBC Information and Archives, Sept. 2002



## Project Justification

- “Most of the complaints we get are due to the way users search – they use the wrong keywords.” Must search stink?, Forrester, 2000
- “40% of search failures come from customers and information providers using different terms.” The Business Benefits of Taxonomy, Judi Vernau, SchemaLogic, Oct. 2005



## Project Justification

- “Knowledge workers spend 35% of their productive time searching for information online, while 40% of the corporate users report that they cannot find the information they need to do their jobs.” Working Council of CIOs, Business Week, Feb. 27, 2001



## Process of CV Design

- Understand user and organizational needs
- Define the subject scope
- Identify sources of 'raw' vocabulary
- Harvest terms (wordstock) that are likely to be search terms in the field
- Group the terms into broad categories, subcategories and sub-subcategories
- Establish relationships
- Collect feedback and revise until stable



11.1, <sup>10</sup>p. 90

If it looks at all likely that you may develop a controlled vocabulary, begin tracking all decisions *immediately*.

Who are the users?

- How expert are they in the field?
- Do they understand the use of controlled vocabularies?
- Do they prefer natural or controlled vocabulary?



## Who are the users?

- How expert are they in the field?
- Do they understand the use of thesauri?
- Do they prefer natural or controlled vocabulary?

11

Your users are very important!

They will serve as a source of wordstock, they will help you revise your structure by bouncing ideas off of them, and they are the people you require buy-in from to make the thesaurus a living, breathing, useful document.

Your prospective users might be people you work with, or they might be someone you need to consult with from the outside. How you approach them is up to you, but you do want to find out what their current practices are, in order to make the thesaurus work for them.

If you yourself are a prospective user that gives you some advantages and some disadvantages. You have intimate knowledge of the material, but you may be so entrenched that you assume implicit knowledge is being made explicit. It might be good to find a colleague or collaborator who is not within your field to make the implicit explicit.

Recording this information up front will help you remember it later on when you are trying to remember why you made the decision to only cater to X group of people.



## Defining Scope

- Core topics immediately pertinent to the main subject (most terms, more specific)
- Fringe topics supportive of the main area (fewer terms, less specific)
- Other
  - Include proper names?
  - Display?
  - Printed? Online?

12

This step can be completed in conjunction with the next three steps, dealing with identifying sources, collecting, and grouping your word stock.

Again, think about your user group and the subject your focusing on.

Defining the scope will eventually help you map out your topic, which will in turn help you group your topics.

The definition of your scope will go into the introductory material of your thesaurus.

Always remember to continuously revisit your central topic, and be prepared to hold onto any material that is not central, for later review. You may find later that they really do have a place, then again you may not.

It is very easy to want to expand your scope, but there is a trade-off between the size of the scope and how well your thesaurus will reflect the topic in question.





## Literary vs. User Warrant

- Identify sources for the raw vocabulary
- Sample literature read (etc.) by users
- Sample general literature (conference proceedings, core journals)
- Sample user questions
- Existing thesauri, dictionaries, encyclopedias
- Titles, abstracts of articles in the field

13

It is up to you to decide what counts as a source for your thesaurus, people or texts or both.

Within the field only published material is considered valid, in which case you should keep to document-able sources. In this case people within the field could be used to find the most reputable or established sources, for which they would definitely deserve some mention in your acknowledgements!

Once you make this decision you should include it in the policy for the creation and up-keep of your thesaurus.

Keep our sources of wordstock can also be sources for structure. Tables of Contents will give you a very good idea of logical organization of a given field or topic.



## Refine the raw vocabulary

- Be sure terms are acceptable to users
- Distinguish among words with different meanings
- Give instruction, clarify how terms should be used, add scope notes
- Construct and identify facets
- Check and add relationships

14

Are these terms the ones that users would actually use? If not, then you might want to find out what terms they actually use!

The development of a thesaurus is an iterative process, where you are constantly bringing your product back to the users for a reality check.

This is where having a clear definition of your user and scope is also very important, because you will not be able to please everyone, but hopefully these initial decisions about the content of your thesaurus will make it apparent who is the highest priority and maintain your focus throughout the project.

This second point may seem obvious, but it betrays one of the main reasons we create thesauri -- to show term relationships and equivalencies. Some of your words may be synonyms, and it will make your sorting and organizing job easier if you put those synonyms together up front -- giving you fewer terms to organize.

The use of scope notes is up to the designer, but from the new users' perspective they can be invaluable. A thesaurus is a knowledge management and transfer system, which make implicit relationships in a domain explicit. Thus this is the perfect resource for people entering the field to get up to speed.

Scope notes can also record your design decisions for you, so that you are able to understand well after the fact, why you thought this word has a whole-part relationship with the broader term.

Scope notes will also make this iterative back and forth with the user more productive and focused, because they will be able to validate your idea of the scope of a term, but if your scope is in your mind only they may not ever see it, and if they cannot see it they cannot validate it.

We will go into faceting your word stock in a moment. But for the purposes of refining, you need to know that how you break up the words into facets may not be how everyone or even one of your users will group the terms.

There are a number of ways to see if the structure makes sense to the user. Again you can use your users as for iterative refinement -- they can normally see problems right away. As with any project, the deeper and deeper you get into development, the more you will need to seek out the opinions of your users or of colleagues who can question your motives behind the structure.

Another way to check the structure is to see if relationships you know exist hold across the structure. This can be an iterative process as well and sometimes you will need to make compromises between the structure and the relationships.

While all of this is going on you should be adding your relationships. They are what make a thesaurus so powerful, and the more feedback you can get from your users, the more powerful they will be.



## Involving Users

- More user involvement = better suited to use
- Take every opportunity to involve users
  - Start from user search logs to find commonly used terms
  - User experience focus groups
  - Prototyping
  - Solicit community feedback
  - Online discussion groups
  - Surveys
  - Observation
  - Term submissions

15

These are some of the ways you can involve the users in the development of your controlled vocabulary. We need to stress the iterative nature of CV design, constantly cycling between developing and feedback. Once the CV is stable, you will continue to collect input from users in the form of term submissions.

Your users are very important!

- They will serve as a source of wordstock, they will help you revise your structure by bouncing ideas off of them, and they are the people you require buy-in from to make the thesaurus a living, breathing, useful document.
- Your prospective users might be people you work with, or they might be someone you need to consult with from outside. How you approach them is up to you, but you do want to find out what their current practices are, in order to make the thesaurus work for them.
- If you yourself are a prospective user that gives you some advantages and some disadvantages. You have intimate knowledge of the material, but you may be so entrenched that you assume implicit knowledge is being made explicit. It might be good to find a colleague or collaborator who is not within your field to help make the implicit explicit.



## Interoperability

- Searchers want to search multiple databases at once
- Indexers want to use a vocabulary they are familiar with to index objects in a different domain
- Content producers want to merge multiple databases indexed using different vocabularies
- User communities want a single thesaurus that spans multiple domains
- International organizations want a single vocabulary that supports searching in multiple languages



10.1, <sup>16</sup>p. 82

Interoperability (Z39.19 section 10 & Appendix D)

As we've seen in the exercises in this course, controlled vocabularies are most effective when developed for and used within a clearly defined, limited domain. Searchers, indexers and the information objects themselves often want to straddle multiple thesauri.



## Thesauri can differ in:

- Specificity
- Treatment of synonyms
- Pre- vs. post-coordination
- Relationships
- Warrant
- Scope




10.2, <sup>17</sup>p. 83

- Because of inherent differences in the specificity or granularity of the terms used to index content, searching a more specific or granular term selected from one controlled vocabulary may not return the desired results in another database because the indexing was done at a broader level.
- Different vocabularies will often treat synonyms differently, with one thesaurus treating a concept as an entry term, while the other treats it as a separate preferred term, while another doesn't include the term at all.
- Vocabularies are often, if not always, developed to operate within a specific search environment. One may be developed to assist navigation in a system that uses faceted browses of the search results, refining the results by progressively segmenting the previous result set using additional terms. An entirely different vocabulary may be developed for use in a system in which users are expected to create a very precise, specific search using multiple terms.
- The presence or absence of hierarchical relationships, and Related Terms, will also make merging thesauri more difficult.
- Vocabularies developed for different user groups will often reflect differences in literary, organizational and user warrants. The result can vocabularies which have little commonality among terms, or which could assign different meanings to the same term without using distinguishing qualifiers.
- The omission of some areas of a subject field can also make searching across thesauri less successful.


What is the point of a thesaurus? To organize terms in a way that reflects a specific domain.

Discussion: When you start combining these different domain views of the same terms, does it make the thesaurus more powerful or more distracting?



## Methods of Integration

- Mapping
- Switching language
- Integration
  - Unified Medical Language System's (UMLS) 3 main components:
    - Metathesaurus ↔ concepts
    - Semantic Network ↔ categories
    - SPECIALIST Lexicon ↔ indices
- Super-language
- Merging



App. D, p. 142

Ref: Z39.19 Appendix D

- Mapping – linking terms directly between thesauri. There is a map, or crosswalk, available that allows LCSH search terms to be converted to AAT terms. This approach generally requires extensive human intellectual analysis of the terms in both the source and target controlled vocabularies.
- Vocabularies can also be integrated by selecting a common switching, or intermediate language and mapping the terms in the vocabularies to the same intermediate language. The intermediate language is often a classification system such as *Dewey Decimal System*. Each term in the separate vocabularies are mapped to the closest equivalent classification number in Dewey, which then allows the term in one language to be converted to a term in a different vocabulary. This approach allows efficient later integration of additional vocabularies. It also allows display of terms from all related vocabularies clustered under the concepts as organized in the switching language.
- Individual controlled vocabularies have been integrated to form a larger controlled vocabulary that is applicable to all of the specific domains. The *Unified Medical Language System* (UMLS) is an example of this approach. The UMLS is comprised of over 1 million biomedical concepts and 5 million concept names, all of which are from over 100 controlled vocabularies and classification systems used in patient records, bibliographic, administrative health data and full text databases.
- A generic controlled vocabulary structure can be defined, and then used as a template for creating multiple, highly specific controlled vocabularies to meet the needs of diverse content domains. The Library of Congress used this approach in developing the *Thesaurus for Graphic Materials* (TGM) and the *Thesaurus for the Global Legal Information Network*.
- Another method is simply to merge the terms in multiple vocabularies and create a single alphabetical list of all the combined terms, usually suppressing duplicates.



## Simple Knowledge Organization System

Term:	Economic cooperation
UF:	Economic co-operation
BT:	Economic policy
NT:	Economic integration European economic cooperation European industrial cooperation Industrial cooperation
RT:	Interdependence
SN:	Includes cooperative measures in banking, trade, industry etc., between and among countries.

19

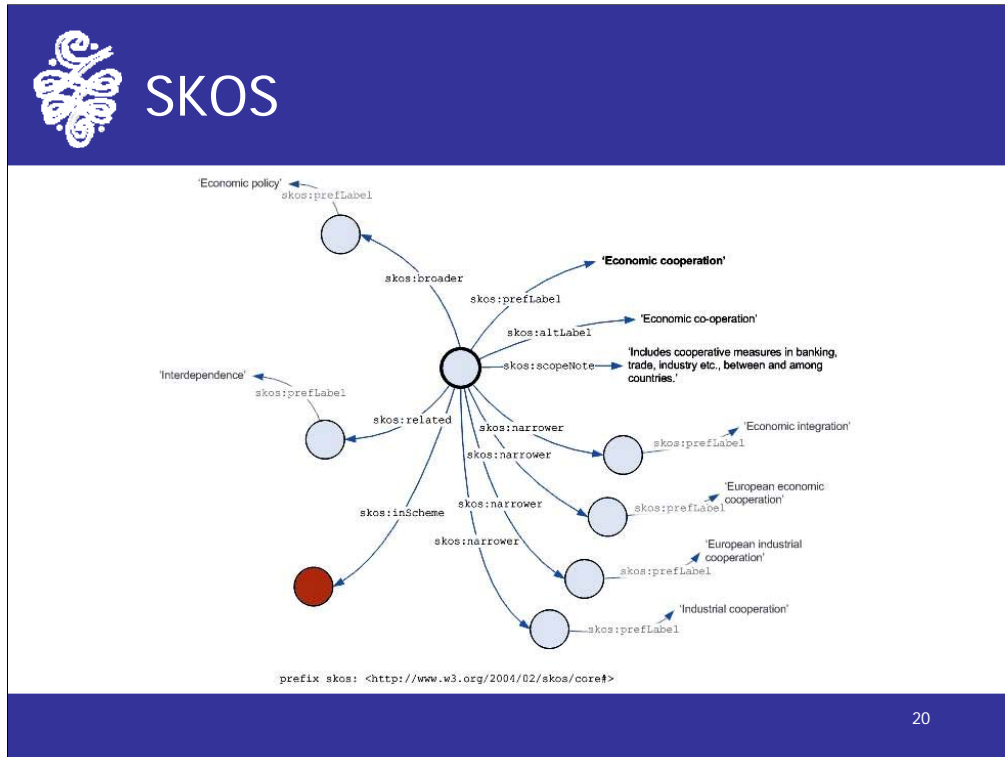
The Semantic Web, based on the Resource Description Framework (RDF), is often described as providing a common framework that allows data to be shared and reused across application, enterprise, and community boundaries. Using RDF is intended to allow data to be linked to and/or merged with other RDF data by semantic web applications.

SKOS Core is a model for expressing the structure and content of concept schemes (thesauri, classification schemes, subject heading lists, taxonomies, terminologies, glossaries and other types of controlled vocabulary) in a machine-understandable way, within the framework of the Semantic Web. SKOS stands for Simple Knowledge Organization System.

There are three methods of encoding SKOS data, XML, N3/Turtle, and N-Triples. We will look at how a SKOS term record is encoded in XML.

This term records is an extract from the UK Archival Thesaurus (UKAT) using standard thesaurus display format for the term 'Economic cooperation'.

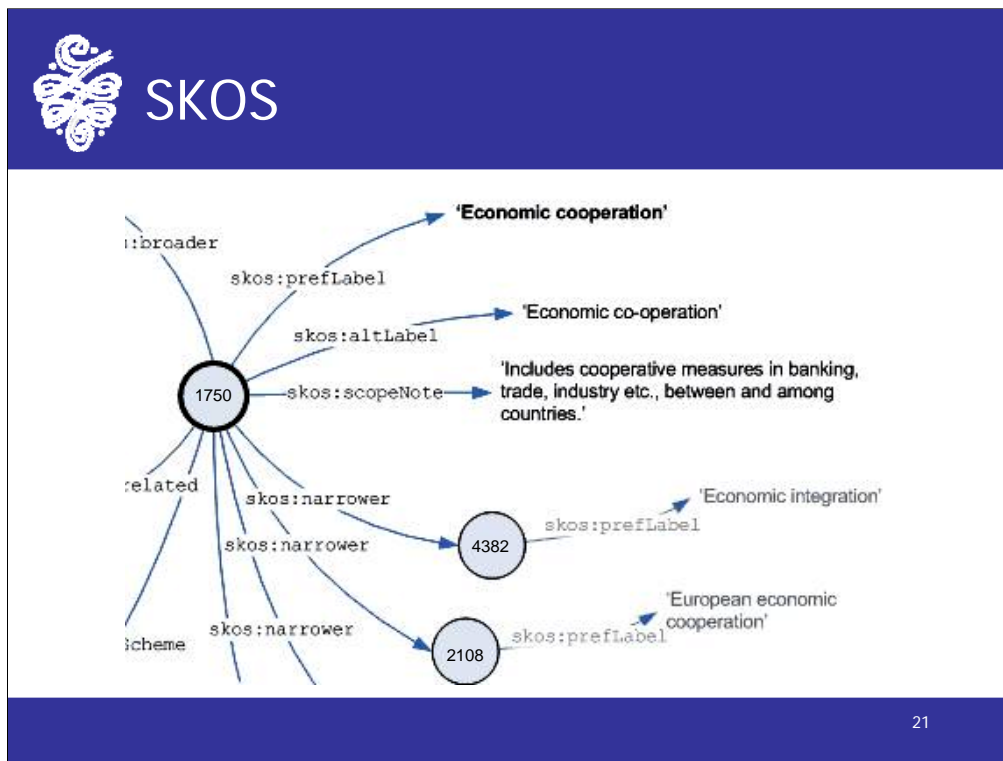
SKOS is a way of looking at how the semantic web will be able to alleviate some of the content integration issue, but it will again come down to implementation choices and the intellectual/conceptual work of organization in different domains.



This is how the same term record is displayed using an RDF graph. The central circle represents our concept which has the 'preferred label' of 'Economic cooperation'. It also shows that our concept is part of the skos defined UKAT schema.

You can also point out narrower terms, the broader term, related term, and 'alternate label', or entry term.





In this blow up of the RDF graph, we can see that concepts are actually linked using digital identifiers, or URIs. Each concept from the UKAT has a URI assigned, which is used to refer to a concept unambiguously from any context. For example, the URI:

<http://www.ukat.org.uk/thesaurus/concept/1750>

... denotes the concept from the UKAT whose preferred term is 'Economic cooperation'.

In the RDF graph, the 'skos:' prefix is short hand representing the actual SKOS prefix:

`<http://www.w3.org/2004/02/skos/core#>`

So another method of referring to our 'Economic cooperation' UKAT term record is:

`<http://www.w3.org/2004/02/skos/core1750>`



SKOS

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#">

  <skos:Concept rdf:about="http://www.ukat.org.uk/thesaurus/concept/1750">
    <skos:prefLabel>Economic cooperation</skos:prefLabel>
    <skos:altLabel>Economic co-operation</skos:altLabel>
    <skos:scopeNote>Includes cooperative measures in banking, trade, industry etc.,
      between and among countries.</skos:scopeNote>
    <skos:broader rdf:resource="http://www.ukat.org.uk/thesaurus/concept/4382"/>
    <skos:narrower rdf:resource="http://www.ukat.org.uk/thesaurus/concept/2108"/>
    <skos:narrower rdf:resource="http://www.ukat.org.uk/thesaurus/concept/9505"/>
    <skos:narrower rdf:resource="http://www.ukat.org.uk/thesaurus/concept/15053"/>
    <skos:narrower rdf:resource="http://www.ukat.org.uk/thesaurus/concept/18987"/>
    <skos:related rdf:resource="http://www.ukat.org.uk/thesaurus/concept/3250"/>
    <skos:inScheme rdf:resource="http://www.ukat.org.uk/thesaurus"/>
  </skos:Concept>

