

## Museum Data Exchange CCO Evaluation

### Criteria for Scoring

Patricia Harpring  
22 July 2009

#### Principles of scoring

The scoring methods described in this document were used to aid in the analysis of the top 20 unique data values in core CCO/CDWA fields, which were derived from the data sets supplied through OCLC by nine participating institutions. The goal of the scoring study is in part to establish useful criteria for scoring and judging compliance with the CCO and CDWA standards. The principles for scoring are derived from the rules in CCO and CDWA Lite (which is compliant with CCO and big CDWA).

The nine participating institutions comprise a significant representative cross section of different types of museums and different scopes of collecting. Studying the data of such a cross section will provide conclusions and suggested courses of action that should also be applicable to the broader museum and art information community.

The standards used in this study to judge compliance are the following (CCO is in agreement with the related CDWA guidelines and harvesting protocol):

Baca, Murtha, Patricia Harpring, Elisa Lanzi, Linda McRae, and Ann Baird Whiteside. *Cataloging Cultural Objects: A Guide to Describing Cultural Works and Their Images*. Chicago: American Library Association, 2006.

*CDWA Lite: XML Schema Content for Contributing Records via the OAI Harvesting Protocol*. J. Paul Getty Trust, 2005.  
[http://www.getty.edu/research/conducting\\_research/standards/cdwa/cdwalite.html](http://www.getty.edu/research/conducting_research/standards/cdwa/cdwalite.html) (accessed 22 July 2009).

Baca, Murtha and Patricia Harpring. *Categories for the Description of Works of Art*. Los Angeles: J. Paul Getty Trust and the College Art Association, 2000. Revised 2009.  
[http://www.getty.edu/research/conducting\\_research/standards/cdwa/](http://www.getty.edu/research/conducting_research/standards/cdwa/) (22 July 2009).

CCO and CDWA describe best practice in cataloging art and related materials as expressed by consensus of representatives from the communities that provide and use art information: museum curators and registrars, art historians, visual resource professionals, art librarians, archivists, information managers, and technical specialists.

#### Results in brief

The nine sets of data analyzed for compliance in this study scored quite well. Many of the points deducted in scoring were due to mapping and parsing issues that could be easily corrected. The most frequent issues concerned having multiple terms in one field or missing data that is 1) probably actually available in the institution's local data base (e.g., a missing Work ID) or 2) could be filled using suggested default values (e.g., globally supply "artist" or "maker" for missing Creator Role).

## Museum Data Exchange: Criteria for Scoring

Most institutions have no, or very little, subject data. The CCO and CDWA require at least one subject term for each work; a short list of generic subject terms is listed in CCO and CDWA, and could be used to satisfy the minimum requirement. With a bit of analysis, strategies could be employed to globally designate some subject terms, using the title of the work or the work type for utilitarian objects to map to subject terms. For example, the title "Portrait of Thomas Jefferson" includes the generic subject "portrait" and a specific subject "Thomas Jefferson." A Title (which itself repeats the Object/Work Type) "Amphora" could be mapped to a generic subject term "object (utilitarian)" or "ceremonial object," and could even repeat the term "amphora" as subject.

The scoring in this study was done by hand, but in order to be practical, the group will need to analyze how to best score CCO compliance without extensive human intervention. Many aspects of compliance would be difficult to judge by algorithm and thus require human judgment. I suggest a combination of human analysis and automatic scoring. Perhaps a human could analyze a sampling of a given data set in a manner similar to that used in this study, then suggest changes to make the data compliant. When changes in mapping, parsing, or additions of missing data have been supplied as far as is possible, the data set could be run against an automatic scoring algorithm. After both tests are completed, the resulting score should be a fairly good indicator of compliance.

The scoring and ranking in the current study was done by calculating percentages of compliance on the following scale:

90-100% = A  
80-89% = B  
70-79% = C  
60-69% = D  
0-59% = F

### Caveats and issues regarding the scoring

**Scoring applicable to the nine data sets being analyzed:** The system for assignment of points, as outlined in this document, is based on the nine data sets being analyzed for the purpose of this particular study. Although I also tried to imagine how other data sets may possibly depart from full CCO compliance, the criteria and assignment of points described here may need to be expanded over time as future data sets supplied by other institutions are tested. In time, after many data sets have been analyzed, a fuller understanding of how data may depart from the guidelines will eventually inform a comprehensive set of criteria for scoring.

**Scoring unique values vs. total number of records:** This study considers the top 20 most often used unique data values for each required element. As such, this study is useful in analyzing the types of issues involved in achieving compliance. However, in a final scoring strategy it would be more useful to assign compliance levels based on the total number of records in a data set that are compliant, rather than examining only the top 20 most often used unique values.

## Museum Data Exchange: Criteria for Scoring

The reason for this recommendation is best explained by example. Consider the Title element in a data set of 10,000 records. When scoring the top 20 unique values, the score for the value "Untitled" (for which points are subtracted because it is not descriptive) would be of equal weight to the other 19 fully compliant values in the top 20 unique values; the score would be around 95%. However, if 9,900 records in the data set have the value "Untitled" and only 100 records have meaningful descriptive titles, 95% is a deceptively high score when judging overall compliance. The scoring could be skewed in the reverse direction as well. Consider the Current Location field: if 9,900 records have a valid value "Art Museum (Our Town, Indiana, USA)" but the other 100 records have "[no data]" in that field, when each unique value is scored equally, the scoring is skewed deceptively low at only 50% compliance.

