Aberystwyth University

A tale of two images: the quest to create a story-based image indexing system
Rafferty, Pauline; Albinfalah, Fawaz

A Tale of Two Images: The Quest to Create a Story-based Image Indexing System

Purpose – This conceptual paper considers the possibility of designing a story-based image indexing system based on users' descriptions of images. It reports a pilot study which uses users' descriptions of two images.

Design/methodology/approach – Eight interviews were undertaken to investigate storytelling in user interpretations of the images. Following this, storytelling was explored as an indexing input method. Twenty-six research subjects were asked to create stories about the images, which were then considered in relation to conventional story elements and in relation to Rafferty and Hidderley's 2005 image modality model.

Findings – The results of the semi-structured interviews revealed that the majority of interpretations incorporated story elements related to setting, character, plot, literary devices, and themes. The fifty-two image stories included story elements identified in the first part of the project, and suggested that the image modality model is robust enough to deal with the 'writerly' images used in this study. In addition, using storytelling as an input method encourages the use of verbs and connotative level responses.

Originality/value –User indexing is generally based on paradigmatic approaches to concept analysis and interpretation in the form of tagging; the novelty of the current study is its exploration of syntagmatic approaches to user indexing in the form of story-telling. It is a pilot, proof of concept study, but it is hoped that it might stimulate further interest in syntagmatic approaches to user indexing.

Keywords – Image indexing, Stories in images, Semantic image indexing,

Paper type – Conceptual paper

Introduction

The prospect of designing a story-based image retrieval system has not received attention from scholars in the image retrieval field. Jörgensen (2003) acknowledged the existence of story elements within some images but remarked that scholars have been reluctant to examine this research area because "most image indexers have considered addressing interpretive or abstract elements within an image to be beyond the scope of their job" (p. 251). The process of designing a story-based image indexing system presents a challenge to indexers, not least because humans differ in their educational and cultural backgrounds and in their perceptions, as well as in prior knowledge and beliefs (Berger, 1972, p. 8), thus rendering any attempt to create a consistent system for describing images difficult. Despite this problem, time and retrieval systems have moved on since 2003, and the time might now be right, given the current interest in user-based indexing, the social semantic web, and an emphasis on dialogic approaches to indexing documents, to explore the potential for story-based indexing of non-text documents.

An earlier project reported in 2011 (Rafferty), investigated the relationship between modality and interpretation, where modality refers to a relationship between the sign and the real world, using a model first proposed by Hidderley and Rafferty (2005). Four images were chosen from publicly available websites to represent four categories: high modality contemporary images; high modality historical images; low modality contemporary images; low modality historical images. Distance learning students from the Department of Information Studies at Aberystwyth University were invited to assign descriptive, or denotative tags, and/or associative, or connotative tags to the images. The results suggested that the Rafferty/Hidderley model was a relatively robust categorisation model that might offer possibilities for designing indexing templates or information representation records, however, the high modality images included in that study were very obviously "readerly" images. The current project adds to the earlier work by focusing on two high modality "writerly" images. The terms "readerly" and "writerly" derive from Barthes (1973) who used them to describe different types of texts, underpinning his textual theory with a constructivist view of interpretation and meaning-making. The "readerly" text is one in which the potential space for reader interpretation is tightly controlled by the author, while the "writerly" text is one in which the space for interpretation is often actively sought in an ambiguous text. The current project, which aims to explore story-telling as a basis for image

indexing, tests the tentative suggestions relating to high modality contemporary images in the Rafferty/Hidderley model, which are that:

"A non-textual information object of high modality decoded at the same historical moment, and within the same culture and logonomic system as the encoding moment, would be expected to evoke a limited range of denotational meanings. There may be a broader range of connotational level interpretations. Textual anchorage could help fix the intended encoded meaning, but decoding might evoke negotiated or oppositional interpretations. Interpretation within a different cultural and logonomic system would potentially evoke a broader range of connotational and denotational meanings". (Rafferty and Hidderley, 2005, p125).

This study is partly underpinned by Schmidt and Stock's (2009) investigation into the possibility of creating an image indexing system based on emotions. In the course of actualising the aim, this study critically evaluates the potential to draw on users' interpretations of images in the design of a story-based image indexing system. The concern is with the input side of the user-based indexing process.

Tagging offers both possibilities and drawbacks in relation to information discovery. On the one hand, it opens up indexing to allow for multiple voices, many approaches, efficient updating, modern terminology and language: on the other hand, studies of tagging practice have reported that terms can be random, vague, and uninformative. One way of categorizing tags is to distinguish between private and public tagging, that is, tags are meant for the taggers' benefit alone, and tags that are assigned with a view to helping others to discover relevant information. Most websites that facilitate tagging allow taggers free rein about the terms they include. This project considers whether it would be helpful to structure taggers' input through story-telling. Earlier work using a democratic indexing template designed to capture informational tags at the level of pre-iconographic, iconographic and iconological meaning suggested that a template has its uses for improving the exhaustivity of indexing, however, the template was too rigid and unattractive to indexers (Rafferty and Hidderley, 1998). Story-telling, is a pervasive and generally pleasurable form of human communicative practice. In addition, story-telling offers a syntagmatic approach to user-based indexing input. This project explores syntagmatic input in

the form of the story as a method to facilitate creativity in tagging while at the same time channelling and disciplining it.

Meaning and interpretation

The project, and our approach to the nature of meaning and interpretation, is underpinned by Hodge and Kress's social semiotics, which explores the construction of meaning as a social activity, acknowledging the role of the reader, located in specific material logonomic systems, in "making meaning" (1988). Where Saussure's semiotics focused on langue, rather than parole, on synchrony rather than diachrony, on paradigms rather than syntagms, and signifiers rather than signified (Hodge and Kress, 1986), social semiotics sees the meaning of the text as being formed from the interplay of polysemic signs and a range of reading positions. This means that interpretation at the point of reception does not necessarily equate to the intended meaning of that text at the time of production. Many writers in the field take the view that interpretation in relation to images is complex, particularly in relation to images that do not necessarily have text to anchor meaning, not least because the production of meaning through interpretation may be subjective (e.g. Burke 1999; Enser 1995; Krause 1988; Shatford 1986; Shatford-Layne 1994; Svenonius 1994).

