

Albums in the attic. An investigation of photographic metadata

Vinayak Das Gupta

Maynooth University, Ireland,
E-mail: vinayak.dasgupta@gmail.com

Abstract: This article investigates the challenges of creating online research collections, specifically for photographic objects. This examination considers the material aspects of the photograph and considers the manner in which they may be made accessible within a web resource. The article explores the challenges of attributing metadata for visual resources in an attempt to address some of the key issues in the production of resources in the Digital Humanities.

Keywords: photographs; images; art history; metadata; material culture; ontology.

In the late nineteen-eighties, my father owned a camera which would imprint the time of the photographic moment on the film. Every print would bear a sequence of faded, red numbers on its edge. Perhaps the camera had a fault, or perhaps my father never bothered to change the batteries, but the imprinted time would invariably be incorrect. Those numbers had their own logic, one that, to this day, I haven't quite managed to understand. But, as a child I would marvel at the manner in which even that incorrect inscription would guide the way in which the photograph was perceived. Those faded red numbers on the bottom-right corner of the photographic paper would enter the image into a series that would potentially enable me to understand the sequence in which the photographs were taken. This was my first encounter with metadata, even if it was one without realisation. It is now common to expect the digital camera to record a host of information from shutter speed to the geo-location of the photograph. These machine-recorded data guide the way in which the image is perceived. Metadata not only shapes the way the object is understood but also aids in

its search and retrieval within digital repositories. The metadata of each individual object affects the experience of the entire collection. Metadata, then, provides a significant way in which descriptions can be provided for objects that are not self-describing.¹ This article contends with, specifically, the possible ways of articulating the material aspects of the photographic image. It is argued that beyond the content image (the indexical quality of the photograph), contextual clues derived from inspections of its physical properties are important for a coherent understanding of the object. Following an examination of the physical nature of the photograph, its presentational forms are inspected; the content image, the material existence, and the presentational forms of the photograph, together, construct a consistent body of knowledge that affords a deeper understanding of the object. An accurate digitisation of the physical object into a digital image (specifically for research purposes), then, must present both the content image and the physicality of the object with sincerity. While current digitisation techniques have found accurate methods of copying the content image, the description of its materiality remains a challenge. If the physicality of the photograph is central to its understanding, this article inspects the possibilities of providing the digital referent with the material information. It presents an examination of both the materiality of the photographic image (and its transformation into a digital object), and the means through which the presentational forms of the original may be inscribed in the digital referent.

1. Photographic materiality

While the image imprinted on photographic paper is important, it would be an error to assume that the photograph may only be understood as an image; scholarly study of the photographic image, specifically with the 'material turn' in anthropology and cultural studies,² has attempted to inspect the photograph as a physical object. This 'material turn' expresses the complexity of the social existence of objects and allows the investigation of the photograph as a physical object whose understanding is augmented by its form. Photographs have specific modes of circulation, production, and consumption, and their inspection has potential beyond the critiques of representation

¹ These are objects that cannot articulate any information about themselves. Unlike the printed book which reveals the author, the date, the publication house etc. (inherent in the structure of the book), visual objects or pieces of music for instance, cannot reveal any information about their nature. So, they are not 'self-describing.'

² Material cultural analysis, from an anthropological position, questions the assumed superiority of language over other forms of expression (Miller 1987: 9). In the context of photography, Elizabeth Edwards (2002: 67) suggests the need to 'break, conceptually, the dominance of the image content and look at the physical attributes of the photograph.'

alone. The inherent bias of the indexical qualities of the image over its material properties may overlook the social and cultural contexts within which the photograph is born and used. '[A]s material objects whose "currency" and "value" arise in certain distinct and historically specific social practices' (Tagg 1988), the stability of photographic meaning remains a descriptive challenge. Arjun Appadurai (1986) explains how objects and the practices within which they are embedded are interwoven, cannot be read independently of each other. He writes (1986: 5), 'things have no meaning apart from those that human transactions, attributions, and motivations endow them with.' Geoffrey Batchen (1997: 6) contends that the meanings of photographic objects are dependent on the context in which they find themselves at any particular instance. He argues that, '[a] photograph can mean one thing in one context and something else entirely in another' and that the identity of the object is contingent on its 'use' in the physical world. Gillian Rose (2010: 18) echoes similar sentiments when she says, 'what people do with photographs is not an optional analytical extra; it is fundamental to exploring photographs' effects in the world.' The examination of photographs, as material objects, appears to present two inter-related concerns: the first discusses the 'plasticity of the image itself, its chemistry, the paper it is printed on, the toning, [and] the resulting surface variations' (Edwards and Hart 2004). The rationale here is that the choices made in the making of photographs are rarely random. As Schwartz (1996: 58) expounds, 'the choice of ambrotype over paper print implies a desire for uniqueness, the use of platinum over silver gelatine intimates an awareness of status, the use of gold toning a desire for permanence'. The second concern is that of the presentational forms — *cartes de visite*, cabinet cards, albums, mounts and frames³ are intrinsically linked with photographs and have constituted a major consumer market since the twentieth century, especially after the Kodak revolution⁴ of the early twentieth century (Edwards and Hart, 2004).

