



Seeing Images: Metadata and Mediation in the Digital Archive

By

Sonya Petersson & Anna Näslund Dahlgren

Abstract

In the cultural heritage digital archive, descriptive metadata makes images (re)searchable. Text-based searches seek terms that match metadata terms or terms referring to aspects of images that have previously been considered essential to select and describe in metadata terms. Such considerations are bound up with historically changing institutional agendas, ideas about user preferences, and implementation of metadata standards. This study approaches image accessibility from a different perspective. It aims to investigate how the infrastructure of the digital archive, comprising metadata and interface, intervenes with, circumscribes as well as enables, the images' visibility and knowledge-producing capacity. The starting points are: first, that images in digital archives, exemplified by the online image collections in Alvin and DigitaltMuseum, are mediated, mediating, and "mixed" media objects that simultaneously represent the past and the present; second, that the digital archive in a media history of images functions as both a tool and an object of research. Using the platforms as tools of research, this study is based on test searches that aim to find viable search strategies for mixed media objects. The chosen search terms represent media-historically significant and common traits such as images that are combined with text and images that represent and/or mediate other images. The study discloses that the platforms give both false negatives and false positives. They do not support searches that focus media terms and relations between media elements. These problems are further related both to heterogenous metadata practices and to the simultaneously restricted and broad image concept behind them. As objects of research, both platforms are considered in relation to a future construction of a media history of images, where the digital archive is a particular node. The study demonstrates how the "hypermedial" environment associated with new media is prefigured by

Keywords: Image, Mediation, Metadata, Media history, Online image collections

Petersson, Sonya & Näslund Dahlgren, Anna: "Seeing Images: Metadata and Mediation in the Digital Archive", *Culture Unbound*, Volume 13, issue 2, 2021: 104-132. Published by Linköping University Electronic Press: <http://www.cultureunbound.ep.liu.se>



media interrelations in analog images – or images that are accessible as mediated through the archive’s interface and as policed by the archive’s metadata structure.

Introduction

In the open access cultural heritage platform DigitaltMuseum, the image of the late nineteenth-century magic lantern slide (Fig. 1) is mediated by a computer screen. The digital representation of the slide is slightly cropped. It appears against a grey background with its metadata record below, including media classifications such as “Type: Photograph.” Such descriptive metadata determines the image’s retrieval.¹ Search terms need to match metadata terms if an image is to be found, with the consequence that not everything that could be attributed to the image is searchable. Descriptive metadata also make claims for what the image shows and what it is about. According to the record, the slide depicts “painting tools and a picture” rather than, for instance, a “still life” or an “image within an image.” In other words, descriptive metadata is selective; it refers to some but not other properties or aspects of the image. This article focuses such dead spots in metadata for images, and the implications these have for precision and recall in searches in image collections.

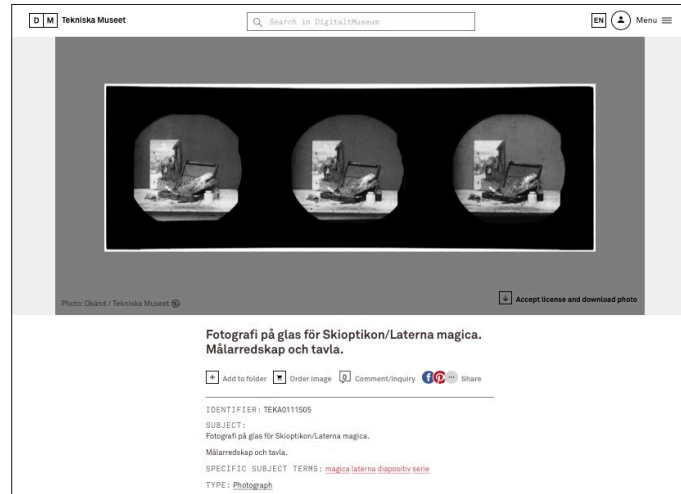


Figure 1. Screen shot of the magic lantern slide with record (cropped) below, late nineteenth-century. Photography on glass with (cropped) wooden frame, 7.8 × 23 cm. Accessible through DigitaltMuseum: <https://digitaltmuseum.se/021016341802/fotografi-pa-glas-for-skiptikon-laterna-magica-malar-redskap-och-tavla>. Identifier: TEKA011505. License: CC PDM. The record’s reference to The Swedish National Museum for Science and Technology (Tekniska museet) as owner is wrong. The object was a deposit from the Swedish Film Institute (Svenska filminstitutet) that is now returned.

That metadata practices vary across different standards, individual institutions, and even between individual information specialists is well established in previous research (Gilliland 2016; Waldron et al. 2018; Seeman & Dean 2019). The variety in interest and search practices among external users of image collections are likewise extensively researched. However, with a focus on accessibility, these studies have primarily been concerned with different vocabularies and search methods and what has been labelled “the semantic gap” between different parties in the field (Eakins 2002; Hollink et al. 2004; Beaudoin 2007). They typically do not investigate images *as images*, in the sense of mediated, mediating, and “mixed” media objects that are simultaneously showing and telling, representing the past and the present.

This study aims to investigate how images in the cultural heritage digital archive are mediated by digital interfaces, represented by descriptive metadata, and made (re)searchable online, in order to analyze how the digital infrastructure intervenes with – circumscribes as well as enables – the images’ knowledge-producing capacity, indeed their very visibility. Our concern is with how the digital archive can serve both as a tool for researching the media history of images and as an object of study in that history (cf. Benardou et al. 2018: 1–14). What is at stake is the digital archive’s double function of provider/gatekeeper *and* mediating apparatus: how its metadata determines the image’s retrieval and how its interface mediates the image once retrieved.

DigitaltMuseum and the similar platform Alvin are our examples of online platforms for cultural heritage material. Both assemble image collections and metadata from Swedish libraries, archives, and museums (LAM). Alvin and DigitaltMuseum were chosen both for their rich image collections and as representative of the larger phenomenon of cultural heritage collections migrating online. DigitaltMuseum, launched in 2009 and curated by KulturIT with support from the Norwegian Arts Council (Kulturrådet), can be compared to Europeana, the European Union’s online platform for LAM institutions across Europe. Alvin, launched in 2014 and curated by Uppsala University Library, can be compared to Gallica, the digital archive of the French National Library (Bibliothèque nationale de France).²

In order to test image searchability on these platforms, we performed a number of different search operations. We started with a search on a traditional iconographic subject. From the result of this search, we selected two primary examples that highlight media-historically significant and common aspects of image mediality. We then translated into search terms aspects of image mediality that put the existent search function to the test. Rather than using default terms such as producer’s name, we searched for images with textual elements and images that represent other images. To do so, we did seven different single-word and

Boolean searches through the search function on the platforms.³ We tested the system with different combinations of media words, i.e., the word “image” (“bild” in Swedish) and text (with some alternates) in different search fields, with the aim of trying out the platforms sensitivity to media relations (see further Table 1). The searches have been paired with studies of individual images and a scrutiny of the online interface where they appear. Our two primary examples generated by the first iconographic search have further been paired with two supportive examples, known through previous research (Dahlgren 2013). In sum, this study combines attentiveness to visual detail common to art historical studies with a simultaneous consideration of the mediating platform (cf. Dahlgren 2018).

Our study is based on recent developments in image studies and media studies. These fields have opened up new avenues for investigating and understanding the mediality of images. One frequently cited text on the topic is W. J. T. Mitchell's 2005 article “There Are No Visual Media” (reprinted in Mitchell 2015; see also Belting 2005; Elkins 2008; Moxey 2013), but there are, especially from the fields of intermedial and multimodal studies, many more elaborated discussions (Bateman et al. 2017; Elleström 2019; Wildfeuer et al. 2019), which informs our study. The field of word and image studies has likewise furthered the understanding of the image's mediality by extensively researching different types of relations between image and text in various historical and theoretical contexts (Heffernan 2006; Bateman 2014; Williams et al. 2019). These studies are important because they have in different ways dug into the image's mediated and “mixed media” character. Within one and the same image there are not only different layers of texts (such as inscriptions and iconographic narratives) but also different layers of mediation.

