

Introduction to Art Image Access

Issues, Tools, Standards, Strategies

Glossary

about-ness

An expression used in the context of subject description to refer to thematic or symbolic meaning intended by a work of art. *About-ness* corresponds to Erwin Panofsky's *interpretation*, the third level of subject description specified in *Categories for the Description of Works of Art*. For example, "The Medusa has been interpreted as a symbol of everything from the female mysteries to the creativity and destruction inherent in Nature. A more recent manifestation is the character played by Glenn Close in the 1987 film *Fatal Attraction*, where the Medusa figure is seen as the devourer and potential destroyer of men." Compare *of-ness*.

access point

A database field or metadata category (for example, Title, Subject, Creator) designed to be searchable and retrievable by an end-user, as distinguished from information that is intended primarily for display. Compare *display data*.

algorithm

A formula or set of steps devised to solve a particular problem. In an automated environment, an algorithm is a set of rules, expressed in programming language, to accomplish a particular task.

associative relationship

In a thesaurus, a link to a concept peripherally related to the concept at hand. For example, *arriccio* and *intonaco* (types of plaster) are linked to *fresco* (a type of wall painting) in the AAT via an associative relationship rather than a *hierarchical* or *equivalence relationship*.

authority (also called authority file)

A file or set of terms extrinsic to the object or document being described. When used for retrieval, this file or set of terms is more efficiently recorded in separate linked files rather than in individual records about the object itself. The advantage of storing ancillary information in an authority is that this information needs to be recorded only once, as it may then be linked to all appropriate object/work records. An authority record typically contains the preferred name or term for a particular concept, as well as synonyms and additional information.

authority control (also called vocabulary control)

The process of governing the form and content of information in an object record, including proper names and other terminology, typically by requiring the use of designated syntax or a *preferred term*, with the goal of ensuring that all citations and references to a given subject or concept are the same. See also *controlled vocabulary*.

Boolean operators

Specific methods of logic, syntax, and terms used in programming and retrieval in databases to manipulate TRUE/FALSE values. For example, "SELECT term FROM term table WHERE value LIKE "Magi" OR ("Three" AND "Kings") OR ("Wise" AND "Men")." In this example, the query is asking for all records with the words "Magi" or "Three Kings" or "Wise Men" but only if the words "King" and "Men" are part of text strings that include the words "Three" and "Wise." Boolean operators, originally formulated in 1847 by George Boole of England, are derived from the symbolic system of mathematical logic that represents relationships among entities.

cataloguing

The process of recording information about a work of art or architecture and/or its visual surrogate, usually including a systematic description, subject analysis, and the assignment of a classification notation and/or terms from one or more vocabularies.

content-based image retrieval (CBIR)

A computer technique for retrieving images based on automatically derived features such as color, texture, and shape.

controlled vocabulary

An established list of terms from which an indexer or cataloguer may select when assigning descriptors or subject headings to a record representing a work. See also *authority control*.

crosswalk

A chart or table that represents the semantic mapping of fields or data elements in one metadata schema to fields or data elements in one or more other metadata schemas. Crosswalks make it possible to search heterogeneous information resources simultaneously with a single query as if they were in a single system (*semantic interoperability*), and to convert data from one metadata schema to another. See also *metadata mapping*.

data structure

The interrelationships of information in a database or information system. At the broadest level, *data structure* includes definitions of entire entities or files, the relationships among entities, the categories or attributes that define the entities, and the relationships of categories to other categories. Examples of database structures include the *flat file structure* and the *relational file structure*.

decorative arts

Artworks that serve utilitarian as well as esthetic purposes, typically involving the decoration and embellishment of household utilitarian objects, including furniture and eating utensils.

descriptive metadata

Metadata used to describe or identify an information object or information resource.

display data

Data that are typically intended to be read as part of a record that has been retrieved by an end-user; not necessarily subject to terminology control, and not intended for use in retrieval. Compare *access point*.

entity

A thing of significance—something distinct or self-contained—about which information needs to be known, recorded, or stored. *Entity* may also refer to principal conceptual entities in an information system, whether or not they are actually reflected in the database structure. To a cataloguer, a corporate body—that is, any group of people working together, not necessarily a legally incorporated *entity*—is considered an entity. In another sense, a record in a catalogue is *metadata* about the entity it describes.

equivalence relationship

The one-to-one relationship between concepts that are equal, as, for example, the relationship between terms or names that represent the same concept in a thesaurus, as *truck* (U.S.) and *lorry* (U.K.).

field

A basic unit of a record, typically a space reserved for the recording of a particular predetermined category of information. The placing of information in fields imposes a data structure on the contents of a file, which makes them more readily searchable and hence retrievable. Database fields often correspond to individual elements of a particular *metadata schema*.