It would be useful at this point to outline the analytical categories drawn from social semiotics that will be used in this paper to explore the operation of meaning and interpretation in image indexing. The sign is the basic unit of analysis for semiotics. The Saussurean sign is composed of the *signified* and the *signifier* which together form the sign. Following Peirce, modern semioticians often refer to a typology of signs which generally includes reference to the *indexical* sign, which is somehow connected with the signifier, for example a cartoon of a politician, the *iconical* sign, which is a sign resembling the signified, and the *symbol*, which is a conventional sign, for example traffic lights or language (Chandler, 2002, p.37).

For Saussure and semioticians who followed, the construction of meaning is governed by operations on two planes, the paradigmatic plane and the syntagmatic plane. A *paradigm* is a set of associated signifiers or signifeds which are members of some defining category, but this is a set in which each member is significantly different from the others (Chandler, 2002, p.80). *Syntagms* are the combinations or chains of elements which form a meaningful whole within a

text, in the broadest sense (Chandler, 2002, p. 81). In text or speech, syntagm might refer to the rules of grammar which operate within human language, while Kress and Van Leeuwen suggest that images operate their own visual grammar based on spatial relations (1996).

Semioticians, following Barthes, often distinguish between *denotation* and *connotation* in relation to the construction of meaning in signs. While images can be interpreted, or decoded (Hall, 2001 [1973]) in a straightforward, or common-sense way (the denotational level in Barthes's terminology), at another level, images are more open to cultural and social interpretation (the connotative level in Barthes's typology). Text can be used to anchor and fix meaning, so that a picture of a child expressing an emotion with the caption "lost her dolls" is different from an image captioned "delighted at finding her doll" (Duncum, 2004, p. 256). In both instances, the words of the caption either reinforce or nullify the meaning of an image, depending on the text–image relationship (Duncum, 2004, p. 256). The importance of context in changing the meaning of an image is supported by Brink, Squire and Bigelow's (2006) findings that users interpret an image differently if it is grouped with other images (p. 716). The role of context in potentially altering the interpretation of an image was expanded by Becker (1998), who argued that the concept of context in relation to the meaning of an image should include the physical place in which an image is situated, and people's discussion around the space of the image (p. 88).

The literature of image retrieval literature sometimes refers the semantic gap which exists in two aspects: firstly, the gap between the machine's ability to decode image data and the translation of these data into objects that can be described, and secondly, the gap between the user's cognitive image interpretation and the actual description of the image itself (Enser, Sandom, Hare & Lewis, 2007, p. 469). Panofsky's iconography theory has been, either extensively cited by image retrieval scholars (e.g., Enser, 2000, p. 201; Svenonius, 1994, p. 602; Shatford, 1986, p. 45), or reintroduced under a different name (e.g., Barthes, 1977, pp. 52–59). However, some scholars have questioned the extent of its applicability to the study of image indexing and retrieval. For example, Burford, Briggs and Eakins (2003) argued that Panofsky's theory was applicable to art history and cannot be generalised to include other disciplines (p. 128).

Shatford's model (1986), which was based on Panofsky's model (1955), reinterpreted the pre-iconographical and iconographical levels as "of" and "about" something, with "of" indicating a "general description of objects and events" such as "people, places, objects, conditions, and actions that have a physical manifestation" (p. 45). The "about" aspect of Shatford's model deals with "symbolic meanings, and abstract concepts that are communicated by images in the picture [such as] love, sorrow…truth, honor" (p. 45). Other scholars (e.g., Jörgensen, Jaimes, Benitez & Chang, 2001, p. 940; Shatford-Layne, 1994, pp. 584–585; Rafferty & Hidderley, 2005, p. 179) have experimented with practical templates for indexing images and have concluded that these models require further development.

Image indexing systems

Eakins (2002) noted that an optimal retrieval system is measured by its capacity to return search results that satisfy users' needs, a principle he acknowledged is difficult to apply to images, unless an understanding of their contents has been established (pp. 5–6). Despite Eakins's (2002) assertion of the difficulty of classifying images, scholars have suggested two possible systems for image retrieval, namely content-based and concept-based image retrieval systems.

Although content-based systems can enhance the precision of retrieving images (Lew, Sebe, Djeraba & Jain, 2006, p. 1; Zhang, Wenyin & Hu, 2003, p. 339), they have limitations which include: the gap between the visual system of humans and that of machines (Deb & Zhang, 2004, p. 59), the inability of some users to draw their image queries (Choi & Rasmussen, 2003, p. 499), the machine's inability to identify complex image elements (Eakins, 2002, p. 5) and the difficulty of retrieving abstract concepts such as emotions from an image (Burford, Briggs & Eakins, 2003, p. 148).

Concept-based systems which describe image content using text (Choi, Cho, Park & Kim, 2003, p. 79) are sometimes advanced as ideal for conveying abstract levels in an image (Eakins and Graham, 1999; Rorissa, 2008). However, Olson and Wolfram (2006) argued that, because concept-based systems depend on textual annotation, they lack indexing consistency and hence diminish the ability to unify image description. In addition, these systems suffer from the inability of words to capture the true meaning of an image (Enser, 2000), thus making the system unlikely to return search results that are precise enough to satisfy users.