On the basis of these research fields we use an image concept that ties together mediation, representation, and media history. The relationship between the analog image and its digital representation is a mediating one. Mediation includes transfer of some, but not all, properties of the mediated object, as well as transformation of the properties transferred (Elleström 2014), with the effect that mediation will always, to some extent, affect how the image is represented. The important point is that when the image is mediated from, for instance, painting to printed photograph to digital file for display on screen, it is at each stage *both* transferred and transformed in itself, *and* transformed by its insertion into a new media environment. In the new environment of the digital archive, it takes on new relations to surrounding parts, including metadata and search terms, the concepts they recall, and the neighboring images they recollect. In the cultural heritage context, the digital archive is also a place where “old” and “new” media converge (cf. Jenkins 2006; Ernst 2013). Media convergence takes place in the present; it is posited on the fundamental difference between the analog and the digital image (cf. Dellman 2019). The digital image can only represent a past media culture from

its present manifestation. It is exactly this mediated state, bound to the (re)search possibilities and pitfalls of the archive, that is now open to study.

Metadata Conventions in Practice

Metadata practices differ between libraries and museums. The library sector has typically a more elaborate and consistent catalog of metadata standards and best practices (Gilliland 2016: 4). Even though they belong to different parts of the cultural heritage sector, what Alvin and DigitaltMuseum have in common is that their metadata practices are underpinned by institutional ideas about what an image is. Such ideas are indicated both by instructions for image indexing and by descriptive metadata itself.

In Alvin, pictures are recorded and indexed with reference to standards developed by the Library of Congress and the Swedish National Library (Kungliga biblioteket) (Andersson & Andersson 2020: web). These standards adapt the description of pictures according to a bibliographic tradition which privileges not only producer names (authors) but also titles and origins (MARC field 245). Titles differ from producer names in that, in image indexing guidelines, they often serve the function of anchoring what the image depicts.⁴ The principle put forth by the Swedish National Library privileges titles or captions written or printed on the picture, rather than titles known through other sources. Where there is no known title, one is created (and marked as such) based on the image's "motif (motiv)" (Kungliga biblioteket 2013: web). Further, in the library's instructions for additional notes (MARC field 520), the field is recommended when neither "subject words or titles are sufficient for describing the motif of the resource." The field can be used both for describing what the image resource "depicts," what it is an image "of," and what it is "about" (cf. Beaudoin 2007: 26). It is also recommended that the cataloguer considers the relation between any textual "content" in the picture, the picture's original function (e.g., as a political poster), and its "subject" (Kungliga biblioteket 2013: web).⁵

The same distinction between depiction and content recurs in the Library of Congress's principles for image indexing:

By their very nature, most pictures are "of" something; that is, they depict an identifiable person, place, or thing. [...] In addition, pictorial works are sometimes "about" something; that is, there is an underlying intent or theme expressed in addition to the concrete elements depicted. [...] Indexers should examine images, their captions, and accompanying documentation carefully to determine both the most salient concrete aspects (what the picture is "of") and any apparent themes or authorial

intents (what the picture is “about”), taking care not to read into the images any subjective aspects which are open to interpretation by the viewer. (Library of Congress 1995: web)

The quote above thus ties “of” to what would be, given broad and intersubjective frames of (Western) references, easily recognizable (i.e., resembling) depictions of already known entities (whether external, fictive, mythological, etc.). The “about” part is notably restricted by “apparent themes or authorial intents,” whereas “apparent” points to the indexer’s frames of references. “Authorial intents” more specifically points to (knowledge about) the object’s context of production, as distinct from the context of reception (whether understood as including the picture’s meeting with its first public or its later institutional care).

The image concept embedded in the bibliographic tradition is thus both restricted and broad. First, it is based on depiction, and therefore dependent on the idea of resemblance but more limited than its cognate concepts mimesis and iconicity. These concepts all express the idea of resemblance that is part and parcel of Western thinking about images (Mitchell 1986: 7–46; Jappy 2013). But whereas mimesis and iconicity include all kinds of similarity, depiction is here limited to more or less easily recognizable (figurative) elements, “the [picture’s] most salient concrete aspects.” Further, the “aboutness” of images is delimited by being tied to inscriptions, titles, known intentions, and themes related to the context of production. At the same time, judging from the online collections in Alvin, this restricted image concept includes a broad range of pictorial genres (“object types”), from maps to postcards to scientific illustrations to artist’s drawings to photographic negatives (Table 3).

Unlike Alvin, the records in DigitaltMuseum reveal no uniform cataloguing standard. Every participating institution (mostly museums), is responsible for selecting the objects and related information they submit to DigitaltMuseum, resulting in a platform that represents a wide range of old and new cataloguing traditions and institutional interests (cf. DigitaltMuseum no year: web). For example, the Swedish National Museum for Science and Technology (Tekniska museet) has not indexed its images in any iconographic detail, while the National Museum of the Fine Arts (Nationalmuseum) has. Notably, many of the museums represent the pictures in their collections as part of a period’s or region’s material culture at large, in metadata terms that consider the “thingness” and cultural use of the picture rather than its image content. A painting indexed with terms that denote what it depicts could very well also be categorized as “home furnishing” and “wall decoration.” In effect, the restricted image concept actualized by the bibliographic tradition is, in DigitaltMuseum, often paired with ideas of images as material objects, which has some problematic implications. For one, the

bibliographic tradition focuses a single named originator. This focus privileges images where such information is typically noted (e.g., art, press photography) over others, commonly not attributed to one person and often cooperatively produced (e.g., advertising, information images, book illustrations, popular visual culture at large). The focus on images as material objects, on the other hand, is limited in the sense that it typically disregards the images as medial objects (medium, image technique) but rather bases their objecthood on associated cultural uses.

The amount of metadata in DigitaltMuseum varies significantly, from only a title, a few subject terms, and mandatory information on institutional location and intellectual property rights, to extensive histories of production and physical descriptions (cf. Loukissas 2019: 62–63 about the same type of heterogeneous metadata structure in the Digital Public Library of America (DPLA)). DigitaltMuseum also invites users to add free tags to images, much like FlickrCommons, but this is not yet common practice. The descriptive metadata structure that is visible in the online interface includes ten different fields in DigitaltMuseum and nine in Alvin (Table 2). As an effect of the bibliographic context, Alvin has more fields related to authorship (persons, origin, subject/persons) than DigitaltMuseum. Conversely, DigitaltMuseum has more metadata fields related to the objecthood of the resource (designation, type, dimensions value, materials, techniques) than Alvin. A similar pattern can be seen in the search fields in Alvin and DigitaltMuseum respectively (Table 3).

The metadata practices of Alvin and DigitalMuseum and the image concept that underpins them both enable and circumscribe the image's retrievability. Text-based searches on image properties (type or form/genre terms) and representational aspects (subject terms and titles) include an act of translation between image and text. Text-based searches build on first having an idea of the property or aspect sought, and then finding suitable search terms. Suitable search terms are those that fit both the image and its metadata, which means that they must rely on shared ideas of matches between visual properties and denotation of terms (cf. Pringle 2010: 21–24). Besides the risk of mismatches inherent in all such intersubjective guessing, it is well-known that text-based searches on image content present an iconographic problem that increases with the abstraction of the search term. From the cataloguer's point of view, Julia Thomas, the researcher responsible for the Database of Mid-Victorian Illustration (DMVI), notes that terms such as "time" or "death," in metadata parlance denoting the abstract conceptual level (Holink et al. 2004), are difficult to employ in a uniform way across a pictorial corpus, since they may not only be differently visualized, but also range from the explicit to the implicit (Thomas 2011: web; cf. Lager Vestberg 2013: 476, 478).

