

Naming Names: The Ethics of Identification in Digital Library Metadata

Dean Seeman

Queen Elizabeth II Library, Memorial University of Newfoundland, St. John's,
Newfoundland & Labrador, Canada A1B 3Y1, <dseeman@mun.ca>

Dean Seeman is a Cataloguing and Metadata Librarian at Memorial University of Newfoundland in St. John's, Newfoundland and Labrador, Canada. His responsibilities include managing metadata activities for Memorial's digital library, the Digital Archive Initiative, and cataloguing rare books.



Seeman, Dean. **Naming Names: The Ethics of Identification in Digital Library Metadata**. *Knowledge Organization*. 39(5), 325-331. 17 references.

ABSTRACT: In many digital libraries, visual objects are published and metadata attached to allow for search and retrieval. For visual objects in which people appear, names are often added to the metadata so that digital library users can search for people appearing in these objects. Although this seems straightforward, there are ethical implications of adding names to metadata for visual objects. This paper explores the impact of this action and discusses relevant ethical issues it raises. It asserts that an individual's right to privacy and control over personal information must be weighed against the benefit of the object to society and the professional ethic to authentically represent a resource through its metadata. Context and an understanding of the major ethical issues will inform the practical decision of whether to keep objects online and add metadata to them, but items should generally be published unless there are clear ethical violations or a community relationship is in jeopardy.

Received 16 July 2012; Accepted 16 July 2012

1.0 Introduction

Visual objects (still and moving images) make up a substantial proportion of many digital libraries, archives, and repositories. Visual objects are described through metadata in order to allow for efficient search and retrieval. Metadata captures basic information about the visual object, such as title, publisher, physical description, and subject. One of the most semantically interesting pieces of metadata for visual objects, however, is recording the names of people who appear in them. Although this seems straightforward, putting a visual object online and identifying people within it through its metadata carries with it certain ethical implications.

Digital libraries generally do not add objects if the creators or rights holders of the object have not given their permission. For the most part, little attention is

given to people who appear in these objects. What of their rights, particularly if the object pre-dates the dissemination made possible by the Internet? Is it ethical for a metadata practitioner to identify people in visual objects when these objects can be found and accessed from anywhere in the world? Are metadata professionals comfortable making the link between a visual object and a person who may not be aware this object is being put online at all? Is the metadata practitioner comfortable publishing that link by depositing this information online?

These questions arose while working with a particular collection of video objects on Memorial University of Newfoundland's digital library, the Digital Archive Initiative, which contains a variety of videos from the past 50 years of varying nature (from documentaries to community-based broadcasting, community interest, talk shows, and university course

videos). The various videos show people in a wide variety of contexts, which required facing the decision to associate a video with someone's name. This paper was born out of the uncertainty of whether people would want to be identified in association with an object or want the object depicting them online at all. It is hoped that an exploration of the surrounding ethical issues may provide guidance to answer these questions. First, though, we turn to the issue of what is at stake when we add visual objects to online repositories and add metadata to them.

2.0 The impact of putting visual objects online

When an institution chooses to add a digital object to its online library, archive, or repository, the item is made public. The first notable aspect of putting a digital object online is that it immediately becomes globally accessible. Through the Internet, anyone could conceivably find it. Once someone has accessed the object, it then becomes interactive (Ess 2009). That is, the intended and presumed use of the object cannot be enforced in any way. In the case of digital libraries, research and community interest are primary reasons for an object to be added to a digital library. However, there is no guarantee that it will be used that way. Two adjectives used to describe a digital object online are "global" and "greased" (Ess 2009): global in the sense that it is now accessible to anyone with Internet access, greased in that the object cannot be limited to a particular use. One example of this is the case of *Chang v. Virgin Mobile USA* (Gagnier 2011). In this case, Justin Wong, a church youth group leader, took a photograph of Alison Chang, one of the youth group's members, and posted the photograph on the photo sharing site Flickr. The photograph was given a Creative Commons Attribution license, which allows anyone else to use the photograph in any way, even commercially, as long as the photographer is given credit (Creative Commons 2012). The photograph was eventually used by Virgin Mobile in an ad campaign in Australia. Alison Chang found out about its use and sued Virgin Mobile. Although her case against Virgin was unsuccessful due to lack of jurisdiction, it highlights the greasy nature of online digital objects. The photo was put online for a particular reason (to document a social event) and used in an entirely different way by a commercial body. What this demonstrates is an ultimate lack of control by people appearing in digital objects. To add objects to the Internet is to add further to the mass of online information about a person and alter the online identity of those de-

picted, regardless of their consent. However, the image is not findable and cannot be brought together with other pieces of online information about people if it lacks textual representation. That is the function of metadata.