Approaches to the indexing of images

The literature suggests three general indexing approaches. The first approach depends on an expert performing the indexing by consulting a system, often based on some form of controlled vocabulary, designed to enhance the recall and precision of search results (Rowley & Hartley, 2008). Lancaster (as cited in Chowdhury, 2010) favoured the expert-led indexing approach, on the basis of its consistency and the agreement on indexing terms that is likely to result among indexers (p. 156) however, others argue that it is "expensive and time consuming" and "single minded" (Rafferty & Hidderley, 2007, p. 399). The second approach is one based on authors' interpretation of their work. In a study that tested the authors' capacity to assign rich metadata to their work, Greenberg et al (2001) found that authors were capable of assigning good terms to their work (p. 43). The current project takes a constructivist view of meaning as being produced in and through the interaction of text and reader. Within this model of interpretation, following Hall's encoding-decoding model, it might be that personal, private meanings attached by readers or viewers can transform the document from the "preferred" meaning that might be inscribed within it by the author (Hall, 2001 [1973]) to other possible meanings.

The third indexing approach is user-based tagging, which has become popular as an outgrowth of the Internet. Tagging might be merged with controlled vocabularies for enriched indexing (Chopin, 2008). Moreover, Merholz (2004) hypothesised that, over time, tagging might lead to desire-lines which others would use for navigation. However, this theory assumes that users will continue to add meaningful terms to the indexing language, a position refuted by others who argue that, when the number of people in a group increases, the process "might lead to the organic production, not of a dominant controlled vocabulary, but of many splintered controlled vocabularies" (Rafferty & Hidderley, 2007, p. 401). In addition, Nesta and Mi (2011) have expressed concerns that most users repeat keywords, or provide ones that have limited value for retrieval purposes, while McDonnell and Shiri (2011) cautioned that tagging lacks the organised structures of controlled vocabularies. In short, the literature is divided as to whether or not tagging offers a potential solution, and this debate is expected to persist because of the dynamic nature of the online environment.

Users and stories of images

Greisdorf and O'Connor (2002b) argued that users can provide rich descriptions of images exceeding what is available in controlled vocabularies systems (p. 383). Such descriptions might enhance indexing exhaustivity and inform indexers' understanding of users' seeking behaviour (O'Connor, O'Connor & Abbas, 1999, p. 682), thereby leading to a narrowing of the semantic gap between the indexer and the user. The importance of user-generated description of images has been documented by Enser, Sandom, Hare and Lewis (2007), who observed that "a richly-descriptive caption" is essential for retrieving images that meet users' needs (p. 477). This view is of concern to Jörgensen (2003) who takes the view that users may lack the expertise to assign appropriate terms to images for the purposes of indexing and subsequent retrieval (p. 256).

O'Connor, O'Connor and Abbas (1999) reported that users employ stories to describe the content of images (p. 684) and tend to use a narrative style for their descriptions as they become accustomed to the viewing experience offered by an image (pp. 687–688). The possibility of using these stories in image indexing has not been considered by scholars because of the "lack of a widely accepted conceptual framework within which to make indexing decisions" (Jörgensen, 2003, p. 252) among experts. This study takes a first step towards exploring the potential of using narratives to design a story-based image indexing system.

An overview of stories

For Barthes (1970), narrative can be divided into story and discourse, where story refers to the "chronological account of the observations and experiences being communicated" (Brophy, 2009, p.35), while discourse refers to the selection of what is told, its genre and its style. Brophy's model, which draws on Barthes and on Riceour's framework, amongst others, is used to underpin this study's understanding of narrative structure. It seems an appropriate model as it derives from the information science domain. Brophy outlines eight elements that might be considered key to understanding narrative structure. These are: setting, causality, plot, style, character, point of view, narrative presence, and hermeneutic and proairetic codes. In addition, Ingermanson & Economy (2010) suggest including literary devices and themes. For Monteleone (2004), the setting of a story includes five key components: place, time, weather, social conditions, for example social norms, mannerisms and social structure, and mood, meaning the "tone and the style" of the story (Monteleone, 2004, p. 95).

Following Propp, the plot of a story is generally seen as composed of six elements: the protagonist, the antagonist, the protagonist's goal, conflict, suspense, and resolution. Plots lead to three possibilities: "protagonist wins, protagonist loses, antagonist wins" (Monteleone, 2004, pp. 71–72). Characters in stories can be either "fully developed and lifelike" or flat in that they represent only "minor action" in the story (Ingermanson & Economy, 2010, p. 41).

Brophy's examples of style, by which he means the way in which the narrative is presented, include register, for example using colloquial language rather than formal language. This element relates to Barthes's notion of discourse. The point of view refers to the narrator(s) standpoint, focus or angle, while narrative presence Brophy explains as referring to "the listener's relationship to the story, which in the extreme can involve feelings of being transported into the action, of being part of the narrative" (p46). The notion of hermeneutic and proairetic codes derives from Barthes, who used them to describe the two different ways in which suspense can be created. The hermeneutic code refers to those elements of the plot that raise unanswered questions, while the proairetic code refers to the actions which occur and which will have consequences. Authors use a combination of similes and metaphors as literary devices in a story. The theme is the "deep meaning" of a story (Ingermanson & Economy, 2010, p. 203) and can include messages such as, good always triumphs over evil, love endures in the end, truth prevails, or justice exists in the world. Some of these elements will have more value in terms of developing indexing templates than others.

Methodology

This study is in two parts. The first part used interviews to explore users' interpretations of two "writerly", high modality images, and to establish the existence of story-telling elements in participants' descriptions. Following this, 26 MSc distance learner students from Aberystwyth University were invited to write stories about the images with a view to testing the Rafferty-Hidderley model in relation to "writerly" images.

In the first part of the study, a purposive theoretical sampling technique targeted information rich cases. Participants had degree qualifications in a range of disciplines, permitting the emergence of interpretations. The choice of sampling technique had limitations, as the findings reflect the chosen population, however, the sampling served the purpose of discovering the variety of

interpretations of images from the perspectives of a select group of users. Eight participants were interviewed. Five participants were based in Aberystwyth and three in the Kingdom of Bahrain. All participants had a high level of familiarity with English.