Consequently, it is easily assumed that the more conventionalized and established the iconography attached to a particular abstract term, the more

successful it will be as a search term. We therefore commenced our test of the platforms with an iconographic search. It comes as no surprise then that our search on “Crucifixion,” representing one of the most well-known configurations in Western painting, yielded 92 and 19 unique hits (a lesser number of duplicates deducted) in DigitaltMuseum and Alvin, respectively (Table 1: Search A).⁶ The result is not impressive in quantity, but it is precise; the hits – at first sight – show the expected standard iconography of the biblical story of Christ on the cross.

Of greater significance than sheer quantity is the fact that this search exemplifies the benefits of transgressing the ordering units prescribed by the metadata categories that tie images to producers, types, and the regional/temporal borders pertaining to the context of production. In our group of search hits, the Crucifixion scene can be studied as migrating (Mitchell 2015: 65–77) from reliefs on late Middle-Age vessels, to seventeenth-century tapestry, to eighteenth-century snuffboxes, nineteenth-century school posters, and as tattooed on the chest of the “Sailor Johan,” as well as on broadsides, lithographs, reproductions of paintings, and altarpieces. The most important result of this search on Crucifixion images is that all hits bear witness to the diversity contained in one and the same iconographic subject. Among the search results, the scene appears as embroidered on cloth and engraved in metal and as richly varied in iconographic details (cf. the soldiers’ dice, the dragon, the hourglass, and the skull in Fig. 2 as compared to the single cross in Fig. 6). The scene can likewise be studied as migrating between popular and elite image cultures, and between the (digitized) material traces of media spectacles, such as magic lantern slides, and vernacular cultural practices, such as broadsides pasted inside trousseau chests.

In the following sections, two of the individual images that surfaced through our iconographic Crucifixion search will serve as the foundation for testing the platforms’ limits of image retrievability, and further be compared to parallel image examples in the same platforms. These images are far from unique – neither historically or in the archives. But to what extent and in which respects are they searchable online? As our next sections will demonstrate, that is a completely different question.

Image and Text

Our first example is a Crucifixion painting (Fig. 2) that might be of art-historical interest both for its iconographic variations of the scene and its deviance from academism. What is striking in the image is, however, something unindexed in its record: the scene is enclosed on three sides by a frame of painted inscriptions retelling the story of the Crucifixion. The frame is not part of the scene as is the inscription “I.N.R.I.” on the cross. While the latter is located inside the pictorial world, the frame lies outside it and surrounds it.



Figure 2. Framed Crucifixion scene, eighteenth-century. Oil on canvas, 88 × 85 cm. Accessible through DigitaltMuseum: <https://digitaltmuseum.se/011023537199/ol-jemalning>. Identifier: NM.0095636. License: CC BY-NC-ND.

This image is pivotal because it in different ways exemplifies image and text relations that are both inherent in the digital archive and in the analog image cultures of the past. Apart from the traditional way of referring to iconographic subjects as “textual content” behind images, the painting shows text in a way that is directly based on the broadside, a genre of prints where image and text combinations of the same kind are paradigmatical (cf. Algee-Hewitt et al. 2018: 5). These types of sheets still abound online: a simple image googling of “broadside and crucifixion” gives many examples illustrating how the broadside has found its way into digital culture. The first point here is that the painting represents image and text combinations that were common in the print era (parenthetically it can also be noted that similar paintings, which emulate prints, are not historically exceptional even if for historiographical reasons rarely researched). The second point is that when the painting is displayed against the “second frame” of DigitaltMuseum’s interface, with its standard grey background and metadata record below (Fig. 3), the image and text combinations of the screen, characteristic of the digital age (cf. Bateman 2014: 14), echo the image and text combinations of the eighteenth-century painting representing prints.

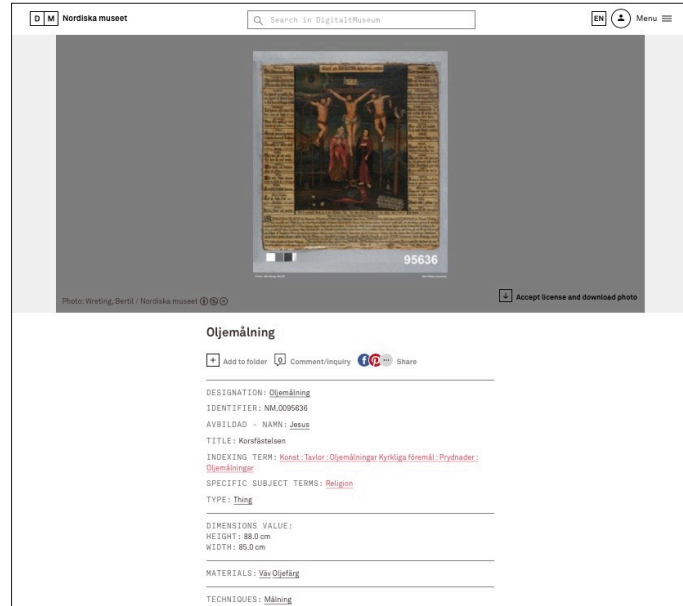


Figure 3. Screenshot of the framed Crucifixion scene in DigitaltMuseum: <https://digitaltmuseum.se/011023537199/oljemalning>.

A parallel example of how image and text combinations recur in different media landscapes and through different media technologies is the nineteenth-century photographic montage of actor portraits and text clippings (Figs. 4–5), originally distributed as a souvenir of the theater event “Min ros i skogen!” The text on the paper clippings carries information about this event. Such information, along with the printed characters’ and actors’ names at the bottom of the frames, contextualizes the portraits as actor portraits. And as accessible through DigitaltMuseum, the montage is also “framed” by the grey background and the text of the metadata record.



Figure 4. Photographic montage of portraits and paper clippings, 1867. Photograph in carte-de-visite format. Accessible through DigitaltMuseum: <https://digitaltmuseum.se/011013855502/visitkortsbild-fran-teateruppsattningen-min-ros-i-skogen-pa-dramatiska>. Identifier: NMA.0063463. License: CC PDM.

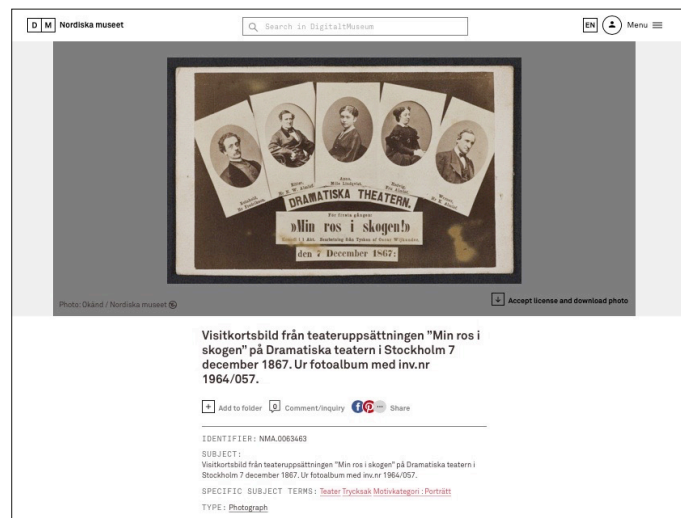


Figure 5. Screen shot of the photographic montage in DigitaltMuseum: <https://digitaltmuseum.se/011013855502/visitkortsbild-fran-teateruppsattningen-min-ros-i-skogen-pa-dramatiska>.

