

From Spectator to Annotator: Possibilities offered by User-Generated Metadata for Digital Cultural Heritage Collections

Seth van Hooland, Université Libre de Bruxelles *

september 2006

Abstract

The changing role of the user, that gradually shifts from a passive consumer of information towards a pro-active user that reorganises and manipulates data, has an increasing impact on traditional information retrieval. A multitude of practical and methodic questions rise as popular web-applications such as blogs, RSS and social bookmarking tools allow users to create and share metadata about online resources. This article tackles these issues in the particular domain of visual cultural heritage. Online image databases increasingly offer users possibilities to annotate and comment on images of interest to them. But what is the pertinence of these user contributions? How can their quality be evaluated? Concretely, our article starts with an introduction to the phenomenon of user-generated metadata by presenting the social tagging of cultural heritage images and the practice of publishing users comments. Secondly, a case study presents an analysis of users comments within the image database of the National Archives of the Netherlands. Based on these empirical data, conclusions and generalizations outside our specific case study are formulated.

This paper will be presented by the author at the CILIP Cataloguing Indexing Group Annual Conference (13-15 September 2006, University of East Anglia) and published in "Catalogue and Index" following the conference.

1 User-generated metadata: an introduction

Despite advances in the field of content-based image retrieval, the effective retrieval of high level semantic meaning within cultural heritage image databases

*The author would like to thank the Nationaal Archief for its cooperation and Isabelle Boydens and Pascal Francq for their corrections and recommendations.

still relies on human indexing. The combination of the high costs of this operation and the increasing amount of digital content created often results in very minimal descriptions of the online published images. The high diversity of the content of large image collections makes it also practically impossible for an institution to ensure sufficient in-house knowledge for the description of such heterogeneous collections.

Technological developments of the collection registration tools at the end of the 1990's provided to some extent a solution towards these problems. The wide-spread use of broadband internet and the maturation of web-applications allowed the breakthrough of web-based collection management systems. Local installation of a collection management software is no longer required as access to the database can be ensured by an internet connection and a secure login. In this sense, web-based applications provided the first conditions towards distributed collection management. Theoretically, the task of cataloging and indexing can now easily be outsourced to external experts if insufficient inhouse knowledge is available or if the sheer quantity of the images to index is too large. The database back-end can administer read/write/update access on the level of an individual object and its individual descriptive fields. Large scale digitalization projects such as the scanning of hundreds of thousands of historical postcards, newspaper clips or photographs can now recruit a large number of temporal employees, that receive a specific training for the description, that can directly work in the database, independently of their location.¹

But presently, the distribution of the description of digital resources is evolving in a further direction. The rise of web2.0 technologies increasingly enables users to interact with online resources. Visitors of certain cultural image databases can not only create their own personal account and subselection of images within the database, but also post their opinion, comment or description of specific images within the database. Some cultural content providers, such as the Metropolitan Museum of Art and the Guggenheim Museum are now taking interactivity with users even a step further by adopting a form of collaborative indexing [3]. This approach has been directly inspired by web applications such as the social bookmarking services Del.icio.us (<http://del.icio.us/>) and the image sharing site Flickr (<http://www.flickr.com/>), two emblematic applications of the web2.0 movement. These allow the attribution of metadata in the form of one-word descriptors or tags to shared content. These applications introduced the notion of peer-to-peer (P2P) within information retrieval. Just as P2P computer networks are based on the computing power and bandwidth of