Digital objects (both born-digital and digitised artefacts) present a conceptual challenge: Joanna Sassoon contends that '[b]y the direct conversion of light into a digital format to create a stable image "photographs" that only exist in the digital form can be seen in one context as a truer version of photography (writing with light) than

³ Photographs were, and still are, used in the production of many curiosities. For instance, see Susan Legêne's (2004) account of playing cards illustrated with photographs (though the production of these were in the first half of the twentieth century).

⁴ The technological changes, specifically with the introduction of the Kodak camera, resulted in the device being smaller, portable, and cheaper. This allowed enthusiasts to practice the form and the portability of the camera allowed them to take landscapes and views with relative ease. This, to some extent, accounts for the larger range of photographic subjects in the early twentieth century. See Mia Fineman's 'Kodak and the Rise of Amateur Photography.' *Heilbrunn Timeline of Art History*. New York: The Metropolitan Museum of Art, 2000. Web. 21 November 2013.

those that require the creation of a physical intermediary to view the image in a material form' (2004). The central rationale of Sassoon's argument takes premise on the assumption that digital objects have no material existence. This reveals a fallacy in our inspection of digital objects: the digital artefact, without question, is a material object. To observe this materiality, a separation between content and carrier of the object is required. The photographic image is printed on paper, and this paper is the carrier of the content image. Similarly, the content of the book is carried in the physical form of the book — the paper, glue, and ink that hold it together. In the physical object (as opposed to its digital counterpart) the content and the carrier are closely inter-linked to the point where their separation is difficult. However, the carrier may change at different moments of time, which may provide the object with different contexts: consider an image that was first printed on photographic paper, then printed in a newspaper, and then, perhaps, in a coffee shop book, displayed on a bookshelf, never quite read.⁵ The different material existences of the image provide contexts that locate it within different points in history. The digital object is, similarly, carried by electronic circuits — a complex creation of technology that is able to house a large amount of data within a tiny space. That the digital object has materiality, then, is undeniable. The problem, perhaps, in identifying this materiality lies in the degrees of separation between the circuits and the perception of the object. To view an electronic image, a screen — an enabler — is required. The experience of the object then is governed by an intermediary system. The dislocation between the carrier of the object and the experience of the object is perhaps the source of the material fallacy. If it is acceptable to think of this separation between the content and the carrier, it is possible to argue that the digital object is merely one iteration of a different carrier. The digitised image, then, is the original image, in a new material form. However, it must also be recognised, that the materiality of the digitised object does not mirror that of the original photograph. Within our current technological limitations, the only possible method of attributing the material specifics of the original photograph to the digital referent is through textual means. The metadata of the digital object relates the materiality of the original. The process of digitisation focuses on the content image, converting the original image into a series of pixels,⁶ while the carrier of the photograph is described through

⁵ The circulation of the photograph moving between carriers and put into new uses, derive new contexts. Brent Harris (1998), for instance, in his investigation of photographs from Africa (1850-1950) provides a strong argument for examining the circulation of the photograph. His examination reveals the contested meaning at various stages of the photograph's production, circulation, and consumption.

⁶ The process of digitisation, from a computational perspective, is quite fascinating. A researched description, specifically investigating the process, may be found in Jähne, 1995.

the metadata of the object. Sassoon's investigation of the process of digitising photographs addresses some major concerns. She writes (2004):

[I]t is no longer an accepted canon that a photograph is merely a print on paper, nor is it a simple and uncomplicated translation between reality and its mechanical representation. [...] the digitising process can no longer be seen as merely changing the physical state of a photograph from the material to the pixel. If a photograph can be seen as a more complex object than simply an image, digitising can be seen as more than simply a transformation of state, or a transliteration of tones. The process of digitising involves a more complex cultural process of translation — or a change between forms of representation.

What Sassoon identifies in this transformation is the complexity that lies in articulating details about both the content image and the carrier, even if this distinction is not explicitly made. The metaphor of translation, or even transliteration, that she uses, is perhaps not entirely convincing; the transformation from physical to digital is a re-configuration of state, one that is consistent with any other similar transformation — from the shellac disc to the audio tape or from celluloid to the video cassette. Translation, even if the word is used loosely in Sassoon's work, indicates an inherent change in content *and* carrier. As discussed, digitisation (or effective digitisation) is focused on preserving an accurate content image, even if an accurate representation of materiality is beyond its purview.⁷ What the word 'translation' does point towards is, in fact, the tension between original and the copy, in the accurate representation of the source. What she stresses on repeatedly in her essay is the importance of cultural contexts within which these material objects reside. Any transformation of state, whether physical or digital, creates a new context for the object; the digital state has its own context, while our attempt is to record the context of the source within the digitised.

