



AAT



TGN



ULAN



CONA

THE GETTY VOCABULARIES TECHNICAL UPDATE

Joan Cobb
Gregg Garcia
Information Technology Services
J. Paul Getty Trust

International Working Group Meetings
January 7-10, 2013

International Terminology Working Group



August 23-26, 2010

November 7-10, 2011



Monday

Overview of the technical aspects that apply to all four of the vocabularies.



AAT



TGN

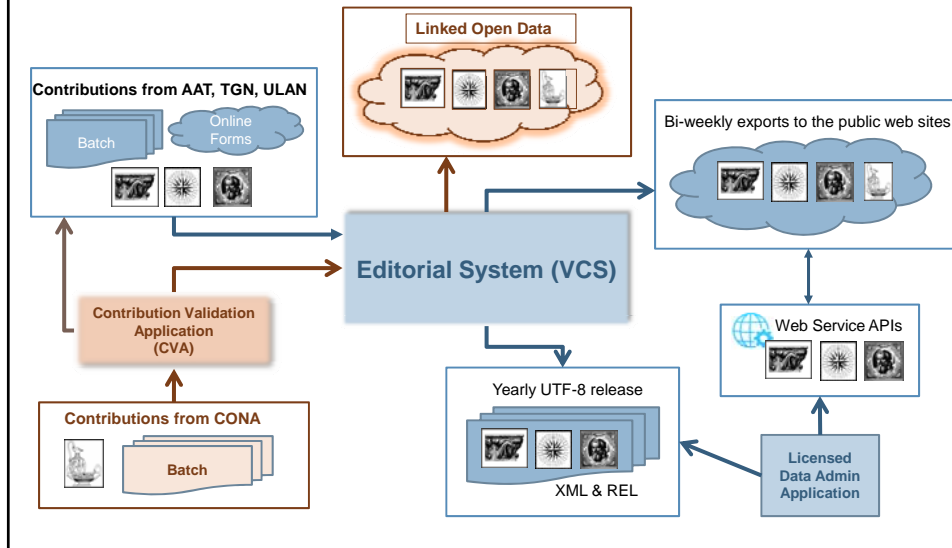


ULAN




CONA

Technical solutions



Japanese and Hebrew


Library of Congress Authorities online




Research Home » Tools » Union List of Artist Names » Full Record Display

Union List of Artist Names® Online Full Record Display

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

Click the  icon to view the hierarchy.

ID: 500060426 **Record Type:** Person

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

Note: He is famous for works of ukiyo-e ('pictures of the floating world') in painting and printmaking. He had a long career of over 70 years in diverse media and styles. He studied painting of Ming period (1368-1644) in China to the styles of the Kanō school, Suni-yoshi school, Rinpa painting, and his contemporaries of the Edo period (1600-1868) in Japan; he also learned from Western style painting. Object types and media included nihonkoku (polychrome or ink paintings); surimono ('printed things'; small-edition woodblock prints) and nishiki (polychrome prints); woodblocks for eihon (illustrated books) and kyōka ehon (illustrated books of poems called kyōka); and printed book illustrations for kibyōshi ('yellow cover' books, often moralizing tales and adventures) and yomihon ('reading books', sometimes historical novels). He was one of the main shunga (erotic picture) artists of the Edo period. Hokusai is thought to have made in all at least 30,000 drawings and the illustrations for 500 books. Since the late 19th century, his work has had a significant impact on Western artists, including Gauguin and van Gogh.

Names:

Hokusai, Katsushika (preferred,V,index)	
Katsushika Hokusai (V,display)	
Hokusai (V) name taken by the artist in 1798
葛飾 北斎 (V,Japanese,NA,U)	
Shunrō (V) gō (artist's name), used in his years of training, when painting hosoban (narrow prints)
Sōri (V) gō (artist's name), used in early career, named taken from his Rinpa-school master Tawaraya Sōri
Kakō (V)	
Tatsumasa (V)	
Cakyōjin (V)	
Taitō (V) name used since 1810, when creating illustrated picture books
Iichi (V)	
Manji (V)	
Tokitarō (V,BN)	
時太郎 (V,BN)	
葛飾北斎 (V)	
葛飾北斎 (V)	
葛飾北斎 (V)	
葛飾北斎 (V)	
葛飾北斎 (V)	
葛飾北斎 (V)	
葛飾北斎 (V)	
葛飾北斎 (V)	
近藤市太郎 (V)	
гогош (V)	
гогош ,апретовск (V)	

Chinese



rK D

Dutch

Spanish





Research Home » Tools » Art & Architecture Thesaurus » Full Record Display

Art & Architecture Thesaurus® Online Full Record Display

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

Click the  icon to view the hierarchy.

ID: 300014916 **Record Type:** concept

 **lacquer (coating)** (<coating by composition or origin>, coating (material), ... Materials (Hierarchy Name))

Note: Term applied broadly to paint or varnish that dries with a high gloss, to a coating made from shellac, or to any coating that dries by the evaporation of solvent. In most lacquers resin is dissolved in a volatile solvent or a drying oil and hardens by evaporation.

Terms:

lacquer (coating) (preferred,C,U,English-P,U,U,N)	
lacker (C,U,English,U,F,U,N)	
漆 (C,U,Chinese (traditional)) P,D,U,U)	
漆器 (C,U,Chinese (traditional)) AD,U,U)	
漆塗 (C,U,Chinese (traditional)) AD,U,U)	
人然漆 (C,U,Chinese (traditional)) AD,U,U)	
大漆 (C,U,Chinese (traditional)) AD,U,U)	
qi (C,U,Chinese (untransliterated Hanyu Pinyin)) P,D,U,U)	
qi (C,U,Chinese (transliterated Pinyin without tones)) P,D,U,U)	
chī (C,U,Chinese (transliterated Wade-Giles)) P,D,U,U)	
lak (C,U,Dutch-P,D,U,U)	
laque (C,U,French-P,D,U,N)	
lacca (vernice) (C,U,Italian-P,D,U,N)	
laca (barniz) (C,U,Spanish-P,U,U,N)	

Additional Notes:

Chinese (traditional) 此漆最廣通用以製成漆器其色澤光亮而耐久，由漆樹液所製成，由各種材料製成之漆器，在人多數的場合中，指點漆器在藝術性的設計或製造中，然後漆器被製成之漆。

Dutch Meestere term voor verf of vernis die met een hoogglans opdroogt en voor schellak of redere andere deklaag die droogt door verdamping van het oplosmiddel. In de meeste laksoorten is hars opgelost in een vluchtig oplosmiddel of drogende olie, waarbij het hars verhardt door verdamping.


Spanish Término aplicado originalmente a la pintura o barniz que se seca y endurece considerablemente, para el revestimiento fabricado con goma laca o para cualquier revestimiento que se endurece por evaporación de su solvente.

Ability to Associate Images


Research
 Research Home • Tools • Art & Architecture Thesaurus • Full Record Display

Art & Architecture Thesaurus® Online
 Full Record Display

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

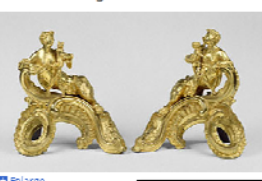
Click the  icon to view the hierarchy.

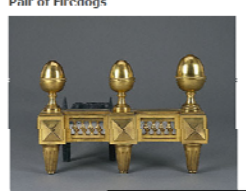
ID: 300161848


 **andirons** (<fire-making and controlling equipment> (Hierarchy Name))


Note: Devices with a horizontal bar on a vertical fireplace.