In the digital space of the archive, these two examples of past images with textual elements, are connected to metadata texts in two ways: technically, by being accessible through metadata, and interconnectively, by being expected to correspond to, being informed by or in other ways standing in relation to the metadata record exposed on screen. Three types of interconnections between metadata and image elements are notable:

First, metadata can be understood as referring to the image of the resource in the traditional sense of providing information about it, as discussed in relation to titles in the previous section and as noted in relation to the content of the text clippings that made the portraits identifiable as actor portraits. When the record pertaining to the montage repeats the text in the clippings as the title of the resource and then again in the metadata entry “Subject,” it effectively uses textual elements to provide information about the image (as representing the theater event that took place at a given point in history). At the same time, metadata is also, like the images of the resources, a communicational media technology that could be situated as a trace of, in these cases, old institutional data transferred online (often without editing). Studied as such, both images and metadata would be employed as examples of communicational, referential, and representational media technologies from different historical and conceptual contexts, which are brought together by the infrastructure and interface of the digital archive, like Jussi Parikka’s idea of old and new media in “parallel lines” (2012).

Finally, comes the distinction between vehicle and content – which is important for pedagogical reasons, but should not be confused with understanding image and text merely as vehicles. John Bateman and fellow multimodality scholars address image and text as “semiotic modes” that are materialized in socially and historically situated media of some kind (Bateman et al. 2017: 124). The “semiotic modes” have a vehicle side in their materialization in a particular media channel and a content side in the meaning evoked by visual configurations, letters, and so on. It should be emphasized that the content is not confined to, for instance, information about the theater event in the clippings; the “vehicles” of image and text could very well also be understood as content, as that which the framed Crucifixion painting and the montage is about. In the first traditional sense, this would be the case if they had subject terms stating “images and text” in their records. It would also be the case in a study inquiring how image and text combinations of the past take part in constituting, with Bolter and Grusin’s (1999) well-used phrase, the “hypermedial” environment of the digital archive. The vehicles of letters and figures would then be treated as expressing a content concerning their medial state as image and text materialized in a particular media channel. Thus, studies of image and text as media elements in the digital archive can provide the impetus to understand the combined communication technologies of the past and the way they are interconnected to and take part in historicizing the combined communication technologies of the present.

The descriptive metadata of the framed painting include “Religion” (as specific subject term), “Crucifixion” (as title), “Oil painting” (as indexing term), and “Artefact” (as type). Besides citing the clippings, the montage’s record adds “Theater,” “Print matter,” and “Category of motif: Portrait” as (rather unspecific)

specific subject terms. None of this is disputable. But it does not provide searchable type, form/genre, or subject terms that captures the fact that both the painting and the montage communicate through image and text. They cannot be found through the search function as examples of images with textual elements.

These considerations raise the question of how media terms function as search tools and how to find images that are *combined* with text. Here, we build on the concept of combined media as defined by intermedia scholar Irina O. Rajewsky. Both image and text, with their own sign vehicles, should be materially present in one media object, and both conventionally expected to contribute to the significance of that object (Rajewsky 2005: 51–52).

Generally, media terms fail as sharp search tools because they occur too often in different sorts of metadata. The Boolean search “image and text” in the general search field renders numerous hits (688 and 88 946) in Alvin and DigitaltMuseum (Table 1: Search B), but is too inclusive. The search precision is weak and results in many false positives. It captures resources that display, for instance, texts with metaphorical titles such as “An Image of xxx.”

To narrow it down, we used the media field, but it did not yield a precise search result.⁷ In DigitaltMuseum, a general search on “inscription” combined with “image” in the media field yields 2 277 hits, whereas “text” combined with “image” yields 201 635 hits (Table 1: Search C and D). Even if numerous, the hits are still mostly of low relevance since they contain too many examples like (a digital photograph of) a vintage candy box accompanied by the metadata description “with white and golden text.” With the concept of combined media in mind and with the Crucifixion painting and the montage as points of reference, the type of object sought for should be both pictorial and textual. A drawing of the vintage candy box would have been a relevant hit, unlike the hit where the candy box is represented as a thing.

When we tested the search terms “inscription” and “text” in the topic field in combination with the term “image” in the media field, manually more manageable numbers of 722 and 729 hits are generated, respectively (Table 1: Search E and F). In the “inscription” case, the majority are photographs of objects other than pictures, such as tombstones with inscriptions (comparable to the candy box). Only a minor number of relevant examples show up (i.e. relevant according to the concept of combined media), mostly consisting of portraits with inscriptions and stamps. In the “text” case, about 200 relevant and unique hits occur, including menus, posters, letters with drawings, paintings with inscriptions, maps, postcards, instructions for needlework, and illustrated pages.⁸ In each case, however, relevant examples that are known to be retrievable within DigitaltMuseum are excluded, such as the Crucifixion painting and the montage, and none of the relevant hits are indexed as either being about image *and* text (it is not possible to conduct a search

with both terms in the topic field) or as instantiating image *and* text as form/genre traits. The same tendency is evident in Alvin but in minor scale (Table 1: Searches C–F). “Image and text” are not consistently employed as subject or form/genre terms. They are common enough in the metadata to give some relevant results, but important cases are left out.

Images in Images

The searches above, aiming at finding image and text combinations, throw light on a related issue of image filtering. On a practical level, the media fields in Alvin and DigitaltMuseum have different functions. In DigitaltMuseum, “Has media/image” means that the resource contains an image (as distinct from containing audiovisual media) while “Resource type/image” in Alvin means that the resource is an image. This difference should not be overstated. Both fields include images that are themselves objects of the resource and images (photographs) that reproduce the object of the resource. Therefore, it is not surprising that our image and text searches include hits such as tombstones with inscriptions. However, on a level concerned with the mediality of images, an interesting problem arises. A (digital) photograph of a tombstone or a candy box is certainly both an image and a picture. It is equally obvious that the tombstone or candy box, the object of the resource, is not a picture. In such cases, the media field does not target the object of the resource, but its “first” mediation. In relation to another one of our Crucifixion hits, the frontispiece in Figures 6–7, this inclusiveness raises the question of how to handle search queries that concern media *in* media, or images in other images.

The frontispiece (1710) by the Swedish engraver and painter Anna Maria Thelott (1683–1710) depicts an artist in the act of painting. On the easel in front of her, an image of the Crucifixion is about to take shape. The frontispiece is signed as no reproduction; both “delineavit” and “sculpsit” are connected to Thelott’s name, which implies that she is the producer of both the engraving and its design.

Like the magic lantern slide (Fig. 1), the frontispiece represents – “is about” – painting as a medium. In both cases, the images *in* the images (i.e., the rural scene on the painting in the slide and the Crucifixion scene on the painting in the frontispiece), are clearly represented as paintings in the sense of pictures, as physical objects separated from their environments by the extension of the canvases. Another set of elements evokes the idea of painting as the result of manual work: the artist by the easel in the frontispiece and the palette, palette knife, and paint tubes. Lastly, the “double exposure” of the slide and the frontispiece highlights the representational function of painting and, by extension, of any other image. The paintings in the slide and the frontispiece *represent* a rural scene

and the Crucifixion. As accessed through the digital archive, there is yet another mediating space that represents the slide and the frontispiece (along with their representations of interior paintings): the interface of the platform. Thereby, the interior paintings, the slide and the frontispiece, and the interface of the archive, mimic each other in a chain of representation.



Figure 6. Anna Maria Thelott, frontispiece to Johann Burchward's *Oratio passionis de crudelissimis Jesu Christi* [...], 1710. Engraving, 20 × 16 cm. Accessible through Alvin: <https://www.alvin-portal.org/alvin/view.jsf?pid=alvin-record:94836>. Identifier: urn:nbn:se:alvin:portal:record-94836. License: CC PDM.