All digitisation activities are carried out for some specific purpose. The purpose might be merely to record the content image for the purposes of the collection. As we have already discussed, the desire of the collector is central to the affect of the collection. For research collections, however, the need to record more than the content seems to be necessary. The focus of digitisation, then, cannot be only on a singular aspect of the photograph — the indexical quality of the image. The relationship

⁷ Beyond the content image of the photograph, its physical peculiarities are also recorded. The back of the photograph, any tears or marks, and any inscriptions are also captured with the indexical image (this depends, though, on institutional policy or individual motivations). The tactile response of the photograph or its smell are more difficult to record. They may be described, but an inevitable bias is very likely.

between the negative and the paper forms of the photograph may reveal visible clues about its technical origins. The physical dimensions of the photograph may reveal details about the kind of camera used, the negative size, and a possible date of production. The tonal ranges, and the photographic texture may indicate the type of process used to develop the image. All these point to the social existence of the object. Viewed on computer screens of different sizes and different calibrations, it is easy to lose sense of the proportions and tonal hues of the original image.⁸ Digitisation effectively alters the way the photograph is viewed; using a loupe to magnify the details of the content image, we are far readily aware of the material aspects of the photograph than that of the digital referent. The use of keyboards and mice might allow new interactions, and the ability to magnify to extreme points, but they fail to offer the tactile response of the physical object. The fundamental difference in viewing the digital object, brought about through mediation (the computer screen) might lead to questions about the perception of the object. Beyond the material peculiarities, the imprint of time on the object — the dirt and the tear, the slow fading of chemicals on paper — is evidence of the lives of the photograph, about the provenance of the image.⁹ Paratextual material such as captions, scribbles on the back of the photograph, hand-colouring of black and white images, are invaluable to its contextual understanding. While research on advanced techniques of image representations point to the future of digitisation, it is rarely available for average consumer use.¹⁰ How are we to record these elements — the material, the contextual, and the paratextual aspects of the image — with the range of technology afforded to us?

⁸ Enablers, for instance computer monitors, present their own problems. The calibration of the screen (brightness, colour fidelity) along with its physical dimensions change the perception of the digitised object. In most digital collections, the digital referent does not possess a reference point against which it can be compared to derive an understanding of the original. Cultural institutions follow strict rules when digitising objects; unfortunately, they have no control over the enabler and the perception of object, for the viewer, may be fundamentally different from what its natural form is.

⁹ Roland Barthes, in his seminal work titled *Camera Lucida: Reflections on Photography*, begins by describing the the material aspects of the image. He writes, 'The photograph was very old. The corners were blunted from having been pasted into an album, the sepia print had faded, and the picture just managed to show two children standing together at the end of a wooden bridge in a glassed-in conservatory, what was called a Winter Garden back then' (1981: 67). The imprint of time on the photograph is clear from his description; the material contextualises the index.

¹⁰ There is a lot of research on object recognition in images. Vinyal et al.'s (2015) paper shows how complex images are provided a dynamically-generated caption. The work uses computer vision and natural language processing to form a complete image description. Their real-world applications haven't reached the average consumer. Also, these are only able to articulate what is in the photographic frame.

2. The Photograph and its Presentational Forms

The photograph has multiple lives; it exists within socio-cultural contexts, and to understand it, the content image must be seen in conjunction with its material form. Since the inception of photography, photographic images have been used in a variety of contexts, and have been presented in a multitude of ways; the carrier often determines the use the content image is put to. Whether preserved in photo-albums (arranged thematically or sequentially), sold as postcards for the curious, or published as documentary evidence the presentational form of the photograph weighs heavy on the readings of the image. Presentational forms, in particular, guide the way in which photographs were used after their inception, and also the way they were understood. It is, here, important to distinguish between the carrier and the presentational form: the carrier is always material, while the presentational form is ideological. The materiality guides the technical production of the image, and bears the imprint of time on it. The presentational form reveals the social, political, cultural, and religious contexts within which the photograph is *used*. The ideological and the material may exist in the same form.

To elaborate on the affect of presentational forms, and to see more clearly the separation between the carrier and the presentational form, it may be prudent to inspect the photographic album. The photographic paper is attached to the album; the album is the presentational form that is imbued with ideology. Photo albums are unique cultural artefacts in the sense that they look beyond the content and the creator of the photographs to the creator of the object in which the photographs reside. Glenn Willumson (2004), in an essay on the displaced materiality of photographic objects writes:

The performance of thumbing through the photographs, selecting and sequencing, and gluing them into an album breaks the bond of the materiality of the photograph from its links in commerce and mass production. In choosing, sequencing, organising and captioning the photographs for the album, the person responsible transforms the meaning of selected images into an intensely individualistic expression. At the moment of creation, the photo album is a personal artefact, a record of people and events that are rich with biography and personal memory.

Photographic albums are embedded with the traces of their owners and their practices. The family album, for example, is immersed in 'mnemonic frameworks' (Langford 2001). It is argued that while textual communication (captions, titles etc.) may aid the reading of photographs, oral communication plays an important role in the experience of family albums. The oral stories surrounding the family albums are 'performatively intertwined' with the photographs (Edwards, 2009: 39) and thus are

construed within the etiquettes of social interactions. From this, it can be observed that presentational forms bring photographs into use, immersing them within social practices and cultural contexts. The travel album — a form so ubiquitous in the nineteenth century that they are sometimes invisible' (Nordstrom, 2004) — exist as a presentational form that provides layered subjects for examination. They bear testimony to the use, and consequently the context, of the photographic images. Though the examples in this thesis are mostly derived from South Asian contexts, the nature of photography and photographic presentation is remarkably similar in other spatial contexts. Nordstrom's study of the Tupper scrapbooks, specifically the material aspects of it, highlights the distinction between the photographs and the album as an object. Made by hand, the albums are visually unremarkable in comparison to other surviving examples from the time. What is remarkable, however, is the precision and detail with which extended captions have been afforded to each photograph. The presentational form has its own history, its own purpose, its own intent: the way a series of photographs are collected and presented in an album can relate a tale on its own accord. At this point, an important question arises: how do we record the material information within the existing structures that digital technology offer us? To coherently articulate the context and the history of the photograph, we must, first, be able to create the distinction between the indexical and the material at the level of its description. The challenge here is not merely to describe the photograph, but to delineate the separation in a form that is comprehensible to the computer; this separation must happen at the level of the metadata of the photograph.