Terms:
andirons (preferred, C,U,I,C,English-P,D,U,PN)
andiron (C,U,English,AD,U,SN)
fire dog (C,U,English,UF,U,N)
fire dogs (C,U,English,UF,U,N)
fire-dogs (C,U,English,UF,U,N)
fire dogs (C,U,English,UF,U,N)
dogs, fire (C,U,English,UF,U,N)
landiers (C,U,English,UF,U,N)
haardijzers (C,U,Dutch-P,D,U,U)
haardijzer (C,U,Dutch,AD,U,U)
chenets (C,U,French-P,D,U,PN)
chenet (C,U,Spanish,UF,U,SN)
 (French,AD,U,SN)
landier (C,U,French,UF,U,N)
morillos (C,U,Spanish-P,D,U,PN)
morillo (C,U,Spanish,AD,U,SN)

Pair of Firedogs

[Enlarge](#)
 Charles-Louis
 French, Paris, about 1730
 Gilt bronze

Pair of Firedogs

[Enlarge](#)
 Unknown
 French, Paris, about 1730
 Gilt bronze, silver tool

Pair of Firedogs

[Enlarge](#)
 Unknown
 French, Paris, about 1700
 Gilt bronze

Pair of firedogs

[Enlarge](#)
 Call of St. Paul Chilly
 Attributed to Pierre Goussier
 French, Paris, about 1770
 Gilt bronze

Multiple Parents



ID: 300011121

Record Type: concept

 **jadeite** (mineral, inorganic material, ... Materials (Hierarchy Name))

Note: Mineral consisting of sodium aluminum silicate, usually containing small amounts of iron, calcium, and magnesium. It constitutes the more valuable variety of jade, ranging in color from pale to dark green that is commonly called jade; nephrite is the other mineral called "jade." Jadeite is highly valued for its translucent glassy appearance; it is found chiefly in China, Myanmar (formerly Burma), and Mexico.

Terms:

jadeite (preferred, C,U,I,C,English-P,D,U,PN)
imperial jade (C,U,English,UF,U,N)
 翠玉 (C,U,Chinese (traditional)-P,D,U,U)
 翠玉 (C,U,Chinese (traditional),AD,U,U)
 翡翠 (C,U,Chinese (traditional),AD,U,U)
ying yu (C,U,Chinese (transliterated Hanyu Pinyin)-P,D,U,U)
ying yu (C,U,Chinese (transliterated Pinyin without tones)-P,D,U,U)
ying yu (C,U,Chinese (transliterated Wade-Giles)-P,D,U,U)
jadeiet (C,U,Dutch-P,D,U,U)
jadeiet (C,U,Dutch,UF,U,N)
jadeite (C,U,French-P,D,U,PN)
jadeit (C,U,German-P,D,U,PN)
jadeita (C,U,Spanish-P,D,U,U)

Facet/Hierarchy Code: MMT

Hierarchical Position:

 Materials Facet
 ... Materials (Hierarchy Name) (G)
 ... materials (matter) (G)
 ... <materials by composition> (G)
 ... inorganic material (G)
 ... mineral (G)
 ... jadeite (G)

Additional Parents:

 Materials Facet
 ... Materials (Hierarchy Name) (G)
 ... materials (matter) (G)
 ... <materials by composition> (G)
 ... inorganic material (G)
 ... rock (G)
 ... <rock by form> (G)
 ... stone (rock) (G)
 ... <stone by quality> (G)
 ... gemstone (G)
 ... <gemstones by specific type> (G)
 ... jadeite (G)

 Materials Facet
 ... Materials (Hierarchy Name) (G)
 ... materials (matter) (G)
 ... <materials by composition> (G)
 ... inorganic material (G)
 ... rock (G)
 ... metamorphic rock (G)
 ... jade (rock) (G)
 ... jadeite (G)

mineral

gemstone

rock

Editorial System
Subject Search



Editorial System - Hierarchy View

Some can be merged automatically during load, others require editorial intervention.

Stat	Mark	Descriptor (Qualifier)	ID
FN		Top of the AAT hierarchies	300000000
IP	+	Materials Facet	300264091
IP	+	Materials (Hierarchy Name)	300010357
IP	+	materials (matter)	300010358
IP		<materials by function>	300014692
IP	+	abrasive	300014693
IP	+	additive (material)	300014701
IP	+	adhesive	300014801
AC		antidotes	300379033
IP	+	antistatic agent	300080663
IP	+	artists' materials	300014842
IP	+	base color	300257037
IP	+	binder (material)	300014720
IP	+	<biological control agents>	300014846
		biomarkers	300379608
		bitstone	300011146
		bleach	300014855
		bluing	300263384
		building materials	300014857
		catalyst	300014902
		chelating agent	300379664
		chrom	300265211
		cleaning compound	300014804
		coating (material)	300014807

Editorial System Top part of a full record data entry screen

Subject ID: 300014916 Review: In Process Merged: Merged Publish: Published
 Parent: 300014908 Record: Concept Candidate: Non Candidate
 Label: lacquer (coating by composition or origin, coating, materials by function, materials, Materials, Materials Facet)

Scope Notes

Note Language: English Term applied broadly to paint or varnish that dries with a high gloss, to a coating made from shellac, or to any coating that dries by the evaporation of solvent. In most lacquers resin is dissolved in a volatile solvent or a drying oil and hardens by evaporation.

Terms/Names

#	P	H	V	Terms/Names	Display Dt	Start Dt	End Dt	Term ID	Disp Nm	AACR2	Flag	Oth
1	P	C	U	lacquer				1000014916	N/A	N/A	N/A	
2	V	C	U	lacker				1000266994	N/A	N/A	N/A	
3	V	C	U	漆				1000523200	N/A	N/A	N/A	
4	V	C	U	塗漆				1000523201	N/A	N/A	N/A	

Term Contrib(s) P

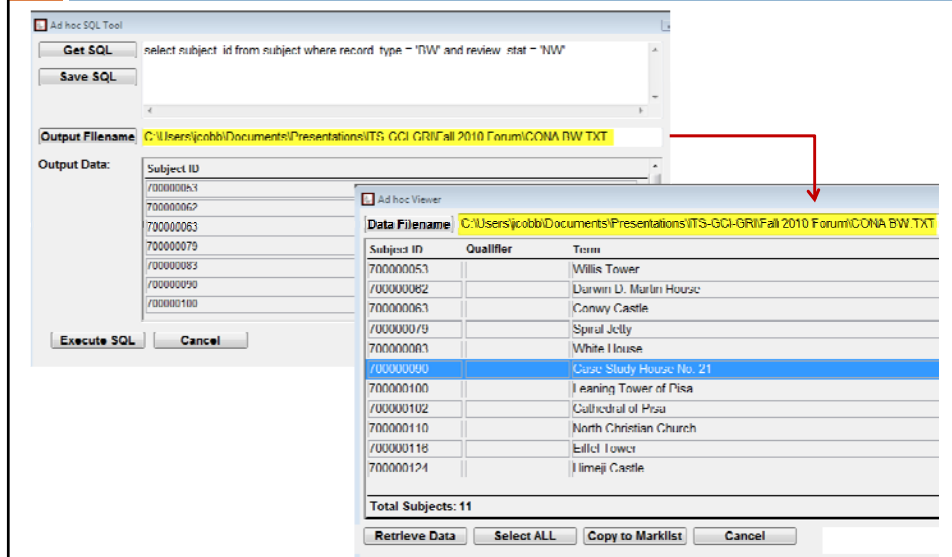
Language(s)	Code	P	Term Type	Qualifier	Term Source(s)	Page			
COBP-DIBAM	N		English	70051 P	Descriptor	coating	AATA database (2002.)	125406	chec
GCI	P						Avery Index (1963.)		
VP	P						Brady and Clauser, Materials Handbook (1986)	430	