Only two images were selected, partly because this study is a preliminary one. The images provide "writerly" text in that they invite the viewer to become "no longer a consumer, but a producer of the text" (Barthes, 1970, p. 4), while the choice of images by one artist ensured a consistent style of photography to which the viewer could relate. A pilot study of randomly selected users (whose views are not presented in this study) suggested that users could describe the images in a variety of ways. This result provided encouragement to continue carrying out the study.

Interviews explored users' interpretations of images, providing "rich insights" into individuals' inner worlds (May, 2001, p. 120), and identifying individuals' special understandings, thus revealing new ways of discovering "the world from the subjects' points of view" (Kvale, 1996, p. 1), while offering the potential to obtain in-depth information from participants. On the debit side, it is acknowledged that participants may resist sharing information with the interviewer, have a difficult time understanding the questions posed, or provide short answers that do not satisfy the researcher (Marshall & Rossman, 1995, p. 81).

Interviews were recorded using note-taking, and face-to-face interviews were combined with online synchronous audio interviews via Google Talk. In the synchronous online audio interviews, images linked on Flickr were sent to participants at the beginning of each interview. One source of concern was that these images had tags and descriptions assigned to them by the images' owner. This problem was addressed by sending a link that contained a large view of the images without the tags and descriptions, thus ensuring that participants' interpretations of the images would not be influenced by the images' captions. For the sake of transparency in this study, the owner's descriptions of these images are provided in the appendix.

During the interviews, participants were invited to describe these two images using keywords or sentences, and encouraged to provide such additional details in relation to these images as they might see proper. Participants were allowed to express themselves freely, without any attempt to

influence their answers. The researcher cross-checked participants' answers with them, in order to verify the accuracy with which their responses were reported.

Descriptive coding was used in data analysis (Saldaña, 2009) (see Table 1 for an example of descriptive coding for Image 1).

Table 1. Descriptive coding example of Figure 1

Description	Coding
"a [1]middle aged [2]relaxed man sitting in a sealed glass [3]jar looking into [4]dark empty space outside the [5]confined space in which he sits" (Participant 5).	[1]physical attributes [2]psychological attributes [3]location [4]mood [5]location

Descriptive codes in Table 1 led to the creation of categories. For example, *physical* and *psychological attributes* subcategories were organised under the category *character*, while the subcategories *location* and *mood* were arranged under the category *setting*. The same coding approach was applied to the remaining participants' descriptions of Figures 1 and 2, leading to the emergence of story elements.

For the second part of the study, distance learner MSc students from DIS, Aberystwyth University were invited to write story interpretations of the two images. Twenty six students participated, of which nineteen were women and seven were men. Of the women, eight were in the 18-30 category, seven in the 30-40 category, three between 40-50 and one over 50. Of the men, two were in the 18-30 category and four were in the 30-40 category. One respondent did not include personal information. Of the fifteen women who specified their first degrees, eleven had arts degrees, one holding a PhD, three had social science degrees and one held a degree in theoretical physics. Three of the men were arts graduates, and one was a social scientist. This activity was undertaken during an introduction to the Information Organisation and Retrieval module at a Distance Learning School. The students were given free rein to record their stories

and responses. Results were analysed in relation to the story element categories and the Rafferty-Hidderley denotation/connotation model.

Results

Overall, the results of the interview stage indicated the existence of story elements in the participants' descriptions. Data analysis allowed for the development of relevant categories, such as "setting", "character", and "plot", within which the construction of the narratives could be understood.

Story elements in Figure 1

Eight participants described the setting of Figure 1 as a location. Three specifically used the term "jar", whereas two participants used the word "bottle". Two participants referred to the location as "confined in space", with only one perceiving it as a "comfortable place". Finally, two participants described the mood of the setting as "light", while one participant indicated that it was "dark".

References to story elements in participants' descriptions of Figure 1

Story element	Number of references	Number of agreements
Setting		
Location	8	7
Mood	3	2
Character		
Physical attributes	2	
Psychological attributes	5	2
Plot		
Protagonist	2	2

Table 2: Seven Story elements in Image 1

Antagonist	2	2
Goal	4	2
Conflict	6	5
Suspense	4	2
Literary devices		
Similes	2	
Metaphors	1	
Themes	8	5

participants described the character's physical and psychological attributes. He was describes as "barefoot" and "middle aged". Five participants specified psychological aspects of the character, describing him as "calm", "afraid", "frustrated", "relaxed" (2).

In relation to plot elements, two interviewees identified the protagonist as a "man", while the rest used the term "he" to refer to the character. The goal of the protagonist was specified by four participants, as "planning" (2) and "isolating" (2). Two participants agreed that the protagonist's goal was to "escape" although they used different words to express it. In terms of conflict, six participants perceived it as consisting of two parts: the protagonist against himself and the protagonist against an outside force. Of these six, one indicated that the protagonist's conflict was internal and accordingly interpreted it as loss of expression. Five participants used sentences expressing a belief that the conflict was external; of these, three agreed that the protagonist was imprisoned by choice, whereas two agreed that he was imprisoned by force. Finally, the suspense element in the descriptions was provided by four participants, the protagonist was "looking at something" (3), one indicating that he was "looking at the light".

Three participants used literary devices in their interpretations of Figure 1, with no commonalities existing among these interpretations. Similes were used in describing the protagonist as a "ballet dancer, a choreographer" and in perceiving the "jar as a prison". Only one metaphor was found in the participants' descriptions, namely, the view of Figure 1 as a 'prison'. Three emphasised the main theme of imprisonment by choice, while two agreed on the theme of imprisonment by force. The remaining three indicated that it represented an "artistic act", an "emotional burden", and mainly a "man sitting in a jar".