Considering only the most often used unique values for each element allows for an organized approach to scoring based on a small number of sample records, and thus was an objective, useful way to carry out the current study. It is hoped that these records are representative of the full data set, and with elements such as Record Type this is true. Note however that for other elements, examining the most often used unique values for an element actually brings the problematic data to the top and distorts the overall compliance of the full set of data. In analyzing Title for example, the best descriptive titles will by definition typically be unique to one or only a few works. However, when sorted with most often used unique values floating to the top, the non-compliant "Untitled" often appears near the top.

**Criteria not scored:** In the list of criteria below, those indicated as "criteria not scored" represent rules and guidelines in CCO and CDWA that were impossible to judge or score in the current data sets. For example, if capitalization is dictated in the rules, it could not be a criterion in scoring data where all terms were expressed in all lower case.

**Fields not included in the study:** Several fields included in the original spreadsheets were not analyzed for scoring because they pertain to images and visual surrogates (Resource-View elements) or to technical or administrative data (Record ID).

**Capitalization and punctuation:** CCO/CDWA outline guidelines for the capitalization of personal names, place names, and other terms and texts; for example, the CCO/CDWA rules advise avoiding the use of all caps for displaying an artist's name. In the current scoring, capitalization was ignored because the data had been transformed to all lower case. When instituting future methods to score overall compliance, some consideration should be given to the original capitalization and punctuation in certain elements; however, overall, lack of compliance regarding punctuation and capitalization would be considered a very minor flaw in determining scores.

**Display vs. indexing:** As described in CCO and CDWA, certain fields are intended for display to end users, while other fields are intended for indexing, and therefore use controlled terminology (e.g., Object/Work Type) or controlled format (e.g., Earliest and Latest Dates). Fields intended for indexing are the most critical fields in retrieval, and they therefore receive the

strictest scrutiny in scoring, regarding both consistency and mapping to authoritative controlled vocabularies.

**Rules for free-text display fields:** Using consistent terminology is especially important for controlled fields that are intended to provide access. Consistency is less important, but still desirable, for a free-text display field than for an indexing, controlled field. Although uncontrolled terminology should be accommodated in display fields, using terminology that is consistent with the terms in controlled fields is recommended for the sake of clarity. Consistent style, grammar, and syntax should be used. However, in scoring for display fields, points were not deducted for inconsistency. At the same time, points were deducted if the display field lacked information required by the rules.

**Rules for indexed, controlled fields:** Use of controlled terminology and other rules were more strictly scored for required elements intended for retrieval. If data used in retrieval is not controlled and consistent, retrieval will be uneven and inadequate.

In some cases, fields for display are required by CCO/CDWA, while the indexing fields for those fields are optional (but recommended). For example, Display Measurements and Display Materials/Techniques are required so that the end users of the data can understand the physical characteristics of the objects. These fields are not intended for use as criteria in retrieval. However, given that most museums capture the indexing fields that go along with such display fields (and in fact often construct the display from these controlled fields), indexing fields for these and other comparable elements should be included if possible. They were not labeled as "required" by CCO/CDWA because certain constituencies, such as visual resources collections, do not always have access to the index fields. Such indexing fields are considered "required if known" because they will be very useful in retrieval.

In a very few cases, important fields indicated in CCO/CDWA were not included in the study. Analysis should be done to double check that all such elements are included in the final compliance test, particularly when it is data that is routinely collected by the repository. For example, it is intended that 4.3. Sub-element: Vital Dates Creator should include not only the display Vital Dates/Display Biography, but that the sub elements indexing the dates are included as well, flagged with Attributes: birthdate, deathdate. While it is true that, in this particular case, if the artist is already in ULAN or another authority, the biographical data could be used in retrieval by linking to the controlled vocabulary. However, if the artist has not yet been contributed to ULAN, important retrieval information such as the artist life dates should be captured.

Analysis should also be conducted to avoid sorting non-required values with required values for scoring for compliance, if possible. It was not possible for me to determine the distinction in the current study, but circumstantial evidence seems to point to non-required values having been grouped together with required values in some elements. For example, in Creator and Location elements, the creator and the current location are required. However, other data is captured in the same elements, distinguished by "type"; for example, the required location is flagged with

type=currentLocation. The values in some of these top 20 unique values appear to be sitters' names (not creators) and discovery locations (not current location). On the other hand, perhaps this distinction is not practical to make when attempting to score data by algorithm rather than by human intervention, because the rules are complex. For example, if the current location is unknown, thus the value flagged type=currentLocation "unknown", it is then required to include the last known or the discovery location, flagged as such for display so it is clear to users. This would require analysis and an "if--then" logic in the compliance scoring algorithm.

**Consistency of scoring:** Every attempt has been made to be consistent and objective in scoring compliance across different institutions' data sets. However, given that all institutions are not contributing data in exactly the same form, it is challenging to maintain complete evenness while scoring similar — but not identical — data.

**When data is unavailable:** In comparing scoring techniques across different fields, scores were designated differently based on whether the data is knowable or impossible to know. This is in keeping with the rules for these elements in CCO and CDWA. For example, in the Creator elements, a value such as "Unknown" or "Unknown Celtic" is considered compliant with the standard because it is common that an individual creator is unknown for certain types of works from given time periods or regions. CCO and CDWA advise using values incorporating the word "unknown" combined with a culture designation for such cases. However, in the Title field, points were subtracted for an analogous value "Untitled," because CCO/CDWA rules require a descriptive title. The standards have guidelines in place for how to construct titles for various types of works that may be untitled, strictly speaking. For example, if the work is a utilitarian object, repeat the Object/Work Type as Title (e.g., "Amphora"). If the work is named "Untitled" by the artist, the guidelines suggest constructing another title that is descriptive (e.g., "Abstract Composition").