¹An exemplary project of this type of outsourcing was carried out for the Central Bureau of Genealogy of the Netherlands, when 50.000 scans of historical newspaper clips were described within 3 weeks by 20 temporally recruited persons which worked from their homes. Several metadata fields in relation to the person, activity/event and location depicted, the date of publication and the newspaper from which the clip was taken had to be transcribed from the scan of the newspaper article. For optimal quality of the metadata, all these fields were two times encoded by different persons. The encoded values were then automatically compared within the database so that different versions of a certain metadata field were traced. In this way typographic and interpretative mistakes were avoided, which is a must in these large scale description projects. This collection can be consulted on <http://www.geheugenvannederland.nl/>

its participants and not on a client-server model, the indexing of content within the above mentioned social software applications does not rely on a centralized limited number of trained information professionals that use a specific vocabulary, but entirely on the user community. Each user can freely apply the tags of his choice in a libertarian manner, in order to organize and retrieve information that he found spread across the internet (e.g. Del.icio.us) or a specific database (e.g. Flickr). This form of information retrieval is commonly referred to as folksonomy. The social or "folk" aspect arises from the possibility to consult the tags assigned by other users. Searches within the library of a user within a P2P network that offered an interesting resource, will most-likely lead to the discovery of other interesting resources. The same informal process of recommendation also appears within folksonomies.

This approach differs fundamentally for several reasons from the traditional resources indexing and cataloguing. Firstly, neither the form or the content of the description is controlled. Within a free-text field, users can insert whichever terms they want to associate with the resource, in a random order. Tags that describe the content or the author of the resource are contained within the same field. No controlled vocabulary from a classification scheme or thesaurus is used. Some systems, like Del.icio.us, do recommend tags that other users already used to tag the same resource. Some parallels can be drawn with the research on the subject of democratic indexing. This approach, specifically developed for image indexing, is based on the idea that the collection of meanings as constructed by the image users should be used to create a subject-based index [11]. Folksonomies implement in an automated manner democratic indexing. When assigning tags to an image in Flickr, the user can choose to use the tags that already have been used. Gradually, a set of common-based descriptors is constituted. This set of the most popular tags can then guide the search process in the form of "tagclouds". These also reflect very intuitively and dynamically the evolving interests of a user-group.

Secondly, the tagging is performed by the entire user community and hence not by a limited number of information professionals. This new "distributed" approach of metadata creation can be related to a more fundamental change in the relationship between a user and a resource. In a networked environment, the user has increasing possibilities to interact with a digital resource. In this context, the unidirectional model of communication, as developed by Shannon-Weaver, could be extended with the destination that automatically acts upon the initial information source. This applies to all kinds of digital resources, but images in particular:

The matter is that in the digital era, the image is not in the first place a digital image, but most of all something completely different: the image has become the processing of an image. In other words: the object has become an action.[2]

This also changes the role of the spectator, who can no longer look innocently at an image but inevitably acts on the image, even in the most minimal way by selecting, panning or zooming. Posting a comment related to a self-made

subselection of an image by users on the Flickr photo sharing site is exemplarily for this trend.

Thirdly, whereas traditional indexing and cataloging theoretically should, but in practice never can, exclude subjective interpretation of images, user-generated metadata such as tags or comments are partially based on information intrinsically related to the person performing the tagging. [10] differentiates seven functions of tags: identifying what (or who) it is about, identifying what it is, identifying who owns it, refining categories, identifying qualities or characteristics that reflect an opinion, self reference and task organizing. The last three functions are strictly related to the personal opinion or use of the resource tagger. The definition of qualities, mainly under the form of adjectives, can be seen as a recommendation system. Self reference tags such as "mycomments" specify the relation of the tagger to the content. Task organizing tags as "toread" help the tagger to group content related to a certain task.

Folksonomies have been widely adopted by technology-aware individuals as a new method for organizing and retrieving online content. But as mentioned previously, some early adaptors within the museum sector are currently implementing folksonomies within their collection registration database. One question immediately rises: do these projects reflect just a current hype, launched by the information industry, or do they announce an evolution that will spread out in the entire cultural heritage sector? ² The intuitive and informal search options as offered by tagclouds could offer users with no specific search need a valuable new tool. But folksonomies can in no sense replace traditional cataloging and indexing, that are based on adequate vocabularies and performed by trained information professionals. Problems related to polysemy, synonymy and basic level variation result in an increased number of false positives and false negatives [10]. More research should be conducted regarding this issue but generally speaking, tags offer an very poor semantic value when used for describing images.³