3. Metadata of the Digitised Photograph

Recent writings on metadata have focused on the more functional aspects of its creation and usage. While examining learning objects and e-print communities of practice, Barton, Currier, and Hey (2004: 5-20) point out the lack of formal investigation of the metadata creation process. While some collection professionals, specifically those working in libraries have written about descriptive practices (Caplan 2003, Haynes 2004, Cole and Foulonneau 2007, Taylor and Joudrey 2008), metadata content standards (Roe 2005, Hillman 2015, Baca 2006) and the use of metadata in digital projects (Kenney and Rieger 2000, Stielow 2003, Hughes 2004) the literature has not yet developed to the point where it affords a complete picture (Park and Tosaka, 2013: 110). A number of existing surveys attempt to provide an overview of metadata practices (Ma 2007, Smith-Yoshimura 2007, Park and Tosaka 2013). Surveys and analyses of the work done by cataloguing and metadata professionals, while very important work in terms of understanding the usage of metadata across institutions, betrays the problems of such approaches. As Ma (2007) points out, the limited number of responses (in comparison to the vast amount of digitised resources), and sometimes,

the veracity of the responses prove problematic in gaining a firm perspective on the issue. The attribution of quality metadata is a process that requires significant time and effort. The results of these surveys are, usually, not very surprising: In Ma's study of the use of metadata schemata, the overwhelming use of Machine-Readable Cataloguing (MARC), a favourite in libraries, followed by Encoded Archival Description (EAD) is what can be expected considering digitised verbal resources far outnumber non-verbal ones. This is confirmed by Smith-Yoshimura (2007: 27-29) and Park and Tosaka (2013: 111) who present remarkably similar results regarding the use of metadata schemata. For digitised resources on the web, Unqualified Dublin Core (DC) is the most popular. If the previous studies had only considered libraries in the United States of America, Palmer, Zavalina, and Mustafoff's study (2007) attempts to widen the research lens to include both research and non-research institutions. They note an overwhelming presence of Dublin Core, with almost half or more of the digital collections using it alone or in combination with other schemata. The criteria in selecting specific metadata schemata are derived from collection-specific considerations of the type of resources, the nature of the collections, and the needs of primary users and communities. Existing technological infrastructure, encompassing digital collection or asset management software, archival management software, institutional repository software, integrated library systems, and union catalogues also greatly affect the selection process. The skills and the knowledge of metadata professionals and the expertise of staff also are significant factors in understanding current practices in the use of metadata schemata and controlled vocabularies for subject access across digital repositories and collections (Ma 2007: 21-28; Park and Tosaka 2013: 113).

Surveys carried out mostly within archival institutions and digital libraries do not contend with the very large number of digital collections produced by individuals who are untrained in preservation methods. Melissa Terras (2010) refers to the growing trend in the creation of amateur online museums, archives, and collections, as an example of how individual endeavour may influence traditional memory institutions in creating useful, interesting repositories. These collections, perhaps not always built along the best practices guidelines, still provide a valuable resource when inspecting our cultural heritage.¹¹ The digitisation of private collections are increasingly becoming more popular. The experience of metadata lies not with the creator, but with the user of the digital resource. Considerations on metadata usage, then, needs also to be made from the perspective of the user who may not be well versed with the protocols

¹¹ With the advancement of technology, it has become easy to commit analog material to the digital space — specifically private collections. These, when committed to the digital space, might not necessarily adhere to the guidelines set by institutions, but provide resources that can be explored for both scholarly and private pursuits.

and vocabularies practiced within the institutions. Kathleen Fear's study, an excellent investigation of metadata from the user's perspective, discusses the value of Dublin Core in digitised image collections, in an attempt to understand the use of metadata. The study, conducted through surveys, focus groups, and search and usability testing, tries to identify the nature of information that non-expert users rely on when searching for images and to locate the vocabularies that best express that information. The study of metadata then has to contend not only with the semantics of expression (with regards to interoperability), but also with a certain lucidity and familiarity that may engage the user (Fear, 2010).

Visual resources are inherently different from textual ones. Their needs, the modes of their description, and the way they need to be conceptualised require a different standard to the long established (and very popular) MARC format. To describe the complexity of visual resources Categories for the Description of Works of Art (CDWA) affords a comprehensive guideline and framework. It has 532 categories and subcategories¹² to include every possible aspect of a visual object. The exhaustive nature of CDWA requires a careful identification of what needs to be described. In other words, collections built using CDWA display the intent of the curator at the very outset. The inherent difference between visual objects and printed verbal resources is that the data related to visual objects, like photographs, is notoriously difficult to articulate and often contradictory. A work of art, like a painting, might have been attributed to one painter at a certain point in time, and another at a later date; it may have multiple titles or multiple dates assigned to it. CDWA, built specifically for art objects, was designed to accommodate these issues (Baca 2001: 3).

