

博物館Metadata III： 知識組織系統

控制詞彙概念：知識組織系統
權威控制專論(Getty)
索引典專論(AAT)
知識本體(Ontology)

控制詞彙的概念

- 以控制詞彙作為知識庫
- 以控制詞彙改善文物的檢索
 - 作為查詢輔助
 - 協助影像資料庫
 - 協助圖書館目錄
 - 協助檔案描述
 - 協助博物館藏品記錄
 - 協助索引
 - 協助瀏覽

知識組織系統之概觀

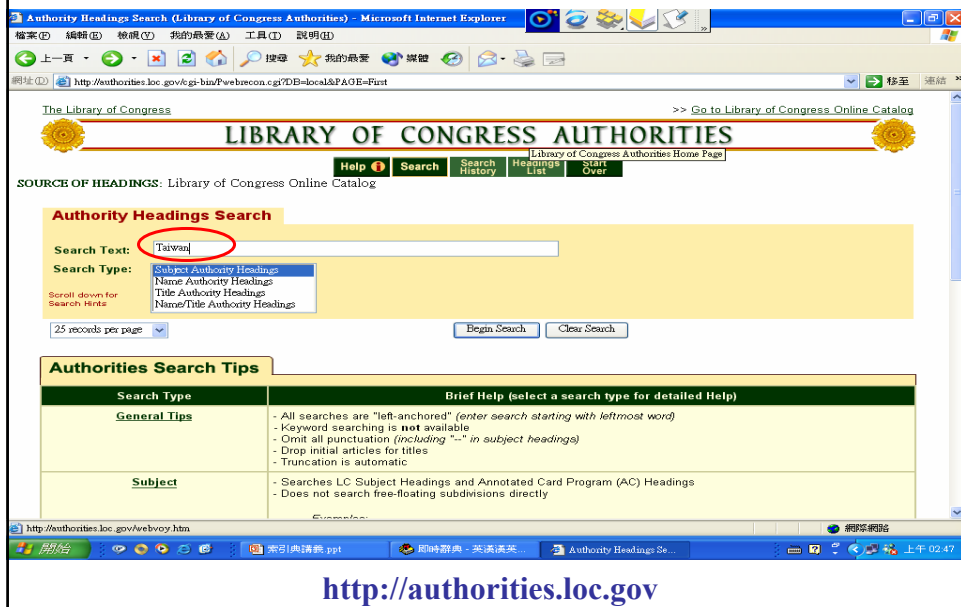
- 用詞清單(Term Lists)
 - 權威檔(Authority files)
 - 詞彙表(Glossaries)
 - 字典(Dictionaries)
 - 地名詞典(Gazetteers)
- 分類與類目(Classification and Categories)
 - 主題標目(Subject headings)
 - 分類表或類目表(Classification Schemes, Taxonomies, and Categorization Schemes)
- 關聯性清單(Relationship Lists)
 - 索引典(Thesauri)
 - 語意網路(Semantic Networks)
 - 知識本體(Ontologies)

權威檔(Authority Files)

- *Authority Files.*

Authority files are **lists of terms** that are used to **control the variant names for an entity or the domain value for a particular field**. Examples include **names for countries, individuals, and organizations**. **Nonpreferred terms may be linked to the preferred versions**. This type of KOS generally does **not include a deep organization or complex structure**. The presentation may be alphabetical or organized by a shallow classification scheme. A limited hierarchy may be applied to allow for simple navigation, particularly when the authority file is being accessed manually or is extremely large. Examples of authority files include the Library of Congress Name Authority File and the Getty Geographic Authority File.

權威檔(Authority Files)



Authority Headings Search (Library of Congress Authorities) - Microsoft Internet Explorer

http://authorities.loc.gov/cgi-bin/Pwebrecon.cgi?DB=local&PAGE=First

The Library of Congress >> Go to Library of Congress Online Catalog

LIBRARY OF CONGRESS AUTHORITIES

Help Search Search History Headings List Start Over

SOURCE OF HEADINGS: Library of Congress Online Catalog

Authority Headings Search

Search Text: Taiwan

Search Type: Subject Authority Headings
Name Authority Headings
Title Authority Headings
Name/Title Authority Headings

25 records per page

Begin Search Clear Search

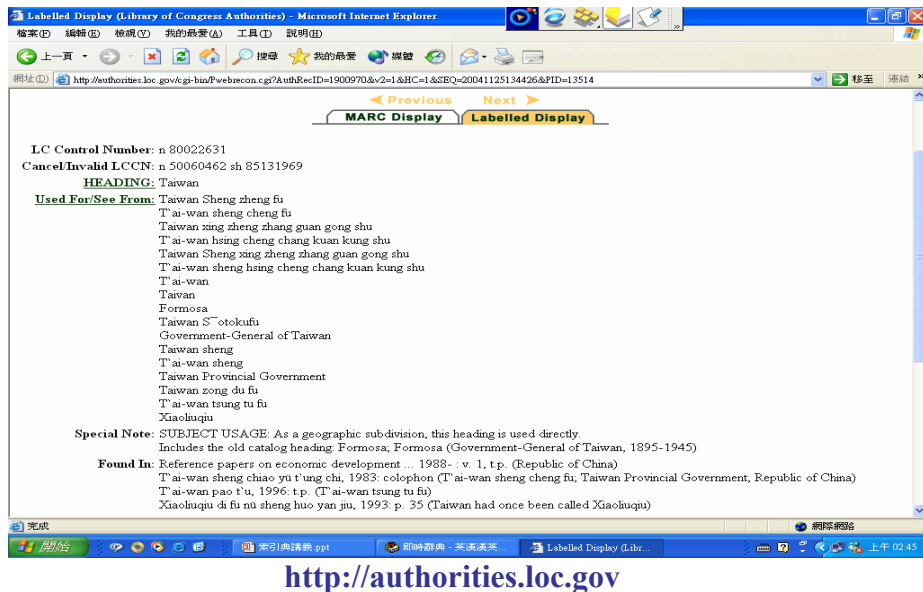
Authorities Search Tips

Search Type	Brief Help (select a search type for detailed Help)
General Tips	<ul style="list-style-type: none">- All searches are "left-anchored" (enter search starting with leftmost word)- Keyword searching is not available- Omit all punctuation (including "-" in subject headings)- Drop initial articles for titles- Truncation is automatic
Subject	<ul style="list-style-type: none">- Searches LC Subject Headings and Annotated Card Program (AC) Headings- Does not search free-floating subdivisions directly

http://authorities.loc.gov/vb/vb.htm

http://authorities.loc.gov

權威檔(Authority Files)



Labelled Display (Library of Congress Authorities) - Microsoft Internet Explorer

http://authorities.loc.gov/cgi-bin/Pwebrecon.cgi?AuthRecID=1900970&v=1&HC=1&SEQ=20041125134426&PID=13514

Previous Next

MARC Display Labelled Display

LC Control Number: n 80022631

Cancel/Invalid LCCN: n 50060462 sh 85131969

HEADING: Taiwan

Used For/See From: Taiwan Sheng zheng fu
T' ai-wan sheng cheng fu
Taiwan xing zheng zhang guan gong shu
T' ai-wan hsing cheng chang kuan kung shu
Taiwan Sheng xing zheng zhang guan gong shu
T' ai-wan sheng hsing cheng chang kuan kung shu
T' ai-wan
Taiwan
Formosa
Taiwan S' otokufu
Government-General of Taiwan
Taiwan sheng
T' ai-wan sheng
Taiwan Provincial Government
Taiwan zong du fu
T' ai-wan tsung tu fu
Xiaoliuqu

Special Note: SUBJECT USAGE: As a geographic subdivision, this heading is used directly.
Includes the old catalog heading: Formosa, Formosa (Government-General of Taiwan, 1895-1945)

Found In: Reference papers on economic development ... 1988- : v. 1, t.p. (Republic of China)
T' ai-wan sheng chiao yu t' ung chi, 1983: colophon (T' ai-wan sheng cheng fu, Taiwan Provincial Government, Republic of China)
T' ai-wan pao t' u, 1996: t.p. (T' ai-wan tsung tu fu)
Xiaoliuqu di fu nu sheng huo yan ju, 1993: p. 35 (Taiwan had once been called Xiaoliuqu)