Story elements in Figure 2

Two participants referred to the location in Figure 2 as a "dried lake" and a "desert". A new category, *weather*, emerged with respect to the setting, four of the participants using the term "clouds" or "cloudy" to describe it. Although four participants identified the mood in Figure 2, no two of them agreed on a description, individual suggestions being that it was "worse than the desert", a "deserted troubled place", a "spoiled landscape", and finally that it was "apocalyptic".

References to story elements in participants' descriptions of Figure 2		
Story element	**Number of references**	**Number of agreements**
Setting		
Location	2	
Weather	4	4
Mood	4	
Character		
Physical attributes	2	
Psychological attributes	5	2
Plot		
Protagonist	2	2
Antagonist	2	2
Goal	4	
Conflict	4	
Suspense	4	2
Literary devices		
Similes	1	
Metaphors	3	

Table 3: Story elements in Image 2

Two participants used the terms "barefoot" and "young" to describe the physical features of the character in Figure 2. Two agreed on the term "given up", while the remaining three used the terms "hopelessness", "peace", and "hope" respectively.

Two participants acknowledged the protagonist in Figure 2 as a "man" and two participants specified the antagonist as "God". Four participants indicated the protagonist's goal, although they differed as to its nature, using descriptions such as "waiting for something", "finding a real love", "wants to be lifted up" and "looking for someone to save him". Four participants mentioned conflict, while disagreeing on its causes. Finally, four participants noted elements of suspense in their narrations, with two agreeing that the man was "dropping the flower" and the others suggesting respectively that the man was "dropping everything" or had "lost his wife".

Four participants used literary devices in their descriptions. Of these, three were metaphors including "life is a desert", "winds of change" and "apocalyptic". The remaining participants used a simile in describing Figure 2, commenting that "it is like a chess board". None of the eight participants agreed on a common theme. The themes that emerged were those of "climate change", "seeking answers", "awaiting love", "contradictions", "loss of a loved one", "apocalypse now", "awaiting to be lifted up" and "blaming God".

Having established the existence of story telling elements, the second part of the study explicitly invited students to write stories about the two images which were then considered in relation to the Hidderley-Rafferty model. In Image 1, there was a relatively narrow range of denotational responses. Respondents referred to the "man" (12), or to "he" (4) or to the "individual" (1). The man was in a "jar" (16) or "bell jar" (1), "glass jar" (1) or "jamjar" (1). Physical aspects of the man were noted: he was sitting cross-legged (2), with knees-up (1), and hands clasped (2), while one respondent described him in some detail as being middle-aged, little hair, dark smart shoes. He was sitting on a dark, flat surface (1). The screwtop lid was referred to four times, while there

were three references to the image as an image: it was a black and white photograph with a clear source of light and was artistically lit. There was only one mention of colour (blue).

At the level of connotation, there was a broader range of interpretations, specifically the image was interpreted as being either a negative and rather frightening image, or in a more positive way. Four respondents interpreted the image as "bottled up emotion" with the man being "miniatured" (1) in the jar. There were nineteen references to the notion of "fear" or "being trapped". The goal was sometimes escape (5) while in one case it was diagnosis. Conflict took the form of fear of the unknown in some way (4) or inner turmoil (6), within which there was reference to mental illness and depression. Suspense was important in these stories: the man was waiting (6), having seen something horrible (2), or being looked at by others (2), and pleading to be released (2). The image provoked some use of highly literary language and simile: "anxiously he pressed his palms together and prayed", the jar was seen as something medical (1) or scientific, with the man being a specimen (1) or a laboratory rat (1). Within the more positive interpretations, the man is seen as resigned (2), contemplative, thinking (7), pensive and meditative. He is content (3) and safe (4), and protected (1), however, although he is safe, safety can be limiting (1) or boring (1). His condition is seen as a status quo, and it might be for this reason that the conventional story elements relating to conflict and suspense are less marked.

In the second image, the term "man" was used 11 times while "he" was used 17 times. Physical aspects were referred to: the man was "slim" (1), "young" (1), wearing a short sleeved shirt (3) which is dark (1), while the trousers (2) are pale (1). The scarf was mentioned nine times, and the overall effect was described as "light clothing" (1) and "casual" (1). The man is bare-footed (10), and looking upwards (2). The man holds a "flower" (12), two respondents specifying that the flower is red. The term "rose" (6) is used to describe the flower although one respondent describes it as a poppy. "Sky" is mentioned eight times, and it is windy (8) with black clouds (4), stormy clouds (1), storms (3) and shadows (2). There were six references to "desert" and earth that is parched (2), dry (5), cracked (4), grey (1). One respondent identified the landscape as the Bolivian Salt Flats.

In terms of connotation, while the range is broader than the range of denotational interpretations, there were many references to trouble and despair: he was a "troubled soul", a "jilted lover", in "despair" (4), with "small hope", "desolate", speaking to God (2), "praying" (1) or "hoping" (1) for rain, having fallen on "bad times". He is lonely (3) and vulnerable (3). For three respondents, this was a story about the end of a love affair or a lost love. There were eight references to "waiting": for rain (1), for something to come down or to leave (1), maybe an aeroplane or a god, or something awful, or a storm. One respondent focused on the suspense element, writing that:

"Even though he seems to be in the middle of nowhere, abandoned and has no way out, he's standing his ground and going to face whatever is coming, in the hope that he can get through it – or perhaps he feels he has no choice and can't see any hope even though it's there (as he's looking up at the sky and the storm, than the flower)".

