



Københavns Universitet



## Rethinking image indexing?

Christensen, Hans Dam

*Published in:*

Journal of the Association for Information Science and Technology

*DOI:*

[10.1002/asi.23812](https://doi.org/10.1002/asi.23812)

*Publication date:*

2017

*Document Version*

Version created as part of publication process; publisher's layout; not normally made publicly available

*Citation for published version (APA):*

Christensen, H. D. (2017). Rethinking image indexing? *Journal of the Association for Information Science and Technology*, 68(7), 1782-1785. <https://doi.org/10.1002/asi.23812>

# Rethinking Image Indexing?

Hans Dam Christensen

Royal School of Library and Information Science, University of Copenhagen, Birkøvej 6, DK-2300, Copenhagen S, Denmark. E-mail: [hansdam@hum.ku.dk](mailto:hansdam@hum.ku.dk)

An abundance of literature on image indexing, visual and image retrieval methods, content-based image retrieval, image tagging, visual information seeking, and so on, is available in information studies. In 2008, decades of development with its great diversity in approaches was summed up as an “evolution” where the literature recently had “grown at a stupendous rate” (Enser, p. 531). In this “evolutionary” process, critical inspections are, however, also needed in specific cases. In the following, three aspects are in concern: (a) the use of an interpretation model created by the renowned art historian Erwin Panofsky (1892–1968); (b) the danger of simplifying theoretical approaches from domains outside information studies, in this case approaches from an art historical methodology; and (c) attention to new “mutations” that might override the old ways of doing things. In particular, the article “Modeling and Analyzing the Topicality of Art Images” (Huang, Soergel, & Klavans, 2015), recently published in this journal, will be targeted as an indication of these aspects.

According to, for example, Enser (2008) and recently Rafferty and Albinfalah (2014), Panofsky figures prominently in the literature on image indexing. In particular, Sara Shatford Layne introduced Panofsky into information studies. In a well-cited text from 1986, she applied Panofsky’s three-level model of interpretation (known as pre-iconography, iconography, and iconology) as a theoretical approach to subject analysis in art images, stating that although he developed the model in terms of Renaissance art, it seemed “possible to apply it to any representational work” (p. 43). In Enser’s words, Shatford Layne, “. . . was instrumental in generalizing Panofsky’s analysis, simplifying the first two modes in terms of ‘generic’ and ‘specific’, and amplifying these by distinguishing between what a picture is ‘of’ and what it is ‘about’” (2008, p. 533).

With due regard for Shatford Layne, Huang et al. (2015) is another contribution with Panofsky as a theoretical primer. In its own words, the article aims at presenting “a

framework for modeling and analyzing image topics and image tagging that improves our theoretical understanding of image topicality” (Huang et al., 2015, p. 1616). The scope of the current text does not allow an in-depth questioning of theoretical understandings in general, but in this article the logic of Panofsky’s model is inverted (again) and, in addition, vernacular art historical practices are taken for granted as rendering a solid domain specific methodology.

First and foremost, it is worth noticing that the representational aspect, stressed by Shatford Layne, has now vanished; pragmatic reflections on image retrieval in general, presented by, for example, Greisdorf and O’Connor (2008), and specific critical reflections on the use of Panofsky, presented by, for example, Burford, Briggs, and Eakins (2003), are also missing. This neglect is indicative of the manipulation of Panofsky’s three strata (1939), which, moreover, are introduced by a new layer, “Layer 1, the visual features of the image”; all four layers are inscribed in the so-called “Facet 1: Layers of image topicality” (Figure 1). Accordingly, the preliminary features mark “visual elements” (colors, shapes, forms, patterns, lines, and styles), “visual constructs” (relationships among visual components as well as the coherent structure held together and ordered by the use of similar shapes, forms, and colors), and techniques used by the artist, “such as pointillism” (Huang et al., 2015, p. 1617).

Taking a closer look at Huang et al.’s (2015) argument, it should be noticed that Pointillism is described as an artist’s technique at the same time that style is part of the separated visual elements in Layer 1. This is confusing: Is Pointillism a technique? Is it not rather a style, such as Rococo, Impressionism, and so on? Probably both in the vernacular of art history, but Huang et al. (2015) are developing a theoretical model, so a more accurate terminology might be expected?