The curation of historical photographs proves a difficult challenge in this respect. In a pragmatic or even manageable sense, one might think of the metadata sheet as the space for setting out the physical and verifiable properties of a digital file. This is certainly so; but the lack of accompanying information about photographs acquired from private collections continually presents difficulties. One, particularly, is with the date of recordings: even if some collectors were assiduous providing details of dates and places, others were content with the slightest of notings on the physical carriers. Of course, in their original location, it was almost invariably the collector's memory that would supply a whole range of details, but in the passage from collection to digital repository, the lack of recorded information results in irretrievable gaps. The problem is compounded when we consider that some of these markings on physical carriers might have been misremembered or misattributed: questions on the veracity

¹² For the enthusiast (as opposed to the expert) there exists a CDWA Lite schema with 19 categories (there are further sub-categories for these) to describe a work of art ('CDWA Lite: Specification for an XML Schema for Contributing Records via the OAI Harvesting Protocol').

of information stored within metadata is substantial. Knowledge of the object, then, becomes paramount in our attempts to create this envelope of tags and markers that we call metadata. The process of curation also has to grapple with the circulation of the historical photograph. The same photograph may appear as an image in the newspaper, or it may also be used as the symbol of a political movement, or, perhaps, it is displayed on the walls of an art gallery. The social life of the historical photograph, then, is determined by the circulation of the image. The use of the photograph — its ideological presentational form — determines its context, and in turn renders new layers of meaning to the same image. How are we, as curators of digital collections, to gain an understanding of the object if its nature is continuously changing? The little that we can articulate about the photographic image is derived from an understanding of the history of the object. How we proceed to classify the image, place it within numerous other photographs is dependent on how we recognise the photographer, the period, the photographic plate, photographic process and the photographic context. The history of the object is vital in our attempts to place it within the structures of a digital archive. A problem that is persistent in issues of digital curation of photographs is related to the provenance of the caption. For the archival image, the caption may be attributed by the curator at the time it is committed to the digital archive or it may have been inserted at some intermediary point in its history. This intermediary caption is of archival value as it may locate the image within the larger discourse of its history; the provenance of the caption determines its archival significance. If the caption, then, is of archival value does it merit a separate level of curation?

The manner in which we write the metadata — the vocabulary — conveys our knowledge of the object as series of facts: the name of the photographer, the date when the photograph was taken etc. The tradition of equating knowledge with facts has exists from a philosophical and scientific perspective that can be dated as far back as Aristotle. This view was augmented through the renaissance and enlightenment in order to systemise knowledge. The expression of knowledge becomes fundamental in our attempts to understand and locate the digital object. Traditionally, the efforts to represent knowledge were largely seen as an attempt to manage collections of facts relating to the physical world. The contemporary interest in ontologies can be seen to originate within this tradition and can be taken as an extension of this monolithic view of knowledge. This view on knowledge has been argued over the centuries: Bacon and Locke can be seen to consider knowledge as a single system of beliefs to which new concepts are added. This view would be challenged by Quine who would consider knowledge to be like a 'field of force', which impinged on experience only along the edges (Quine 1963: 42). However, what can be agreed on, is that a body of formally represented knowledge is based on a conceptualisation: the objects, concepts, and other entities that are assumed to exist in some area of interest and the relationships that hold among them (Genesereth & Nilsson 1987: 9). A conceptualisation is an

abstract, simplified view of the world that we wish to represent for some purpose. Every knowledge base, knowledge-based system, or knowledge-level agent is committed to some conceptualisation, explicitly or implicitly. We use common ontologies to describe ontological commitments for a set of agents so that they can communicate about a domain of discourse without necessarily operating on a globally-shared theory. Pragmatically, a common ontology defines the vocabulary with which queries and assertions are exchanged among agents. Ontological commitments are agreements to use the shared vocabulary in a coherent and consistent manner. The agents sharing a vocabulary need not share a knowledge base; each knows things the other does not, and an agent that commits to an ontology is not required to answer all queries that can be formulated in the shared vocabulary. Thus, we can assume that 'concepts' are the key building blocks and that we manipulate these concepts with words. Ontologies are dependent on human language to represent the world. It is here that we face the first and perhaps the most significant challenge in order to achieve a shared understanding of the humanities.

With reference to ontologies, at the very outset, we are faced with two distinct issues – a problem of metaphysics and a problem of semiotics. The philosophical investigation of ontology seeks to find the necessary building blocks of the world, their properties and their inter-relationships. A starting point could be found in Brentano's notion of intentionality and 'objects of consciousness' (Brentano 1973: 127-128). An ontology must make clear what the nature, necessary conditions, and properties of these objects could be. Formal ontologies combine this goal with a use of logic that is intended to ensure rigour and axiomatizability of postulated results.