For this reason, we propose to focus our analysis of user-generated metadata on the structured sending of free-text comment. As comments are not restrained to a chain of one-word descriptors, they can offer a higher semantic value and have more potential use for implementation within cultural heritage databases. Throughout history, users have annotated and commented upon resources. Let us think for example about the transcribers of manuscripts who frequently added glosses to the original work. And as [17] notes, historiographs in the 17th and 18th century also relied on the feedback of their readers to correct their work. But the appearance of online databases has completely revolutionized the com-

²A lively debate regarding the larger issue of "community curators" can be followed on the Museum Computer Network (<http://www.mcn.edu/>) listserv.

³A good illustration of the poor semantic value of tags is demonstrated by the ESP-game (<http://www.espgame.org/>) which ames at "labeling all the images on the web". It is a so called two-player game, where each player is randomly paired with another player and confronted with the same image. The goal is to guess what tags the other player is assigning to the image, in order to find a common ground tag. Most successful tags are very generic descriptors ("man", "tree", "red") that offer few opportunities for specific image retrieval afterwards.

menting and annotation process, due to the possibilities of storage, publication and, most importantly, retrieval of the comments. However, insufficient research has been undertaken to examine the possibilities offered by these comments for information retrieval. [12] presented an interesting case study on the use of stories told by users to improve the description of historical images. But the main focus of the article was to document the authoring tool which manages the user-generated comment. An in depth analysis of user comment in itself and its pertinence to the user community, as we propose in this article, has not yet been presented.

2 Case study: the Image Database of the National Archives of the Netherlands

The image database of the National Archives of the Netherlands⁴ was launched in 2004 and contains approximately 500.000 images. The collection of the former press agency Anefo forms the backbone of the database. Practically each Dutch news-item (politics, sports, culture, economy) but also the daily life activities from 1945 until 1989 can be illustrated with photographs of this collection. In this respect, the database represents a huge source for the contemporary history of the Netherlands. Most of the images made accessible through the database were digitized in 1997, and no longer respond to current quality standards. The National Archives are now gradually replacing the older, low resolution files with high quality scans, that offer the opportunity to users for zooming and ordering reproductions.

From the detail view, which represents the selected image accompanied with metadata such as description, date, keywords, collection, photographer, press agency, catalogue number, users can click on a "comment" link which directs them to a form whereupon free-text comment can be inserted. The user has to fill in name and e-mailaddress, and indicate whether these informations can be published along with the comment on the website. The comment itself may not exceed 1.000 characters, but no other guidelines or restrictions regarding the content or style of the comment are given. A link on the home-page enables users to consult all the comments that have been published on the site. When a comment has been approved by a reviewer, it is published underneath the existing metadata.

As we just mentioned, the comments are reviewed before publication on the website. A specific employee of the collection management department judges individually the relevance of each comment within the back-end database. A comment can receive a different status, ranging from total deletion (the comment is seen as not pertinent and is never published on the website), provisional (newly arrived, not published), approved/to be treated (the comment is published on the website, but the existing metadata have not yet been updated), approved/treated (the existing metadata have been augmented with the users

⁴<http://beeldbank.nationaalarchief.nl>

comment, but the comment itself is no longer published on the website) and approved/published (the comment has been treated and remains published on the website). These categories have been implemented to organize the treatment of the comments, but do not reflect a consistent analysis of the relevance of the comments. Therefore we have decided to perform our analysis on the raw data as they entered the back-end database.