figural art

Art that includes depictions of human figures.

flat file

A computerized file in which the fields that make up a record are held in a single file. Files of this type tend to be relatively simple to use, but are not well suited to applications where much of the information is hierarchical and/or where multiple occurrences of fields are required. Compare *relational file*.

free text

Text in natural language, typically unstructured and used for description and display. Compare *controlled vocabulary*.

genre scene

A visual representation that depicts a scene from everyday life.

granularity

The level of detail at which an information object or resource is described or displayed.

hierarchical relationship

The relationship between broader and narrower concepts in a thesaurus or information system. Hierarchical relationships may be whole/part, where all linked entities are part of the whole (for example, *Cádiz*, *Granada*, *Málaga*, and *Sevilla* are all part of *Andalucía*), or genus/species, where all links represent entities that are a type of the original entity (for example, *Adirondack chairs*, *bergères*, and *great chairs* are all types of *armchairs*).

iconography

The description, classification, analysis, and interpretation of visual subject matter in works of art. Also, the imagery used to depict a particular subject, including its symbolic meaning. Compare *of-ness*.

iconology

The description, classification, or analysis of meaning or symbolism in the visual arts that takes into account the tradition of pictorial motifs and their historical, cultural, and social meanings. Compare *about-ness*.

indexing

The human computerized process of making a list of terms, names, dates, and other data that are stored in a structured data file and used to enhance access to full information on a given concept or topic. Indexing terms represent the most salient information necessary to retrieve the record; they are often taken from a *controlled vocabulary*.

information object

A digital item or group of items, regardless of type or format, referred to as a unit that a computer can address or manipulate as a single object.

keyword

A word or phrase that can be used as a search term in querying an online resource.

keyword searching

Searching in an automated environment using individual terms or combinations of words, terms, or component parts of multiword phrases, often with the implementation of *Boolean operators*. Keyword searching contrasts with *exact matching*, which is a search technique that seeks phrases that match the user's query word for word, maintaining the original order, and sometimes the punctuation and capitalization of words. In retrieval of Web pages or other documents outside a database, an algorithm is often developed to locate keywords in texts.

metadata

Literally, "data about data" or data categories. Metadata can include data associated with either an information system or an information object for purposes of description, administration, legal requirements, technical functionality, use and usage, and preservation. The present volume focuses on *descriptive metadata*.

metadata mapping

A formal identification of equivalent or nearly equivalent metadata elements or groups of metadata elements within different metadata schemas, carried out to facilitate *semantic interoperability*. Also known as *field mapping*.

metadata schema (plural schemas, schemata)

A set of rules for recording or encoding information that supports a specific metadata element set.

narrative

Subject matter that refers to a story or sequence of events.

nonrepeating field

A field in a database that can have only one occurrence, and hence only one data value—for example, the field for the unique identification number of a work of art. Compare *repeating field*.

nonrepresentational art (also called nonfigurative art, nonobjective art)

Art that presents a visual form with no specific reference to anything outside itself.

normalization

The removal of punctuation, capitalization, and diacritics in the data targeted for retrieval and in the words or text typed by a user in a query.

object

In the context of this publication, *object* refers to an artwork or other museum object, a work of architecture, an archival document, or an archaeological artifact.

of-ness

An expression used in the context of subject description to refer to what a work of art depicts. *Of-ness* may also include a specific identification, as well as a generic description, of what is depicted in a work of art. This expression corresponds to Erwin Panofsky's *identification* and *description*, the first two levels of subject description outlined in *Categories for the Description of Works of Art*. For example, "The artwork is of a woman with snakes for hair" (*description*). "The artwork is of the gorgon Medusa, the creature in classical mythology who turned her beholders to stone" (*identification*). Compare *about-ness*.

optional fields (or optional metadata elements, optional categories)

Categories of information that complement *required or core fields*; when these fields are provided, a fuller, more detailed description or record is achieved.

Panofsky, Erwin

German-born American art historian (1892–1968) best known as a scholar of late medieval and Renaissance art of Northern Europe and Italy, and also known for having developed theories regarding iconography.

precision

In database searching, a measure of the effectiveness of retrieval expressed as the ratio of relevant items retrieved in response to a specific query to the total number retrieved. Compare *recall*. High precision indicates that from all the items retrieved by a query, a large proportion was relevant to the search. Precision in a search yields more accurate results.

preferred term (also called preferred name, descriptor)

A term or phrase that is flagged and may be used consistently to represent a particular concept or entity in records. In many systems, the preferred term is what appears in displays, while variant or nonpreferred terms are used as *access points* "behind the scenes." When cataloguing policy or technical limitations (for example, the lack of a linked authority) require that all variations of a word or name be cross-referenced to one and only one term or name, the preferred term is given favor. In such systems, users must look for the concept or entity under that preferred form if they wish to find all the related information available in the database.