3.0 The impact of adding metadata to visual objects

Current search engines and databases depend almost exclusively on textual description for retrieval of electronic information and material. Visual objects, then, need a textual surrogate in order to be found and accessed. Unless a visual object is happened upon by browsing, its efficient retrieval will be entirely dependent on the metadata accompanying it. In this way, it is separated from a textual object, which has at least the possibility of being discovered through its full text. Metadata, then, gives voice to the visual object. It both enables and restricts the semantic journey of the content and becomes the idiom of exchange and communication. In a search environment, it becomes "the shadow that overtakes the object that casts it" (Kallinkikos and Mariátegui 2011, 291). The metadata practitioner in this scenario is the gatekeeper of information (Barzilai-Nahon 2008) about visual objects. They control access to the object through their decision to textually describe it.

Among the metadata elements to include for visual objects is the identification of people who appear in the object. The metadata practitioner adds names and becomes responsible for making the link between an identity and an object. He or she makes a visual object findable by who appears in it. The person appearing is represented, in image and text, possibly without their knowledge or consent, and their online identity is changed, possibly irrevocably. The metadata practitioner globally exposes the link (if the object is put online) between someone's name and image. If this happens without the consent or knowledge of the person appearing in an object, the power to control personal information would be taken away from the person. Control over disclosure of information is one aspect of informational privacy (Brey 2000). The impact of adding identifying metadata to visual objects, then, is no less than allowing an object to be found and altering the informational privacy of individuals who appear in them.

4. Ethics

It is clear that the impact of putting a digital object online and adding identifying metadata is significant.

The next question to ask is whether these actions, together with their impact, are ethical. To help inform the ethical discussion surrounding the act of identification in visual object metadata, literature from various fields (information science, journalism, academic research ethics, ethics, information technology) as well as legal literature have been consulted as it relates to the present topic. The result of this survey is the identification of several major themes. The first of these is the issue of the rights of the individual versus the rights of society.

4.1 Ethics: individual vs. society

One clear theme highlighted in the literature is the issue of the rights of the individual versus the rights of society. A definition of ethics is given by Marshall (1999, 82) as "guidelines to influence human social behavior in a manner intended to protect and fulfill the rights of individuals in a society." In the present context, this could be taken to signify the rights of the individual to retain control over their online privacy versus the rights of society to access visual material through its metadata.

In the field of information science, this tension is clearly present in codes of professional conduct. Koehler and Pemberton's study of information professional ethical codes and the resulting model suggests five elements for a model ethical code, including to "place the needs of clients above other concerns," to "be sensitive and responsive to social responsibilities appropriate to the profession," and to be "aware of and responsive to the rights of users, employers, fellow practitioners, one's community, the larger society" (Koehler and Pemberton in Beghtol 2002, 514). Following this ethical code, it is an ethical duty to supply useful access points in metadata records so users can find relevant resources. Britz (in Bair 2005) clearly outlines the professional duty to make information available as contributive justice: to contribute to fair and equitable access to information and maintain information that benefits society. This urge to free information to benefit society, however, must be balanced against compromising the informational privacy (control over personal information) of people appearing in visual objects. Froehlich's principles on ethical research require us to "respect the self and others," "seek to minimize harm," "seek justice or fairness," promote "social harmony," and "be faithful to organizational public and professional trust" (Froehlich in Beghtol 2002). This suggests a more cautious approach: one that would value possible concerns

about individual privacy over society's right to access visual objects. Information science ethics, then, require us to balance the beneficial flow of information to society against the rights of the individual to control information about themselves.