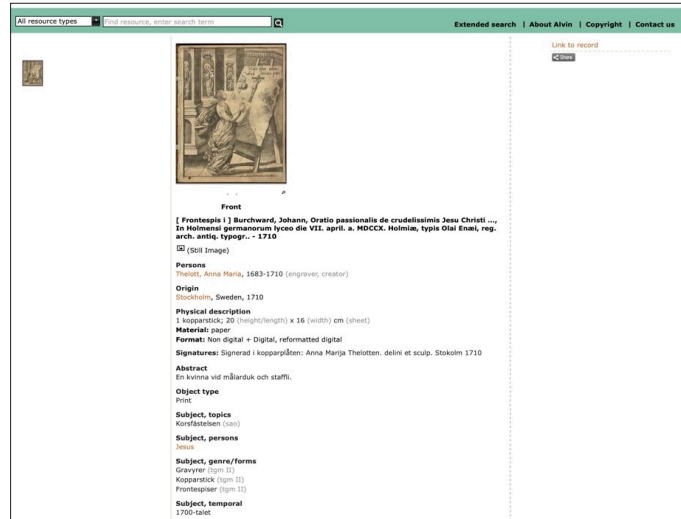


Figure 7. Screen shot of the frontispiece in Alvin: <https://www.alvin-portal.org/alvin/view.jsf?pid=alvin-record:94836>.

Images that represent other images are examples of what intermedia scholar Werner Wolf calls “metareference” (2009), which refers to the cognitive act of becoming aware of media objects as mediating and representational through props such as the image in image chains described above, or props in media objects that refers to media functions. For us, there are two key aspects. The slide and the frontispiece, like the montage, are *past* examples of metareferential images that in the digital environment are further highlighted as such, as mediated once again, and contributing to the specific hypermediacy of the digital interface. This is of some importance because Wolf (2011) and others have made a point of tying metareference to postmodernity. Our point is not to find predecessors – well-known studies of early modern painting by Svetlana Alpers (1983) and Victor I. Stoichita (1997) would make that undertaking unnecessary. Rather, our point for the future is to better understand metareferential images from the double perspective of asking both how present media technologies and media concepts may inform the study of past images and how past metareferential aspects of images are not simply an effect of the archive, but specific traits of past images that can historicize similar phenomena in the present. This is also why we highlighted the difference between digital photographs of tombstones and digital photographs of other images, whether photographs, paintings, drawings, and so on. Basically, *any* photograph displayed via a screen could be argued to partake in the hypermediacy of new media, exactly as an effect of the new media interface. Contrarily, we advocate a perspective that emphasizes the fact that the phenomenon of metareference is already a distinct trait of past images (as

exemplified by the slide, the montage, and the frontispiece), a trait that has the potential to throw new light on the digital media that mediates it.

In order to find images that illuminate past and present metareference by representing other images, the digital archive could function as a great tool of research. One of the keys seems to be to combine “image” as a search term with fields pertaining to subject terms and form/genre terms. In order to throw light on the categories in the search fields and to better understand the mediation of images by the digital archive, a pragmatic distinction between mediation as a purely technological phenomenon and as within the sphere of representation (cf. Elleström 2014: 14, 27–34) seems to be warranted. This can be exemplified by comparing the slide and the frontispiece.

The frontispiece and the slide are alike in that they represent painting as a medium by depicting paintings. But they relate to the depicted paintings in different ways. The slide mediates the brushstrokes and colors of the canvas through the technology of analog, monochrome, photography. The analog photograph is then again technologically transferred to a digital file for future display on screen. These transfers are indexical in the strong sense of implying causal relations between the painting and the analog photograph, and between the analog photograph and its digital counterpart (i.e., the relations are independent of any perceiver’s recognition).

The frontispiece could also be said to expose painting in another materiality and by other sign vehicles: paper instead of canvas, and lines, strokes, and cross-hatchings instead of brushstrokes and paint. But the difference is that the frontispiece relates to painting only by evoking the idea of painting *as represented* within the engraving. There is no technical mediation involved (such as between the frontispiece and its digital representation). The frontispiece is connected to painting as a predecessor by inference. Its depictive qualities form the basis of attributing its representation of painting to a preceding sphere of, for example, artist’s knowledge, cultural contexts, and so on.

Technological mediation is implicitly acknowledged in any record categorizing, for instance, a tombstone as a photograph. Media representation is not systematically acknowledged as a category of its own, or singled out from any other type of representation (tombstones, landscapes, iconographic subjects, etc.). In the frontispiece’s Alvin record, nothing, except the general description “woman by canvas and easel,” reveals that the resource actually contains two representational spaces, where the first image represents another image. The general description may help us find similar images of women painters with second-level images – Thelott has made at least three that are accessible through Alvin – but the artist’s name is not useful for conducting a general search on images in images.

The larger question is which of all the images involved in the examples above enters the metadata record and in what sense, as included in the categories of subject or form/genre. In DigitaltMuseum, we searched on the term “image” in the topic field, combined with “image” in the media field. The search yielded 23 046 hits of overall low relevance, according to the criteria that they do not generally uncover depictions of pictures (Table 1: Search G). The overwhelming majority of the hits are photographs emanating from one institution (Swedish Military Heritage/Sveriges militärhistoriska arv). The search result is therefore skewed by the practice of one particularly comprehensive institution to index its photographs of artifacts in the museum collection with “image” as a subject term. Subject terms generally come in groups such as “Image, Building, Apartment Building, Photograph” (for a 1960s photograph of a small-town street view), where the terms variously denote media properties (first and fourth) and that which is depicted. Among thousands of similar examples, a few relevant hits occur haphazardly. A corresponding search in Alvin on “image” in the subject field renders only 27 random hits (Table 1: Search G), containing examples with “damaged image,” “transparent image,” and “xxx is to the left in the image” as subject terms. In both DigitaltMuseum and Alvin, images such as Thelott’s frontispiece, known to be included within the platforms, are not found among the results. Thereby, the metadata practices behind our attempts does not allow structured searches that presuppose a distinction between first- and second-order mediation and/or representation. Put differently, there is no straightforward way of searching images (photographs, paintings, drawings, etc.) that mediate and/or represent other images. It should also be added that the “type” field in DigitaltMuseum is in theory near equivalent to the “resource type” field in Alvin (Table 3). For our purposes, however, it has not been a viable option since the preselected list of image alternatives is, at once, both too narrow (“Fine art”) and too broad (“Photographs”).

Part of the problem of searching on image resources and expecting second-order images to be indexed as such is that the terms employed for indexing pictorial subjects and objects are both heterogeneous and distributed across many of the metadata fields (cf. Table 4). Words such as “photograph,” “painting,” “hanging,” “drawing,” “landscape,” and “flowers,” denoting both medial properties and representational aspects, all occur as subject terms, while many of them also occur as type or form/genre terms, apart from showing up in titles and general designations. DigitaltMuseum, especially, has a seemingly coherent metadata structure standard, but much less coherence in the vocabularies and data content standards applied. Our search on “image” in the topic field, which we expected to return pictures as the content of a resource, generated hits targeting medial properties of the resource. At the same time, we would not deny that groups

of subject terms such as “Picture, Building, Apartment Building, Photograph” may very well, depending on the search and research objectives, be a pertinent articulation of the *content* of a 1960s photograph of a small-town street view. But they do not support searches where it may be of interest to separate form/genre from subject terms.

Consequently, the media insensitivity of the searching systems prevents the type of result we got from the Crucifixion search. The image of the Crucifixion was found migrating between different pictorial media, periods, and regions. Yet, equivalent migrations could not be found through searches on “image and text” or “image in image” with the same precision. These types of media phenomena, both historically abundant and the target of much humanities research, even beyond obvious fields such as word and image studies and intermedial studies, are therefore to some extent hidden within the systems. That this occurs in a library platform like Alvin is particularly remarkable as the vast majority of the images in such a collection are images combined with text. Every illustrated book, magazine, and daily paper in the library collection is a combined media product.