CCO/CDWA rules suggest defaults for missing data in other fields. For example, if an institution does not capture roles such as "sculptor," "painter," "designer," etc. for an artist of a given work, the default value "artist" or "maker" should be parsed into the Artist Role field in CDWA Lite.

**Mapping data:** When mapping data for compliance to CCO/CDWA, if required values are missing or not controlled, it is not necessarily required to change data in the institutions' local data base. Instead, a mapping may be set up to fill in values in the target data base (in this case, via CDWA Lite). For example, if the institution's database does not include role, the default value "artist," "creator," or "maker" could be placed in the CDWA Lite element. If the institution uses a Role value that is not controlled by a published vocabulary, such as "person depicted," that could be mapped to the AAT term "sitter" for parsing into CDWA Lite.

**Multiple terms in one field:** Other frequently occurring mapping issues in the data sets in this study have to do with multiple values in one field. This applies to fields intended for indexing and retrieval. Points were subtracted in the scoring for multiple values, because this makes it difficult or impossible to then match the values to a standard controlled vocabulary.

In most cases, multiple values were separated by punctuation. Thus, the individual values could be mapped to repeating instances of a field or CDWA Lite sub-element. Where multiple terms were not separated by punctuation, more points were deducted because terms could not be parsed for comparison to controlled vocabulary in retrieval (i.e., an algorithm could not know which values were separate because a term may be comprised of multiple words; so simply separating the words is not effective).

When comparing the values in the spreadsheet to similar values in the institutions' online Web sites, it occasionally seemed to me as if the punctuation may simply be an artifact of parsing into CDWA Lite (or parsing from CDWA Lite into the spread sheets), because there were no such strings of terms separated by punctuation in their online data.

**Broader / narrower contexts and synonyms:** Multiple terms in one field occasionally seem to represent broader/narrower contexts. How the scoring methodology will deal with such cases is a matter for general discussion, probably after the technical possibilities are explored.

In the data scored in this study, two types of broader/narrower contexts were evident: whole/part relationships between objects and thesaural hierarchical relationships between terms.

For example, if Object/Work Type is "book-print," this presumably means the work being cataloged is a print that is part of a book. Ideally, the Object/Work Type should be "print," and the entire record for the print should be linked to a separate record for the book.

In another example, if Object/work Type is "print-lithograph," this presumably means the work is a lithograph, and the broader context of the term "lithograph" is "print" in a thesaurus. Thesaural relationships were also seen for Locations and Nationality, where terms such as "Greek-Attic" and "Italian-Roman" were found.

It must be studied to determine whether or not it is possible distinguish the various types of relationships intended when multiple values are included in one field. If multiple types of implied relationships are expressed without distinction in the same data set, it will presumably be impossible to parse the relationships correctly by algorithm. The safer and easier solution may simply be to parse each term in a separate instance of the field or sub-element; retrieval will still be improved, although subtle distinctions and relationships will be lost in the data.

Finally, there were examples of synonyms being included in one field; for example, "Orley, Nicolas van (Orley, Nikolaus)". In all cases, synonyms should be parsed in separate instances of the sub element or field, not enclosed in parentheses. This must be distinguished from examples where multiple artists are in the same nameCreator field (e.g., Ricci, Marco- Ricci, Sebastiano). There were few examples of the latter in the data sets in this study. Variant names for same artist should be contained in the same cdwalite:indexingCreatorSet, but names for separate artists should be indifferent sets.

**Compound terms:** If terms are not separated by punctuation, it must be considered that they may be compound terms rather than uncontrolled terms. Compound terms are compliant with the rules. The CCO/CDWA rules allow construction in local databases of compound terms, where the terms are derived from separate terms in an authority such as the AAT. For example, "furnishing fabric" is considered compliant, although the terms are not bound as a compound term in the AAT (i.e., "furnishing" and "fabric" are separate terms in the AAT).

**Qualifiers:** Note that qualifiers should be stripped from terms (e.g., *mbulu (reliquary figure)*) in the incoming data and the controlled vocabulary for the purposes of matching for compliance. Some contributors may include the qualifier in the field with the term. This is compliant with CCO/CDWA, but should be ignored in automatic matching.

**Contribute to vocabularies:** Controlled terms should ideally be judged compliant not only based on whether they are currently in a vocabulary, but whether they are of the correct form for and will be contributed to the vocabulary. By contributing such terms to a published controlled vocabulary, the terms will eventually be available for future retrieval of the data.