</rdf:RDF>
```

22

Examine the XML record for the UKAT term record 'Economic cooperation'.

As can be seen, encoding a term record in SKOS is similar to encoding a term record in MARC or any other defined container. When creating systems that need to search across multiple thesauri, the same issues of specificity, scope, treatment of synonyms, etc. remain unresolved. You don't know if 'Economic cooperation' within UKAT is intended to be applied to the same concepts as the same term in a different controlled vocabulary.



## Testing & Evaluation: Methods

- Heuristic Evaluation
  - Evaluation by an expert or a panel of experts
- Affinity Modeling
  - Task a sample of users with organizing your terms
  - Compare to your own organization of the terms
- Usability Testing
  - Holistic evaluation of the information system, including the content, interface, etc.



11.2, <sup>23</sup>p. 94

Ref: Z39.19 section 11.2

Testing methods:

- Heuristic Evaluation – an expert or panel of experts is asked to evaluate the controlled vocabulary. Such an evaluation can be information and qualitative (each expert sharing their reactions and opinions) or formal and quantitative (all experts assign scores according to a shared list of criteria).
- Affinity Modeling – a representative sample of users is asked to sort a collection of cards (i.e., card sorting) each marked with terms from the vocabulary into groups. Users may also be invited to score equivalent and related terms according to level of similarity.
- Usability Testing – holistic evaluation of the information system can yield insight into the efficacy of the controlled vocabulary.

Note: It is often difficult to isolate the quality of a vocabulary from the performance of other elements of an information system such as content, technology, search engine, and interface.



## Testing & Evaluation: Discussion

- Why test a controlled vocabulary?
- What are some useful criteria for evaluating a controlled vocabulary?

24

Two Major reasons for testing and evaluating controlled vocabularies:

- To determine if the controlled vocabulary provides adequate search results (i.e., high relevance and recall) for most users, and
- To determine if the controlled vocabulary matches users' expectations of the terms contained. For example, if the entry vocabulary is not sufficiently rich, users may not discover their desired terms.

To evaluate a controlled vocabulary, it is useful to ask questions such as:

- Does it contain terms that are synonymous or equivalent without indicating the relationship?
- How well does the display reflect the conceptual organization?
- Are history notes provided where needed?
- Are all needed facets included?
- Are the terms up-to-date with accepted usage?
- How complete are the entry term synonyms included in the alphabetical list?
- Are term forms consistent?
- Does term form adhere to common usage?
- Are useful scope notes included?
- Are cross-references active hyperlinks?
- Is there an online term index or search by keyword?
- Is the controlled vocabulary database searchable? How does the search work? What fields are searched?
- Were standard procedures followed in constructing the controlled vocabulary?

Internal consistency is key.

If your scope says that the thesaurus covers world cheese and you have only American cheese, then it is not what it says it is.

If you are using your users for iterative design and feedback, then your thesaurus should evaluate well.



## Upkeep & Maintenance

- Controlled vocabularies as living entities needing
  - New material added
  - Outdated material removed
  - Changes made
- Requires a long-term maintenance plan
  - Institution support and resources
  - Someone who is a maintainer
- Look to your users for input!
  - Term submissions
  - Search logs
- Anticipate change!



11.3, <sup>25</sup>p. 96

[Instructors: bring your own stories to the table here of where plans help, where lack of plans hurt, etc]



## Front Matter Exercise

- Given the thesaurus you devised in the earlier sections, write the appropriate front matter:
  - Title
  - Introduction
  - Users/scope
  - How to use
  - Term submission/schedules
- Please use this sample front matter to start from

26

At this point hand out the sample, or have them turn to the sample if included in the handouts.

Give them time for this -- 20-30 minutes -- and leave roughly 5-10 for discussion...



## Front Matter Discussion

- How does this front matter address the planning aspects of the thesaurus development?
- How does this front matter address the maintenance aspect?

27

Front matter should be the central point around which all effort is rallied. Or if you ever have a question of who the user is or whether a term should be part of the thesaurus, the front matter should be consulted, as these questions can easily be answered by your front matter. It will keep you on track through your development!

The maintenance plan should be laid out in the front matter along with the tools to allow for the thesaurus to grow and remain relevant to the field.

*THIS PAGE INTENTIONALLY LEFT BLANK  
FOR DOUBLE SIDED COPY*



# Controlled Vocabulary & Thesaurus Design

Resources & Future Directions





## Thesaurus Design Software

- Comprehensive list of Thesaurus Software
  - <http://www.willpower.demon.co.uk/thessoft.htm>
- Comparison of Thesaurus Software
  - <http://www.willpower.demon.co.uk/thestabl.htm>




## A Cautionary Note

- Thesaurus software
  - A tool for developing thesauri
  - Analogous to the functionality of a word processor for writing a book
  - Unfortunately will not do all of the work for you
- That said, it is always good to have the right tool for the job!




11.4, p. 99

This is in here because people seem to think that the software will do everything, and this is unfortunately not the case.




## Example 1: TCS-10

 **TCS-10**

TCS-10 provides you with an exhaustive range of tools to develop controlled vocabularies (thesauri, classification schema, topic maps, etc.) inexpensively and efficiently, and easily to transform any controlled vocabularies into browsable and linkable hierarchies. TCS-10 constitutes the first tier (**Build and Transform**) of WebChoir's three-tier end-to-end solution.

more...

 **ViewChoir** for TCS-10

**ViewChoir** easily supports publication of your thesauri and other controlled vocabularies (CVs) on the Internet/intranet/extranet or on a CD, to enable the your clientele to participate in collaborative building of controlled vocabularies with their candidate term submissions to enhance your CVs and to greatly enhance their online searching on the Internet/intranet/extranet through utilization of your browsable CVs.

more..

4

Most software is available for PCs and not Macs, in fact only one seems to be available for Macs.

Note from 9/06 – consider adding sample screen shots from Data Harmony's MAIstro (which combines Thesaurus Master and MAI machine aided indexer programs; <http://www.dataharmony.com>) and SchemaLogic's Enterprise Suite (<http://www.schemalogic.com>)



## Example 1: TCS-10

### **LinkChoir** for TCS-10

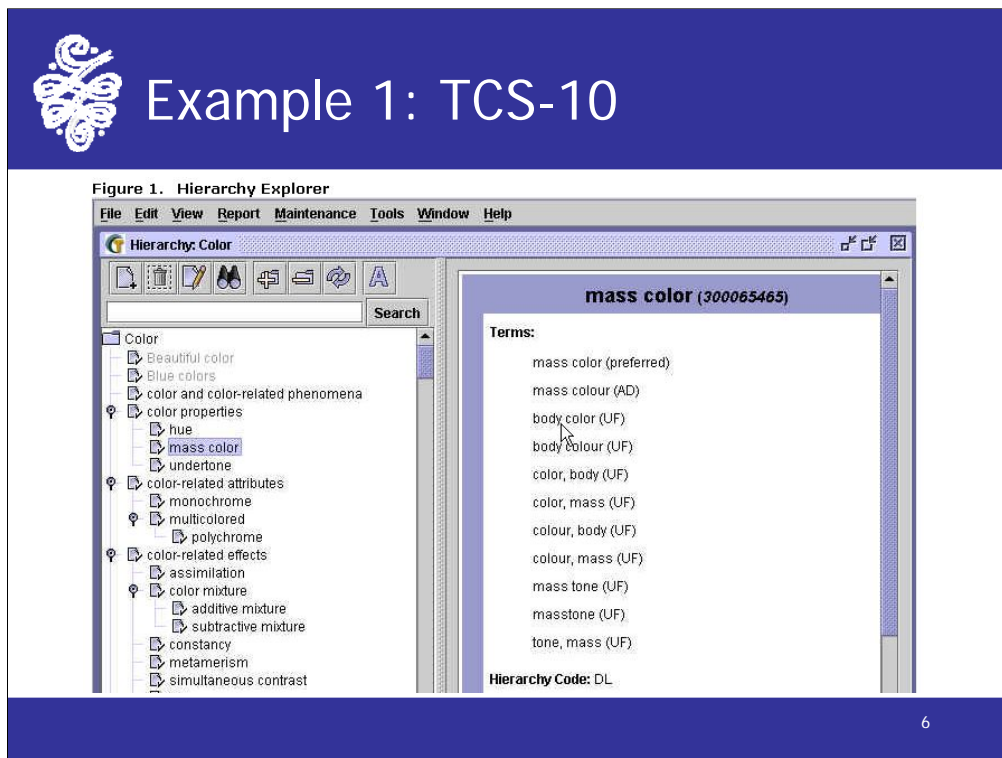
**LinkChoir** enables you to integrate any browsable controlled vocabularies, either created or transformed by TCS-10, with a variety of information sources for classification and indexing by linking the terms in your CVs with fielded records in your data warehouses. **ViewChoir** and **LinkChoir** form the second tier (**Enhance and Index**) of WebChoir's three-tier end-to-end solution.

[more...](#)

### **SeekChoir** for TCS-10

**SeekChoir** enables your clientele to create simple to complex search strategies to retrieve precise information throughout the Internet/intranet/extranet and/or data warehouses using keywords and/or browsable XML-based controlled vocabularies that are either created or transformed by TermChoir. **SeekChoir** makes up the third tier (**Search and Retrieve**) of WebChoir's three-tier end-to-end solution.

5



## Hierarchy Explorer

After you log into the system and open a hierarchy, the system will open the TCS-10 Hierarchy Explorer (HE), the main window of TCS-10. HE is composed of four areas: the Tree Display Area (TDA), Tool Bar Area (TBA), Web Search Area (WSA), and Node Detail Area (NDA) as shown in Figure 1. The TDA displays the descriptors in the selected hierarchy in a tree structure, much like Microsoft Windows products. The TBA contains command buttons you can use to manipulate the hierarchy in the TDA, the WSD allows you to search selected descriptors in the TDA against the Internet using popular search engines, and the NDA shows the detailed information of the highlighted descriptor in the TDA. In addition to these four areas, a popup menu will appear whenever you highlight a descriptor in the tree and click the right-mouse button in the Tree Display Area.

There is a lot of drag-and-drop functionality to these programs as well, allowing you to change the order of hierarchies, or the locations of terms.




## Example 1: TCS-10

The screenshot displays a software interface with a hierarchical tree on the left and a detailed view of a term on the right. The tree includes nodes such as 'simultaneous contrast', 'tint', 'tone', 'continuous tone', 'halftone', 'dropout', and 'colors'. The detailed view shows the following information:

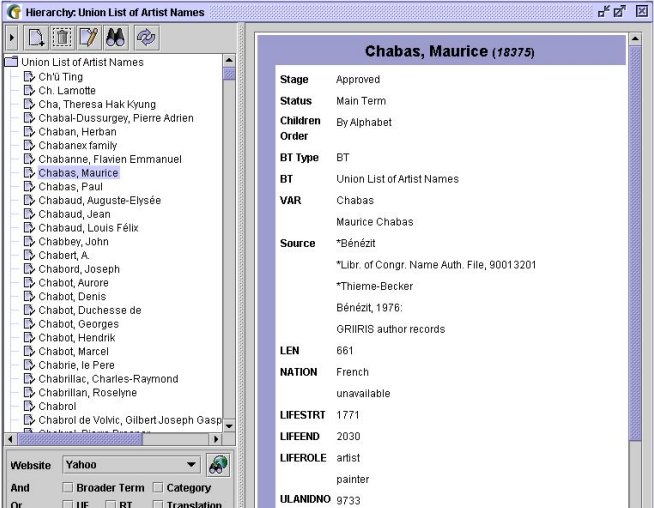
- Hierarchy Code:** DL
- Hierarchy Position:**
  - [Color](#)
  - ... [color properties](#)
  - ..... [mass color](#)
- Class Notation:** DL.00.00.07.02
- Scope Note:**

The hue of a pigment, seen as a pile of dry powder or dispersed in a vehicle and spread out in a thick, opaque film.

At the bottom of the interface, there is a status bar that reads: "'mass color' created on 03/13/2004 and last modified on 03/13/2004."



## Example 1: TCS-10



**Hierarchy: Union List of Artist Names**

Union List of Artist Names

- Ch'ü Ting
- Ch. Lamotte
- Cha, Theresa Hak Kyung
- Chabal-Dussurgey, Pierre Adrien
- Chaban, Herban
- Chabanex family
- Chabarne, Flavien Emmanuel
- Chabas, Maurice**
- Chabas, Paul
- Chabaud, Auguste-Elysée
- Chabaud, Jean
- Chabaud, Louis Félix
- Chabbey, John
- Chabert, A.
- Chabert, Joseph
- Chabot, Aurore
- Chabot, Denis
- Chabot, Duchesse de
- Chabot, Georges
- Chabot, Hendrik
- Chabot, Marcel
- Chabrie, le Pere
- Chabrilac, Charles-Raymond
- Chabrilan, Roselyne
- Chabrol
- Chabrol de Volvic, Gilbert Joseph Gasp

Website: Yahoo

And  Broader Term  Category

Or  UF  RT  Translation

---

**Chabas, Maurice (18375)**

**Stage** Approved

**Status** Main Term

**Children** By Alphabet

**Order**

**BT Type** BT

**BT** Union List of Artist Names

**VAR** Chabas  
Maurice Chabas

**Source** \*Bénézit  
\*Libr. of Congr. Name Auth. File, 90013201  
\*Thieme-Becker  
Bénézit, 1976:  
ORNIIRIS author records

**LEN** 661

**NATION** French  
unavailable

**LIFESTRT** 1771

**LIFEEND** 2030

**LIFEROLE** artist  
painter

**ULANIDNO** 9733

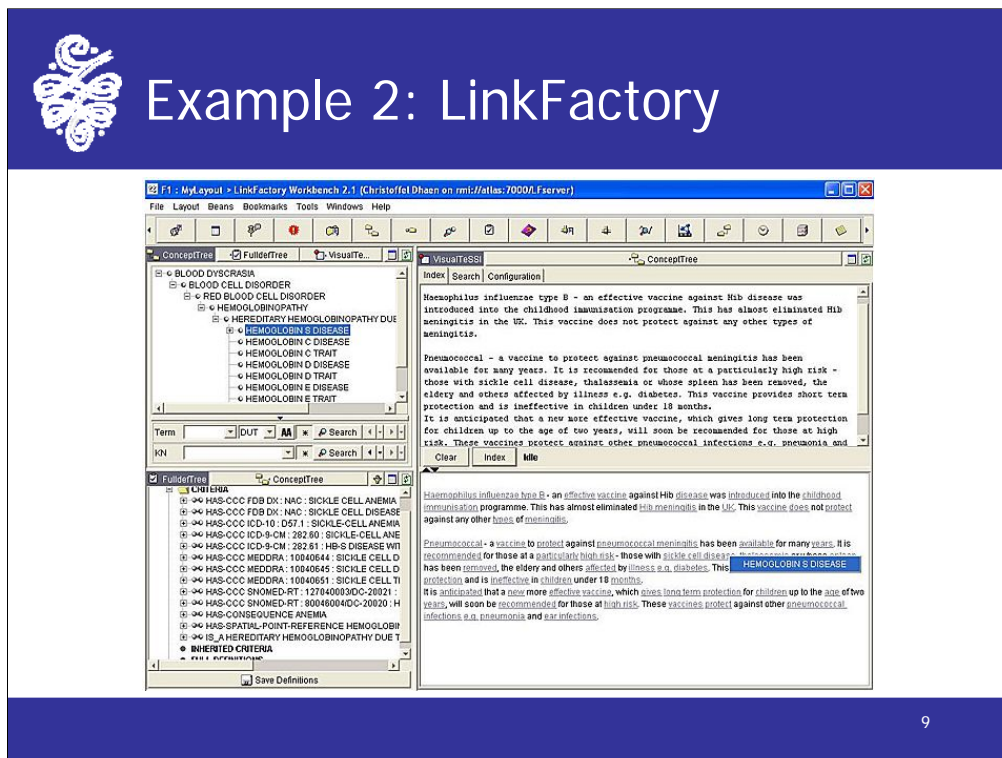
8

Union List of Artist Names – very flat structure

Why would you want to consider a more hierarchical display?

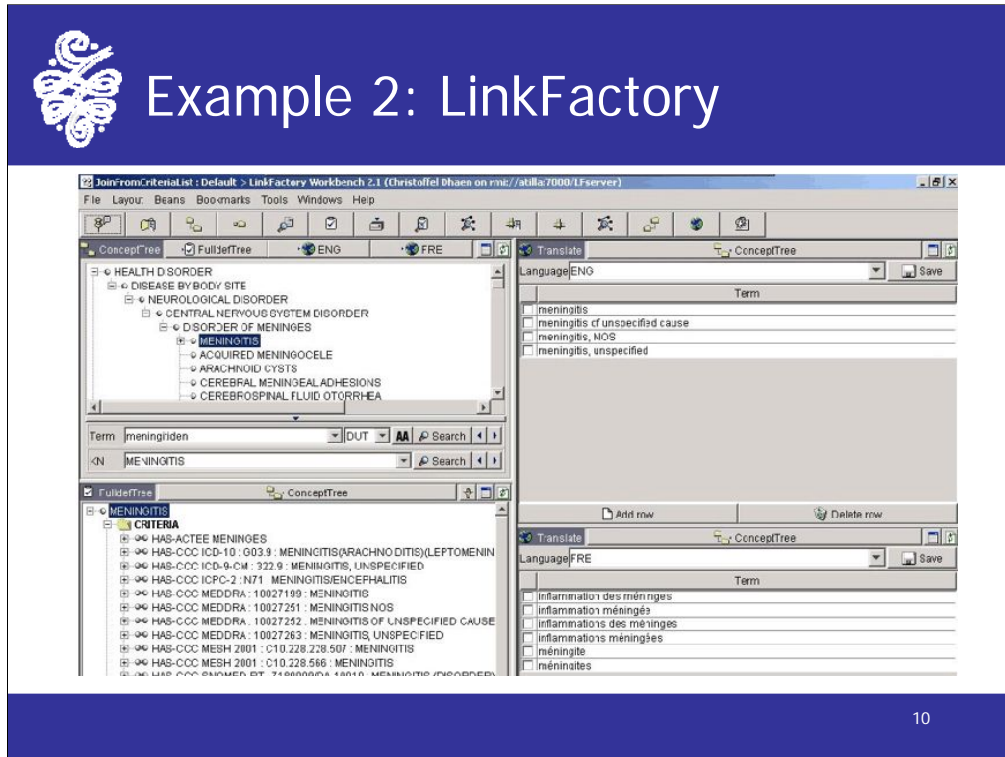
Why would you want to use a flat, alphabetic display?



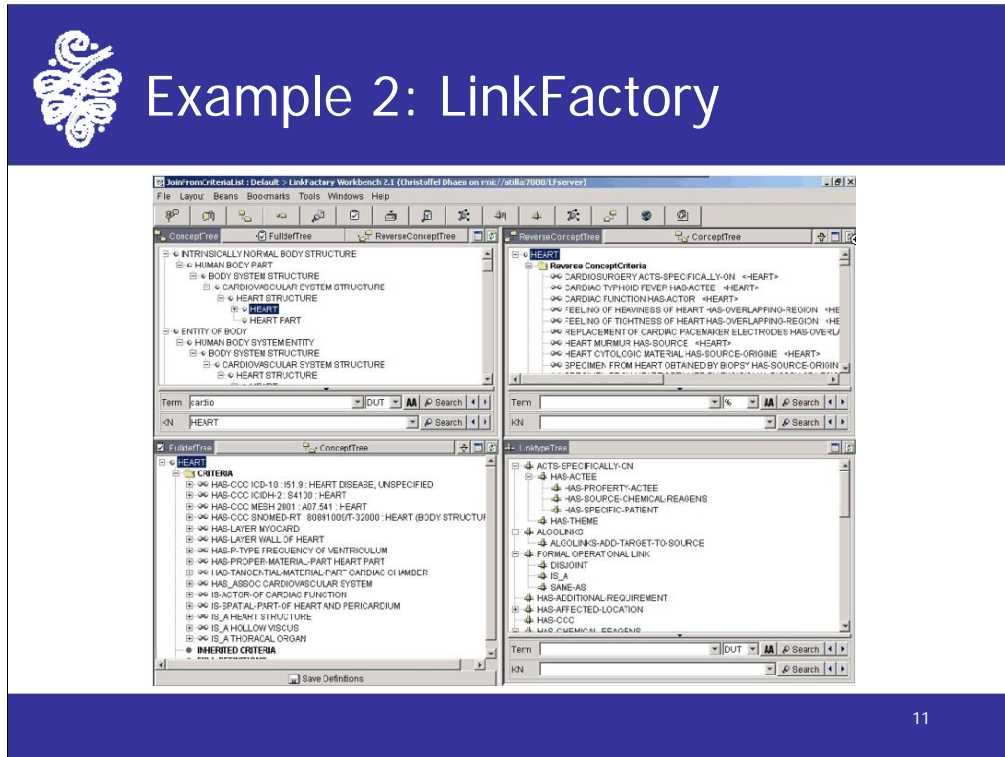


LinkFactory® is a feature-rich tool used for building comprehensive corporate terminology systems and is capable of extracting significant value out of the large stores of unstructured data hidden within corporate documents and databases. LinkFactory® is specifically designed to build ontologies exceeding millions of concepts. It includes advanced features for performing complex functions such as: 1) attaining the conceptual level of language expression, 2) managing a large scale and complex ontology and 3) defining relationships between concepts that extend beyond the simplistic "IS-A", "HAS-A"-type linkages in order to attain a semantic network.

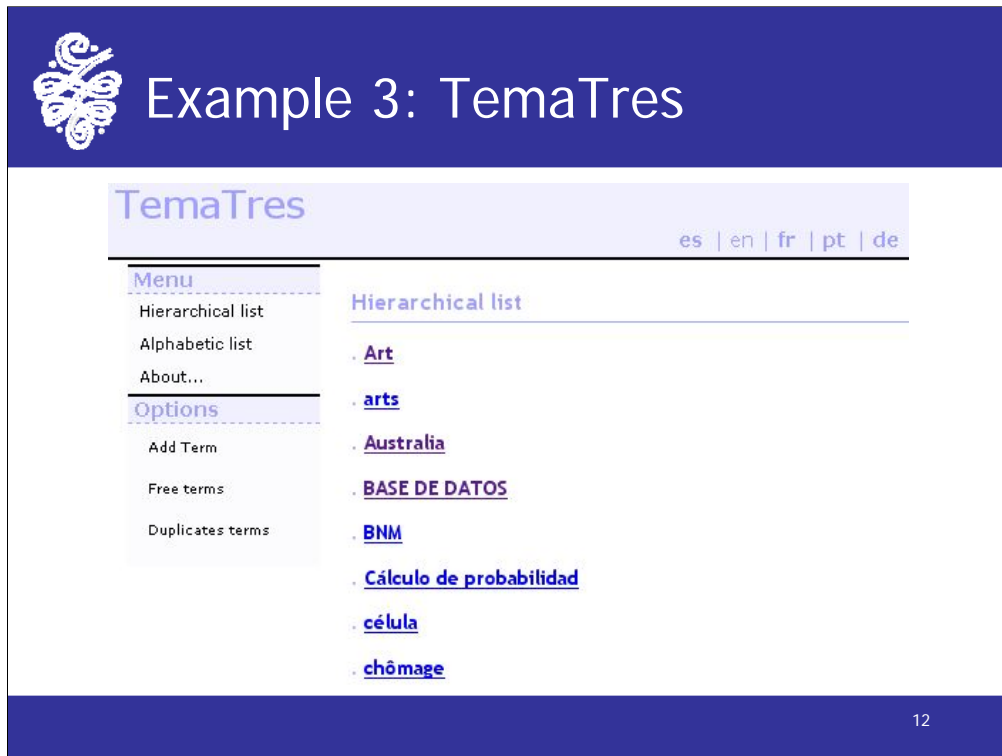
Illustration 1: Visula TeSSI



Interface for Multilingual term management



Ontology construction interface



**Example 3: TemaTres**

TemaTres es | en | fr | pt | de

**Menu**

- Hierarchical list
- Alphabetic list
- About...

**Options**

- Add Term
- Free terms
- Duplicates terms

**Hierarchical list**

- [Art](#)
- [arts](#)
- [Australia](#)
- [BASE DE DATOS](#)
- [BNM](#)
- [Cálculo de probabilidad](#)
- [célula](#)
- [chômage](#)

12

A open source web application for managing documentation languages - especially good at hierarchical thesauri and cataloguing vocabularies but also useful for navigation structures.

Features (version 0.9) include standard thesaurus relations between terms (USE/UF, BT/NT, RT/RT). No limits on number of terms, levels of hierarchy, alternative labels. Systematic or alphabetical navigation. Multilingual interface, in English, Spanish or Portuguese.