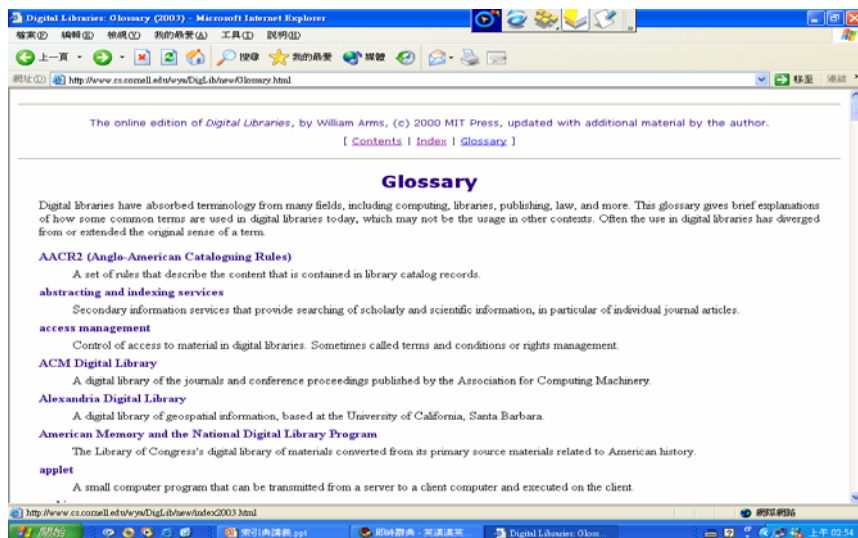
http://authorities.loc.gov

詞彙表(Glossaries)

- *Glossaries*

A glossary is a **list of terms, usually with definitions**. The terms may be from a **specific subject field** or from a particular work. The terms are defined **within a specific environment** and rarely include variant meanings. Examples include the Environmental Protection Agency (EPA) Terms of the Environment.

詞彙表(Glossaries)

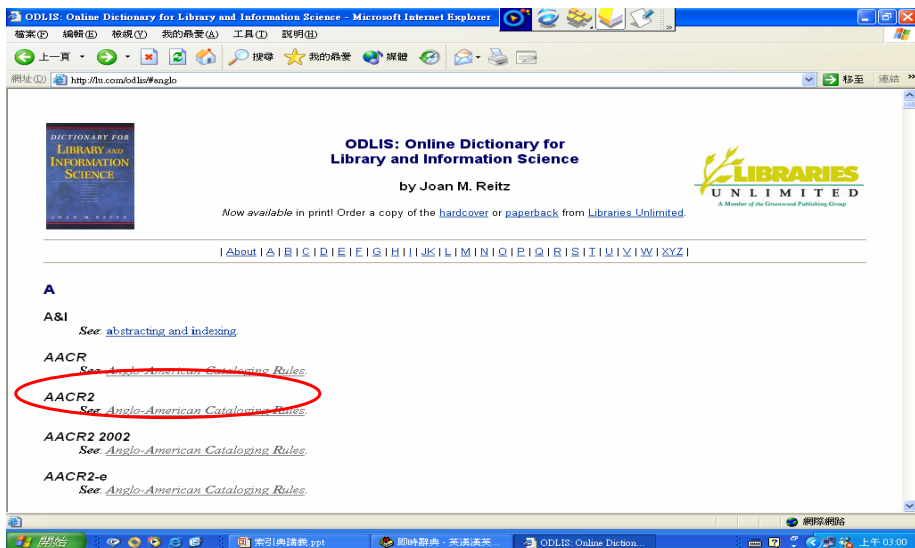


<http://www.cs.cornell.edu/wya/DigLib/new/Glossary.html>

字典(Dictionary)

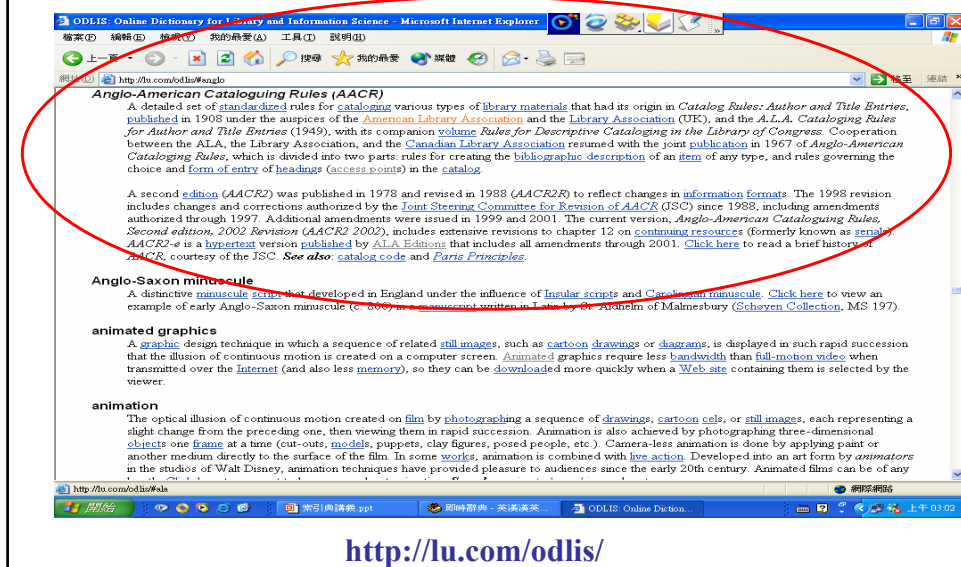
- A **dictionary** is a listing of words and phrases giving information such as spelling, morphology and part of speech, senses, definitions, usage, origin, and equivalents in other languages (bi- or multilingual dictionary)

字典(Dictionary)



<http://lu.com/odlis/>

字典(Dictionary)



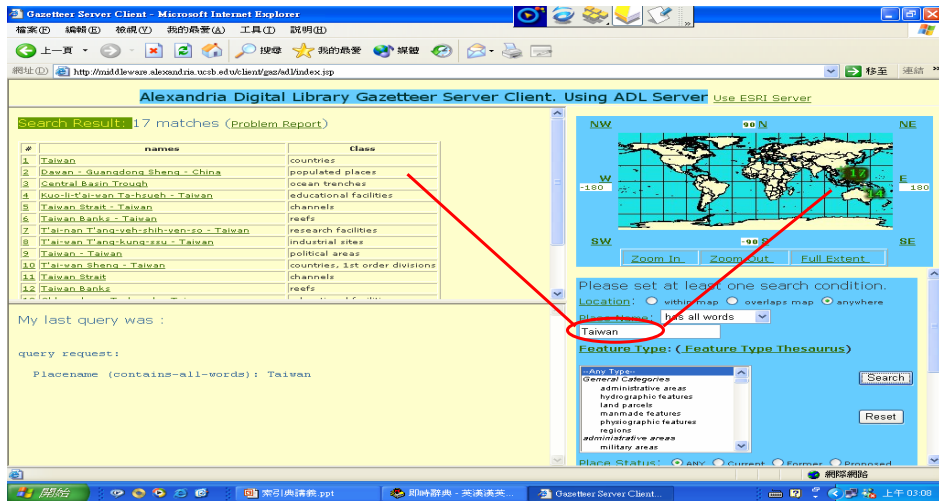
地名詞典(Gazetteers)

- **Gazetteers**

A gazetteer is a **list of place names**.

Traditional gazetteers have been published as books or have appeared as indexes to atlases. Each entry may also be **identified by feature type, such as river, city, or school**. An example is the U.S. Code of Geographic Names. Geospatially referenced gazetteers provide coordinates for locating the place on the earth's surface.

地名詞典(Gazetteers)



<http://middleware.alexandria.ucsb.edu/client/gaz/adl/index.jsp>

主題標目(Subject Headings)

- *Subject Headings*

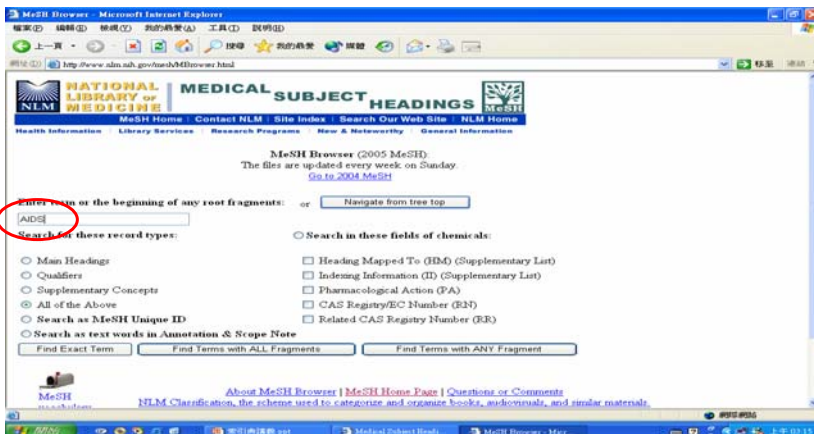
This scheme type provides **a set of controlled terms** to **represent the subjects of items in a collection**. Subject heading lists can be extensive and cover a broad range of subjects; however, the subject heading list's structure is generally very **shallow**, with a **limited hierarchical structure**.