As we have throughout this article presented our examples as “common” and “historically abundant,” it is important to remember the difference between the past media landscape and the analog/digital archive. In relation to the number of negative/positive hits represented in Table 1, ranging from lesser than ten to more than 200 000, it is also worth emphasizing that the hits exclusively represent metadata practices – not quantities in the past media landscape or in the analog/digital archive. In the past media landscape, the Crucifixion scene and the frontispiece, as well as the parallel examples of the slide and the montage, are nonunique, standard, images: the Crucifixion scene as representing the image and text combinations of the ubiquitous broadside; the montage as representing image and text combinations recurring not only in similar photographic cards but also in posters and the illustrated press; and the frontispiece, along with the montage and the slide, as representing a “mise en abyme” of images that feature across media history, from illuminated medieval manuscripts to the present interface of image collections online. As for the holdings of both Alvin and DigitaltMuseum, they clearly are more extensive than indicated by their metadata. This is supported by the fact that examples of images with combined textual and pictorial elements that are known to be included in the digital collections, are not searchable through terms denoting textual and pictorial properties. The latter makes our lesser numbers unreliable, whereas our greater numbers include a significant number of negatives, which first and foremost result from an excessive use of media terms in the metadata (see next section).

The perspective of images as mediated, always entangled in media relations, also challenges the existing writings on the production and use of metadata

for images. It nuances the understanding of what is typically thought of as the “perceptual level” of images as being confined to color, texture, and shape (James & Chang 2000; Choi 2017). On the contrary, we have highlighted how textual and pictorial combinations together can constitute the “perceptual level” of an image. The textual frame in the Crucifixion painting and the text clippings in the montage exemplifies both image and text as objects of visual perception.

Conclusion: For the Future

All in all, text-based searches on images that include media interrelations are problematic. What our searches on “image and text” and “image in image” have in common is that they presuppose relations (“and,” “in”) between media elements translated into search terms, and that these relations should occur within what is conventionally delineated as an individual image resource. Our search problems are related to the fact that metadata is full of media terms. For example, the term “image” appears in almost all metadata fields in DigitaltMuseum and Alvin (Table 4). The excessive use of “image” and other media terms in these platforms actually risks making images invisible. Even the simplest search for all images held in a particular collection or in the whole platform is cumbersome to perform in DigitaltMuseum, while Alvin is, as noted, more consistently curated and therefore more predictable and easier to navigate. Moreover, this overuse of the word “image” (“bild”) results in many different types of false positives. It returns reproduction photographs of objects, indications of the location of a certain detail in an image, and even photographs of people with the personal name “Bild” (Fig. 8).

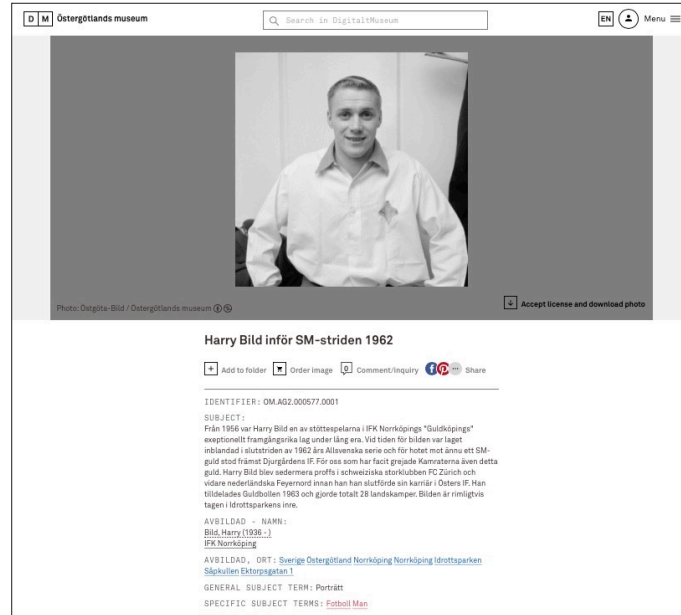


Figure 8. One of the hits when searching on "image (bild)." Football player Harry Bild before the Swedish Championship 1962. Screen shot from DigitaltMuseum. Photograph by Arne Gustafsson. Accessible through DigitaltMuseum: <https://digitaltmuseum.se/021018612736/harry-bild-infor-sm-striden-1962>. Identifier: OM.AG2.000577.0001. License: CC BY-NC.

Our examples of images have been explicit in order to be as eloquent as possible. We have dealt with images with bold inscriptions and with other images as easily recognizable content. The examples have highlighted two particular types of obstacles to systematic searches. As could be expected, image and text relations within images are in DigitaltMuseum searchable only to the extent that the contributing institutions happen to index their images in that way. As in DigitaltMuseum, the result in Alvin is to the point in individual hits, but nonetheless also largely exclusive. In both cases, known examples of image and text combinations in the platforms are not included among the hits. The more important point is that our searches on images in images on both platforms reveals far-reaching variations in what is even indexed as an image, with the implication that our image searches include hits where "image" targets the resource and where it does not target the resource but its photographic mediation. Although this is perfectly logical – who would deny that a photograph is both an image and a picture? – neither platform has viable alternatives for handling image in image searches, whether the image in image relation concern technological mediation or media representation.

The search problems we encountered should be considered in line with Yanni Alexander Loukissas's (2019) discussion of the heterogeneity of metadata in

cultural heritage contexts such as online libraries. As Loukissas, we have noted that “old” – sometimes obsolete – metadata migrate online, just as “old” images. Whether old or new, we have also noted that the metadata in our study is based on an image concept that does not incorporate the last few decades’ results of image studies and media studies, which is part of the reason why our searches on “mixed media” images were not entirely successful. Nonetheless, the restricted image concept that informs (old) metadata is a highly interesting part of the media history of images. On the one hand, it does not, as restricted to figural depiction, allow images to represent media or relations between media, or exemplify such relations. The former would be a case of understanding the subject of a resource as “image and text” or “image in image,” while the latter would be a case of form/genre or type terms. On the other hand, the bewildering variety of pictorial genres recorded as images in the online catalogues transgress traditional categories such as “fine arts,” “folk arts,” and even the inclusive iconographic research tradition in the wake of Aby Warburg. The future study of the digital cultural heritage archive as both a research tool and a research object, a node in the media history of images, is in no way made impossible by our not all too successful searches, even though our results highlight the gatekeeping function of metadata, rather than the door opening function it could have.

Lastly, we wish to highlight some examples that illustrate what we think of as a media history of images that is both unfettered by a linear narrative and does not enclose the images in the past, but attends to them as actualized in the present. James Mussel’s (2012) investigation of the process of digitizing the nineteenth-century press is a good example of careful attention to the interplay of image and text in nineteenth-century journals paired with thorough consideration of how this interplay continues in the online repository *Nineteenth-Century Serials Edition* (ncse.ac.uk). Mussel’s work is concerned with how the printed pages being digitized are mediated and edited again, with ontological (image as code, printed text as digital image) as well as epistemological implications (the senses made of the image and text interplay at different moments in time). Without explicitly mentioning any “media history of images,” Mussel’s understanding of image and text interrelations in the past as well as in the present, unintentionally hits at the heart of such a history.

Further, and in line both with our image in image examples and with Johanna Drucker’s (2014) studies of past and present, printed and digital, graphic interfaces, there is a lot to be made in what we would call a history of the unstable image. How does the digital possibility of making, editing, and reediting endless image copies turn out if framed by the nineteenth-century montage of already mediated actor portraits? The once “fixed,” printed, copy is in the digital archive not only mediated again, but inherently reproducible, editable – and unstable.

But as is evident from much previous bibliographic and print media scholarship, the single item in an analog series is never an exact replica materially, much less epistemologically, since every difference – even slight differences in the amount of ink from copy to copy – would potentially be a different sign, with different effects in different contexts (McKitterick 2003; Batchen 2011). The standard opposition between the dynamic digital image and the fixed analog image, which recurs in Ernst’s writings on the digital archive, may turn out less as an opposition but as different types of instabilities. The reproducibility and editability in the digital present can provide the impetus to scrutinize similar phenomena in past images, which in their turn may give reason to understand the specificities of both better.

