Cultural Objects
Name Authority

Questions and Answers about CONA
The Getty's Cultural Objects Name Authority™ (CONA)

The Cultural Objects Name Authority™ (CONA) is a new Getty vocabulary available for contributions as of 2011. CONA joins the other three Getty vocabularies, the Art & Architecture Thesaurus® (AAT), the Getty Thesaurus of Geographic Names® (TGN), and the Union List of Artist Names® (ULAN), as a tool for cataloging and retrieval of art information. CONA contains titles, current location, and other core information for cultural works. The scope of CONA will include architecture and movable works such as paintings, sculpture, prints, drawings, manuscripts, photographs, ceramics, textiles, furniture, and archaeological artifacts.

The following Q & A grew out of two recent presentations on CONA.

Questions and Answers:

Q&A from

VRA 2012, presentation by editor Jon Ward
April 2012

Q. Will CONA and the Getty Vocabularies have URNs? How can CONA bridge the concept of being an "authority" and structured data in the semantic web?

A. This topic is covered in this presentation:
http://www.getty.edu/research/tools/vocabularies/Linked_Data_Getty_Vocabularies.pdf
Highlights:

- The Getty Vocabularies' data is well equipped to be presented in a variety of RDF formats
- We have started to experiment with SKOS
- We are close to identifying a persistent URN
- We are Investigating which format will best serve the community (i.e., where to invest our limited resources)
- We recognize that Vocabulary data expressed in one of the RDF formats must happen soon, and we now look to the user community to provide partnership and to express their support
- CONA in particular would lend itself to being presented as linked data; CONA is the first vocabulary to actually be linked to the other three vocabularies
- CONA has the potential of creating a critical mass of supervised initial links among multiple institutions upon which the much larger endeavor can be built (linked)
- The GRI recently initiated a collaborative project to transform the data from our vocabulary databases into linked open data: Stay tuned for upcoming developments

---

Q. What can contributors do if they lack data in certain required CONA fields?

A. There are defaults that should be used as values for required fields where the data is unavailable (e.g., dimensions may be "unavailable," the creator name may be "unknown," etc.)

Q. Can VRA users contribute names to a vocabulary using Excel, provided the required fields match?

A. We prefer that you use the prescribed XML contribution format. The second preference would be that you give us data directly from your database (Access, etc.). However, if Excel is the only format you are able to provide, and if you intend to provide a large contribution (several thousand records) that will make it worthwhile for us to process data from Excel: then, yes, you may use Excel. If you intend to submit only small contributions (of several hundred records or less), please use the online contribution form.

Q. How will CONA data be updated? Works change hands often, for example. Will there be a process for repositories to refresh/update existing contributions, in order to keep the information current?

A. The full resolution of this issue is still in development. New data will be loaded and merged with existing CONA records. Changes in location or other information would ideally be contributed by the former or new repository.

Q. Can a contributor Wiki be established to load contributions and track changes for CONA?

A. The Getty cannot establish such a Wiki. But your user community is encouraged to do so. You can forward editorial questions to the Vocabulary Program as you work, and we will respond to you promptly. You can submit contributions in installments.

Webinar sponsored by the Museum Computer Network (MCN), Gallery Systems, and the J. Paul Getty Trust

Danielle Uchitelle, Managing Director, Gallery Systems, was the moderator. Erin Coburn, Manager, Collections Information and Access, Getty Museum, and Murtha Baca, Head of Digital Art History Access at the Getty Research Institute, introduced the topic. Patricia Harpring, Managing Editor of the Getty Vocabulary Program, presented an overview of CONA.
Tuesday, 4 May 2010

Q: Will this webinar be recorded and posted? Where will the PDF of this presentation be available? Thank you - this has been a great meeting!

A: The Webinar has been recorded and will be posted on the Gallery Systems, MCN, and Getty site at www.getty.edu/research/conducting_research/vocabularies/training.html . A printable version of the presentation slides is available on this page:
www.getty.edu/research/conducting_research/vocabularies/cona_intro.pdf

Q: How do I sign up to become a CONA contributor?