## Museum Data Exchange: Criteria for Scoring

### Scoring: Brief Overview

#### 1. Object/Work Type

- Full compliance	5
- Controlled term, but multiple terms per field with punctuation	4
- Term is not controlled by AAT, LC, or Chenhall; a phrase or several terms are included in one field without punctuation	2.5
- Not blank, but contains "N/A" etc.	0.5
- Missing value	0

#### 2. Title

- Full compliance; "untitled" is not full compliance, because title can always be constructed	5
- Title is included, but not descriptive	3
- Not blank, but contains "N/A," "untitled," etc.	0.5
- Missing value	0

#### 3. Display Creator

- Full compliance; "unknown artist", etc. is okay, but culture is preferred for unknowns "unknown Celtic"	5
- Name is included, but biographical context information is missing	3
- Not blank, but contains "N/A" etc.	0.5
- Missing value, e.g., reference to unknown artists are missing	0

#### 4. Name Creator

- Full compliance; "unknown artist", etc. is okay	5
- Controlled, but multiple names in one field	4
- Names are not controlled (should come from a published authority ULAN, Library of Congress, published artist authority)	2.5
- Not blank, but contains "N/A" etc.	0.5
- Missing value	0

#### 5. Nationality Creator

- Full compliance; estimate for unknown artist; use multiple terms for "American or English"	5
- Controlled, but multiple terms per field with punctuation	4
- Term is not controlled by ULAN nationality list, AAT, TGN, or any source; a phrase or several terms are included in one field without punctuation	2.5
- Not blank, but contains "N/A" etc.	0.5
- Missing value	0

#### 6. Vital Dates Creator

- Full compliance, either a strictly CDWA Lite vital date or CCO/CDWA display biography; estimate when uncertain	5
- Incomplete reference to date	2.5
- Not blank, but contains "N/A" etc.	0.5
- Missing value	0



## Museum Data Exchange: Criteria for Scoring

### 7 Role Creator

- Full compliance; "artist" or "maker" are defaults for unknown	5
- Field contains multiple roles, separated by punctuation	4
- Not controlled by AAT or other authority, contains phrase, question marks, other information	2.5
- Not blank, but contains "N/A" etc.	0.5
- Missing value	0

### 8. Display Measurements

- Full compliance	5
- Not blank, but contains "N/A" etc.	0.5
- Missing value	0

### 9. Display Materials

- Full compliance	5
- Not blank, but contains "N/A" etc.	0.5
- Missing value	0

### 10. Display Creation Date

- Full compliance; estimate if exact date is unknown	5
- Not blank, but contains "N/A" etc.	0.5
- Missing value	0

### 11-12. Dates Concatenated (Earliest & Latest scoring)

- Full compliance	5
- Dates are not estimated earlier and later to accommodate <i>ca.</i> , etc.	4
- Standardized date formats or numbers are not used, e.g., question mark, " <i>ca.</i> " etc. are included in this indexing field	2.5
- Not blank, but contains "N/A" etc.	0.5
- Missing value	0

### 13. Location Name

- Full compliance; current location unknown, but former location listed	5
- Geographic location is missing for repository	4
- Data is not controlled	2.5
- Not blank, but contains "N/A" etc.	0.5
- Missing value, e.g., if current location is unknown, this is not indicated	0

### 14 WorkID

- Full compliance	5
- Data is not consistent	2.5
- Not blank, but contains "N/A" etc.	0.5
- Missing value	0

### 15. Subject Term

- Full compliance	5
- Multiple terms in one field, separated by punctuation	4
- At least one subject term is included, but not controlled	2.5
- Not blank, but contains "N/A" etc.;	0.5
- Missing value	0

## Museum Data Exchange: Criteria for Scoring

### 17. Record Type

- Full compliance	5
- Does not use controlled terminology: <i>item, volume, group, subgroup, collection, series, set, multiples, component, whole, part</i>	2.5
- Not blank, but contains "N/A" etc.;	0.5
- Missing value	0

**1. Object Work Type**

**Definition:** The kind of object or work described.

**Possible points for each instance: 5**

	<b>Criteria for scoring:</b>	<b>Score</b>
	Full compliance	5
	Controlled term, but multiple terms per field with punctuation	4
	Term is not controlled by AAT, LC, or Chenhall; a phrase or several terms are included in one field without punctuation	2.5
	Not blank, but contains "N/A" etc.	0.5
	Missing value	0

**Other criteria not scored in this data set:**

Record only one term per field, repeat field as necessary  
 Term should be the most specific appropriate term  
 Use singular for singular work, plural where appropriate  
 Use multiple work types where appropriate  
 Use representative or overall types for groups of works  
 Capitalization and punctuation should follow rules, e.g., lower case except if the term incorporates proper noun

**Location of rules in the standards:**

CCO: 1.2.1.1.1. Recording Work Type

CDWA Light: 1.1. Sub-element: Object/Work

CDWA: 1.2. Object/Work Type

[http://www.getty.edu/research/conducting\\_research/standards/cdwa/1/object.html#Objectwork-Type](http://www.getty.edu/research/conducting_research/standards/cdwa/1/object.html#Objectwork-Type)

	<b>Examples with scores:</b>	
	drawing	5
	judaica	5
	textile	5
	photograph, carte-de-visite [ <i>multiple terms per field w/punctuation</i> ]	4
	decorative framing for a seat cushion [ <i>uncontrolled phrase</i> ]	2.5
	N/A	.5

**2. Title**

**Definition:** Titles, identifying phrases, or names given to a work of art, architecture, or material culture. For complex works, series, or collections, the title may refer to a discrete unit within the larger entity (a print from a series, a photograph in a collection, a panel from a fresco cycle, a building within a temple complex) or it may identify only the larger entity (series, collection, cycle) itself.