The article goes on: “. . . topical relationships on this layer [Layer 1] are explicit and straightforward”; this layer “does not focus on high-level meaning, it relies on minimal context information” (Huang et al., 2015, p. 1617). “Minimal context information” is changed into “minimal cultural knowledge” at the first level (2.1) in the next layer, “Layer 2: Meaning and interpretation,” and this level is identified as Panofsky’s “pre-iconography,” which is “based on factual or expressional objects or events that could be interpreted with minimal cultural knowledge” (Huang et al.,

---

Received May 3, 2016; revised August 11, 2016; accepted September 6, 2016

© 2017 ASIS&T • Published online 0 Month 2017 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/asi.23812

TABLE 1. A comprehensive framework for the modeling and analysis of image topicality.

---

**Facet 1: Layers of image topicality**  
*Layer 1: Visual features of the image*  
*Layer 2: Meaning and interpretation of the image*  
 2.1 Primary or natural subject matter (Pre-iconography)  
 2.2 Secondary or conventional subject matter (Iconography)  
 2.3 Tertiary or intrinsic meaning or content (Iconology)  
*Layer 3: Function of the image*  
 3.1 Effect of the image on the viewer  
 3.2 Use of the image

**Facet 2: Perspectives on image topicality**  
*Perspective 1: Art history and theory.* Emphasis on artist's intent  
*Perspective 2: Viewer's perception.* Emphasis on viewer's response

**Facet 3: Image topicality relationships**

---

FIG. 1. Table 1 from Huang et al. (2015).

2015, p. 1618). At the next sublayer (2.2), “iconography,” analysis is “based on recognition of individual objects and historical events using cultural knowledge.”

In defining the upper iconological sublayer (2.3), Panofsky is cited explicitly: the intrinsic meaning “. . . ‘is apprehended by ascertaining those underlying principles that reveal the basic attitude of a nation, a period, a class, a religious or philosophical persuasion – qualified by one personality and condensed into one work’ . . .” (Huang et al., 2015, p. 1618). This, however, is added on immediately following: “. . . speaking from the Perspective 1, Art history and theory with emphasis on the artist’s intent.” This perspective is, in Huang et al.’s (2015) overall theoretical approach, supplemented by “Perspective 2’, ‘Viewer’s perception. Emphasis on viewer’s response’”; both perspectives, constituting the so-called “Facet 2: Perspectives on image topicality,” can be applied to each level of “Facet 1: Layers of image topicality” (Huang et al., 2015, p. 1617). Huang et al. (2015) also present a “Facet 3: Image topicality relationships,” which, although important for their theoretical approach, is irrelevant in the present critique.

Turning to Panofsky’s 1939 text, matters are a little more complicated than as expressed by Huang et al. (2015). For example, Panofsky is not ignorant of the formal aspects of Layer 1, the visual features of the image, when he writes: “The world of pure forms thus recognized as carriers of primary or natural meanings may be called the world of artistic motifs” (1939, p. 5). Moreover, he equates the pre-iconographical description with a “pseudo-formal analysis” (1939, p. 14).

Indeed, Panofsky considers the “history of style (insight into the manner in which, under varying historical conditions, objects and events were expressed by forms)” as a controlling principle at the level of “pre-iconography” (1939, p. 15). In other words, inventing a new preliminary sublayer does not make sense in Panofsky’s model. The reason for doing this is, quite evidently, Huang et al.’s (2015) discussion on automated image retrieval, but why should Panofsky’s model be part of this?

In terms of discussions about the formal qualities of images beyond Panofsky, things are again more complex than just relying on the above-mentioned minimal context information. “Cultural knowledge,” which Huang et al. (2015) identify as growing in the transfer from the pre-iconographical to the iconological sublayer, might not be in question when it comes to the automated indexing of pixels, as every digitally created tone of a color can, of course, be defined within the system, but a color is not just a color in the image-word negotiation. The human eye can discern an abundance of tones, far beyond the possible wording of colors. Habitus, disciplinary socialization, and so on, determine the range of color signifiers (with a range of connotations) that professionals and laypeople are capable of applying. Put differently, so-called cultural knowledge is very present at this level.

