Introduction to
Controlled
Vocabularies
Featuring the Getty Vocabularies

Patricia Harpring
Managing Editor
Getty Vocabulary Program

Revised June 2023

Getty

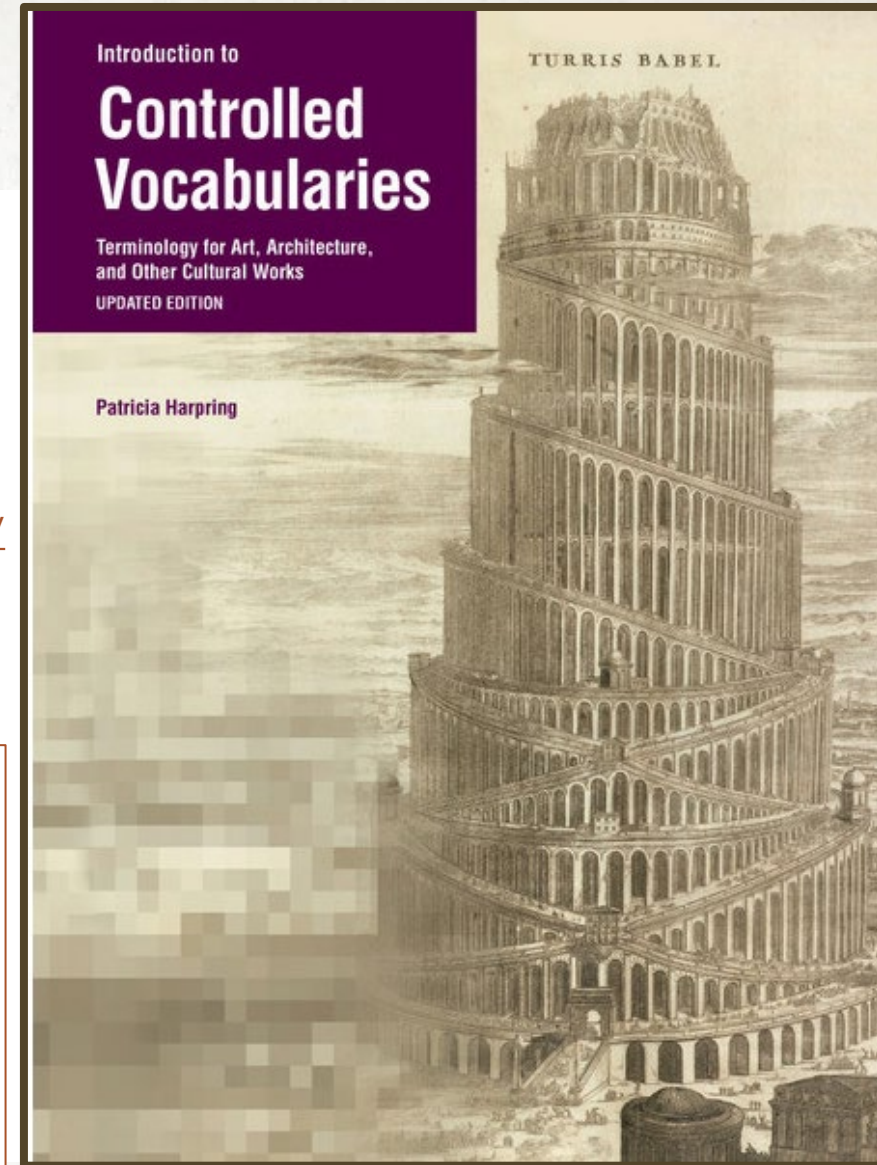# TABLE OF CONTENTS

PREFACE

- This ppt is a companion presentation to the book, *Introduction to Controlled Vocabularies*, revised edition, 2013

- For an online version of the 2010 edition of *Introduction to Controlled Vocabularies* , see www.getty.edu/research/publications/electronic_publications/intro_controlled_vocab/

- Questions? Contributions to the Getty Vocabularies? Send an email to vocab@getty.edu

Patricia Harpring
Series edited by Murtha Baca

This primer on the characteristics, scope, uses, and methods for building and maintaining controlled vocabularies for art and cultural materials explains how vocabularies should be integrated in cataloging systems; utilized for indexing and retrieval; and structured to group synonyms and arrange concepts into categories.

Introduction to
Controlled Vocabularies

Terminology for Art, Architecture, and Other Cultural Works
UPDATED EDITION

Patricia Harpring

TURRIS BABEL

CONTROLLED VOCABULARIES IN CONTEXT

VOCABULARIES FOR VISUAL ARTS

Getty

Audience

- This book deals specifically with controlled vocabularies related to <u>art works</u> — those products of human creativity that have visual aesthetic expression.

- Such vocabularies are employed with the ultimate goal of allowing art works, images of art works, and information about them to be discovered, brought together, and compared for study and appreciation

- The intended audience of this book includes students, academics, and professionals in art museums, art libraries, archives, visual resource collections, and other institutions that catalog the visual arts, architecture, and other cultural objects.

- The audience may also include systems providers who support these communities, as well as consortia or other groups attempting to compile or use vocabularies about art.

- The topics discussed here may be applicable to disciplines outside the arts as well.

What are cultural works?

- Objects representing visual arts and material culture are called **works** in this book.

- *Material culture* refers to art, architecture, and visual arts more broadly comprising the aggregate of physical objects produced by a society or culturally cohesive group.

- *Cultural works* are the physical artifacts of *cultural heritage*, which encompasses broadly the belief systems, values, philosophical systems, knowledge, behaviors, customs, arts, history, experience, languages, social relationships, institutions, and material goods and creations belonging to a group of people and transmitted from one generation to another.

- The group of people or society may be bound together by race, age, ethnicity, language, national origin, religion, or other social categories or groupings.

Getty

Creators of cultural works

- Creators of information about works include museums, visual resources collections, libraries, special collections, archival collections, private collections, and scholars

- Complexity is inherent in art information itself

- Issues surrounding the development and maintenance are further complicated by the diverse spectrum of information creators, including museum professionals, librarians, archivists, visual resource specialists, art and architectural historians, archaeologists, and conservators

- Users of the information may include all of these groups, as well as the general public. While these communities share a vast overlap of required information about works, they also have various requirements and different cataloging and indexing traditions

Getty

Types of
Standards for art information

There are several types of standards used, some used to create and others used to share art information

- **Standards for data values** provide the actual values to be entered in fields, including the vocabulary terms and allowable character sets. Controlled vocabularies are standards for data values.

- **Standards for data structure** dictate what constitutes a record. They define the names, length, repeatability, and other characteristics of fields and their relationships to each other. Examples are the MARC format and CDWA. A Resource Description Framework (RDF) is a language or format for describing things as well as the relationships between things as simple properties and values (known as 'triples'), while things are represented using URIs. Among the most often used standard formats for publishing art vocabularies are the Simple Knowledge Organization System (SKOS) and Web Ontology Language (OWL).

- **Standards for data content** indicate how data should be entered, including cataloging rules and syntax for data. They may refer to standards for data values and standards for data structure. Examples of standards for data content are *Resource Description and Access* (RDA) (replaces AACR2) and *Cataloging Cultural Objects* (CCO).

- **Standards for data releases or exchange** are a different class than the above, but important to mention. They may include rules for both structure and maybe content. The Getty Vocabularies are released in relational tables, XML, APIs, and Linked Open Data (LOD) releases, JSON, RDF, N3/Turtle, N-Triples.

Getty

WHAT ARE CONTROLLED VOCABULARIES?

THE BASICS

Getty

WHAT IS A CONTROLLED VOCABULARY?

- **An organized arrangement of words and phrases that are used to *index* content and/or to *retrieve* content through navigation or a search**

  - Controlled vocabulary typically includes preferred terms and has a limited scope or describes a specific domain

- Synonyms
  - Athens (English) = Athínaí (Greek) = Athenae (Latin)
- Concepts have relationships, provide context
  - Athens is in Greece
  - Epiktetos II is possibly identified as the Kleophrades Painter
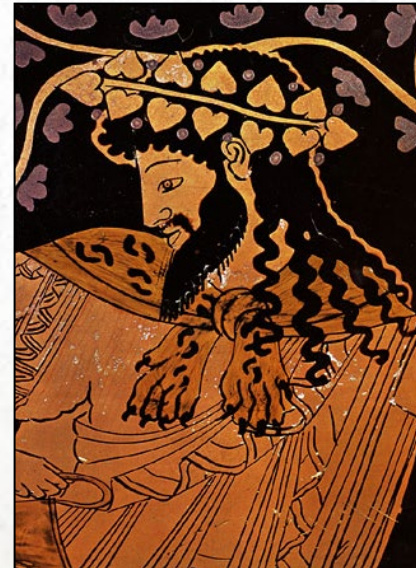
CONTROLLED VOCABULARY

- Controlled vocabulary comprises standardized words and phrases used to refer to ideas, physical characteristics, people, places, events, subject matter, and many other concepts.

- Controlled vocabularies allow for the categorization, indexing, and retrieval of information.

- The purpose of controlled vocabularies is to organize information and to provide terminology to catalog and retrieve information.

- While capturing the richness of variant terms, controlled vocabularies also promote consistency in preferred terms and the assignment of the same terms to similar content.

- One of the most important functions of a controlled vocabulary are to gather together variant terms and synonyms for concepts and to link concepts in a logical order or into categories.

- Enhancing access to visual arts and material culture information for research, education, and discovery

- Users don't always know what a person, place, or thing is called

- Even knowledgeable users or catalogers may use different terms for same person, place, or thing

- Vocabularies gather related terms and other information together to improve access to art and art history information

- In access and retrieval, controlled vocabularies are critical components for increasing reliability and success of research and discovery

Getty

CONTROLLED VOCABULARY

stained glass

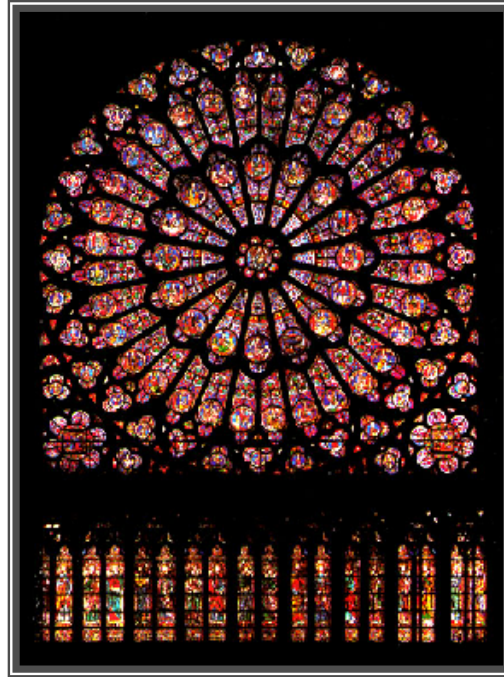pot-metal glass

rose window

Catherine wheel

lancets

Medieval

Gothic

tracery

bar tracery

- Are a *rose window* and a *Catherine wheel* the same thing?
- How *is pot-metal glass* related to the more general term *stained glass*?
- The links and relationships in a controlled vocabulary ensure that these relationships are defined and maintained, for cataloging, retrieval, and discovery.

Getty

PURPOSE OF CONTROLLED VOCABULARIES

Cataloging: Used as sources of standard terminology for use in description, cataloging, and documentation

Retrieval: Used as assistants in online search engines, creating a semantic "road map" that shows links and paths between concepts and terms; linked to other metadata as conduit for research and discovery in LOD

Research: Used as knowledge bases in themselves
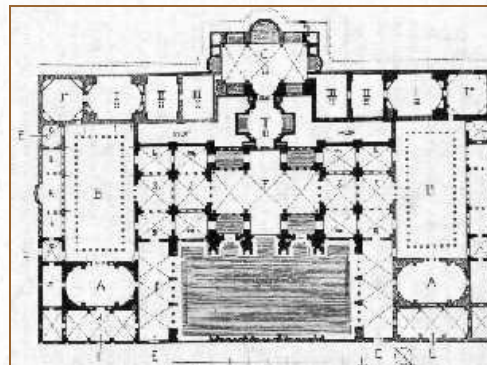
apodyteria
apodyterium
gymnasteria
gymnasterium

Scope Note:
Dressing rooms in ancient Greek and Roman baths and palaestrae.

DISPLAY AND CONTROLLED INFORMATION

- Both descriptive and administrative data must be maintained in ways that will accommodate two categories of information:
  - information intended for display to end users and information intended for retrieval.
  - Information utilized for retrieval should be adapted for controlled vocabularies and controlled format.

Why are display and indexing of information separate issues?

- Art and cultural heritage information provides unique challenges in display and retrieval of information. Information must be displayed to users in a way that allows expression of nuance, ambiguity, and uncertainty.
  - The facts are not always known or straightforward, and it is misleading and contrary to the tenets of scholarship to fail to express this uncertainty.

- At the same time, efficient retrieval requires indexing according to firm rules and controlled terminology.

DISPLAY AND CONTROLLED INFORMATION

- It is often necessary to allow fuzziness in the expression of information that at the same time must be retrievable via controlled vocabulary
- In certain key areas of a work record, this is accomplished by including separate display and indexing fields for the same information.
- Controlled information included vocabulary-controlled data ("Authority" in the example below) and controlled format (e.g., for years)

Creator Description:♦ primary painter and calligrapher was Dai Xi (Chinese, 1801-1860), with additional inscriptions and colophons added by other officials; commissioned by Wu Zhongzhun	Free text

Identity:♦ Dai Xi
   Roles:♦ draftsman
        calligrapher
Identity:♦ Wu Zhongzhun
   Role:♦ patron
	Authority
	Authority
	Authority
	Authority

Creation Date:♦ between 1847 and 1849 **Earliest:**♦ 1847 **Latest:**♦ 1849	Free text Controlled format

Image credits: Dai Xi (Chinese, 1801-1860); *Viewing the Waterfall at Lung-Ch'iu*; 1847; ink on paper; 31.9 x 133.9 cm (12 1/2 x 52 3/4 inches); The Saint Louis Art Museum (Saint Louis, Missouri, USA); 7:19853.