recall

In database searching, a measure of the effectiveness of retrieval expressed as the ratio of the total number of relevant entries or documents in a given database to the number of relevant entries or documents retrieved in response to a specific query. Compare *precision*. High recall indicates that a large number of all the items that were possibly relevant to a query were retrieved. Recall in a search yields a greater number of results.

record

A coherent set of data elements or fields in a computerized file. The term "record" may also be used to refer to a coherent intellectual concept (for example, the concept of "person" or "object").

relational file

A computerized file in which the fields that combine to make up a record are held in a number of different files or tables. This type of structure is well suited to applications that require the ability to relate and combine pieces of information held in a number of files for both search and retrieval. In a relational structure, any hierarchical or other relationships between elements of records can be modeled explicitly and reproduced in the structure. (In a *flat file*, these relationships are implicit.)

relationships

Links between terms in a given record or between records, often stored in relational tables (*a relational file structure*).

repeating field

A field that may occur more than once in a record, with different data values in the different occurrences. For

example, the field corresponding to an artist's life role could repeat with different values, such as *sculptor*, *painter*, *printmaker*, and so on. Compare *nonrepeating field*.

representational art (also called figurative art, objective art)

Art that portrays, even if in altered or distorted form, things that can be perceived in the visual world.

required field (also called core field, core category)

A database field or category of data that is part of the minimum set of information necessary to describe and retrieve information about an object or concept.

retrieval

The process of finding information in an automated environment, typically by using two main techniques: (1) matching words or phrases in the query against the database index (*searching*) and (2) traversing the database with the aid of hypertext or hypermedia links (*navigating*). The retrievability of information is determined to a large extent by the care taken to make it accessible by using such tools as data standards, vocabularies, and the design of the data structure of the particular information system.

search engine

A computer program that allows users to search information resources. In the context of the World Wide Web, search engines enable users to search large indexes of Web pages generated by an automated Web crawler or spider.

semantic interoperability

The ability to search for digital information across heterogeneous distributed databases whose metadata schemas have been mapped to one another.

sight size (also called sight measurement)

The dimensions of an object based upon the area that is visible. *Sight size* is most often used with framed paintings or works of art on paper to refer to the dimensions of the area visible within the window of the frame or mat, as distinguished from measurements taken of the entire canvas, panel, or piece of paper that forms the support of the object.

structured vocabulary

A type of controlled vocabulary in which words, phrases, or other terminology are structured to show relationships between terms and concepts. The primary functions of a structured vocabulary are to provide context for the terms and to allow better retrieval in information systems.

subject (also subject matter)

As defined in *Categories for the Description of Works of Art*, the subject matter of a work of art (sometimes referred to as its *content*) is the narrative, iconic, or nonobjective meaning conveyed by an abstract or a figurative composition. It is that depicted in and by a work of art. *Subject* may also be interpreted as the function or object type of an object or work of architecture that otherwise has no narrative content. Compare *about-ness*, *iconography*, *iconology*, and *of-ness*.

subject access

The process of providing or finding information about an artwork or other object by means of its subject—that is, what is portrayed in or on the object.

subject analysis

The interpretational process by which a work of art or its surrogate image is analyzed and explicit subject data or access points are created for the record representing the work.

subject heading

A word or phrase assigned in a descriptive metadata record to indicate a subject of a work, which serves as an *access point* in a catalogue, index, or database search. Broad subject headings may have narrower subdivisions, which are often concatenated to create more specific descriptors (for example, *Islamic Art-India-Exhibitions*).

syndetic structure

A structure in a *controlled vocabulary* in which logical and semantic relationships between descriptors or headings (for example, *broader term*, *narrower term*, *related term*) are explicitly indicated.

thesaurus

A *structured vocabulary* made up of names, words, and other information, typically including synonyms and/or hierarchical relationships for the purpose of cross-referencing in order to organize a collection of concepts for reference and retrieval. Standards for thesaurus construction are the *Guidelines for the Construction, Format, and Management of Monolingual Thesauri: An American National Standard*, ANSI/NISO Z39.19-1993 (R1998), from the National Information Standards Organization; and the related *Documentation-Guidelines for the Establishment and Development of Monolingual Thesauri*, 2d ed., ISO 2788, dated 1 January 1986, from the International Organization for Standardization.

typology

Classification based on a comparative analysis of structural or other characteristics.

variant (*also called variant term, variant name*)

A nonpreferred term or name that has an *equivalency relationship* to the *preferred term* or *descriptor*.

visual document (*also called surrogate image, visual surrogate*)

A photograph, slide, digital image file, or other reproduction of a work of art.