**Possible points for each instance: 5**

	<b>Criteria for scoring:</b>	<b>Score</b>
	Full compliance; "untitled" is not full compliance, because title can be constructed	5
	Title is included, but not descriptive	3
	Not blank, but contains "N/A," "untitled," etc.	0.5
	Missing value	0

**Other criteria not scored in this data set:**

Title should not include a full inscription rather than only inscribed title  
 Multiple titles should be included where appropriate, e.g., repository title, artist title, inscribed title, former titles  
 For preferred title, abbreviations should be avoided  
 Capitalization and punctuation should comply; use title case rather than sentence case, not all caps  
 For constructed titles, follow guidelines

**Location of rules in the standards:**

CCO: 1.2.2.1.1. Recording Title

CDWA Light: 2.1.1. Sub-element: Title

CDWA: 3.1. Title Text

[http://www.getty.edu/research/conducting\\_research/standards/cdwa/4/titles.html#RTFToC2](http://www.getty.edu/research/conducting_research/standards/cdwa/4/titles.html#RTFToC2)

	<b>Examples with scores:</b>	
	Madonna and Child	5
	Ecce Homo	5
	Self-Portrait	5
	George Washington	5
	Coronation of Napoleon	5
	Basket ( <i>decorative arts may have not iconographical subject, use object type as title</i> )	5
	Photograph of a Tree	5
	Untitled (Abstract Composition) ( <i>okay, parenthetical subject in title is good</i> )	5
	photograph ( <i>not descriptive of subject content</i> )	3
	untitled	0.5

**3. Display Creator**

**Definition:** The name, brief biographical information, and roles (if necessary) of the named or anonymous individuals or corporate bodies responsible for the design, production, manufacture, or alteration of the work, presented in a syntax suitable for display to the end-user and including any necessary indications of uncertainty, ambiguity, and nuance. If there is no known creator, make a reference to the presumed culture or nationality of the unknown creator.

**Possible points for each instance: 5**

	<b>Criteria for scoring:</b>	<b>Score</b>
	Full compliance; "unknown artist" is okay, but culture is preferred for unknowns, e.g., "unknown Celtic"	5
	Name is included, but biographical context information is missing	3
	Not blank, but contains "N/A" etc.	0.5
	Missing value, e.g., reference to unknown artists are missing	0

**Other criteria not scored in this data set:**

If there is an English version of the name, use it  
 Capitalization should follow guidelines  
 The display name should be in natural order  
 Uncertain attributions should be noted  
 Multiple creators should be included where appropriate  
 Corporate body creator should be included as appropriate  
 Patrons, publishers, and other non-artists should be included as appropriate  
 Artists for groups should be included per guidelines  
 Former attributions should be included  
 Qualifiers and extent should be included

**Location of rules in the standards:**

CCO: 2.2.1.1.1. Recording the Creator

CDWA Light: 3. Element: Display Creator

CDWA: 4.1. Creator Description

[http://www.getty.edu/research/conducting\\_research/standards/cdwa/14creation.html#Creation-Creator](http://www.getty.edu/research/conducting_research/standards/cdwa/14creation.html#Creation-Creator)

	<b>Examples with scores:</b>	
	Wilmarth, Christopher (American sculptor, 1943-1987)	5
	photographer: Byron, Percy C. (English photoengraver, 1878-1959); printed by Byron Company (American photography studio, 1888-1942)	5
	unknown Celtic ( <i>or unknown (Celtic), Celtic, artist unknown (France), etc.</i> )	5
	unknown artist ( <i>or anonymous, etc.</i> )	5
	Giovanni Battista Piranesi [ <i>missing display bio</i> ]	3
	N/A	0.5

**4. Name Creator**

**Definition:** Proper names, appellations, pseudonyms, nicknames, or other identifying phrases by which an individual, group of people, or corporate body is known.

**Possible points for each instance: 5**

	<b>Criteria for scoring:</b>	<b>Score</b>
	Full compliance; "unknown artist", etc. is okay	5
	Controlled, but multiple names in one field	4
	Names are not controlled (should come from a published authority ULAN, Library of Congress, published artist authority)	2.5
	Not blank, but contains "N/A" etc.	0.5

**Other criteria not scored in this data set:**

Names should be inverted where appropriate, and punctuation of the inverted name should be consistent (or store first and last in separate fields)

Unknown artists, or name the culture of creation should be included in Nationality

Variant names should be included in the authority, or a published vocabulary (ULAN) used in retrieval

Capitalization should follow guidelines

Multiple creators should be included where appropriate

**Location of rules in the standards:**

CCO: A.1.2.1.1.1. Recording Names

CDWA Light: 4.1.1.1. Sub-element: Name of Creator

CDWA: 28.2. Person/Corporate Body Name

[http://www.getty.edu/research/conducting\\_research/standards/cdwa/28person.html#name](http://www.getty.edu/research/conducting_research/standards/cdwa/28person.html#name)

<b>Examples with scores:</b>		
	Wilmarth, Christopher	5
	unknown Roman ( <i>Roman, unknown artist, anonymous artist, etc.</i> )	5
	anonymous, Italy, Roman-Bolognese, 17th century ( <i>local control for anonymous, okay</i> )	5
	Ricci, Marco; Ricci, Sebastiano	4
	attributed to John Piranese ( <i>not controlled</i> )	2.5
	N/A	.5

**5. Nationality Creator**

**Definition:** The national, geopolitical, cultural, or ethnic origins or affiliation of the person or corporate body.

**Possible points for each instance: 5**

	<b>Criteria for scoring:</b>	<b>Score</b>
	Full compliance; estimate for unknown artist; use multiple terms for "American or English"	5
	Controlled, but multiple terms per field with punctuation	4
	Term is not controlled by ULAN nationality list, AAT, TGN, or any source; a phrase or several terms are included in one field without punctuation	2.5
	Not blank, but contains "N/A" etc.	0.5
	Missing value	0