Output in SKOS-Core, Dublin Core and ZThes formats.



# Example 3: TemaTres

TemaTres es | en | fr | pt | de

**Menu**

- Hierarchical list
- Alphabetic list
- About...

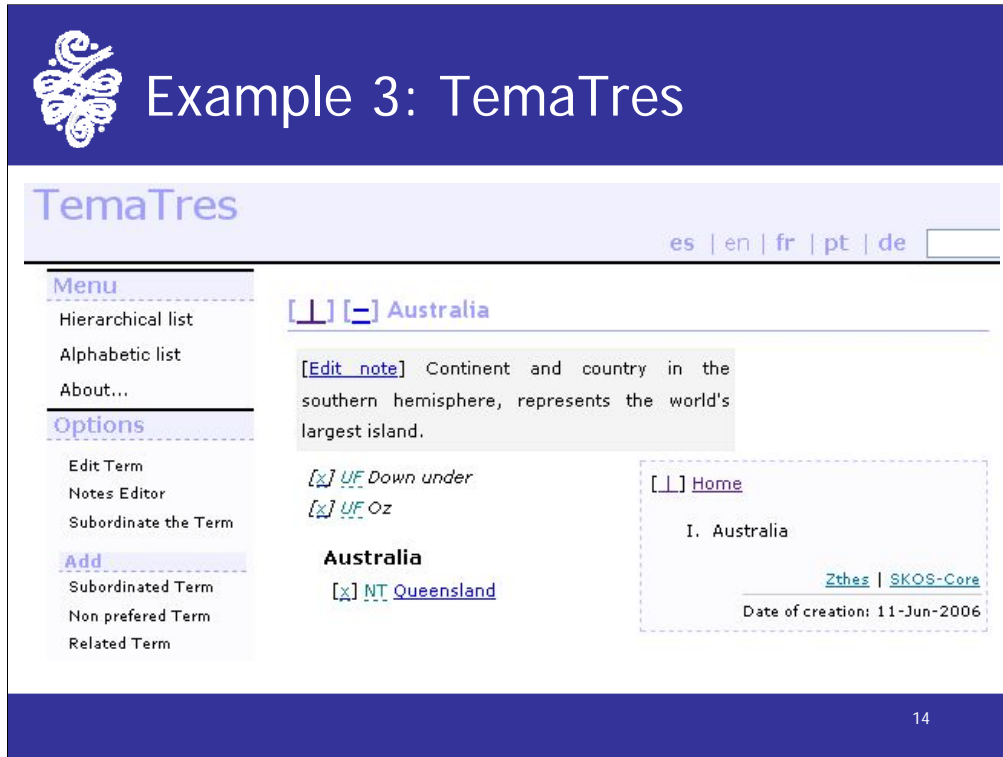
**Alphabetic list**

A · B · C · D · E · F · G · H · I · J · K · L · M · N · O · P · Q · R ·  
S · T · U · V · W

Letter A: 68 terms.

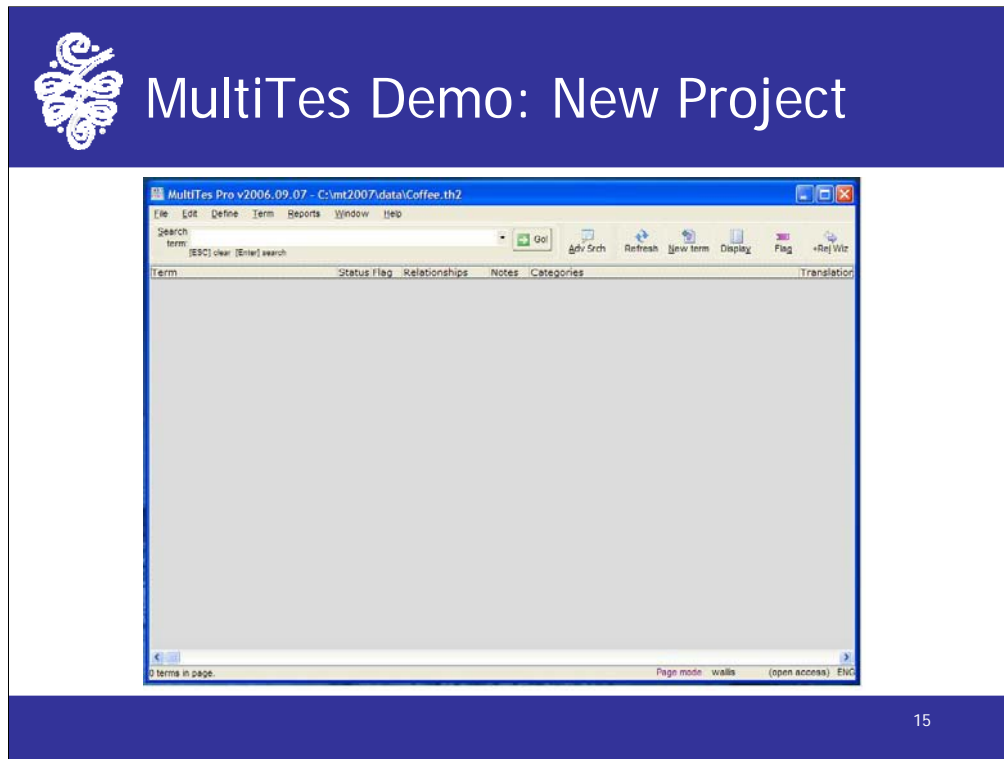
- . [acento](#)
- . [Ácidos y bases](#)
- . [Actividades bursátiles](#)
- . [Actividades económicas](#)
- . [Actividades grupales](#)
- . [Adjetivos](#)
- . [Administración de la educación](#)
- . [ADMINISTRACION](#)
- . [ADMINISTRACION DE BIBLIOTECA](#)
- . [ADMINISTRACION DE LA DEMENDA](#)

13



The screenshot displays the TemaTres web interface. At the top left is a logo of a stylized butterfly. The main header is "Example 3: TemaTres". Below this is a navigation bar with the text "TemaTres" and language links "es | en | fr | pt | de" followed by a search box. The main content area is titled "Australia" and includes an "Edit note" section with the text: "Continent and country in the southern hemisphere, represents the world's largest island." Below the note are two relationships: "[UF] Down under" and "[UF] Oz". To the right is a "Home" section with the text "I. Australia" and links for "Zthes" and "SKOS-Core". At the bottom right of the Home section is the text "Date of creation: 11-Jun-2006". On the left side, there is a "Menu" section with options: "Hierarchical list", "Alphabetic list", and "About...". Below the menu is an "Options" section with: "Edit Term", "Notes Editor", and "Subordinate the Term". At the bottom of the menu is an "Add" section with: "Subordinated Term", "Non preferred Term", and "Related Term". The page number "14" is visible in the bottom right corner.

<http://www.r020.com.ar/tematres/index.en.html>

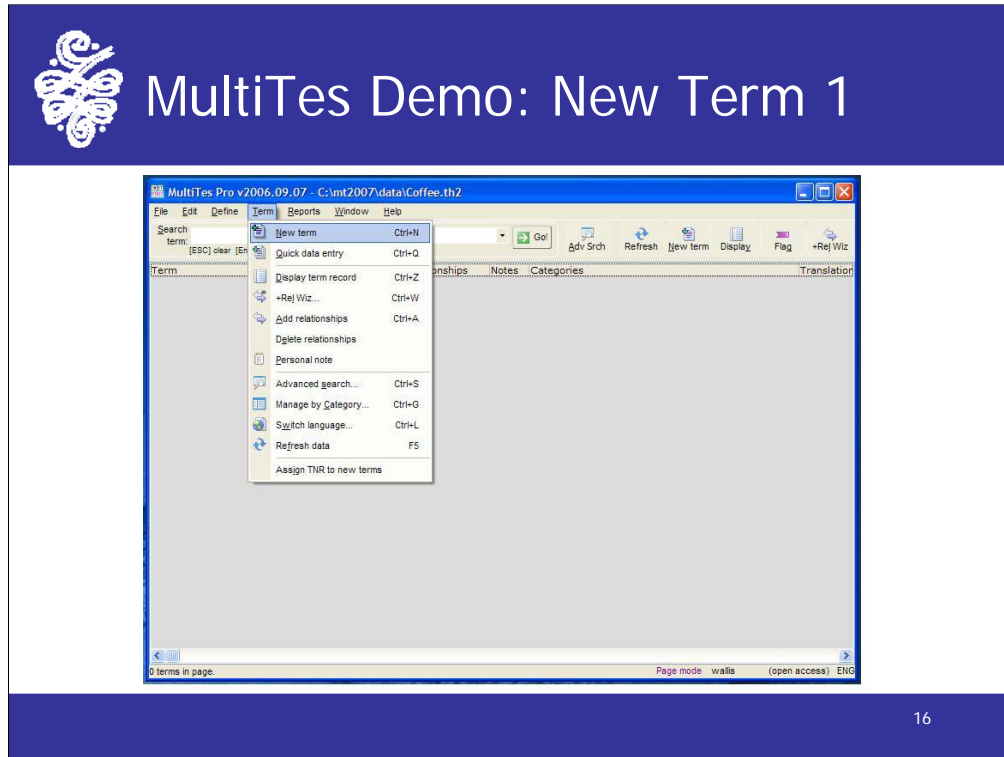


What you see when you open MultiTes. Notice that it has been saved as Coffee.th2. This is to be thesaurus of coffee terms.

Lay of the land -- large table for terms, and some tools to add new terms, or to Display/Report them or to set relationships.

NOTE: It should be noted that we are not advertising/advocating the use of MultiTes, but we do see it as an example of thesaurus softwares. It has its benefits and its drawbacks, and this is all part of the discussion.

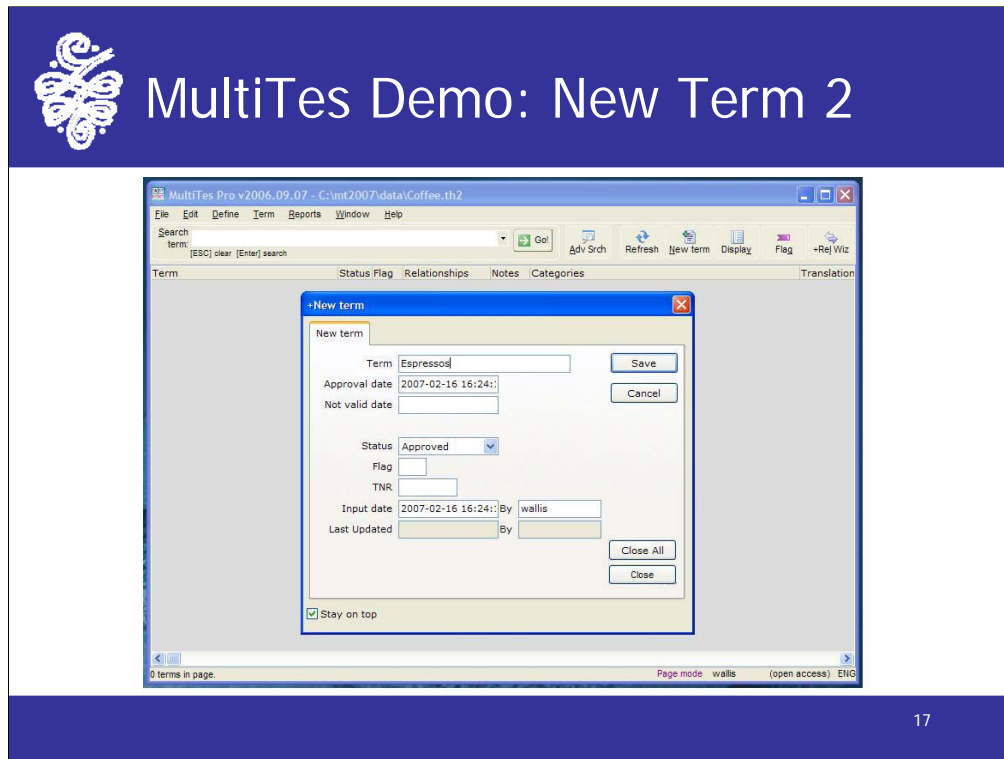
NEXT: we will add a single term by going to the "Term" menu...



16

From the Term menu we select New Term...

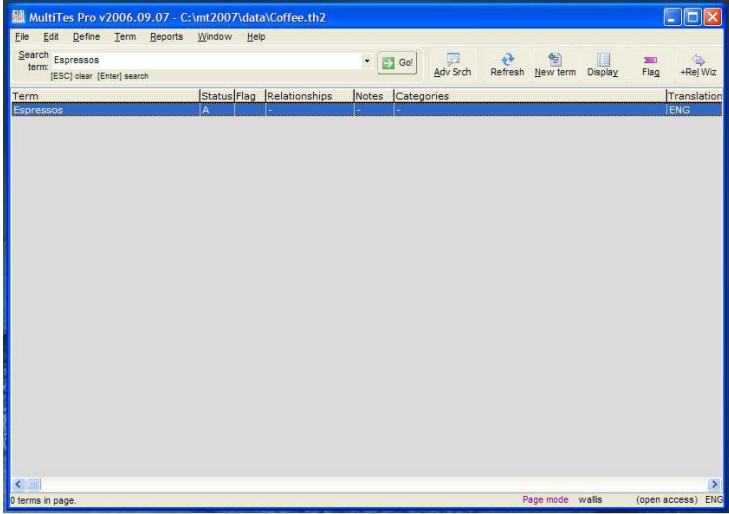




17

This opens a new window, where you are to enter the new term as well as info about the term if available, but we will skip for the moment, allowing to auto-fill the places it wants to.

Now we will save the term...

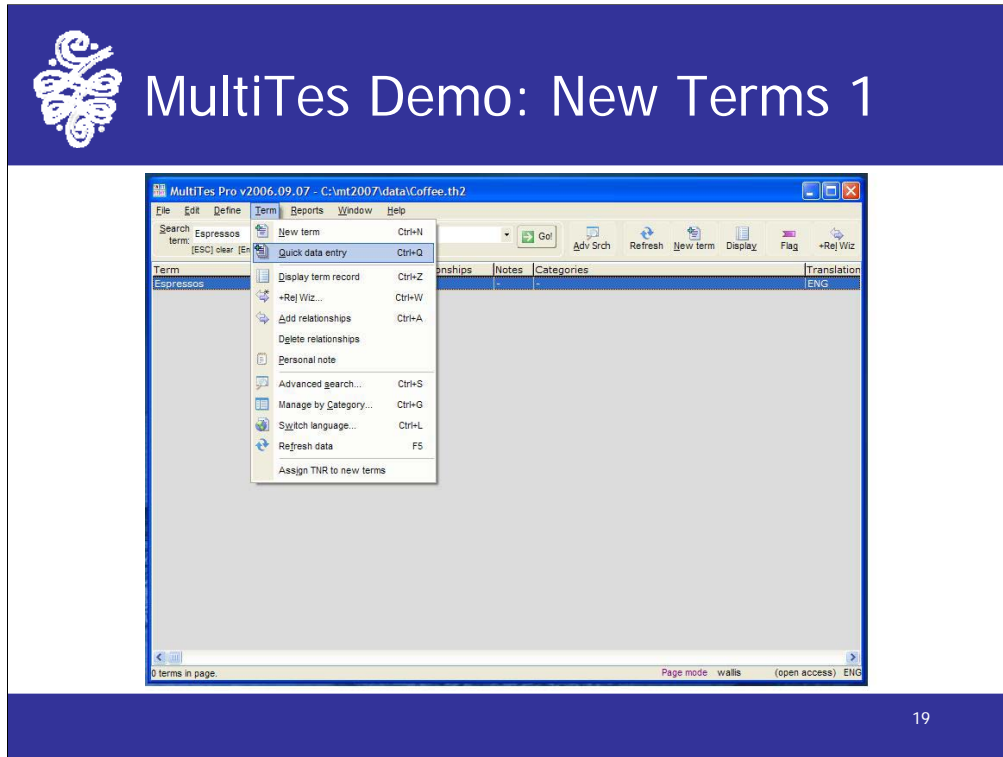


The screenshot shows the MultiTes Pro v2006.09.07 application window. The title bar reads "MultiTes Pro v2006.09.07 - C:\vnt2007\data\Coffee.th2". The menu bar includes "File", "Edit", "Define", "Term", "Reports", "Window", and "Help". The search bar contains "Espresso" and "term:". Below the search bar is a table with the following columns: "Term", "Status", "Flag", "Relationships", "Notes", "Categories", and "Translation". The table contains one row with the following data: "Espresso", "A", "-", "-", "-", "-", "IENG". The status bar at the bottom indicates "0 terms in page." and "Page mode walls (open access) ENG".

18

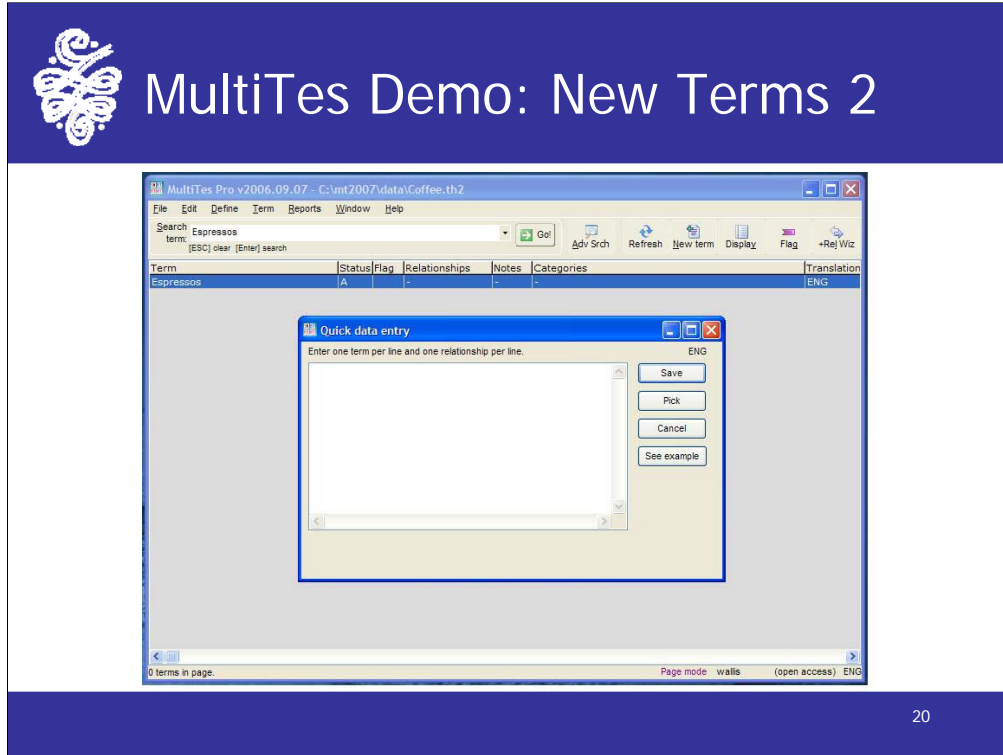
Note that the term has been added to our very short list.

There is another way to add terms, which allows you to add multiple terms. Going to the "Term" menu again...



The screenshot displays the MultiTes Pro v2006.09.07 application window. The title bar reads "MultiTes Pro v2006.09.07 - C:\vnt2007\data\Coffee.th2". The menu bar includes "File", "Edit", "Define", "Term", "Reports", "Window", and "Help". The "Term" menu is open, showing options such as "New term", "Quick data entry", "Display term record", "+Rej Wiz...", "Add relationships", "Delete relationships", "Personal note", "Advanced search...", "Manage by Category...", "Switch language...", "Refresh data", and "Assign THR to new terms". The "Quick data entry" option is highlighted. The main window area shows a table with columns for "Term", "Relationships", "Notes", "Categories", and "Translation". The "Term" column contains the word "Espresso". The "Translation" column contains "ENG". The status bar at the bottom indicates "0 terms in page" and "Page mode: walls (open access) ENG".

Now we select Quick data entry...



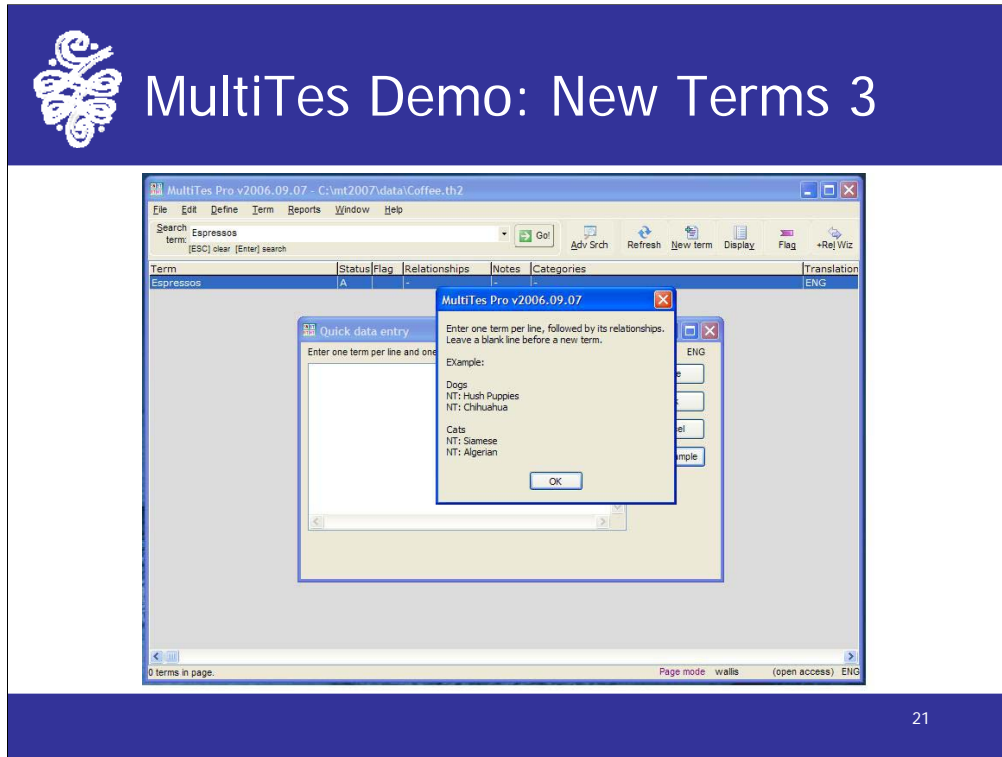
The screenshot displays the MultiTes Pro v2006.09.07 software interface. The main window shows a search for 'Espressos' with a table of results. A 'Quick data entry' dialog box is open, prompting the user to 'Enter one term per line and one relationship per line.' The dialog box includes a text input area, a 'Save' button, a 'Pick' button, a 'Cancel' button, and a 'See example' button. The background window shows a table with columns: Term, Status, Flag, Relationships, Notes, Categories, and Translation. The table contains one row with the term 'Espressos', status 'A', and translation 'ENG'.

MultiTes Demo: New Terms 2

20

This opens a wide open box, but no hints as to how the terms should be entered.

We click on the see example button...



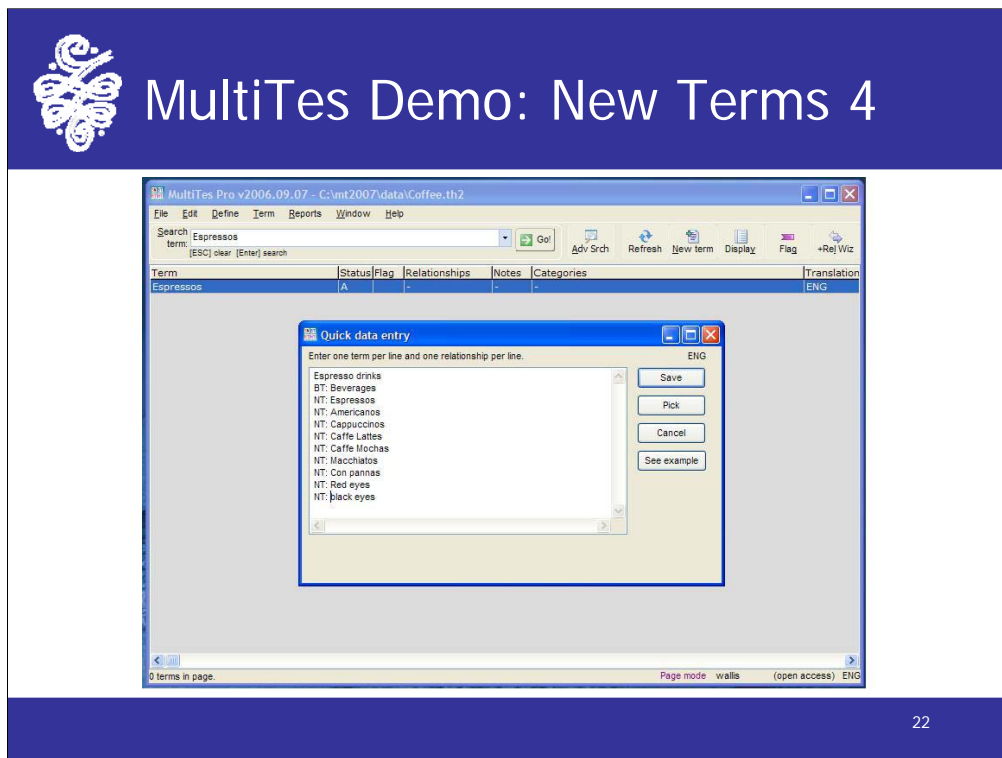
The screenshot displays the MultiTes Pro v2006.09.07 software interface. The main window shows a search for 'Espresso' with a table of results. A 'Quick data entry' dialog box is open, providing instructions on how to enter terms and relationships. The example text in the dialog box is as follows:

Enter one term per line, followed by its relationships.  
Leave a blank line before a new term.  
Example:  
Dogs  
NT: Hush Puppies  
NT: Chihuahua  
Cats  
NT: Siamese  
NT: Algerian

The dialog box also includes an 'OK' button and a list of relationship types (ENG, el, mple) on the right side.

21

And they give us an example of how a term and some term relationships can be added...



22

So then we add the term “Espresso drinks”, as well as the Broader and narrower terms associated with this term.

The program will automatically generate term records for those terms that have not yet been entered, so in this case everything but Espressos (which we created earlier) will be created when we hit “save”...

NOTE: Some potential drawbacks of this bulk entry method is that spelling counts, and you can easily find yourself creating multiple variety terms by accident using this auto-term creation functionality.

If anyone asks:

Cappuccino is espresso shots with equal parts steamed milk and foam

Americano is espresso and hot water

Latte is espresso shots with steamed milk and a touch of foam


Mocha is a latte with cocoa

Macchiato is espresso with a dab of foam

Con panna is espresso with a dab of whipped cream

Red eye is coffee with a shot of espresso (technically a coffee drink and an espresso drink)

Black eye is coffee with 2 shots of espresso (technically a coffee drink and an espresso drink)



## MultiTes Demo: New Terms 5

MultiTes Pro v2006.09.07 - C:\mnt2007\data\Coffee.th2

File Edit Define Term Reports Window Help

Search term: Espresso Go! Adv Srch Refresh New term Display Flag +Rej Wiz


Term	Status	Flag	Relationships	Notes	Categories	Translation
Americanos	A		BT(1)	-	-	ENG
Beverages	A		NT(1)	-	-	ENG
black eyes	A		BT(1)	-	-	ENG
Caffe Lattes	A		BT(1)	-	-	ENG
Caffe Mochas	A		BT(1)	-	-	ENG
Cappuccinos	A		BT(1)	-	-	ENG
Con pamas	A		BT(1)	-	-	ENG
Espresso drinks	A		BT(1), NT(9)	-	-	ENG
Espressos	A		BT(1)	-	-	ENG
Macchiatos	A		BT(1)	-	-	ENG
Red eyes	A		BT(1)	-	-	ENG

11 terms in page. Page mode walls (open access) ENG

23

Notice that the relationships column also now has entries.

Now we are going to check out the term record for one of our terms...



## MultiTes Demo: Term Record 1

MultiTes Pro v2006.09.07 - C:\vnt2007\data\Coffee.thz

File Edit Define Term Reports Window Help

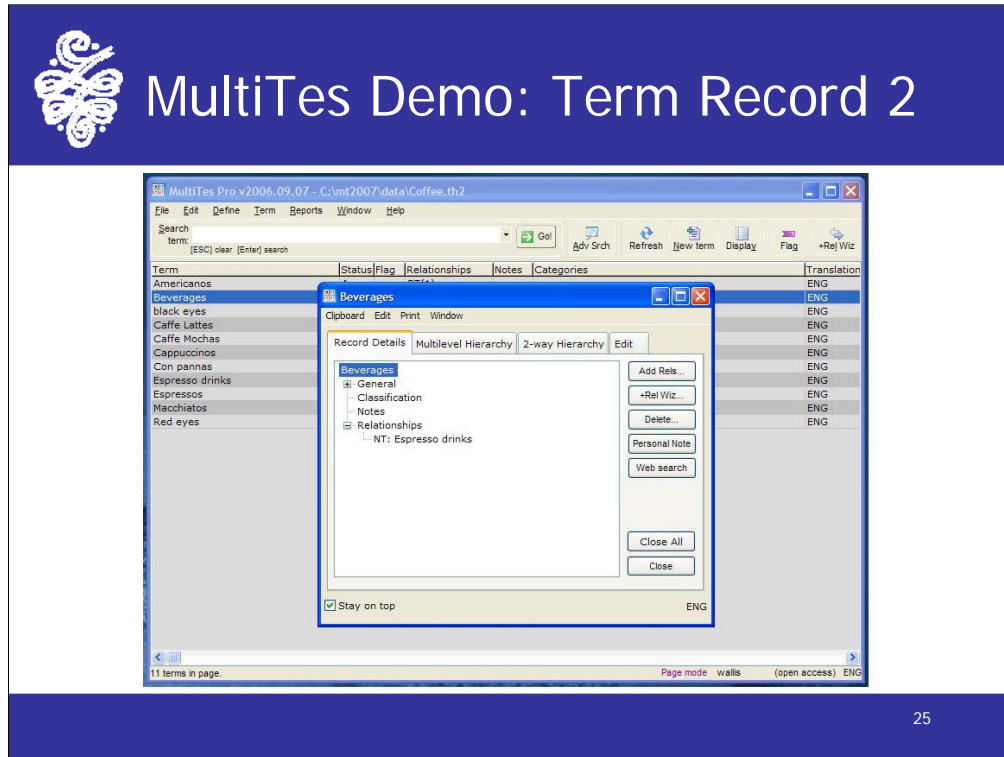
Search term: Espresso Go! Adv Srch Refresh New term Display Flag +Rej Wiz

Term	Status	Flag	Relationships	Notes	Categories	Translation
Americanos	A		BT(1)	-	-	ENG
Beverages	A		NT(1)	-	-	ENG
black eyes	A		BT(1)	-	-	ENG
Caffe Lattes	A		BT(1)	-	-	ENG
Caffe Mochas	A		BT(1)	-	-	ENG
Cappuccinos	A		BT(1)	-	-	ENG
Corr panias	A		BT(1)	-	-	ENG
Espresso drinks	A		BT(1), NT(s)	-	-	ENG
Espressos	A		BT(1)	-	-	ENG
Macchiatos	A		BT(1)	-	-	ENG
Red eyes	A		BT(1)	-	-	ENG

11 terms in page. Page mode walls (open access) ENG

We are going to double click on Beverages, or click on Beverages and go to the Display tool...





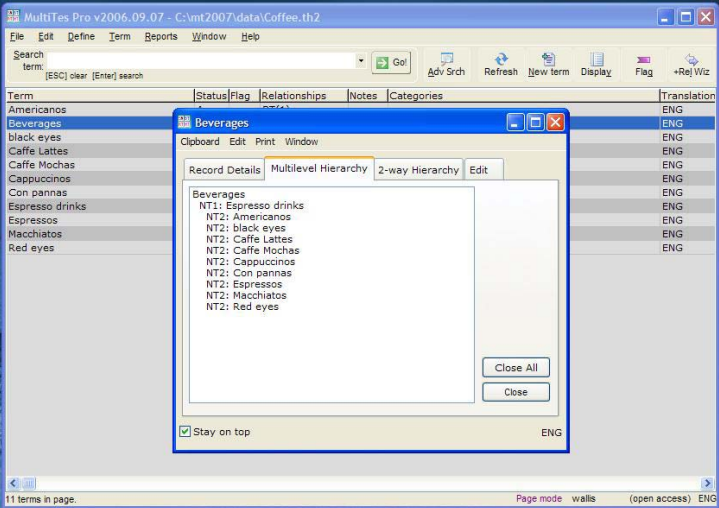
25

This brings up the term record in a new window.

Take a look at the features available -- add relationship, personal note, etc, as well as different views of the term -- hierarchy, etc.

NOTE: if anyone asks -- the checked box "Stay on top" is to make sure the little pop-up windows stay on top of the Main window.

NEXT: Clicking through the Multi-level hierarchy tab...



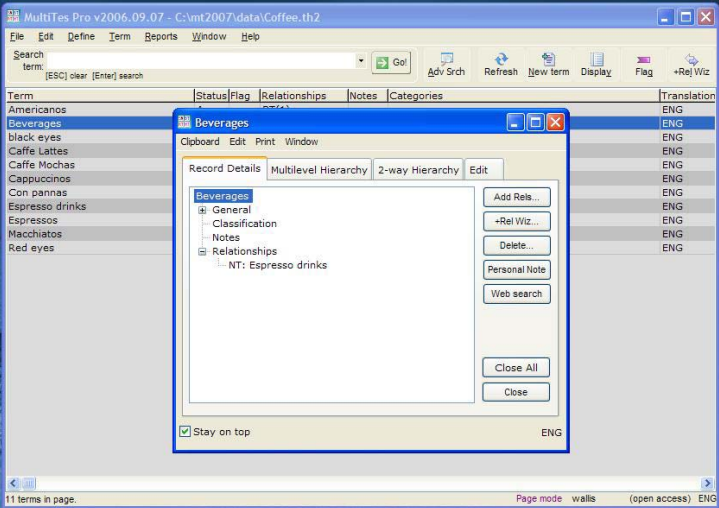
The screenshot displays the MultiTes Pro v2006.09.07 interface. The main window shows a list of terms under the 'Beverages' category, including Americanos, black eyes, Caffe Lattes, Caffe Mochas, Cappuccinos, Con paninas, Espresso drinks, Espresso, Macchiatos, and Red eyes. A 'Beverages' dialog box is open, showing a hierarchy of terms: Beverages, NT1: Espresso drinks, NT2: Americanos, NT2: black eyes, NT2: Caffe Lattes, NT2: Caffe Mochas, NT2: Cappuccinos, NT2: Con paninas, NT2: Espresso, NT2: Macchiatos, and NT2: Red eyes. The dialog box has tabs for 'Record Details', 'Multilevel Hierarchy', '2-way Hierarchy', and 'Edit'. The 'Multilevel Hierarchy' tab is selected. The dialog box also has 'Close All' and 'Close' buttons, and a 'Stay on top' checkbox.

26

This hierarchy shows not only the term below, but the terms below this narrower term.

Notice the NT1 NT2 convention.

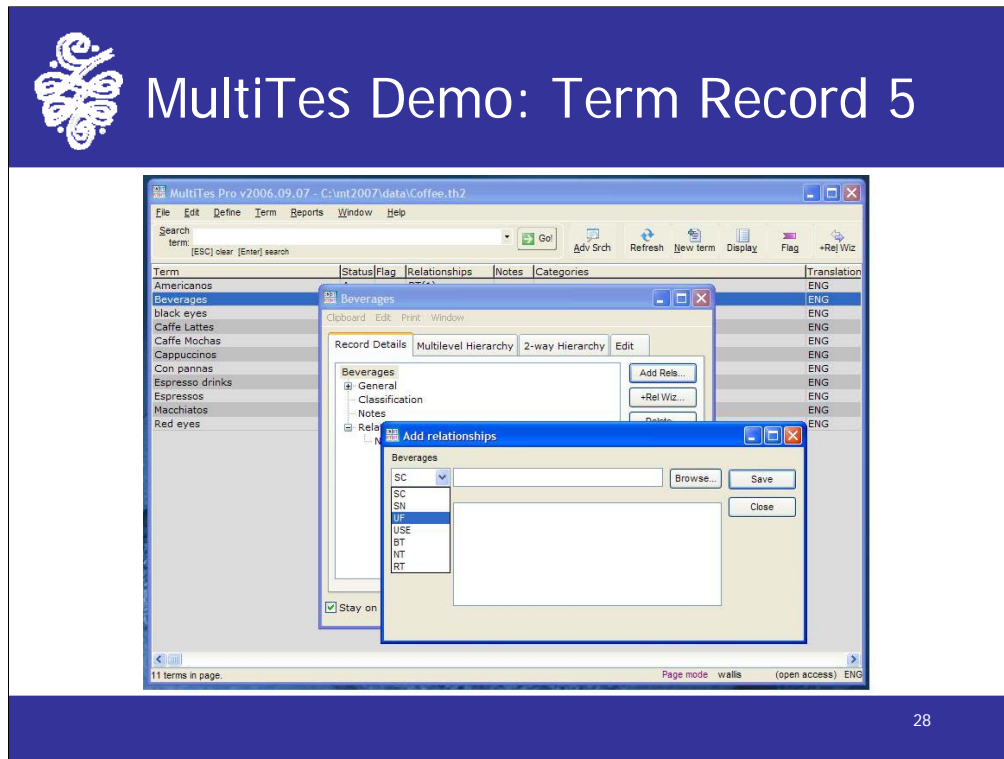
NEXT: Back to the record details...



The screenshot displays the MultiTes Pro v2006.09.07 interface. The main window shows a list of terms with columns for Term, Status, Flag, Relationships, Notes, Categories, and Translation. The term 'Beverages' is selected. A 'Record Details' dialog box is open over the 'Beverages' record, showing a tree view of the record's structure: 'Beverages' (expanded) contains 'General', 'Classification', 'Notes', and 'Relationships'. Under 'Relationships', there is a sub-entry 'NT: Espresso drinks'. The dialog box also features buttons for 'Add Rel...', '+Rel Wiz...', 'Delete...', 'Personal Note', 'Web search', 'Close All', and 'Close'. A 'Stay on top' checkbox is checked at the bottom left of the dialog. The status bar at the bottom of the main window indicates '11 terms in page' and 'Page mode walls (open access) ENG'.

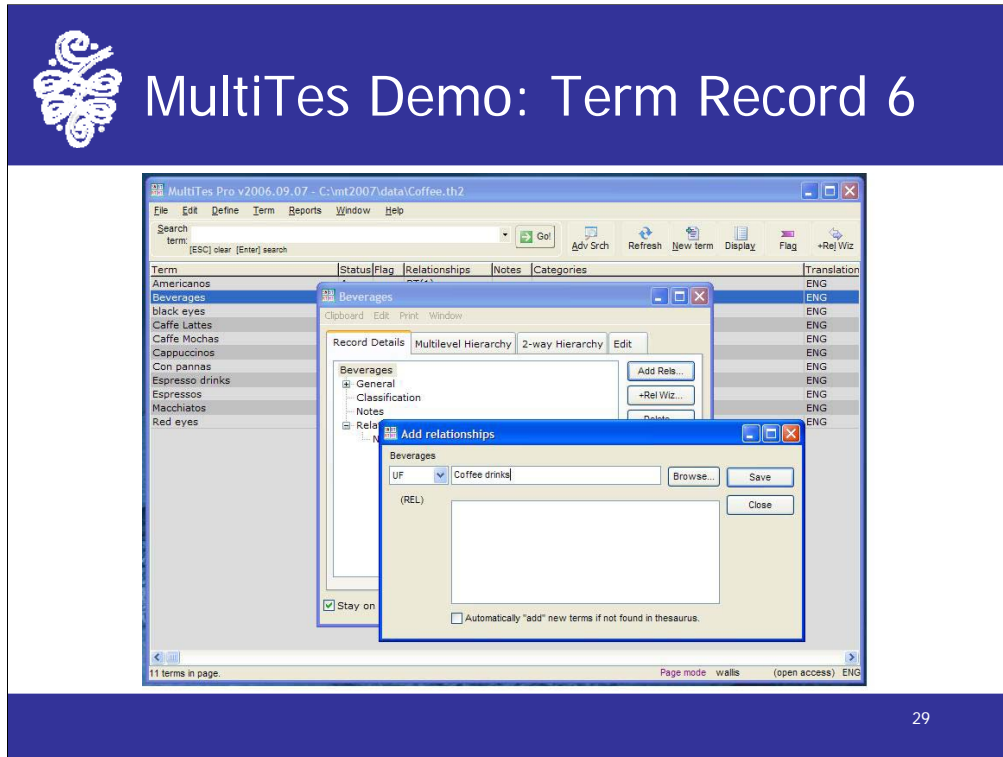
27

NEXT: Clicking the “add rels” button...



28

This opens a new window, which allows you to select the relation type, in this case the UF or Use for relationship...



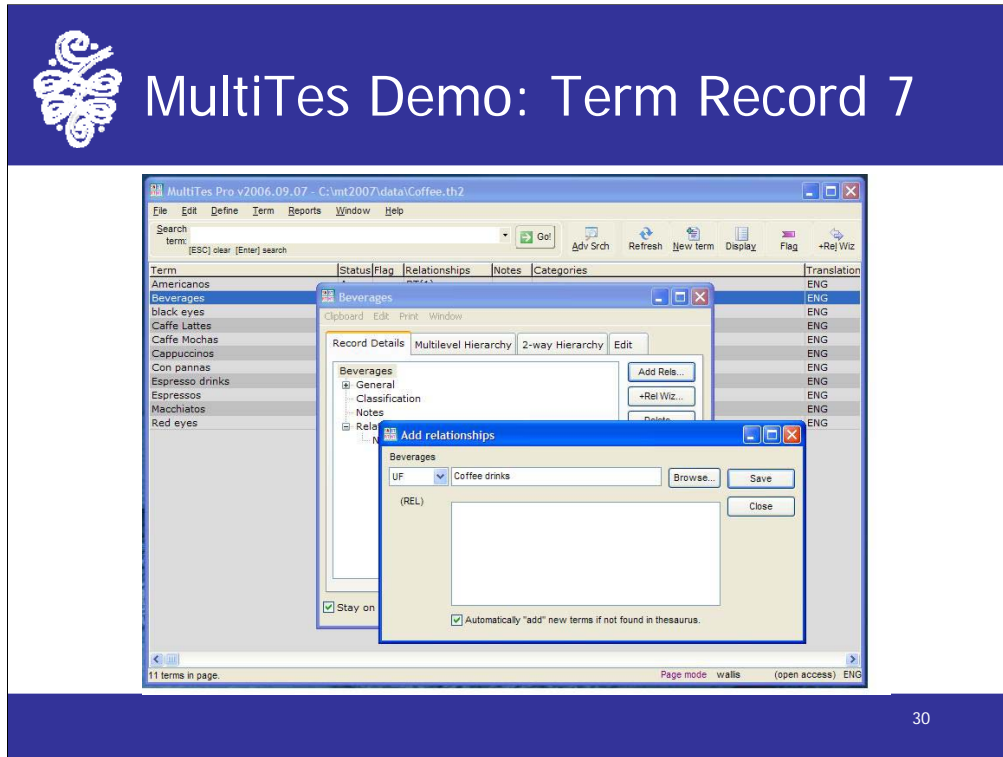
The screenshot displays the MultiTes Pro v2006.09.07 software interface. The main window shows a list of terms under the 'Beverages' category, including 'Americanos', 'black eyes', 'Caffe Lattes', 'Caffe Mochas', 'Cappuccinos', 'Con paninis', 'Espresso drinks', 'Espressos', 'Macchiatos', and 'Red eyes'. A 'Beverages' dialog box is open, showing 'Record Details' and 'Add relationships' options. The 'Add relationships' dialog box is further open, showing a search for 'Coffee drinks' and a 'Save' button. The status bar at the bottom indicates '11 terms in page' and 'Page mode walls (open access) ENG'.

29

And here we enter our term, which is “Coffee drinks”.

So to recap the main entry term is Beverages and you would use it for Coffee drinks.

We could also browse for a term if the synonym had already been entered...



The screenshot displays the MultiTes Pro v2006.09.07 software interface. The main window shows a list of terms on the left, including 'Americanos', 'Beverages', 'black eyes', 'Caffe Lattes', 'Caffe Mochas', 'Cappuccinos', 'Con panmas', 'Espresso drinks', 'Macchiatos', and 'Red eyes'. The 'Beverages' term is selected. A 'Record Details' dialog box is open, showing a tree view with 'General', 'Classification', 'Notes', and 'Relationships'. The 'Add relationships' dialog box is also open, showing a dropdown menu for 'UF' set to 'UF', a text field containing 'Coffee drinks', and a 'Browse...' button. Below the text field is a '(REL)' field and a 'Save' button. At the bottom of the dialog, there is a checkbox labeled 'Automatically "add" new terms if not found in thesaurus' which is checked. The status bar at the bottom of the main window indicates '11 terms in page' and 'Page mode walls (open access) ENG'.