Acknowledgement

The research for this article was conducted within the project *The Politics of Metadata* at the Department of Culture and Aesthetics, Stockholm University funded by the Swedish Research Council (grant no 2018-01068).

¹ About metadata see Gartner (2016), Gilliland (2016) Gill (2016), and Riley (2017).

² DigitaltMuseum contains about five million objects from about 250 museums in Norway and Sweden, of which our study concerns the Swedish part comprising 64 museums and more than 3 million objects (KulturIT no year: web). Alvin is smaller in scope, containing about 300,000 objects, mostly from Uppsala University Library (Alvin no year a: web). Both are utilized mostly in Sweden and Norway. DigitaltMuseum had nearly 7 million visits in 2019, while Alvin had nearly 170,000 visits in 2018 (KulturIT no year: web; Alvin no year: web). For statistical surveys of DigitaltMuseum’s user preferences and user demography, see Gran et al. (2019).

³ While the metadata categories and search fields in the platforms have both a Swedish and an English interface, all entries are in Swedish. Therefore, all searches in this study have been made in Swedish. All search terms in the running text are our translations from Swedish, while the indicated search fields and metadata categories equal the English terms of the platform. However, as the translations are not straightforward we have indicated the exact meaning in brackets in the overviews in Tables 2 and 3.

⁴ “Anchoring” as in Barthes (1977: 39), when a textual element (caption, heading, etc.) directs the perception and understanding of the image’s possible meanings.

⁵ In the reports that prepare the digitization of the Swedish National Library, “motif” is defined by distinguishing between depiction, representation, illustration, and symbolization as different types of relations between the “motif” and its referent. These more nuanced discussions are, however, not represented in the cataloguing guidelines (cf. Kungliga biblioteket 2003: 45, 206–207; also Kungliga biblioteket 2000).

⁶ Our search term “Crucifixion” has in DigitaltMuseum been combined with the field “Has media/image” and in Alvin with “Resource type/image.” These search functions are not equivalent. In DigitaltMuseum “Has media/image” indicates that the resource has an image attached to it, whereas in Alvin “Resource type/image” indicates that the resource is an image. It should also be noted here that ICONOCLASS, perhaps the most well-known controlled vocabulary of iconographic subjects, originating in the 1940s and accessible online since the 1990s, has not, judging from metadata, been consistently employed by the institutions connected to DigitalMuseum. The search term “Crucifixion” is solely motivated by expected recognition.

⁷ See note 6.

⁸ The result is dominated by the work of the photographer Hilding Mickelsson (1919–2002), known for his photographs of nature, but with a keen interest in pictorial-textual combinations in folklore, contemporary everyday culture, and children's drawings. Of the total 729 hits, 500 represent Mickelsson's work through the indexing of Hälsingland's museum. Of those hits that represent image and text combinations, Mickelsson's work stands for the majority.

References

- Algee-Hewitt, Mark et al. (2018): *Interacting with Print: Elements of Reading in the Era of Print Saturation*. Chicago: University of Chicago Press.
- Alpers, Svetlana (1983): *The Art of Describing: Dutch Art in the Seventeenth Century*. Chicago: University of Chicago Press.
- Alvin (no year a): "Background and History." <https://info.alvin-portal.org/about-alvin/background-and-history/> (accessed 1 December 2020).
- Alvin (no year b): "Advanced Web Statistics." <http://www.alvin-portal.org/stats/www/2018/awstats.alvin.html> (accessed 1 December 2020).
- Andersson, Kent-Inge & Stefan Andersson (2020): "Katalogiseringsregler." <https://wiki.epc.uu.se/display/alvininfo/Katalogiseringsregler> (accessed 1 December 2020).
- Barthes, Roland (1977): *Image Music Text*, trans. Stephen Heath. London: Fontana.
- Batchen, Geoffrey (2011): *Repetition och skillnad: Fotografins re/production*, trans. Cecilia Grönberg & Jonas Magnusson. Hågersten: CO-OP.
- Bateman, John A. (2014): *Image and Text: A Critical Introduction to the Visual/Verbal Divide*. London: Routledge.
- Bateman, John et al. (2017): *Multimodality: Foundations, Research and Analysis: A Problem-Oriented Introduction*. Berlin: De Gruyter Mouton.
- Belting, Hans (2005): "Image, Medium, Body: A New Approach to Iconology." *Critical Inquiry*, 31:2, 302–318.
- Beaudoin, Joan E. (2007): "Visual Materials and Online Access: Issues Concerning Content Representation." *Art Documentation*, 26:2, 24–28.
- Bernadou, Agiati et al. (2018): "Introduction: A Critique of Digital Practices and Research Infrastructures." Agiati Bernadou et al. (eds.): *Cultural Heritage Infrastructures in Digital Humanities*. Abingdon, Oxon: Routledge, 1–14.
- Bolter, Jay David & Richard Grusin (1999): *Remediation: Understanding New Media*. Cambridge: The MIT Press.
- Choi, Youngok (2017): "Nature of Tags in a Knowledge Organization System of Primary Visual Resources." *Journal of Library Metadata*, 17: 1, 37–53.
- Dahlgren, Anna (2013): *Ett medium för visuell bildning: Kulturhistoriska perspektiv på fotoalbum 1850–1950*. Stockholm: Makadam.
- Dahlgren, Anna (2018), *Travelling Images: Looking Across the Borderlands of Art, Media and Visual Culture*, in the series Rethinking Art's Histories. Manchester: Manchester University Press.

- Dellman, Sarah (2019): "Analogue Objects Online: Epistemological Reflections on Digital Reproductions of Lantern Slides." *Early Popular Visual Culture*, 17:3–4, 322–340.
- DigitaltMuseum (no year): "About DigitaltMuseum." <https://dok.digitaltmuseum.org/en/about> (accessed 1 December 2020).
- Drucker, Johanna (2014): *Graphesis: Visual Forms of Knowledge Production*. Cambridge: Harvard University Press.
- Elkins, James, ed. (2008): *Visual Literacy*. New York: Routledge.
- Eakins, J. P. (2002): "Towards Intelligent Image Retrieval." *Pattern Recognition*, 35:1, 3–14. doi: 10.1016/S0031-3203(01)00038-3.
- Elleström, Lars (2014): *Media Transformation: The Transfer of Media Characteristics Among Media*. Basingstoke: Palgrave Macmillan.
- Elleström, Lars (2019): "The Modalities of Media II: An Expanded Model for Understanding Intermedial Relations." Lars Elleström (ed.): *Beyond Media Borders: Intermedial Relations among Multimodal Media*, vol. I. Cham: Springer International Publishing, 3–91.
- Ernst, Wolfgang (2013): *Digital Memory and the Archive*, ed. and intro. Jussi Parikka. Minneapolis: University of Minnesota Press.
- Gartner, Richard (2016): *Metadata: Shaping Knowledge from Antiquity to the Semantic Web*. Switzerland: Springer.
- Gill, Tony (2016): "Metadata and the Web." Murtha Baca (ed.): *Introduction to Metadata*. Los Angeles: Getty Research Institute, 21–37.
- Gilliland, Anne J. (2016): "Setting the Stage." Murtha Baca (ed.): *Introduction to Metadata*. Los Angeles: Getty Research Institute, 1–20.
- Gran, Anne-Britt et al. (2019): "A Digital Museum's Contribution to Diversity: A User Study," *Museum Management and Curatorship*, 34:1, 58–78.
- Heffernan, James A. W. (2006): *Cultivating Picturacy: Visual Art and Verbal Interventions*. Waco: Baylor University Press.
- Hollink, Laura et al. (2004): "Classification of User Image Descriptions," *International Journal of Human-Computer Studies*, 61:5, 601–626.
- Jappy, Tony (2013): *Introduction to Peircean Visual Semiotics*. London: Bloomsbury.
- Jenkins, Henry (2006): *Convergence Culture: Where Old and New Media Collide*. New York: New York University Press.
- James, Alejandro & Shih-Fu Chang (2020): "Conceptual Framework for Indexing Visual Information at Multiple Levels." *SPIE Internet Imaging*, 3964.
- KulturIT (no year): "DigitaltMuseum." <https://kulturit.org/digitaltmuseum> (accessed 1 December 2020).
- Kungliga biblioteket (2000): *Plattform för bilddatabaser*. Stockholm: Kungliga biblioteket.
- Kungliga biblioteket (2003): *Bilddatabaser och digitalisering – Plattform*