The general programme of ontology relies on it being possible to uncover properties that could not fail to be as they are for the world to be as it is. Existing ontologies have been concerned with the organisation and the structuring of human knowledge of reality rather than with reality itself. However, to engage with an ontology at a level deeper than this – with specific focus on the conceptual framework – it needs to be epistemologically adequate. Some form of accepted constraints on modelling decisions agreement over conceptual ontology construction is required. The main issue with creating these constraints is, of course, in defining the required ontological level. Since this level has to include accounts of basic objects and basic relations independently of our knowledge of them, it is necessary for the account to define how such objects and relations may be put together in order to reveal an understanding of the world. As argued by Heidegger (Heidegger, 1962 1927), and later by Schutz (1966: 82) Wittgenstein (Wittgenstein, 1953), and others, the world of human being is essentially committed and inter-subjective. That is, the world which human beings have access to is already organised ontologically in inter-subjective terms of human interest. Creating a committed view of the world from a 'God's eye-view' neutral perspective of necessity appears to be extremely difficult.

The semiotic problem (Bateman 1993: 5) is derived from a non-theoretical understanding of language that hinders an appropriate construction of ontologies. The underlying conception of language is that it places an emphasis on the world as a source of its decisions concerning ontology construction without a prior analysis of what is meant by the world. It compounds the problem by driving attention away from natural language as it is inadequate and restricted. The relationship between a sign and its meaning is only arbitrary for the most trivial of possible sign-types – that between linguistic form and phonetic substance. A semiotically richer view can capture the fact that more complex signs are strongly and non-arbitrarily related to their social purpose (Hodge and Kress 1988: 82).

With these considerations, we can posit that ontology or knowledge representation is a surrogate standing for the objects and relations outside in the world. The fidelity of the representation depends on what the ontology captures from the real thing and what it omits. Perfect fidelity is impossible. A simplistic view would say that an ontology is a model of the world which can be used to reason about it. One of the major claims made in favour of ontologies is that can facilitate the interchange of knowledge between agents, or the reuse in different systems. However, if each ontology is modelled around an imperfect universe, knowledge sharing would increase or compound errors which were not visible in the initial use of the ontology. Again, an ontology is a set of ontological commitments. The choice of ontology is also a 'decision about how and what to see in the world' (Davis et al., 1993). This is unavoidable when we consider that representations are imperfect; however, at the same time, the purpose-built ontology has its advantages as it focuses on what is relevant or interesting within the boundaries of the domain. These choices allow us to cope with the overwhelming complexity and detail of the world. Consequently, the content of the representation provides a particular perspective on the world. The way a knowledge representation is conceived reflects a particular insight or understanding in human reasoning. The selection of any of the available representational technologies commits one to the fundamental views on the nature of intelligent reasoning and consequently very different goals and definitions of successes. An ontology must allow for computational processing, and consequently issues of computational efficiency will inevitably arise. Since all ontologies depend on a propositional view of knowledge in order to begin to be computationally tractable, already a very restricted view of what it is possible to represent has arisen. All forms of knowledge representation including ontologies are both media of expression for human beings and ways for us to communicate with machines in order to tell them about the world.

The criticism levelled at ontologies focuses on the fact that they are unsuited to the world of applications once they get beyond a certain level of complexity. While some ontologies are acceptable there is always a trade-off between expressivity,

usability and accuracy. Further arguments can be made (on a more pragmatic level) about the difficulty of maintaining ontologies and reify a particular point of view of the domain knowledge. Ontologies can be seen to be struggling to keep pace with the dynamic, complex world of knowledge bodies and knowledge-sharing. One of the most basic issues facing the users and developers of ontologies is its degree of complexity. Folksonomies are comparatively easier to use and maintain while offering a flexible and personalised perspective; however, their use is limited due to two reasons – (a) their quality of concepts involved does not match that of ontologies and (b) their reliability cannot be compared to that of an ontology. On the other hand, formal ontologies, such as the Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE), the General Formal Ontology (GFO), the Web Ontology Language (OWL), or the Resource Description Framework (RDF), require specialised knowledge to build and use them, and are more challenging to maintain. They are also more rigid than the ubiquitous folksonomies and thesauri, and less adaptable to changing applications and user perspectives (Brewster et al. 2007: 563-568).

The realisation of the obvious technical challenge of creating both human readable and machine readable data is heightened when we consider that over the last few years there has been a steady increase in localised digitisation projects by individuals often unaffiliated to archival and preservation institutions. The digitisation and curation of collections outside the walls of the archival institution bring with it its own set of challenges. Projects carried out without specific funding and done with consumer-level instruments, are often discouraged by the archival community. Katrina Dean (2014: 172), in a defence of traditional archival practices, contends that the practical problems of creating adequate descriptives for digital objects for the purposes of creating digital repositories are intensified by the digital economy. She argues that 'shifting content from public domains to commercial ownership, either through the domination of commercially generated content or the commercial ownership of and exploitation of third party and user-generated content' reflects poorly on the value of information and heightens its fragility. The act of publishing archival documents online, thus bringing them out of a controlled and moderated public sphere and into a digital economy brings with it issues of intellectual property, and consequently ownership of cultural property in general. The level of descriptive, technical, and administrative metadata required to manage, preserve, and discover digitised collections generally exceeds the level of metadata required to manage and make accessible their physical counterparts. Increased standards of explicit evidential value, and compliance required to bring archival collections online merely reflect on the inadequacies of meeting those standards in the traditional archive, in the first place. Traditional archives, for Dean, are about relationships; for their evidence and informational value to be fully explored, the objects must reveal relationships between contexts and records, and

