



Subject Cataloging of Images in Museums and Photo Collections

Angela Kailus

RSC Satellite Meeting, Frankfurt/Main, 4.11.2016

Deutsches Dokumentationszentrum für Kunstgeschichte
Bildarchiv Foto Marburg – Philipps-Universität Marburg



German Documentation Centre for Art History – Bildarchiv Foto Marburg

- Central facility at Philipps-University Marburg, Germany
- Documentation and research institute for visual cultural assets
- Task: documenting Europe's monumental cultural heritage for scientific research and enabling access to it (art history and related disciplines)
- Collection: 2.2 m original photos on art and architecture in Europe and adjoining countries
- Maintenance of expert web portals publishing content from approx. 100 LAM partners
- Enhancing documentation standards and data interoperability in the LAM sector:
Participating in data quality initiatives in Europeana and German Digital Library contexts



First of all, I would like to introduce my institution. The German Documentation Centre of Art History – Bildarchiv Foto Marburg is a documentation and research institute for visual cultural assets. It is a central facility at Philipps-University Marburg, Germany.

One of its tasks is to document Europe's monumental cultural heritage for scientific research in art history and the related disciplines, and enabling access to it.

The Bildarchiv has a collection of 2.2 m original photos on art and architecture in Europe and adjoining countries, dating from as early as 1860, and increasingly growing by born-digital stocks.

Documentation of art and architecture, standardization and enabling data interoperability has been an ongoing mission for more than 25 years.

In this field we have collaborated with institutions from various sectors, namely museums, libraries, research institutions, image libraries, archives and state offices for historical monuments.

There have been numerous joint digitization projects. The Bildarchiv runs expert web databases publishing content from approx. 100 partners from heritage institutions. These are *Image Index of Art and Architecture*, *Manuscripta Mediaevalia*, and *Digital Portrait Index*.

A priority is the development of the LIDO XML harvesting format, which aims at achieving interoperability for museum collection data in data exchange and publication contexts. Here, we are increasingly involved in data quality projects and working groups in the context of Europeana and the German Digital Library.

My role here is precisely in this field: data management and web publication, standardization and interoperability.

What is an image?



paintings



drawings



graphic art works



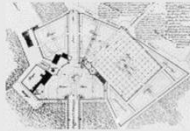
book illustrations



manuscript
illuminations



posters



architectural plans
and drafts



pamphlets



exlibris



photographs / image
files



postcards



emblems



sketches



silhouettes



portraits

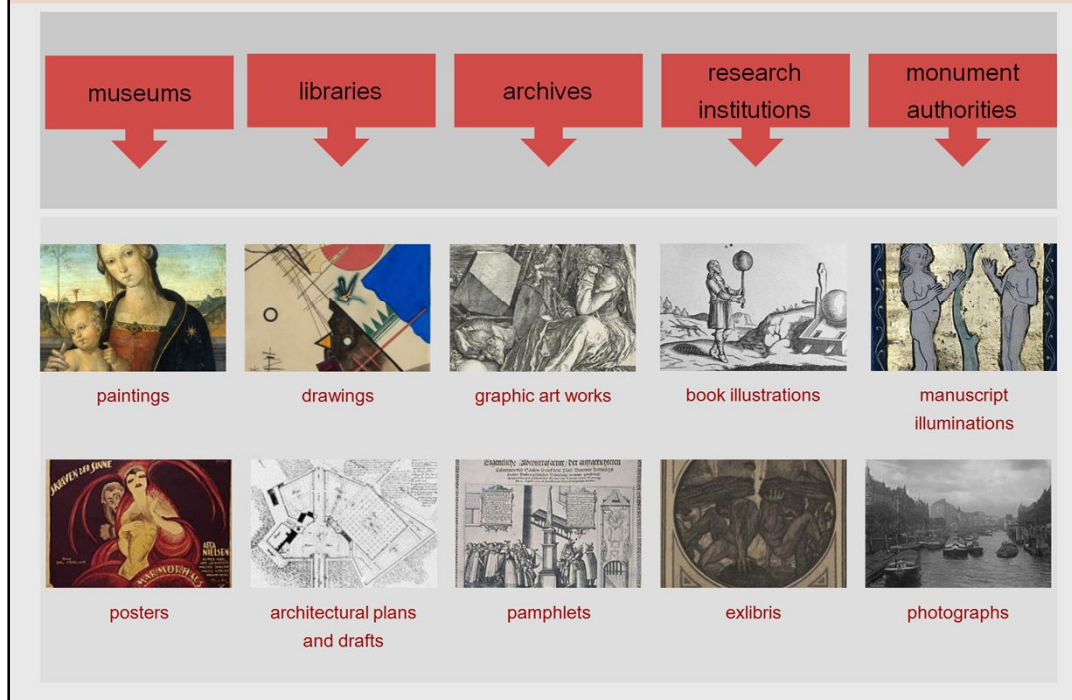
What is an image? In collections, you will come across them in various forms. Some will come to mind instantly as self-contained art works, such as paintings, drawings, fine art prints or art photographs.