What is this type of cultural knowledge more precisely, with regard to Panofsky? Huang et al.’s (2015) reasoning about the increase in cultural knowledge reduces the complexity in Panofsky’s interchanges between the pre-iconographic, iconographic and iconological strata. In short, his model mirrors the process of a continuous mode of interpretation, where a lower stratum might be informed by a higher stratum which again qualifies the understanding of the lower stratum and, further, might inform and displace meaning at the higher strata, and so on.

In other words, Panofsky’s model reflects a hermeneutic interpretation, very much inspired by Wilhelm Dilthey (e.g., 1883) in combining an intuitive hypothesis-formation and a comparative method reviving and securing the “objectivity” of the interpretative process. In a note, Panofsky also uses the term “circulus methodicus” as a parallel to the so-called “hermeneutic circle” (1939, p. 11). Panofsky (1939) was published in a slightly expanded second version in 1955. In the added parts, Panofsky even parallels the interpretative process with the solving of the Riddle of the Sphinx (1955, p. 57). That is, the iconological analysis is a demanding process: metaphorically speaking, the interpreter is risking his or her life, as in the myth of the Sphinx and Oedipus.

So, cultural knowledge with regard to Panofsky is the result of a hermeneutical interpretative process, not the random level of the viewers’ cultural knowledge. Thus, Huang et al.’s (2015) introduction of the above-mentioned “Facet 2: Perspectives on image topicality,” which can be applied at each layer, seems odd in terms of Panofsky’s model. In addition, a sentence such as “a viewer may experience the intrinsic meaning [at the iconological layer] as perceived by him or her without any external reference” (Huang et al., 2015, p. 1618), is meaningless with regard to this model. In their argument, Huang et al. (2015) refer to aesthetic response theory, which is something quite different.

The use of the “artist’s intent” is problematic as well. As can be read from Huang et al.’s (2015) explicit quotation of Panofsky, they consider the intrinsic meaning as “a basic attitude of a nation, . . .”, etc., “qualified by one personality and condensed into one work.” This cannot, however, be equated with the artist’s intent, or “the artist’s intention” or

“the specific topic(s) originally intended by the artist” (Huang et al., 2015, p. 1620), but lies beyond the conscious action of the artist in question. In general, “intentionality” cannot be reduced to just something intended by the artist. It is a disputed philosophical term; for example, Panofsky uses the term “intention” in a self-declared “scholastic” sense (1940, p. 194), which cannot be elucidated in the present context.

From a more pragmatic point of view, it may suffice to say that, considering the overwhelming number of paintings which have survived few sources are left to document the artists’ intent or the specific contexts in which the artworks were originally created. Thus, although the use of the term “artist’s intent” is widespread in vernacular art history, the use of it as “perspective” in a theoretical model on image indexing is questionable.

It is time to ask why this vigilance regarding the use of Panofsky as a model, when apparently it has been very productive in developing new approaches to image indexing in information studies. First, it should be clear by now that the model was conceived as the very opposite in comparison to its current use within this field. Instead of attention to the hermeneutical process, the purpose has been to segregate Panofsky’s strata and constitute distinct categories.

This makeover is further accentuated by the fact that Panofsky’s model might be replaced by more relevant theories, for example, image semiotics in the structuralistic vein; in particular, Barthes’ article “La rhétorique de l’image” (1964) comes to mind. In her seminal texts on subject analysis in images (Shatford, 1984; Shatford Layne, 1986, 1994), Shatford Layne never really turned to semiotic thinking. In particular, her distinction between of-ness and about-ness which, in part, is developed with reference to Gottlob Frege’s (1892) distinction between Sinn (sense) and Bedeutung (reference), quite obviously invites such an approach.