The flower is referred to as signifier of "hope" (7), "happiness" (2), "transformation" (1), "romantic aspirations" (1) and "new life ahead" (1). Another set of interpretations views the man as "confused", "lost", "stranded" or perhaps "dropped there" (3). This image signifies normality interrupted: perhaps he is about to be "taken away"(3), by an strange craft (1), UFO (1), "little green men" (1) or "aliens" (1). The image might signify the end of the world:

"Maybe he was just going about his daily business and the world started ending. That's why there's nothing around him and it's desolate. He wasn't expecting it. He's looking up and maybe that's because the end of the world is caused by extreme weather and he's looking at it, or he sees God or a higher being (or whatever) that's the cause of it maybe he just picked the rose up because it just happened to be left when his town/place of work/life got carried away."

These stories included elements of setting, of physical and psychological attributes of the characters, and elements of plot and literary devices. In addition, the stories suggest that the modality model is fairly robust when used with "writerly" images.

Discussion

In describing the two images, participants provided different narratives, which included themes such as imprisonment by choice, emotional burden, loss of a loved one and seeking answers. The presence of these themes supported Greisdorf and O'Connor's (2002a) findings that users are capable of providing rich descriptions when viewing an image, while existing indexing tools, however sophisticated, are limited in their ability to capture this richness (p. 7). Even when a description, such as a "man waiting to be lifted up" seemed irrational, this type of descriptive narrative revealed that users can complement expert-based and author-based descriptions. It also reinforced the concern of Rafferty and Hidderley (2007) that expert indexing is "single minded" (p. 399) because of its inability to capture the myriad detail that users might assign to images.

In addition, disagreements between participants were apparent throughout their descriptive narratives, thus supporting the findings of Taylor and Joudrey (2009) that there is possibly more of an element of ambiguity in the interpretation visual materials as compared to non-visual materials (p. 308), although the model of image modality might suggest that such a claim needs to be a little more nuanced and specific. This was evident in some participants' perception of the character in Figure 1 as scared and frustrated, while others thought he was calm and relaxed. Such variations illustrate the challenge of finding common perceptions among opposing points of view and creating a consistent system of indexing images.

This study could not verify the view of Shatford (1986) that individuals from different professional spheres attach different meanings to images (p. 42) although participants came from a variety of disciplines. It would be worth investigating this point by asking questions such as: was the participant's interpretation of Figure 2 as representing an artistic act influenced by prior interest or qualification in art? Or was another participant's view that Figure 2 represented an apocalypse the result of a deep interest in natural disasters? In addition, the researcher could not verify the propositions that the meaning of an image changes with time (Rafferty & Hidderley, 2005, p. 116), context (Duncum, 2004, p. 256), and the physical space in which the image is located (Becker, 1998, p. 88); or Brink, Squire and Bigelow's (2006) suggestion that the grouping of images can lead to different interpretations (p. 716).

The descriptions in both parts of the study reflected participants' inclination to look for abstract levels of meaning, or for the aboutness of an image (Shatford, 1986, p. 45), and this confirms the

findings of Jörgensen (2003), who believed that users seek images at the abstract level (pp. 97–98). It has been suggested that any image system that cannot meet this requirement of users would be rendered ineffective (Enser, Sandom, Hare & Lewis, 2007, p. 468).

Reflections on the study

This study explored story-telling in relation to indexing practice. The first part of the study examined a small number of responses in some detail. This investigation established that story telling plays an important role in how people interpret images, and suggested that incorporating elements of storytelling in to the indexing process might be valuable in relation to indexing exhaustivity. Current tagging practice tends to be in the form of inputting individual terms or short phrases, in other words, it operates mainly on the paradigmatic plane. It may be that operating at the syntagmatic plane – through sentences and stories – would allow us to capture a broader range of interpretations.

The second part of the study explored storytelling as an input method, in addition, it also tested the image modality model (Rafferty and Hidderley, 2005). The results of this study suggest that the model is robust in relation to "writerly" high modality contemporary images as well as the more "readerly" high modality contemporary images reported in the 2011 study. In addition, the results of this study tentatively suggest that a storytelling approach appears to encourage connotative responses. Comparing the results of the second part of this study with earlier work documented in Rafferty (2011), suggests that moving towards storytelling at the input stage produces more verbs in the responses and encourages a greater number, and so potentially a greater range, of connotative interpretations. One of the challenges in tagging is to encourage creative input while at the same time disciplining input. Earlier templates (e.g. Rafferty and Hidderley, 2005) have had limited success in practice. It might be that the ubiquitous form of the story would allow the development of a relatively structured but intuitive method of inputting interpretative content. If storytelling were to be developed as a method of indexing, it might be that analytical algorithms based on the conventions of the story could be used to identify and process story elements as individual terms.

The goal for many scholars has been to design image retrieval systems which can capture meaning and develop indexing approaches at pre-iconographic, iconographic and iconological levels. The modality model might make a contribution in this regard. For high modal, contemporary images, denotation is relatively narrow, and this might suggest that at the denotative level the majority view is an appropriate one to use. Differences in interpretation are likely to come at the connotative level, and at this level, storytelling might offer a particularly rich input method, as it appears to encourage creative and more varied responses. It might be at this level that we should allow for a range of interpretative approaches to be made available. Following on, there is a question about whether it might be useful to use syntagmatic displays of text at output level, in other words, to show not just individual terms in search results but possible the stories themselves, or suitably edited versions.

How might this aid retrieval and access to information? Such an approach might improve exhausitivity of indexing and might help with information seeking activities which are more serendipitous and possibly creative in nature. Informal studies undertaken in the classroom with popular music suggest that storytelling approaches might also work with music. In short, as the Web has allowed a greater number of people to search for more information, more easily and for more reasons, so we should be looking to developing creative approaches to information access.