STANDARDS FOR CONTROLLED VOCABULARIES

- Controlled lists, synonym ring lists, taxonomies, and thesauri are covered by ISO and NISO standards

ANSI/NISO Z39.19-2005: Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies. Bethesda: National Information Standards Organization, 2005.

ISO 25964-1:2011: Information and documentation: Thesauri and interoperability with other vocabularies: Part 1: Thesauri for information retrieval. Geneva: International Organization for Standardization, 2011.

ISO 25964-2:2013: Information and documentation: Thesauri and interoperability with other vocabularies: Part 2: Interoperability with other vocabularies. Geneva: International Organization for Standardization, 2013.
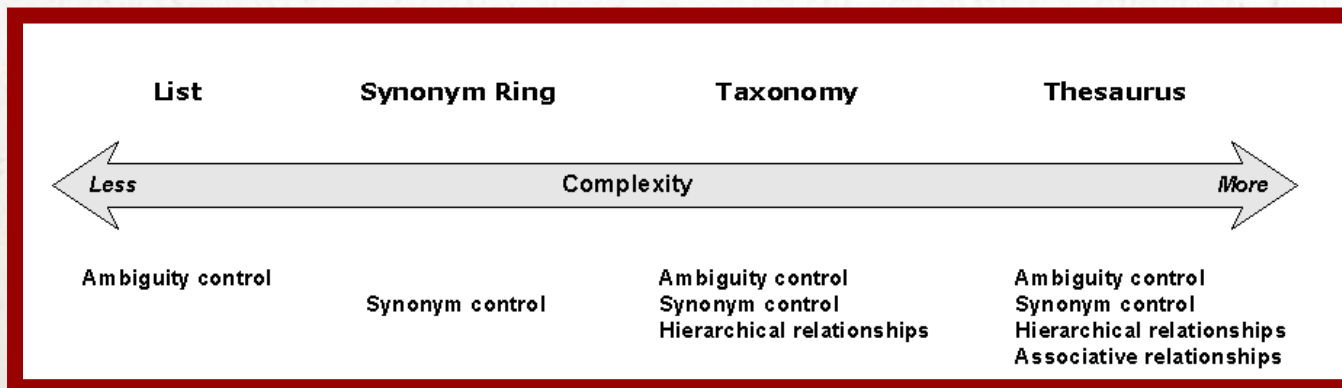
Illustration from ANSI, illustrating controlled vocabularies in terms of increasing complexity

TYPES OF CONTROLLED VOCABULARIES

- There are several types of controlled vocabularies
  and related specifications relevant for art information
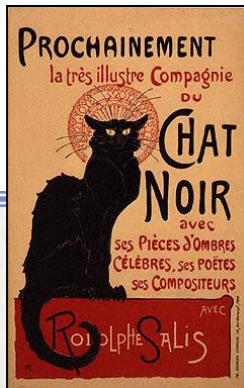
- Subject heading lists
- Controlled lists
- Synonym ring lists
- Authority Files
- Taxonomies
- Alphanumeric classifications
- Thesauri
- Ontologies
- Folksonomies

The Getty Vocabularies are thesauri

Types of vocabularies

- **Subject heading list:** uniform words or phrases assigned to books and articles (or other materials) to describe the subject or topic and to group them with materials having similar subjects

  - usually in alphabetical order
  - precoordination of terminology is a characteristic of subject headings)
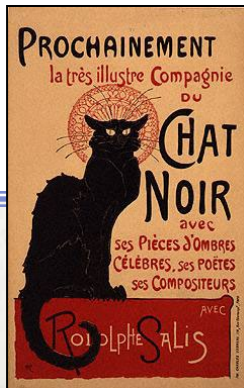
Cat family (Mammals)--Literary collections
Cat Sacred - Ancient Egypt

Types of vocabularies

- Subject heading list: uniform words or phrases assigned to books and articles (or other materials) to describe the subject or topic and to group them with materials having similar subjects

- subject headings typically combine several unique concepts together in a string as in the Library of Congress Subject Headings (LCSH)

Cat family (Mammals)--Literary collections
Cat Sacred - Ancient Egypt

Types of vocabularies

- **Controlled list:** A simple list of terms used to control terminology

- In a well-constructed controlled list: ▪each term must be unique; ▪ terms should all be members of same class; ▪ terms should not be overlapping in meaning; ▪ terms should be equal in granularity/specificity; ▪ and terms should be arranged alphabetically or in another logical order

anteaters
armadillos
bats
bears
cats
cows
deer
dogs
elephants
goldfish
horses
ostriches
pandas
pigs
zebras

Getty

Types of vocabularies

- Controlled list: A simple list of terms used to control terminology

- May include terms from other controlled vocabulary resources

- For some elements or fields in the database, a controlled list may be sufficient to control terminology, particularly where the terminology for that field is limited and unlikely to have synonyms or ancillary information

anteaters
armadillos
bats
bears
cats
cows
deer
dogs
elephants
goldfish
horses
ostriches
pandas
pigs
zebras

Getty

Controlled list: A simple list of terms used to control terminology

Example of a controlled pick list for Classification

Pick List

manuscripts
miscellaneous
paintings
photographs
sculpture
site Installation
texts
vessels

Types of vocabularies

- **Authority file:** A set of established names or headings and cross-references to the preferred form from variant or alternate forms. The Library of Congress name authority file is a well-known authority file.
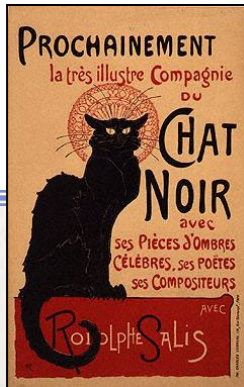
- Vocabulary types are not mutually exclusive; a given vocabulary can be an authority file and a thesaurus.

- Building local authorities is recommended by CDWA and CCO.

LC Control Number: n 79003969

HEADING: Moses, Grandma, 1860-1961

```
000  00578cz a2200193n 450
001  1418836
005  19910703055707.6
008  790117n| acannaab |a aaa
010  __  |a n 79003969
035  __  |a (DLC)n 79003969
040  __  |a DLC |c DLC |d DLC-R
100  10  |a Moses, |c Grandma, |d 1860-1961
400  00  |a Grandma Moses, |d 1860-1961
400  10  |w nna |a Moses, Anna Mary Robertson, |d 1860-1961
400  10  |a Mōzesu, |c Guranma, |d 1860-1961
670  __  |a Her Grandma Moses ... 1946.
670  __  |a Her Guranma Mōzesu ten, 1990: |b t.p. (Grandma Moses)
952  __  |a RETRO
953  __  |a xx00 |b zz00
```

Getty

Types of vocabularies

- **Synonym ring list:** A list comprising sets of terms that are considered equivalent for retrieval

- No preferred term

- Not necessarily true equivalents (= quasi-synonyms or generic postings)

- Generally used for search and retrieval, providing access to content that is represented in natural, uncontrolled language

domestic cat → Felis domesticus → cat family

house cat ← kittens
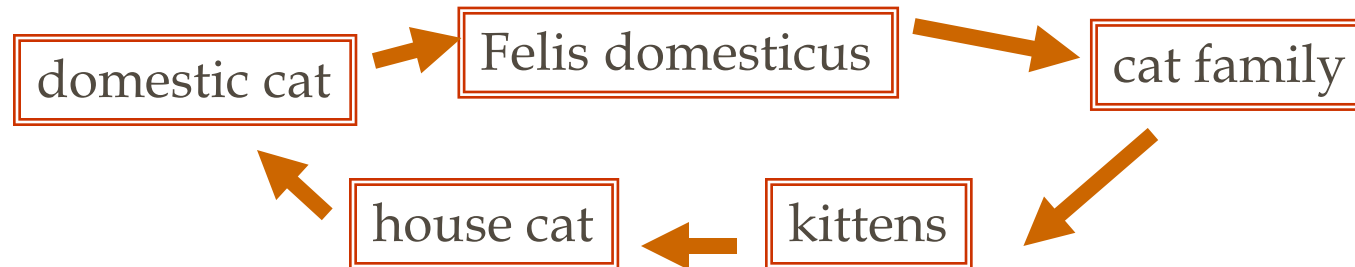
Types of vocabularies

- **Authority File:** An authority file is a set of established names or headings and cross-references to the preferred form from variant or alternate forms.
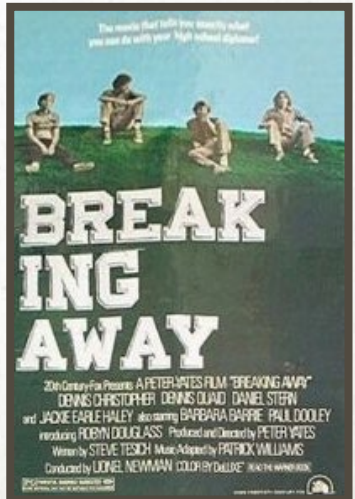
LIBRARY OF CONGRESS AUTHORITIES

Using *Library of Congress Authorities*, you can browse and view authority headings for [...] [...]hority records in [...] [...]free of charge.

LC Control Number: 90004394

HEADING: Bicycle racing United States

```
000  047nz  a22001169n 450
001  4833743
005  19901012154012.1
```

Subject headings in LC authority file

[...]on books and other items in the collections of the Library of Congress.

Types of vocabularies

- **Taxonomy:** An orderly classification for a defined domain
- Vocabularies that organize a body of knowledge into conceptual categories are classifications and taxonomies

Common name	Human	Lion	House Cat
Kingdom	Animalia	Animalia	Animalia
Phylum	Chordata	Chordata	Chordata
Class	Mammalia	Mammalia	Mammalia
Order	Primate	Carnivora	Carnivora
Family	Homonidae	Felidae	Felidae
Genus	Homo	Panthera	Felis
species	sapiens	leo	domesticus

Getty

Types of vocabularies

- **Alphanumeric classification scheme:** Controlled codes (letters or numbers, or both letters and numbers) that represent concepts, headings, or textual correlates

- They generally have an implied taxonomy that can be surmised from the codes

- The Dewey Decimal System and Iconclass are examples

Dewey #	10 Main Classes	Kinds of Books
000-099	General Works	encyclopedias, almanacs, record books, such as Guinness
100-199	Philosophy and Psychology	paranormal phenomena, such as ghosts, ethics, how we think
200-299	Religion	mythology, religions
300-399	Social Science	government, holidays, folklore, fairy tales, education, community
400-499	Language	English and foreign languages, sign language, dictionaries
500-599	Natural Science	math, chemistry, biology, weather, rocks, plants, animals in nature
600-699	Applied Science	inventions, health, drugs, transportation, cooking, pets
700-799	Fine Arts and Recreation	crafts, art, drawing, painting, music, games, TV, movies, sports
800-899	Literature	short stories, poetry, plays, jokes, riddles (fiction could be here)
900-999	History and Geography	countries, flags, historical events, biographies (92 or 920)

Iconclass

9 Classical Mythology and Ancient History
......94 the Greek heroic legends (I)
.... 94L (story of) Hercules (Heracles)
..................94L1 early life, prime youth of Hercules
..................94L2 love-affairs of Hercules
..................94L3 most important deeds of Hercules:
            the Twelve Labours

Thesaurus

- Thesaurus: A semantic network of unique concepts

- Thesauri may be monolingual or multilingual

- Thesauri may have the following three relationships:

    - Equivalence Relationships
    - Hierarchical Relationships
    - Associative Relationships

stirrup cups
coaching glasses
hunting glasses

associative

sturzbechers
Sturzbecher
stortebekers

distinguished
from

hierarchical

Objects Facet
.... Furnishings and Equipment
.... Containers
.... <culinary containers>
.... <vessels for serving / consuming food>
.... rh

rhyta
rhyton
rhytons
rhea
rheons
ritón
萊坦酒杯

equivalence

Tools related to or utilizing vocabularies

- **Ontology:** Formal, machine-readable specifications of a conceptual model

- Concepts, properties, relationships, functions, constraints, and axioms are all explicitly defined

- Not a controlled vocabulary, but uses one or more controlled vocabularies for a defined domain and expresses the vocabulary in a representative language that has a grammar for using vocabulary terms to express something meaningful

- Ontologies are used in the semantic web, artificial intelligence, software engineering, & information architecture as knowledge representation about a particular domain of knowledge.

- Ontologies generally divide their world into the following areas: individuals, classes, attributes, relations, and events

- The grammar of the ontology links these areas together by formal constraints that determine how the vocabulary terms or phrases may be used together

Folksonomy: An assemblage of concepts, which are represented by terms and names (called tags) that are compiled through social tagging.

- Social tagging refers to the decentralized practice and method by which individuals and groups create, manage, and share tags (terms, names, etc.) to annotate and categorize digital resources in an online "social" environment.
- Not a controlled vocabulary, but could use vocabularies; used for retrieval
- Social tagging = decentralized practice where individuals and groups create, manage, and share tags (terms, names, etc.) to annotate digital resources in an online "social" environment
- Typically without hierarchical structure, typically have no preferred term for a concept, and may not even cluster synonyms
- By definition not applied to indexing by professional indexers, generally characterized by non-standard, idiosyncratic tagging

Image from weblogs.elearning.ubc.ca/.../FolksonomyCast.jpg

Unique Identifiers

Unique Identifier: A number or other string that is associated with a record or piece of data, exists only once in a database, and is used to uniquely identify and disambiguate that record or piece of data from all others in the database.