Others are part of a complex work, such as book illustrations or illuminations of medieval manuscripts. Some appear in works linked to more specific functional contexts, such as posters, postcards, pamphlets or exlibris. Or their representation of a subject may serve a specific purpose with regard to another object, as it is with drafts, sketches, plans or documentary photos.

When referring to images in this talk, I would like to restrict myself to two-dimensional still images which can be documented as part of a collection. Their physical form can be either analog or digital.

Thus, moving images are not included. In addition, 'image' here does not refer to a complex figural artifact, for which the term stands in visual culture studies or 'Bildwissenschaft'.

Two-dimensional visual works in collections



Images can be found in collections of all sectors. Fine art museums of course collect paintings and graphic art, but you will also find a broad variety of image material in museums of cultural, technical or natural history.

Libraries have illustrated books and manuscripts, or their own special collections of prints and drawings. Archives acquire estates containing image material. Monument authorities hold extensive photographic collections on buildings and urban development. In research institutions, you will find large image collections supporting research and teaching.

Whether these institutions value subject access to their images at all, and how they provide it, varies widely according to the documentation practices of the respective sectors.

Identifying the subject of an image: a complex process

- Cataloguing: transforming visual code into textual code
- With help of additional information
- Crucial: interpretation by cataloguer
- There is no „correct“ or „complete“ interpretation



How to identify the subject of an image? This can be a complex process.

In visual works, the content is not primarily pre-structured by language. So, subject cataloguing means to transform visual code into textual code, possibly into controlled terminology suitable for retrieval.

Subject cataloguing relies on additional information for contextualization, for instance a given title, inscriptions, accompanying materials. In any case, information will have to be interpreted by the indexer. Is it correct? Is it complete? Sometimes he or she will have to consult more sources to understand the depicted scene or some elements of it. His or her experience and historical knowledge will play a major role for the description of the image's content.

For determining the subject of a work of art, iconographic knowledge is necessary to identify the imagery, typical motifs and their meaning.

Subject cataloguing to a large extent is interpreting the image.

This applies in particular to works of art. No single interpretation is exhaustive of the meaning of a work. There can be different, competing, and contradictory interpretations of the same artwork. No interpretation is complete.

What does the example image show?