主題標目 (Subject Headings)



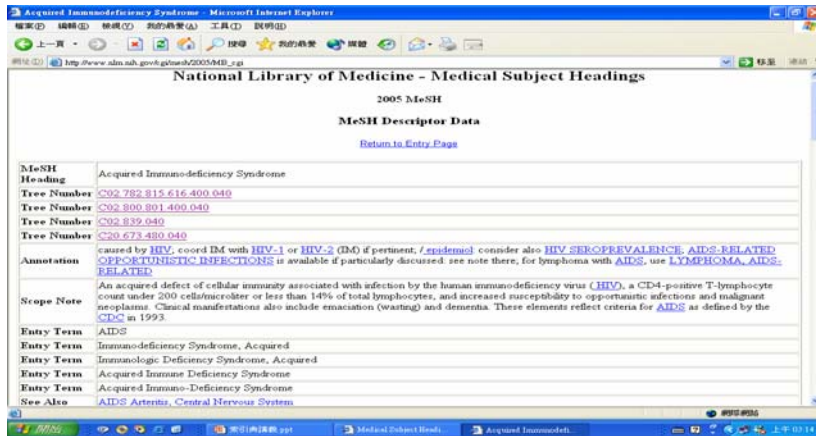
<http://www.nlm.nih.gov/mesh/meshhome.html>

主題標目 (Subject Headings)



<http://www.nlm.nih.gov/mesh/meshhome.html>

主題標目 (Subject Headings)



<http://www.nlm.nih.gov/mesh/meshhome.html>

分類表 (Classification Schemes)

- *Classification Schemes, Taxonomies, and Categorization Schemes.*

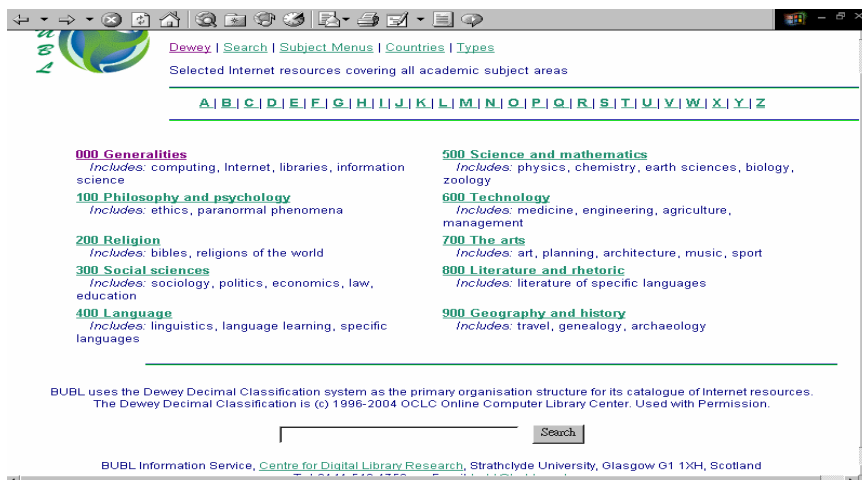
These terms are often **used interchangeably**. Although there may be subtle differences from example to example, these types of KOSs all provide ways to separate entities into "buckets" or broad topic levels. Some examples **provide a hierarchical arrangement of numeric or alphabetic notation to represent broad topics**. These types of KOSs may **not follow the rules** for hierarchy required in the ANSI NISO Thesaurus Standard (Z39.19) (NISO 1998), and they **lack the explicit relationships presented in a thesaurus**. Examples of classification schemes include the Library of Congress Classification Schedules (an open, expandable system), the Dewey Decimal Classification (a closed system of 10 numeric sections with decimal extensions), and the Universal Decimal Classification (based on Dewey but extended to include facets,

分類表(Classification Schemes)

- *Classification Schemes, Taxonomies, and Categorization Schemes.*

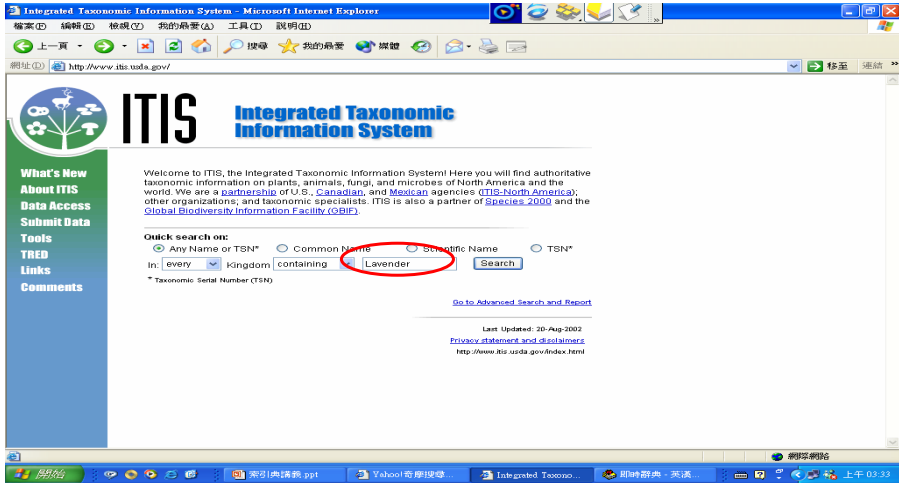
Subject categories are often used to group thesaurus terms in broad topic sets that lie outside the hierarchical scheme of the thesaurus. **Taxonomies** are increasingly being used in object-oriented design and knowledge management systems to indicate any grouping of objects based on a particular characteristic. (The science of naming things is called **taxonomy**)

分類表(Classification Schemes)



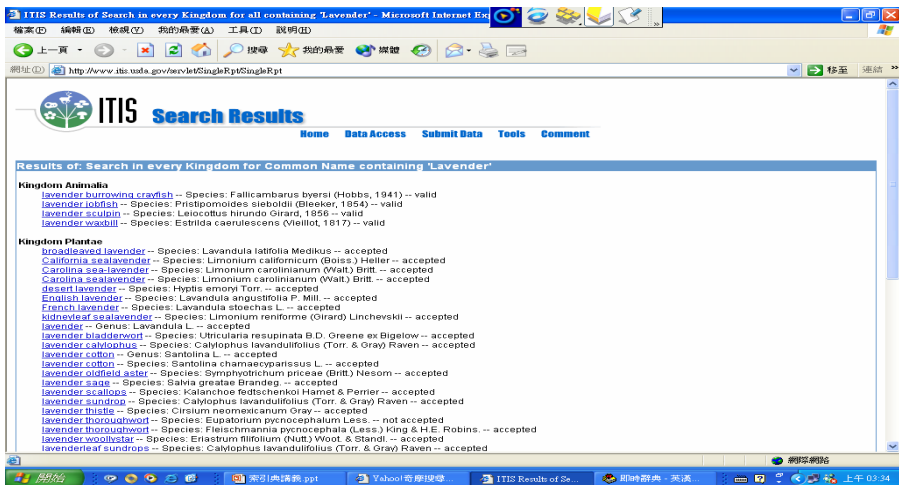
<http://www.bubl.ac.uk/link/index.html>

分類(Taxonomies)



<http://www.itis.usda.gov/>

分類(Taxonomies)



<http://www.itis.usda.gov/>

索引典(Thesauri)

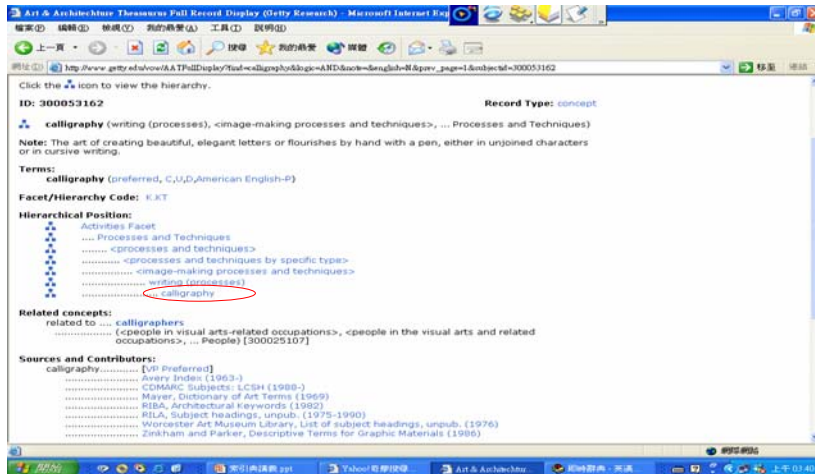
- A **thesaurus** is a structure that manages the complexities of terminology and provides conceptual relationships, ideally through an embedded classification/ontology.
- A thesaurus may specify descriptors authorized for indexing and searching. These descriptors form a **controlled vocabulary (authority list, index language)**.
- A **monolingual thesaurus** has terms from one language, a **multilingual thesaurus** from two or more language.