Subject_ID: 300311452
Record Type: concept
Label: **cinnabar (mineral) (mineral, inorganic material, ... Materials (Hierarchy Name))**
Scope Note: A soft, dense, red, native ore composed of mercuric sulfide, found in deposits in veins near volcanic rocks or hot springs around the world. [...]
Terms: cinnabar (mineral) (preferred,C,U,English-P,D,U,N) **Term_ID: 1000403760**
           cenobrium (mineral) (C,U,English,UF,U,N) **Term_ID: 1000403756**
           natural vermilion (mineral) (C,U,English,UF,U,N) **Term_ID: 1000403762**
           cinabre (mineral) (C,U,French-P,D,U,N) **Term_ID: 1000403757**
           cinabrio (mineral) (C,U,Spanish-P,D,U,N) **Term_ID: 1000403758**
           cinabro (minerale) (C,U,Italian-P,D,U,N) **Term_ID: 1000403760**
           cinábrio (mineral) (C,U,Portuguese-P,UF,U,N) **Term_ID: 1000403774**
           cinnaber (mineraal) (C,U,Dutch-P,D,U,U) **Term_ID: 1000539139**
           Zinnober (Mineral) (C,U,German-P,D,U,N) **Term_ID: 1000403760**
           Zinnoberersatz (Mineral) (C,U,German,UF,U,N) **Term_ID: 1000403771**
           tan-sha (C,U,Chinese (transliterated)-P,D,U,N) **Term_ID: 1000403776**
           shinsha (C,U,Japanese (transliterated)-P,D,U,N) **Term_ID: 1000403775**
Hierarchical Position:
    Materials Facet **Subject_ID 300264091**
    .... Materials (Hierarchy Name) (G) **Subject_ID 300010357**
    .... materials (matter) (G) **Subject_ID 300010358**
    .... <materials by composition> (G) **Subject_ID 300212963**
    .... inorganic material (G) **Subject_ID 300010360**
    .... mineral (G) **Subject_ID 300011068**
    .... cinnabar (mineral) (G) **Subject_ID 300311452**
Related concepts: **Code 2818** source for .... cinnabar (pigment) **Subject_ID 300400883**

RELATIONSHIPS IN CONTROLLED VOCABULARIES

EQUIVALENCE, HIERARCHICAL, ASSOCIATIVE

Getty

Equivalence relationship

- **Equivalence relationships:** The relationships between synonymous terms or names for the same concept.

- A good controlled vocabulary should include terms representing different forms of speech and various languages where appropriate.

- Ideally, all terms that share an equivalence relationship are either true synonyms or lexical variants of the preferred or another term in the record.

Getty

Equivalence relationship

- **True synonyms:** Synonyms may include names or terms of different linguistic origin, scientific terms and common terms, dialectical variants, and names in different languages.

- Synonyms are names or terms for which meanings and usage are identical or nearly identical in a wide range of contexts.
- Note that true synonyms are rare in natural language.
- In many cases, different terms or names may be interchangeable in some circumstances, but they should not necessarily be combined as synonyms in a single vocabulary record

- **Near synonyms:** *Near synonyms*, also known as *quasi-synonyms*, are terms with meanings that are regarded as different, but the terms are treated as equivalents for the purposes of retrieval in the controlled vocabulary.

Ancestral Puebloan
Ancestral Pueblo
Basketmaker-Pueblo
Anasazi

Getty

Equivalence relationship

- In a vocabulary, the equivalence relationship may apply to generic terms and also to proper names
- E.g., names of people or names of places

Kahlo, Frida
Frida Kahlo
Kahlo de Rivera, Frida
Rivera, Frida
Kahlo y Calderon,
Magdalena Carmen Frida
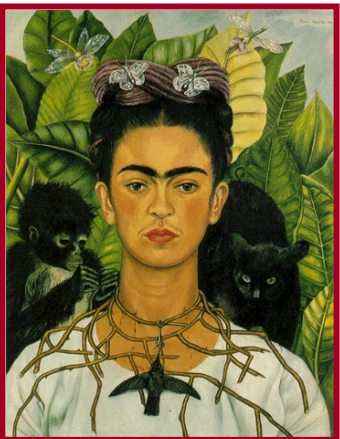
Strasbourg
Strassburg
Straßburg
Strossburi
Estrasburgo
Strateburgum
Argentoratum

sparver beds
sparver bed
sparver-beds
sparver bedsteads
sperver beds

Equivalence relationship

- **Lexical variants:** Although they are grouped with synonyms for practical purposes, lexical variants technically differ from synonyms in that synonyms are different terms for the same concept, while lexical variants are different word forms for the same expression.

- Lexical variants may result from spelling differences, grammatical variation, and abbreviations. Terms in inverted and natural order, plurals and singulars, and the use of punctuation may create lexical variants.

- In a controlled vocabulary, such terms should be linked via an equivalence relationship.

dimidiating rhyta
dimidiating rhyton
dimidiating rhytons
rhyta, dimidiating

watercolor (paint)
water color (paint)
watercolour (paint)
water-colour (paint)

Ancient Greek (language)
grc

Equivalence relationship

- **Homographs:** A *homograph* is a term that is spelled identically to another term, but the meanings of the terms are different.
- For example, *drums* can have at least three meanings: They can be components of columns, membranophones, or walls that support a dome.
- Homographs exist whether or not the terms are pronounced alike.

- **Qualifier:** A word or phrase used to distinguish a term in a vocabulary from otherwise identical terms that have different meanings. A qualifier is separated from the term, generally displayed within parentheses.

drums (column components)

drums (membranophones)

drums (walls)

DOME

DRUM
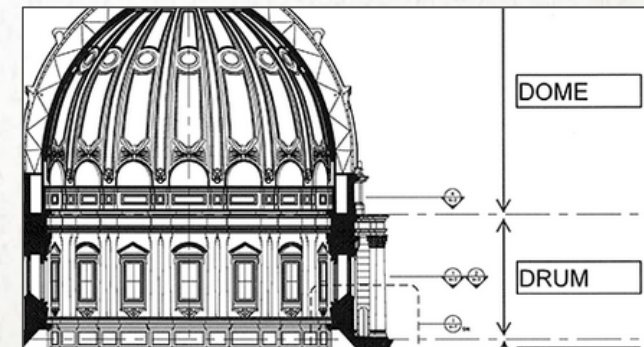
Equivalence relationship

preferred term
The term designated among all synonyms or lexical variants for a concept to be used as the default term to represent the concept in displays and other situations. In a monolingual thesaurus, the preferred term is also the only descriptor in the record. In a multilingual thesaurus, there may be a descriptor for every language, but there is often only one preferred term for the record as a whole.

jades (objects) (record-preferred, English-P,D)
     (objetos) Spanish-P,D)
jade (object) (English,AD)
     (objeto) Spanish,AD)
玉器 (Chinese (traditional)-P,D)
yù qì (Chinese (transliterated Hanyu Pinyin)-P,UF)
jaden beeldhouwwerken (Dutch-P,D)
jaden beeldhouwwerk (Dutch,AD)
Jadekünste (German,D)
Jadekunst (German-P,AD)

descriptor (D)
In a thesaurus, the term recommended to represent the concept in displays and indexing. A multilingual thesaurus may have multiple descriptors (one in each language represented), but may possibly have only one preferred term for use as default in displays.

alternate descriptor (ALT or AD)
A variant form of a descriptor available for use; usually a singular form or a different part of speech than the descriptor.

used for term (UF)
A term that is not a descriptor nor an alternate descriptor. If the thesaurus is being used as an authority, a used for term is not authorized for indexing. *Used for* terms typically comprise spelling or grammatical variants of the descriptor or have true synonymy with the descriptor

Equivalence relationship

Cross-language equivalence

When terms are in different languages in a multilingual thesaurus, each language may have descriptors, alternate descriptors, and used for terms in its own language; the relationship of terms across languages is known as *cross-language equivalence*.

vases (preferred,C,U,LC,English-P,D,U,PN)

vases (French-P,D,U,PN)

vase (C,U,English,AD,U,SN)

vase (French,AD,U,SN)

花瓶 (C,U,Chinese (traditional)-P,D,U,U)

陶瓶 (C,U,Chinese (traditional),UF,U,U)

huā píng (C,U,Chinese (transliterated Hanyu Pinyin)-P,UF,U,U)

hua ping (C,U,Chinese (transliterated Pinyin without tones)-P,UF,U,U)

hua p'ing (C,U,Chinese (transliterated Wade-Giles)-P,UF,U,U)

vazen (C,U,Dutch-P,D,U,U)

vaas (C,U,Dutch,AD,U,U)

vasos (C,U,Spanish-P,D,U,PN)

vaso (C,U,Spanish,AD,U,SN, Italian-P, AD)

vasi (C,U,Italian,D,U,PN)

Hierarchical Relationship

Hierarchical relationships: The broader and narrower (parent/child) relationships between records. The hierarchical relationship is the primary feature that distinguishes a thesaurus or taxonomy from simple controlled lists and synonym rings.

Hierarchical relationships are referred to by genealogical terms: *child, children, siblings, parent, grandparent, ancestors, descendants,* etc.

Getty

Hierarchical relationship

Types of hierarchical relationships

Genus/Species Relationships

- The genus/species or generic relationship is the most common relationship in thesauri and taxonomies because it is applicable to a wide range of topics.
- All children in a genus/species relationship should be a type of the parent, (e.g., *bronze* is a type of *metal*).

- The placement of a child may be tested by the all/some argument.
- In the example of bronze, all architectural bronze is bronze, but only some bronze is architectural bronze.

bronze

ALL ↑ ↓ SOME

architectural bronze
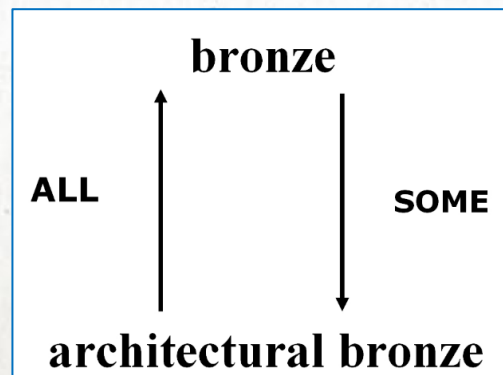
Top of the AAT hierarchies
... Materials Facet
...... Materials (Hierarchy Name)
.... materials (matter)
.... <materials by composition>
.... inorganic material
.... metal
.... <metal by composition or origin>
.... nonferrous metal
.... <copper and copper alloy>
.... copper alloy
.... bronze (metal)
.... architectural bronze
.... bell metal
.... Corinthian bronze
.... gunmetal (casting bronze
.... lead bronze
.... manganese bronze [N]
.... nickel bronze
.... phosphor bronze
.... statuary bronze

Getty

Hierarchical relationship

Types of hierarchical relationships

Whole/Part Relationships

- *Whole/part*, also called a *partitive relationship*, typically applied to geographic locations, parts of corporate bodies, parts of the human body, and other types of concepts that are not readily placed into genus/species relationships.

- Each child should be a part of the parent and all the other ancestors above it. (e.g., *Acomayo* is a part of *Cusco* region).

Top of the TGN hierarchy (hierarchy root)
... World (facet)
...... South America (continent)
.... Peru (nation)
.... Cusco (region (administrative division))
.... Acomayo (inhabited place)
.... Anta (inhabited place)
.... Calca (inhabited place)
.... Ccapi (inhabited place)
.... Chanapata (deserted settlement)
.... Chaullay (inhabited place)
.... Choquequirao (deserted settlement)
.... Conchopata (inhabited place)
.... Cuzco (inhabited place)
.... Espinar (inhabited place)
.... Huarocondo (inhabited place)
.... Lares (inhabited place)
.... Llactapata (deserted settlement)
.... Llusco (inhabited place)
.... Machupicchu (inhabited place)
.... Machupicchu (deserted settlement)

Hierarchical relationship

Types of hierarchical relationships

Instance Relationships

- In addition to the whole/part and genus/species relationships, some vocabularies may utilize a third type of hierarchical relationship, the instance relationship.

- Children are an instance or example of the parent, most commonly seen in vocabularies where proper names are organized by general categories of things or events

- For example, if the proper names of mountains and rivers were organized under the general categories mountains and rivers.

Top of the ULAN list / hierarchy
.... Non-Artists (ULAN facet)
.... Abarca de Bolea, Pedro Pablo, conde de Aranda (Spanish
.... count, statesman, 1719-1798)
.... Abati Olivieri, Giulia degli (Italian aristocrat, 1630-1718)
.... Abbas, M. A. (Hong Kong author and critic, born 1942)
.... Abbey, John Roland (English collector, 1894-1969)
.... Abbott, Frederick (British army officer, 1805-1892, active
.... in India)
.... Abbott, Helen (American editor and writer, contemporary)
.... Abd el-Maksond, Mohamed (Egyptian author, contemporary)
.... Abdülmecit I, Sultan of the Turks (Ottoman sultan, 1823-1861)
.... Abelard, Peter (French theologian, philosopher, 1079-1142)
.... Aberson, Frederick Nicholas Lucretius (Dutch military
.... officer, 1779-1859)

mountains
.... Alps
.... Apennines
.... Himalayas
rivers
.... Amazon River
.... Colorado River
.... Mississippi River
.... Nile
.... Ohio River

Getty

Hierarchical relationship

Record types for thesauri

facet

A fundamental, homogeneous, and mutually exclusive category in a thesaurus (for example, the AAT has seven facets: *Associated Concepts, Physical Attributes, Styles and Periods, Agents, Materials, Activities*, Objects; an eighth facet, *Brand Names*, was recently added).

guide term

Also called a *node label*. A record represented by a term or phrase that is created as a hierarchical level where no concept is appropriate as the level, to provide order and structure to thesauri by grouping narrower terms according to a given logic. Guide terms are not used for indexing and are often enclosed in angled brackets (for example, *<photographs by form>*).

concept

In the context of the AAT and other thesauri comprising generic terms, the subject of the vocabulary record (i.e., the concept to which the terms refer).