Journalism is a field that also makes images of people public and may also textually identify them. Often, a "greater good" argument is invoked, which insists that images of an individual are published in order to benefit society (Bersak 2006). This can be likened to an Act Utilitarianism approach, which dictates that the "an action is good if its net effect (over all affected beings) is to produce more happiness than unhappiness" (Quinn 2011, 75). Nissenbaum, for one, agrees that the action of registering public information is important (Nissenbaum in Brey 2000). Legally, publishing images of people requires a balance between privacy and free speech so that an individual's rights are not violated while ensuring that information flows out into society to its benefit. The notions of private and public space will be addressed in the next section, but this helps inform the discussion of what images may be published. Even if an image is deemed to come from a private space, however, it still may be published if it can be shown to be newsworthy (i.e., it has inherent social value). If this is the case, it can be published without fear of liability (Blackman 2008). Journalism is not a direct corollary to publishing visual objects in digital libraries, but it does provide a framework which highlights the tension between the rights of an individual and the rights of society. If a visual object benefits society, it may be acceptable to publish it and minimize an individual's right to control information about him or herself.

4.2 Ethics: public vs. private

Another major ethical consideration helpful to the present discussion is whether an image can be considered private or public. This could inform the decision of whether to publish visual objects and add identifying metadata to them. Ess (2009) states that, in contexts and spaces where people can legitimately expect privacy, information about behavior in those spaces should be able to be controlled. The photojournalist Bersak (2006) agrees, but approaches the issue from a different angle, stating that, if a photograph has no expectation of privacy, no invasion of privacy is possible. Attempting to precisely define what is inherently public and inherently private can be problematic, however.

The issue has been raised in the legal literature in the United States and Canada. Two major themes are

publicly disclosing private facts and intruding upon seclusion (Blackman 2008). These insist that there are indeed private facts and private situations; however, these may be breached to various degrees in exchange for social value. Canadian legal literature describes this as a reasonable expectation of privacy: “a person shall have no reasonable expectation of privacy in what he or she knowingly exposes to the public, or to a section of the public, or abandons in a public place” (Scassa 2010, 194). The insistence is that there is a distinct public sphere, and images should be able to be captured of people in those spaces unless they can reasonably expect privacy. Private situations, people’s homes or other intimate surroundings, would carry with them the expectation that privacy would be protected. Images captured within that context would not be published without the explicit consent of the subject or, as mentioned above, they possess a clear social value.

A further consideration is whether, as technology changes, expectations of privacy also evolve. Technological changes (particularly the ubiquity of image capturing devices and the seamless ability to publish images) could change societal definitions of the public sphere and private sphere (Brey 2000) and alter notions of reasonable expectations of privacy in a public space (Scassa 2010). It may also affect what information is considered personal and influence how much control people insist on having, or expect, over their personal information. To Sun Microsystems co-founder Scott McNealy, technology has already changed the privacy landscape whether we like it or not when he stated, “You have no privacy, get over it” (Ess 2009, 30).

The visual object, then, must be probed for whether it was captured in a private or public environment. If captured in a private environment, the metadata practitioner will have to decide whether to treat the object and identifying information as private (meaning taking the object offline). If the image was captured in a public environment, this could allow the object and statements about it (its metadata) to be treated as public information.

4.3 Ethics: authentic representation

Another relevant ethical consideration is the ethical and professional duty as metadata practitioners to authentically represent resources. Bair (2005, 17) expresses this in the following way: “catalogers should work with honesty and integrity to represent the truth about each resource in regard to its subject area, or

‘aboutness,’ the identity of those responsible for the content, and accurate description.” This requires that the information professional respect the object and represent it as truthfully as possible. Brody (2003) calls this fair representation. The practitioner has a responsibility to truthfully represent an object in order to help users find the material, but also because metadata affects the way users understand the object or resource itself (Unsworth 2009). A failure to authentically represent an item in its metadata, then, could be construed as eroding the professional ethic to fairly represent objects to society. This ideal pulls against general ethical concerns that control over personal information is eroded by adding identifying metadata to an object. By adding identifying metadata, the rights of users are satisfied at the expense of individuals. Informational privacy is taken away so that society can benefit through the discovery of visual objects. This makes the professional ethic to fairly represent resources through identifying metadata non-neutral since the implication is clearly the loss of personal informational privacy. Whether this violation should prevent authentic representation is arguable, but it is another ethical tension that must be reckoned with.