Identifying the subject of an image: a complex process

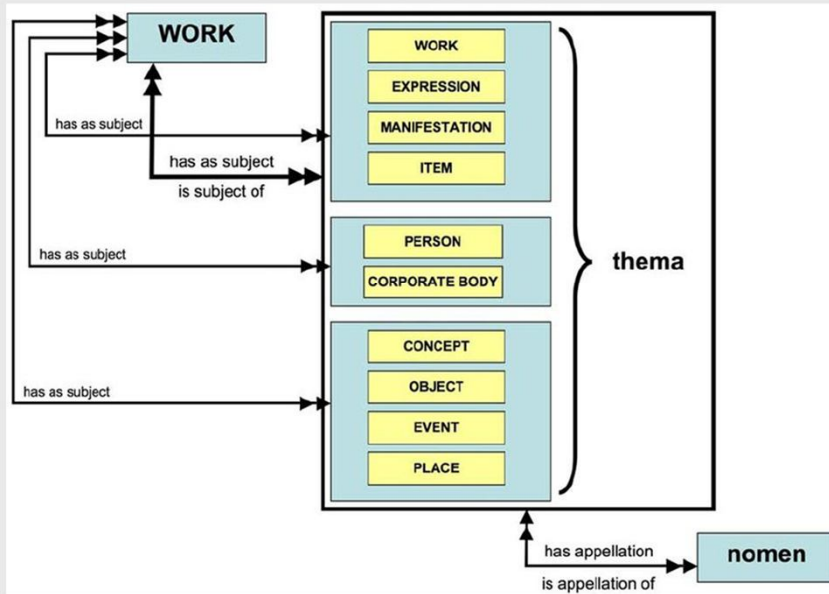
- Cataloguing: transforming visual code into textual code
- With help of additional information
- Crucial: interpretation by cataloguer
- There is no „correct“ or „complete“ interpretation



Evangelist Mark, by Friedrich Sustris (attrib.), Art Collection of Göttingen University, photo: Bildarchiv Foto Marburg, http://www.bildindex.de/document/obj15510695?medium=fmbc25149_08

This drawing by Friedrich Sustris is a representation of the Evangelist Mark, identifiable by his attributes, the halo, the book, the lion.

Subject cataloguing of visual works according to FRSAD



FRSAD's relation to FRBR (with the addition of FRAD entity family)

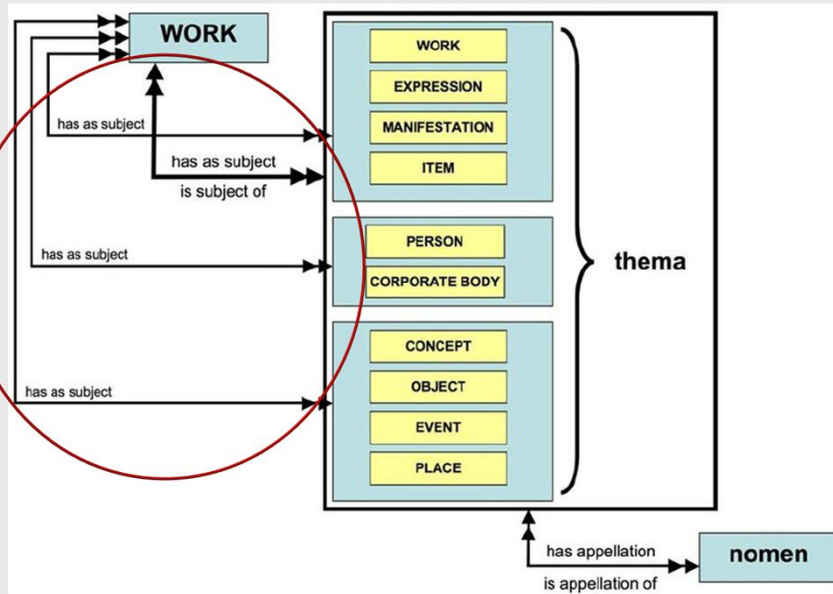
Functional Requirements for Subject Authority Data (FRSAD). A Conceptual Model, Final report 2010, p. 15

You will all be very familiar with this diagram explaining the relation between a work and its *thema*, which can incorporate different entities described in FRBR and FRAD.

Principally, this semantic relation is also correct if the work is an image, under the condition that it has a figurative content. The *thema* can include any of the entities mentioned here.

For an abstract image, we would „recognize“ any of the mentioned entities only with the help of additional information, for instance a title.

Subject cataloguing of visual works according to FRSAD



FRSAD's relation to FRBR (with the addition of FRAD entity family)

Functional Requirements for Subject Authority Data (FRSAD). A Conceptual Model, Final report 2010, p. 15

For images, the established relation between a work and its *thema*, the „has as subject / is subject of“ contains several levels of complexity. We should take a closer look.

Panofsky's three levels of art-historical understanding

Has as subject / is subject of ...