30

Note that we are going to select the “automatically add new term if not found...” so that it will create the term record for Coffee drinks.

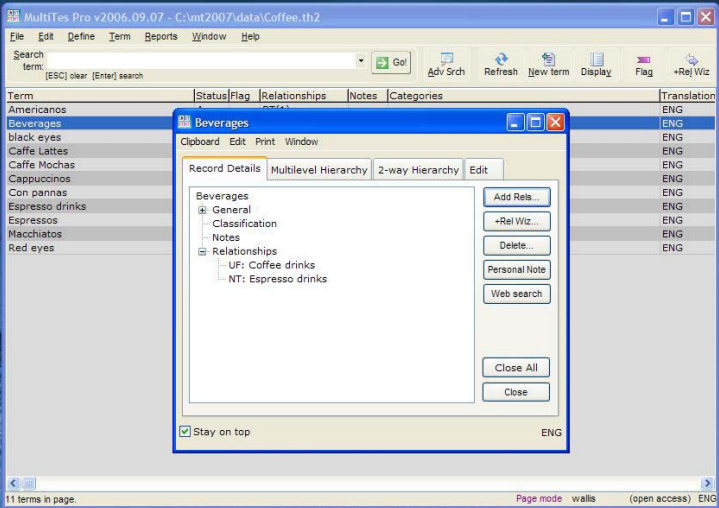
NEXT: we hit save button...

The screenshot shows the MultiTes Pro v2006.09.07 software interface. The main window displays a list of terms under the 'Beverages' category, including Americanos, black eyes, Caffe Lattes, Caffe Mochas, Cappuccinos, Con paninis, Espresso drinks, Espresso, Macchiatos, and Red eyes. A 'Beverages' dialog box is open, showing 'Record Details' and 'Add relationships' options. The 'Add relationships' dialog box is further open, showing a list of terms to be added, with 'Coffee drinks' selected. The 'Add relationships' dialog box has a 'Uf' dropdown menu, a 'Browse...' button, and a 'Save' button. The 'Add relationships' dialog box also has a 'Close' button. The 'Add relationships' dialog box has a checkbox for 'Automatically add new terms if not found in thesaurus' which is checked. The status bar at the bottom of the window shows '11 terms in page' and 'Page mode walls (open access) ENG'.

31

And now the new term has been added, and you can add another UF term if you have one.

NEXT: click through “close” as we are done with this...



The screenshot displays the MultiTes Pro v2006.09.07 interface. The main window shows a list of terms under the 'Beverages' category, including Americanos, black eyes, Caffe Lattes, Caffe Mochas, Cappuccinos, Con paninis, Espresso drinks, Espresso, Macchiatos, and Red eyes. A 'Beverages' dialog box is open, showing a relationship hierarchy under 'Relationships':

- UF: Coffee drinks
  - NT: Espresso drinks

The dialog box also includes buttons for 'Add Rel...', '+Rel Wiz...', 'Delete...', 'Personal Note', 'Web search', 'Close All', and 'Close'. The 'Stay on top' checkbox is checked. The status bar at the bottom indicates '11 terms in page' and 'Page mode walls (open access) ENG'.

32

Now we can see the new term and relationship reflected in the term record, and we hit close...



**MultiTes Demo: Term Record 10**


Term	Status	Flag	Relationships	Notes	Categories	Translation
Americanos	A		BT(1)	-	-	ENG
Beverages	A		[NT(1), UF(1)]	-	-	ENG
black eyes	A		BT(1)	-	-	ENG
Caffe Lattes	A		BT(1)	-	-	ENG
Caffe Mochas	A		BT(1)	-	-	ENG
Cappuccinos	A		BT(1)	-	-	ENG
Coffee drinks	A		[USE(1)]	-	-	ENG
Con pannas	A		BT(1)	-	-	ENG
Espresso drinks	A		BT(1), NT(9)	-	-	ENG
Espressos	A		BT(1)	-	-	ENG
Macchiatos	A		BT(1)	-	-	ENG
Red eyes	A		BT(1)	-	-	ENG

12 terms in page. Page mode: walls (open access) ENG

Now the UF term Coffee drinks has been added to the term list and the relationships on Beverages has been updated.

NOTE: If anyone asks, the numbers next to the relationships is the number of that type of relationship this term has. Espresso drinks had 9 narrow terms and 1 broader term, etc.

NEXT: We are going to set up relationships using the Relationships Wizard tool...



## MultiTes Demo: Relationships 1

MultiTes Pro v2006.09.07 - C:\mt2007\data\Coffee.th2

File Edit Define Term Reports Window Help

Search term: [Go]

[ESC] clear [Enter] search    Adv Srch    Refresh    New term    Display    Flag    +Rel Wiz

Term	Status	Flag	Relationships	Notes	Categories	Translation
Americanos	A		BT(1)	-	-	ENG
Beverages	A		NT(1), UF(1)	-	-	ENG
black eyes	A		BT(1)	-	-	ENG
Caffe Lattes	A		BT(1)	-	-	ENG
Caffe Mochas	A		BT(1)	-	-	ENG
Cappuccinos	A		BT(1)	-	-	ENG
Coffee drinks	A		USE(1)	-	-	ENG
Con pannas	A		BT(1)	-	-	ENG
Espresso drinks	A		BT(1), NT(9)	-	-	ENG
Espressos	A		BT(1)	-	-	ENG
Macchiatos	A		BT(1)	-	-	ENG
Red eyes	A		BT(1)	-	-	ENG

12 terms in page.    Page mode: walls    (open access)    ENG

34

We are going to start with the term Caffe Lattes, which so far only has one BT -- Espresso drinks.

NEXT: We are going to click on the +Rel Wiz...

The screenshot shows the MultiTes Pro v2006.09.07 interface. The main window displays a list of terms with columns for Term, Status, Flag, Relationships, Notes, Categories, and Translation. The 'Caffe Lattes' term is selected. A 'Relationship Wizard' dialog box is open, titled 'Caffe Lattes'. The dialog box has a tab labeled 'STEP 1: Select Relationship Type'. Below the tab is a list of relationship types: SC (Subject Category), SN (Scope Note), UF (Used For), USE (Use), BT (Broader Term), NT (Narrower Term), and RT (Related Term). The 'UF' option is currently selected. At the bottom of the dialog box are three buttons: 'Cancel/Close', '< Back', and 'Add UF >'. The status bar at the bottom of the main window indicates '12 terms in page', 'Page mode: walls', '(open access)', and 'ENG'.

35

This opens a new window, with Step 1 of adding a relationship. We will again select the Use for relationship and click on “Add UF”...

The screenshot displays the MultiTes Pro v2006.09.07 interface. The main window shows a list of terms with columns for Term, Status, Flag, Relationships, Notes, Categories, and Translation. The 'Caffe Lattes' term is selected. A 'Relationship Wizard' dialog box is open, titled 'Caffe Lattes'. The dialog has a tab labeled 'STEP 2: Enter/Select Related Term'. It contains a text input field with 'Flat whites' entered, an 'Adv. Select' button, and a 'Show top 10 matches' checkbox which is checked. Below the input field is a large empty box for displaying search results. To the right of the input field, there are radio buttons for 'Add term' (selected) and 'Prompt'. At the bottom of the dialog are 'Cancel/Close', '< Back', and 'Save' buttons. The status bar at the bottom of the main window indicates '12 terms in page.' and 'Page mode: walls (open access) ENG'.

36

This opens step 2 which is to input the term, or select the term. Here we are adding “Flat whites” as another term for a caffe latte.

NEXT: we will save this...

The screenshot shows the MultiTes Pro v2006.09.07 interface. The main window displays a list of terms with columns for Term, Status, Flag, Relationships, Notes, Categories, and Translation. A 'Relationship Wizard' dialog box is open, titled 'Caffe Lattes'. The wizard is at 'STEP 1: Select Relationship Type'. A list of relationship types is shown on the left, with 'UF Used For' selected. The right side of the dialog shows a list of terms, with 'Flat whites' selected. The dialog has buttons for 'Cancel/Close', '< Back', and 'Add UF >'. The status bar at the bottom indicates '12 terms in page', 'Page mode: walls', '(open access)', and 'ENG'.


37

And it brings us back to step 1.

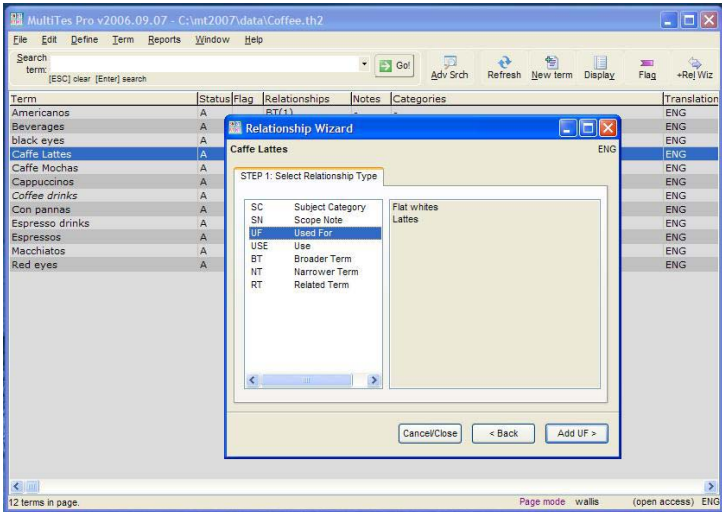
We want to add another UF, so we will again go through the Add UF button...

The screenshot displays the MultiTes Pro v2006.09.07 software interface. The main window shows a list of terms with columns for Term, Status, Flag, Relationships, Notes, Categories, and Translation. The 'Caffe Lattes' term is selected. A 'Relationship Wizard' dialog box is open, titled 'Caffe Lattes'. The dialog box has a tab labeled 'STEP 2: Enter/Select Related Term'. It contains a text field labeled 'UF - Used For' with 'Lattes' entered. Below this is a checkbox labeled 'Show top 10 matches' which is checked. To the right of the text field is an 'Adv. Select' button. Below the text field is a section labeled 'If term not found...' with two radio buttons: 'Add term' (selected) and 'Prompt'. At the bottom of the dialog box are three buttons: 'Cancel/Close', '< Back', and 'Save'. The status bar at the bottom of the main window indicates '12 terms in page.' and 'Page mode walls (open access) ENG'. The number '38' is displayed in the bottom right corner of the slide.

Now we will add the term Lattes, as another word for Caffe Lattes, and again save...



## MultiTes Demo: Relationships 6




The screenshot shows the MultiTes Pro v2006.09.07 interface. A 'Relationship Wizard' dialog box is open for the term 'Caffe Lattes'. The wizard is at 'STEP 1: Select Relationship Type'. The 'UF' (Used For) relationship type is selected in the list on the left. The right pane shows 'Flat whites' and 'Lattes' as related terms. The 'Cancel/Close' button is highlighted.

39

Now that we have added all our relationships for this term we will hit the Cancel/Close button.

Is this last step confusing? YES!

You need to think about it as cancelling the next step 1/step 2 sequence or closing the wizard, but yes, you are not the only ones confused by this...



## MultiTes Demo: Relationships 7

MultiTes Pro v2006.09.07 - C:\mt2007\data\Coffee.tht2

File Edit Define Term Reports Window Help

Search terms: [ESC] clear [Enter] search [Go]

Adv Srch Refresh New term Display Flag +Rel Wiz

Term	Status	Flag	Relationships	Notes	Categories	Translation
Americanos	A		BT(1)	-	-	ENG
Beverages	A		NT(1), UF(1)	-	-	ENG
black eyes	A		BT(1)	-	-	ENG
Caffe Lattes	A		BT(1), UF(2)	-	-	ENG
Caffe Mochas	A		BT(1)	-	-	ENG
Cappuccinos	A		BT(1)	-	-	ENG
Coffee drinks	A		USE(1)	-	-	ENG
Con pannas	A		BT(1)	-	-	ENG
Espresso drinks	A		BT(1), NT(9)	-	-	ENG
Espressos	A		BT(1)	-	-	ENG
Flat whites	A		USE(1)	-	-	ENG
Lattes	A		USE(1)	-	-	ENG
Macchiatos	A		BT(1)	-	-	ENG
Red eyes	A		BT(1)	-	-	ENG

14 terms in page. Page mode: walls (open access) ENG

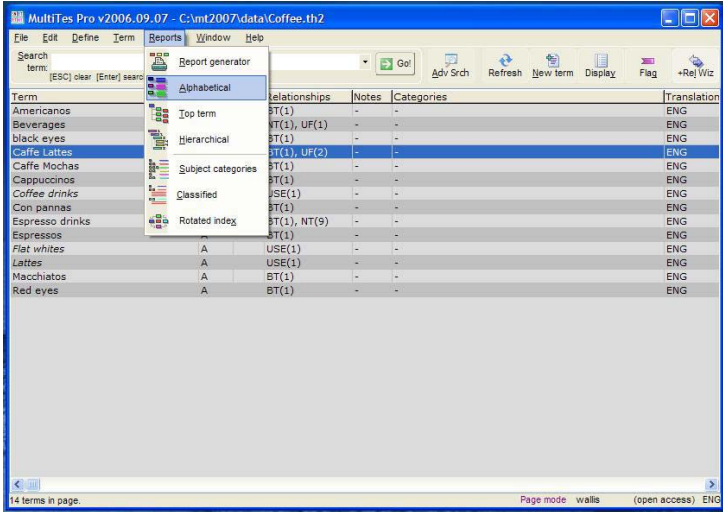
40

We can see that our term list has been updated with the new use for terms and relationships.

Looking at this, can anyone see the shape of the thesaurus? Kind of like an alphabetical output, difficult to see the hierarchy from just the number of NTs or BTs, so let us see how it looks when output...

NEXT: We are going to the Reports menu...



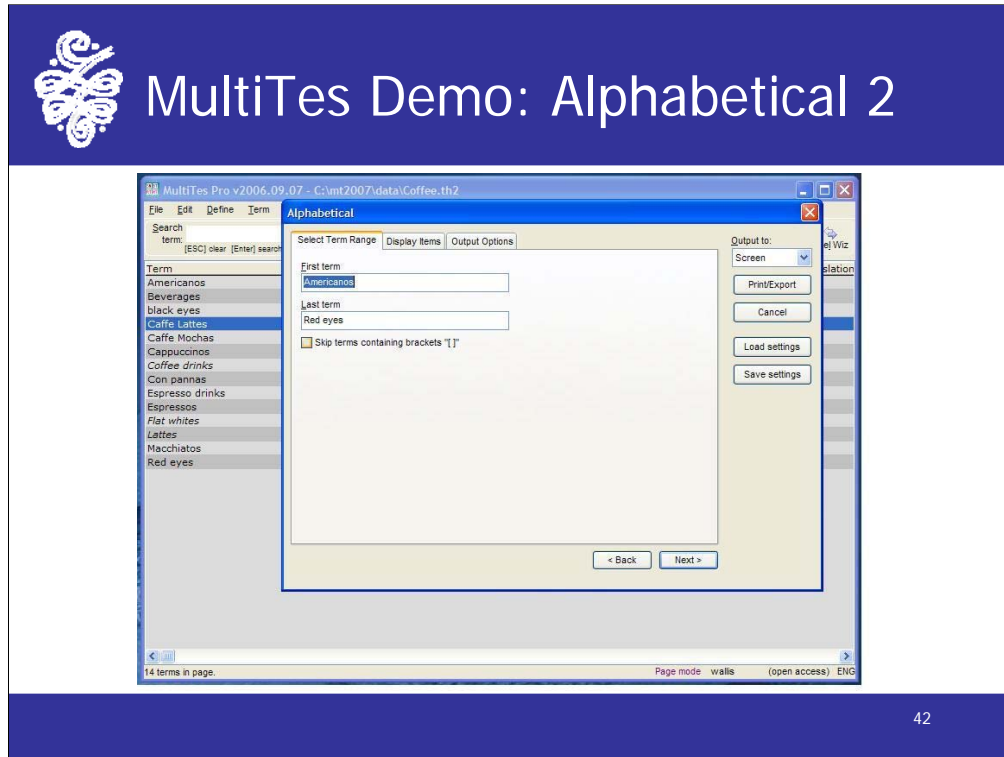


The screenshot shows the MultiTes Pro v2006.09.07 interface. The 'Reports' menu is open, and the 'Alphabetical' option is selected. The main window displays a table of terms and their relationships.

Term	relationships	Notes	Categories	Translation
Americanos	BT(1)	-	-	ENG
Beverages	BT(1), UF(1)	-	-	ENG
black eyes	BT(1)	-	-	ENG
Caffe Lattes	BT(1), UF(2)	-	-	ENG
Caffe Mochas	BT(1)	-	-	ENG
Cappuccinos	BT(1)	-	-	ENG
Coffee drinks	ISE(1)	-	-	ENG
Con pannas	BT(1)	-	-	ENG
Espresso drinks	BT(1), NT(9)	-	-	ENG
Espressos	BT(1)	-	-	ENG
Flat whites	A	USE(1)	-	ENG
Lattes	A	USE(1)	-	ENG
Macchiatos	A	BT(1)	-	ENG
Red eyes	A	BT(1)	-	ENG


41

And select the Alphabetical reporting option...



42

This brings up a new window asking the range of terms to be reported, this range defaults to everything, so we just hit Print/Export, and it will go to a screen export...



## MultiTes Demo: Alphabetical 3

**-27.-ts - Notepad**  
File Edit Format View Help

Americanos  
BT: Espresso drinks

Beverages  
UF: Coffee drinks  
NT: Espresso drinks

black eyes  
BT: Espresso drinks

Caffe Lattes  
UF: Flat whites  
Lattes  
BT: Espresso drinks

Caffe Mochas  
BT: Espresso drinks

Cappuccinos  
BT: Espresso drinks

Coffee drinks  
USE: Beverages

Con pannas  
BT: Espresso drinks

Espresso drinks  
BT: Beverages  
NT: Americanos  
black eyes  
Caffe Lattes  
Caffe Mochas  
Cappuccinos  
Con pannas  
Espressos  
Macchiatos  
Red eyes

Espressos  
BT: Espresso drinks

Flat whites  
USE: Caffe Lattes

Lattes  
USE: Caffe Lattes

Macchiatos  
BT: Espresso drinks

Red eyes  
BT: Espresso drinks

43

In a little Notepad file. And we can see the alphabetical list of terms and their brief records.

NEXT: On to the Hierarchical display...

**MultiTes Demo: Hierarchical 1**

The screenshot shows the MultiTes Pro v2006.09.07 interface. The 'Reports' menu is open, and 'Hierarchical' is selected. The main window displays a table of terms and their relationships.

Term	relationships	Notes	Categories	Translation
Americanos	BT(1)	-	-	ENG
Beverages	BT(1), UF(1)	-	-	ENG
black eyes	BT(1)	-	-	ENG
Caffe Lattes	BT(1), UF(2)	-	-	ENG
Caffe Mochas	BT(1)	-	-	ENG
Cappuccinos	BT(1)	-	-	ENG
Coffee drinks	USE(1)	-	-	ENG
Con pannas	BT(1)	-	-	ENG
Espresso drinks	BT(1), NT(9)	-	-	ENG
Espressos	BT(1)	-	-	ENG
Flat whites	A	USE(1)	-	ENG
Lattes	A	USE(1)	-	ENG
Macchiatos	A	BT(1)	-	ENG
Red eyes	A	BT(1)	-	ENG

14 terms in page. Page mode: walls (open access) ENVC

Again we are in the Reports menu and select Hierarchical this time...

MultiTes Demo: Hierarchical 2

MultiTes Pro v2006.09.07 - C:\vnt2007\data\Coffee.thz

File Edit Define Term

Search  
term: [ESC] clear [Enter] search

Term  
Americano  
Beverages  
black eyes  
Caffe Lattes  
Caffe Mochas  
Cappuccinos  
Coffee drinks  
Con pannas  
Espresso drinks  
Espresso  
Flat whites  
Lattes  
Macchiatos  
Red eyes

Hierarchical

Select Term Range Output Options

First term  
Americano

Last term  
Red eyes

Skip terms containing brackets [ ]

Output to:  
Screen

Print/Export  
Cancel  
Load settings  
Save settings

< Back Next >

14 terms in page. Page mode walls (open access) ENG

45

Again it gives us a range and we hit the Print/Export button...

**MultiTes Demo: Hierarchical 3**