The current study was limited to two images from the same artist. The images allow for an investigation of the high modality, contemporary "writerly" image, but the image was still highly modal and iconic. It would be interesting in future to experiment with images that are more clearly indexical or symbolic, such as caricatures, abstract fine art and three-dimensional photos. In many ways, this study has raised more questions than answers. What drives users to assign a particular description to an image? What is the role of experience, education, gender, or prior knowledge in the process of interpreting an image? Would it be possible to relate this area of research to other disciplines such as psychology, philosophy, behavioural sciences and computer science? How is meaning constructed in the human mind? How do we make sense of our world through viewing images?

This paper is an invitation to scholars and future students to investigate the possibilities of creating a story-based image system. As this topic is relatively new and interdisciplinary, there can be multiple approaches to it. Story telling adds richness to human inter-subjective communicative practice: it nudges us towards empathy, supports multivalent interpretation, and leads to an appreciation of other people's opinions. It is our way of making meaning and sense of the world, and our place in the world, and its very fundamental nature might offer an always-already template, or structure, for deep, connotative image indexing.

References

Barthes, R. (1970), *S/Z* (R. Miller, Trans.), Hill and Wang, New York, NY.

Barthes, Roland, 1973 [1957], *Mythologies*. Paladin, London.

Barthes, R. (1977), *Image-music-text* (S. Heath, Trans.), Fontana, London.

Becker, H. S. (1998), "Visual sociology, documentary photography, and photo journalism: It's (almost) all a matter of context", in J. Prosser (Ed.), *Image-based Research: A Sourcebook for Qualitative Researchers,* Falmer, London, pp.84-96.

Berger, J. (1972), *Ways of Seeing,* BBC and Penguin, Harmondsworth.

Brink, W., Squire, D. M., & Bigelow, J. (2006), "The meaning of an image in content-based image retrieval", in Thibaud Latour and Michaël Petit eds., *2nd International Workshop on Philosophical Foundations of Information Systems Engineering (PHISE 2006), in conjunction with the 18th International Conference on Advanced Information Systems Engineering (CAiSE'06)*, Luxembourg, pp. 710-719.

Brophy, P. (2009), *Narrative-based Practice*, Ashgate, Aldershot.

Burke, M. A. (1999), *Organization of Multimedia Resources: Principles and Practice of Information Retrieval.* Gower, London.

Chandler, Daniel, (2002), *Semiotics: The Basics,* Routledge, London.

Choi, J., Cho, M., Park S. H., & Kim, P. (2003), "Concept-based image retrieval using the new semantic similarity measurement" in *Proceedings of the 2003 International Conference on Computational Science and its Applications, 2667/2003,* pp. 79–88. doi: 10.1007/3-540-44839-X_9

Choi, Y., & Rasmussen, E. M. (2003), "Searching for images: The analysis of users' queries for image retrieval in American history", *Journal of the American Society for Information Science and Technology,* Vol 54 No 6, pp. 498–511. doi: 10.1002/asi.10237

Chopin, K. (2008), "Finding communities: Alternative viewpoints through weblogs and tagging", *Journal of Documentation,* Vol 64 No 4, pp.552–575.
doi: 10.1108/00220410810884075

Chowdhury, C. G. (2010), *Introduction to Modern Information Retrieval* (3rd ed.). Facet, London.

Daele, V. S. (2010), *Bottled up emotions* [image], available at:
http://www.flickr.com/photos/renaissancestudios/5184287516/

Daele, V. S. (2010), *In Flander's Field* [image], Available at:
http://www.flickr.com/photos/renaissancestudios/5160655215/in/photostream/

Deb, S., & Zhang, Y. (2004), "An overview of content-based image retrieval techniques", in *Proceedings of the 18th International Conference on Advanced Information Networking and Application,* Vol 1, pp. 59–64.
doi: 10.1109/AINA.2004.1283888

Duncum, P. (2004), "Visual culture isn't just visual: Multiliteracy, multimodality and meaning", *Studies in Art Education,* Vol 45, No 3, pp. 252–264, available at:
http://www.jstor.org/stable/1320972 (16 May 2013)

Eakins, J., & Graham, M. (1999), *Content-based Image Retrieval* (Report No.39). JISC, Manchester, available at: http://www.jisc.ac.uk/media/documents/programmes/jtap/jtap-039.pdf (16 May 2013)

Eakins, J. P. (2002), "Towards intelligent image retrieval", *Pattern Recognition,* Vol, *35* No 1, pp. 3–14. doi: 10.1016/S0031-3203(01)00038-3

Enser, P. (2000), "Visual image retrieval: Seeking the alliance of concept-based and content-based paradigms", *Journal of Information Science,* Vol 26 No 4, pp.199–210. doi: 10.1177/016555150002600401

Enser, P., Sandom, J. C., Hare, J. S., & Lewis, P. H. (2007), "Facing the reality of semantic image retrieval", *Journal of Documentation,* Vol 63 No 4, pp.465–481. doi: 10.1108/00220410710758977

Greenberg, J., Pattuelli, M. C., Parsia, B., & Robertson, W. D. (2001), "Author-generated Dublin Core metadata for Web resources: A baseline study in an organization", *Proceedings on the International Conference on Dublin Core and metadata applications, Japan, 2*, pp. 38–46, available at:

http://www.dublincore.info/archives/2001/10/public_html/proceedings/product/paper-06.pdf

Greisdorf, H., & O'Connor, B. (2002a), "Modelling what users see when they look at images: a cognitive viewpoint", *Journal of Documentation,* Vol 58 No 1, pp. 6–29. doi: 10.1108/00220410210425386

Greisdorf, H., & O'Connor, B. (2002b), "What do users see? Exploring the cognitive nature of functional image retrieval" in *Proceedings of the 65th Annual Meeting of the American Society for Information Science and Technology, New Jersey,* Vol 39 No 1, pp. 383–390. doi: 10.1002/meet.1450390142

Hall, S. (2001 [1973]), "Encoding/decoding", in Durham, M.G. and Kellner, D.M.,(Eds), *Media and Cultural Studies: Keyworks,* Blackwells, Oxford, pp.166-177.