Pre-iconographical description

Primary or natural subject matter, relating to facts or expressions.
necessary: familiarity with depicted matter

Iconographical description

Secondary or conventional subject matter, which forms the realm of imagery, anecdotes and allegories.
Necessary: familiarity with literary or visual traditions, history

Iconographical interpretation

Tertiary or intrinsic meaning or content, which forms the realm of symbolic values (iconology)
necessary: familiarity with history of ideas under changing historical conditions

Erwin Panofsky, *Meaning in the Visual Arts*, 1955

By 1955, the art historian Erwin Panofsky had developed his model of the three levels of art-historical understanding. For art history, his work is still pivotal in understanding the methodology of subject analysis and is still widely recognized.

On the first level, pre-iconographical description deals with primary or natural subject matter. It relates to facts and expressions depicted. This first level is the most basic understanding of a work. A correlation to any added cultural knowledge will not be attempted.

The second level is iconographic analysis. It deals with secondary or conventional subject matter (iconography) such as imagery, anecdotes and allegories. And with historical context. Therefore it requires cultural and iconographic knowledge

On the third level, the iconographical interpretation explores the intrinsic meaning or content: This level aims at a complex understanding of the work as a product of a historical environment, in the history of ideas. The researcher can ask questions like "why did the artist choose to represent the subject in this way?" Essentially, it is the question "what does it all mean?"

Shatford: Ofness and Aboutness

Has as subject / is subject of ...

Ofness

Description of the elements an images depicts picture consists: people, objects, activities, events, places, times.

Improved access by using specific and generic terms in cataloguing

Aboutness

Interpretation of the depicted content by relating it to imagery, allegory and symbols, to historic context

FRSAD (p.10) "... any statement on the "ofness" of a work can be a subject statement and is likely to be the target of a catalogue user's search. The FRSAD model is therefore applicable to situations where ofness statements (depiction) are included in subject access."

Sara Shatford: Analyzing the Subject of a Picture: A Theoretical Approach. Cataloging & Classification Quarterly. 6 (3), 1986, p. 39-62

In the library context, Sara Shatford's adaption of Panofsky's ideas has been widely received. In her work *Ofness and Aboutness* (1986), she assigns the descriptive level of *Ofness* to the depicted people, objects, places, activities or events in an art image and suggests to apply specific and generic descriptive terms to enable subject access.

But, to make sense of an image, frequently it is not really sufficient to look at what it is *of*. In asking what it is *about*, Shatford addresses the level of meaning beyond the depicted matter. When works of art are designed as allegories or symbolic expressions, when a work relies heavily on its historical context, *aboutness* is an essential element of subject analysis.

But FRSAD states that "... any statement on the "ofness" of a work can be a subject statement and is likely to be the target of a catalogue user's search. The FRSAD model is therefore applicable to situations where ofness statements (depiction) are included in subject access." (p.10)

For scholarly searches, subject access on the *ofness* level is not sufficient. By searching for certain nameable and common visual motifs, or illustrations of legendary or literary subjects, scholars expect to find images referring to certain levels of *aboutness*.

But a Panofsky level 3 interpretation generally is beyond the scope of cataloguing. Apart from the fact that this is a time-consuming effort, the results cannot properly be expressed in controlled vocabulary. They will vary widely, and so will rather not support the aim of enabling user access in a consistent manner.

Ofness and Aboutness: Example 1



Theodor Rehbenitz, The Visitation, drawing, 1820, The J. Paul Getty Museum;
<https://www.google.com/culturalinstitute/beta/asset/hwHnL51nhwNiQw>

Level 1 - *of*: two women embracing each other cordially; a wanderer is watching; another man at desk, reading

Level 2 - *about*: Visitation - Mary, pregnant with Jesus, visits her relative Elisabeth, pregnant with John the Baptist. Joseph and Zacharias are present.

Level 3: pictorial tradition of the motif and relation to Christian doctrine; the artist's interpretation of a major medieval motif ...

The subject description of this example would indeed miss the point if it included no reference to the *aboutness* of the depicted scene.

What is this image *of*? The drawing is of two women embracing each other cordially in the hall of a house, two men are visible in the background.