索引典(Thesauri)



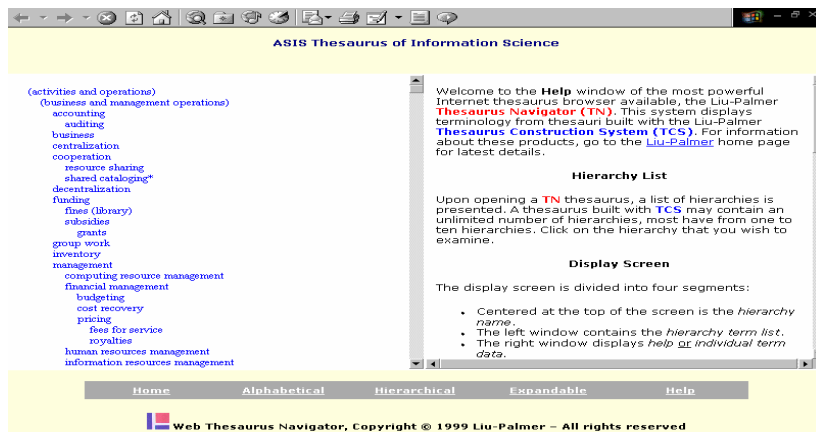
http://www.getty.edu/research/conducting_research/vocabularies/aat/index.html

索引典(Thesauri)



http://www.getty.edu/research/conducting_research/vocabularies/aat/index.html

索引典(Thesauri)



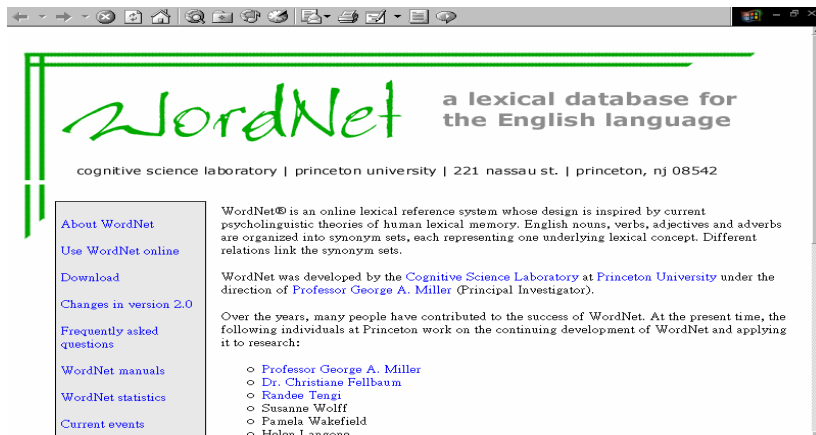
<http://www.asis.org/Publications/Thesaurus/isframe.htm>

語意網路(Semantic Network)

- *Semantic Networks.*

With the advent of natural language processing, there have been significant developments in semantic networks. These KOSs **structure concepts and terms not as hierarchies but as a network or a web.** Concepts are thought of as **nodes**, and **relationships branch** out from them. The relationships generally go beyond the standard BT, NT, and RT. They may include specific whole-part, cause-effect, or parent-child relationships. The most noted semantic network is Princeton University's WordNet, which is now used in a variety of search

語意網路(Semantic Network)



<http://www.cogsci.princeton.edu/%7Ewn/>

語意網路(Semantic Network)

WordNet 2.0 Search

Search word: Find senses

Overview for "washington"

The **noun** "washington" has 5 senses in WordNet.

1. Washington, Washington D.C., American capital, capital of the United States -- (the capital of the United States in the District of Columbia and a tourist mecca; George Washington commissioned Charles l'Enfant to lay out the city in 1791)
2. Washington, Evergreen State, WA -- (a state in northwestern United States on the Pacific)
3. Capitol, Washington -- (the federal government of the United States)
4. Washington, George Washington, President Washington -- (1st President of the United States; commander-in-chief of the Continental Army during the American Revolution (1732-1799))
5. Washington, Booker T. Washington, Booker Taliaferro Washington -- (United States educator who was born a slave but became educated and founded a college at Tuskegee in Alabama (1856-1915))

Search for of senses

☒ Show glosses
☐ Show contextual help

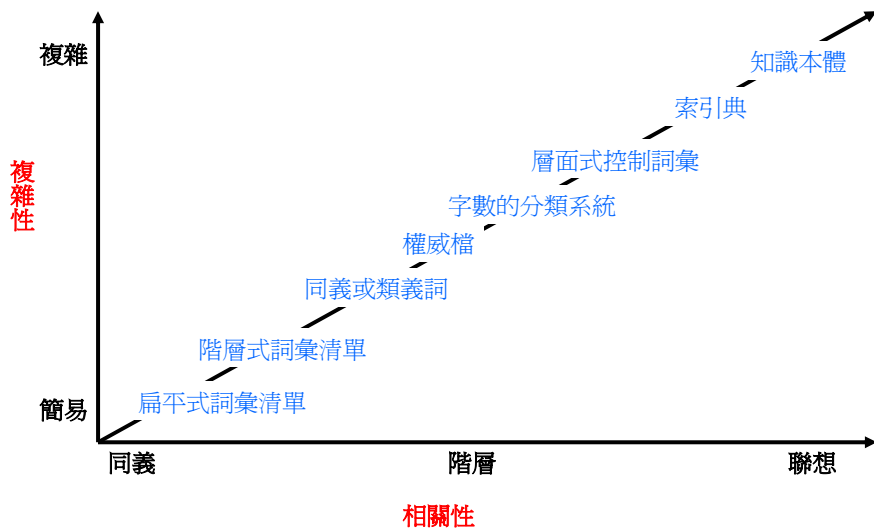
<http://www.cogsci.princeton.edu/%7Ewn/>

知識本體(ontology)

- *Ontologies.*

Ontology is the newest label to be attached to some knowledge organization systems. The knowledge-management community is developing ontologies as specific concept models. They can **represent complex relationships among objects, and include the rules and axioms missing from semantic networks.** Ontologies that **describe knowledge in a specific area** are often connected with systems for data mining and knowledge management.

控制詞彙的類型



權威控制專論

Getty人名，地名，概念，主題
權威檔

創作者權威檔

- 定義：
 - 有關負責設計或產生藝術作品或建築的藝術家，建築師及其他個人或團體之資訊。
 - 同時也作為儲存對作品重要的顧客或團體之資訊
- 結構：

姓名(核心元素)		國籍／文化／種族(核心元素)	
異名		國籍／公民資格	
日期／地點(核心元素)		文化	
		種族／民族	
出生日期 (核心元素)		性別	
死亡日期 (核心元素)		生命角色(核心元素)	
活動日期一起 (核心元素)		相關人物	
活動日期一迄		關係類型	
出生地		姓名	
死亡地		備註	
活動地		引用文獻	

地理名稱權威檔

- 定義：
 - 對於藝術作品，建築或創作者重要的地理名稱之資訊
 - 包括行政管理實體，如：國家或城市；及實質特徵，如：河流或陸地
- 結構：

地點名稱	
其他地點名稱	
時間	
	起
	迄
座標	
地點類型	
相關地點	
	關係類型
	名稱
備註	
引用文獻	

概念權威檔

- 定義：
 - 對於藝術作品，建築或創作者重要的一般性概念
 - 包括文物類型，質材，相關的活動，風格，及其他的屬性或該藝術家的角色或地名
- 結構：

描述語	
非描述語	
時間	起
	迄
相關一般概念	
	關係類型
	名稱／描述語
備註	
引用文獻	

主題權威檔

- 定義：
 - 藝術或建築作品內所描述的主題之資訊
 - 包括文學，神話或宗教的角色之名稱，或論文，事件，動物或任何其他圖像主題
- 結構：

主題名稱	
其他主題名稱	
時間	
	起
	迄
標引詞彙	
相關主題	
	關係類型
	名稱
備註	
引用文獻	

索引|典專論

索引|典與Metadata的關係
索引|典的基本原理
實例解析：AAT

索引典與Metadata的關係

- Metadata
 - 資料結構：Metadata格式
 - 如：Dublin Core, EAD, MARC, LOM…等
 - 資料內容
 - 資料值：資料結構內每個元素或欄位的值
 - 如：free text, 關鍵字, 自然語言, 控制詞彙

或稱vocabulary control詞彙控制
索引典，是詞彙控制的一種方式

索引典的基本原理

- 通則(概念的類型)

- 具體實體(concrete entities)

- 東西與其物質上的部分(如：鳥、微縮複製品、山、油畫、小熊)
 - 材料/原料(如：膠黏劑、氧氣、繪畫顏料、水)