**Other criteria not scored in this data set:**

Should be specific at least to level of nation, or more specific if known  
 Should be a historical nationality, if appropriate  
 Multiple nationalities should be included if appropriate  
 Capitalization: Initial caps should be used

**Location of rules in the standards:**

CCO: A.1.2.2.2.1 Recording Nationality

CDWA Light: 4.2. Sub-element: Nationality Creator

CDWA: 28.8. Person Nationality/Culture/Race  
[http://www.getty.edu/research/conducting\\_research/standards/cdwa/28person.html#nationality](http://www.getty.edu/research/conducting_research/standards/cdwa/28person.html#nationality)

	<b>Examples with scores:</b>	
	Italian	5
	Sioux	5
	American (Sioux) <i>[two terms in one field]</i>	4
	probably Sioux <i>[or with question mark, e.g., Inuit ?; Switzerland, born Germany]</i>	2.5
	N/A	.5

**6. Vital Dates Creator**

**Definition:** A description of the lifespan of the person or the existence of the corporate body, using "ca." and any other expressions of uncertainty or nuance. For Birth and Death date attributes, record years of birth and death, estimated where necessary. For a corporate body, use birthdate and deathdate to record the dates of founding and dissolution. The dates may be combined with nationality and/or life roles in a Display Biography; this is also compliant with the rules.

**Possible points for each instance: 5**

**Criteria for scoring:**

Full compliance, either a strictly CDWA Lite vital date or CCO/CDWA display biography; estimate when uncertain	5
Incomplete reference to date	2.5
Not blank, but contains "N/A" etc.	0.5
Missing value	0

**Other criteria not scored in this data set:**

Should include controlled, indexed Birth Date and Death Date with Vital Dates, per CDWA Lite. This dataset did not include Birth and Death Date. Should be like this: <cdwalite:vitalDatesCreator birthdate="1435" deathdate="1525"> ca. 1440-after 1522</cdwalite:vitalDatesCreator>

**Birth Date:** The date when the person was born, or when a corporate body came into existence. If only the date of first activity is known, or if the date of birth is uncertain or approximate, this is the earliest possible or estimated year when the artist was born or the corporate body came into existence. **Death Date:** The date when the person died, or when a corporate body ceased to exist. If only the date of last activity is known, or if the date of death is uncertain or approximate, this is the latest possible or estimated year when the artist died or the corporate body ceased to exist.

Birth and Death Date should be formatted as dates or numbers per rules in CCO and CDWA

Estimate earlier and later years for Birth and Death Date to accommodate *ca.*, etc.

**Location of rules in the standards:**

CCO: A.1.2.2.1.1 Recording Display Biography;  
A.1.2.2.3. Rules for Birth/Start Dates and Death/End Dates

CDWA Light: 4.3. Sub-element: Vital Dates Creator

CDWA: 28.3 Display Biography [http://www.getty.edu/research/conducting\\_research/standards/cdwa/28person.html#biography](http://www.getty.edu/research/conducting_research/standards/cdwa/28person.html#biography)

28.4. Birth Date [http://www.getty.edu/research/conducting\\_research/standards/cdwa/28person.html#birth](http://www.getty.edu/research/conducting_research/standards/cdwa/28person.html#birth)

28.5. Death Date [http://www.getty.edu/research/conducting\\_research/standards/cdwa/28person.html#death](http://www.getty.edu/research/conducting_research/standards/cdwa/28person.html#death)



Museum Data Exchange: Criteria for Scoring

	<b>Examples with scores:</b>	
	American architect, 1877-1923	5
	French, probably 1766-died 1821/1822	5
	1498-1544	5
	1780	2.5
	N/A	0.5

**7. Role Creator**

**Definition:** The role or activity performed by a creator or maker in the conception, design, or production of a work.

**Possible points for each instance: 5 \***

<b>Criteria for scoring:</b>		
Full compliance; "artist" or "maker" are defaults for unknown		5
Field contains multiple roles, separated by punctuation		4
Not controlled by AAT or other authority, contains phrase, question marks, other information		2.5
Not blank, but contains "N/A" etc.		0.5
Missing value		0

**Other criteria not scored in this data set:**

Express terms in lower case  
Use the most specific term applicable

**Location of rules in the standards:**

CCO: 2.2.2.1.1. Choosing the role

CDWA Light: 4.1.1.1. Sub-element: Name of Creator

CDWA: 4.1.4. Creator Role

[http://www.getty.edu/research/conducting\\_research/standards/cdwa/14creation.html#Creator-Role](http://www.getty.edu/research/conducting_research/standards/cdwa/14creation.html#Creator-Role)

<b>Examples with scores:</b>		
painter		5
artist		5
designer		5
architect; draftsman		4
designed house (or other non-role words and phrases such as "after," "printer?", etc.)		2.5
N/A		0.5

**8. Display Measurements**

**Definition:** Information about the dimensions, size, or scale of the work, presented in a syntax suitable for display to the end-user and including any necessary indications of uncertainty, ambiguity, and nuance. It may include the scale of the work. It may also include the number of the parts of a complex work, series, or collection.

**Possible points for each instance: 5**

	<b>Criteria for scoring:</b>	<b>Score</b>
	Full compliance	5
	Not blank, but contains "N/A" etc.	0.5
	Missing value	0

**Other criteria not scored in this data set:**

Data is not consistent in syntax, terminology, etc.