The scope of this text does not allow an in-depth elaboration of Barthes’ image theory; fortunately, Barthes is not unknown to information studies (see, e.g., Buckland, 1997; Cárcamo Ulloa, Marcos Mora, Cladellas Pros, & Castelló Tarrida, 2015; Cronin, 2001; Ibekwe-SanJuan, 2012; Matines-Avila, Smiraglia, Lee, & Fox, 2015; Murphy & Rafferty, 2015; Raber & Budd, 2003; Rafferty & Albinfalah, 2013). However, a semiotic turn might without further detail just be affirming a trajectory of indexing not entirely in accord with the reality of digital and analog picture collections. For example, Huang et al. (2015) operationalize several “facets,” “layers,” and “perspectives” that are further interweaved and expounded in their article. Other theoretical models embrace a similar complexity (e.g., Jörgensen, James, & Chang, 2001, p. 938; Jörgensen, 2003, p. 235). That is (and without doing justice to the strengths of these examples), they require professional users or indexers if they were to have practical implications. In light of this, their possible or future practical implementation in image collection practice appears labor intensive and time-consuming in terms of large picture collections.

This impression is further accentuated by the fact that many picture collections contain huge numbers of objects, inaccurately indexed, inadequately indexed or indexed using older methods that need to be converted into new modes. In addition, new (and partly commercial) platforms and photo-sharing services exist, and will probably continue to emerge in the future, with enormous amounts of digital objects, far beyond the scale and contents of traditional cultural heritage collections; in addition, in particular the social platforms will be far beyond the controlled vocabulary of professional indexers. Instead, users might even radicalize “metadata creation by democratizing it, intrinsically offering a critique of the normative symbolizing systems and structures used in content analysis, as well as the assumed need for consistency, experience and authority in how items are described” (Murphy & Rafferty, 2015, p. 478). That is, instead of a top-down controlled enlightenment project with professional image indexers trying to improve access to large image bases by way of theoretical models, metadata creation might be “democratized” bottom-up.

This is not an entirely new approach. In 2008, Enser briefly touched on “social tagging” as a challenge to professional subject representation but also pointing at “opportunities for beneficial enhancement of both exhaustivity and specificity in subject indexing” (p. 534). Subsequently, for example, Rorissa (2010) and Jörgensen, Stvilia, and Wu (2014) highlighted relevant issues in terms of relationships between controlled index terms and uncontrolled tagging, although they both, more or less, also “express a need to ‘tame’ tags” (Jörgensen et al., 2014, p. 847).

Thus, on one side, new methods for image indexing, conducted by professionals without the assistance of crowdsourcing, seem to be so demanding of resource that practical implementation seems unlikely. A new route in image indexing should probably turn to a less-is-more approach instead of intricate theoretical models, if they are aiming at a generalized usage. This invites, however, us to question the role of indexing when classificatory hierarchies and authority are destabilized and access to digital objects is unlimited in terms of user-generated metadata; at least when the question is addressing the public audience, where services to professionals might still be improved. On the other side, there seem to be at least two obstacles on this route: user motivation and the apparent development in tagging from a descriptive to a more social language, at least if, for example, the observations of two authoritative voices within the museum field are taken for granted (Bernstein, 2014; Simon, 2014).

Last, processes of information-seeking facilitated by way of technological systems, also have to be taken into consideration. The possibilities of adjusting or training automatic indexing algorithms in accordance with previous search terms has, of course, proven itself valuable in commercial use. Taking this route, the makeover of Panofsky’s (1939) model might come to an end very fast.

First and foremost, this text was prompted by the prevalent misuse of Panofsky’s (1939) interpretation model in

information studies. Although his common-sense model has seemingly been productive in developing new approaches to image indexing, it was conceived as the very opposite in comparison to the current implementation. If professional image indexing is to be advanced along this trajectory, image theory in the structuralistic vein will probably be more promising as this theory deliberately seeks to segregate layers of meaning and processes of signification in the image instead of striving for Panofsky's (1939) "circulus methodicus." However, considering the emergence of social media platforms and portals to image collections, and the current status of existing public and private picture collections containing huge numbers of objects inaccurately indexed, and so on, social tagging and automatic indexing algorithms might improve picture indexing and seeking in a more fruitful way.