- 抽象實體(abstract entities)

- 活動或製作過程(如：酸化、油畫、水彩畫、養育、縫紉)
 - 抽象實體，及東西、材料與活動的屬性及特性(如：人的意識、人格特質、[材料的]構造與質地、[事件]生日、南北戰爭、革命)
 - 訓練、學科或科學(如：人類學、資訊科學、藝術史)
 - 測量的單位(如：赫茲[頻頻單位]、公里)

- 個別實體(individual entities)

- 具獨特的實體，多為專有名稱。(如：亞里斯多德、地球、7月4日、10月10日、921地震、世界衛生組織、台灣)

索引典的基本原理

- 索引典的基本關聯屬性

- 等同關係(The equivalence relationship)
 - 層級關係(The hierarchical relationship)
 - 聯想關係(The associative relationship)

關係類型	關係指標	縮寫符號
等同關係 (同義詞)	USE	None or U
層級關係	USED FOR	UF
	BROADER TERM	BT
	NARROWER TERM	NT
聯想關係	RELATED TERM	RT

等同關係

- 參照符號
 - USE
 - UF
- 包括二類型的詞
 - 同義字
 - 不同語言來源的詞
 - 常用名稱與科學名稱
 - 普通名詞與商業名詞
 - 新概念的不同名稱
 - 流行或偏好的詞 vs. 過時或不贊成的詞
 - 不同的寫法
 - 縮寫與全稱
 - 半同義字



A = B

等同關係

- USE
- UF
 - Aves USE **birds** ; **birds** UF Aves
 - Outline USE **shape** ; **shape** UF outline
- 等同關係的類型(同義字)
 - a) 不同語言來源的詞
 - Examples:*
 - cats / felines(貓科動物)
 - freedom / liberty
 - sodium(鈉)/ natrium(【化】鈉)
 - sweat(汗水—緊張) / perspiration (汗水—賣力)

等同關係

- 等同關係的類型(同義字)

b) 常用名稱與科學名稱

Examples:

aspirin / acetylsalicylic acid
gulls / Laridae
salt / sodium chloride

c) 普通名詞與商業名詞

Examples:

petroleum jelly(石油凝膠) / Vaseline(凡士林)
Photocopies / Xeroxes
refrigerators / Frigidaires
tissues / Kleenex
發酵乳類品 / 養樂多

d) 新概念的不同名稱

Example:

Hovercraft (氣墊船) / air cushion vehicles(運載工具)
後設資料 / 詮釋資料

等同關係

- 等同關係的類型(同義字)

e) 流行或偏好的詞 vs. 過時或不贊成的詞

Examples:

Poliomyelitis(小兒麻痺症) / infantile paralysis(兒童的癱瘓)
developing countries / underdeveloped countries

f) 通俗名詞與俚語或行話

Examples:

Helicopters(直升飛機) / whirlybirds (直升飛機)
Psychiatrists(精神病醫師) / shrinks(美)(俚語) (精神科醫師)

g) 方言的變體

Examples:

elevators / lifts
subways / undergrounds

等同關係

- 等同關係的類型(半同義字)

Examples:

Wetness / dryness

Smoothness / roughness

– Generic Posting

Examples:

waxes plant waxes

UF plant waxes USE **waxes**

furniture

UF beds beds USE **furniture**

UF chairs chairs USE **furniture**

UF desks desks USE **furniture**

UF tables tables USE **furniture**

層級關係

- 參照符號

- BT

- NT

- 包括四類型

- 屬種關係

- 整部關係

- 身/主體的全部與器官

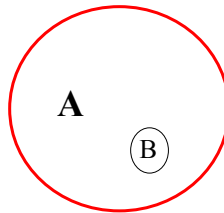
- 地理位置

- 論述的學科領域

- 層級的社會結構

- 實例關係

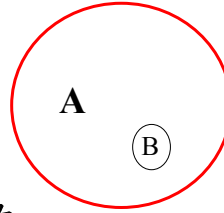
- 多層級關係



層級關係

- 參照符號

- BT (Broader Term)
- NT (Narrower Term)



- 例子1

哺乳動物 脊椎動物
BT 脊椎動物 NT 哺乳動物

- 例子2

- 解剖學 **VS.** 中樞神經系統
 - 中樞神經系統 **VS.** 大腦
- 身體一部分的研究領域

層級關係

- 屬種關係



- 代碼符號

BTG = Broader term (generic)

NTG = Narrower term (generic)

例子

老鼠 齧齒目動物
BTG 齧齒目動物 **NTG** 老鼠

層級關係

- 整部關係
 - systems and organs of the body
 - Example:*
 - nervous system
 - central nervous system
 - brain
 - spinal cord
 - Geographic locations
 - Example:*
 - Canada
 - Ontario
 - Ottawa
 - Toronto

層級關係

- 整部關係
 - Disciplines or fields of discourse
 - Example:*
 - science
 - biology
 - botany
 - zoology
 - Hierarchical organizational, corporate, social, or political structures
 - Example:*
 - countries
 - states/provinces
 - cities

層級關係

- 整部關係的代碼符號
 - BTP = Broader term (partitive)
 - NTP = Narrower term (partitive)

Example:

Central nervous system	nervous system
BTP nervous system	NTP central nervous system

層級關係

- 實例關係

Example:

mountain regions — class —	state capitals
Alps — instances —	Albany 美國紐約州的首府
Himalayas	Trenton 美國新澤西州的首府

- 實例關係的代碼符號
 - BTI = Broader term (instance)
 - NTP = Narrower term (instance)

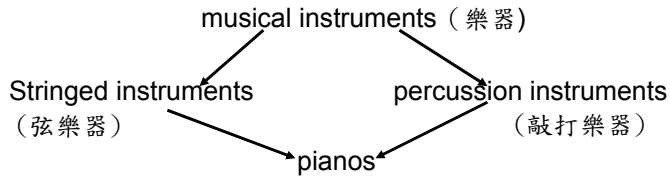
Example:

Fairy tales
NTI Cinderella

層級關係

- 多層級關係

Example:



- 多層級關係的節點標示(node labels)

Example:

cars

by purpose

racing cars

sports cars

聯想關係

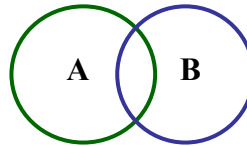
- 參照符號

– RT

- 包括二 類型

– 相同範疇

– 不同範疇



- A discipline or field of study and the objects or phenomena studied
- An operation or process and its agent or instrument
- An action and the product of the action
- An action and its patient
- Concepts related to their properties
- Concepts related to their origins
- Concepts linked by causal dependence
- A thing and its counter agent
- A concept and its unit of measurement
- Syncategorematic phrases and their embedded nouns

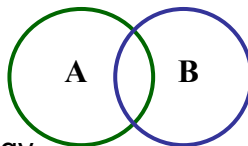
聯想關係

- 參照符號

- RT (related term)

Example:

cells	cytology
RT cytology	RT cells



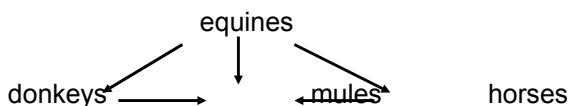
Example:(相同範疇)

boats	ships
BT vehicles	BT vehicles
RT ships	RT boats

聯想關係

- 相同範疇的聯想關係(衍生關係)

Example:(圖形展現)



Example:(字順展現)

donkeys	horses
BT equines	BT equines
RT mules	RT mules
equines	mules
NT donkeys	BT equines
NT houses	RT donkeys
NT mules	RT houses

聯想關係

- 不同範疇的聯想關係

- Disciplines or fields of study and the objects or phenomena studied, or the discipline's practitioners

Example:

mathematics	mathematicians
RT mathematicians	RT mathematics
neurology	nervous system
RT nervous system	RT neurology
botany	plants
RT plants	RT botany

聯想關係

- 不同範疇的聯想關係

- Operations or processes and their agents or instruments

Example:

temperature control	thermostats
RT thermostats	RT temperature
hunters	hunting
RT hunting	RT hunters

- An action and their products

Example:

scientific research
RT scientific inventions
publishing
RT music scores

聯想關係

- 不同範疇的聯想關係
 - An action and its patient
 - Example:*
 - data analysis
 - RT data
 - teaching
 - RT students
 - Concepts related to their properties
 - Example:*
 - liquids
 - RT surface tension
 - women
 - RT femininity

聯想關係

- 不同範疇的聯想關係
 - Concepts related to their origins
 - Example:*
 - water
 - RT water wells
 - information
 - RT information sources
 - Concepts linked by causal dependence
 - Example:*
 - injury
 - RT accidents
 - infections
 - RT pathogens