Capitalization and syntax follows guidelines, including use of lower case, h x w x d, etc.

Extra credit awarded for indexing these to allow retrieval

Display both metric and imperial units, if possible. If using only one type of unit of measurement, use metric units because they are more universally understood. If your audience is solely American, use feet and inches

**Location of rules in the standards:**

CCO: 3.2.1.1.1. Recording measurements

CDWA Light: 5. Element: Display Measurements

CDWA: 6.1. Dimensions Description

[http://www.getty.edu/research/conducting\\_research/standards/cdwa/7/measurements.html#RTFToC2](http://www.getty.edu/research/conducting_research/standards/cdwa/7/measurements.html#RTFToC2)

	<b>Examples with scores:</b>	
	overall: 50.8 x 76.3 cm (20 x 30 1/16 in.)	5
	codex: 20.32cm x 15.71cm x 6.98cm	5
	16 3/4 x 12 5/8 in. (42.55 x 32.07 cm) (portfolio)	5
	various	5
	unmeasured	0.5

**9. Display Materials Technique**

**Definition:** An indication of the substances or materials used in the creation of a work, as well as any implements, production or manufacturing techniques, processes, or methods incorporated in its fabrication, presented in a syntax suitable for display to the end-user and including any necessary indications of uncertainty, ambiguity, and nuance. For works on paper, descriptions of watermarks may also be included.

**Possible points for each instance: 5**

	<b>Criteria for scoring:</b>	<b>Score</b>
	Full compliance	5
	Not blank, but contains "N/A" etc.	0.5
	Missing value	0

**Other criteria not scored in this data set:**

Display syntax is not consistent  
 Extra credit for indexing these for retrieval, using AAT or other controlled vocabularies to provide access

**Location of rules in the standards:**

CCO: 3.2.2.1.1. Recording Materials and Techniques

CDWA Light: 7. Element: Display Materials/Techniques

CDWA: 7.1. Materials/Techniques Description  
[http://www.getty.edu/research/conducting\\_research/standards/cdwa/8materials.html#RTFToC2](http://www.getty.edu/research/conducting_research/standards/cdwa/8materials.html#RTFToC2)

	<b>Examples with scores:</b>	
	bronze	5
	soft-paste porcelain, painted in enamels and gilt	5
	oil on canvas	5
	gelatin silver print (mounted to board)	5
	N/A	0.5

**10. Display Creation Date**

**Definition:** A concise description of the date or range of dates associated with the creation, design, production, presentation, performance, construction, or alteration of the work or its components, presented in a syntax suitable for display to the end-user and including any necessary indications of uncertainty, ambiguity, and nuance.

**Possible points for each instance: 5**

	<b>Criteria for scoring:</b>	<b>Score</b>
	Full compliance; estimate if exact date is unknown	5
	Not blank, but contains "N/A" etc.	0.5
	Missing value	0

**Other criteria not scored in this data set:**

Record the date of design or creation  
Follow guidelines lower case and other format as described in guidelines

**Location of rules in the standards:**

CCO: 4.2.3.1.1. Recording Date

CDWA Light: 12. Element: Display Creation Date

CDWA: 4.2. Creation Date

[http://www.getty.edu/research/conducting\\_research/standards/cdwa/14creation.html#CREATION-DATE](http://www.getty.edu/research/conducting_research/standards/cdwa/14creation.html#CREATION-DATE)

	<b>Examples with scores:</b>	
	1765	5
	late 19th century	5
	ca. 55 BCE	5
	begun 1301/1306	5
	contemporary	5
	N/A	0.5

---

**11-12: Dates Concatenated: Earliest and Latest Dates scored together**

---

**11. Earliest Date**

**Definition:** The earliest possible date when the work of art or architecture was created. For works that were created over a span of time, this is the year when the work was designed or when execution was begun. For uncertain or approximate dates, this is the earliest possible year when the work could have been begun or designed.

**12. Latest Date**

**Definition:** The latest possible date when the work of art or architecture was created. For works that were created over a span of time, this is the year when the work was completed or when the structure was dedicated. For uncertain or approximate dates, this is the latest possible year when the work could have been completed.

**11-12. Dates Concatenated (Earliest & Latest scoring)**

<b>Criteria for scoring both together:</b>		
Full compliance		5
Dates are not estimated earlier and later to accommodate <i>ca.</i> , etc.		4
Standardized date formats or numbers are not used, e.g., question mark, " <i>ca.</i> " etc. are included in this indexing field		2.5
Not blank, but contains "N/A" etc.		0.5
Missing value		0

**Other criteria not scored in this data set:**

**Location of rules in the standards:**

CCO: 4.2.3.1.1. Recording Date

CDWA Light: 13.1.2. Sub-element: Earliest Date

CDWA Light: 13.1.3. Sub-element: Latest Date

CDWA: 4.2.1. Earliest Date [http://www.getty.edu/research/conducting\\_research/standards/cdwa/14creation.html#CREATION-EARLIEST](http://www.getty.edu/research/conducting_research/standards/cdwa/14creation.html#CREATION-EARLIEST)

CDWA: 4.2.2. Latest Date [http://www.getty.edu/research/conducting\\_research/standards/cdwa/14creation.html#CREATION-LATEST](http://www.getty.edu/research/conducting_research/standards/cdwa/14creation.html#CREATION-LATEST)

<b>Examples with scores:</b>		
Earliest: 1876 Latest: 1876		
Earliest: 1800 Latest: 1900		
Earliest: -323 Latest: -304		
N/A		0.5

**13. Location Name**

**Definition:** The name of the repository that currently houses the work and the geographic place where the work is located, including the location of the repository, the site of a building other large work, or the place where performance took place. If the work is lost, stolen, or destroyed, this subcategory identifies the last known repository or geographic location, or discovery location, and states that the work has been lost, stolen, or destroyed, or that its current repository is unknown.