2.1 Evaluating the quality of the comment

Our goal is to evaluate the quality of the user comments. But how can we define information quality in our particular context? The objective "correctness" of the comments as a criteria can not be easily applied to our situation. Despite the specific and very concrete nature of the comments (see below), it is not feasible to investigate the truthfulness of the user comments. Following the "fitness for purpose" information quality definition, our quality measurement will be the relevance of the comment towards the user community.⁵ So firstly, we need to determine the needs of the user community of the image database. The National Archives have not conducted any studies regarding the specific public they serve with the image database, nor on the specific user requests, so we needed to perform supplementary research regarding these issues. As query analysis has been widely adopted within the image research domain as an effective method to define user needs [5], we decided to adopt this method. The mapping of the analysis of the user queries with the analysis of the user comments should give us an idea of the overlap and hence the pertinence of the comments towards the user needs. Nevertheless, we should also consider that the process of image searching on the web has a iterative character, a factor that is not reflected in our sample population. In order to facilitate the mapping, we have used the faceted classification of Shatford [15] (see table 1) as a framework for the categorization of both the queries and the comments. The process of categorizing heterogeneous content from user queries and comments into a limited set of abstract categories always implies some degree of violation of reality, but considerable preliminary testing with existing and new developed classification schemes proved the Shatford classification to be adequate for our purpose, as it focusses both on the level of specificity and the subject foci. Moreover, as this classification scheme already has been used in previous studies [1, 5] on the subject of user queries in image collections, it allows us to compare our results with previous research.

2.1.1 Mapping user queries to the Shatford classification

Several search options are offered on the website. Visitors can either submit their query using the simple search field on the homepage, or they can use the

⁵As noted by Svenonius [16], the principle of user convenience stands central within Anglo-American cataloging literature. But we also have to acknowledge that this notion is somehow problematic, since there is no such thing as a single, homogenous public.

	Iconography (Specifics)	Pre-iconography (Generics)	Iconology (Abstracts)
Who?	individually named person, group, thing (S1)	kind of person or thing (G1)	mythical or fictitious being (A1)
What?	individually named event, action (S2)	kind of event, action, condition (G2)	emotion or abstraction (A2)
Where?	individually named geographical location (S3)	kind of place: geographical, architectural (G3)	place symbolized (A3)
When?	linear time: date or period (S4)	cyclical time: season, time of day (G4)	emotion, abstraction symbolized by time (A4)

Table 1: Shatford faceted classification

advanced search mode which allows them to search simultaneously within specific fields. Our analysis was based on the queries using the simple search field, as this is the most popular search option. During the period 05/04/2004 (date when the site went online) - 01/03/2006 (date when export was made) 465.124 searches were launched. Using a confidence interval of 5 percent and a confidence level of 95 percent, we obtained a sample population of 384 comments, which the author analyzed and mapped onto the Shatford classification. This analysis resulted in the following results: S1: 17,50%, S2: 5,5%, S3: 57%, S4: 2,5%, G1: 9%, G2: 8,5% (G3, G4, A1, A2, A3 and A4 are not represented). The majority of users want to retrieve images related to a specific geographical location. Secondly, searches regarding specific individuals, groups or objects are also very popular. On the other hand we can state a total lack of use of abstract query terms.

The two studies of P.G.B. Enser [1] regarding image requests in large non-domain-specific image databases confirm these outcomes. The larger part of queries refer to specific instances and unique items as object names and geographical locations, whereas more general and abstract concepts are not included. Studies focussing on requests within newspaper image archives reaffirm these results.

2.1.2 Mapping user comments to the Shatford classification

For our analysis we rely on an export from the administrative back-end database that stores all the metadata concerning the image database. The file contains 4647 comments, sended in by users between the 5th of April 2004 and the 1st of March 2006. Each comment is accompanied by the name of the person that sent it in, the date of receipt and the status of the comment within the database, as

described in the previous paragraph. Using a confidence interval of 5 percent and a confidence level of 95 percent, we obtained a sample population of 355 comments, which the author analyzed.⁶ As some users sent the same type of comment regarding a series of photographs after each other and the population of comments we worked on was sorted chronologically, we decided to perform a systematic sampling method, which ensured the spreading of the sample evenly over the population. As the comments are quite lengthy, they can easily incorporate more than one category of the Shatford classification, e.g a comment can contain information regarding an individually named person, group, thing (S1) and individually named geographical location (S3). The analysis gave the following results: S1: 67,61%, S2: 18,87%, S3: 30,70%, S4: 20,56%, G1: 6,29%, G2: 1,71%, G3: 0,57% , G4: 0,29%, A2: 2,86% (A1, A3 and A4 are not represented).