Associative Relationship

Relationship Type/Code	Entity	Related Entity/Term	Display Dt	Start Dt	End Dt	H	Subject ID	Candidate
2415 - activity/event needing	Concept	lacquering					C 300053796	N
2501 - material used/product	Concept	lacquerware					C 300264021	N

Editorial System

Ad hock query tool



Batch loading of data

- Information about how to contribute is available on our Web site
 - <http://www.getty.edu/research/tools/vocabularies/contribute.html>
- Bulk contributions
 - ▣ Can range from a few thousand to several million records
 - ▣ Every new contribution requires close coordination
- Online Web candidate contribution forms
 - ▣ URL: vocabcontrib.getty.edu
 - ▣ Uses the same XML schema as bulk contributions
 - ▣ Harvested on the 1st of every month

Contribution XML schemas

- Basic Schemas
 - ▣ Getty_Vocabulary_AAT.xsd
 - ▣ Getty_Vocabulary_TGN.xsd
 - ▣ Getty_Vocabulary_ULAN.xsd
 - ▣ Getty_Vocabulary_CONA.xsd
- Schemas containing specific enumeration values are updated every month
 - ▣ AAT : associative_type, contrib, language, lookup_lists, source
 - ▣ TGN: associative_type, contrib, language, lookup_lists, source plus place_type
 - ▣ ULAN: associative_type, contrib, language, lookup_lists, source plus event, nationality, place, and role
 - ▣ CONA: associative_type, contrib, language, lookup_lists, source plus event, nationality, role, and class

Online contribution forms for AAT, ULAN & TGN

Research Home » Conducting Research » Learn about the Getty Vocabularies » Candidate contribution form

Art & Architecture Thesaurus® Online
Candidate contribution form

Save Draft Open Draft Clear Form Submit Candidate Log Out Search AAT

All fields marked with an (*) are required. Click on field name to see rules. NOTE: You can not edit the record after candidate is submitted.
Draft in use: **None**

CONTRIBUTOR INFORMATION

Contributing Institution *

Contributor ID *

Editor Name

Editor Email

New or Comment

Record Type

Preferred Parent

Additional Parent

TERMS

1. Preferred Term *

2. Variant Term *

Contributor ID is assigned by the Vocabulary Program

Loading the data

- Programs to load the data are very complex
 - ▣ Several rounds of testing are necessary
 - ▣ Actual load takes from several hours to several days



- Merging
 - ▣ Rules are different for each vocabulary
 - ▣ Need to accommodate differing standards
 - ▣ Some of the merge routines can be skipped with translations because the matching records are already known
- Unicode
 - ▣ Contributions can be submitted in any encoding

Loading the data

Multi-lingual data

- Worked through a number of issues related to ordering and classifying terms in different languages
- Adjusted the data structure to make the term qualifier, term type and part of speech language specific
- Modified the rules combining terms
- Updated the way the information is presented to users
- Changes to the core editorial system require changes to everything else

Yearly Exports

Sample Data available in UTF-8

XML

Relational Tables

Download sample records from the Art & Architecture Thesaurus (AAT). The full AAT is a hierarchical vocabulary of around 34,000 records, including 134,000 terms, descriptions, bibliographic citations, and other information relating to fine art, architecture, decorative arts, archival materials, archaeology, and other material culture. The full XML data file for the AAT is 20,425 kilobytes in size. The sample available here is a small subset of the AAT. The next data release will take place in July 2012. For further information, see [About the AAT](#).

AAT Sample Data
Format options:
 XML UTF-8 (2.38 MB)
 Relational Table UTF-8 (40.4 MB)

Data Dictionary for the AAT Data Release
 XML format PDF
 Relational Table format PDF

Download sample records from the Union List of Artist Names (ULAN). The full ULAN is a vocabulary of around 162,000 records, including 453,000 names and biographical and bibliographic information for artists, architects, firms, shops, and art repositories, including a wealth of variant names, pseudonyms, and language variants. The full XML data file for the ULAN is 97,383 kilobytes in size. The sample available here is a small subset of the ULAN. The next data release will take place in July 2012. For further information, see [About the ULAN](#).

ULAN Sample Data
Format options:
 XML UTF-8 (14 MB)
 Relational Table UTF-8 (92 MB)

Data Dictionary for the ULAN Data Release
 XML format PDF
 Relational Table format PDF

Download sample records from the Getty Thesaurus of Geographic Names (TGN). The full TGN is a hierarchical vocabulary of around 895,000 records, including 1.1 million names, place types, coordinates, and descriptive notes, focusing on places important for the study of art and architecture. TGN is not a GIS; coordinates are included for many records, but they are for finding purposes only (i.e., to find the place on a map). The full XML data file for the TGN is 133,312 kilobytes in size. The sample available here is a small subset of the TGN. The next data release will take place in July 2012. For further information, see [About the TGN](#).

TGN Sample Data
Format options:
 XML UTF-8 (1.10 MB)
 Relational Table UTF-8 (7.86 MB)

Data Dictionary for the TGN Data Release
 XML format PDF
 Relational Table format PDF

<http://www.getty.edu/research/tools/vocabularies/obtain/download.html>

Web Services - User's Instructions

<http://www.getty.edu/research/tools/vocabularies/obtain/download.html>

Instructions include:
Examples
Data Element Descriptions

The following information and examples may guide those who wish to explore licensing options for the AAT, ULAN, or TGN. (CONA is in development and not yet available for licensing.) The Getty vocabularies change and grow over time; thousands of records are added or changed each year. Implementers are encouraged to refresh the data over time.

Licensing: The Getty vocabulary data is copyrighted by the J. Paul Getty Trust. All rights reserved. Your institution must sign a license and pay the required fees in order to use the data. The data is compiled from various contributors using published sources, which must be cited along with the J. Paul Getty Trust when the data is displayed.

The data files are released annually, in July or slightly earlier each year (barring unforeseen technical difficulties). If you wish to license multiple vocabularies, you must sign a separate license for each one. Licensing agreements are for a fixed five-year term upon payment of a fee. Contact us at vocab@getty.edu, subject line: "to learn the amount of the fees and terms of licenses. Please include an action of how you intend to use the data and whether your institution is for-profit or not-for-profit."

Formats of the releases: The vocabulary data is released in XML UTF-8 and Relational Table formats. Currently, AAT and ULAN are also released in MARC format. The XML and Relational Table formats are available in both UTF-8 and ASCII. The ASCII files use our six digit codes. Note: Both the MARC format and the ASCII releases will be phased out in one or two years, when all three vocabularies will be released only in XML UTF-8 and Relational Tables.

The data is released in data files only. No software is provided. Only full sets of data are available for licensing; we do not provide subsets of the data.

Web services: Data files may also be accessed through Web services APIs (application programming interfaces), which provide access to data that is refreshed every two weeks. The APIs are available to any institution that holds a valid license. Please read details in the Web Services User's Instructions (PDF).

Changes to the data structure: Changes have been made in the structure of all three vocabularies to allow for the seamless integration of multilingual terms. One significant change is that the Qualifier field is now linked to language (which has multiple occurrences per term), to better accommodate qualifiers in multilingual data. Typically,

Related Sections

- Learn about the Getty Vocabularies
- Obtain the Getty Vocabularies
- Contribute to the Getty Vocabularies
- Frequently Asked Questions
- Editorial Guidelines
- Training Materials for Vocabularies and Standards
- Email Vocabulary Program

Additional Information

- To view and print Portable Document Format (PDF) you will need Adobe Reader.