Hodge, R. and Kress. G. (1988), *Social semiotics,* Polity, Cambridge.

Ingermanson, R., & Economy, P. (2010). *Writing fiction for dummies.* Hoboken, NJ: Wiley.

Jörgensen, C., Jaimes, A., Benitez, A. B., & Chang, S. (2001), "A conceptual framework for classifying visual descriptors" *Journal of the American Society for Information Science and Technology,* Vol 52 No 11 pp. 938–947. doi: 10.1002/asi.1161

Jörgensen, C. (2003), *Image Retrieval: Theories and Research,* Scarecrow, Lanham, MD.

Krause, M. (1988), "Intellectual problems of indexing picture collections". *Audiovisual librarian*, Vol 14 pp.73-81.

Kress, G and Van Leeuwen, T. (1996), *Reading Images: The Grammar of Visual Design*, Routledge, London.

Kvale, S. (1996), *Interviews: An Introduction to Qualitative Research Interviewing. Thousand Oaks*, Sage, London, Thousand Islands, CA.

Lew, S. M., Sebe, N., Djeraba, C., & Jain, R. (2006), "Content-based multimedia information retrieval: State of the art and challenges" in *ACM Transactions on Multimedia Computing, Communications and Applications,* Vol *2* No 1 pp.1–9. doi: 10.1145/1126004.1126005

Marshall, C. & Rossman, G. B. (1995), *Designing Qualitative Research* (2nd ed.). Sage, London, Thousand Oaks, CA.

May, T. (2001), *Social research: Issues, Methods and Process* (3rd ed.), OUP, Buckingham.

McDonnell, M., & Shiri, A. (2011), "Social search: A taxonomy of, and a user-centred approach to, social web search", *Electronic Library and Information Systems,* Vol 45 No 1 pp.6–28. doi: 10.1108/00330331111107376

Merholz, P. (2004), *Metadata for the Masses,* available at: http://www.adaptivepath.com/ideas/e000361 (16 May 2013)

Monteleone, T. (2004), *The Complete Idiot's Guide to Writing a Novel: Presenting Techniques for Laying the Foundation of a Great Novel*. Penguin, Harmondsworth.

Nesta, F., & Mi, J. (2011), "Library 2.0 or Library III: Returning to leadership", *Library Management,* Vol 32, No1/2 pp. 85–97. doi: 10.1108/01435121111102601

Novakovich, J. (2008), *Fiction Writer's Workshop* (2nd ed.), Writers Digest Books, Cincinnati, OH.

O'Connor, B. C., O'Connor, M. K., & Abbas, J. M. (1999), "User reactions as access mechanism: An exploration based on captions for images", *Journal of the American Society for Information Science,* Vol 50 No 8 pp. 681–697. doi: 10.1002/(SICI)1097-4571(1999)50:8<681::AID-ASI6>3.0.CO;2-J

Olson, H. A., & Wolfram, D. (2006), "Indexing consistency and its implications for information architecture: A pilot study", *ASSIS&T Information Architecture Summit,* Canada, available at: http://www.asis.org/~iasummit/2006/files/175_Presentation_Desc.pdf

Panofsky, E. (1955), *Meaning in the Visual Arts*. Penguin, London.

Rafferty, P. (2011), "Informative tagging of images: the importance of modality in interpretation"', *Knowledge Organization*, Vol 38 No4 pp. 283-298

Rafferty, P., & Hidderley, R. (2005), *Indexing Multimedia and Creative Works: The Problems of Meaning and Interpretation,* Ashgate, Aldershot.

Rafferty, P., & Hidderley (2007), "Flickr and Democratic Indexing: Dialogic Approaches to Indexing, *New Information Perspectives,* Vol 59 No 4/5 pp.397–410. doi: 10.1108/00012530710817591

Rorissa, A. (2008), "User-generated descriptions of individual images versus labels of groups of images: A comparison using basic level theory", *Information Processing and Management,* Vol 44 No 5 pp.1741–1753. doi:10.1016/j.ipm.2008.03.004

Rowley, J., & Hartley, R. (2008), *Organizing Knowledge* (4th ed.), Ashgate, Aldershot.

Saldaña, J. (2009). *The Coding Manual for Qualitative Researchers,* Sage, London

Schmidt, S., & Stock, W. G. (2009), "Collective indexing of emotions in images: A study in emotional information retrieval (EmIR)", *Journal of the American Society for Information Science and Technology,* Vol 60 No 50 pp. 863–876.
doi: 10.1002/asi.21043

Shatford, S. (1986), "Analyzing the subject of a picture: A theoretical approach", *Cataloging & Classification Quarterly,* Vol 6 No 3 pp. 39–62.
doi: 10.1300/J104v06n03_04

Shatford-Layne, S. (1994), "Some issues in the indexing of images" *Journal of the American Society for Information Scienc*e *and Technology,* Vol 45, pp.583-588, available from:
http://polaris.gseis.ucla.edu/gleazer/462_readings/Layne_1994.pdf   (16 May 2013)

Svenonius, E. (1994), "Access to nonbook materials: The limits of subject indexing for visual and aural languages", *Journal of the American Society for Information Science and Technology,* Vol 45 No 8 pp. 600–606. Retrieved from
http://polaris.gseis.ucla.edu/gleazer/462_readings/Svenonius_1994.pdf

Taylor, A. G., & Joudrey, D. N. (2009), *The Organization of Information* (3rd ed.). Libraries Unlimited, Westport, CT.

Zhang, YL. H., Wenyin, L., & Hu, C. (2003), "Joint semantics and feature based image retrieval using relevance feedback", *IEEE Transactions on Multimedia,* Vol 5 No 3 pp. 339–347. doi: 10.1109/TMM.2003.813280