2.1.3 Mapping of user queries with user comment

Figure 1 demonstrates a general correlation between the content of the queries and the comments. Both queries and comments are highly motivated by interests in specific terms, use few generic terms and hardly any or no abstract notions. But differences in popularity among the subcategories of specific, generic and abstract exist however. The most popular queries concern individually named geographical locations (S3), whereas the most prevailing comments relate to individually named persons, groups or objects (S1). This can be explained by the interest of users to find images of their hometown, street or even house, whereas users will most likely comment upon specific persons, groups or objects they know. But in spite of these and other differences, we can conclude that the mapping of queries and comments demonstrated similarities in their content. The comments thus help to fulfill a concrete information need from the user community.

2.2 Further analysis of the comments

Apart from the mapping of the comments and queries onto the Shatford classification, we also decided to deduct the most recurrent characteristics of the user comment. The same sample population of comments was used as with the Shatford mapping. Our analysis of this population of comments resulted in the following categories:

- correcting of the displayed metadata (regarding spelling, identification of persons, event/action and geographical and temporal location): 34,13%
- including narrative elements in relation to the image: 18,87%
- linking of the user's personal history regarding the image: 4,29%

⁶Theoretically, the confidence interval is too high to be applied to small categories such as g2, g3, g4 and a2, but practical constraints did not permit a larger sample size[8]. Future research on this topic will be based using adequate sample sizes regarding all categories.

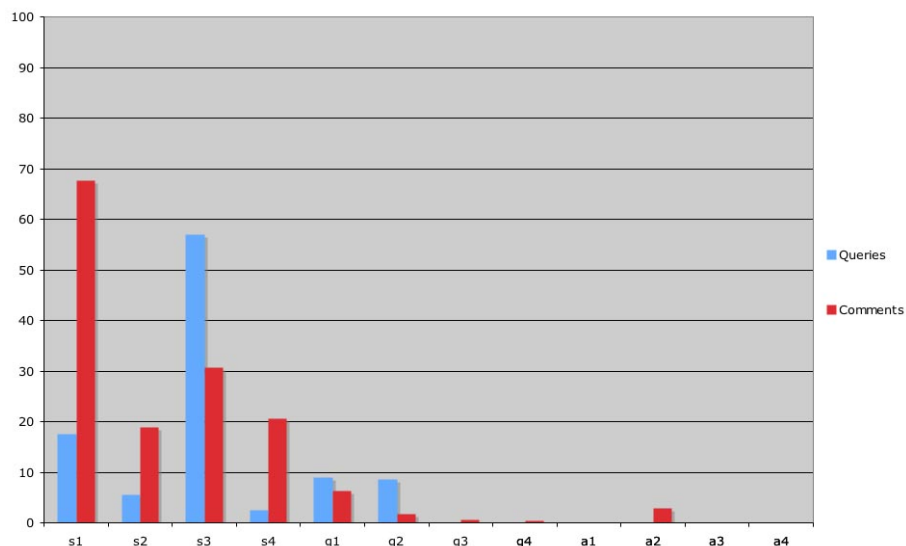


Figure 1: Correlation between user queries and user comment

- mentioning a false or inadequate display of the image: 3,14%
- stating an opinion or judgement: 2,86%
- engaging of a dialog with the institution or other users, under the form of a question: 1,15%

These categories are neither exclusive or inclusive, meaning that a comment can belong to more than one category, but does not necessarily belong to one. Although the process of categorizing is somehow subjective, we do believe it allows a deeper understanding of the nature of user comments and their relevance towards the public.