4.4 Ethics: protection from harm

Another important ethical concern is protecting people from harm. In its simplest form, this suggests a Kantian duty ethic to treat other people with respect (Brey 1999). The difficulty of this ethical tenet is that it is not always clear what harm is. Determining this would require establishing the effect on an individual of putting a visual object online and adding identifying metadata to it on an individual. The overall effect is not easily ascertainable, but may be related to the level of consent an individual has given to a digital library to publish an image and identify people within it. The problem with this in a digital library context is that it is often unknown whether a person appearing in an object would give their consent. Metadata practitioners cannot gauge consent or the possible harm associated with naming an individual in metadata because they do not necessarily know how people feel about their image being put online and them being identified with it. The person being identified could be pleased, elated, indifferent, embarrassed, or outraged. In addition, even if consent was given at the time to make the object somewhat public, could the people appearing in these objects, especially if it is over twenty years old, have known that it would eventually be made instantly and globally accessible

when it was put online? To add another layer of complexity, views on fair treatment of individuals and entire conceptions of privacy can differ between cultures. While privacy is of primary importance to those in the West since it implies having the freedom to choose one's own identity, this may not be as important in some African or Asian cultures (Ess 2009). This brings us back to the notion of unpredictability. Metadata practitioners can guess, or assume, whether an individual or someone from a particular culture would not want to be identified and their image put online, but they do not know it for certain.

So, if harm and consent are difficult to establish, what would a protection from harm ethic mean? The instruction given here is that it would be entirely driven by context, but there are some general guidelines to allow this general ethic to be followed. The first is that there is a group of people in society who are not able to fully protect, or speak for, themselves: the vulnerable. As an ethic in society, the more vulnerable someone is, the greater the responsibility society bears to protect them (Ess 2009). In the case of digital libraries, this requires making sure the vulnerable in visual objects are not published and named if in a vulnerable situation. The exact definition of vulnerability would depend on the context of a particular video, but awareness of this issue is important in order to allow for ethical practice. Other than this group, a context-driven caution to treat others with respect and with dignity would be part of the ethic to protect others from harm. An example of this is the ethic to protect the rights, dignity, privacy and well-being of people involved in academic research (Wiles et al., 2008).

4.5 Ethics: discussion

The ethical themes discussed above leads to an awareness of several issues: the benefit of information to society and what this means for an individual's right to control their personal information, the discussion of public and private, the ethical duty of a metadata practitioner to authentically represent items, and the need to consider whether people in a visual object need to be protected from the possible harm incurred as a result of visual object being published and metadata added. The only clear conclusion that can be reached is that context often determines which ethical consideration should take precedence. The argument to make material generally available, however, is a convincing one, unless a clear and egregious ethical violation is committed in order to do so. This benefits so-

ciety and allows the metadata practitioner to authentically represent resources. This should be tempered by questioning whether the object was captured in public or private and whether a clear harm is being inflicted. Although this may seem bold, digital library visual material is filled with too many other unknowns to allow for much to be published at all if a more cautious approach is taken. Besides this inclination, little clarity can be imposed. The most helpful proposal is to keep these issues in mind simultaneously while working with visual objects in order to come to an appropriate decision.

5. Practicalities

The discussion of ethical considerations (the individual and society, public and private, authentic representation, and protection from harm) allows us to attempt to apply this discussion to metadata practice. When dealing with an ethically questionable visual object, the practical choice could elicit two responses. The first would be to make the decision not to name someone in the metadata if their appearance was ethically questionable. This approach would make the object considerably less findable, but still allow that a person would appear in the object. Their textual identity would be censored, but their image, or visual identity, would remain. This would mitigate an ethical problem somewhat, but would not solve it completely. The second option is to remove the object completely. This is the suggested course of action as it would remove both the visual and textual identity, thereby eliminating the ethical problem.

In support of building an applied ethic, several practical questions can be asked of a visual object, informed by the ethical discussion above. First, was the image captured in a private setting or public setting? If in a private setting, the ethical and legal repercussions of publishing an image need to seriously be considered. If in a public setting, publishing an image should be legal, but is there any indication that the person appearing in the image would expect the image to be kept private? Another consideration, if it can be ascertained easily, is whether there was an expectation for dissemination when the image was captured. Was there an expectation that this image would be made public? If so, how widely would the person appearing in the object think it would have been made available? If there is evidence to suggest that it would have been published or made available to a very select and personal audience, the object's inclusion in a digital library may need to be reconsidered.

Another question is whether someone in the video is in what might be termed a vulnerable state. If they are, the metadata practitioner has a responsibility to act, most likely requiring the object to be taken down completely. Good practical examples of this might be medical patients or, depending on the nature of the object, children. A final question relates to when the object was captured. If an object was created more recently, the impact on someone's identity and privacy would most likely be greater. If the object is more historical in nature, the impact could be far less. If the person appearing is deceased or the image is quite old, objections to the object being made public may be less likely, but the potential damage to a family, community, or reputation could also be a factor. Although older objects may present less immediate ethical issues, they do not guarantee an ethically uncomplicated response.