```

~2B.~ts - Notepad
File Edit Format View Help
Americanos
  BT1: Espresso drinks
  BT2: Beverages
Beverages
  UF: Coffee drinks
  NT1: Espresso drinks
  NT2: Americanos
  NT2: black eyes
  NT2: Caffe Lattes
  NT2: Caffe Mochas
  NT2: Cappuccinos
  NT2: Con pannas
  NT2: Espessos
  NT2: Macchiatos
  NT2: Red eyes
black eyes
  BT1: Espresso drinks
  BT2: Beverages
Caffe Lattes
  UF: Flat whites
  Lattes
  BT1: Espresso drinks
  BT2: Beverages

Caffe Mochas
  BT1: Espresso drinks
  BT2: Beverages
Cappuccinos
  BT1: Espresso drinks
  BT2: Beverages
Coffee drinks
  USE: Beverages
Con pannas
  BT1: Espresso drinks
  BT2: Beverages
Espresso drinks
  NT1: Americanos
  NT1: black eyes
  NT1: Caffe Lattes
  NT1: Caffe Mochas
  NT1: Cappuccinos
  NT1: Con pannas
  NT1: Espessos
  NT1: Macchiatos
  NT1: Red eyes
  BT1: Beverages
Espessos
  BT1: Espresso drinks
  
```

46

Here we have the notepad output in hierarchical form.

Note that it is even a multi-level hierarchical display.

NEXT: on to Rotated index...

**MultiTes Demo: Rotated 1**

The screenshot shows the MultiTes Pro v2006.09.07 interface. The 'Reports' menu is open, and 'Rotated index' is selected. The main window displays a table with the following data:

Term	Relationships	Notes	Categories	Translation
Americanos	BT(1)	-	-	ENG
Beverages	BT(1), UF(1)	-	-	ENG
black eyes	BT(1)	-	-	ENG
Caffe Lattes	BT(1), UF(2)	-	-	ENG
Caffe Mochas	BT(1)	-	-	ENG
Cappuccinos	BT(1)	-	-	ENG
Coffee drinks	USE(1)	-	-	ENG
Coh pannaas	BT(1)	-	-	ENG
Espresso drinks	BT(1), NT(9)	-	-	ENG
Espressos	BT(1)	-	-	ENG
Flat whites	A	USE(1)	-	ENG
Lattes	A	USE(1)	-	ENG
Macchiatos	A	BT(1)	-	ENG
Red eyes	A	BT(1)	-	ENG

47

Again we are in the Reports menu and this time we select the Rotated index...

MultiTes Demo: Rotated 2

MultiTes Pro v2006.09.07 - C:\mt2007\data\Coffee.th2

Rotated index

Select Term Range | KWOC Options | Output Options

First word: A

Last word: ZZZ

Skip terms containing brackets [ ]

Output to: Screen

Print/Export

Cancel

Load settings

Save settings

< Back Next >

14 terms in page. Page mode walls (open access) ENG

48

It gives us a range and we go ahead with it...



**MultiTes Demo: Rotated 3**

~2D.~ts - Notepad  
File Edit Format View Help

Americanos  
Americanos  
Beverages  
Beverages  
black  
black eyes  
Caffe  
Caffe Lattes  
Caffe Mochas  
Cappuccinos  
Cappuccinos  
Coffee  
Coffee drinks  
Con  
Con pannas  
drinks  
Coffee drinks  
Espresso drinks

Espresso  
Espresso drinks  
Espresso drinks  
Espresso drinks  
black eyes  
Red eyes  
Flat  
Flat whites  
Lattes  
Caffe Lattes  
Lattes  
Macchiatos  
Macchiatos  
Mochas  
Caffe Mochas  
pannas  
Con pannas  
Red  
Red eyes  
whites  
Flat whites

49

The rotated index lists every word found in the term list.



## MultiTees Demo: Discussion

- Would this product make the conceptual organization of terms an easy task?
- Who is this software made for?
- What could make this a better tool?
- What do you need from your thesaurus software?

50

My answer to the first question is NO! This software is best for someone who has already mapped out their terms and the organization in some other program!

Making it better... Drag and drop functionality would make it easier to use it for term organization. Being able to switch the display to a hierarchy from the alphabetical list would also be handy. Anything else people can come up with...

Finding the right tool for themselves is something they need to think about and compare options...



## More Cautionary Notes

- The true work of controlled vocabulary design is the collection and intellectual organization of terms!
- Developing a controlled vocabulary is like taking a still photo, while reality is a movie!
- Indexer consistency studies
- Controlled vocabulary vs. free-text search studies
- Folk classification



11.4, <sup>51</sup>p. 99

Some indexing consistency studies:

- Chan, L. M. (1989). Inter-indexer consistency in subject cataloging. *Information Technology and Libraries*, 8(4):349-58.
- Cooper, W.S. (1969). Is interindexer consistency a hobgoblin? *American Documentation*, 20:268-278.
- Leonard, L. E. (1977). Inter-indexer consistency studies, 1954-1975: A review of the literature and summary of study results. Techreport, University of Illinois, Graduate School of Library Science, Champaign, IL.
- Sievert, M. C. & J. J. Andrews (1991). Indexing consistency in information science abstracts. *Journal of the American Society for Information Science*, 42(1):1-6.

Even professional indexers are subject to the vocabulary problem. Except in rare instances and limited collections, free text searching has been found to be about as good as controlled vocabulary...and is faster, and requires less effort from searchers.

Folk classification – looking at how people classify and organize information in their daily lives, e.g., recipe cards, digital music playlists, websites and bookmarks (del.icio.us). Much different than AACR2! Answerbag is a folksonomy (<http://www.answerbag.com>) where users submit new categories for questions...



## The Future of Controlled Vocab

- Integrated ILS and thesauri
- The semantic web
  - Goals
    - Make semantic relationships machine-readable
    - Distributed database platform
  - Combined sets of semantic relationships = Semantic Web
  - Some XML based technologies
    - RDF - Resource Description Framework
    - OWL - Web Ontology Language
    - RSS - Really Simple Syndication
  - Ex. MARCXML

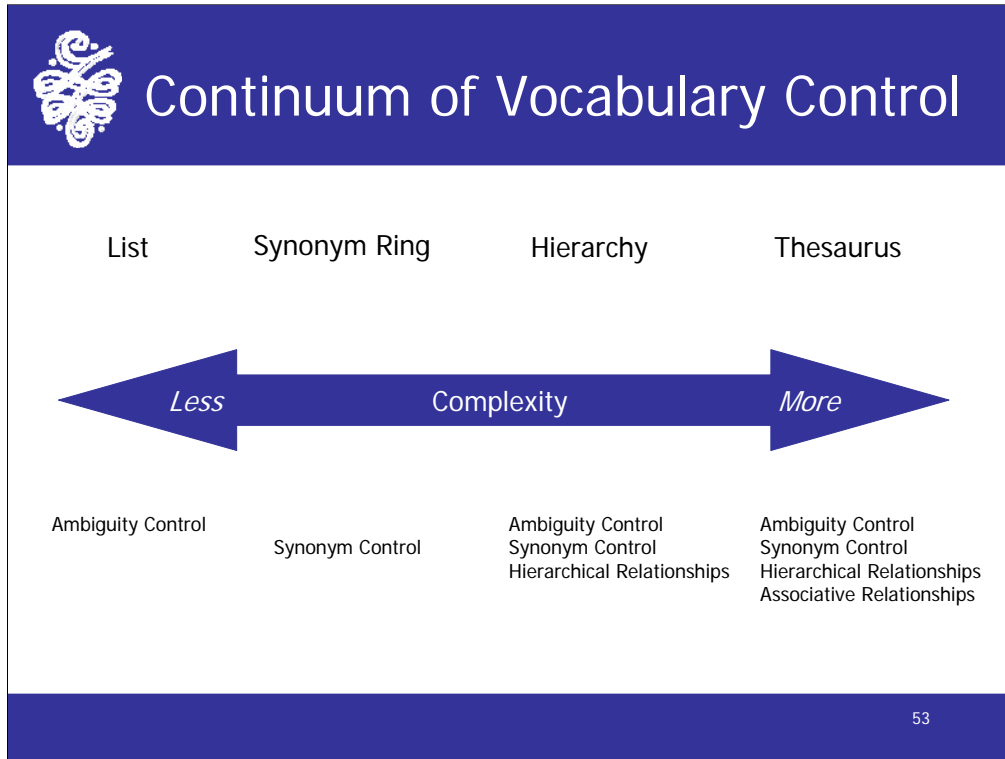
52

Items are labeled in a given ontology, but do not need to be stored with the ontology, just as long as they call the namespace and be explicit about what parts of the ontology are used.

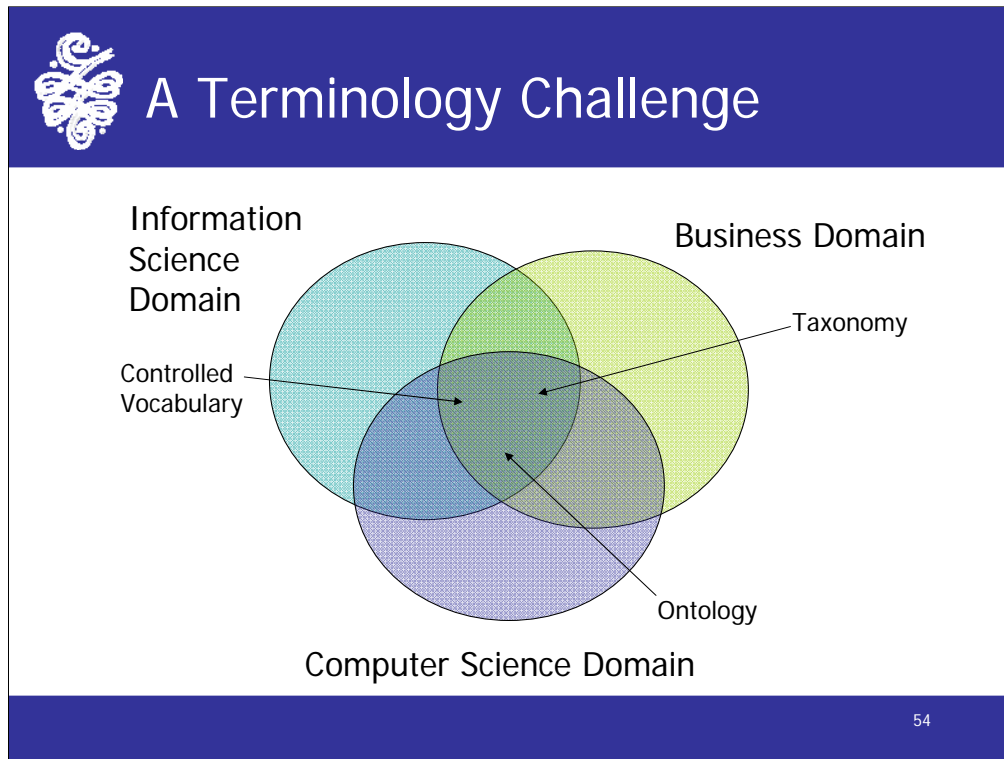
Not really something that can be covered in an hour or even a day. I took an entire course on it and I do not understand all of the various pieces of the Semantic Web.

You need to be aware of them when talking to your IT people who are more aware of this terminology. Perhaps it is better to refer back to the continuum and use the

*Note from 9/06: Try not to do a slow demo (i.e., a demo using a slow connection, or a demo requiring a lot of navigation) at the end of the second day (students are tired by this point).*




We apologize for the use of this diagram which we pulled from the standard. Use these functionalities instead...



Using the different terms for the same thing.

At this point each of these domains are using a different term to describe the same thing.



## Changing Definitions

- When the standard was published
  - Ontology ≈ Taxonomy = Hierarchy
- Now...
  - Taxonomy = Controlled Vocabulary = Ontology
- Expect it to continue to change!

55

This confusion is really our fault, we were using the illustrations from the standard and since it was published in 2005, the term definition has changed. We should have instead used the term “hierarchy”.

Talk about keeping up with the terms being used by your users today!



## Taxonomies are Everywhere

- Especially in product websites
- Examples include
  - Yahoo!, Amazon, HomeDepot, etc.
- To see them you need to just look around

56

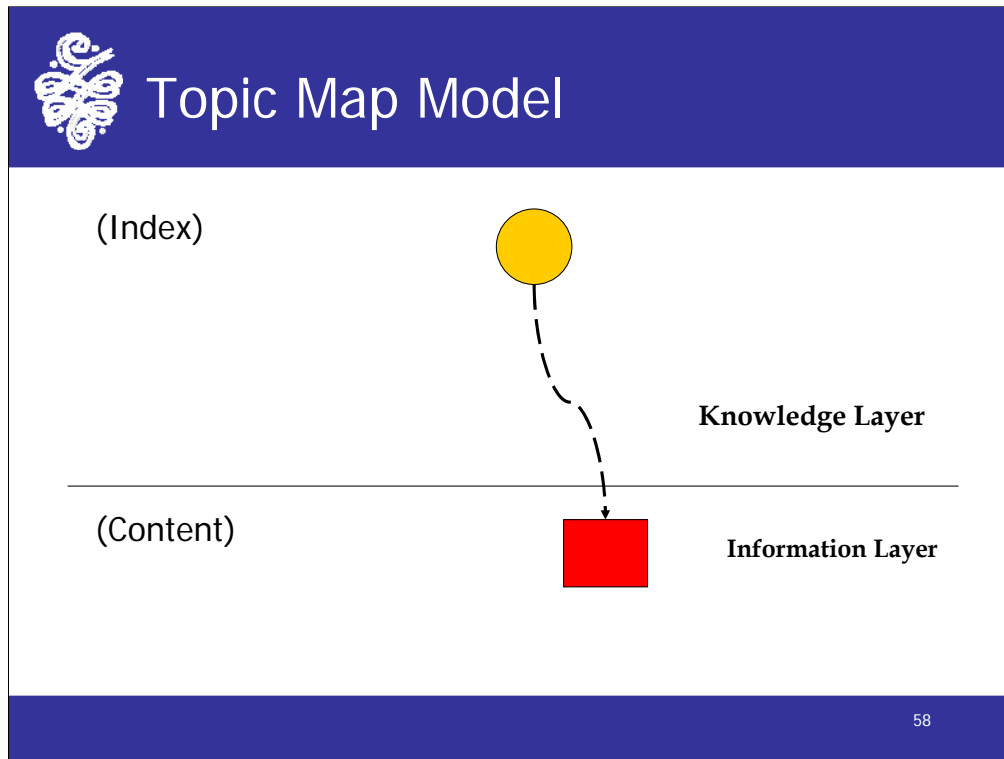
These are taxonomies because they are all in the business idiom.





## Ontologies

- Part of the Semantic Web suite of technologies
- Ontologies are:
  - Published in a Namespace (like a URL)
  - Consist of Objects, Associations, and Instances
- Completely analogous to Controlled Vocabularies
  - Terms, Relationships, Application of the term to some Thing

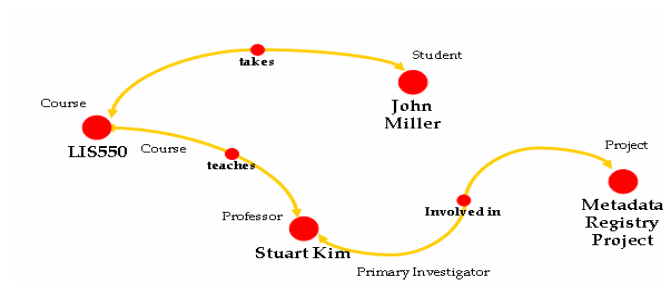


A topic map is a type of semantic web construction. With 2 layers, the information layer and the knowledge layer. The objects found in each and the relationships between them need to be defined.



## The Knowledge Layer

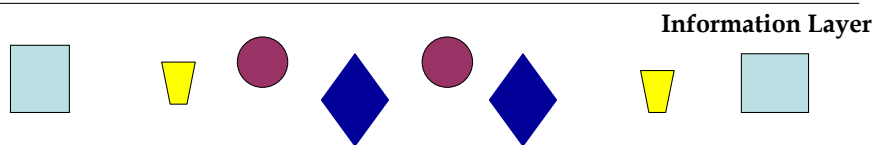
- The upper layer consists of topics and associations
  - Topics represent the subjects that the information is about
  - Associations represent relationships between those subjects

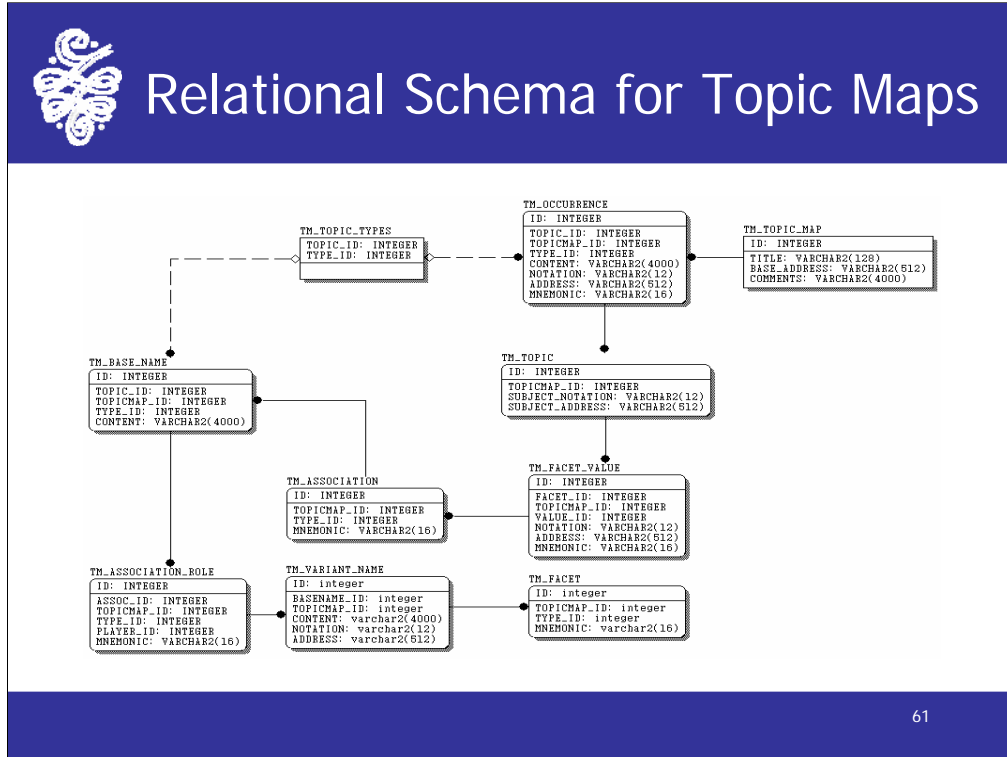


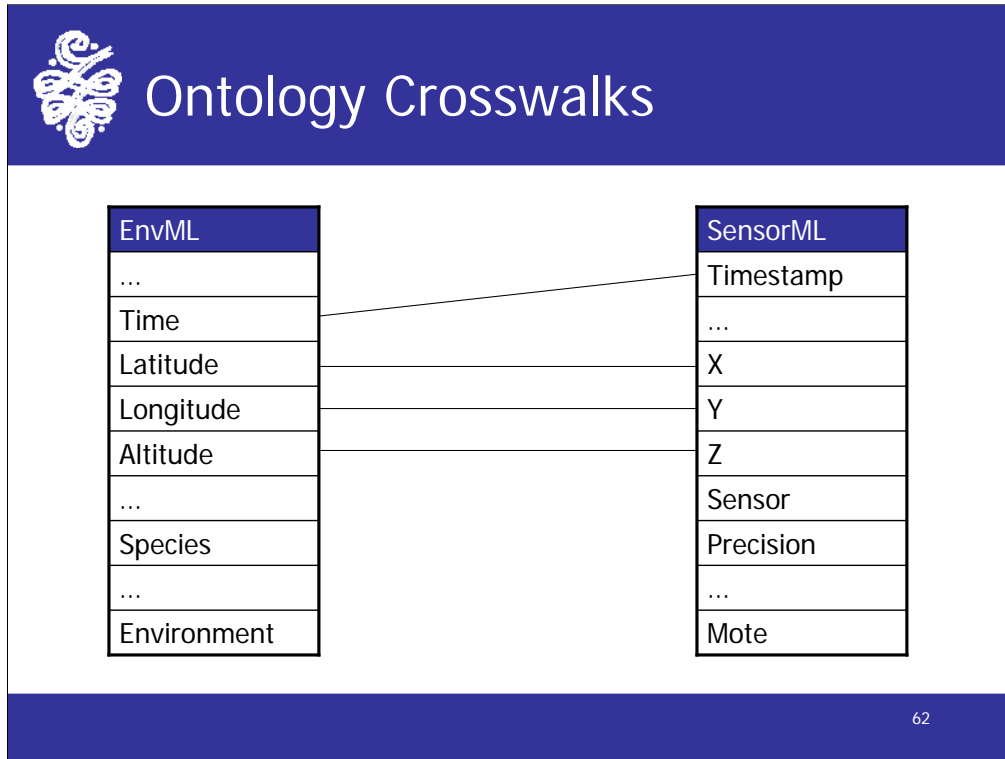


## The Information Layer

- The lower layer contains the content
  - Usually digital, but need not be
  - Can be in any format or notation
  - Can be text, graphics, video, audio, etc.



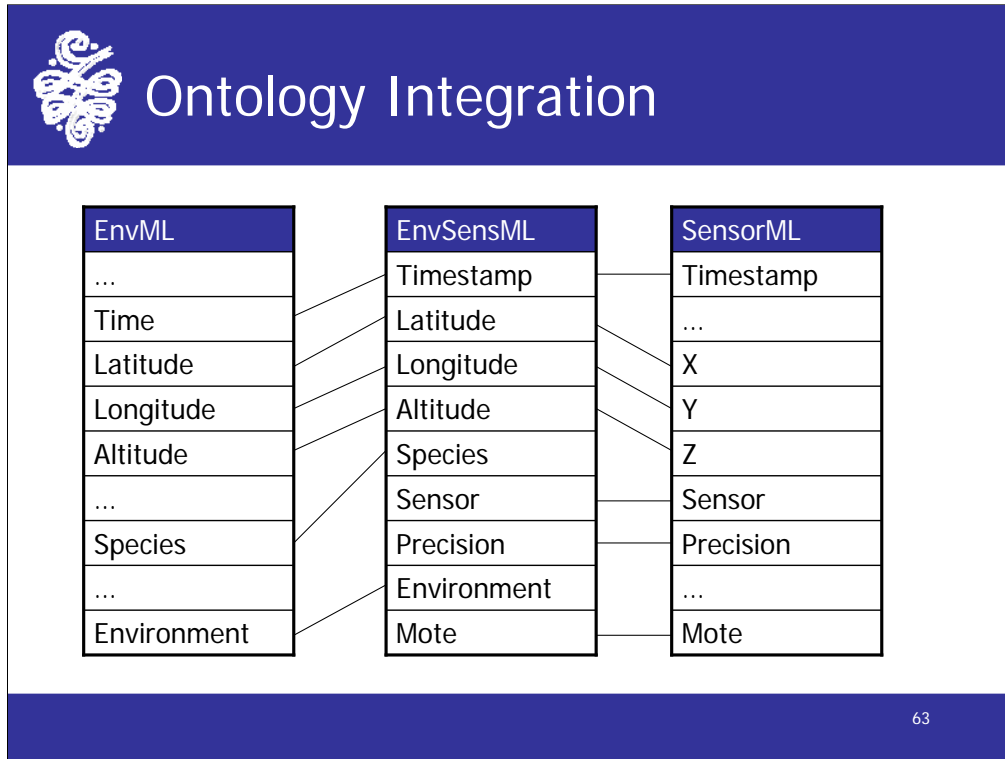




Saying one field is equivalent to the other in a document that serves as a map between the two.

What does this look like?

Synonym Ring

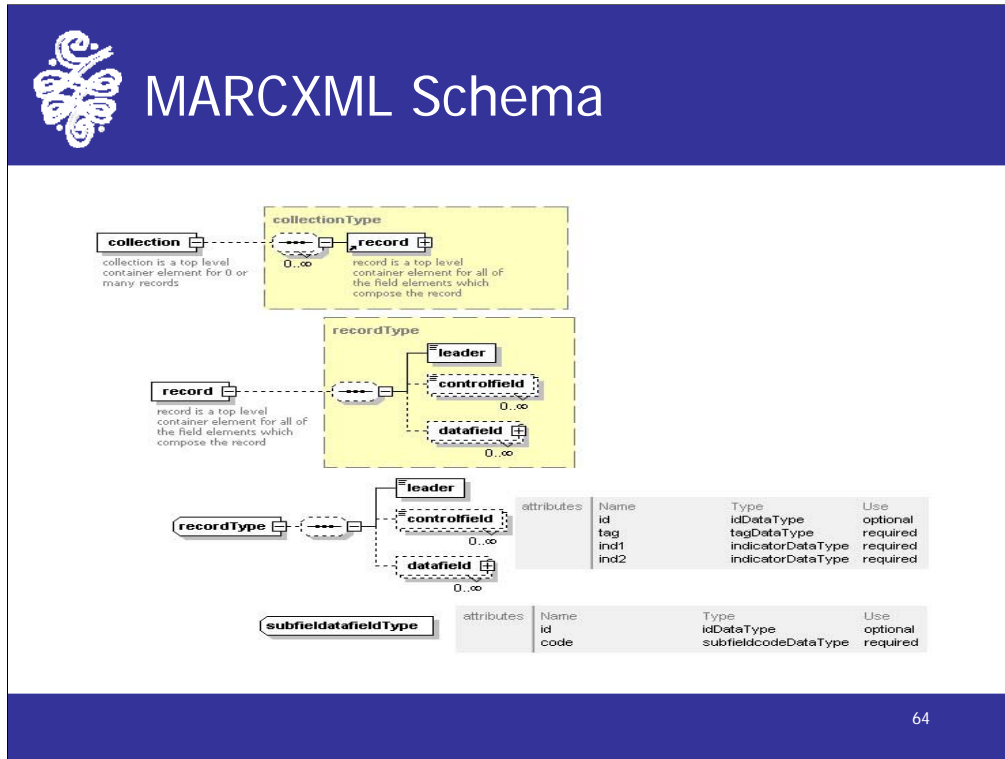


What does this look like?

Pulling together you applicable thesauri to create some new thesaurus...

Are we noticing a pattern here?

These are all analogous to CV things, with different terminology.







## Resources Available

- List of thesauri
  - <http://www.lub.lu.se/metadata/subject-help.html>
- Thesaurus construction guide
  - <http://www.willpower.demon.co.uk/>
- Course materials - for updated slidesets
  - <http://www.moebiusstrip.org/CV/>
- More on the Semantic Web
  - Spinning the Semantic Web, by Fensel et al  
<http://w5.cs.uni-sb.de/teaching/ss03/SemanticWebHTML/Vorlesung%20SemanticWebSS03/Introduction.pdf>



## Course Goals

- Understand and apply fundamental concepts of controlled vocabulary and thesaurus design, and why they are important
- Understand and apply diverse types of term relationships to structure descriptive terms
- Understand and apply both basic rules and best practices from existing thesauri to the construction and maintenance of thesauri and controlled vocabularies
- Develop a basis for exercising individual judgment for making thesaurus and controlled vocabulary decisions



## Wrapping up

- Any last questions?
- Course evaluations

*THIS PAGE INTENTIONALLY LEFT BLANK  
FOR DOUBLE SIDED COPY*

### Exercise 1 Handout



A 1921 view of the Citrus Exchange Building (611-622 Franklin Street)/Maas Brothers department store (612-622 Franklin Street) at the corner of Franklin and Zack Streets in Tampa, Florida. Maas Brothers department store occupies the lower floors of what is primarily an office building in the downtown area. Cars are parked in varying configurations on the two streets and a bicycle and three pedestrians are in view. The top of a streetlight is in the foreground. A black-and-white photograph in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-00006557



The 1921 home of Max C. Gill at 710 South Delaware Street in Tampa, Florida. The house is a brick bungalow in the craftsman style with a porch, tile roof, porte cochere, and detached garage. A lamppost stands at the driveway entrance. A black-and-white photograph in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-00005715



The Stovall Office Building at 416 Tampa Street, Tampa, Florida, 1950. Retail business occupy the ground floor. Automobiles and pedestrians are in view. A black-and-white photograph in the Burgert Brothers Photograph Collection. Unique Identifier: B29-00053051.



Pedestrians and motor traffic move through the intersection of Franklin and Zack Streets in downtown Tampa, Florida, 1929. A three story office and retail business building is featured and a traffic light is visible. Businesses in the view are Don Thompson Radio Fixtures and Appliances, Home City News Stand, Peeler's Men Shop, and William's Corner. A black-and-white photograph in the Burgert Brothers Photograph Collection. Unique Identifier: B29-rel-00000242



The Economical Drug Store on the ground floor of the Stovall-Nelson Building at 817 Franklin Street in Tampa, Florida. The view shows the main entrance from a busy sidewalk busy with pedestrian traffic. One sign advertises Coca-Cola and another proclaims “This is not a chain store”. 1929. A black-and-white photograph in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-00029362



Three women and a small dog look across Tampa Bay from the Venetian Apartments courtyard in Tampa, Florida. The courtyard is enclosed by an open work cement wall with a decorative wrought iron gate leading to a boat slip. Ornate lamp posts top the columns on either side of the gate. A black-and-white photograph in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-00022457



The Home of Val Antuono. (4024 Bay to Bay Boulevard cor Clark Street) in Tampa. Note the Spanish-style architecture, paved street, and street lights. A black-and-white photo in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-00032392



St. Mary's Catholic School-Sinai Hebrew Synagogue, 209 North Henderson Street, Jefferson, Marion County, TX. Interior of synagogue looking toward balcony and main entrance. Call Number: HABS TEX,158-JEF,14-. Collection: Historic American Buildings Survey (Library of Congress). Reproduction Number: HABS TEX,158-JEF,14-



A House in Parkland Estates on Lisbon Avenue later known as MacDill Avenue. Photo taken August 7, 1941; house owned by Zola E. Hagadorn (804 MacDill Avenue). A black and white photograph in jpeg format contributed by: Robertson and Fresh (Firm). Tampa System number R05-00015728



The Elks Club (425 Florida Avenue) and the on Union Bus Station (703 Cumberland Street) ca. 1920s when both streets were paved with brick. The automobiles are parked in front of the bus station. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-v-00000361



Traffic rushes past the Hotel Hillsboro on Florida Avenue in Hillsboro, Fla. Businesses shown include Floridan Flower Shop (905 Florida Avenue), Hillsboro Hotel (512 Florida Avenue corner of Twiggs Street), Powell's Flower Shop (505 Florida Avenue), and the Sunshine Laundry (503 Florida Avenue). A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-



The Hyde Park Presbyterian Church (later the Tampa Heights Presbyterian Church) (602 East Palm Avenue). A black and white photograph in jpeg format contributed by: Robertson and Fresh (Firm). Tampa System number R05-v-00000526 The Italian Club, or L'Unione Italiana Club (1725 East 7th Avenue) in Ybor City, Fla. Also shown are the street railroad tracks and power lines, brick paving, and automobiles. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-v-00000365



The Italian Club, or L'Unione Italiana Club (1725 East 7th Avenue) in Ybor City, Fla. Also shown are the street railroad tracks and power lines, brick paving, and automobiles. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-v-00000365



The Tribune Building on Tampa Street (602 Tampa Street cor Twiggs Street) on March 15, 1923. Note the delivery truck at the loading dock. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-00008758



The Times Building on Franklin and Washington Streets in Tampa on January 30, 1923. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-00008511

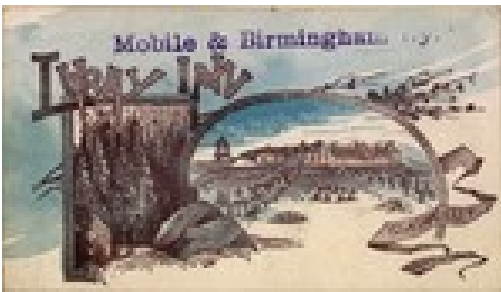


A house in Parkland Estates subdivision on Simms Boulevard on August 3, 1941. The house (2701 Simms Boulevard) is owned by Harry C. Quarles. A black and white photograph in jpeg format contributed by: Robertson and Fresh (Firm). Tampa System number R05-00015733





The German-American Club (6011 North Rome Avenue) in Ybor City, Fla. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-00004247



Luray Inn (Shenandoah Valley, Va.) color advertisement poster from early 1900s. Source Collection: Advertising Ephemera Collection. Reproduction number: A0596



Exterior view of the Heitler Resort, a house of prostitution at 310 North Peoria Street (formerly 169 Peoria Street), in the Near West Side community area of Chicago, Illinois. Cite as: DN-0008171, Chicago Daily News negatives collection, Chicago Historical Society. Gelatin 5 x 7 in. dry plate negatives. Reproduction number: DN-0008171



Exterior view of the Friends meeting house at Lake Forest, Illinois, with three women and two children on the porch and stairs. Photo taken in 1910. Gelatin dry plate 4 x 5 in. negative. Cite as: DN-0051770, Chicago Daily News negatives collection, Chicago Historical Society. Reproduction number: DN-0051770  
The Sears Roebuck and Company of Tampa, Florida exhibit at the 1938 Ruskin, Florida Tomato Festival at night. Merchandise available from the company is shown: bicycles, refrigerators, kitchen stoves, tires, a washing machine, a lawn chair, furniture, and a radio. A black-and-white photograph in jpeg format in the Robertson and Fresh Photograph Collection. Unique Identifier: R05-00011727

**Exercise 2 Handout**



View of a trolley driving on the Milwaukee and Racine electric road between Lake Bluff and Libertyville, Illinois in 1908. This image shows the trolley driving through a rural landscape with many electrical wires overhead. This image was taken as part of a Chicago Railways Company Trolley Trip. Cite as: DN-0006821, Chicago Daily News negatives collection, Chicago Historical Society.

Madison Drugs, Resnick's Cafeteria, Alvin Magnon Jewelers, Tampa Theater, ~~Robertson and Fresh Brothers~~, and Darling Department Store. A black-and-white photograph in jpeg format in the Robertson and Fresh Photograph Collection. Unique Identifier: R05-00015791



The Orange Pharmacy on Grand Central Avenue in Tampa, Florida viewed on August 14, 1921. In addition to prescription services, the pharmacy sells liquor and has a luncheonette. The building's façade has a Spanish architectural details with tile accents. A black and white photograph in jpeg format in the Robertson and Fresh Photograph Collection. Unique Identifier: R05-00015791.



The intersection of Franklin and Twig Streets in downtown Tampa, Florida during the 1940s. Trolley car tracks are shown in an intersection busy with foot and vehicular traffic. Includes views of these retailers:



Women dressed for La Verbena Del Tabaco Festival outside a window at the El Pasaje Hotel and Restaurant (1318-18 1/2 9th Avenue) in Ybor City, Fla. on October 22, 1928. Automobiles and commercial buildings are in the background.

A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-00027736



The Mecca Café at 1701 East 7<sup>th</sup> Avenue in the Ybor City district of Tampa, Florida, 1929. The café is a two story building with a covered porch wrapping around the second story. Several men stand on the sidewalk near the corner entryway to the café. Several automobiles are parked on the avenue as well as a delivery truck from La Joven Francesca Panderia . A black-and-white photograph in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-00030195



The Interior of the Las Novedades Caf, (1416 North 7th Avenue) in Ybor City, Fla. on June 19, 1929. The stamped tin ceiling, lighting and ceiling fans can be seen over the diners and staff. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-0002961



The Hillsborough Pharmacy and Liquors store (5317-19 Nebraska Avenue) on August 28, 1941. A black and white photograph in jpeg format contributed by: Robertson and Fresh (Firm). Tampa System number R05-00015839



The Dining room of the Davis Island Country Club (455 Bosphorus Avenue) in Tampa, Fla. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-000001234



The Dining room at the Hillsboro Hotel (512 Florida Avenue corner of Twiggs Street). A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-00035281



Rose Esperante dances for a couple in the patio room at the Columbia Restaurant in Tampa, Florida during the La Verbena del Tabaco Festival. The couple in evening clothes are seated at a table with cloth and full setting. The dancer wears a shawl and a Spanish style dress featuring large polka dots and ruffles. In the background a bartender makes cocktails behind a small bar decorated with ornate Spanish tiles. The room has a high ceiling and a small balcony is in view. n.d. A black-and-white photograph in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-v-00001771.



The Gasparilla King is honored in the banquet hall of the Tampa Terrace Hotel on October 20, 1953. Attendees were formal evening clothes. A black and white photograph in jpeg format in the Robertson and Fresh Photograph Collection. Unique Identifier: R05-6757



A Couple in a local bistro during the La Verbena del Tabaco Festival in 1937. Note the Spanish arched windows, the beautifully laid table, and the attentive waiter. A black-and-white photo in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-v-00001765