2.2.1 Critical comments

The most important incentive for users to send comment is to express their disapproval with the published metadata on the website. 34% of all the comments criticize the existing metadata and propose a correction. Within this type of comment, 40,34% is relevant to incorrect spelling of the metadata (with the larger part spelling of names of persons and locations), 10,92% to the identification of persons, 24,37% to the identification of an object, 15,97% to geographical location and 8,40% relevant to time. The National Archives recognize on the website that some image descriptions may be incomplete or contain errors, and therefore explicitly invite users to indicate potential errors and to send possible corrections. We do not possess precise information regarding the validation of

the corrections, but according to the National Archives, more than half of this type of comment is effectively used to correct the existing metadata.

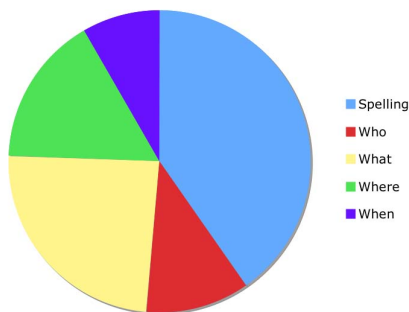


Figure 2: Categories of user proposals to correct existing metadata

2.2.2 Including narrative elements in relation to the image

Another recurrent feature of the comments is their narrative character. Almost 20 % of the comments relate separate events into a narrative form or a story that enables a deeper understanding of the image. The sequence of metadata fields related to the identification of persons/objects depicted, location, date, etc that are published by the National Archives on the site does not imply an understanding of images. User comment can introduce some narrative aspects within the database by making connections between events, persons, locations and temporal specifications. We can relate here to Lev Manovich, who noted the rise of the database as a cultural form and its tension with narrative:

As a cultural form, the database represents the world as a list of items, and it refuses to order this list. In contrast, a narrative creates a cause-and-effect trajectory of seemingly unordered items (events). Therefore, database and narratives are natural enemies. Competing for the same territory of human culture, each claims an exclusive right to make meaning out of the world.[14, pg 225]

Here it is not so much the content of the information, but its form that is of interest. This article does not allow us to go deeper into the subject, but recently, a whole new research area emerged that studies the impact and possible benefits of narrativity and storytelling within the domain of information representation.⁷

⁷For a more elaborate vision on the impact of narrativity within the representation of cultural heritage, consult for example [9].

used by other users as an informal recommendation system. Analogies can be made with the use of user bookreviews on Amazon (<http://www.amazon.com>).



Figure 5: Example of a user comment illustrating a personal opinion

2.2.5 Engagement of a dialog with the institution or other users, under the form of a question

A fraction of the comments contain questions or invitations toward the institution or other users to help identifying an image. Sometimes dialogues between users take place, which transforms the comments page in a forum where users can interact with one-another. But even if the technology for implementing a forum within a website is widely available at a low cost, there a very few heritage institutes that offer such a feature.⁸



Figure 6: Example of a user asking question

2.2.6 Mentioning a false or inadequate display of the image

A last category of comments consists of users mentioning an incorrect display of images. A recurrent problem with high-volume scanning projects of photo negatives is the appearance of mirrored images, which are very hard to detect. When no text is displayed on the image, one has to personally know the scene or object depicted in order to detect the mirrored publication.

⁸Consult <http://www.archimuse.com/mw2003/papers/bowen/bowen.html> for an overview of the use of forums within the cultural heritage sector.

3 Conclusions and future work

Heritage institutions are overflowed with a digital production stream they can not process the way they have been doing for the last century. The distributed indexing and cataloging of digital images among the user community could to some extent offer a solution to this problem. With the help of a case study we have presented an introductory evaluation of the quality of user-generated metadata, guided by the "fitness for purpose" definition of metadata quality. Using the Shatford classification scheme, we confronted the content of samples of user queries with samples of user comments in order to determine if the comments matched the users interests. We observed that both comments and queries are highly motivated by interests in specific terms, use few generic terms and hardly any or no abstract notions. We can thus conclude that the content of the comment lies within the search interests of the user community and by this accords to the "fitness for purpose" criterium. Secondly, we drew up a typology which groups six main characteristics of our sample population of user comment and related them to the notion of metadata quality. The most recurrent type of comment, the posting of corrections of the existing metadata, clearly offers possibilities to enhance the correctness and precision of metadata, just as the notification of an inadequate display of the images helps to clear the database from errors. It is however less evident to assess the pertinence within a historical image database of narrativity, the inclusion of personal experiences, opinions and the dialog between users.