## References

- Barthes, R. (1964). *La rhétorique de l'image*. *Communications*, 4, 40–51.
- Bernstein, S. (2014). Clear choices in tagging. Retrieved from <http://www.brooklynmuseum.org/community/blogsphere/2014/07/22/clear-choices-in-tagging>
- Burford, B., Briggs, P., & Eakins, J.P. (2003). A taxonomy of the image: On the classification of content for image retrieval. *Journal of Visual Communication*, 2, 123–161.
- Cárcamo Ulloa, L., Marcos Mora, M.-C., Cladellas Pros, R., & Castelló Tarrida, A. (2015). News photography for Facebook: Effects of images on the visual behaviour of readers in three simulated newspaper formats. *Information Research*, 20, paper 660.
- Cronin, B. (2001). Hyperauthorship: A postmodern perversion or evidence of a structural shift in scholarly communication practices? *Journal of the American Society for Information Science and Technology*, 52, 558–569.
- Dilthey, W. (1973). *Einleitung in die Geisteswissenschaften* (1883). In B. Groethuysen (Ed.), *Wilhelm Dilthey: Gesammelte Schriften* (vol. 1). Stuttgart: Teubner.
- Frege, G. (1892). Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, 100, 25–50.
- Enser, P. (2008). The evolution of visual information retrieval. *Journal of Information Science*, 34, 531–546.
- Greisdorf, H.F., & O'Connor, B.C. (2008). *Structures of image collections. From Chavet-Pont-D'Arc to Flickr*. Westport, CT: Libraries Unlimited.
- Huang, X., Soergel, D., & Klavans, J.L. (2015). Modeling and analyzing the topicality of art images. *Journal of the Association for Information Science and Technology*, 66, 1616–1644.
- Ibekwe-SanJuan, F. (2012). The French concept of information science: 'une exception française'? *Journal of the Association for Information Science and Technology*, 63, 1693–1709.
- Jørgensen, C. (2003). *Image retrieval: Theory and research*. Lanham, MD: Scarecrow Press.
- Jørgensen, C., James, A., & Chang, S.F. (2001). A conceptual framework and empirical research for classifying visual descriptors. *Journal of the American Society for Information Science and Technology*, 52, 938–947.
- Jørgensen, C., Avila, B., & Wu, S. (2014). Assessing the relationships among tax, semantics, and perceived usefulness. *Journal of the Association for Information Science and Technology*, 65, 836–849.
- Matines-Avila, D., Smiraglia, R., Lee, H.L., & Fox, M. (2015). What is an author now? Discourse analysis applied to the idea of an author. *Journal of Documentation*, 71, 1094–1114.
- Murphy, H., & Rafferty, P. (2015). Is there nothing outside the tags? Towards a poststructuralistic analysis of social tagging. *Journal of Documentation*, 71, 477–502.
- Panofsky, E. (1939). *Studies in iconology. Humanistic themes in the art of the renaissance*. Oxford: Oxford University Press.
- Panofsky, E. (1940). The history of art as a humanistic discipline. In T. Greene (Ed.), *The meaning of the humanities* (pp. 89–118). Princeton, NJ: Princeton University Press.
- Panofsky, E. (1955). *Iconography and iconology: An introduction to the study of renaissance art. In meaning in the visual arts: Papers in and on art history by Erwin Panofsky* (pp. 26–54). Garden City, NY: Doubleday Anchor Books.
- Raber, D., & Budd, J.M. (2003). Information as sign: Semiotics and information science. *Journal of Documentation*, 59, 507–522.
- Rafferty, P., & Albinfalsh, F. (2014). A tale of two images: The quest to create a story-based indexing system. *Journal of Documentation*, 70, 606–621.
- Rorissa, A. (2010). A comparative study of Flickr tags and index terms in a general image collection. *Journal of the American Society for Information Science and Technology*, 61, 2230–2242.
- Shatford, S. (1984). Describing a picture: A thousand words are seldom cost effective. *Cataloging & Classification Quarterly*, 4, 13–30.
- Shatford Layne, S. (1986). Analyzing the subject of a picture: A theoretical approach. *Cataloging & Classification Quarterly*, 6, 39–62.
- Shatford Layne, S. (1994). Some issues in the indexing of images. *Journal of the American Society for Information Science*, 45, 583–588.
- Simon, N. (2014). Tagging in Museums #blowingupperpersonal #notwhat-weplanned. Retrieved from <http://museumtwo.blogspot.dk/2014/03/tagging-in-museums-blowingupperpersonal.html>