Top of the AAT hierarchies
... Objects Facet
...... Visual and Verbal Communication (Hierarchy Name)
.... Visual Works (Hierarchy Name)
.... visual works (works)
.... <visual works by material or technique>
.... photographs
.... <photographs by form>
.... negatives (photographs)
.... color separation negatives
.... glass plate negatives
.... halftone negatives
.... internegatives
.... <negatives by color>
.... <negatives by process>
.... paper negatives
.... positives (photographs)
.... direct positives
.... holograms
.... photographic prints
.... photographic transparencies

Getty

Hierarchical relationship

Polyhierarchy

- Some concepts logically belong to more than one broader context. To accommodate this, the data structure should allow polyhierarchical relationships, meaning that each child may be linked to multiple parents.

- Polyhierarchical relationships may exist in whole/part, genus/species, and instance relationship models.

- In a polyhierarchy, each record exists only once in the vocabulary, but may be linked to multiple parents and can thus appear in multiple hierarchical views.

- In the example here, Siena is part of the modern nation of Italy, but it was also part of the ancient confederation of Etruria. Dates and Historical flag clarify this is a former relationship

World (facet)
....Europe (continent) (P)
........Italy (nation) (P)
............Tuscany (region (administrative
            division)) (P,Current) *part of new kingdom
            of Italy from 1861*
................Siena (province) (P)
...................Siena (inhabited place) (P)

Additional Parent:
World (facet)....Europe (continent) (P)
..........Italian Peninsula (peninsula) (P)
............Etruria (former group of
            nations/states/cities) (P,Historical)
................Siena (inhabited place) (P,Historical)
            *sites near Siena date to the Late Etruscan
            period*

Getty

Associative Relationship

Associative Relationships: Relationships between records that are closely related conceptually, but where the relationship is not equivalent and not hierarchical.

- Only clear and direct associative relationships should be recorded.

  - If a thesaurus is bound together by too many associative relationships between entities that are only loosely or indirectly related, the value of the relationships in retrieval is lost.

  - These direct relationships are typically current, but occasionally may be historical.

Getty

ASSOCIATIVE RELATIONSHIP

When to make associative relationships

- Associative relationships may be made between records in the same or different hierarchies.
  - There may occasionally be relationships between overlapping siblings, for example where the meanings are similar and the terms are occasionally (but not generally) used as synonyms.

- In general, terms that are mutually exclusive do not require associative relationships, particularly if they cannot be confused with one another.

- There should be associative relationships between terms that are intended to be used as separate concepts but may be confused by users.
  - Barring that situation, do not make associative relationships between homographs simply because the terms are spelled the same

Final Neolithic (transitional periods Stone Age to Bronze Age, three-age system, ... Styles and Periods)

Note: Refers to the last stage of a Neolithic culture, which often partially overlaps with the Early Bronze Age. It is generally held to have occurred from around 4,000 BCE to around 3,000 BCE in Europe, lasting until around 2,000 BCE in parts of the British Isles.

Related concept:
meaning/usage overlaps with
  Early Bronze Age (Bronze Age, three-age system, ... Styles and Periods)

Types of associative relationships

- Most basic type of associative relationship is simply *related to*.

- In some vocabularies, more specific types of associative relationships may be designated.

- Associative relationships are reciprocal.

  - For some relationships, the relationship type is the same on both sides of the link (e.g., "related to")

  - For others the associative relationship is different depending upon which record is the focus.

student of
teacher of ⟹ <name of teacher>
<name of student>

Hokusai, Katsushika (Japanese
    printmaker and painter, 1760-1849)

teacher of
    Taito, Katsushika, II (Japanese
    printmaker, active ca. 1820-1850)

Getty

VOCABULARIES FOR CULTURAL OBJECTS

BRIEF OVERVIEW OF THE GETTY VOCABULARIES

Getty

Vocabularies for cultural objects

Types of vocabulary terms

Classification: antiquities   `AAT`
Work Type: amphora
Creators:
Painter of the Wedding Procession *(Greek vase painter, active ca. 360s BCE)*   `ULAN`
Nikodemos *(Attic potter, active ca. 362 BCE)*
Repository: J. Paul Getty Museum
Geographic Locations:
Los Angeles, California, USA *(repository location)*
Athens, Greece *(creation location)*   `TGN`
Culture: Greek (Attic)   `AAT`
Materials & Techniques: `AAT`
terracotta, turning, sintering, vase painting
Specific Subjects:
Athena Promachos
Nike   `IA`
boxers (athletes)
prize
competition (event)   `AAT`
Panathenaia
kionedon (writing style)

- The types of terms that will be necessary for art and architecture include personal names, corporate body names, geographic names, object names, names of iconographic subjects, and genre terms, among others.
- Values may be provided from a variety of sources, flagged in the illustration from the Getty Vocabularies

Vocabularies for cultural objects

The primary vocabularies

Getty Vocabularies
AAT, TGN, ULAN, CONA, and IA are authorities in the form of thesauri, intended to provide terminology and other information about the objects, artists, concepts, and places important to various disciplines that specialize in art, architecture and material culture.

Nomenclature
Nomenclature 4.0 for Museum Cataloging is a revised and expanded version of Robert Chenhall's system for classifying man-made objects first published in 1978 as a cataloging tool for historical organizations.

Library of Congress Authorities
The U.S. Library of Congress Authorities include subject, name, and title authority records created by or for the Library of Congress; comprise a tool used by librarians to establish forms of names for persons, places, meetings, and organizations, as well as titles and subjects (i.e., topics) indexed in bibliographic records.

Thesaurus for Graphic Materials
Developed from a list of terms for visual images used by the U.S. Library of Congress Prints and Photographs Division, including subject terms and descriptive terms.

Iconclass
An alphanumeric classification scheme designed for the iconography of art, focusing primarily on the religious and mythological stories and themes in Western art.

Getty

THE GETTY VOCABULARIES
ENABLING DIGITAL ART HISTORY

- *Art & Architecture Thesaurus* ®
  AAT = includes generic terms, relationships, sources, and notes for work types, roles, materials, styles, cultures, and techniques
  (e.g., amphora, oil paint, olieverf, peintures à l'huile, acetolysis, 玉器, Jadekünste, sintering, orthographic drawings, Olmeca, Rinascimento, Buddhism, watercolors, asa-no-ha-toji)

- *Getty Thesaurus of Geographic Names* ®
  TGN = includes names, relationships, and coordinates for current places and special focus on historical cities, nations, empires, archaeological sites, lost settlements, and physical features; through LOD, TGN may be linked to GIS and maps
  (e.g., Diospolis, Acalán, Ottoman Empire, Mogao, Ch'ien-fu-tung, Ganges River, गंगा नदी)

- *Union List of Artist Names* ®
  ULAN = includes names, relationships, notes, sources, and biographical information for artists, architects, firms, studios, repositories, patrons, sitters, and other individuals and corporate bodies, both named and anonymous
  (e.g., Mark Rothko, Cai Xiang, 葵襄, Crevole Master; Altobelli & Molins, Rajaraja Museum)

- *Cultural Objects Name Authority* ®
  CONA = rich metadata such as title, artist attribution, patron, materials, and location for works of art, architecture, and other cultural works, to allow a conduit for research and discovery, particularly for built works, multiples (e.g., prints), works depicted in other works, works designed but not built, and lost works *(e.g., Chayasomesvara Temple, Hagia Sofia, Αγία Σοφία ,The Lacemaker, La Dentellière, Merlettaia, Lion Throne Room,神奈川沖浪, Great Wave, Die große Welle)*

- *Getty Iconography Authority*
  IA = includes proper names, relationships, and dates for iconographical narratives, religious or fictional characters, historical events, names of literary works and performing art, scope is global and multilingual *(e.g., Viaggio dei Re Magi, Flood of Deucalion, French Revolution, Olouaipipilele, Xibalba, Niflheim, शिव, Shiva, Bouddha couché)*
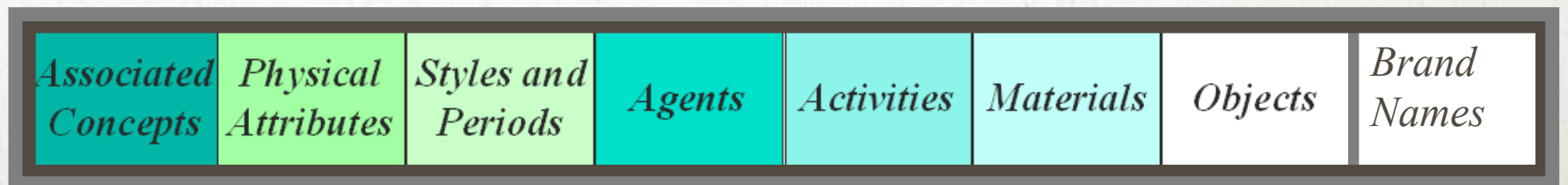
amfoor
amphorae
arc-boutant
brittleness
Buddhism
calligraphy
emulsifier
façades
flying buttress
飛扶壁
lithographer
litógrafos
luchtbogen
orthogonal plan
Macedonian
millwork
Rinascimento
ῥυτόν
stupa
Sturzbecher
travertine
weft wrapping
yastiks
zeppelin
Zigula
zinc oxide

AAT, the Art & Architecture Thesaurus®

The AAT is a thesaurus containing generic terms, dates, relationships, sources, and notes for work types, roles, materials, styles, cultures, techniques, and other concepts related to art, architecture, conservation, other cultural heritage

- Generic terms, not proper names
  oil paint, olieverf, acetolysis, sintering, orthographic drawings, Olmeca, Rinascimento, Buddhism, watercolors, asa-no-ha-toji, sralais
- AAT is multilingual; large translation projects are underway
- Conceptually organized from terms to describe abstract concepts to generic terms for concrete, physical artifacts
- Facets are the upper levels of the AAT structure
- AAT is not organized by subject matter or discipline

Associated Concepts	Physical Attributes	Styles and Periods	Agents	Activities	Materials	Objects	Brand Names

AAT, the Art & Architecture Thesaurus®

Sample record

ID: 300132869

Terms:
bobbin lace (pref, en)
bone lace (en)
cushion lace (en)
梭心蕾絲 (zh)
線軸編織花邊 (zh)
kloskant (nl)
dentelle aux fuseaux (fr)
encaje de bolillos (es)
encaje de bolillo (es)
Klöppelspitze (de)
Klöppelspitzen (de)

Associative Relationships
requires ... lace pillows (<textile
   fabricating tools and
   equipment>...Objects Facet)
   [300132869]

Hierarchical Relationships (poly)
.Objects Facet
... Visual & Verbal Communication
..... Visual Works
.... visual works (works)
.... <visual works by material >
.... needlework (visual works)
.... lace (needlework)
.... bobbin lace

Note: With "needle lace," one of two
primary types of handmade lace. It is
characterized by being made by ...

Chinese (traditional) ..... 與「針織蕾絲
   （needle lace）」同為主要的手工蕾絲,
其特色是以纏繞於線軸或梭心...

Dutch ..... Een van de twee belangrijkste
soorten met de hand vervaardigde kant;
'naaldkant' is ...

German ..... Zusammen mit der
"Nadelspitze" eine der wichtigsten...

Spanish ..... Junto a "encaje a aguja",
uno de los dos tipos principales de ...

Contributors: VP,CHIN,AS,RKD,IfM-SMB-PK,CDPB-DIBAM
Sources: Earnshaw, Clabburn, Needleworker's Dictionary (1976); Identification
of Lace, 2d ed. (1984); Ginsburg, Illustrated History of Textiles (1991)

Getty

Contributors are cited

AAT records are merged information from multiple contributors

Getty Vocabularies grow through contributions from the expert user community

All information in a Getty Vocabulary record is attributed to a contributor and cites the published sources

bobbin lace (**preferred**,C,U,LC,English-P,D,U,PN)
bone lace (C,U,English,UF,U,N)
cushion lace (C,U,English,UF,U,N)
lace, bobbin (C,U,English,UF,U,N)
pillow lace (C,U,English,UF,U,N)
梭心蕾絲 (C,U,Chinese (traditional)-P,D,U,U)
線軸編織花邊 (C,U,Chinese (traditional),UF,U,U)
骨架花邊 (C,U,Chinese (traditional),UF,U,U)
枕墊花邊 (C,U,Chinese (traditional),UF,U,U)
suō xīn lěi sī (C,U,Chinese (transliterated Hanyu
suo xin lei si (C,U,Chinese (transliterated Pinyin
so hsin lei ssu (C,U,Chinese (transliterated Wad
kloskant (C,U,Dutch-P,D,U,U)
dentelle aux fuseaux (C,U,French-P,D,U,PN)
merletto a tombolo (C,U,Italian-P,D,U,PN)
encaje de bolillos (C,U,Spanish-P,D,U,PN)
encaje de bolillo (C,U,Spanish,AD,U,SN)
Klöppelspitze (C,U,German,AD,U,SN)
Klöppelspitzen (C,U,German-P,D,U,PN)

- **Getty Conservation Institute** (Los Angeles, California)
- **Academia Sinica** (Nankang, Taiwan)
- **Netherlands Institute for Art History** (RKD: Bureau AAT) (The Hague, The Netherlands)
- **Centro de Documentación de Bienes Patrimoniales** (Dirección de Bibliotecas, Archivos y Museos) (Santiago, Chile)
- **Staatliche Museen zu Berlin Preussischer Kulturbesitz** (Berlin, Germany)
- *[Canadian Heritage Information Network (CHIN) (Gatineau, Quebec, Canada), Istituto Centrale per il Catalogo e la Documentazione (Rome, Italy)(ICCD)] and other contributors of partial translations; future translations in French, Italian, Portuguese, other languages*

Getty

Achilles Painter
Abasi, Riza
رضا عباسی
An Qi
安岐
Bartolo di Fredi
Galleria degli Uffizi
Gilbert & George
Katsushika Hokusai
Kicking Bear
Master of Boí
Méraud, Pierre-Antoine, père
Ngapatari, Hakaraia
O'Keeffe, Georgia
Pei, I. M.
Rothko, Mark
Sa'īd, Mahmūd
Skidmore, Owings & Merrill
unknown Mende
Vikramaditya I
Wren, Christopher
Zeuxiades
Ζευξιάδης

ULAN, the Union List of Artist Names®

ULAN contains names, relationships, notes, sources, and biographical information for artists, architects, firms, studios, repositories, patrons, sitters, and other individuals and corporate bodies, both named and anonymous

- Named people and corporate bodies
  Bartolo di Fredi, Xueshi Bai, 白雪石, National Palace Museum (Taipei)
- Anonymous creators known by appellation: *Santa Eufemia Master*
- Scope also includes "Non-Artists" (e.g., sitters and patrons)
- "Unknown People by Culture" (e.g., *unknown Aztec*)
- "Unidentified Named People" (e.g., known from archival documents)

- ULAN includes current and historical associative relationships
  (e.g., students to teachers; firm and studios to members)
- Contributions to ULAN are from experts at authorized institutions (e.g., museums, special collections, art libraries, cataloging projects, bibliographic projects, etc.)