Web Services

Access to basic information

- Information that is available on our public web site is also available via our Web Services
 - ▣ Term Match – uses the same criteria
 - AAT – Term and Note
 - TGN – Name, Place Type, Nation
 - ULAN – Name, Role, Nationality
 - ▣ Full record for a particular subject ID
 - ▣ Preferred and non-preferred parent hierarchies for a particular subject ID
 - ▣ Immediate preferred and non-preferred children for a particular subject ID

Web Service

Ability to synchronize

- List of records that have been merged along with their old and new IDs
- Current ID for a given subject ID, accounting for whether or not a record has been merged
- List of records that have been modified based on a date range and one of the following parameters
 - ▣ Overall subject record edits
 - ▣ Added, deleted, modified terms
 - ▣ Scope note edits
 - ▣ Moved records
 - ▣ New records

Tuesday



CONA

Overview of the technical aspects
that are particular to the
Cultural Objects Name Authority®

Editorial System

Full record data entry screen

Subject Edit - Term:Hagia Sophia, S:700000014, P:700000010

Subject ID: 700000014 Review: New VP Merged: Not Merged Publish: Not Published
 Parent: 700000010 Record: Built Work Candidate: Non Candidate
 Label: Hagia Sophia (temp.parent/test Built Work)

Scope Notes

Note Language: English
 Nothing remains of the earlier two basilicas on this site: a 4th-century church and a 5th-century one built by Constantius, the son of Constantine the Great, which burned during the Nika riots of 532. The current church was rebuilt under the personal supervision of emperor Justinian I; it was rededicated on 27 December 537. Isidore of Miletus and Anthemius of Tralles, who were scientists and mechanical engineers but not architects per se, were chosen by

Contribs

VP

Sources

Fletcher, History of Architecture

Terms/Names

#	P	H	V	Terms/Names	Display Dt	Start Dt	End Dt	Term ID	Disp Nm	AACR2 Flag	Oth
1		C	V	Hagia Sophia				1000000025	N/A	Yes	desc
2	V	C	V	Holy Wisdom				1000000027	N/A	N/A	trans
3	V	C	V	Ayasofya				1000000028	N/A	N/A	N/A
4	V	C	V	Santa Sofia				1000000029	N/A	N/A	N/A

Term Contrib(s) P

VP P

Language(s) Code P

English 70051 P

Term Type Qualifier

Undetermined

Term Source(s)

Encyclopedia Britannica Online (2002-)

Fletcher, History of Architecture (1996)

Page

CONA-VCS: Associate an image

Subject Edit - Term: Case Study House No. 21, 2006.M.30, S:700000120, P:700000121

Subject ID: 700000120 Review: New VP Merged: Not Merged Publish: Not Published
 Parent: 700000121 Record: Movable Work Candidate: Non Candidate
 Label: Case Study House No. 21, 2006.M.30 (Pierre Koenig papers and drawings, 1925-2007, Movable Work)

WorkType/Class/Level Creator Display Related Person/Corp Creation Date Location Physical Characteristics


Person/Corp Label	ULAN ID	Role	Extent
Koenig, Pierre (American architect, active in Los Angeles, 1902-1982)	500096720	architect/30018	

Depicted Subject Outside Iconography Culture/Style Other Displays Events Subject Continuation/Sources Notes Media Adult Subject Index

Order File Name Copyright Info Caption Info

1	Koenig drawing 1.jpg (P) J. Paul Getty Trust. Used with permission. P Pierre Koenig (American, 1902-2007)		
---	---	--	--

Order Image URL [Go](#)



VCS: Associative Relationships

Subject Edit - Term: Case Study House No. 21, S:700000090, P:700000001

Subject ID: 700000090 Review: New VP Merged: Not Merged Publish: Not Published
 Parent: 700000001 Record: Built Work
 Label: Case Study House No. 21 (Built Work)

Note Language Case Study Houses were commissioned by Arts & Architecture Architecture new building materials, technology the middle class. Case Study No. 21 architect to be his masterpiece. The design in an open-plan steel-and-house is surrounded by five reflecting pools.

Note Language English

Scope Notes

The drawing is an axonometric projection bird's-eye view created as a presentation drawing for Koenig's Case Study No. 21 (Los Angeles, California). The point of view is facing southeast, showing the house and certain elements of the surrounding landscaping.

Contribs VP

Sources CDWA online (1995-2010) Getty Research Institute

Titles/Names

#	P	H	V	Terms/Names	Display Dt	Start Dt	End Dt	Term ID	Disp Nm	AAC
1	P	C	U	Case Study House No. 21				1000000263	N/A	N/A
2	V	C	U	Bailey House				1000000264	N/A	N/A
3	V	C	U	CSH #21						

Term Contrib(s) P Language(s) Code P

VP	P	English	70051	P
----	---	---------	-------	---

Associative Relationship

Relationship Type/Code	Entity	Related Entity/Term	Display Dt	Start Dt	End Dt	H	Subject ID
4325 - depicts	Built Work	Case Study House No. 21				C	700000090

Relationship Type/Code Entity Related Entity/Term Display Dt Start Dt End Dt H Subject ID

4326 - depicted in	Movable Work	Case Study House No. 21				C	700000120
--------------------	--------------	-------------------------	--	--	--	---	-----------

CONA-VCS: Live links to AAT, TGN & ULAN

Subject ID: 700000120 Review: New VP Merged: Not Merged Publish: Not Published
 Parent: 700000121 Record: Movable Work Candidate: Non Candidate
 Label: Basic Study House No. 21, 2006 M.30 (Pierre Koenig papers and drawings, 1925-2007, Movable Work)

Work Type/Class/Level Creator/Depictor Related Person/Corp Creation Date Location Physical Characteristics

Person/Corp Label	ULAN ID	Role	Extent
Koenig, Pierre (American architect, active in Los A)	500006520	architect/30010	

Depicted Subject Outside Iconography Culture/Style Other Displays Events Subject Contribs/Sources Notes Media Add Subject Fields

General Subject

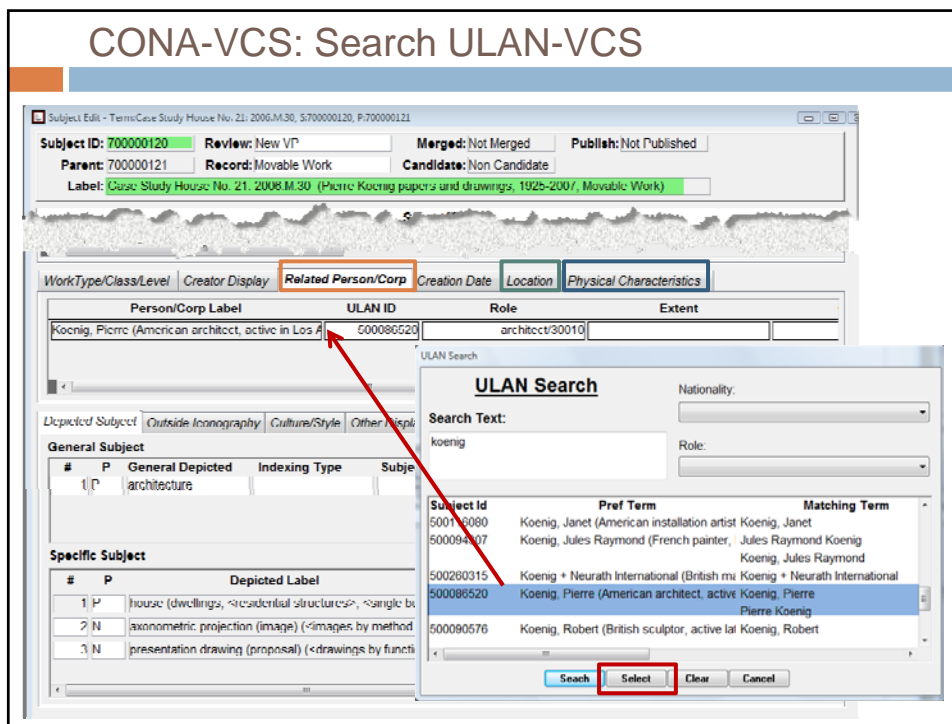
#	P	General Depicted	Indexing Type	Subject Extent
1	P	architecture		