And what is it *about*? It is a specimen of a traditional and very common motif of Christian imagery, the meeting of Mary and Elisabeth, both pregnant, normally referred to as the *visitation*.

Even if the title was not available to the cataloguer, the bible verse at the bottom would provide the key to put the scene into context. If the title was provided, he or she should know that *visitation* is the established name for this motif.

Ofness and Aboutness: Example 2



Level 1: frozen sea; ice jam; HMS Hecla; HMS Griper; ships; 1819

Level 2: North-west-passage; discovery voyage; winter camp

Situation of the H.M.S. Hecla and Griper, 20th September 1819
etching by William Westall after a drawing by F. W. Beechey. In: William Edward Parry: Journal of Voyage for the Discovery of the North-West-Passage from the Atlantic to the Pacific 1819 to 1820, 1821

<https://commons.wikimedia.org/wiki/File:Situation-Hecla-and-Griper-September-20th-1819.jpg>

Sometimes it is not easy to determine the scope of *aboutness* to be covered in cataloguing. The presumable intention for the production of the image should be taken into account.

Here, we have a representation that serves to illustrate a factual historical situation: the winter camp of the two ships of William Parry's voyage to discover the North-West-Passage from 1819 to 1820. The depicted scene is explained in the caption.

The *aboutness* level here should be covered well if the historical context was included.

Ofness and Aboutness: Example 3



Caspar David Friedrich: The Sea of Ice (title given by the artist) / The Wreck of Hope
oil painting, 1823-24, Kunsthalle Hamburg

https://commons.wikimedia.org/wiki/File:Caspar_David_Friedrich_-_Das_Eismeer_-_Hamburger_Kunsthalle_-_02.jpg

Level 1: frozen sea; ice
jam; shipwreck

Level 2: ?

Level 3: permanent
failure; thwarted hope

Caspar David Friedrich's painting *The Sea of Ice*, at first glance, has a similar content, although it is detached from any historical situation. Its allegorical meaning has been so widely recognized that it is reflected in the second title it was given after the artist's death, *The Wreck of Hope* (*Die gescheiterte Hoffnung*).

Friedrich employs no traditional imagery to convey his message. The widely shared interpretation of the scene as an allegory of permanent failure and thwarted hope is a level 3 interpretation. But does level 1 provide a sufficient indexing of the subject?

After highlighting some of the intricacies of subject identification, I will switch to the present practice of subject cataloguing in museums and photo collections.

Subject cataloguing in museums and photo collections: significance and modelling

Lack of universally shared documentation standards!

Museums

- Objective of documentation: inventory of collection for administrative needs
- Title is a core data field, subject is not.
- Distinct subject cataloguing in major museums, for defined stocks

The general situation of documentation differs widely from the library sector.

Although there have been a number of initiatives to establish concepts of standardization in documentation, a generally shared practice has not prevailed. The reason for this is the lack of a central governance and organization structure in the museum sector.

So, after having begun to switch to IT-based cataloguing in the 1990s, still a large number of collections apply local sets of rules and vocabulary. Only recently, after museums started to contribute to Europeana and the German Digital Library, there has been an increasing awareness of the crucial role of standardization in cataloguing to provide satisfactory user access in a portal.

What role does subject cataloguing play? If you check the core data fields usually employed in museum databases, the field ,subject' in many cases will not be among them. A field ,title' or ,description' will bear all textual indications to the contents of the work. This is due to fact that the conventional objective of cataloguing in a museum is to provide an inventory of the collection, to meet the administrative needs.

Nevertheless, a number of museums, namely the major ones which can afford to put expertise and staff resources into a more ambitious level of documentation, do subject cataloguing at least for selected parts of their stocks.

Subject cataloguing in museums and photo collections: significance and modelling

Lack of universally shared documentation standards!