A Gasparilla Night Parade float sponsored by Sears and Roebuck is shown as it passes L. B. Chapman Grocery, IGA Supermarket in the Ybor City section of Tampa, Florida. The float has a nautical theme featuring sailboats with four women in swimsuits and a masked man. n.d. A black-and-white photograph in jpeg format in the Robertson and Fresh Photograph Collection. Unique Identifier: R05-z-00007207



The Shortstop Market on February 23, 1941. A black and white photograph in jpeg format contributed by: Robertson and Fresh (Firm). Tampa System number R05-00015154



A window display at White's Home Supply Store on Broadway in Tampa, Florida. The display, viewed from the street, features Health Club Baking Powder and Rumsford Baking Powder. n.d. A black-and-white photograph in jpeg format in the Robertson and Fresh Photograph Collection. Unique Identifier: R05-unk-00000011



The Temple Terrace Estates neighborhood near the Temple Terrace Country Club in Temple Terrace, Florida photographed on September 23, 1926. A black and white photograph in jpeg format in the Burger Brothers Photograph Collection. Unique Identifier: B29-00021114



Image of a man standing on a trolley car clearing the trolley wire in Chicago, Illinois, during a Chicago City Railway strike (Chicago City Railway Company workers strike, Chicago, Ill., 1903). Residential buildings are visible in the background. Cite as: DN-0001406, Chicago Daily News negatives collection, Chicago Historical Society.



Interior view of the Butler Brothers general merchandise company terrace and lunch room on July 28, 1913. Butler Brothers was located at the southeast corner of North Canal and West Randolph Streets in the Near West Side community area of Chicago, Illinois. Glass (gelatin dry plate) 4 x 5 in. negative. Cite as: DN-0060844, Chicago Daily News negatives collection, Chicago Historical Society. Reproduction number: DN-0060844



Image of an adult and a baby elephant standing on a dirt floor in the middle of an empty auditorium. Two men are standing with the elephants. Text on the image reads: Barnum Bailey Circus. Photo taken in 1926. Glass 4 x 5 in. negative. Cite as: DN-0080898, Chicago Daily News negatives collection, Chicago Historical Society.



View on Illinois State Fair Grounds, Springfield, Ill. on April 2, 1909.  
Gelatin silver prints. Georg, Victor, copyright claimant, photographer. Gelatin silver 7.5  
x 44.5 in. print. Part of: Panoramic photographs (Library of Congress) .



Panoramic view of Piccadilly [sic] Circus, London in 1909. Gelatin silver 10 x 54 in. print.  
Reproduction numbers: LC-USZ62-121634 DLC (b&w film copy neg. of left section),  
LC-USZ62-121635 DLC (b&w film copy neg. of right section). Part of: Panoramic  
photographs (Library of Congress).

### Exercise 3 Handout



A Coca-Cola delivery truck with driver and passenger is parked on a street in Tampa, Florida, 1921. The truck is parked near a building having Spanish architectural detail. Two bicycles lean against the building. Vehicles parked on an intersecting street are in view. A black-and-white photograph in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-00005298



Boat on a street flooded during the 1921 hurricane. Three people navigate a small boat near a building on a flooded street in Tampa, Florida after the 1921 hurricane. Spectators view the scene from the porch of a three story hotel or residence. Utility poles and an industrial smokestack are predominant. A black-and-white photograph in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-00006674.



A 1921 view of the Citrus Exchange Building (611-622 Franklin Street)/Maas Brothers department store (612-622 Franklin Street) at the corner of Franklin and Zack Streets in Tampa, Florida. Maas Brothers department store occupies the lower floors of what is primarily an office building in the downtown area. Cars are parked in varying configurations on the two streets and a bicycle and three pedestrians are in view. The top of a streetlight is in the foreground. A black-and-white photograph in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-00006557



The 1921 home of Max C. Gill at 710 South Delaware Street in Tampa, Florida. The house is a brick bungalow in the craftsman style with a porch, tile roof, porte cochere, and detached garage. A lamppost stands at the driveway entrance. A black-and-white photograph in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-00005715





The Columbia Restaurant float during the Gasparilla Festival Parade at the Florida Fair Grounds, 1932. The float is a representation of an open oyster shell. Three women wearing mermaid costumes sit in the shell. Signage on the float says “Columbia Restaurant, Gem of Spanish Restaurants”. A black-and-white photograph in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-f-00003366.



The Sears Roebuck and Company of Tampa, Florida exhibit at the 1938 Ruskin, Florida Tomato Festival at night. Merchandise available from the company is shown: bicycles, refrigerators, kitchen stoves, tires, a washing machine, a lawn chair, furniture, and a radio. A black-and-white photograph in jpeg format in the Robertson and Fresh Photograph Collection. Unique Identifier: R05-00011727



Five women in swimming costumes are posed standing in a large cigar box to promote La Verbena del Tabaco Festival, 1937. The central figure, Rose Esperante, holds a cigar above her head. The box decoration depicts a woman in Spanish dress, with mantilla, holding a fan. The box is labeled La Verbena del Tabaco Queens. A black-and-white photograph in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-v-00001763.



Rose Esperante dances for a couple at the Columbia Restaurant in Tampa, Florida during the La Verbena del Tabaco Festival. The couple in evening clothes are seated at a table with cloth and full setting. The dancer wears a shawl and a Spanish style dress featuring large polka dots and ruffles. In the background a bartender makes cocktails behind a small bar decorated with ornate Spanish tiles. The room has a high ceiling and a small balcony is in view. n.d. A black-and-white photograph in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-v-00001771.



Miss Latin American XII with her court, Latin American Festival, 1948. Seated under an arch, the festival queen is surrounded by eleven female attendants wearing Spanish style dress including headdress. A black-and-white photograph in jpeg format in the Robertson and Fresh Photograph Collection. Unique Identifier: R05-z-0003214



A Gasparilla Night Parade float sponsored by Sears and Roebuck is shown as it passes L. B. Chapman Grocery, IGA Supermarket in the Ybor City section of Tampa, Florida. The float has a nautical theme featuring sailboats with four women in swimsuits and a masked man. n.d. A black-and-white photograph in jpeg format in the Robertson and Fresh Photograph Collection. Unique Identifier: R05-z-00007207



A Gasparilla Parade float sponsored by Walgreen's Drugs ca. 1932. The float is adorned with representations of fountain drinks, cigars, oranges and a castle. Three women on the float wear pirate themed costumes. A black-and-white photograph in jpeg format in the Robertson and Fresh Photograph Collection. Unique Identifier: R05-00005024



A Gasparilla Parade float sponsored by Maas Brothers department store is shown traveling Bayshore Boulevard on a rainy festival day. The float decorated with representations of seahorses and palm trees and features men and women in different costumes. Davis Island is seen in the background. A black-and-white photograph in jpeg format in the Robertson and Fresh Photograph Collection. Unique Identifier: R05-z-00003047



A window display at White's Home Supply Store on Broadway in Tampa, Florida. The display, viewed from the street, features Health Club Baking Powder and Rumsford Baking Powder. n.d. A black-and-white photograph in jpeg format in the Robertson and Fresh Photograph Collection. Unique Identifier: R05-unk-0000011



Pedestrians and motor traffic move through the intersection of Franklin and Zack Streets in downtown Tampa, Florida, 1929.

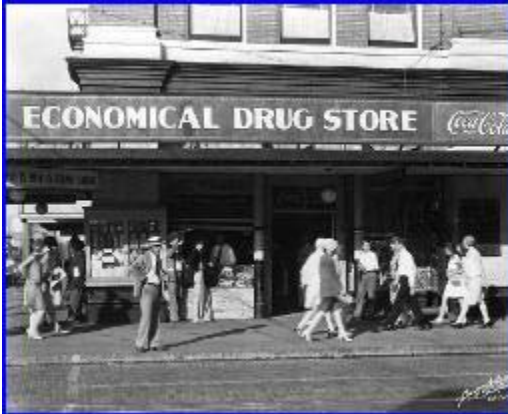
A three story office and retail business building is featured and a traffic light is visible. Businesses in the view are Don Thompson Radio Fixtures and Appliances, Home City News Stand, Peeler's Men Shop, West Coast Fertilizer Company, and William's Corner. A black-and-white photograph in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-rel-00000242



The Stovall Office Building at 416 Tampa Street, Tampa, Florida, 1950. Retail business occupy the ground floor. Automobiles and pedestrians are in view. A black-and-white photograph in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-00053051.



Traffic moves through the intersection of Bayshore Boulevard and Platt Street in Tampa, Florida. The view includes Tampa General Hospital on Davis Island and the approach to the Platt Street Bridge. 1939. A black-and-white photograph in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-v-00002068-p



The Economical Drug Store on the ground floor of the Stovall-Nelson Building at 817 Franklin Street in Tampa, Florida. The view shows the main entrance from a busy sidewalk busy with pedestrian traffic. One sign advertises Coca-Cola and another proclaims “This is not a chain store”. 1929. A black-and-white photograph in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-00029362



The Mecca Café at 1701 East 7<sup>th</sup> Avenue in the Ybor City district of Tampa, Florida, 1929. The café is a two story building with a covered porch wrapping around the second story. Several men stand on the sidewalk near the corner entryway to the café. Several automobiles are parked on the avenue as well as a delivery truck from La Joven Francesca Panderia . A black-and-white photograph in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-00030195



The intersection of Franklin and Twigg Streets in downtown Tampa, Florida during the 1940s. Trolley car tracks are shown in an intersection busy with foot and vehicular traffic. Includes views of these retailers: Madison Drugs, Resnick’s Cafeteria, Alvin Magnon Jewelers, Tampa Theater. A black-and-white photograph in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-v-00002376



Three women and a small dog look across Tampa Bay from the Venetian Apartments courtyard in Tampa, Florida. The courtyard is enclosed by an open work cement wall with a decorative wrought iron gate leading to a boat slip. Ornate lamp posts top the columns on either side of the gate. A black-and-white photograph in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-00022457



A “tin can tourist” camp at DeSoto Park near the shoreline of Tampa Bay in the 1920s. Tents are pitched near cars under palm and oak trees. Several men, women, and children are in view. A black and white photograph in jpeg format in the Burger Brothers Photograph Collection. Unique Identifier B29-00004868



Tin can tourists camp at Six Mile Creek in October 1925. A woman prepares a meal as a man washes his face. A table is set and children play nearby. Tents are pitched and cars are parked under trees draped with Spanish moss. A black and white photograph in jpeg format in the Burger Brothers Photograph Collection. Unique Identifier B29-00015531.



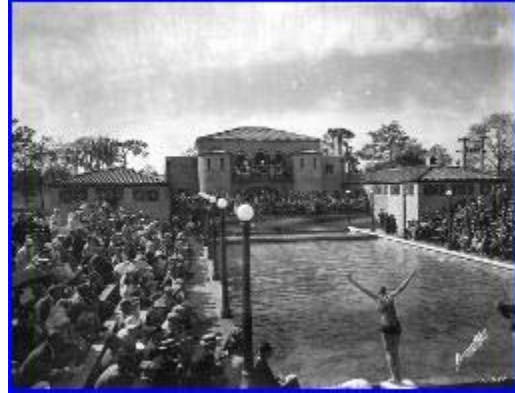
Tin can campers are shown at Palmetto Beach on the edge of Tampa Bay in the 1920s. Campers have pitched their tents and parked their cars under the palms near water's edge. A black and white photograph in jpeg format in the Burger Brothers Photograph Collection. Unique Identifier B29-v-00000220



Players and spectators at a Temple Terrace Golf and Country Club (119 North Glen Arven Avenue) golf tournament with views of the club grounds and buildings in this undated photograph. A black and white photograph in jpeg format in the Burger Brothers Photograph Collection. Unique Identifier: B29-00003327



A large crowd watches a February 4, 1926 promotional boxing match featuring Jack Dempsey at the Forest Hills Country Club in Tampa, Florida. A black and white photograph in jpeg format in the Burger Brothers Photograph Collection. Unique Identifier: B29-00017448



Many spectators attend a swimming meet at the Temple Terrace swimming pool. Lamp posts topped with spherical lights border the pool. Buildings in Spanish architectural style with tile roofs surround the pool. A black and white photograph in jpeg format in the Burger Brothers Photograph Collection. Unique Identifier: B29-00013290



The Temple Terrace Estates neighborhood near the Temple Terrace Country Club in Temple Terrace, Florida photographed on September 23, 1926. A black and white photograph in jpeg format in the Burger Brothers Photograph Collection. Unique Identifier: B29-00021114



King Gasparilla XV and his court ride his float in the Gasparilla Festival Parade at the Florida State Fair Grounds on February 4, 1924. A black and white photograph in jpeg format in the Burger Brothers Photograph Collection. Unique Identifier: B29-s-00000095



King Gasparilla XV (Marian Harvey) poses next to a piano in his chambers, 1924. The king is dressed in full costume including cape and crown. Two dolls in similar costumes sit on the piano and several cut out figures stand on the floor. A black and white photograph in jpeg format in the Burger Brothers Photograph Collection. Unique Identifier: B29-00009341-b



The Gasparilla King is honored with a banquet at the Tampa Terrace Hotel on October 20, 1953. Attendees were in formal evening clothes. A black and white photograph in jpeg format in the Robertson and Fresh Photograph Collection. Unique Identifier: R05-00024369



A man and a woman in a canoe on the Hillsborough River near Tampa, Florida ca. 1922. The canoe is near the riverbank under trees draped with Spanish moss. A black and white photograph in jpeg format in the Burger Brothers Photograph Collection. Unique Identifier: B29-00008462



The Orange Pharmacy on Grand Central Avenue in Tampa, Florida viewed on August 14, 1921. In addition to prescription services, the pharmacy sells liquor and has a luncheonette. The building's façade has Spanish architectural details with tile accents. A black and white photograph in jpeg format in the Robertson and Fresh Photograph Collection. Unique Identifier: R05-00015791.



In the shade of some pine trees adults and children share a meal in a picnic area near the waters of Tampa Bay in 1948. The picnic spot is on the Courtney Campbell Causeway. A black and white photograph in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-unk-00000020



Palm trees line the beach at Palmetto Shore near Rocky Point on Tampa Bay. A black and white photograph in jpeg format in the Burger Brothers Photograph Collection. Unique Identifier: B29-v-00000672



People gather under a stand of palm trees to enjoy a barbecue near Tampa, Florida, April 18, 1925. A black and white photograph in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-00014184



The disastrous effect of the 1921 hurricane on citrus crops is seen in this photograph. A tree is has been split and toppled and citrus is scattered on the ground. A black and white photograph in jpeg format in the Burger Brothers Photograph Collection. Unique Identifier: B29-00006715





A View down Davis Boulevard in Tampa, Florida with the Davis Island Country Club on the left and automobiles down the street. Also in view are the Palace of Florence (43 Davis Boulevard) and Sea-Born School (30 East Davis Boulevard). A black-and-white photograph in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-v-00002233



A Trolley car on Lafayette Street passing Hyde Park Avenue. in Tampa, Florida in 1918. Plant Park (West Lafayette Street opp Hyde Park Avenue) can also be seen. A black-and-white photograph in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-rel-00000185



An Automobile as it passes the Davis Island Country Club on palm-lined Davis Boulevard in Tampa, Florida. A black-and-white photograph in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-v-00001961



A View of Grand Central Avenue during the war the Second World War. Billboards advertising 7 up and Buy War Bonds, the Grand Central 5 and 10 Cent Store (704 Grand Central Avenue, Hyde Park, Rosemary Hotel (706 1/2 Grand Central Avenue), Shea and Prange Pharmacy (702 Grand Central Avenue), and Snow Park Service Station (602-604 Grand Central Avenue) can also be seen. A black-and-white photograph in jpeg format in the Robertson and Fresh Photograph Collection. Unique Identifier: R05-00016715



An Aerial view of a dirt street (later Hyde Park Avenue) passing through a residential area of Hillsborough Bay, Fla. Houses have already begun appearing in this area at this time of the picture. A black-and-white photo in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-rel-00000217



A Fred Herman's Sausage tractor trailer delivery truck in front of the Fred Herman's Meats (101 Governor) in Tampa, Fla. on April 11, 1942. A black-and-white photo in jpeg format in the Robertson and Fresh Photograph Collection. Unique Identifier: R05-00016337



The Intersection of Brevard Avenue, Grand Central Avenue, and Lafayette Street in the Hyde Park section of Tampa, FL. Note the gas street lights, sidewalks, homes, and trees lining these early Florida avenues. A black-and-white photo in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-rel-00000135



A Horse and delivery wagon carrying Turner and Wilson Coffee. The photo was taken ca. 1912 on Franklin Street in Tampa. A black-and-white photo in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-00000033



The Sulphur Springs Hotel and Apartments (8122 Nebraska Avenue) in Sulphur Springs, Fla. Also visible are automobiles, on-street parking, Kilgore Hardware Company (8440 Nebraska Avenue), Maves 5-10-25 Cent Stores (8138 Nebraska Avenue), Sulphur Springs Drug Store (8120 Nebraska Avenue), and Sulphur Springs Home Supply Store (8132-8134 Nebraska Avenue). A black-and-white photo in jpeg format in the Robertson and Fresh Photograph Collection. Unique Identifier: R05-00015829



The Home of Val Antuono. (4024 Bay to Bay Boulevard cor Clark Street) in Tampa. Note the Spanish-style architecture, paved street, and street lights. A black-and-white photo in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-00032392;



The Living room of the Casimiro Hernandez home (3108 Avon Avenue) in August 1942 showing the chandelier, furniture, fireplace, carpet and lamps. A black-and-white photo in jpeg format in the Robertson and Fresh Photograph Collection. Unique Identifier: R05-00016589



White's Home Supply Store on Broadway Avenue (4928 East 7th Avenue) in Tampa in 1938. A black-and-white photo in jpeg format in the Robertson and Fresh Photograph Collection. Unique Identifier: R05-00011676



The Reed Funeral Home (500 Twigg Street) on April 1, 1928. Note the brick-paved street, street-railroad tracks and utility poles. A black-and-white photo in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-rel-00000211;



A 1930's party at the spanish-themed Davis Island Country Club (455 Bosporus Avenue) during the La Verbena Del Tabaco Festival. A black-and-white photo in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-00027732



A House on Marion and Jackson Streets in 1922 with the Tampa City Hall (315 Lafayette Street) in the distance and automobiles in the foreground. A black-and-white photo in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-rel-00000189



Dancer Rose Esperante dressed in Spanish attire for the 1937 La Verbena del Tabaco Festival in Tampa. A black-and-white photo in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-v-00001767;



Spanish Dancers in 1940 in Tampa. A black-and-white photo in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-00058645