- för ABM-samverkan: Ett samverkansprojekt mellan Kungl. Biblioteket, Nationalmuseum, Riksantikvarieämbetet och Riksarkivet 2001–2002. Stockholm: Kungliga biblioteket.
- Kungliga Biblioteket (2013): “Katalogisatörens verktygslåda: Bilder.” <https://katalogverk.kb.se/katalogisering/Katalogisering/bildkatalogisering/index.html> (accessed 1 December 2020).
- Lager Vestberg, Nina (2013): “Ordering, Searching, Finding.” *Journal of Visual Culture*, 12:3, 472–489.
- Library of Congress (1995): “Thesaurus for Graphic Materials I: Indexing Images.” <https://www.loc.gov/rr/print/tgm1/iib.html> (accessed 1 December 2020).
- Loukissas, Yanni Alexander (2019): *All Data are Local: Thinking Critically in a Digital Age*. Cambridge: The MIT Press.
- McKitterick, David (2003): *Print, Manuscript and the Search for Order, 1450–1830*. Cambridge: Cambridge University Press.
- Mitchell, W. J. T. (1986): *Iconology: Image, Text, Ideology*. Chicago: University of Chicago Press.
- Mitchell, W. J. T. (2015): *Image Science: Iconology, Visual Culture, and Media Aesthetics*. Chicago: University of Chicago Press.
- Moxey, Keith (2013): *Visual Time: The Image in History*. Durham: Duke University Press.
- Mussel, James (2012): *The Nineteenth-Century Press in the Digital Age*. Basingstoke: Palgrave Macmillan.
- Parikka, Jussi (2012): *What is Media Archaeology?* Cambridge: Polity Press.
- Pringle, Mike (2010): “Do a Thousand Words Paint a Picture?” Chris Bailey (ed.): *Revisualizing Visual Culture*. Farnham: Ashgate, 11–28.
- Rajewsky, Irina O. (2005): “Intermediality, Intertextuality, and Remediation: A Literary Perspective on Intermediality.” *History and Theory of the Arts, Literature, and Technologies*, 6, 43–64.
- Riley, Jenn (2017): *Understanding Metadata: What is Metadata, and What is it For?* Baltimore: National Information Standards Organization (NISO).
- Seeman, Dean & Heather Dean (2019): “Open Social Knowledge Creation and Library and Archival Metadata.” *KULA: Knowledge Creation, Dissemination, and Preservation Studies*, 3:1, article 13.
- Stoichita, Victor I. (1997): *The Self-Aware Image: An Insight into Early Modern Meta-Painting*, trans. Anne-Marie Glasheen. Cambridge: Cambridge University Press.
- Thomas, Julia (2011): “Iconographic Description in DMVI.” *Database of Mid-Victorian Illustration*. <https://www.dmvi.org.uk/guide.php?page=11> (accessed 1 December 2020).
- Waldron, Zoe et al. (2018): “Cataloging and Metadata Practices Survey Report.

December 20, 2017.” *VRA Bulletin*, 45:1, article 3.

Wildfeuer, Janina et al. eds. (2019): *Multimodality: Disciplinary Thoughts and the Challenge of Diversity*. Berlin: De Gruyter.

Williams, Keith et al., eds. (2019): *Art and Science in Word and Image*. Leiden: Brill Rodopi.

Wolf, Werner (2009): “Metareference Across Media: The Concept, its Transmedial Potentials and Problems, Main Forms and Functions.” Werner Wolf et al. (eds.): *Metareference Across Media: Theory and Case Studies*. Amsterdam: Rodopi, 1–85.

Wolf, Werner (2011): “Is There a Metareferential Turn, and If So, How Can It Be Explained?” Werner Wolf (ed.): *The Metareferential Turn in Contemporary Arts and Media*. Amsterdam: Rodopi, 1–47.

Appendix

Table 1. Performed Searches

DigitaltMuseum (advanced search)

Search fields	Search terms						
	Search A	Search B	Search C	Search D	Search E	Search F	Search G
search for	korsfästelse*	bild OCH text	inskriftion	text			
type							
topic					inskriftion	text	bild
place							
time							
media	har bild		har bild	har bild	har bild	har bild	har bild
producer							
No. of hits	92	88 946	2 277	201 635	722	729	23 046

Alvin (extended search)

Search fields	Search terms						
	Search A	Search B	Search C	Search D	Search E	Search F	Search G
resource type (from list)	bild		bild	bild	bild	bild	bild
free text	korsfästelse*	bild OCH text	inskriftion	text			
person							
role (from list)							
title							
year							
object type							
subject					inskriftion	text	bild
format							
technique							
material (from list)							
No. of hits	19	688	38	1 501	9	8	27

Table 2. Fields for Descriptive Metadata in DigitaltMuseum and Alvin
Examples of entries are taken from Figures 2 and 6 in this article in the online interfaces.

DigitaltMuseum

Swedish interface	English interface [our literal translation]	Examples of entries translated into English
benämning	designation	oil painting
avbildad-namn	avbildad-namn [depicted person]	Jesus
titel	title	Crucifixion
ämnesord	indexing term [subject term]	Art, paintings, oil paintings, religious objects
motiv-ämnesord	specific subject terms [motif-subject term]	religion
beskrivning	about [description]	
typ	type	thing
mått	dimensions value	H 88 cm, W 85 cm
material	materials	cloth, oil paint
teknik	techniques	painting

Alvin

Swedish interface	English interface	Examples of entries translated into English
personer	persons	Thelott, Anna Maria, 1683–1710 (engraver, creator)
tillkomstinformation	origin	Stockholm, Sweden, 1710
ämnesord, personnamn	subject, persons	Jesus
ämnesord, allmänna	subject, topics	Crucifixion
abstract/beskrivning		A woman by a canvas and easel
object category	object type	Print
ämnesord, genre/former	subject, genre/forms	Engraving, copperplate, frontespiece
format	format	nondigital + digital, reformatted digital
material	material	paper

Table 3. Search Fields Pertaining to Descriptive Metadata in Digitalt Museum and Alvin

DigitaltMuseum (advanced search)

Swedish interface	English interface	Examples from pre-selected list
sök efter	search for	
typ	type	photograph, thing, fine art, media, architecture
tema (filtrera på ämne)	topic	
plats	place	
tid	time	
media	media	has media, has no media (photographic reproduction)
tillverkare	producer	photographer, artist, designer, author

Alvin (extended search)

Swedish interface	English interface	Examples from pre-selected list
resurstyp	resource type (select from list)	image, map, book, video
fritext	free text	
person	person	
roll	role (select from list)	artist, author, book designer, copyright holder
titel	title	
år	year	
objektkategori	object type	album, art reproduction, banknote, comic strip
ämnesord	subject	
format	format	digital, non digital
teknik	technique	ink, watercolor, ballpoint pen
material	material (select from list)	glass, paper, ivory, wood

Table 4. Metadata Fields where the Term “Image” may Occur

DigitaltMuseum	
designation	x
avbildad-namn	x
title	x
indexing term	x
specific subject term	x
about	x
type	x
dimensions value	
materials	x
techniques	x

Alvin	
persons	
origin	x
subject, persons	x
subject, topics	x
abstract	x
object type	x
subject, genre/form	x
format	
material	