ULAN, the Union List of Artist Names®

Sample record

ID: 500060426

Names:

Hokusai, Katsushika (pref,en,ja-trans,de)
Katsushika Hokusai (en,ja-trans)
Hokusai (en,ja-trans)
葛飾 北斎 (ja)
Shunrō (en,ja-trans)
Tawaraya Sōri (en,ja-trans)
Kakō (en,ja-trans)
Tatsumasa (en,ja-trans)
Gakyōjin (en,ja-trans)
Taito (en,ja-trans)
Iichi (en,ja-trans)
Manji (en,ja-trans)
Tokitarō (en,ja-trans)
時太郎 (ja)
葛飾北濟 (ja)

Nationalities:

Japanese (pref)

Roles:

artist
printmaker
painter
designer
calligrapher
draftsman
landscapist
marine artist
figure artist
ukiyo-e artist

Gender: male

Birth and Death Places:
Born: Tokyo (Kanto, Japan)) (inhab place)
Died: Tokyo (Kanto, Japan)) (inhab place)

Events:
active: Tokyo (Kanto, Japan)) (inhab place)
.... in Edo, modern-day Tokyo

Associative Relationships:
teacher of Taito, Katsushika, II
.............(Japanese printmaker, active ca. 1820-1850)
child of Nakajima Ise
.............(Japanese mirror maker, 18th century)
possibly identified with ... Tōshūsai Sharaku
.... (Japanese printmaker, active 1794-1795)

Contributors:
VP,Avery,GRL,BHA,CCA
Sources: Bowie, Drawings of Hokusai (1964);
Grove Dictionary of Art online (1999-2002); Library
of Congress Authorities database (n.d.);
Metropolitan Museum of Art [online] (2003-);
Bouquillard, Hokusai: First Manga Master (2007)

AAT

AAT

TGN

Getty

Relationships and linking are built into and among the Getty vocabularies

Merging data in one record

- Getty vocabularies merge contributed records representing the same concept
- Example from ULAN: Three contributions; are any of these records for the same person?
  Yes, 1st and 3rd rows
- Merged by algorithm, if possible; if not, merged by editor4s

ULAN ID	Matching name	Display biography	Roles (AAT-controlled)	Nationality (AAT-controlled)	Estimated birth date	Estimated death date
500017409	Guardi, Francesco	Venetian painter, 1712-1793	artist painter landscapist	Venetian Italian	1712	1793
500312969	Guardi, Francesco	Italian noble, born 1514	noble	Italian	1514	1614
500412666	Guardi, Francesco	Italian artist, ca. 1712-1793	artist	Italian	1707	1793

Getty

A name matches exactly; role matches; a nationality matches; dates within range?

Publishing data

- Contributors' data is loaded and processed by the Getty Digital technical team and the Vocabulary Program editorial team
- Data is merged, normalized, and published
- Refreshed for users every two weeks
- Releases: Online search, XML, relational tables, APIs, LOD

ULAN record is a merged, homogenous entity

Union List of Artist Names® Online
Full Record Display

🔍 New Search   ◀ Previous Page   ❓ Help

Click the 🔀 icon to view the hierarchy.

Semantic View (JSON, JSONLD, RDF, N3/Turtle, N-Triples)   ▪ Representa

🔀 **Guardi, Francesco** (Italian painter, 1712-179

Note: Venetian painter, during the first half of his
on decorations and church paintings; later becam
vedutista (view painter) of Venetian urban life.

Names:
Guardi, Francesco (**preferred**,V,index,LC)
Francesco Guardi (V,display)
Gardi, Francesco (V)
Garde, Francesco (V)
Gardis, Francesco (V)
Gauda, Francesco (V)
Guada, Francesco (V)
Guarde, Francesco (V)
Guardet, Francesco (V)
Gvardi, Franchesko (V)
Guardi, François (V)
Guardi, Francisco (V)

Contributors
AVERY, BHA, GRISC, GRL, GRLPA, PROV, VP, WCI, WCP, WL-Courtauld
Sources
Avery Authority files (1963-)
Harald Szeemann papers, Finding Aid, GRI Special Collections (2011-)
J. Paul Getty Museum, collections online (2000-)
Library of Congress Authorities database (n.d.) n 50074509
M. Knoedler & Co. records, Finding Aid, GRI Special Collections (2012-)
Provenance Index Databases, Authority file (1985-)
RILA/BHA (1975-2000)
Witt Library, Authority files

Merged in one record

Attribution to contributor and source is maintained, but merged and unified, and various "preferred" data is flagged

Introduction to Controlled Vocabularies

ULAN record is a merged, homogenous entity

Union List of Artist Names® Online
Full Record Display

🔍 New Search   ◀ Previous Page

Click the ⚎ icon to view the hierarchy.

Semantic View (JSON, JSONLD, RDF, N3/Turtle, N-Triples)   ▪ Representa

⚎ **Guardi, Francesco** (Italian painter, 1712-179

Note: Venetian painter, during the first half of his
on decorations and church paintings; later becam
vedutista (view painter) of Venetian urban life.

Names:
Guardi, Francesco (**preferred**,V,index,LC)
Francesco Guardi (V,display)
Gardi, Francesco (V)
Garde, Francesco (V)
Gardis, Francesco (V)
Gauda, Francesco (V)
Guada, Francesco (V)
Guarde, Francesco (V)
Guardet, Francesco (V)
Gvardi, Franchesko (V)
Guardi, François (V)
Guardi, Francisco (V)

Contributors
AVERY, BHA, C
Courtauld
Sources
Avery Author
Harald Szeem
Collections (2
J. Paul Getty M
Library of Co
50074509
M. Knoedler &
Collections (2
Provenance In
RILA/BHA (1
Witt Library,

Nationalities:
Italian (**preferred**)
Venetian   **AAT**

Roles:
artist (**preferred**)
painter
landscapist   **AAT**

Gender: male

Birth and Death Places: **TGN**
Born: Venice (Venezia province, Veneto, Italy) (inhabited place)
Died: Venice (Venezia province, Veneto, Italy) (inhabited place)

Related People or Corporate Bodies:
member of .... Guardi family
(Italian painters, active 18th century) [500060865]
parent of .... Guardi, Giacomo
(Italian painter, 1764-1835) [500023608]
sibling by marriage (in-law) of .... Tiepolo, Giovanni Battista
(Venetian painter, 1696-1770)
[500018523]   **ULAN**
Cecilia Guardi married Tiepolo in 1719
sibling of .... Guardi, Antonio
(Italian painter, 1699-1760) [500025094]
sibling of .... Guardi, Niccolò
(Italian painter, 1715-1785) [500029088]

Merged information,
redundancy eliminated

Note links to other
ULAN records, to AAT
and TGN

Implementing the Getty Vocabulary data

- Vocabulary data is licensed and used by hundreds of institutions and vendors
- We contribute to VIAF, collections management systems (e.g., Gallery Systems)
- Now linking vocabulary data in LOD

Getty

http://viaf.org/

The VIAF® (Virtual International Authority File) combines multiple name authority files into a si[...]
authority files by matching and linking widely-used authority files and making that information a[...]
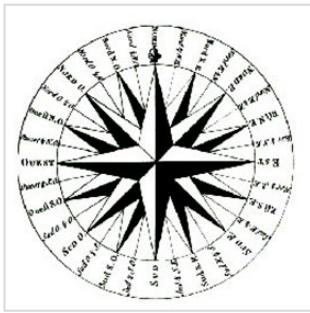
VIAF Contributors

Library of Congress/NACO
  National Library of Mexico
  British Library
  National Agricultural Library (U.S.)
  National Library of Medicine (U.S.)
  National Library of New Zealand

Vatican Library
Swiss National Library
Library and Archives Canada
Union List of Artist Names [Getty Research Institute]
NUKAT Center of Warsaw University Library

y, Hungary
f Western Switzerland

Louis XIV, roi de France, 1638-1715

Louis XIV, King of France, 1638-1715

Louis XIV, koning van Frankrijk, 1638-1715

Ludwik XIV (król Francji ; 1638-1715).

Lluís XIV, rei de França, 1638-1715.

Luis XIV, Rey de Francia

Louis 1638-1715 King of France XIV

Ludwig XIV. Frankreich, König 1638-1715

Ludvig XIV, kung av Frankrike, 1638-1715

Louis XIV roi de France et de Navarre de 1643 à 1715

Louis XIV, re di Francia, 1638-1715

Ludvikis XIV, Francijas karalis, 1638-1715

Ludvík XIV., francouzský král, 1638-1715

לואי 14, מלך צרפת, 1638-1715

Louis 14世, 1638-1715, フランス国王

Louis <re di Francia ; 14.>

Louis XIV, King of France (French monarch, collector, and patron, 1638-1715)

- GRI (with ULAN and TGN) contributes to VIAF, the Virtual International Authority File.
- Link to other sources in Linked Open Data

TGN, the Getty Thesaurus of Geographic Names®

TGN focuses on places relevant to art, architecture, and related disciplines, recording names, relationships, place types, dates, notes, and coordinates for current and historical cities, nations, empires, archaeological sites, lost settlements, and physical features

Abthugni
Angel Falls
Athínai
Αθήναι
Bhārat
Býčí Skála
Bora-Bora
Casablanca
Chichén Itzá
Cueva Cerro Sota
Donner Pass
Dos Cabezas Site
Firenze
Fort Rouillé
Gdansk
Genna Bogai, Arcu
Святой Престол
Jabal Mayyah
Kush
Luxor
מישׂראל
Mare Cantabricum
Nan Madol Ruins
Rapa Nui
臺北

- Places relevant to art history
  Thebes, Diospolis, Ottoman Empire, Mogao Caves, Ch'ien-fu-tung, Ganges

- TGN is a structured vocabulary
- TGN is a thesaurus compliant with ISO standards

- Focus on historical regions, archaeological sites, lost settlements, built upon a skeleton of the modern world from NGA/NIMA and USGS data sets

- TGN is not GIS, is a thesaurus focused on names and links. Why invent the wheel over again? Many sources of geographic data exist. However, TGN may be linked to GIS, maps, and other geographic resources

TGN, the Getty Thesaurus of Geographic Names®

Sample record   Linked to other vocabularies

ID: 7011179

Names:
Siena
(it,en,es,fr,nl,de)
Σιένα   (el)
Siyenah
锡耶纳   (zh)
シエーナ   (ja)
Sienese (en-adj)
Sienna (h-en)
Sena Julia (h-la)
Sanna   (h)
Saena   (h)

Hierarchical Rels. (polyhier.)
World (facet)
.. Europe (continent)
... Italy (nation)
.... Tuscany (region)
...... Siena (province)
.... Siena (inhabited place)

World (facet)
.. Europe (continent)
... Italian Peninsula (peninsula)
..... Etruria (former confederation)
.... Siena (inhabited place)
            Historical: Late Etruscan

Associative Rels
ally of Pisa ..(inhabited place) [7006082]
    Ghibelline allies, 13th-14th centuries

Place Types
inhabited place   settled by Etruscans
      (flourished by 6th century BCE)
city
archiepiscopal see   **AAT**
commune (administrative) since 1125
cultural center
university center
world heritage site

Lat: 43 19 07 N (43.3180)
Long: 011 19 50 E (11.3300)
Elevation: 1056.430 feet (322.0000 meters)

Note: Siena was founded as an Etruscan hill town; later was the Roman city of Sena Julia (3rd century BCE). It flourished under the Lombard kings (6th century CE) and was Medieval self-governing commune ...