A: To be added to the list of potential contributors, please write to vocab@getty.edu, including your name and institutional affiliation. We will add this information to a database of contributors that we maintain. The XML schema for contributions is on this page:
http://www.getty.edu/research/tools/vocabularies/contribute.html
When the data dictionaries and other documents are ready, we will announce this on the VRA and ARLIS listserves, and we will send emails to potential contributors.

Q: How will CONA be made available? How often will the published data be updated?

A: The schedule and formats of publication are not yet established, since CONA is still under development. Our current plans are to make CONA available in the same ways and at the same schedule as we do for AAT, ULAN, and TGN, which are currently released 1) annually as licensed files in relational tables and in XML format, and via Web services in which the data is refreshed every two week, and 2) in

an online search tool in which data is refreshed every two weeks, as currently exists for example TGN and the other vocabularies: www.getty.edu/research/conducting_research/vocabularies/tgn/ . To see samples of data files currently available for the other Getty vocabularies, see this page: www.getty.edu/research/conducting_research/vocabularies/download.html .

Q: Will CONA be available via Cataloger's Desktop?

A: Since CONA is still under development, we have not investigated this possibility yet.

Q: Who do you envision will be the contributors to CONA?

A: Contributors to CONA will likely include museums, special collections, and other repositories of art and cultural works; visual resource collections; archives; art libraries; consortia of any of the above; bibliographic projects such as the Avery Index; and systems vendors or other entities that collect object/work data for clients.

Q: May CONA be used by non-contributors? If so, could you go over the advantages for contributors that will not apply to non-contributors?

A: As with the other vocabularies, CONA will be available to contributors and non-contributors alike. The fees and terms for licensing CONA are not yet established. However, they are likely to be the same or similar to those in place for the other Getty vocabularies. You can read about conditions currently in effect for the AAT, ULAN, and TGN on this page: www.getty.edu/research/conducting_research/vocabularies/download.html . Contributors of significant numbers of records to AAT, ULAN, and TGN are currently given a fee-free license for the data. However, the primary advantage of contributing to CONA and the other Getty vocabularies – outside of the altruistic goal to share data, which is often mandated by the contributing cataloging institution's mission – is the practical advantage to a contributing institution of having their own data included in a compiled data set, making that data set particularly useful for them in retrieval and cataloging. For repositories of objects/works, an additional advantage of contributing to CONA is the assurance that CONA authority records for their objects/works contain accurate data. When visual resource collections or other contributors to CONA must rely upon second-hand sources to fill in values for these objects/works, the information may not be as up-to-date as the repository's own records.

Q: When will CONA be available to users? Do you already have a basic set of content?

A: CONA is scheduled to be available for contributions in 2011. It is currently still in development. We will make an announcement on our Web site, to list serves, and in emails to potential contributors when an exact date is known. It is still unknown when CONA will be available for licensing; that will largely depend upon the timing and the number of contributions received for CONA. Currently, we anticipate

that the data sets to be used in testing the CONA system, and then published in CONA, will include work records from the Getty Museum, the Avery Index, and the BHA data files.

Q: How do you envision CONA will be used? By whom will it be used?

A: As with the AAT, ULAN, and TGN, CONA will be used as an authority for indexing and a source of values and relationships for retrieving. In the case of CONA, the unique numeric identifier and the titles of objects/works (along with other identifying information for the objects/works) will provide important access points for information about works of art. The anticipated audience for CONA includes visual resources collections, repositories (museums and special collections), archives, art libraries, consortia, and the technical support and systems providers for these groups.

Q: How will you weigh the holding museum's contributions against information for the same art works provided by a visual resources collection or other contributor?

A: The information from the repository of the object/work will be considered more authoritative than information coming from other contributors who have had to rely on second hand sources for the same information. Where only one value is included – for example, for dimensions – the repository's dimensions would be preferred over the dimensions provided by another contributor for the same object/work. Where a field is repeating – for example, Title/Name – titles from both contributors would be included, with the contributor for each tile noted. In this case, titles are a critical access point, and all titles by which the work may be known should be included. End-users will be able to see which is preferred by the repository through both a link to the contributor's acronym, and by a "Title Type" field that will be flagged "repository title."