**Possible points for each instance: 5**

	<b>Criteria for scoring:</b>	<b>Score</b>
	Full compliance; current location unknown, but former location listed	5
	Geographic location is missing for repository	4
	Data is not controlled	2.5
	Not blank, but contains "N/A" etc.	0.5
	Missing value, e.g., if current location is unknown, this is not indicated	0

**Other criteria not scored in this data set:**

Rules of capitalization and punctuation should be followed

**Location of rules in the standards:**

CCO: 5.2.1.1.1. Types of Location

CDWA Light: 14.1.1. Sub-element: Location/Repository Name

CDWA: 21.2. Current Repository/ Geographic Location

[http://www.getty.edu/research/conducting\\_research/standards/cdwa/26current.html#repository\\_name](http://www.getty.edu/research/conducting_research/standards/cdwa/26current.html#repository_name)

	<b>Examples with scores:</b>	
	Philadelphia Museum of Art (Philadelphia, Pennsylvania, USA)	5
	Italy, Rome	5
	Urban Art Museum ( <i>geographic location is missing</i> )	4
	Modern Art Museum, Gift from the Collection of J. Mellows, 34.567.33 ( <i>not controlled</i> )	2.5
	N/A	0.5

**14. Work ID number**

**Definition:** Any unique identifiers assigned to a work by the current or last known repository.

**Possible points for each instance: 5**

	<b>Criteria for scoring:</b>	<b>Score</b>
	Full compliance	5
	Data is not consistent	2.5
	Not blank, but contains "N/A" etc.	0.5
	Missing value	0

**Other criteria not scored in this data set:**

**Location of rules in the standards:**

CCO: discussion of *Current Location* at 5.1.1. Discussion

CDWA Light: 14.1.2. Sub-element: Repository Work Identification Number

CDWA: 21.3. Current Repository Numbers

[http://www.getty.edu/research/conducting\\_research/standards/cdwa/26current.html#Repository\\_Numbers](http://www.getty.edu/research/conducting_research/standards/cdwa/26current.html#Repository_Numbers)

	<b>Examples with scores:</b>	
	PD-1987-21	5
	09.74a,b	5
	1.1965.1014	5
	N/A	0.5



**15. Subject Term**

**Definition:** Indexing terms that characterize what the work depicts or what is depicted in it, including generic terms and proper names.

**Possible points for each instance: 5**

	<b>Criteria for scoring:</b>	<b>Score</b>
	Full compliance. At minimum, should list a general subject: <i>advertising and commercial, allegory, animal, apparel, architecture, botanical, cartographic, ceremonial object, cityscape, didactic and propaganda, funerary art, genre, history and legend, human figure, interior architecture, landscape, literary theme, machine, military, mixed motif, nonrepresentational art, object (utilitarian), performance, portrait, religion and mythology, seascape, still life</i>	5
	Multiple terms in one field, separated by punctuation	4
	At least one subject term is included, but not controlled	2.5
	Not blank, but contains "N/A" etc.;	0.5
	Missing value, e.g., if current location is unknown, this is not indicated	0

**Other criteria not scored in this data set:**

Terms are lower case, proper names are initial caps

**Location of rules in the standards:**

CCO: 6.3.1.1.1. Recording the Subject

CDWA Light: 15.1.2. Sub-element: Indexing Subject Term

CDWA: 16.2. Subject Indexing Terms

[http://www.getty.edu/research/conducting\\_research/standards/cdwa/18subject.html#RTFToC3](http://www.getty.edu/research/conducting_research/standards/cdwa/18subject.html#RTFToC3)

	<b>Examples with scores:</b>	
	landscape	5
	Madonna and Child	5
	U.S. Civil War	5
	The Sleeping Beauty	5
	Admiral Lord Horatio Nelson	5
	landscape; portraits	4
	Woman with a sleeping child on her lap ( <i>not controlled format</i> )	2.5
	N/A	0.5

**17. Record Type**

**Definition:** An indication of the level of cataloging represented by the record, based on the physical form or intellectual content of the material.

**Possible points for each instance: 5**

	<b>Criteria for scoring:</b>	<b>Score</b>
	Full compliance	5
	Does not use controlled terminology: <i>item, volume, group, subgroup, collection, series, set, multiples, component, whole, part</i>	2.5
	Not blank, but contains "N/A" etc.;	0.5
	Missing value, e.g., if current location is unknown, this is not indicated	0

**Other criteria not scored in this data set:**

Use term appropriate to what is being cataloged  
Capitalization: use lowercase.

**Location of rules in the standards:**

CCO: 1.2.1.2.5 Record Type

CDWA Light: 21.2. Sub-element: Record Type

CDWA: 1.1. Catalog Level

[http://www.getty.edu/research/conducting\\_research/standards/cdwa/1object.html#RTFTOC2a](http://www.getty.edu/research/conducting_research/standards/cdwa/1object.html#RTFTOC2a)

	<b>Examples with scores:</b>	
	item	5
	group	5
	N/A	0.5