Contributors: BHA, FDA, GRLPSC, VP
Sources: Annuario Generale (1980); Canby, Historic Places (2007) 2/861; Dizionario Corografico Toscana (1977) 1508; Library of Congress Subject Headings (2002-); Milanesi, Storia dell'Arte Senese (1969) passim; NGA/NIMA database (2003-) -129709; TCI: Toscana (1984) 479 ff; Times Atlas of the World (1992) 183; UNESCO World Heritage Centre [online] (1992-) accessed 20 July 2010; Webster's Geographical Dictionary (2007) 1114
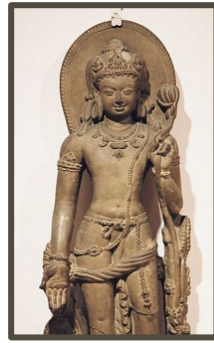
Getty

Adorazione dei Magi
Apedemak
Baby Jaguar
Death and the Miser
Ehecatl
Ganesh
Hades
Hercules
Herakles
Mithras
Niflheim
Noli me tangere
Olouaipipilele
Prajnaparamita
Penance of Indra
Reclining Buddha
Bouddha couché
涅槃仏
Storming the Bastille
The Divine Comedy
Tumatauenga
Virgin Hodegetria
World War I
Xibalba

IA, the Getty Iconography Authority ™

IA is a thesaurus that covers topics relevant to art, architecture, and related disciplines; includes multilingual proper names, relationships, and dates for iconographical narratives, religious or fictional characters, themes, historical events, and named literary works and performing arts

- Includes the proper names of subjects not covered by other Getty Vocabularies
- Scope is multicultural and multilingual, grows through contributions
- Includes links to other sources where there is overlap, such as Iconclass and US Library of Congress subject authorities
- The IA is compliant with the Subject Authority of CDWA and CCO (*Categories for the description of Works of Art* and *Cataloging Cultural Objects*)
- The IA has a thesaural structure; it includes equivalence, associative, and hierarchical relationships
- The IA is linked to the other Getty Vocabularies

IA, the Getty Iconography Authority ™

Sample record

Names in multiple languages

Users may choose the name appropriate for their needs

ID: 1000085

Names:
Avalokiteshvara
Avalokiteśvara
अवलोकितेश्वर
སྤྱན་རས་གཟིགས
Spyan ras gzigs

AAT links
role/characteristic is ... bodhisattva
symbolic attribute is ... lotus
culture/religion is ... Mahayana (Buddhism)
culture/religion is ... Theravada (Buddhism)

AAT TGN ULAN CONA

Associative Relationships
associated with .... Krishna (Hindu iconography)
counterpart is …. Guanyin (Buddhist iconography)

Note The bodhisattva of infinite compassion and mercy; embodies the compassion of all Buddhas. This bodhisattva is portrayed in different cultures as either female or male…

Hierarchical Relationships
Top of the IA hierarchies
   Legend, Religion, Mythology
      Buddhist iconography
         Buddhist characters
            Avalokiteshvara (Buddhist iconography)

Contributors & sources
[VP] Encyclopedia Britannica Online (2002-); Bowker, Oxford Dictionary of World Religions (1997); Huntington, Art of Ancient India (1985); *LC: LC control no.: sh 85010492*

Getty

Bhimesvara Temple
Bradshaw rock
   paintings
Case Study House 21
Château de Versailles
Empire State Building
En liberté comme toi
Himeji Castle
姫路城
Irises
Kashmir Gate
Mona Lisa
La Gioconda
Model for the Façade
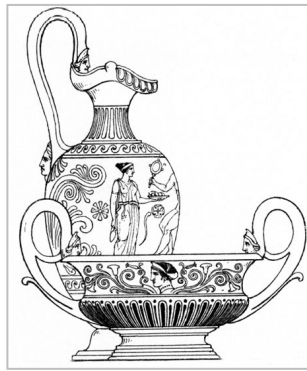   of San Lorenzo
MS Ludwig XV
Pantheon
Portrait of Napoleon
Ronda di notte
Square Ting Vessel
   of Duke Kang
西周早期 康侯方鼎

CONA, the Cultural Objects Name Authority®

CONA compiles titles/names and other metadata for works of art, architecture, and other cultural works, current and historical, documented as items or in groups, whether works are extant, destroyed, or never built

- In development, may be used to record works depicted in visual surrogates and for other purposes

- CONA compiles titles, attributions, depicted subjects, and other metadata about works of art, architecture, and cultural heritage, both extant and historical works, works never built, disassembled works, conceptual works for multiples
  Hagia Sophia, Mona Lisa, Thirty-six Views of Mount Fuji, 富嶽三十六景

- Scope is multicultural and multilingual; grows through contributions
- CONA is linked to images; CONA is linked to the AAT, TGN, ULAN, and IA
- Through this rich metadata and links, CONA may provide a powerful conduit for research and discovery for digital art history

THE SIMPLIFIED ENTITY-RELATIONSHIP DIAGRAM FOR CONA IS THE SAME AS CDWA / CCO

ULAN

TGN

AAT

Iconography Authority

CONA
(Work Records)

Source Records

Getty
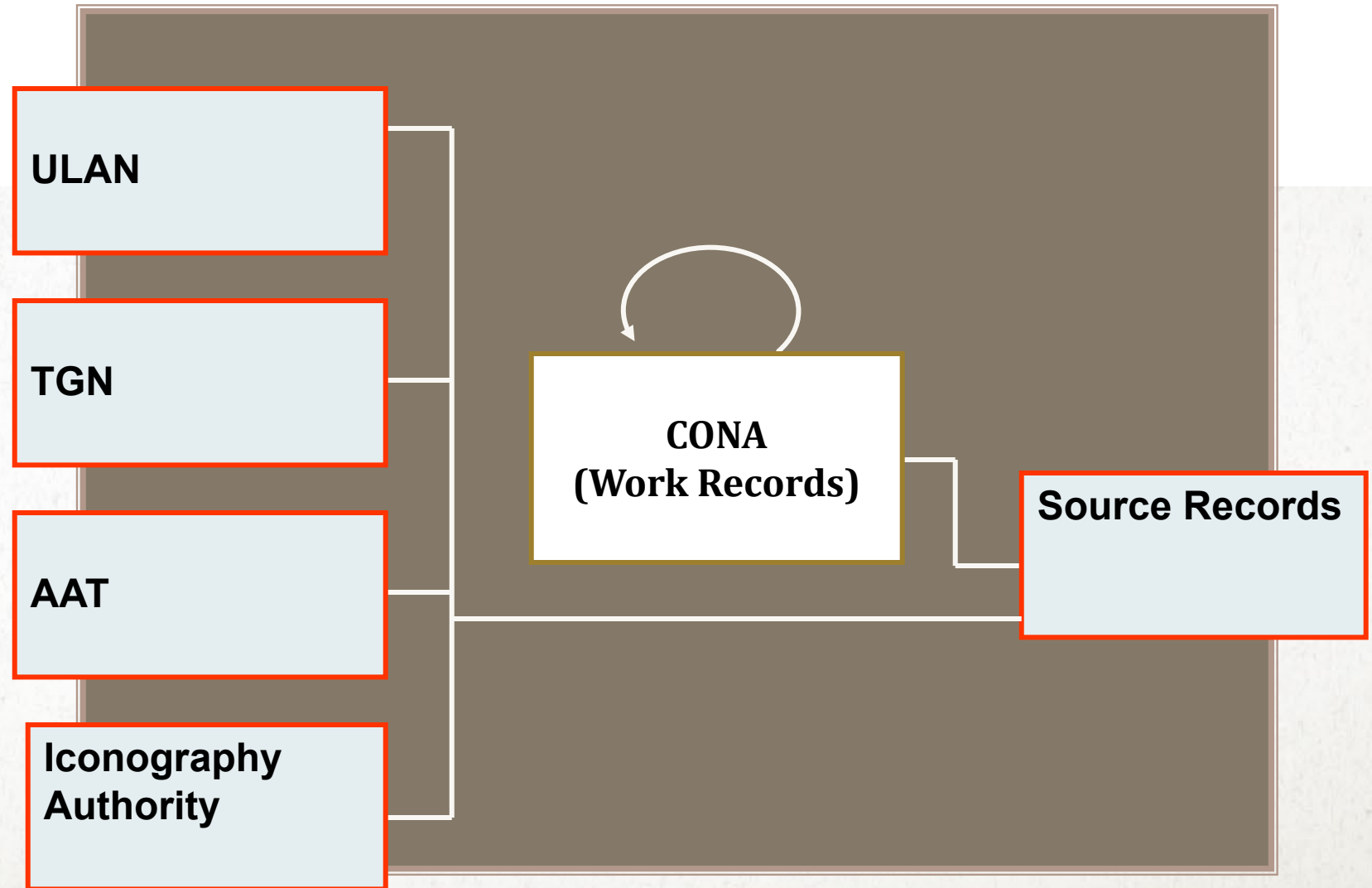
CONA, the Cultural Objects Name Authority®

Sample record

ID: 700001950

Titles:
Shiva met de maan in het haar (nl)
Shiva with the Moon in his Hair (en)

Catalog Level: item

Work Types:
sculpture (visual work)   AAT

Classifications:
sculpture (preferred)

Creation Date: 1000/1200

Creator Display:
anoniem
sculptor: unknown Chola   ULAN

Locations: ULAN
Current: Rijksmuseum (Amsterdam, North Holland,   TGN
Netherlands) AK-MAK-1291; RM001.collect.910; Bruikleen
van de Vereniging van Vrienden der Aziatische Kunst
Creation: Tamil Nādu (India)

Materials: bronze founding   AAT

Dimensions: 40 cm (height) x 24cm (width) x 10.5 (depth)

Events: exhibition: Metamorfoze/Geheugenproject Willem
Witsen   AAT

Cultures:
Indian   Chola   AAT

General Subject:
human figure(s) (preferred)
religion and mythology

AAT
TGN
ULAN
CONA
CONA IA

Specific Subjects:
Shivá (Hindu iconography)
mudrā (pose, <visual and representational concepts>)
[link to Iconclass]

Contributors & Sources: [Rijksmuseum;VP]; Rijksmuseum XML file

CONA, the Cultural Objects Name Authority®

Sample record

ID: 700000141

Titles/Names:

Hagia Sophia (en,de,nl)
Ayasofya (tr)
Αγία Σοφία (el)
Agia Sofia
Haghia Sophia
Sainte-Sophie (fr)
Santa Sofía (es)
Santa Sofia (it)
Sancta Sophia (la)
Sancta Sapientia (la)
Holy Wisdom (en)
Saint Sophia (en)
聖蘇菲亞大教堂 (zh)

Work Types: church; mosque; museum **AAT**

Date: original structure dated from 4th century CE; present structure built 532-537 CE; rebuilt in 12th century

Creators: Anthemios of Tralles (Byzantine architect, ca. 474-ca. 534) and Isidoros of Meletus, the Elder (Byzantine architect, active mid-6th century); **ULAN**
Patron: Justinian I (Byzantine emperor, ca. 482-565)

Built Work
... Hagia Sophia
Associative Relationship
depicted in *View of the Hagia Sophia in Constantinople*; Eduard Hildebrandt; watercolor; ca. 1852; Hermitage Museum (Saint Petersburg, Russia) [700008345]

Location: İstanbul (Marmara, Turkey) **TGN**
Address Note: 41.008548°N; 28.979938°E

Materials: system bearing masonry, **AAT** centralized plan; interior surfaces are sheathed with polychrome marble, porphyry, and mosaics
Dimensions: central dome: diameter 31 m (102 feet); height 56 m (184 feet)

General Subject: architecture **AAT TGN ULAN CONA CONA IA**
Specific Subject: **Holy Wisdom** Religion/Mythology)

Note: Domed basilica was co... by the Emperor Justinian; built in Constantinople in 6th century...

Contributors: Avery, BHA, FDA, VP
Sources: Fleming, Penguin Dictionary of Architecture (1999); Herrin, Judith, Byzantium (2007); Maidstone, Hagia Sophia (2002); Theoharidou, Architecture of Hagia Sophia (1988); Xydis, Chancel Barrier of Hagia Sophia (1947)

Getty

Movable works linked to architecture depicted

CONA ID 700000120
Work Type: architectural drawing
Title:
Creator:
Current:
Co:
An:

CONA ID 700000090   Class. architecture
Work Type: house
Title: Case Study House No. 21
Title: Bailey House Title: CSH #21
Creator: architect: Pierre Koenig
Creation Date: 1956-1958; renovated 1998
Dimensions: 1 story, 1320 square feet, 4 rooms and 2 baths, on a 110 x 160 foot lot
Mat & Tech: steel frame and flat roof deck
Current Location: Los Angeles (California, USA)
Address Note: 34.116634 -118.391623; 038 Wonderland Park Avenue, Hollywood Hills, Los Angeles, 90046

LINKS

preparatory for

based on
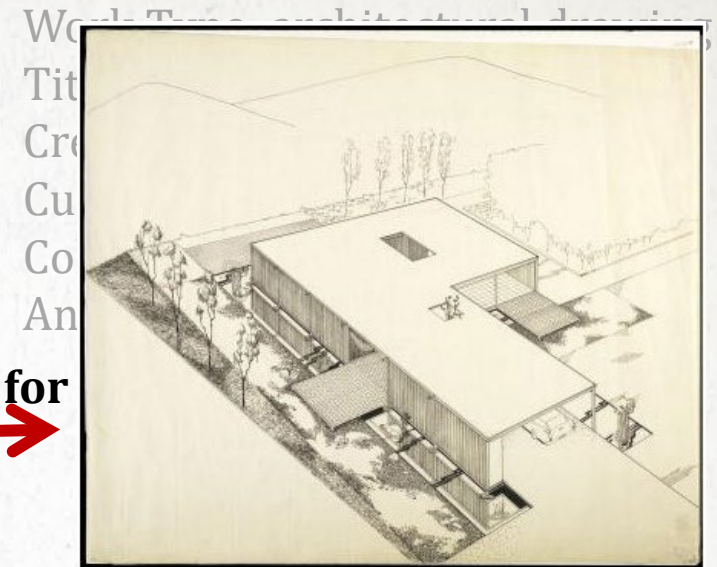
depicts

depicted in

CONA ID 700008539 Work Type: photograph
Title:
Creator:
Current:
Co:
An:
26

Getty

Possible implementations of CONA

Merging information from multiple contributors

- CONA ID uniquely identifies the work in CONA
- Links to other IDs, such as LOC, VIAF, BWR