Specific Subject

#	P	Depicted Label	CONA ID	ULAN ID	TGN ID	AAT ID	Iconogr
1	P	house (dwellings, <residential structures>, <single bui				300005133	
2	N	axonometric projection (image) (<images by method c				300034021	
3	N	presentation drawing (proposal) (<drawings by funcio				300034461	

Examples of New Web Services for In-house Reports and CONA CVA

- **GetChildCount** – returns an indexing alpha character along with the count of the number of subject records with a preferred term starting with that letter
- **GetChildrenReports** – returns basic child record information given parent ID, alpha index character, language code and candidate status
- **GetMergedSubjectIDsReports** – returns the count, subject IDs and merge dates of all records merged into the given current subject ID
- **GetMergedSubjectReports** – returns full record information for a merged subject record
- **GetParentHierarchy** – returns total parent levels along with pertinent parent/child information for each level
- **GetSourcesReports** – returns full source information based on the date of edits or creation, alpha index, and editor
- **GetSubjectBriefReports** – returns a limited set of full record data for a current subject ID
- **GetSubjectIDsDiacritics** – returns a list of subject IDs given a comma delimited list of Getty diacritic codes; also allows user to input a hierarchy parent ID and candidate status
- **GetSubjectIDsReports** – returns a list of subject IDs given a broad range of criteria: creation/edit date, alpha indexing character, editor, candidate status, parent hierarchy ID, and contributor
- **GetSubjectReports** – returns full record information given a current subject ID
- **CreateSubject** – creates a candidate record for a given preferred term and parent ID
- **TGNGetSubjectIDsCoordinates** – returns a list of subject IDs given a set of coordinate points, parent key and candidate status
- **CONAIconographyMatch** – term search matching service for the Iconography module within CONA
- **CONACreateIconography** – creates a candidate Iconography record given a preferred term



Loading the data

- Programs to load the data are very complex
 - ▣ Several rounds of testing are necessary
 - ▣ Actual load takes from several hours to several days



- Merging
 - ▣ Rules are different for each vocabulary
 - ▣ Need to accommodate differing standards
 - ▣ Some of the merge routines can be skipped with translations because the matching records are already known
- Unicode
 - ▣ Contributions can be submitted in any encoding

A CONA contribution may also be a contribution to AAT, TGN & ULAN

The screenshot shows a record in a database with the following details:

- Subject ID:** 700000120
- Review:** New VP
- Merged:** Not Merged
- Publish:** Not Published
- Parent:** 700000121
- Record:** Movable Work
- Candidate:** Non Candidate
- Label:** Case Study House No. 21: 2006 M.30 (Pierre Koenig papers and drawings, 1925-2007, Movable Work)

Below the record, there are several tabs and tables:

- Work Type/Class/Level:** Creator, Deleter, Related Person/Corp, Creation Date, Location, Physical Characteristics
- Person/Corp Label:** Koenig, Pierre (American architect, active in Los A)
- ULAN ID:** 500006520
- Role:** architect/30010
- Extent:**
- General Subject:**

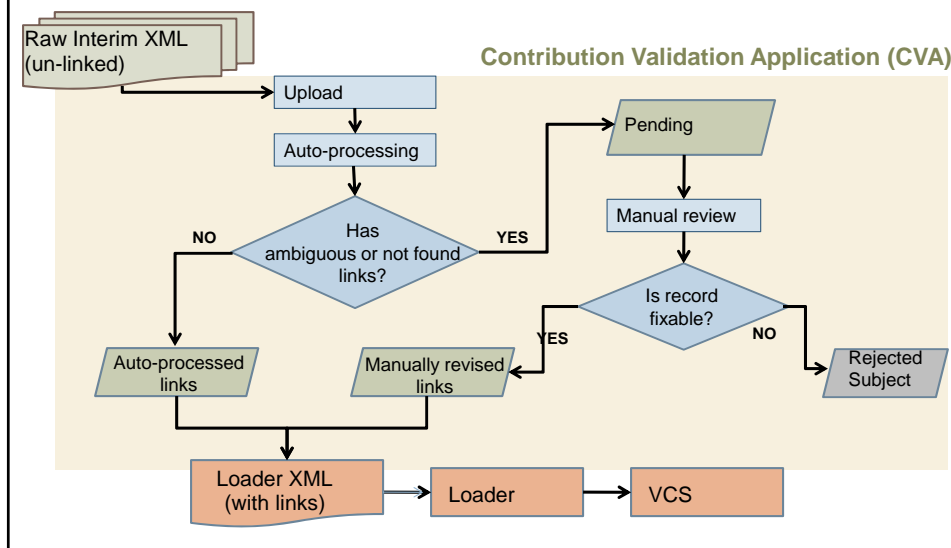
#	P	General Depicted	Indexing Type	Subject Extent
1	P	architecture		
- Specific Subject:**

#	P	Depicted Label	CONA ID	ULAN ID	TGN ID	AAT ID	Iconogr
1	P	house (dwellings, <residential structures>, <single bui				300005133	
2	N	axonometric projection (image) (<images by method c				300034021	
3	N	presentation drawing (proposal) (<drawings by funcio				300034461	

Arrows on the left indicate connections to external vocabularies:

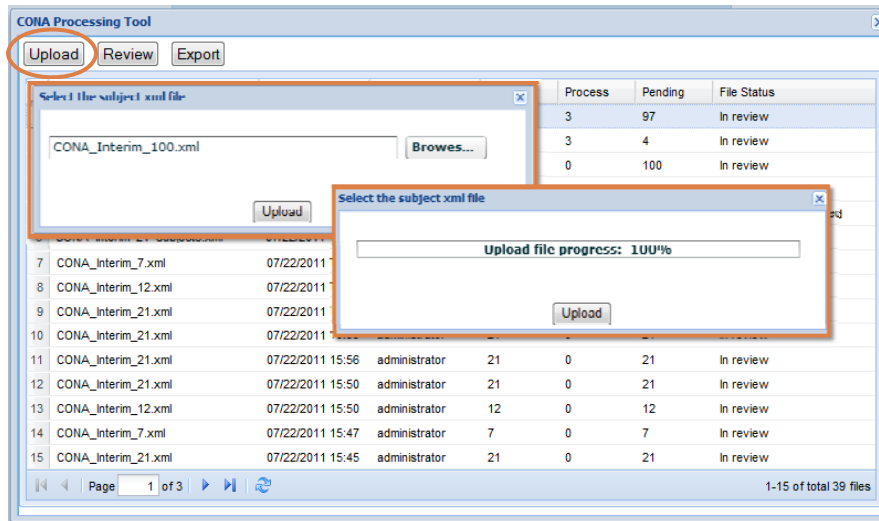
- ULAN:** Points to the 'ULAN ID' field.
- TGN:** Points to the 'TGN ID' field.
- AAT:** Points to the 'AAT ID' field.
- CONA:** Points to the 'CONA ID' field.