In summary of the practical considerations, a few things should be emphasized. First, a balance must be struck between the rights of the individual and the obligation to society and the community. The metadata practitioner should be aware of the effects of publishing images and metadata, but still ensure that images are not prevented from being published unnecessarily, due to caution, so that society is still able to benefit. The second is that performing rights clearance and attempting to secure individual consent from all people appearing in a visual object is very difficult, if not impossible. This would involve securing permission from a large number of people or their estates and the effort of this exercise would likely be sufficiently taxing to prevent many objects from being made public at all. A practical solution is to add objects to a digital library and remove objects if questioned, keeping in mind the ethical discussion presented in this paper to eliminate objects that are clearly ethically dubious. Partnered with this boldness, a digital library should liberally take objects down if people object to them, particularly people appearing in those objects. There is a need to keep in mind that continued good relations with the community (Wiles et al., 2008) are essential to the work of most digital libraries. This flexibility allows people appearing in visual objects to control their identity should they choose to exercise it.

6. Conclusion

Naming names in the metadata record has a tremendous impact on people who appear in visual objects. No less than personal privacy and control over identity are at stake, which must be balanced against the

rights of society and other ethical considerations. Like most ethical discussions, there are few easy and clear cut answers when bringing an ethic into practice. At the very least, discussion of these broad ethical themes should broaden our understanding of the relevant issues, resulting in a more ethical practice.

References

Bair, Sheila. 2005. Toward a code of ethics for cataloging. *Technical services quarterly* 23n1: 13-26.

Barzilai-Nahon, Karine. 2008. Toward a theory of network gatekeeping: a framework for exploring information control. *Journal of the American Society for Information Science and Technology* 59: 1493-1512.

Beghtol, Clare. 2002. A proposed ethical warrant for global knowledge representation and organization systems. *Journal of documentation* 58: 507-32.

Bersak, Daniel R. 2006. "Ethics in photojournalism: past, present, and future." Master's thesis, Massachusetts Institute of Technology. Available <http://web.mit.edu/drB/Public/PhotoThesis/>.

Blackman, Josh. 2008. Omniveillance, Google, privacy in public, and the right to your digital identity: a tort for recording and disseminating an individual's image over the Internet. *Santa Clara law review* 49: 313-88.

Brey, Philip. 1999. The ethics of representation and action in virtual reality. *Ethics and information technology* 1: 5-14.

Brey, Philip. 2000. Method in computer ethics: towards a multi-level interdisciplinary approach. *Ethics and information technology* 2: 125-29.

Brody, Roberta. 2003. Information ethics in the design and use of metadata. *IEEE technology and society magazine* 22n2: 34-39.

Creative Commons. 2012. "Attribution 3.0 license." Available <http://creativecommons.org/licenses/by/3.0/legalcode>.

Ess, Charles. 2009. *Digital media ethics*. Malden, Mass.: Polity.

Gagnier, Christina. 2011. On privacy: liberty in the digital revolution. *Journal of high technology law* 11n2: 229-79.

Kallinikos, Jannis, and Mariátegui, José-Carlos. 2011. Video as digital object: production and distribution of video content in the Internet media ecosystem. *Information society* 27: 281-94.

Marshall, Kimball P. 1999. Has technology introduced new ethical problems? *Journal of business ethics* 19: 81-90.

Quinn, Michael J. 2011. *Ethics for the information age*. Boston: Addison-Wesley.

Scassa, Teresa. 2010. Information privacy in public space: location data, data protection and the reasonable expectation of privacy. *Canadian journal of law and technology* 7: 193-220.

Unsworth, Kristene. 2009. Ethical concerns of information policy and organization in national security. *Cataloging & classification quarterly* 47: 642-56.

Wiles, Rose, Prosser, Jon, Bagnoli, Anna, Clark, Andrew, Davies, Katherine, Holland, Sally, and Renold, Emma. 2008. *Visual ethics: ethical issues in visual research*. Southampton: ESRC National Centre for Research Methods. Available <http://eprints.ncrm.ac.uk/421/1/MethodsReviewPaperNCRM%2D011.pdf>.