Museums

- Objective of documentation: inventory of collection for administrative needs
- Title is a core data field, subject is not.
- Separate subject cataloguing in major museums, for defined stocks

Documentary photo collections

- Objective of documentation: includes information needs of external users (find, identify, select, obtain)
- Core information on depicted entities generally: who, what, when, where?
- Core information on depicted artifacts: maker, location, work type and subject, location of the artifact
- Overlap: information on depicted entity = information on the photo
- Relevance of separate structured subject cataloguing is recognized
- Application in varying degrees of structured subject cataloguing (from descriptive text to fully formed entity-relation-models)

In photo collections, you see the same lack of generally recognized standards, but the situation is somewhat better. The photo, as a placeholder for the real-world thing and a documentary medium, can only be used if access to its contents is provided.

Core data today generally identify the persons, objects, time and location of the depicted subject.

If photos document works of art, architecture or cultural interest, the cataloguer will try to identify the maker, location, time of production, work type and subject of the depicted artifact.

So, assertions on the *isness/ofness/aboutness* of the artifact become statements on the *ofness/aboutness* of the photo. The photo as a separate medium and collection asset tends to vanish behind its content, it inherits the properties of the depicted entity.

Some photo collections just put that information into a free text field. But generally, there is a wider acceptance of distinct and structured subject cataloguing in databases and a readiness to act accordingly.

Subject cataloguing: vocabularies and authorities

In-house vocabularies and local authority files

Integrated Authority File (GND)

Advantages:

- cross-sectional application of published LOD authority file
- authority managed by reliable institution; sustainable

Problems with use of GND topical headings:

- semantic inconsistencies through use of motif-keywords for depicted entities
- Generic terms: inconsistent
- *GND Systematik*: sometimes not appropriate for art-historical image subject access

- Restrict use of GND topical headings to image material sufficiently indexed on the *offness* level?

If there is controlled subject cataloguing, which vocabularies and authority files are being used?

As it is to be expected in an environment that has only begun to recognize standards, they use in-house vocabularies and local authority files, sometimes mapped to published vocabularies and authorities.

As collections have recognized the need to link up their documentation to published authority files, they have also started to use the Integrated Authority File (GND). – Of course it is welcome and forward-thinking to use the GND beyond libraries. After having started to apply the facet for persons, some museums also started to use the topic headings facet (*Sachschlagworte*) for subject indexing. However, the application practice in some collections is not without problems. Let me explain.

- The GND holds several thousands of topic headings for motifs in literary and art. The frequently observed use of these motif-headings for depicted matter in image cataloguing is a problem. It is based on a disregard of the semantic relations specified in library cataloguing rules for topics (*Regeln für den Schlagwortkatalog*). Works of art carry one instance of a motif, but they are not about a motif. The *visitation* only becomes a visual motif through repeated depiction in a recognizable manner.
- It also has an impact that the GND is a cross-disciplinary authority file. So a number of specific concepts for topics common in visual narratives are still missing (annunciation, wise and foolish virgins, wedding of Cana, cloth of St Veronica, cyclop). Generic search terms (*Oberbegriffe*) are somewhat inconsistent, but that could improve over time.
- Every GND Topic heading is assigned to one or more categories of the *GND Systematik*, a classification of subject categories. This is a proven method of contextualizing subjects of library materials. When used in image cataloguing, the contexts provided by the *GND Systematik* sometimes appear as ahistorical, inappropriate or unspecific. Examples: The *Seven deadly sins* is a concept of Catholic moral doctrine. Several of the enlisted vices (lust, greed, wrath, envy, pride) are solely categorized under psychology, whereas gluttony and sloth are categorized under theological anthropology. - Palm branches in an image are typical attributes of martyrs, the GND just lists them under plant anatomy.

It appears that the GND topic headings might be a good choice for some kinds of image material which have no or an insignificant level of *aboutness*. Portraits, plans, documentary photos of architecture and events could belong to that group.

But they are less appropriate for the specific forms of visual imagery present in art works.

In cross-sectional search contexts, the misalignment of the GND topic headings for this purpose will impair precision and recall of retrieved results.

There is an renowned authority specially developed for subject cataloguing of visual cultural content, this is Iconclass.