聯想關係

- 不同範疇的聯想關係
 - A thing or action and its counter agent

Example:

害蟲

RT 殺蟲劑

腐蝕

RT 腐蝕抗化劑

- A raw material and its product

Example:

(拌水泥用的)粒料

RT 混凝土

獸皮

RT 皮革製品

聯想關係

- 不同範疇的聯想關係
 - An action and a property associated with it

Example:

精確測量

RT 準確性

- A concept and its opposite

Example:

單身

RT 已婚者

寬容

RT 偏見

- 多層級關係的節點標示 (node labels)

Example: Books

RT

<operations>

binding

printing

AAT簡介

- The Art & Architecture Thesaurus (AAT)
- 藝術與建築索引典
 - 資料值標準，提供編目，索引及記錄實務方面使用．與資料結構標準(如：CDWA)及資料內容標準(如：CCO)一起搭配使用的效益最好
 - 遵循標準所建立的索引典，包括：ISO, NISO及索引典相關建置實務之規則
 - 作為索引與檢索之用．意在成為索引者與檢索者之間語言的橋樑．

AAT簡介

- 促使不同類型藏品之間的分享．如：AAT可用來描述圖書館的書籍、博物館的物品、檔案館的文書記錄，及網路上的影像等之題材
- 獨立於應用環境之外，因此，AAT可以應用在多樣化電子環境(如：資料庫，搜索引擎)，及人工索引系統(如：卡片檔)
- 一個不斷發展與成長中的工具．

AAT的詞彙範圍

- 並不只於藝術與建築類，其嘗試提供文化資產資訊的描述，記錄與檢索所需之詞彙
- 地理含概範圍方面，目前主要以西歐及北美地區的文化資產為重點，東方方面較不齊全。
- 詞彙包括的種類
 - 藝術作品(如：書畫，雕塑品，混合媒材)
 - 建築作品(如：建築物，自然環境)
 - 文化的傳統(如：事件)
 - 質材文化(如：傢俱，服裝，用具)
 - 形式與種類(如：文件類型，記錄)

AAT的詞彙範圍



個案：Mackintosh製作的椅子

這是什麼？	High-backed <i>chair</i>
它是由什麼作成的？	<i>Oak, horsehair</i>
它是如何作的？	<i>Upholstered, stained, pierced</i>
誰製作了它？	Charles Rennie Mackintosh, <i>architect</i>
它是何時被製作的？	1989-99
它是屬於什麼風格？	<i>Arts and Crafts</i>
它是什麼的一部分？	<i>Tea room</i>
它如何被使用？	<i>Dining</i>
它以什麼為題材？	<i>Anthropomorphic</i>
它從那裡來的？	Miss Cranston's Argyle Street Tea Rooms
它現在那裡？	Glasgow School of Art, Glasgow

AAT的結構

- 基礎架構是由層面(facets)，層級(hierarchies)及詞彙(terms)組合而成
- 層面結構的概念，源起於1930年代印度圖書館學家阮根納桑，提出將知識的分類歸入層面式的類目中。AAT以七個層面作為AAT的層級結構之最上層
- AAT層面的排序，由抽象至具體
- 每一個層面可再細分為數個層級，目前AAT具有32個層級。詞彙(或稱描述語)則是藉由概念展現在層級之內

AAT的結構 七個層面

- 相關概念層面 (Associated Concepts)
- 物質屬性層面 (Physical Attributes)
- 風格與時期層面 (Styles and Periods)
- 人與組織團體層面 (Agents)
- 活動層面 (Activities)
- 質材層面 (Materials)
- 物件層面 (Objects)

AAT的結構

32個層級

層面	層級	範例	層面	層級	範例
相關概念			風格與時期		
	相關概念(BM)	symmetry, folk art, patronage		風格與時期(FL)	Mombasa, Romanesque, Fluxus
層面	層級	範例	層面	層級	範例
物質屬性			人與組織團體		
	屬性與特性(DC) ATTRIBUTE AND PROPERTIES	無殼，寶石，循環的		人物 PEOPLE (HG)	裝訂師，表演藝術家，園丁
	條件與結果(DE) CONDITIONS AND EFFECTS	瑕疵，蟲害，指紋		組織團體 ORGANIZATIONS (HN)	協會，公園委員會，研究院
	設計元素 (DG) DESIGN ELEMENTS	鋸齒形雕飾，阿拉伯式花紋，小圓點			
	顏色(DL) COLOR	橄欖褐，單色（黑，白），鮮紅			

AAT的結構

32個層級

層面	層級	範例
活動		
	學科 DISCIPLINES (KD)	美學，民族考古學，都市設計
	職務 FUNCTIONS (KG)	輸入，出版，展覽
	事件 EVENTS (KM)	加冕典禮，暴風雨，朝聖
	物理的活動 PHYSICAL ACTIVITIES (KQ)	射箭運動，划船，下棋
	製程與技術 PROCESSES AND TECHNIQUES (KT)	塗清漆，壁畫，枝編與塗鴉
層面	層級	範例
質材		
	質材 MATERIALS (MT)	亞麻纖維，熟鐵，蛋彩畫法(用蛋白或膠水畫由調和顏料的一種畫法)

AAT的結構

32個層級

層面	層級	範例
物件	物件群組與系統 OBJECT GROUPINGS AND SYSTEMS (PC)	茶具、灌溉系統
	物件種類 OBJECT GENRES (PE)	贗品、自動化(技術)、供奉
	組件 COMPONENTS (PJ)	把手、前蓋、鉛字面
	定居與風景 SETTLEMENTS AND LANDSCAPES (RD)	小村莊、邊疆城市、森林
	建築設施與行政區 BUILT COMPLEXES AND DISTRICTS (RG)	大莊園、校園、歷史街區
	單一建築作品 SINGLE BUILT WORKS (RK)	大教堂、陶器廠、木條泥(草)屋
	開放空間與環境成分 OPEN SPACES AND SITE ELEMENTS (RM)	修菊(樹木)或裝飾形式的、自行車道、廣場
	傢俱 FURNISHINGS (TC)	印第安人椅子、藏比松地毯
	服裝 COSTUME (TE)	和服、帶柄望遠鏡、背心
	工具與儀器 TOOLS AND EQUIPMENT (TH)	織布機、磨刀石、調色刀

層面	層級	範例
物件	武器與彈藥 WEAPONS AND AMMUNITION (TK)	鯊魚槍、短彎刀、地雷
	測量儀器 MEASURING DEVICES (TN)	日規、經緯儀、星盤
	容器 CONTAINERS (TQ)	簍、筐、骨鐺、墨水池
	音頻儀器 SOUND DEVICES (TT)	中提琴、說話鼓、汽笛
	娛樂的工藝品 RECREATIONAL ARTIFACTS (TV)	紙牌、擲馬蹄鐵遊戲、萬花筒
	運輸工具 TRANSPORTATION VEHICLES (TX)	十五、六世紀西班牙大型帆船(galleons)、(由二馬牽引的)古代雙輪戰車 (bigas)、消防車
	視覺作品 VISUAL WORKS (VC)	名家素描(master drawings)、(簡式的)玻璃版照相
	交換媒介 EXCHANGE MEDIA (VK)	代幣、輔幣、郵票
	資訊形式 INFORMATION FORMS (VW)	導航圖、日記、電影劇本

描述語(descriptor)

- 特徵：
 - 每個描述語代表一個單一的概念(concept)
 - 可能來自任何層級
 - 可能是單一或多元文字的詞彙
 - 排列於層面(facets)中，再由層面根據概念與概念間的關係，編排其語法的角色(syntactic role)
 - 若干基本的描述語(各代表單一概念)組合後，建構成一個修正的描述語(modified descriptor)，以形成複合概念

描述語(descriptor)

範例：洛可可鍍金的雕刻木椅

Rococo carved gilded wood chairs

Facet	Descriptor no. and text		
F Styles and Periods	FL3265	Rococo	modifier
K Activities	KT911	carved	modifier
K Activities	KT139	gilded	modifier
M Materials	MT2670	wood	modifier
T Objects	TC449	chairs	focus

如此，藉由基本描述語，組成複合概念，可讓有限的基本詞彙，表達許多特別的物件描述

再深入表達，可將描述語組合成為字串：

洛可可鍍金的雕刻木椅－收藏(collecting)

文件(paper)－修復(restoration)－檔案館員 (由檔案館員修復的文件)

知識本體(Ontology)專論

為何發展本體論？
何謂本體論？
建構知識本體的方法論

為何發展本體論？

- 共享人們或軟體代理間資訊結構的共通理解
(to share common understanding of the structure of information among people or software agents)
- 促使領域知識的再度使用
(to enable reuse of domain knowledge)
- 使領域知識的假設條件更清楚而明確
(to make domain assumptions explicit)
- 區分領域知識與操作知識
(to separate domain knowledge from the operational knowledge)
- 分析領域知識
(to analyze domain knowledge)

何謂本體論？

- 正式而明確的詳細說明，論域(論述領域)內的概念(classes(有時又稱concepts))，描述此概念各種特色與屬性的每一個概念之特性(slots(有時又稱role或properties))，以及特性的限制(facets(有時又稱role restrictions))
an **ontology** is a formal explicit description of concepts in a domain of discourse (**classes** (sometime called **concepts**)), properties of each concept describing various features and attributes of the concepts (**slots** (sometimes called **roles** or **properties**)), and restrictions on slots (**facets** (sometimes called **role restrictions**)).
- 本體論伴隨一組類別的個別實例(instances)，建構出知識庫。
an ontology together with a set of individual instances of classes constitutes a **knowledge base**.