CONA ID: 700002923
Catalog Level item **Classification** architecture
Work Type building complex | temple
Title/Name Temple of Heaven (en)
   Tian Tan (zh-tr) | T'ien-t'an (zh-tr) | 天坛 (zh)
   | 天壇 (zh) | 천단 (ko)
Creator [display] first built by Ming Chengzu (1360-1424); extended by Ming Shizong (1507-1567); renovated by Qianlong (1711-1799)
Creation Date [display]: 1530, during Ming dynasty
General Subject architecture **Type:** isness
Specific Earth [TGN 7030638] **Extent:** dedication
Current Location Beijing (Beijing Shi (municipality), China) (nation) **Address Note**: 39°57′05″N 116°24′36″E
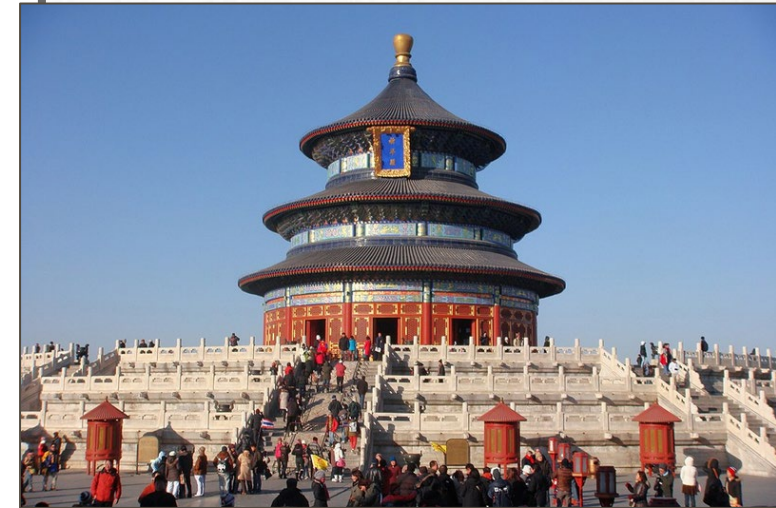**Dimensions** [display] entire complex: 237 hecatres
Mat & Tech [display] wood and stone [wood AAT 300011914 | stone AAT 300011176]
Style: Chinese [AAT 300018322] | Taoist [AAT 30014366
Contributors JPGM VP GRISC BWR
Sources Built Works Registry **BWR3792244**; Library of Congress Authorities online (2002-) **n2011019454**; Fodor's Travel (2016); J. Paul Getty Museum, collections online (2000-); Great Buildings Online (1997-)

LINKS

Getty

Possible implementations of CONA

CONA ID: **700000178**
Catalog Level item
Classification paintings
Work Type panel painting | altarpiece
Title Adoration of the Magi
　　Adorazione dei Magi **(repository title)**
Creator Bartolo di Fredi (Sienese painter, active by 1353, died 1410)
Creation Date [display]
1380s/1390s; some scholars date it ca. 1400 or later
execution　　　**Start Date 1380** **End Date 1399**
alternative date　**Start Date 1395** **End Date 1409**
General Subject religion and mythology | human figures
Specific | Adoration of the Magi **Extent:** foreground | Siena | horses | Holy Family | gold | frankincense | myrrh | Journey of the Magi **Extent:** background | king | deity | death | Epiphany | Three Ages of Man
Current Location Pinacoteca Nazionale di Siena (Siena, Italy);
Repository No **no. 104**
Dimensions 195 x 158 cm (76 x 61 5/8 inches)
Mat & Tech tempera on panel
LINKS
Contributors BHA VP CD PNS GRI-PA
Sources Pinacoteca Nazionale di Siena (2002-); Boucher and Fiorini, Magi Reconstructed (2012); Cole, Sienese Painting (1985); Freuler, Bartolo di Fredi (1994); Harpring, Bartolo di Fredi (1992)

- Opinions vary, all may be included in the CONA record

Getty

Possible implementations of CONA

Works linked to subjects depicted

OCLC survey, of core fields missing is <u>subject depicted</u>
Mining titles and text for subjects depicted, missing from museum data
Controlled by AAT, ULAN, TGN, CONA, and IA

LINKS

CONA ID 70000012 Class: Pre-Colombian art

General Subject:
• religion and mythology [General Subject ID 31801]

Specific Subjects:
• Xibalba (Maya underworld) [IA 1000045]
• Baby Jaguar (Maya character) [IA 901002211]
• Chahk (Maya deity) [IA 901002210]
• man (male human) [AAT 300025928]
• skeleton (animal component) [AAT 300191778]
• ax (weapon) [AAT 300024664]
• altar [AAT 300003725]
• death [AAT 300151836]
• celebration [AAT 300152441]

IA ID 1000045
Names
• Xibalba
• Place of Fear

AAT ID 300191778
Terms
• skeleton [English]
• beendergestel [Dutch]
• esqueleto [Spanish]
• squelette [Italian]
• 骨架 [Chinese]

Note Source: Metropolitan Museum of Art on line

Works depicted in visual surrogates

- What is advantage to linking to CONA for work depicted?
- Access for all surrogates for same work
- Multilingual access, links, etc.

Surrogate in Photo Archive

Velázquez et al copy in the Uffizi after Rube[...]

LINKS

specific subject

Enough information in surrogate records to make minimal record for work depicted

CONA_ID: 700008694   Class.: paintings
Titles: Equestrian Portrait of Philip IV
                Filippo IV, re di Spagna
                Felipe IV a caballo
Work Types: painting (visual work)   **AAT**
Creation Date: ca. 1645
Creator Display: Diego Velázquez and assistants; after Peter Paul Rubens, now lost   **ULAN**
Locations: Galleria degli Uffizi (Flore[...]ze province, Tuscany, Italy) [500125191]   **TGN**
Repository Numbers: 1890:792   **ULAN**
Address Note: Room 41: Ruben[...]
Display Materials: oil on canvas   **AAT**
Dimensions: 338 x 267 cm
General Subject: portraits
Specific Subjects:
        equestrian portrait [300403976] (AAT)   **AAT**
        Philip IV, King of Spain (Spanish king, 1605-1665) [500282775] (ULAN)   **ULAN**
Related Works:
        pastiche copy after .... Equestrian Portrait of Philip IV [700008695] painting (visual work); Peter Paul Rubens; now lost; 1628-1629
Sources and Contributors:
[VP] [GRIPA]   **CONA**
.... Fossi, Uffizi Gallery (2001)
.... GRI Photo Archives database (1974-)

CO[...]
surrog[...]
Titles: Equestrian Portrait of Philip IV
Work T[...]
[30040[...]
Creatio[...]
Creato[...]

(Flemis[...]
depicte[...]
Locatio[...]
Reposito[...]
Display Ma[...]
Dimension[...]
General Su[...]
Specific Su[...]

Velázquez[...]
List/Hiera[...]
..... Visual S[...]
.... Study[...]
1600)
Sources an[...]
[VP]
 GRI Photo[...]

.... Fossi, Uffizi Gallery (2001)
.... GRI Photo Archives database (1[...]

HOW ARE VOCABULARIES USED?

- **Traditional use of vocabularies for retrieval focuses on the names**
- *If a user asks for one name/term, all of the synonyms can be used for retrieval*
- *For example, if a user searches on "Giovanni da Bologna"...*

Giambologna
Female Figure;
Marble; H: 48 in.; 2008 J. Paul Getty Museum; 82.SA.37
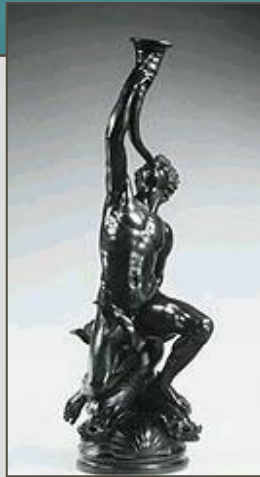
Giovanni Bologna;
Triton, 16th century (1560–70); Bronze; H. 36 in; Metropolitan Museum of Art; Bequest of Benjamin Altman, 1913 (14.40.689)

Jean Boulogne;
Mercure volant; Bronze; H 1,7 m; L 0,57m; Pr 1,4 m; Musee du Louvre, Collection du duc de Brissac (saisie en 1794)?; MR
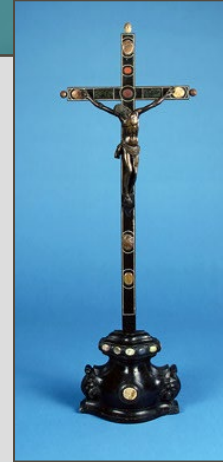
Bologna, Giovanni da; *Christ Crucified,* probably before 1588; bronze, corpus, height including foot rest: .371 x .254 x .078 m ; National Gallery of Art, Washington, DC; Gift of Mr. and Mrs. John R. Gaines in memory ...

From ULAN
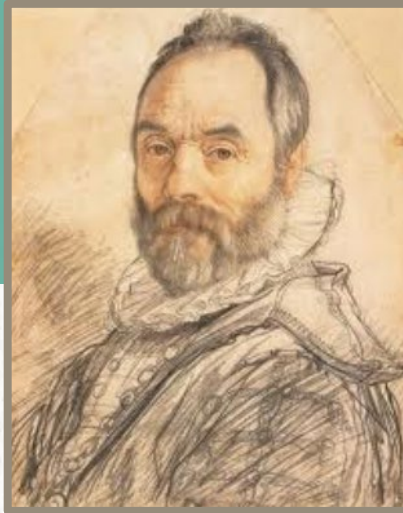
Giambologna

Giovanni da Bologna

Giovanni Bologna

Jean Boulogne

Jean de Boulogne

Bologna, Giovanni da

"giambologna" OR "giovanni da bologna" OR "giovanni bologna" OR "jean boulogne" OR "bologna, giovanni da"

Getty

HOW ARE VOCABULARIES USED?

Associative relationships

- New methods and LOD can incorporate hierarchical and associative relationships, and other links or inferred relationships to answer more complex questions and aid serendipitous discovery

Giambologna *(Flemish sculptor and architect, 1529-1608, active in Italy)*

Marie de Medici *(French queen, patron, 1573-1642)*

Tacca, Pietro *(Italian sculptor and architect, 1577-1640)*

had patron

patron of
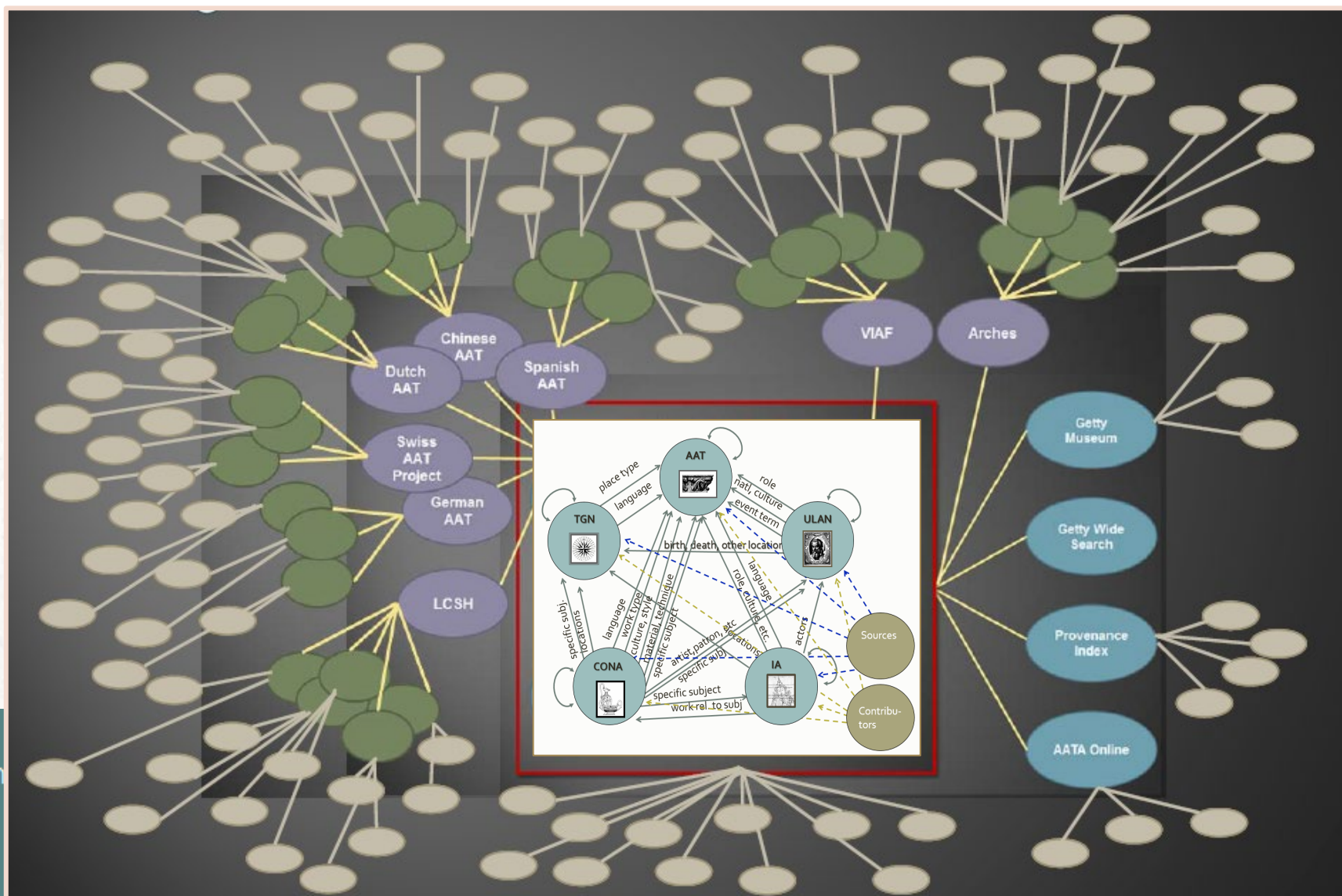
teacher of

student of

Display Date: from 1592
Start Date: 1592   End Date: 1608

Getty

WHAT IS LINKED OPEN DATA (LOD)?