Subject cataloguing: vocabularies and authorities

Iconclass


- Classification system developed by Dutch art historians to describe, classify and examine the subjects of visual works, primarily of Western art
- 28.000 hierarchically ordered definitions for concepts, objects, persons, literary and legendary subjects, ordered into 10 main divisions
- Definitions: alphanumerical codes with textual explanation of the subject
- 14.000 keywords help to identify the image subject

Iconclass is a classification system for the purpose of describing, classifying and examining the subjects of visual works, mainly of Western art, developed by art historians.

The model for it was Dewey Decimal Classification. It offers 28.000 hierarchically ordered definitions for concepts, objects, persons, literary and legendary subjects, ordered into 10 main divisions.

Definitions consist of an alphanumerical code, accompanied by a textual explanation of the subject. 14.000 specific keywords help the cataloguer to find the most appropriate definition and later the user to find images with the subjects he or she is interested in.

Subject cataloguing: vocabularies and authorities - Iconclass



0 Abstract, Non-representational Art
 1 Religion and Magic
 2 Nature
 3 Human Being, Man in General
 4 Society, Civilization, Culture
 5 Abstract Ideas and Concepts
 6 History
 7 Bible
 73 New Testament
 73A (scenes from the life of) John the Baptist and Mary
 73A6 Visitation (possibly Joseph and/or Zacharias present) (Luke 1:39-56)
 73A62 Mary's arrival in Judah
 73A622 *Mary and Elisabeth, both pregnant, embracing*
Elisabeth (St.) · Joseph (St.) · Judah (kingdom) · Luke 01 · Mary (Virgin) · New Testament · Visitation · Zacharias · arrival · bible · embracing · pregnancy
 73A622(+0) (+ variant)
 73A622(+1) (+ Holy Trinity)
 73A622(+2) (+ Mary)
 73A622(+3) (+ angel(s))
 73A622(+4) (+ devil(s))
 73A622(+5) (+ donor(s), supplicant(s), whether or not with patron saint(s))

8 Literature
 9 Classical Mythology and Ancient History

<http://www.iconclass.org/>

Here you see how the subject of our example presents itself within the system.

The definition can be retrieved by any of the keywords in *Italics*. If the cataloguer hadn't identified the biblical context or the traditional motif, a keyword derived from the *ofness*-level would have lead him or her to this definition ("*embracing*"). There is a separate definition for each variety of the *visitation* motif, of which this example shows just one.

See the broader context for this specific scene on the upper levels of the hierarchy. Each level is a definition with keywords in its own right. The cataloguer just copies the alphanumeric code to his or her data. Cataloguing can be as specific as possible, as generic context will be inherited from the upper levels, and can be augmented by software.

Subject cataloguing: vocabularies and authorities - Iconclass

- Good scientific coverage of its field: with adequate and differentiated specific vocabulary on the subjects of western art
- Hierarchical classification ensures comprehensive contextualisation of each definition
- Multilingual: English, German, French, Italian, Finnish
- Published on the web (user-friendly / RDF / JSON)
- Open license: unrestricted re-use and distribution of data
- Maintenance and development by a reliable institution (Netherlands Institute for Art History - RKD)
- Increasingly used in art and history museums and photo collections internationally