Q: Will there be a process for repositories to refresh/update existing contributions, in order to keep the information current?

A: The full resolution of this issue is still in development. For the current vocabularies, new data is loaded and merged with the existing vocabulary records. If the repository were to indicate that the newly contributed CONA data should replace their previously contributed CONA data, one possible solution could perhaps be that the loader could be set to accommodate that.

Q: What are the editorial rules? Are they available now? What if our data is missing a required field – would we be prohibited from contributing?

A: They are available on this page:
http://www.getty.edu/research/tools/vocabularies/guidelines/index.html
The CONA editorial rules, required fields, and requirements for controlled values are drawn from CDWA (*Categories for the Description of Works of Art*,
www.getty.edu/research/conducting_research/standards/cdwa/) and CCO (*Cataloging Cultural Objects*) rules, both of which are published. For CONA, there will be default values and suggestions for how to supply

data for CONA's required fields when it is missing in contributors' records. For example, for a CONA display field, data for this display could perhaps be parsed from several controlled fields in the contributor's database. However, given that CDWA and CCO represent best common practice already in place in repositories, by extension, most contributors will typically already have data for most CONA required fields.

Q: Insofar as many contributing databases are formatted in VRA Core, how mappable would such data be with the CONA version of CDWA? It looks pretty close. Will you be creating a transform or support for VRA Core 4 validated XML?

A: You will be able to map data elements between VRA Core and CDWA Lite or the CONA import format. There is currently a mapping between CDWA Lite and VRA Core (and other standards) on this page (the CONA import format is not yet available): www.getty.edu/research/conducting_research/standards/intrometadata/crosswalks.html . We do not have specific plans to create a transform or support for VRA Core 4 validated XML, however we will note that this would be desirable and add it to our technical wish list. In any case, I would think that users or a systems vendor would perhaps do so in any case, and share the information among themselves. One thing to note about VRA Core: CONA collects records for the objects/works depicted in the visual resources, not information about the image per se.

Q: Will CONA link to ULAN, so there will be no need to re-enter all of the creator biographical information in an individual CONA record? How will new artist names, those not already in ULAN, be incorporated in CONA? What about TGN names in CONA? Would creator names or geographic names not already in ULAN or TGN need to be entered first into ULAN or TGN?

A: Yes, the artist names and place names in CONA will be controlled by ULAN and TGN, so there is no need to re-enter all of the peripheral creator information, etc. in the CONA contribution. Currently, controlled lists in the other vocabularies (those controlled by AAT or TGN) are reconciled with values in the contributor's data in pre-processing. If the CONA contributor has actually captured the ULAN, TGN, or AAT unique identifier in their record, this mapping is relatively easy. If the contributor has only captured text values in a field (with no numeric ID), then automatic matching can be done; but there may be some homographs that require a human to make decisions about mapping. For CONA, we realize that there will be sometimes be new artists that must first be integrated into ULAN. We believe the process will be similar to that currently used for the other vocabularies, but our technical developers are still working on ways to accommodate this efficiently. Once it's worked out, we can give you more details. A bulk load of your new artists to ULAN prior to your CONA contribution would probably be one way to deal with this.

Q: What will be done when duplicate records are submitted? Will you combine the information somehow?

A: We will use the same process as we currently do for AAT, ULAN, and TGN. When multiple contributors submit records for the same object/work, they will be "merged" into a single record in CONA, with certain information linked to the contributor's acronym or name in the merged record. You can see an example of a merged record in ULAN, for Le Corbusier (the contributor acronym appears at the bottom of the page, after the name they contributed):
www.getty.edu/vow/ULANFullDisplay?find=Le+Corbusier&role=&nation=&prev_page=1&subjectid=500027041 . We will find matching records in a way similar to the way we do it for the current vocabularies. Currently with the AAT, ULAN, and TGN, this is done by algorithm in preprocessing or when a contributor's data set is loaded into the vocabulary editorial system. Matches are automatically made by comparing key fields, for example, the name, birth and death dates, nationality, and roles of a contributed ULAN record. If the incoming data is missing values, or if values do not match due simply to the common variation possible in such data, the system is still able to recognize near-matches, which are distributed to the Vocabulary Program editors for further investigation. We estimate that 70 to 80% of records can be matched automatically for ULAN. The rate of matches will probably be higher for CONA, if most of the required data is provided by contributors.