A man serenading a seniorita (Rose Esperante) at the Davis Island Country Club (455 Bosphorus Avenue) during the 1937 La Verbena del Tabaco Festival. The audience really appreciated the entertainers. A black-and-white photo in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-v-00001752



A Couple in a Spanish restaurant during the La Verbena del Tabaco Festival in 1937. Note the Spanish arched windows, the beautifully laid table, and the attentive waiter. A black-and-white photo in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-v-00001765



A Spanish guitarist and Louise Esperante in Spanish attire during the La Verbena del Tabaco Festival on September 11, 1938 at the Davis Islands Country Club (455 Bosphorus Avenue). The dinner was delicious and the entertainment lively. Unique Identifier: B29-v-00001952



A "Jack and the Bean Stalk" float passing in front of Haber's Department Store (613 Franklin Street) during the 1956 Children's Gasparilla Parade in Tampa. A black-and-white photo in jpeg format in the Robertson and Fresh Photograph Collection. Unique Identifier: R05-00030188



A marching band and some women in uniform round the bend from Franklin Street to Lafayette Street near Postal Telegraph Company (406 Franklin Street) during a parade. A black-and-white photo in jpeg format in the Robertson and Fresh Photograph Collection. Unique Identifier: R05-z-00002570



A Banbury Cross float passing Exchange National Bank (601 Franklin Street cor Twiggs Street) at the Children's Gasparilla Parade in 1956. A black-and-white photo in jpeg format in the Robertson and Fresh Photograph Collection. Unique Identifier: R05-00030231



Crowds watch the Jose Gaspar sail up the river, through an open bridge, and into downtown Tampa during the 1938 Gasparilla Pirate Invasion. A black-and-white photo in jpeg format in the Burgert Brothers Photograph Collection. Unique Identifier: B29-f-00003138



Tampa's Bay Shore Boulevard in 1922 when it was a brick road. A blank and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-v-00000426



A House in Parkland Estates on Lisbon Avenue later known as MacDill Avenue. Photo taken August 7, 1941; house owned by Zola E. Hagadorn (804 MacDill Avenue). A black and white photograph in jpeg format contributed by: Robertson and Fresh (Firm). Tampa System number R05-00015728



A School bus transversing a flooded road during a Tampa flood. Date unknown. A black and white photograph in jpeg format contributed by: Robertson and Fresh (Firm). Tampa System number R05-z-00003011



Poinsettia Dairy Products (3319 Florida Avenue), Incorporated on Florida Avenue, showing their delivery trucks and their affiliation with Borden's Ice Cream Dairy. Photo taken February 2, 1947. A black and white photograph in jpeg format contributed by: Robertson and Fresh (Firm). Tampa System number R05-00020562



Elkes Pontiac Company (1101 Florida Avenue corner of Harrison) on October 2, 1953. A delivery truck is shown in front of the service station. A black and white photograph in jpeg format contributed by: Robertson and Fresh (Firm). Tampa System number R05-00024299



Traffic rushes past the Hotel Hillsboro on Florida Avenue in Hillsboro, Fla. Businesses shown include Floridan Flower Shop (905 Florida Avenue), Hillsboro Hotel (512 Florida Avenue corner of Twiggs Street), Powell's Flower Shop (505 Florida Avenue), and the Sunshine Laundry (503 Florida Avenue). A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-00043518



The Elks Club (425 Florida Avenue) and the on Union Bus Station (703 Cumberland Street) ca. 1920s when both streets were paved with brick. The automobiles are parked in front of the bus station. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-v-00000361



Night time on 7th Avenue in Ybor City, Fla. ca. 1927. Shown are automobiles in front of Henry Clay Shoe (1602 East 7th Avenue) and Smith and Smith Chiropractic. Also included are the street railroad tracks in the 1920s brick paving. A black and white photograph in jpeg format contributed by: Burgert Brothers. System number B29-unk-00000021





The Broadway Service Station (901 East 7th Avenue) on East Broadway Avenue in Ybor City, Fla. ca. 1926 soon after the paving of East Broadway with brick. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-00022326



1903 Road construction on Florida Avenue in front of the Tampa Furniture Company (707-709 Florida Avenue) in Hillsborough, Fla. Seen are a crew of construction works in a water main ditch, mixed brick and stone paving, and the street railroad tracks. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-00042357



A View down 7th Avenue from 15th Street in the Ybor City section of Tampa, Fla. in 1920. A Tampa Trolley car can be seen near Katz's Adams and Company (1430 East 7th Avenue). A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa--Ybor City System number B29-rel-00000319



The Ybor City Presbyterian Mission Church Assembly Hall (953 11th Avenue) on April 5, 1942. A black and white photograph in jpeg format contributed by: Robertson and Fresh (Firm). Tampa System number R05-00016317



The Hyde Park Presbyterian Church (later the Tampa Heights Presbyterian Church) (602 East Palm Avenue). A black and white photograph in jpeg format contributed by: Robertson and Fresh (Firm). Tampa System number R05-v-00000526



Image of many children assembled next to and in a cable car during a Chicago Daily News Fresh-Air Fund outing in Chicago, Illinois ca. 1902. Some of the children are standing on the cable car's running boards and several are leaning out of the car's windows. Cite as: DN-0000090, Chicago Daily News negatives collection, Chicago Historical Society.



The Steamship Pokenoket sails past the Tampa Bay Hotel (West Lafayette and Hyde Park Avenue) on the Hillsborough River (Polk County--Hillsborough County, Fla.) in 1925. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-00015079



Near view of a bridge and locks on the drainage canal at Lockport, Illinois. This image was taken as part of a Chicago Railways Company Trolley Trip in 1908. Cite as: DN-0006895, Chicago Daily News negatives collection, Chicago Historical Society.

**Exercise 4 Handout**

Women dressed for La Verbena Del Tabaco Festival outside a window at the El Pasaje Hotel and Restaurant (1318-18 1/2 9th Avenue) in Ybor City, Fla. on October 22, 1928. Automobiles and commercial buildings are in the background.

A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-00027736



The Interior of the Las Novedades Caf, (1416 North 7th Avenue) in Ybor City, Fla. on June 19, 1929. The stamped tin ceiling, lighting and ceiling fans can be seen over the diners and staff. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-0002961



Sailing ships docked at Lee Terminal (Lee Terminal and Warehouse Corp. (Ella Mae Avenue cor 13th Street)) at the Port of Tampa, located on the Hillsborough River (Polk County-Hillsborough County, Fla.). Photo taken October 13, 1919. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-rel-00000143



The Italian Club, or L' Unione Italiana Club (1725 East 7th Avenue) in Ybor City, Fla. Also shown are the street railroad tracks and power lines, brick paving, and automobiles. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-v-00000365



Roy Rogers, the famous motion picture actor, and his horse Trigger at the Florida State Fair in 1959. They are in front of Trigger's personal trailer. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-00076640



Women dressed for the La Verbena Del Tabaco Festival in front of the Centro Asturiano (1913 Nebraska Avenue) in Ybor City. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-00027734



La Verbena Del Tabaco Festival at the Centro Asturiano (1913 Nebraska Avenue) in Ybor City, Fla. ca. 1931. Dancers, diners and spectators can all be seen under the electric lights, fans and tin ceiling. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-00027740



The Cargo ship Neches at the Tampa Union Terminal (1004 Finley Street) in the Sparkman Channel of the Tampa Harbor. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-00028444



Boats on the Hillsborough River near Plant Park (West Lafayette Street opp Hyde Park Avenue) on the Hillsborough River (Polk County-Hillsborough County, Fla.). A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-v-00000732



An Aerial view of Tampa above Lafayette Street in 1914. Structures in photo include Hillsborough County Court House (401 Frankilin Street bet Madison and Lafayette Streets) and Bay View Hotel (208 Jackson Street). Also shown are trolley cars, street-railroads and automobiles. Tampa System number B29-00001186



Brick paved Memorial Highway built on sand through a forest grove. Photo taken ca. 1920s. A black and white photograph in jpeg format contributed by: Burgert Brothers.



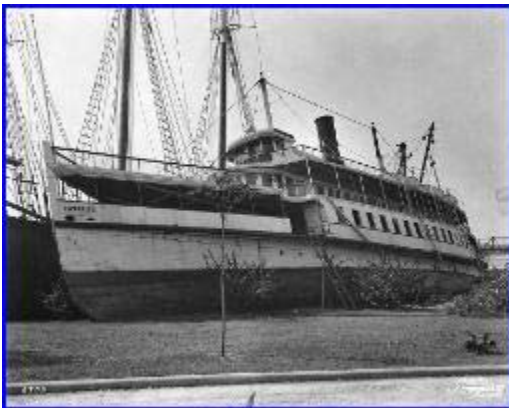
View of the Imperial Cafe (1501 6th Avenue) on the left. Cross street is 4th Avenue. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-00000410



The German-American Club (6011 North Rome Avenue) in Ybor City, Fla. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-00004247



The Times Building on Franklin and Washington Streets in Tampa on January 30, 1923. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-00008511



The Ferry boat the Favorite on dry land after the Hurricane of 1921. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-00006733



The Tribune Building on Tampa Street (602 Tampa Street cor Twiggs Street) on March 15, 1923. Note the delivery truck at the loading dock. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-00008758



A Glass truck crossing the Gandy Bridge on Tampa Bay. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-00012520



The lobby and check-in desk of the Floridan Hotel (905 North Florida Avenue ne cor Cass Street) in 1926. Note hotel employees in front of check-in desk. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-00022486



The Dining room of the Davis Island Country Club (455 Bosphorus Avenue) in Tampa, Fla. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-00013950



Cars passing the Las Novedades Cafe (1427 East 7th Avenue) and Broadway Cafe (1427 East 7th Avenue) in Ybor City ca. 1920s. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-00033628



Women boxing cigars at Cuesta Rey and Company (2015 North Howard Avenue) in 1947. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-00034993



The rides and concession stands at the 1924 Florida State Fair in Tampa. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-f-00000330



The Dining room at the Hillsboro Hotel (512 Florida Avenue corner of Twiggs Street). A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-00035281



An Atlantic Coast Line Railroad steam locomotive on display at the Atlantic Coast Line Railroad Company (101 Lafayette corner of Ashley Street) in 1929. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-f-00000946





An National Airlines, Inc. airplane on the tarmac of Peter O. Knight Airport (700 West Davis Boulevard) on Davis Island in Tampa, Fla. on December 20, 1940. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-v-00002219



A house in Parkland Estates on Simms Boulevard on August 3, 1941. The house (2701 Simms Boulevard) is owned by Harry C. Quarles. A black and white photograph in jpeg format contributed by: Robertson and Fresh (Firm). Tampa System number R05-00015733



Men working on a banana shipment at Port Tampa. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-v-00002426



The bar of the Colombia Spanish Restaurant (2117 East 7th Avenue) on April 19, 1942. A black and white photograph in jpeg format contributed by: Robertson and Fresh (Firm). Tampa System number R05-00008076



The Eastern Air Lines Ticket Office (415 Florida Avenue) on October 16, 1941. A black and white photograph in jpeg format contributed by: Robertson and Fresh (Firm). Tampa System number R05-00015951



The Shortstop Market on February 23, 1941. A black and white photograph in jpeg format contributed by: Robertson and Fresh (Firm). Tampa System number R05-00015154



The Hillsborough Pharmacy and Liquors store (5317-19 Nebraska Avenue) on August 28, 1941. A black and white photograph in jpeg format contributed by: Robertson and Fresh (Firm). Tampa System number R05-00015839



The Tampa International Airport (700 West Davis Boulevard) on November 20, 1952. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-v-00002953



The Interior of the Las Novedades Café (1416 North 7th Avenue) in Ybor City on June 19, 1929. Note the display case, diners and staff under the stamped tin ceiling, lighting, and ceiling fans. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-00029611



The Pool and garden of the Davis Islands Country Club (455 Bosphorus Avenue). A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-00032399



Looking down 7th Avenue at a trolley car as it passes the Las Novedades Café (1416 North 7th Avenue), Ritz Theater (1503 East 7th Avenue) and Ybor City Chamber of Commerce (1710 East 7th Avenue) in 1937. A black and white photograph in jpeg format contributed by: Burgert Brothers. Tampa System number B29-v-00001864



Image of a man helping a girl climb down from a delivery wagon carrying passengers during the June 1916 street car strike in Chicago, Illinois. Other passengers wait their turn to climb down. Men carrying umbrellas are visible walking across the street in the background, and street car tracks are visible on the street. Glass 4 x 5 in. negative. Cite as: DN-0064592, Chicago Daily News negatives collection, Chicago Historical Society. Reproduction number: DN-0064592



Interior view of the Schultz Drug Co. on 16th (Sixteenth) Street in Denver, Colorado ca. 1890s?. Men and a woman stand behind the display cases and African American men, who include Clarence F. Holmes, Sr., stand behind the lunch counter. Photoprint (4 x 6 in.) on mat board. Source: Clarence F. Holmes MSS Collection. Reproduction number: X-22356



Image of a lunch room on the river dock in Chicago, Illinois in 1906. Two men are standing in the doorway of the lunch room, and a sign painted on the exterior wall reads, lunches put up at quick notice. Glass 5 x 7 in. negative. Cite as: DN-0003793, Chicago Daily News negatives collection, Chicago Historical Society. Reproduction number: DN-0003793



Stereoscopic views of the Metropolitan Sanitary Fair and other exhibitions, New York City ca. 1864-1875?. Cite as: Robert Dennis Collection of Stereoscopic Views.

**\* Sample Thesaurus Front Matter \***

## **The Thesaurus for Nonprofit Information**

### **1. Scope, Purpose and Users**

The Thesaurus for Nonprofit Management (TNM) provides a set of terms that may be used in the cataloging and organization of materials specific to management of nonprofit organizations. TNM may also be used in creating web page FAQs and manual indexing systems such as vertical files. The main concepts in this thesaurus cover nonprofit management, finance and human resources. Peripheral concepts include service fields, regulations, and organizational change. The thesaurus does not include terms to describe the different service fields of nonprofit organizations or types of grants available for nonprofit organizations, as a thesaurus covering these terms is available through the Foundation Center and the National Taxonomy of Exempt Entities. (See Appendix A for links to these sources.)

TNM contains a controlled vocabulary of terms used to describe activities in the operation of nonprofit organizations and those performing these activities. *Guidelines for the Construction, Format, and Management of Monolingual Thesauri* (ANSI/NISO Z39.19 – 2003) and *Thesaurus construction and use: a practical manual* by Aitchison et al were used in the construction of this thesaurus. Every attempt has been made to adhere to the standards and guidelines noted in these sources.

As a general rule TNM does not include proper nouns, but there are exceptions. These terms are indicated through the use of capitalization.

TNM is intended for librarians or others developing resource material for use by nonprofit organizations. Management support organizations (MSO) are also target users as they provide direct technical assistance to nonprofit organizations.

### **2. Source Information and Software**

In addition to existing thesauri, experts in the field of nonprofit management and other classification systems were consulted. (See Appendix B for a list of resources consulted.) Other sources of “word stock” came from back-of-the-book indexes of selected nonprofit management books and an accumulation of FAQs. TNM was created using TCS-8 WebChoir software.

A general overview of our treatment of source material is outlined here. For specific information on the entry policy used in TNM see Section 4. Using electronic resources of classification systems in the domain, we concatenated the files and did a word frequency count to create a basis for our word stock list. We also analyzed each of the major divisions used in the C-MAP classification systems and found similarity in the major categories and organization. These terms were used to help create our preliminary set of descriptors. After analyzing the major divisions, we entered the UNESCO terms into TCS-8 and used these descriptors as a guide during construction. Once our word stock list was entered, we identified several “main” concepts including management operations, persons involved in nonprofit management and

services provided by nonprofit organizations. We divided these main concepts into main facets similar to the fundamental facets mentioned in Aitchison. We determined that creating facets that mirrored the life cycle of a nonprofit organization would be the most logical organization of the terms. As such, we start with startup of a nonprofit organization and follow it through to the dissolution. The remainder of our work involved creating relationships between our descriptors and adding scope notes where necessary. The number of word stock terms and resources for this area were overwhelming. However, once we completed the facet analysis and started arranging the descriptors according to the main facets, the work became manageable and we had a systematic way to effectively populate the thesaurus.

### 3. Layout and Relationships

TNM currently contains 199 preferred terms and 65 non-preferred or cross-reference terms. The relationships included in TNM are equivalence relationships, hierarchical relationships and associative relationships. The reference structure for each authorized term includes:

- broader terms (BT) – a preferred term which indicates the more general class to which the term belongs
  - Accounting
  - BT Financial administration
- narrower terms (NT) – a preferred term which indicates the narrower scope and is a member of a broader class
  - Board committees
  - NT Audit committees
- related terms (RT) – a preferred term that indicates a concept that is closely related to another preferred term
  - Bylaws
  - RT Board of directors
- use for terms (UF) – a non-preferred term that is provided to help identify meaning or scope of term; (the reciprocal is USE)
  - Consultants
  - UF Contractor
- use for alternate (UF+) – an alternate language term that is provided to assist users from a broader scope; (the reciprocal is SEE)
  - Labor market
  - UF+ Labour market
- use terms (USE) – a cross reference term that is used to direct catalogers to a preferred term; (the reciprocal is UF)
  - Chief executive officers
  - USE Executive directors

- see terms (SEE) – a cross reference term that is used to direct cataloguers to an alternate language term; (the reciprocal is UFA)
  - Labour relations
  - SEE Labor relations
- scope notes (SN) – a note that specifies the meaning of a term
  - Employees
  - SN Includes full time and part time workers that are employed by an organization
- source notes (SRC) – a note that specifies the source material of a term
  - Employment opportunities
  - SRC UNESCO

Since we are using TCS-8, reciprocal markings are made automatically.