Loading CONA data



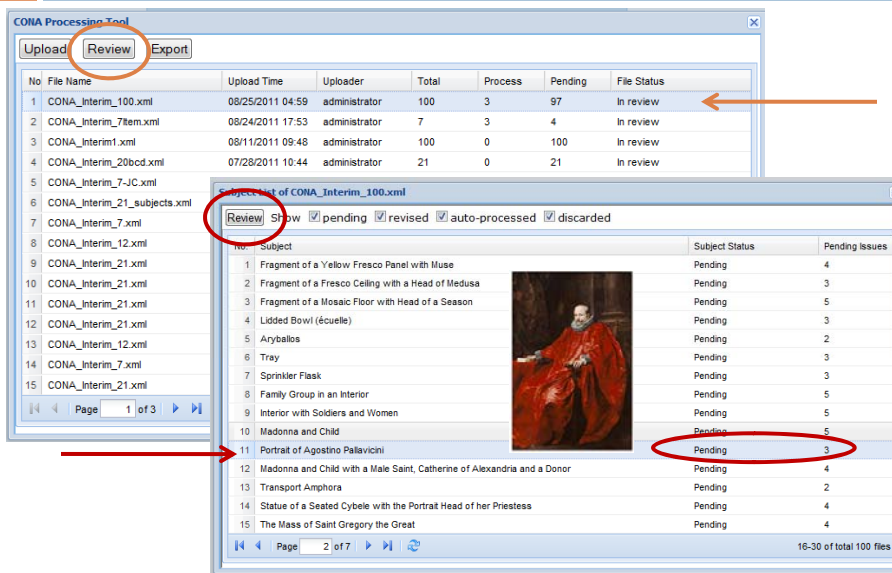
CONA from CVA to Publication

Upload Raw Interim XML file



CONA from CVA to Publication

Manually review the list



CONA from CVA to Publication

Review the record

Editors can define candidate records for AAT, TGN and ULAN as needed.

Editors can define rules for batch processing of the set as needed.

CONA from CVA to Publication

Approve the record

Tip: Approve Successful

Subject List of CONA_Interim_100.xml

No.	Subject	Subject Status	Pending Issues
1	Fragment of a Yellow Fresco Panel with Muse	Pending	4
2	Fragment of a Fresco Ceiling with a Head of Medusa	Pending	3
3	Fragment of a Mosaic Floor with Head of a Season	Pending	5
4	Lidded Bowl (écuelle)	Pending	3
5	Arybalos	Pending	2
6	Tray	Pending	3
7	Sprinkler Flask	Pending	3
8	Family Group in an Interior	Pending	5
9	Interior with Soldiers and Women	Pending	5
10	Madonna and Child	Pending	0
11	Portrait of Agostino Pallavicini	Revised	0
12	Madonna and Child with a Male Saint, Catherine of Alexandria and a Donor	Pending	4
13	Transport Amphora	Pending	2
14	Statue of a Seated Cybele with the Portrait Head of her Priestess	Pending	4
15	The Mass of Saint Gregory the Great	Pending	4

Page: 2 of 7

16-30 of total 100 files

CONA from CVA to Publication

Ready to load into our editorial system

CONA Processing Tool

Upload Review **Export**

No	File Name	Upload Time	Uploader	Total	Process	Pending	File Status
1	CONA_interim_100.xml	08/25/2011 01:59	administrator	100	100	0	All subjects processed
2	CONA_interim_780m.xml	08/24/2011 17:53	administrator	7	3	4	In review
3	CONA_interim_7.xml	08/11/2011 09:48	administrator	100	0	100	In review
4	CONA_interim_710bcd.xml	07/18/2011 10:44	administrator	21	0	21	In review
5	CONA_interim_7-JC.xml	07/20/2011 00:52	administrator	7	0	7	All subjects processed
6	CONA_interim_21_subjects.xml	07/22/2011 18:32	administrator	21	0	21	In review
7	CONA_interim_7.xml	07/22/2011 18:18	administrator	7	0	7	In review
8	CONA_interim_12.xml	07/22/2011 17:11	administrator	12	0	12	In review
9	CONA_interim_21.xml	07/22/2011 16:34	administrator	21	0	21	In review
10	CONA_interim_21.xml	07/22/2011 16:33	administrator	21	0	21	In review
11	CONA_interim_21.xml	07/22/2011 15:56	administrator	21	0	21	In review
12	CONA_interim_21.xml	07/22/2011 15:50	administrator	12	0	12	In review
13	CONA_interim_21.xml	07/22/2011 15:47	administrator	7	0	7	In review
14	CONA_interim_21.xml	07/22/2011 15:45	administrator	21	0	21	In review

1-15 of 159 files

CONA XML with links

AAT Stubs **TGN Stubs** **ULAN Stubs**

CONA from CVA to Publication

Edit the record in the editorial system if necessary

Subject Edit - Term: Portrait of Agostino Pallavicini, S:700000190, P:700000002

Subject ID: 700000190 Review: **Accepted Candidate** Merged: Not Merged Publish:

Parent: 700000002 Record: Movable Work Candidate: Non Candidate

Label: Portrait of Agostino Pallavicini (painting (visual work); Anthony van Dyck (Flemish, 1599 - 1641) (Flemish Artist);

Scope Notes

Note Language	Contribs	Sources	Page

Titles/Names

#	P	H	V	Terms/Names	Display Dt	Start Dt	End Dt
1	P	C	U	Portrait of Agostino Pallavicini			

Term Contrib(s) P

Term	Contrib(s)	P	Language(s)	Code	P	Term Type	Qualifier	Term Source(s)	Page
JPGM	P		English	70051	P	N/A		J. Paul Getty Museum database for collections	

Associative Relationship

CONA from CVA to Publication

The links established in CVA are already there

Subject ID: 700000190 Review: Accepted Candidate Merged: Not Merged Publish: ☐

Parent: /00000000/ Record: Movable Work Candidate: Non Candidate

Label: Portrait of Agostino Pallavicini (painting (visual work); Anthony van Dyck (fl. 1599 - 1641) (fl. 1599-1641))

Titles/Names

#	P	H	V	Terms/Names	Display Dt
1	P	C	U	Portrait of Agostino Pallavicini	

WorkType/Class/Level Greater Display Related Person/Corp Creation Date Location Physical Characteristics

Materials:

Display Materials: Oil on canvas

Term	AAT ID	Role	Flag	Ma
oil paint (paint) (spread by composition or t)	700015104	medium	material	100
canvas (textile materials by process or id)	300014078	support	material	300

Dimensions:

Display Dimensions: Unframed: 216.2 x 141 cm (85 1/8 x 55 1/2 in.); Framed: 264.2 x 185.1 x 10.2 cm (104 x 72 7/8 x 4 in.)