Q: Are you using a scheme or a framework for managing persistent ids? How does it change during the merge process? How many digits will be used for the CONA ID numbers? There were six digits in the example on your slide.

A: We'll use the same sort of scheme that we use for the other three vocabularies. Currently, for ULAN, AAT, and TGN, there is a unique and persistent ID for each record, which is automatically generated, never re-used. If a new record is loaded for the same concept, it is merged with the existing record and the ID of the existing vocabulary record is retained. If a rare situation would require the generation of a new ID, the licensee would receive a list of such changed IDs with the updates of the data files. The only situation that I can think of that would generate a new ID is the following, for example in ULAN: If two records for the same artist were in ULAN, but not merged, and both records were published in one year, but merged during the following year, one of the IDs would be unavailable to end users: In the next data release, the implementer would use the file of merged IDs to see that the old ID from their previous data set had now become part of a merged record. The same methodology will apply to merged CONA records: the unique and persistent ID can change in only the rarest of situations, in which case implementers are notified. Regarding the number of digits for the CONA ID, our technical team will allow a range of numbers large enough to accommodate the vast potential of growth for CONA. The slide was created by me for illustration purposes only, since CONA is still under development.

Q: How will you disambiguate records when all values are the same for multiple works? For example, how will end users differentiate between the many works with a generic title and similar creator data?

A: For systems, the unique numeric ID will disambiguate records. For display lists for end users, in addition to creator and title, enough information will be supplied to disambiguate one object/work from another, including tombstone information and the repository for movable works.

Q: Are thumbnail images going to play a role in disambiguation? Is a thumbnail included in the CONA record and as part of the contribution? Are images required for contributions?

A: The role played by thumbnail images in CONA is yet to be finalized. We recognize that thumbnail images could be very useful for CONA, even more than for AAT, ULAN, and TGN. For the existing vocabularies, you can see an example of how we have linked to images for a limited number of records, as for the "representative images" for *rhyta* in the AAT:
www.getty.edu/vow/AATFullDisplay?find=rhyton&logic=AND&note=&english=N&prev_page=1&subjectid=300198841 .
Thumbnail images would not be required of contributors. Whether and how thumbnail images would be included in the CONA import format is still being developed.

Q: How will CONA accommodate LC subject headings? Will it accommodate names from the Oxford Dictionary of National Biography or other published sources?

A: For indexing of the Subject Matter fields of the CONA object/work, there will be a set of fields for recording the LC identifier and text string of the LC (Library of Congress) subject heading. Iconclass or other subject indicators may also be recorded here. For other data, for example artists in the CONA Creator fields, the names are controlled by ULAN. However, ULAN editorial rules require the LC name to be included and flagged in the ULAN record. Thus in theory, the CONA contributor could locate the LC name in the ULAN. However, not all contributors to ULAN have been able to flag the LC names. Thus in practical application, the mapping between some LC personal names in a contributor's data and the ULAN records controlling CONA will need to be worked out in pre-processing. The same is true for names taken from published biographical dictionaries: they will be mapped to ULAN values before loading.

Q: Would Iconclass be included or linked in some way to the iconographic authority?

A: Fields are provided to enter Iconclass text strings and codes as auxiliary fields, but not actually incorporated into the Subject Authority per se of CONA. In the future, we will possibly explore with the Iconclass team the possibility of loading the Iconclass values in the CONA Subject Authority.

Q: Are art works from India, Africa, or other non-Western art in scope for CONA?

A: Yes. There are no geographic or chronological limits to coverage in CONA.