It is clear that more research needs to be undertaken to explore the new possibilities offered by networked technologies for the indexing of digital cultural heritage. Personally, we would like to focus on the two following main issues within this research context. User participation can be considered as one of the key notions of digital culture [6]. But where do we have to draw the line between user-generated metadata as added-value or as the expression of the "narcissism of the viewer" [7] that deconstructs the authority of the expert?

Secondly, if heritage institutions infinitely allow users to add metadata to resources, how should these new metadata be incorporated into the existing metadata scheme? Cataloging and indexing has long been considered as a one-time operation, but as resources and their context evolve in time, the metadata too need to be updated. In order for the metadata scheme to support frequent updating, an "event-awareness" should be incorporated [13]. On the other hand, the user comment clearly demonstrates the inherent weakness and danger of metadata: they are ever extendible [4]. Research should thus focus on temporality and longevity of metadata. In conclusion, we should also investigate to which extent the conclusions drawn from our specific context of digital cultural heritage can be generalized to other domains.

References

- [1] L. H. Armitage and P. G. Enser. Analysis of user need in image archives. *Journal of Information Science*, 23(4), 1997.
- [2] J. Baetens. Notes on photo analysis (<http://www.photherel.net/notes/>).
- [3] D. Bearman, J. Trant, S. Chun, M. Jenkins, K. Smith, R. Cherry, T. Leonhardt, D. Uchitelle, D. Hiwiller, W. Lee, B. Wyman, R. Shah, B. Gawronski, and M. Morgan. Social terminology enhancement through vernacular engagement: exploring collaborative annotation to encourage interaction with museum collections. Technical report, Steve Museum, 2005.
- [4] I. Boydens. *Informatique: normes et temps*. Bruylant, 1999.
- [5] Y. Choi and E. M. Rasmussen. Searching for images: The analysis of user's queries for image retrieval in american history. *Journal of the American Society for Information Science and Technology*, 54(6):498–511, 2003.
- [6] M. Deuze. Participation, remediation, bricolage: considering principal components of a digital culture. *The Information Society*, 22(2):63–75, 2006.
- [7] H. M. Donald. The met's triumphant democratic elitism. *City Journal*, 2001.
- [8] J.-J. Drosbeke. *Eléments de statistique*. Editions de l'ULB, 2001.
- [9] W. M. Duff and V. Harris. Stories and names: archival description as narrating records and constructing meanings. *Archival Science*, 2:263–285, 2002.
- [10] S. A. Golder and B. A. Huberman. Usage patterns of collaborative tagging systems. *Journal of Information Science*, 32(2):198–208, April 2006.
- [11] R. Hilderly and P. Rafferty. *Indexing multimedia and creative works. The problems of meaning and interpretation*. Ashgate, 2005.
- [12] G. Howard, J. Pratty, and M. Stapleton. Storymaker: user-generated content - worthy or worthwhile? In *Museums and the web 2005*, 2005.
- [13] C. Lagoze. Business unusual: how "event-awareness" may breathe life into the catalog? In *Conference on bibliographic control in the new millenium*, 2000.
- [14] L. Manovich. *The language of new media*. MIT press, 2001.
- [15] S. Shatford. Analyzing the subject of a picture: a theoretical approach. *Cataloguing and Classification Quarterly*, 5(3), 1986.
- [16] E. Svenonius. *The intellectual foundations of information organization*. MIT press, 2000.

- [17] T. Verschaffel. *De hoed en de hond. Geschiedschrijving in de Zuidelijke Nederlanden 1715-1794*. Verloren, 1998.