何謂本體論？

- 本體論的組成
 - 類別(概念) (concepts)(classes)
 - 特性(properties)(slots)(roles)
 - 描述概念的特徵(features) · 屬性(attributes)
 - 特性(slot)的限制(facets)(role restrictions)
- 知識庫的組成
 - 本體論
 - 個別實例(instances)

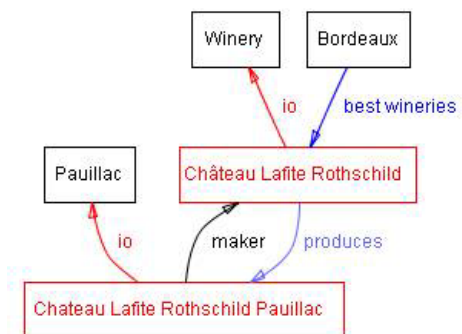
何謂本體論？

- 類別(classes)描述此領域內的概念(concepts)
 - 如：葡萄的class呈現所有的葡萄酒。特定的葡萄酒是該class的實例。你面前杯中的Bordeaux葡萄酒就是Bordeaux葡萄酒這個class的一個實例
 - 一個class可以有subclasses(下層類別)，subclasses呈現的概念比superclass(上層類別)更特定。例如，我們可以將所有葡萄酒的class區分成紅葡萄酒、白葡萄酒以及玫瑰紅的葡萄酒。另一方面，也可以將所有葡萄酒的class區分成起泡的葡萄酒與未起泡的葡萄酒

何謂本體論？

- Slots：描述classes與實例的特性(properties)
 - 如：Chateau Lafite Rothschild Pauillac葡萄酒是全醇的；該酒是由Chateau Lafite Rothschild葡萄酒廠所生產的。我們有兩個slots來描述此範例中的葡萄酒：
 - 醇度slot：有百分之百的資料值(value)
 - 製造商slot：有Chateau Lafite Rothschild葡萄酒廠的資料值。
- 在class層次，葡萄酒class的實例將具有描述他們的香味、醇度、甜度以及葡萄酒的製造商等的slots

何謂本體論？



- 葡萄酒領域內的某些類別，實例以及關係
- 黑色代表：類別
- 紅色代表：實例
- 直鏈代表：slots
- 曲鏈代表：instance-of 與 subclass-of

簡易的知識工程方法論

基本規則

- 沒有獨一方式去模型化一個領域
- 本體論的發展是一個反覆的過程
- 本體論中的觀念，必須接近關心的領域之物件(實體或邏輯)及關係

簡易的知識工程方法論

七項步驟

- 步驟一 決定本體論的領域與範圍
- 步驟二 考慮是否使用現存的本體論
- 步驟三 列舉本體論中的重要用詞
- 步驟四 界定類別與類別層級
- 步驟五 界定類別的特性 – slots
- 步驟六 界定slots的層面
- 步驟七 創造實例

簡易的知識工程方法論

步驟一 決定本體論的領域與範圍

- 回答下述基本問題，以界定此本體論模型的範圍：
 - 本體論要含概的領域為何？
 - 我們為什麼要使用本體論？
 - 本體論的資訊須回答什麼類型的問題？
 - 誰將使用與維護此本體論？
- 例：葡萄酒與食物的本體論
 - 食物與葡萄酒的展現(representation)就是本體論的領域
 - 若本體論將用於支援葡萄酒雜誌中文章的自然語言處理：此本體論的概念將需包括同義字與其詞類資訊
 - 若本體論將用於協助餐館顧客決定需要訂購那類的葡萄酒：需要包括零售價的資訊
 - 若本體論將用於酒商儲備葡萄酒於酒窖中：批發價與採購管道的流通性可能就是必要的概念。
 - 若此本體論維護者採用不同本體論使用者的語言來描述該領域：需要提供不同語言之間的比對

簡易的知識工程方法論

步驟一 決定本體論的領域與範圍

- 能勝任的問題(Competency questions)
 - 決定本體論範圍的其中一種方式，就是列出一份問題清單，該問題清單是基於此本體論上的知識基礎能夠回答的問題，也就是能勝任的問題(competency questions)
 - 例：葡萄酒與食物的領域
 - 挑選葡萄酒時，我需考慮酒的哪些特徵？
 - Bordeaux是紅葡萄酒還是白葡萄酒？
 - Cabernet Sauvignon適合跟海鮮搭配嗎？
 - 吃烤肉最好要選擇何種葡萄酒？
 - 葡萄酒的何種特徵會影響它跟一道菜的搭配？
 - 特定的葡萄酒的香味或者醇度會隨著釀造年份而改變嗎？
 - Napa Zinfandel出產的好的葡萄酒有哪些？
 - 據上述問題清單判斷，該本體論包含的資訊有：葡萄酒的各種特徵與葡萄酒的各種類型、釀造年代、優良與不佳的釀造、跟選擇搭配的葡萄酒有關的食物分類、葡萄酒與食物的建議組合。

簡易的知識工程方法論

步驟二 考慮是否使用現存的本體論

- 可先考慮與測試他人已完成的本體論，是否能夠加以精鍊與擴展現，以適用於自己特定的領域與工作。
- 若我們的系統需和已完成的特定本體論或控制詞彙的應用互動時，再度使用現存的本體論可能必需的。可再度使用的本體論，包括：
 - Ontolingua (<http://www.ksi.stanford.edu/software/ontolingua>)
 - DAML (<http://www.daml.org/ontologies/>)
 - UNSPSC(www.unspsc.org)
 - RosettaNet (www.rosettanet.org)
 - DMOZ (www.domz.org)
- 例：法國葡萄酒的知識基礎可能已經存在
 - 若輸入這些知識基礎與其為基礎的本體論，不但可以具備法國葡萄酒的分類知識，也可以初步通過區別與描述法國葡萄酒的酒性特徵之分類。酒類特性的清單可以從商業網站獲取，如顧客會考慮用來購酒的網站www.wines.com

簡易的知識工程方法論

步驟三 列舉本體論中的重要用詞

- 寫下一份我們想要陳述或想向使用者解釋的所有詞彙的清單
- 那些是我們想談論的用詞？那些用詞有那些特性？我們想要述說這些詞彙的什麼？
- 例：重要的葡萄酒相關的用詞，包括：
 - 葡萄酒、葡萄、葡萄酒廠、位置、葡萄酒的色澤、醇度、香味以及含糖量；
 - 不同類型的食物，如魚與紅肉；葡萄酒的次類型(subtypes of wine)，諸如白葡萄酒等。
- 先獲得一份完整的用詞清單，無需擔心呈現的概念間是否重疊、詞彙之間的關係、概念具有的特性，或這些概念是類別或者特性。

簡易的知識工程方法論

步驟四 界定類別與類別層級

- 三種發展類別層級的方式
 - 從上至下的發展過程(A top-down development process)
 - 先界定領域內最普遍的概念，及隨之而來的特殊化概念。如：
 - 先建立類別(classes)，代表葡萄酒與食物最普遍的概念
 - 接續建立葡萄酒的次類別(subclasses)，將葡萄酒的類別特殊化，包括：白葡萄酒、紅葡萄酒以及玫瑰紅的葡萄酒
 - 後續可進一步再將「紅葡萄酒」類別，分類為 Syrah、Red Burgundy、Cabernet Sauvignon等

簡易的知識工程方法論

步驟四 界定類別與類別層級

- 三種發展類別層級的方式
 - 由下而上的發展過程(A bottom-up development process)
 - 先界定最特殊的類別，
 - 再將這些特殊類別編組成更普遍的概念
 - 例：先界定代表Pauillac與Margaux兩種葡萄酒的類別；再建立一個共通的superclass – Medoc – 以代表這兩項類別；換言之，此superclass便是 Bordeaux的subclass。