among sources. While in the traditional archive, these relationships were implicit in the architecture of the buildings, in the storage configurations, within the old registers and the administrative files, and in the curatorial knowledge, its digital counterpart is still searching for those connections. 'Short of digitising whole collections and transposing these contexts into metadata, it seems unlikely that collections in their present configurations will be transmitted into the future knowledge economy' (172). While Dean speaks about digital repositories created by institutions and by individuals in the same breadth, there is a difference: the value of archival objects digital and curated to preservation standards lies in the expectation that they would stand the test of time. Projects done in individual capacities, mostly out of interest or curiosity, have a different approach and a different value. These projects accumulate private collections, objects that would otherwise have not entered the archival institutions. Most of these are freely available and resources that may be explored for research purposes. These collections often do not meet the required standards of preservation, or the quality that an institution would demand: their value lies singularly in their availability.

The creation of digital collections requires more than the availability of scanners, cameras, and a knowledge of metadata schemata. The primary problem of creating digital resources, specifically with historical artefacts, lies in the process of their curation. Beyond the technical challenges of metadata attribution, curatorial expertise is required to foster some understanding of the digital object. This is faced severely in the case of historical objects, specifically non-textual objects, which require a degree of curatorial more expertise to interpret. The ubiquitous nature of digital objects has made metadata a primary concern for all engaged in work in the digital space. To aid identification of the object, the machine requires textual markers; this become more necessary for non-textual objects where the machine cannot adequately look through the content of the object. The framework of metadata is vital to the creation of digital repositories. Beyond search and retrieval, metadata guides the conceptual understanding of an object within the collection.

Bibliography

- Appadurai, Arjun. *The Social Life of Things*. Cambridge: Cambridge University Press, 1986. Print.
- Baca, Murtha, et al. *Cataloging Cultural Objects: A Guide to Describing Cultural Works and Their Images*. Chicago: American Library Association, 2006. Print.
- Baca, Murtha. 'A Picture Is Worth a Thousand Words: Metadata for Art Objects and Their Visual Surrogates.' *Cataloging the Web: Metadata, AACR, and Marc21*. Ed. Wayne Jones, Judith R. Ahronheim, and Josephine Crawford. 2001.

- Barthes, Roland. *Camera Lucida: Reflections on Photography*. New York: Farrar, Straus, and Giroux. 1981. Print.
- Batchen, G. *Photography's Objects*. Albuquerque: University of New Mexico Art Museum, 1997. Print.
- Bateman, J. A. 'Ontology Construction and Natural Language.' *Proceedings of the International Workshop on Formal Ontology*. 1993. May 2012. PDF file.
- Brentano, F. *Psychology from an Empirical Standpoint*. Trans. A. C. Rancurello, D. B. Terrell and L.L. McAlister. London: Routledge & Kegan Paul. 1973. Print.
- Brewster, C. and K. O'Hara (2007). 'Knowledge representation with ontologies: Present challenges—Future possibilities', *International Journal of Human-Computer Studies* 65, Elsevier. 4 May 2012. PDF file.
- Caplan, Priscilla. *Metadata Fundamentals for All Librarians*. Chicago: American Library Association. 2003. Print.
- Carole Palmer, Oksana Zavalina, and Megan Mustafoff, 'Trends in Metadata Practices: A Longitudinal Study of Collection Federation.' *Seventh ACM/ IEES-CS Joint Conference on Digital Libraries, Vancouver, British Columbia, Canada, June 18–23, 2007*. <http://hdl.handle.net/2142/8984>. Web. Jan. 24, 2015.
- Cole, Timothy W. and Muriel Foulonneau. *Using the Open Archives Initiative Protocol for Metadata Harvesting*. Westport, CT: Libraries Unlimited, 2007. Web. 21 February 2014.
- Davis, R., H. Shrobe, and P. Szolovits. 'What is a knowledge Representation'. *AI Magazine* 14.1. 1993.
- Dean, Katrina. 'Digitizing the Modern Archive.' *Archives and Manuscripts*. 2014. PDF file.
- Edwards, Elizabeth and Janice Hart. 'Photographs as Objects'. *Photographs Objects Histories*. Ed. Elizabeth Edwards and Janice Hart. London and New York: Routledge, 2004. Kindle file.
- Edwards, Elizabeth. 'Material beings: objecthood and ethnographic photographs.' *Visual Studies* 17.1: (2002) 67-75. PDF file.
- Fear, Kathleen. 'User Understanding of Metadata in Digital Image Collections: Or, What Exactly Do You Mean by 'Coverage'?.' *The American Archivist* 26:(2010). JSTOR Journals. Web. 24 Jan. 2015.
- Fineman, Mia. 'Kodak and the Rise of Amateur Photography.' *Heilbrunn Timeline of Art History*. New York: The Metropolitan Museum of Art, 2000. Web. 21 November 2013.
- Genesereth, M. R. and N. Nilsson. Mateo, CA: Morgan Kaufmann Publishers: San. 1987. Print.
- Harris, Brent. 'Photography in Colonial Discourse: The Making of "the Other" in Southern Africa.' *The Colonising Camera: Photographs in the Making of Namibian History*, 20-24. Ed. Wolfram Hartmann, Jeremy Silvester, and Patricia Hayes. South Africa: University of Cape Town Press Ltd., 1998.