Value	Unit	Type	Extent	Qualifier	Scale Type
216.2	centimeters	height	undetermined	undetermined	undetermined
140.8703	centimeters	width	undetermined	undetermined	undetermined
10.16	centimeters	depth	undetermined	framed	undetermined

Depicted Subject Outside Iconography Culture/Style Other Displays Events Subject Credits/Sources Titles Med

General Subject

#	P	General Depicted	Indexing Type	Subject Extent
1	P	portrait		

CONA from CVA to Publication

Records are published bi-weekly to the web

Research

Research Home » Tools » Cultural Objects Names Authority » Full Record Display

Cultural Objects Names Authority® Full Record Display

TO: 700000190 Record Type: Movable Work

Images: 1

Portrait of Agostino Pallavicini (painting (visual work); Anthony van Dyck (fl. 1599 - 1641) (fl. 1599-1641))

Titles:

Portrait of Agostino Pallavicini (preferred CVA/English/UA)

Catalog Level: Item

Work Types:

painting (visual work) [600000190] (preferred)
 (Object)-act, Visual and Verbal Communication (Hierarchy Name), Visual Works (Hierarchy Name), <visual works (Genre Term)>, <visual works by medium or technique>

Classifications:

paintings (preferred)

Creation Date: about 1621

Creator Display:

Anthony van Dyck (fl. 1599 - 1641) (fl. 1599-1641) (fl. 1599-1641) (fl. 1599-1641)
 Dyck, Anthony van (fl. 1599-1641), active in England (1500-1641)

Current:

Repository Number: 6000190
 The Getty Museum (Los Angeles, Los Angeles County, California, United States) (500000000) (The Getty Trust, Corporate Archive (Corp. Arch))
 Repository Number: 6000190

Display Materials:

oil paint (paint) [600015104]
 (Material)-act, Materials (Hierarchy Name), materials (matter), <materials by function>, coating (material), coating by form, paint (coating), <paint by composition or origin>
 Canvas [300014078]
 (Material)-act, Materials (Hierarchy Name), materials (matter), <materials by form>, <materials by physical form>, <fiber and fiber products>, <textile materials by process or technique>

Dimensions:

Unframed: 216.2 x 141 cm (85 1/8 x 55 1/2 in.); Framed: 264.2 x 185.1 x 10.2 cm (104 x 72 7/8 x 4 in.)

Style/Period/Group/Movement:

Baroque [600011177] (preferred)
 (Style and Period)-act, Styles and Periods, <styles and periods by region>, European, <European style and period>, <Renaissance Baroque style and period>

Wednesday



Overview of some of the basic concepts relating to Linked Open Data.

A brief glossary – to start us out on the same page

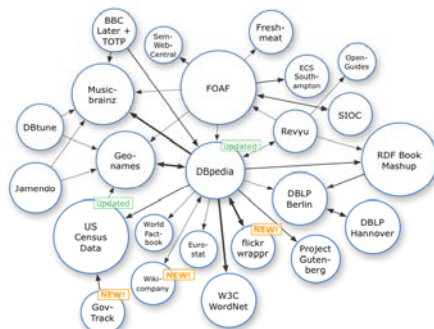
- Semantic Web
- Linked Open Data (LOD)
- Resource Description Framework (RDF)
- LOD Ontologies & Vocabularies
- Uniform Resource Identifier (URI)
- Web Ontology Language (OWL)
- Simple Knowledge Organization System (SKOS)
- Triplestore
- SPARQL Protocol RDF Query Language (SPARQL)
- Creative Commons (CC)

Semantic Web

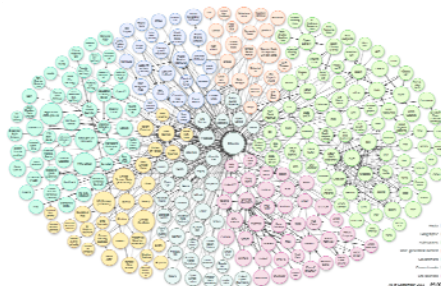
- A collaborative movement led by the World Wide Web Consortium (W3C),
- Promotes common data formats on the Web.
- Involves publishing in languages specifically designed for describing data: RDF, OWL, XML.
- Aims at converting the current web dominated by unstructured and semi-structured documents into a web of linked data,
- History
 - ▣ Concept of the Semantic Network Model was formed in the early 60s as a form to represent semantically structured knowledge.
 - ▣ Term 'Semantic Web' was coined by Tim Berners-Lee, the director of the W3C.
 - ▣ The original 2001 *Scientific American* article by Berners-Lee described an expected evolution of the existing Web to a Semantic Web.
 - ▣ He defined the Semantic Web as "web of data that can be processed directly and indirectly by machines."

LOD: Doubled in size every 10 months since 2007

View of LOD as of October 2007
2 billion RDF triples
Interlinked by over 2 million RDF links



View of LOD as of September 2011
31 billion RDF triples
Interlinked by over 504 million RDF links



Linked Open Data (LOD)

- Linked Data
 - ▣ A method of publishing structured data so that it can be interlinked and become more useful.
 - ▣ It builds upon standard Web technologies such as HTTP and URIs.
 - ▣ But, rather than using them to serve web pages for human readers, it extends them to share information in a way that can be read automatically by computers.
- Open data
 - ▣ The idea that certain data should be freely available to everyone to use and republish as they wish, without restrictions from copyright, patents or other mechanisms of control.
- Linked Open Data (LOD)
 - ▣ A community project, sponsored by the Semantic Web Education and Outreach group of the W3C, to extend the Web with a data commons by publishing various open datasets as RDF on the WEB and by setting RDF links between data items from different data sources.
 - ▣ Both linked and open.

Resource Description Framework (RDF)

- One of the key ingredients of Linked Data.
- Provides a generic graph-based data model for describing things, including their relationships with other things.
- Three column format – known as a *triple* – *forms the fundamental building block of semantic representations.*
 - ▣ *Subject* - corresponds to the entity – a ‘thing’
 - ▣ *Predicates* – property of the entity - names, birth/death dates
 - ▣ *Objects* - Subjects in another triple OR Literal values
- Conceptualized anything (and everything) in the universe as a resource.
- Specifically supports evolution of schemas over time without requiring all the data consumers to be changed
- RDF has varying representations (e.g., XML, text, JSON)

LOD Ontologies & Vocabularies

- Ontology
 - ▣ Used to capture knowledge about some domain of interest. An ontology describes the concepts in the domain and also the relationships that hold between those concepts.
- Vocabulary
 - ▣ A vocabulary in the sense of knowledge systems or ontologies are controlled vocabularies. They provide a way to organize knowledge for subsequent retrieval.
- Semantic Web Ontology
 - ▣ This is a formal specification of a shared conceptualization of the information to be published.
 - ▣ Examples of the vocabulary used in ontologies: class, object property, role, category, property, resource, etc.

Uniform Resource Identifier (URI)

- Identifies a resource either by location (URL), or name (URN). It usually includes some type of unique and persistent identifier.
- URIs are not URLs but every URL is a URI
- The URI for a resource represented in an RDF statement is called the URI reference (URIref) for that graph node.
- URIs are used as 'strong keys' that are assigned to each of the nodes so that we can refer to them consistently across all the triples that describe their relationships.
- The URN that will be the base URI for AAT will be <http://vocab.getty.edu/aat>.

Web Ontology Language (OWL)

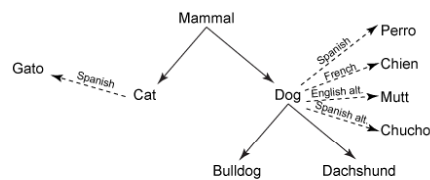
- Developed by the W3C
- Built as an extension to RDFS
- Provides a precise vocabulary that enables more powerful reasoning and inference over relationships with which to knowledge can be represented.
- Can be seen as a social contract between a data provider and a data consumer.