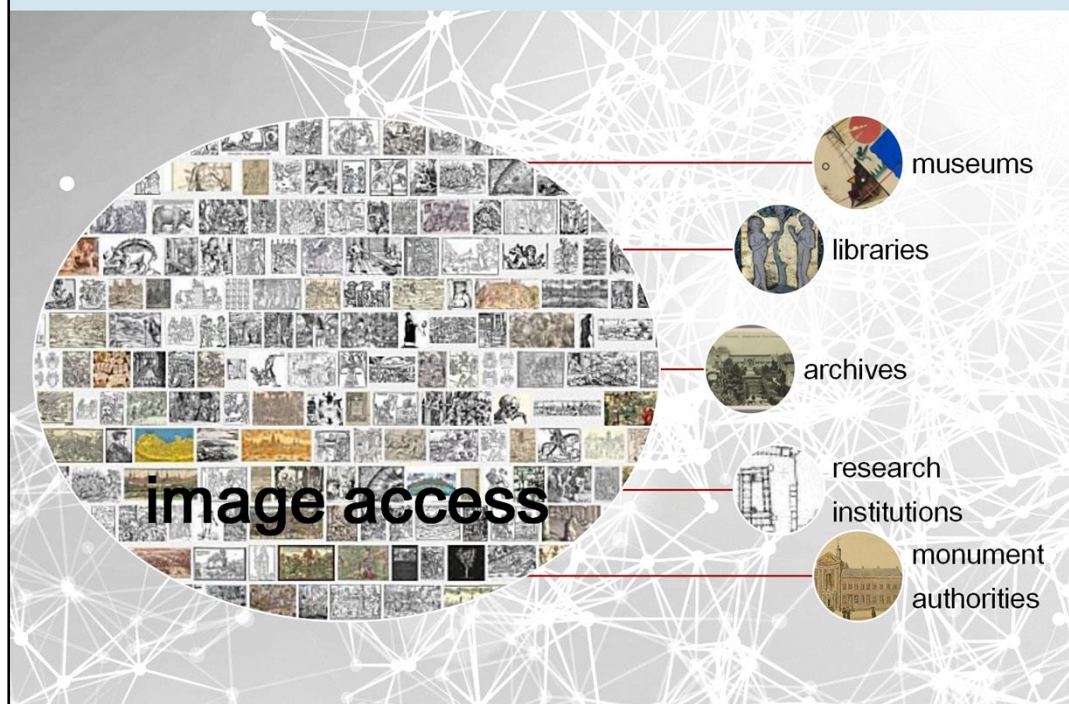
With adequate and specific vocabulary on the subjects of western art, Iconclass offers a good scientific coverage of its field. The hierarchical classification ensures a comprehensive contextualisation of each definition.

There are other features recommending Iconclass as a sustainable and future-oriented authority, suitable for semantic web applications:

- The system is available in English, German, French, Italian and Finnish. More translations are under way.
- It is published freely accessible on the web as the Iconclass Browser, also under an open license as Linked Open Data in RDF and JSON.
- Maintenance and development lies with a reliable institution, the Netherlands Institute for Art History (RKD).

Among those visual collections which invest their precious resources into a structured subject cataloguing, Iconclass is increasingly in use for all kinds of image material.

Future challenge: Cross-sectional access to images as primary research data



Images are valuable primary research data. For art history and visual culture studies, images are representations of a real-world object or primary objects of research. For the other disciplines of historical research, increasingly, they are sources and pieces of evidence. Qualified subject cataloguing opens up access to essentially non-textual materials.

It is encouraging that with RDA the library sector aims at establishing a comprehensive cataloguing practice for all kinds of collection materials in heritage institutions. Even if the other sectors will perhaps not embrace RDA readily, over time a variety of library materials will be catalogued in a much more interoperable way.

This provides the possibility of opening access to a wealth of image materials in libraries which are presently not accessible to users. For cross-sector discovery systems or semantic web applications, the underlying entity-relationship data models presented in the FRBR family can be a good basis for building semantic crosswalks to other standards and formats, to enhance the discoverability of image-related metadata.

For museums and image archives, the Conceptual Reference Model (CRM) has received recognition as a data modelling standard. So, with FRBRoo, a model which harmonizes FRBR and the CRM already exists, and this will help significantly in aligning data from the library and the museum communities.

Museums and photo collections have to meet a challenge as well: they have to adjust their documentation strategies to new web-based user communities. This asks for a much more comprehensive standardization of cataloguing than we see at present, to enable technically new ways for qualified cross-sectional discoverability.



Thank you for your attention!

Angela Kailus

kailus@fotomarburg.de

Deutsches Dokumentationszentrum für Kunstgeschichte
Bildarchiv Foto Marburg – Philipps-Universität Marburg



Illustration: Bruegel, Jan, d. Ä.: Allegorie des Gesichtssinns, Madrid, Museo del Prado, um 1617-18,
http://prometheus.uni-koeln.de/pandora/image/show/halle_kg-dcc4251aa9cb8313010ee13fc83bcc826e1d298f