- The AAT, TGN, and ULAN are now available as LOD
- They are published under the Open Data Commons Attribution License (ODC-By) 1.0

- When data is linked and open, it means that data is structured and published according to the principles of Linked Data, so that it can be both interlinked and made openly accessible and shareable on the Semantic Web. The goal of linked open data is to allow data from different resources to be interconnected and queried.

- In order for data to be understood and processed automatically by computers, data in records or about resources must be expressed in a standard format. Each thing (for example, a museum object, a place, or a person) must be represented by a persistent identifier (known as a Uniform Resource Identifier (URI)). A Resource Description Framework (RDF) is a language or format for describing things as well as the relationships between things as simple properties and values (known as 'triples'), while things are represented using URIs.

{   "Subject" : {   "type" : "uri",   "value" : "http://vocab.getty.edu/aat/300198841"   },   "Predicate" : {   "type" : "uri",   "value" : "http://www.w3.org/2000/01/rdf-schema#label"   },   "Object" : {   "xml:lang" : "el-latn",   "type" : "literal",   "value" : "rhyta"   }   }, {   "Subject" : {   "type" : "uri",   "value" : "http://vocab.getty.edu/aat/300198841" },   "Predicate" : {   "type" : "uri",   "value" : "http://www.w3.org/2000/01/rdf-schema#label"   },   "Object" : { "xml:lang" : "es",   "type" : "literal",   "value" : "rhyta"   }   }, {   "Subject" : {   "type" : "uri",   "value" : "http://vocab.getty.edu/aat/300198841"   },   "Predicate" : {   "type" : "uri",   "value" : "http://www.w3.org/2000/01/rdf-schema#label"   },   "Object" : {   "xml:lang" : "en",   "type" : "literal",   "value" : "rhyton"   }   }, {   "Subject" : {   "type" : "uri",   "value" : "http://vocab.getty.edu/aat/300198841"   },   "Predicate"

Getty Vocabularies are linked to each other

{   "Subject":{   "typ...   ...dicate":{   "type" :
"uri",   "value" :"http:/...   ...: "el-latn",   "type" :
"literal",   "value" : "rhy...   ...ty.edu/aat/300198841"
},   "Predicate":{   ...   "Object":{
"xml:lang" : "es",   "typ...   "value" :
"http://vocab.getty.ed.../...   ...dicate ...
"http://www.w3.org/2000/01/rdf-schema#label"   },   "Object":{   "xml:lang" : "en",   "type" :"literal",   "value" :
"rhyton"   } },{   "Subject":{   "type" : "uri",   "value" : "http://vocab.getty.edu/aat/30019884...   "Predicate"

slide by Joan Cobb

{ "Subject":{ "type" : "uri", "value" :"h...latn", "type" : "literal", "value" ...u/aat/300198841" }, "Predicate":{ "Object":{ "xml:lang" : "es", ...ue" : "http://vocab.getty.edu/aat/300198841" }, "Predicate":{ "type" : "uri", "value" : "http://www.w3.org/2000/01/rdf-schema#label" }, "Object":{ "xml:lang" : "en", "type" : "literal", "value" : "rhyton" } },{ "Subject":{ "type" : "uri", "value" : "http://vocab.getty.edu/aat/30019884..." "Predicate"

USING MULTIPLE VOCABULARIES

INTEROPERABILITY

Getty

Using Multiple Vocabularies

Multiple vocabularies are required

- Catalogers of art information will require multiple vocabularies

- No single vocabulary will provide the full set of terminology needed to catalog or index a given set of cultural heritage data

- A combination of vocabularies will be necessary for indexing

- Separate vocabularies may be required for retrieval

- Ideally, retrieval vocabularies would be based on indexing vocabularies, but may be optimized and applied differently for this different purpose

- Strategies for using vocabularies for indexing and retrieval will differ

Interoperability of Vocabularies

What is interoperability?

- *Interoperability* in the context of controlled vocabularies refers to the ability of two or more vocabularies and their systems or components of their systems to map to each other's data with the goal of exchanging information or enhancing discovery.
- Deals with the two conflicting demands, both require interoperability:
  - specialized vocabularies should be developed for a certain community, such as the art and cultural heritage community reflecting the specific terms and concepts needed by catalogers to index and classify that material; however, no single vocabulary will be comprehensive
  - end users who want to use a single search to find resources in federated settings across resources in different domains and created by different communities.

- Mappings between vocabularies may be used to facilitate faster indexing and better retrieval
  - When the user selects a term from the first vocabulary, the system can respond by offering corresponding terms from the second vocabulary

INTEROPERABILITY OF VOCABULARIES

Achieving interoperability

Interoperability may involve merging or adapting two or more controlled vocabularies to actually or virtually form a new controlled vocabulary that will combine all the concepts and terms contained in the originals. It could also involve merging or adapting two or more resources that have been indexed using different controlled vocabularies. Various methodologies for direct mapping and switching may be used.

- **Direct mapping:** Matching of terms one-to-one in each controlled vocabulary. The vocabularies need not be the same size (one may be smaller or larger) or cover exactly the same content, but there should be significant overlap in content. This technique assumes that where overlap exists, there will be the same meaning and level of specificity between the two terms in each controlled vocabulary. Mapping may be done by algorithm or using human mediation, but often both methods are employed together.

- **Switching:** Use of a third vocabulary, a *switching vocabulary*, that itself can link to terms in each of the two original controlled vocabularies. As with direct mapping, this type of mapping also assumes that the meaning of the terms can be reconciled, in this case between all three terms: the original two controlled vocabulary terms and one switching term. The advantage of this method is that the scope and format of the switching term may be made broad enough compensate for differences between the two original terms. Another application of switching occurs when the third vocabulary provides notations or a classification scheme under which terms from both original controlled vocabularies may be grouped. This approach enables a single, unifying hierarchical display for terms that originated in multiple sources.

Getty

Interoperability of Vocabularies
Factors for successful interoperability

- **Scope of mapping:** The greater the number of elements included in the mapping, the more difficult the mapping will become. At minimum, a mapping between vocabularies should match terms to terms.

- **Similarity of content:** The more there is similarity in the content of each of the vocabularies and of the resources being searched, the more likely it is that successful interoperability may be achieved. If the content is not similar, differences and variability in terminology, meaning, and syntax will hamper cross-domain interoperability.

- **Intended audience**: If the purposes or intended audiences of the resources or vocabularies are very different, mappings of vocabularies will be difficult or impossible and search results will be uneven.

- **Format and hierarchical structure:** The more there is similarity in the format and hierarchical structure of the vocabularies, the more likely interoperability between them will be successful.

- **Pre- and post-coordination:** Differences in the application of pre- and post-coordinated terminology in the vocabularies will complicate mapping efforts if one vocabulary contains headings while the other contains unique terms. A related issue concerns the differences in pre- and post-coordination expected in the search methodologies of the resources being searched.

- **Granularity and specificity:** The differences in degree of specificity or granularity of the controlled vocabularies themselves, and of the indexers' applications of the vocabularies in the target resources, may result in uneven results in indexing and retrieval. But if indexers of both resources have used the same vocabulary, use = relatively successful because the broader and narrower terms are logically linked.

- **Synonymy and near-synonymy:** Differences in how synonyms and near-synonyms are handled will affect the ability to make a successful mapping between vocabularies.

- **Authoritativeness:** If vocabularies differ in the level of authoritativeness by which they are developed, mapping them will be difficult.

Interoperability of Vocabularies

Semantic mapping

- A semantic network comprises relationships between terms and concepts based on their meanings or the nature of the relationships. The semantic relationships are sometimes derived from the vocabularies.
  - In other cases, they are extrapolated from the target content databases.

- A semantic network may be used to map terms from one or many controlled vocabularies according to a defined underlying organizational structure or conceptual scheme.
  - The relationships may range from a simple hierarchical structure with generic broader/narrower relationships to a more complex set of carefully defined relationships, such as contained in, agent for, process is, etc.
  - The relationships may be categorized to indicate the degree of closeness between linked terms, for example exact synonyms, near synonyms, closely related terms, loosely related terms, and antonyms.

Interoperability of Vocabularies

Interoperability across languages

- Multilingual controlled vocabularies are sometimes treated as a special case of interoperability.

- If unique vocabularies have been developed independently by groups in separate nations using different languages, utilizing the two together as a multilingual controlled vocabulary is generally not effective without extensive human intervention in the mapping process.

- This is due to the problems and idiosyncrasies of translation and usage of terms in various languages, which are not resolved with the simple employment of an automated dictionary or data mining.

- For translating vocabularies, see Harpring, *Guidelines for Multilingual Equivalency Work: Translating the Art & Architecture Thesaurus®* [anticipated publication date, 2019] and *ISO 25964-1-2:2011 & 2013: Information and documentation: Thesauri and interoperability with other vocabularies.*

Getty

Interoperability of Vocabularies

Issues of multilingual terminology

Exact equivalence: The most desirable match involves terms in each language that are identical, or nearly identical, in meaning and scope of usage in each language.

Inexact and partial equivalences: In cases where a suitable preferred term with the exact meaning and usage of the original term is not available in the second language, terms are sometimes linked as equivalents when they have only inexact or partial matches in scope and meaning.

Single-to-multiple term equivalence: If there is no match in scope and meaning between terms, sometimes a concept in one vocabulary is matched to multiple descriptors in the second language.

Non-equivalence: Sometimes there is no exact match, no term in the second language has partial or inexact equivalence, and there is no combination of descriptors in the second language that would approximate a match. The AAT would advise using a coined term, a literal translation, or a borrowing of the source-language term (without literary warrant) rather than leaving a gap in the target language.

Interoperability of Vocabularies

Dominant languages

- In a completely multilingual vocabulary, all languages would be treated equally, with none serving as a so-called dominant language.

- However, in practical applications, it is often necessary to treat one language as the default dominant language, particularly when the vocabulary is rich and complex.

  - An example is the AAT, in which each concept record has over 100 fields in addition to the term itself. With such vocabularies, it would be impractical to maintain not only the terms, but also the values of flags, notes, dates, hierarchies, and other subsidiary information in several languages.
  - For the AAT, English is the dominant language, although terms and the scope note may be in multiple languages.
  - In addition, if every term in the original source language has not been assigned equivalents in all other target languages, the status of the other languages is not equal to that of the source language, and they are known as secondary languages.
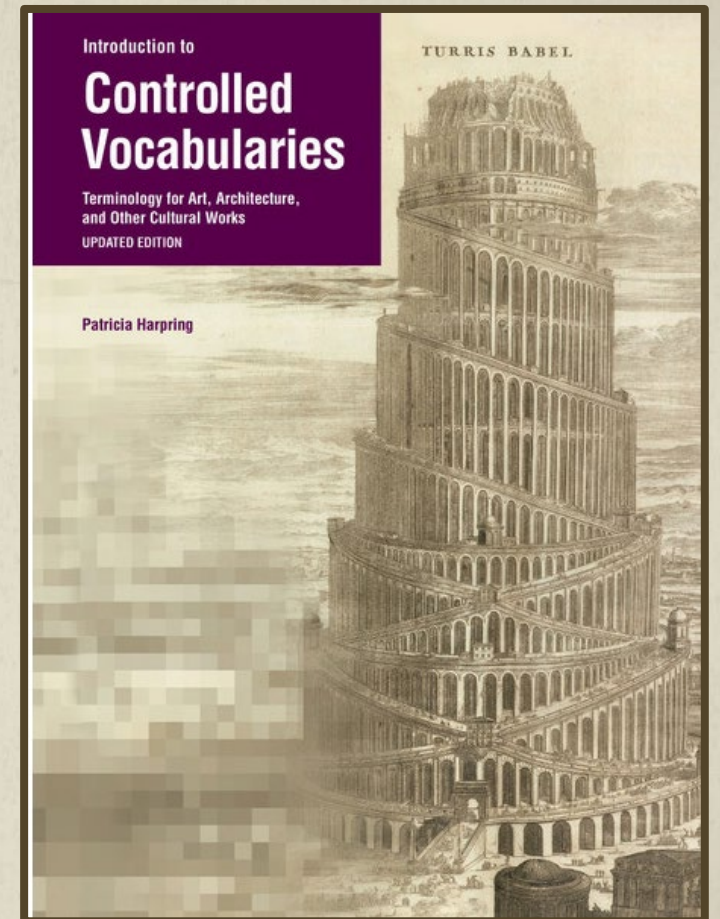
Getty

INTEROPERABILITY OF VOCABULARIES

Satellite and extension vocabularies

- Satellite and extension vocabularies may be considered *microcontrolled vocabularies* (also known as *microthesauri*), because they are specialized vocabularies that may be fit into the structure of a larger, broader, or more generic controlled vocabulary.

- Characterized by having been constructed with the goal of being interoperable with an existing vocabulary, e.g., a narrow specialty vocabulary that is intended to be linked to the superstructure of a larger vocabulary. The satellite may be linked at multiple points to the original vocabulary.

- Node or leaf linking is a closely related method that links a specialized vocabulary to a node in the hierarchical structure of a broader controlled vocabulary, so that the specialized vocabulary becomes a virtual new branch to the original vocabulary.

- With either approach, the resulting family of controlled vocabularies should be consistent in structure, term format, and editorial oversight.
    - By using satellite or extension vocabularies, specialized users may have access to the desired levels of specificity in the new controlled vocabulary without swamping the original controlled vocabulary with detail that may not be needed by most users.
    - It can allow a particular set of users to access only the specialized vocabulary terms that apply to their indexing needs, thus excluding the full original vocabulary from these users, while ensuring that their specialized terms are still compatible with the full vocabulary in retrieval.

Getty

Other topics covered in
Introduction to Controlled Vocabularies

- Local authorities
- Constructing a vocabulary
- Indexing with controlled vocabulary
- Retrieval using controlled vocabulary

Thank you.

Patricia Harpring
Managing Editor
Getty Vocabulary Program

pharpring@getty.edu
vocab@getty.edu
https://www.getty.edu/research/tools/vocabularies/