Simple Knowledge Organization System (SKOS)

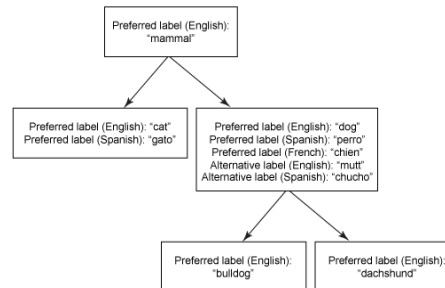
- A language designed for representation of thesauri, classification schemes, taxonomies, subject-heading systems, or any other type of structured controlled vocabulary.
- Based on RDF
- SKOS is currently developed within the W3C framework
- Can be used on its own, or in combination with other ontologies.

LOD: Simple SKOS example

Sample concepts relationship in a pre-SKOS taxonomy



Sample concepts relationship in SKOS



<http://www.ibm.com/developerworks/xml/library/x-skostaxonomy/index.html>

Triplestores

- Are purpose-built databases for the storage and retrieval of the type of unstructured information express in RDF triples that makes up the Semantic Web.
- Are not relational databases.
- Query language used with triplestores is typically SPARQL
- In addition to queries, triples can usually be imported/exported using RDF and other formats.
- Some triplestores store billions of triples.

SPARQL Protocol RDF Query Language

- Was made a standard by the RDF Data Access Working Group (DAWG) of the W3C.
- Is considered one of the key technologies of the semantic web.
- Allows users to write unambiguous queries.
- SPARQL endpoint is a service that accepts SPARQL queries and returns results.
- Queries can be distributed to multiple SPARQL endpoints.

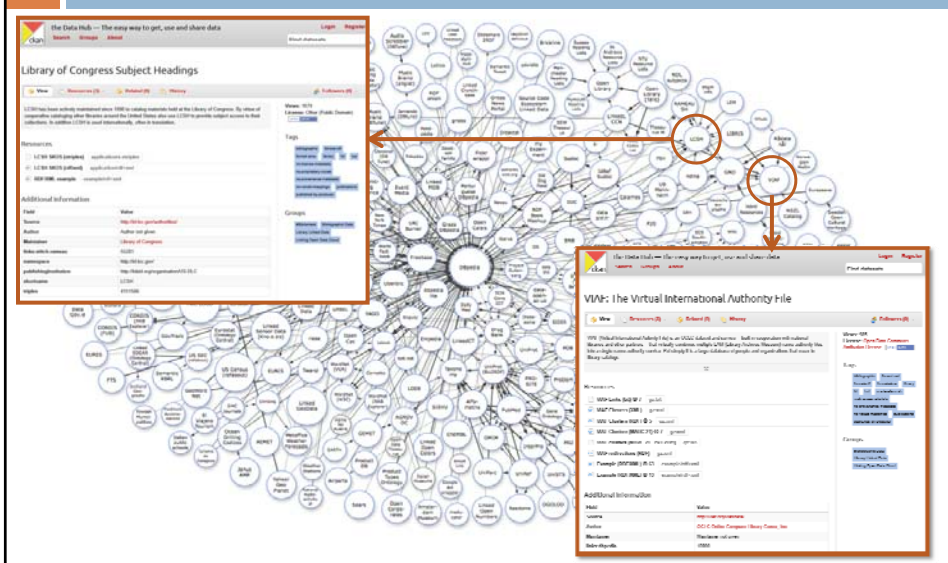
Why do we care

- The way data is being published on the web is currently in transition.
- A current trend in managing art information is to increasingly make data art, architecture, and cultural heritage objects available as Linked Open Data.
- This applies to the information directly describing the objects, but also the vocabularies used in the descriptions.
- For years now, the Getty Vocabulary team has been receiving legitimate requests to publish our thesauri as LOD.
- It is time to stop asking “why” and to start exploring “how”.

Explore the LOD Cloud Diagram

- This web page is the home of the diagram.
- The image show datasets that have been published in Linked Data format, by contributors to the Linking Open Data community project and other individuals and organizations.
- Clicking on the image will take you to an image map where each dataset is a hyperlink to its homepage
- URL: <http://richard.cyganiak.de/2007/10/loa>

Interactive LOD Cloud Diagram



Explore Existing SKOS Datasets

URL: <http://www.w3.org/2001/sw/wiki/SKOS/Datasets>

W3C Semantic Web

navigation

- [Main Page](#)
- [Recent changes](#)
- [Tools](#)
- [Books](#)
- [Validators](#)

other w3c resources

- [Activity news](#)
- [Publications](#)
- [Logos, buttons](#)
- [Activity home page](#)

w3c rss feeds

- [Activity newsfeed](#)
- [W3C blogs](#)
- [Use cases, case studies](#)

account request

- [W3C Member](#)

page discussion view source history

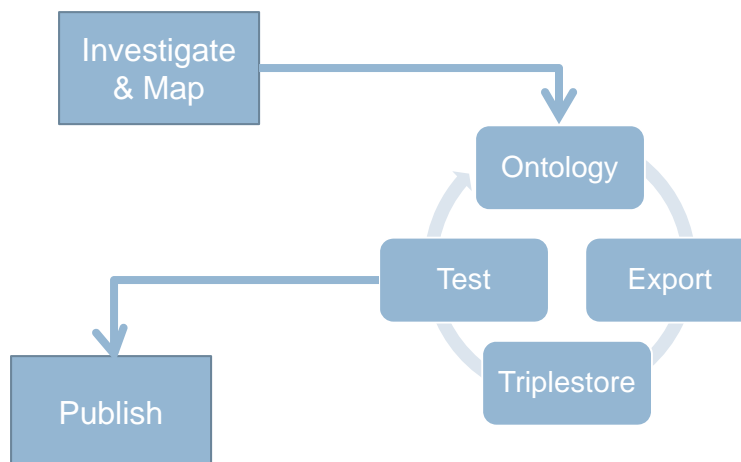
SKOS/Datasets

< [SKOS](#)

Contents [hide]

- 1 Find & Share SKOS Data
 - 1.1 STW Thesaurus for Economics
 - 1.2 Thesaurus for the Social Sciences
 - 1.3 Country Codes
 - 1.4 Drug Administration Forms
 - 1.5 Vocabularies for GeoSciML Geoscience information interchange
 - 1.6 DDC Dewey Decimal Classification
 - 1.7 Libris' vocabularies
 - 1.8 National Széchényi Library's vocabularies
 - 1.9 Library of Congress' vocabularies
 - 1.10 German national library subject headings
 - 1.11 French National Library's subject headings
 - 1.12 VIAF person authorities
 - 1.13 wikipedia categories
 - 1.14 New York Times subjects
 - 1.15 IVOA astronomy vocabularies
 - 1.16 NASA taxonomy
 - 1.17 GEMET General Multilingual Environmental Thesaurus
 - 1.18 AGROVOC Agricultural Thesaurus
 - 1.19 Linked Life Data
 - 1.20 Taxonconcept
 - 1.21 UK Public sector vocabularies

Our process



Miscellaneous hints from various sources about semantic data models

- No one gets it right the first time. Be happy if it seems like you are getting closer by the third time around.
- Semantic data modeling is very different from relational data modeling.
- It is a collection of ontologies that work together to represent the slice of the world you are focusing on.
- They are extremely open-ended and highly collaborative.
- They borrow vocabularies whenever possible from other ontologies. The more you borrow, the more portable and extensible your data becomes.

Contact Information

Joan Cobb: jcobb@getty.edu
Gregg Garcia: ggarcia@getty.edu

Information Technology Services
J. Paul Getty Trust