- Haynes, David. *Metadata for Information Management and Retrieval*. London: Library Association Publishing, 2004.
- Heidegger, M. *Being and Time*. Trans. J. Macquarrie and E. Robinson, Oxford: Basil Blackwell. 1962.
- Hillmann, Diane I. *Using Dublin Core. Dublin Core*. Web. 20 September 2015.
- Hodge, R. and G. Kress *Social Semiotics*. Cambridge: Polity Press. 2004. Print.
- Hughes, Lorna M. *Digitizing Collections: Strategic Issues for the Information Manager*. London: Facet, 2004. Print.
- Jähne, Berndt. 'Pixels.' *Digital Image Processing: Concepts, Algorithms, and Scientific Applications*. 77-99. Berlin, Heidelberg: Springer. 1995.
- Jane Barton, Sarah Currier, and Jessie M. N. Hey, 'Building Quality Assurance into Metadata Creation: An Analysis Based on the Learning Objects and E-Prints Communities of Practice' *Dublin Core Conference: Supporting Communities of Discourse and Practice—Metadata Research & Applications*. Seattle, Washington: eprints in library and information science. 2003.
- Kenney, Anne R., and Oya Y. Rieger. *Moving Theory into Practice: Digital Imaging for Libraries and Archives*. Mountain View, CA: Research Libraries Group, 2000. Print.
- Langford, Martha. *Suspended Conversations: The Afterlife of Memory in Photographic Albums*. Montreal: McGill University Press. 2001.
- Legêne, Susan. 'Photographic playing cards and the colonial metaphor.' *Photographs Objects Histories*. Ed. Elizabeth Edwards and Janice Hart. London and New York: Routledge, 2004. Kindle file.
- Ma, Jin. 'Metadata.' *SPEC Kit 298*. Washington, DC: Association of Research Libraries, 2007. Print.
- Miller, Daniel. *Material Culture and Mass Consumption*. Oxford: Basil Blackwell, 1987. Print.
- Nordstrom, Alison. 'Making a Journey: The Tupper Scrapbooks and the Travel they Describe.' *Photographs Objects Histories*. Ed. Elizabeth Edwards and Janice Hart. London and New York: Routledge, 2004. Kindle file.
- Park, Jung-ran, and Yuji Tosaka. 'Metadata creation practices in digital repositories and collections: Schemata, selection criteria, and interoperability.' *Information Technology and Libraries* 29.3 (2013): 104-116.
- Roe, Kathleen D. *Arranging and Describing Archives and Manuscripts*. Chicago: Society of American Archivists, 2005. Print.
- Rose, Gillian. *Doing Family Photography: The Domestic, The Public and The Politics of Sentiment*. Kent, UK: Ashgate. 2010. Print.
- Sassoon, Joanna. 'Photographic materiality in the age of digital reproduction.' *Photographs Objects Histories*. Ed. Elizabeth Edwards and Janice Hart. London and New York: Routledge, 2004. Kindle file.

- Schutz, A. 'The problem of intersubjectivity in Husserl'. *Collected Papers III: Studies in Phenomenological Philosophy*. Ed. Ilse Schutz. The Hague: Nijhoff. 1966. Print.
- Schwartz, Joanne M. "We make our tools and our tools make us": lessons for photographs from the practice, politics and poetics of diplomatics.' *Archivaria*, 40: (1996). 40-74. Print.
- Smith-Yoshimura, Karen. Rlg Programs Descriptive Metadata Practices Survey Results. Dublin, Ohio: OCLC Programs and Research, 2007. Internet resource.
- Stielow, Frederick J. *Building Digital Archives, Descriptions, and Displays: A How-To-Do-It Manual for Librarians and Archivists*. New York: Neal-Schuman, 2003.
- Tagg, John. *The Burden of Representation: Essays on Photographies and Histories*. Minneapolis: University of Minnesota Press. 1988. Print.
- Taylor, Arlene G., with Daniel N. Joudrey. *The Organization of Information*. Westport, CT: Libraries Unlimited, 2008. Print.
- Terras, M. 'Digital curiosities: resource creation via amateur digitization'. *Literary and Linguistic Computing* 25.4: (2010). 425-438.
- Vinyals, Oriol, Alexander Toshev, Samy Bengio, and Dumitru Erhan. 'Show and Tell: A Neural Image Caption Generator.' *arXiv*: 1411.4555: (2014). Web. 12 June 2015.
- Willumson, Glenn. 'Alfred Hart: Photographer of the Central Pacific Railroad.' *History of Photography*. Vol. 12, No. 1, 61-75. 1998.
- Wittgenstein, L. *Philosophical Investigations*. Trans. G.E.M. Anscombe. Oxford: Basil Blackwell. 1953.