TNM is displayed in several formats, including an alphabetical format and hierarchical format. Preferred and non-preferred terms are interfiled in the alphabetical listing. These two displays are meant to support the user in different ways. The alphabetical display allows the user the ability to look up a known term to determine if it is the preferred term or to find the appropriate preferred term for non-preferred terms. The hierarchical display supports subject browsing by the user to find the appropriate term for the materials to be assigned subject terms.

We would have liked to include a two-way hierarchical term display in lieu of two separate displays, because it allows movement both up and down the hierarchy while maintaining the ease of known term access that is afforded by an alphabetical display. The potential benefits of this display were eventually outweighed by concerns that this type of display would require more long-term maintenance, lack of display of related terms, and the potential for each term to exceed space limitations.

The use of a permuted index would also have been an option, because both preferred and non-preferred terms are listed in alphabetical order for ease of access, and because of the way the words organize themselves within the permuted structure. The permuted index separates all of the terms into separate word entries, so that access is by a single word even for compound terms. Because of this, there are term groupings that occur around a given entry that have a relationship with the term. This decreases the use of related terms, and also preserves some of the hierarchical relationships of main entry terms. A permuted indexing display may be provided in future versions of TNM.

#### **4. Term Selection and Syntax**

The ANSI standards were followed as close as possible in determining word usage. Natural language and standard American English were used. It is important to note that the UNESCO terms often contained British English. In these cases, American English was preferred.

##### ***4.1. Single Word Terms***

Headings for concrete concepts were entered in the plural (eg. Managers). Headings for abstract concepts were entered in the singular form (eg. Management). Heading for activities (operations, procedures) were entered in the gerund form (eg. Planning) whenever possible.

#### ***4.2 Compound Word Terms***

When a concept could not be expressed using a single term, a compound term consisting of two or more words was used. Words in compound entries were made using natural word order and not inverted.

#### ***4.3 Parenthetical Qualifiers***

Parenthetical qualifiers were added to terms in order to clarify meaning, as in the case of homographs or when a different sense of the same word was used.

- Liabilities (economic)

#### ***4.4 Capitalization***

The first word in each term is capitalized, unless the term is a proper name, in which case all of the appropriate words are capitalized within the term.

- Financial statements
- Internal Revenue Service

#### ***4.5 Scope Notes***

Scope notes were used for terms in order to clarify meaning, or to clarify usage and scope of a term.

## **5. Using TNM in Searching and Indexing**

TNM was designed primarily to be used as an indexing tool. How the thesaurus will be used will depend on the implementation of TNM and the intended outcomes. For example, web pages of FAQs would tend to use the higher levels of TNM hierarchy in order to organize and provide access to their materials. For cataloging of materials general cataloging principles should be used to determine what the materials are "about" and apply the appropriate terms. These terms should also be applied using the rule of specificity, where the most "specific" term possible to be applied to the material over the more "generic".

## **6. TNM Updating Procedures**

As with other thesauri, TNM will be updated to:

- correct an existing term
- change the status of a term
- remove a term
- add a term, and
- make changes to the hierarchy.

Some cases in which TNM will have to be edited include changes to existing laws or legislation that affects nonprofit organizations or new reporting procedures or requirements. In addition to



legal changes, the introduction of new office technologies may also need to be reflected in TNM. We have envisioned using TNM for cataloging, creating vertical files and organizing FAQs. There may be other applications yet unknown which could affect TNM and require changes. Literary precedent for new terms will also justify changes to TNM. Users are encouraged to provide feedback as often as possible to TNM Management Staff.

As TNM will be used by different catalogers, a worksheet (See Appendix D) will be used to note and / or request changes. The process is generally as follows: report filed with TNM Management Staff; report reviewed and decision made; TNM updated by TNM Management Staff. Request changes should be forwarded to Monique Leahey-Sugimoto and Jillian Wallis.

### **Appendix A: External Thesaurus Referenced Within TNM**

[http://fdncenter.org/research/grants\\_class/thesaurus.html](http://fdncenter.org/research/grants_class/thesaurus.html)

### **Appendix B: Word stock Sources**

ARDI Nonprofit Management and Leadership Taxonomy (1997) included in the paper: A Research Report of the Applied Research & Development International (ARDI) Access to Management Assistance in the United States.

C-MAP, Humboldt: The Humboldt Area Foundation –Humboldt (Subject Files)

C-MAP, Los Angeles: Center for Nonprofit Management (Vertical File Listing, Print Collection Labels)

C-MAP, Orange County: Nonprofit Resource Center (Vertical File Listing)

C-MAP, Sacramento: Nonprofit Resource Center – Sacramento (Print Collection labels, Vertical File Listing,)

C-MAP, San Diego: Nonprofit Management Solutions (Vertical File Listing)

C-MAP, San Francisco: CompassPoint Nonprofit Services (Webpage directory)

C-MAP, Santa Barbara: Nonprofit Support Center of Santa Barbara (Vertical File Listing)

C-MAP, Sonoma: Center for Volunteer and Nonprofit Leadership of Marin (Vertical File Listing)

C-MAP, Sonoma: Resource Center for Nonprofits - Sonoma County (Vertical File Listing)

Foundation Center Library Classification Scheme

Thesaurus of ERIC Descriptors (Various areas)

Thesaurus of Sociological Indexing Terms 14<sup>th</sup> Edition (Mangement)

Topical list from the Free Management Library ([www.managementhelp.org](http://www.managementhelp.org))

UNESCO Thesaurus (6.7-6.85) [www.ulcc.ac.uk/unesco](http://www.ulcc.ac.uk/unesco) (Finance & Trade, Organization & Management, Personnel, Labour)

### **APPENDIX C: Reference Sources**

Aitchison, J., Gilchrist, A., Bawden, D. (2000) *Thesaurus construction and use: A practical manual*. London; Aslib IMI.

ANSI/NISO Z39.19-2003. *Guidelines for the Construction, Format, and Management of Monolingual Thesauri*.

**Appendix D: TNM Term Review Request Form** (Modified from ANSI/NISO Z39.19-2003)

TNM Term Review Request Form	
Date: ____/____/____	
Submitted By: _____ Title: _____	
Term:	
<b>ACTION REQUESTED:</b>	
<input type="checkbox"/> Add As New Term <input type="checkbox"/> Preferred <input type="checkbox"/> Non-preferred  <input type="checkbox"/> Cross-Reference Add:                      Delete: <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> Change Term To: <input type="checkbox"/> Preferred <input type="checkbox"/> Non-preferred  <input type="checkbox"/> Delete Term: Delete Only:           Delete and Transfer: <input type="checkbox"/> <input type="checkbox"/> _____
<b>AMPLIFYING INFORMATION</b> Scope Note Information	
<b>CROSS REFERENCE STRUCTURE</b>	
USE(U):	USED FOR (UF):
BROADER TERM (BT):	NARROWER TERM (NT):
RELATED TERMS (RT):	
<b>COMMENTS:</b>	
<b>REVIEW AUTHORITY</b>	
REVIEW COMMENTS:	
<b>ACTION:</b> <input type="checkbox"/> Approved <input type="checkbox"/> Disapproved	REVIEWER SIGNATURE: _____ DATE: _____

## Resources

### Meta Sites

American Society of Indexers.

<http://www.asindexing.org/site/index.html>

Cover Pages. Resource Description and Classification.

<http://www.oasis-open.org/cover/classification.html>

Queensland University of Technology : Controlled Vocabularies

[http://sky.fit.qut.edu.au/~middletm/cont\\_voc.html](http://sky.fit.qut.edu.au/~middletm/cont_voc.html)

University of British Columbia, School of Library, Archival and Information Studies. Indexing Resources on the WWW.

<http://www.slais.ubc.ca/resources/indexing/database1.htm>

### Bibliographies

Queensland University of Technology : Controlled Vocabularies.

[http://sky.fit.qut.edu.au/%7Emiddletm/cont\\_voc.html#Bibliography](http://sky.fit.qut.edu.au/%7Emiddletm/cont_voc.html#Bibliography)

University of British Columbia, School of Library, Archival and Information Studies. Indexing Resources on the WWW.

<http://www.slais.ubc.ca/resources/indexing/database1.htm>

Willpower Information. Publications on thesaurus construction and use.

<http://www.willpower.demon.co.uk/thesbibl.htm>

### Controlled Vocabulary – Construction

American Society of Indexers : How do I build a thesaurus?

<http://www.asindexing.org/site/thesbuild.shtml>

Craven, T. Introductory Tutorial on Thesaurus Construction. London, Ontario, Canada: University of Western Ontario, Graduate School of Library and Information Science.

<http://instruct.uwo.ca/gplis/677/thesaur/main00.htm>

HILT : High-Level Thesaurus.

<http://hilt.cdlr.strath.ac.uk/Sources/index.html>

Hoy, M. (1998). Understanding official government terminology: natural language searching and government thesauri. Canberra: National Archives of Australia.

[http://www.naa.gov.au/recordkeeping/gov\\_online/agift/gov\\_term/intro.html](http://www.naa.gov.au/recordkeeping/gov_online/agift/gov_term/intro.html)

International Organization for Standardization. (fee)

ISO 2788:1986 Documentation -- Guidelines for the establishment and development of monolingual thesauri. (charge)

ISO 5964:1985 Documentation -- Guidelines for the establishment and development of multilingual thesauri

<http://www.iso.org/iso/en/ISOOnline.frontpage>

Leise, Fred, and Karl Fast, Mike Steckel. All About Facets & Controlled Vocabularies, December 9, 2002 (series of articles).

[http://www.boxesandarrows.com/view/all\\_about\\_facets\\_controlled\\_vocabularies](http://www.boxesandarrows.com/view/all_about_facets_controlled_vocabularies)

Morville, Peter. Building a Synonymous Search Index. Dr. Dobb's Portal, January 1, 2002.

<http://www.ddj.com/dept/architect/184412944>

Morville, Peter. How Do You Build a Thesaurus?, January 1, 2002.

<http://www.ddj.com/dept/architect/184412945>

National Archives of Australia. Developing a Functions Thesaurus, Revised July 2003.

[http://www.naa.gov.au/recordkeeping/control/functions\\_thesaur/contents.html](http://www.naa.gov.au/recordkeeping/control/functions_thesaur/contents.html)

National Information Standards Organization (U.S.)(2005). Guidelines for the construction, format, and management of monolingual controlled vocabularies. (ANSI/NISO Z39.19-2005)

<http://www.niso.org/standards/resources/Z39-19-2005.pdf>

NISO/ASI/ALCTS Workshop on Electronic Thesauri: planning for a standard, November 1999, Washington DC, USA.

[http://www.niso.org/news/events\\_workshops/thesau99.html](http://www.niso.org/news/events_workshops/thesau99.html)

Rosenfeld, Lou. Cuisinarts, E-Commerce, and ... Controlled Vocabularies, January 1, 2002.

<http://www.ddj.com/dept/architect/184412772>

State Records NSW. Guideline 15 Developing and implementing a keyword thesaurus.

[http://www.records.nsw.gov.au/recordkeeping/guideline\\_15\\_keyword\\_thesaurus\\_7273.asp](http://www.records.nsw.gov.au/recordkeeping/guideline_15_keyword_thesaurus_7273.asp)

and

[http://www.records.nsw.gov.au/recordkeeping/dirks-thesaurus\\_1858.asp](http://www.records.nsw.gov.au/recordkeeping/dirks-thesaurus_1858.asp)

University of British Columbia, School of Library, Archival and Information Studies. Indexing Resources on the WWW.

<http://www.slais.ubc.ca/resources/indexing/database1.htm>

Will, L. Thesaurus principles and practice.

<http://www.willpower.demon.co.uk/thesprin.htm>

## Software Lists/Reviews

American Society of Indexers. Thesaurus Management Software.

<http://www.asindexing.org/site/thessoft.shtml>

Betz, Anne. Thesaurus Management Software.

[http://www.fbi.fh-koeln.de/institut/labor/Bir/thesauri\\_new/thsoften.htm](http://www.fbi.fh-koeln.de/institut/labor/Bir/thesauri_new/thsoften.htm)

Thesaurus Construction Software

<http://www.webindexing.biz/articles/ThesaurusConstruction.htm>

University of British Columbia, School of Library, Archival and Information Studies. Indexing Software.

<http://www.slais.ubc.ca/resources/indexing/software.htm>

Willpower Information. Software for building and editing thesauri.

<http://www.willpowerinfo.co.uk/thessoft.htm>

## **Lists of Controlled Vocabularies**

American Society of Indexers. Thesauri Online.

<http://www.asindexing.org/site/thesonet.shtml>

Betz, Anne. Classification schemes and thesauri on-line.

[http://www.fbi.fh-koeln.de/institut/labor/Bir/thesauri\\_new/indexen.htm](http://www.fbi.fh-koeln.de/institut/labor/Bir/thesauri_new/indexen.htm)

BUBL LINK Catalogue of Internet Resources. Thesauri.

<http://bubl.ac.uk/link/types/thesauri.htm>

Controlled Vocabulary. Example Controlled Vocabularies and Thesauri available online.

<http://www.controlledvocabulary.com/examples.html>

HILT : High-Level Thesaurus. A-Z of Thesauri.

<http://hilt.cdlr.strath.ac.uk/Sources/thesauri.html>

Lutes, Barbara. Web Thesaurus Compendium.

<http://www.ipsi.fraunhofer.de/%7Elutes/thesoecd.html>

Queensland University of Technology : Controlled Vocabularies

[http://sky.fit.qut.edu.au/~middletm/cont\\_voc.html](http://sky.fit.qut.edu.au/~middletm/cont_voc.html)

Taxonomy Warehouse.

<http://www.taxonomywarehouse.com/>

Thesaurus Construction Software

<http://www.webindexing.biz/articles/ThesaurusConstruction.htm>

University of British Columbia, School of Library, Archival and Information Studies. Indexing Resources on the WWW : Links to Online Thesauri, Term Lists & Classification Schemes.

<http://www.slais.ubc.ca/resources/indexing/database1.htm>

University of Toronto, Faculty of Information Studies. Subject Analysis Systems (SAS) (Repository).

<http://www.fis.utoronto.ca/content/view/386/134/>

*THIS PAGE INTENTIONALLY LEFT BLANK  
FOR DOUBLE SIDED COPY*

## Controlled Vocabulary and Thesaurus Design Course bibliography

### Background readings

Aitchison, J., Gilchrist, A., and Bawden, D. (2000; 4<sup>th</sup> ed.). *Thesaurus construction and use: A practical manual*. London: ASLIB.

Bates, Marcia (1996). Indexing and access for digital libraries and the Internet: Human, database, and domain factors. Available at:  
<http://www.gseis.ucla.edu/faculty/bates/articles/indexdlib.html>

Cimino, James J. (2000). From data to knowledge through concept-oriented terminologies: experience with the medical entities dictionary. *Journal of the American Medical Informatics Association* 7, 288-297.

Ezzo, Jeannette (2005). Bella and Yakov and Tillie's panties: What I learned in "Construction and Maintenance of Indexing Languages and Thesauri." *Bulletin of the American Society for Information Science and Technology* 31(4).

Ganzmann, Jochen (1990). Criteria for the evaluation of thesaurus software. *International Classification* 17 (3/4), 148-157. Available at:  
<http://www.willpower.demon.co.uk/ganzmann.htm>

Hudon, Michele (1997). Multilingual thesaurus construction: Integrating the views of different cultures in one gateway to knowledge and concepts. *Knowledge Organization* 24 (2), 84-91.

Lancaster, F.W. (1986). *Vocabulary control for information retrieval*. Arlington, VA: Information Resources Press.

Lopez-Huertas, Maria (1997). Thesaurus structure design: A conceptual approach for improved interaction. *Journal of Documentation* 53(2): 139-177.

Miller, Uri (1997). Thesaurus construction: problems and their roots. *Information Processing & Management* 33, 481-493.

Purat, Jacek (1998). The world of multilingual environmental thesauri. Available at:  
[http://www.sims.berkeley.edu/~purat/world\\_multilingual\\_environmental\\_thesauri.htm](http://www.sims.berkeley.edu/~purat/world_multilingual_environmental_thesauri.htm)

Spiteri, Louise (1999). The essential elements of faceted thesauri. *Cataloging & Classification Quarterly* 28 (4), 31-52.

Willpower Information. Thesaurus principles and practice.  
<http://www.willpower.demon.co.uk/thesprin.htm>

### **Online thesauri**

Thesaurus for Graphic Materials-Subject Terms (TGM-I)  
<http://www.loc.gov/rr/print/tgm1/>

Medical Subject Headings (MeSH)  
<http://www.nlm.nih.gov/mesh/2002>

Getty Art & Architecture Thesaurus  
<http://www.getty.edu/research/tools/vocabulary/aat/>

Unified Medical Language System (UMLS).  
<http://www.nlm.nih.gov/research/umls/>

Thesaurus of ERIC Descriptors. Education Resources Information Center. <http://www.eric.ed.gov/>

### **Thesaurus construction standards**

ANSI/NISO Z39.19 (2005). National Information Standards Organization. Guidelines for the Construction, Format, and Management of Monolingual Thesauri (Z39.19). NISO Press.

ISO 2788 (1986). International Organization for Standardization. Guidelines for the establishment and development of monolingual Thesauri.

ISO 5964 (1985): International Organization for Standardization. Guidelines for the establishment and development of multilingual Thesauri.



## Evaluation Form

### Principles of Controlled Vocabulary and Thesaurus Design

Your evaluation of this workshop is very important to the future development of this course and other similar courses. Your honest, candid answers to the following questions will assist us in providing quality programs.

Please rate the following aspects of today's workshop by checking the box that best reflects your evaluation:

1. The overall content of the workshop:

	5	4	3	2	1	
a. was extremely valuable						was of little value
b. provided enough detail						was too general
c. was current & relevant						was outdated
d. was cohesive & logical						was fragmented/difficult to follow
e. was appropriate to my needs						was not at all appropriate
f. met its stated objectives						did not meet objectives

2. Presenter:

	5	4	3	2	1	
a. was knowledgeable						was unsure of the material
b. had good presentation skills						had poor presentation skills
c. encouraged participation						discouraged participation
d. addressed my level of understanding						did not consider my level
e. answered questions directly						did not answer questions
f. was prepared						was not prepared
g. understood the audience dynamics						ignored audience dynamics

3. Presenter:

	5	4	3	2	1	
a. was knowledgeable						was unsure of the material
b. had good presentation skills						had poor presentation skills
c. encouraged participation						discouraged participation
d. addressed my level of understanding						did not consider my level
e. answered questions directly						did not answer questions
f. was prepared						was not prepared
g. understood the audience dynamics						ignored audience dynamics

4. The handouts:

	5	4	3	2	1	
a. are excellent						are poor
b. followed course content						are disjointed/out of sequence
c. are valuable for future reference						are of no value

5. The PowerPoint slides:

- a. were clear and easy to read
- b. were well organized
- c. illustrated concepts clearly
- d. covered an appropriate amount of information
- e. were visually effective
- f. were enhanced by and supported the presenter's remarks

5	4	3	2	1

- were hard to read
- were poorly organized
- were confusing
- contained too much or not enough information
- were not effective
- were poorly related to the presenter's remarks

**Please give the following information about yourself:**

- 6. Your level of knowledge in the subject of this workshop before today: expert 5 4 3 2 1 novice
- 7. Your level of experience in the subject of this workshop before today: very experienced 5 4 3 2 1 beginner
- 8. Other comments:

**Comments on specific sessions:**

THANK YOU!