簡易的知識工程方法論

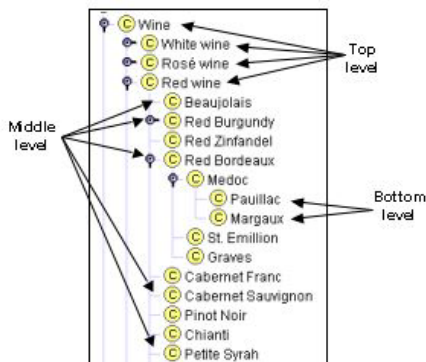
步驟四 界定類別與類別層級

- 三種發展類別層級的方式
 - 結合的發展過程(A combination development process)
 - 「從上而下」與「由下而上」兩種方法的結合
 - 先界定較顯著的概念，將之適當地普遍化與特殊化
 - 如: (1) 先從一些上層次的概念，如葡萄酒；與一些特殊的概念，如Margaux開始；(2) 隨後將它們有聯繫到一個中等層次的概念，如Medoc；(3) 產生來自法國所有區域葡萄酒的類別，也就是製作出中等層次的概念

簡易的知識工程方法論

步驟四 界定類別與類別層級

- 葡萄酒分類的不同層次



- 葡萄酒是最普遍的概念
- 紅葡萄酒、白葡萄酒以及玫瑰紅的葡萄酒是普遍的上層概念Pauillac與Margaux是層級中最特殊的類別(或底層的概念)

簡易的知識工程方法論

步驟四 界定類別與類別層級

- 上述三種方法沒有那一個是最好的
 - 不論選擇那種方式，通常皆需從界定類別(classes)開始
 - 根據步驟三所建立的清單，選擇物件已經獨立存在的用詞
 - 這些用詞將成為本體論中的類別
 - 藉由提問物件是否為一個類別的實例，將類別組成一個層級式分類表，而物件必須是(藉由定義)某個其他類別的實例
- 若類別A是類別B的superclass，那麼類別B的每一個實例也是類別A的一個實例
- 如：每種Pinot Noir 葡萄酒必然是一種紅酒。因此，Pinot Noir類別就是紅葡萄酒此類別的subclass

簡易的知識工程方法論

步驟五 界定類別的特性— slots

- 只有類別尚無法提供足夠資訊去回答步驟一提出的能勝任的問題，還必須描述概念的內在結構。
- 從步驟三所建立的用詞清單中選出類別後，剩下的多數用詞就像是這些類別的特性(properties)。如：葡萄酒的色澤、醇度、香味和糖含量，以及葡萄酒廠的位置。
- 我們必須決定清單中的每一個特性，是描述那一個類別，這些特性便成為附著於類別的slots。
- 因此，葡萄酒此類別將有以下 slots：色澤、醇度、香味以及糖份。而葡萄酒廠這個類別將有一個位置slot。

簡易的知識工程方法論

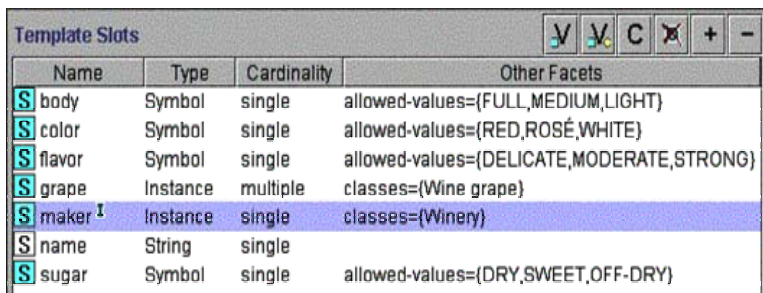
步驟五 界定類別的特性—slots

- 一般而言，有數種物件特性的類型，可以成為本體論中的slots：
 - 「**內在的**(intrinsic)」特性，如：葡萄酒的香味；
 - 「**外在的**(extrinsic)」特性，如：葡萄酒的名稱與其產地；
 - 部分**，若該物件是結構性的，這些**部分**可以是具體或抽象的「部分」(如：一餐的數道菜)
 - 與其他個體的關係**，這些關係是在該類別的個別成員與其他項目之間的關係。(如：葡萄酒的**製造商**，代表葡萄酒與葡萄酒廠的關係，也代表葡萄酒與製作該酒的葡萄之間的關係)

簡易的知識工程方法論

步驟五 界定類別的特性—slots

- 因此，除了先前已界定的特性，我們必須給葡萄酒的類別增加**slots**，包括：名稱、地區、製造商、葡萄。
- 葡萄酒此類別的**slots**:



Name	Type	Cardinality	Other Facets
S body	Symbol	single	allowed-values={FULL,MEDIUM,LIGHT}
S color	Symbol	single	allowed-values={RED,ROSÉ,WHITE}
S flavor	Symbol	single	allowed-values={DELICATE,MODERATE,STRONG}
S grape	Instance	multiple	classes={Wine grape}
S maker ¹	Instance	single	classes={Winery}
S name	String	single	
S sugar	Symbol	single	allowed-values={DRY,SWEET,OFF-DRY}

簡易的知識工程方法論

步驟五 界定類別的特性—slots

- 一個類別的所有subclasses皆繼承(inherit)該類別的所有slots。
 - 如：葡萄酒此類別所有slots將會傳承給葡萄酒的所有subclasses，包括紅葡萄酒與白葡萄酒。此外，於紅葡萄酒此類別，增加一個附加的特性，單寧酸濃度(低、中度或高度)。單寧酸濃度的slot將會由所有代表紅葡萄酒的類別所繼承(如Bordeaux與Beaujolais)。
- 一個slot需附著在具有該特性的最普遍之類別上。
 - 如：葡萄酒的醇度與色澤需附著在葡萄酒此類別上，因為此類別是擁有醇度與色澤的最普遍的類別。

簡易的知識工程方法論

步驟六 界定slots的層面

- slots可以有不同層面(facets)來描述資料值類型、容許的資料值、資料值的數量(基數)，以及slot可用的其他資料值特徵
 - 如：名稱slot的資料值(「葡萄酒的名稱」)是一個字串，因此，名稱是一個帶有字串資料值類型的 slot；生產slot(「葡萄酒廠生產這些葡萄酒」)可能有多重的資料值，且這些資料值會是葡萄酒此類別的實例
- 共通的層面，包括：
 - Slot cardinality：界定一個slot可以有多少資料值。某些系統只區分單一基數(允許最大的一個資料值)與多重基數(允許多個任何資料值)。葡萄酒的醇度是單一基數slot(葡萄酒只能擁有一個醇度)
 - Slot-value type：描述什麼資料值的類型可以填寫此slot。包括：字串(String)，數字(Number)，布林邏輯 (Boolean)，列舉式 (Enumerated)，實例類型 (Instance-type)
 - Domain and range of a slot：就實例類型的slot而言，認可的類別(allowed classes)通常稱作一個slot的範疇 (a range of a slot)

簡易的知識工程方法論

步驟七 創造實例

- 界定類別的個別實例，需要
 - 選擇一個類別
 - 建立該類別的個別實例
 - 填寫slot資料值。
- 如：建立一個個別實例Chateau-Morgon-Beaujolais來表示Beaujolais葡萄酒的特別類型，Chateau-Morgon-Beaujolais是代表所有Beaujolais葡萄酒的Beaujolais類別的一個實例，該實例具有下述所界定的slot資料值：
 - 醇度：淡(Light)
 - 色澤：紅
 - 香味：清淡
 - 單寧酸濃度：低
 - 葡萄：Gamay(釀製葡萄酒的葡萄類別的實例)
 - 製造商：Chateau-Morgon(葡萄酒廠類別的實例)
 - 產區：Beaujolais(葡萄酒產地類別的實例)
 - 糖分：不甜(Dry)

簡易的知識工程方法論

步驟七 創造實例

- Beaujolais類別的實例定義：此實例是來自Beaujolais地區的Chateaux Morgon Beaujolais，由Chateau Morgon葡萄酒廠使用Gamay葡萄釀製。此種葡萄酒的醇度淡、香味清淡、色澤紅以及單寧酸的含量低。它是一種不甜的葡萄酒。

The screenshot shows a software window titled "Chateau Morgon Beaujolais (Beaujolais)". It contains several input fields and dropdown menus for defining the instance's properties:

- Name:** A text field containing "Chateau Morgon Beaujolais".
- Area:** A dropdown menu showing "Beaujolais region".
- Body:** A dropdown menu showing "LIGHT".
- Color:** A dropdown menu showing "RED".
- Maker:** A dropdown menu showing "Chateau Morgon".
- Flavor:** A dropdown menu showing "DELICATE".
- Sugar:** A dropdown menu showing "DRY".
- Grape:** A dropdown menu showing "Gamay grape".
- Tannin Level:** A dropdown menu showing "LOW".

問題與討論