Q: Names of portrait sitters are not in ULAN. How can their names be controlled in the Subject Matter area of CONA, which seems to be linked to ULAN in your slide?

A: Afacet in ULAN for non-artists has been created in ULAN, which could contain information for the sitters needed for CONA and other implementations using ULAN. Actually, we have long had such a facet, but was unpublished and formerly contained primarily authors who are not visual artists; they came into ULAN via various past contributions.

Q: Where can I find a copy of the General Subject terms? Secondly, with what authority do the subject terms comply?

A: The general subjects are listed here
http://www.getty.edu/research/tools/vocabularies/guidelines/cona_3_6_3_subject_authority.html#3_6_3_1
and are also found in CCO and in CDWA, under "authority and controlled list" at
www.getty.edu/research/conducting_research/standards/cdwa/18subject.html#RTFToC3 .

Q: Any comments about the difficulty of assigning subject matter to contemporary art?

A: This is an important topic, but out of scope for this discussion. At minimum for CONA, you could list a general subject, such as *nonrepresentational art* or *performance*. Hopefully the topic of subject access can be addressed at a conference or workshop in the next year or so. I heard a rumor that such a workshop may possibly be proposed for the 2011 VRA conference. Examples and outcomes of such a discussion could be included in the CONA manual.

Q: Our database is totally in Dutch. How can this data fit in CONA?

A: We welcome contributions in languages other than English. Much of the minimum required CONA data is the same in both Dutch and English, or could be mapped to controlled values in CONA, which are often in turn controlled by the multilingual AAT, ULAN, or TGN. For example, the AAT will soon include Dutch terms, so this will help map to controlled English values (indeed, Dutch museums probably already have access to the Dutch AAT). Other fields are repeatable and values in multiple languages are accepted, for example Title/Name and Descriptive Note. An English title is strongly encouraged. We would be happy to have an in-depth discussion with you concerning practical details of how your Dutch database could be contributed to CONA.

Q: Can data submissions include information in other alphabets, like Greek?

A: CONA is Unicode compliant and will accept values in multiple languages and alphabets in two key fields, Title/Name and Descriptive Note. For other fields, the Greek data would need to be translated into the Roman alphabet, or mapped to English controlled values. This is the same contribution policy in place for the other Getty vocabularies. Given that it would impractical for us to allow every single field to exist in every possible language, we focus on gathering multilingual information in the key fields required for access or of most importance to end users. We would welcome an in-depth discussion with you concerning practical details of how Greek data could be contributed to CONA.

Q: For those of us working in VR collections, we gather information from multiple sources to complete a data record. Is there a place in CONA to record the data source for each piece of information? Another asks: Is warrant required? For contributing museums, is the museum itself considered warrant?

A: Sources are linked to key fields of the CONA record, including Title/Name, Descriptive Note, and the record as whole. There is a separate controlled source authority to capture sources. To see how these sources are linked to an online display of a current Getty vocabulary, see the bottom of the AAT record on this page, where sources and contributors for the terms, note, and full record are linked: www.getty.edu/vow/AATFullDisplay?find=concrete&logic=AND&note=&english=N&prev_page=1&subjectid=300010737 . For repositories of objects/works (museums and special collections), the warrant of the repository's own catalog record is sufficient.

Q: Will CONA's online interface link back to the contributor's record online?

A: Yes. Among the caveats is the practicality of maintaining the currency of the many thousands of resulting live links. At minimum, the contributor will be able to cite the URL in the Source area of a CONA record.

Q: Are provenance and exhibition history included in CONA?

A: A designated note field is set aside for Provenance in CONA. Exhibition history is not included as a separate field, but you could include it in the Descriptive Note.

Q: How do the components of "Current Location" work? Is it required to enter them separately, i.e. repository and its geographic components, or is all of the information entered in a single field?

A: You will not have to enter the components of Current Location separately. Current Location is controlled by two authorities: ULAN for repositories and TGN for geographic places. In ULAN, the

geographic place of the repository is included. So for CONA, you will enter or map to only one controlled entity, and all of the component parts will come with it from the authority (e.g., the geographic place "Washington, DC" would be linked to the controlled repository "National Gallery of Art." You need only link to the NGA record, not separately to DC).

Q: Is a repository accession or inventory number required? This would be one of the best ways of uniquely identifying objects with similar titles and creators. How is Repository ID number factored into the Current Location field?

A: Yes, the repository's ID for the work is required, when known. It is a field separate from, but linked to, the location fields.

Q: Is there a GIS component for built works?

A: The geographic coordinates for built works could be included a note field, which could also contain street address, or shelf location for a movable work. There is no fielded GIS component in CONA, although through the link to geographic location, the coordinates of the general location would be linked.

Q: What terms do you use to classify installation works? What about video projections? What about new media that is permanently or semi-permanently installed in one location? What about born-digital and digitized objects?

A: All of these types of work are within scope of CONA. We would welcome examples to help us work out the issues regarding Object/Work Type and Classification for the manual (send them to vocab@getty.edu). The work types in CONA are controlled by AAT. You could also study the AAT in the meantime, and suggest new terms if necessary (include literary warrant).

Q: How do you use repeating fields to catalog installations that get reinstalled?

A: That is a good issue for the editorial manual. It could be accommodated through repeating fields (for example, Creation Location), or – if the installations were significantly different from one another – even by making each installation a separate work, linked to the others through hierarchical or associative relationships. In the meantime, if you have good examples, please send them to vocab@getty.edu.

Q: At the start of the presentation it was stated that music and films were not intended for CONA, but among the examples was a piece of film. When are films within scope of CONA?

A: CONA is not intended to be a comprehensive catalog of commercial films or the world's music. However, at the same time CONA must accommodate the types of cultural works collected by visual arts institutions. Therefore, films about art, films by a director who is considered a visual artist, or films otherwise contained in the collections of such institutions are welcome. If your institution collects primarily works of visual art, and a film is part of your collection, then the film would most likely be within scope of CONA.

Q: If you are cataloguing ethnographic works, e.g., African, Asante, etc., would the creator information go into Culture or the Creator fields?

A: For CONA, it can go in both places. Creator is required. Therefore, if you locally record it only in Culture, you or your programmer (for bulk loads) can add the word "unknown" to the value in your culture field in order to fill in the CONA Creator field. ULAN has a facet containing such "unknown" creators, which will control values in CONA. If you notice in the meantime that ULAN is missing a value that you need, please let us know. See example at
www.getty.edu/vow/ULANFullDisplay?find=unknown+asante&role=&nation=&prev_page=1&subjectid=500202882 and the hierarchy on the following page (click on the "U" to see the values)
www.getty.edu/vow/ULANHierarchy?find=&role=&nation=&page=1&subjectid=500125081 .

Q: Are copper printing plates for etchings, or photographic negatives or transparencies considered objects/works in CONA?

A: Yes, these may be included in CONA, and linked to the prints made from them.

Q: Will there be a way to sequence component records related to a parent work, e.g. pages in a folio?

A: Yes. You can list the page in "Relationship Number or Qualifier" (a note where you can express nuance, e.g., *34 v*.), and you can indicate how the siblings should sort under a parent in the hierarchy with "Sort Order among siblings."

Q: Will there some type of a demo available for CONA? Will there be training for cataloging and using CONA?

A: We plan to have training materials on the Vocabulary Program Web pages. We also plan to conduct workshops at conferences when CONA is available for contributions.

Q: The authority records have unique IDs - are these URIs that could be used to create semantic references? What about the relationships like part/whole - could those have URIs?

A: This is in development. Higher priority development projects for CONA currently on our plate include the editorial system, the publication systems that include annual exports, Web Services, and the online "browsers", reports for quality control and statistics, the loader, the online contribution form, and various other systems.

Q: Could we use the CONA namespace to connect our field definitions with CONA's, like DC?

A: We realize that this is an issue and will re-visit it after CONA is further developed.

** Answers were written by Patricia Harpring, with input from other Webinar participants and the Getty ITS development team.