OPEN ACCESS

University of BRISTOL

Tonkin, E. L., & Tourte, G. J. L. (2016). Using the crowd to update cultural heritage catalogues. 1-6. Paper presented at Involving the CROWD in future MUSEUM experience design - CHI 2016 Workshop, San Jose, CA, United States.

Peer reviewed version

License (if available):
Unspecified

Link to publication record in Explore Bristol Research
PDF-document

University of Bristol - Explore Bristol Research

General rights

Using the Crowd to Update Cultural Heritage Catalogues

Emma L. Tonkin
University of Bristol
Bristol, United Kingdom
`e.l.tonkin@bristol.ac.uk`

Gregory J. L. Tourte
University of Bristol
Bristol, United Kingdom
`g.j.l.tourte@bristol.ac.uk`

ABSTRACT

It is a familiar observation that digital cultural heritage brings with it new challenges. One such challenge is the effect of age on digital objects held within heritage databases, and on the array of materials that surround and support access to these resources. In this position paper, we discuss effects of long-term societal change on data preservation in digital cultural heritage, and present a means by which ongoing user modelling processes drawing on contemporary resources can support 'just-in-time' preemptive review of material to be presented to the public, as well as feeding into enhancement of data retrieval processes. We remark that similar issues and principles apply in contemporary information access contexts: for example, the processes of information sharing between expert practitioners and non-expert members of the public may exhibit similar effects.

ACM Classification Keywords

H.5.2 Information Interfaces and Presentation (e.g., HCI): User Interfaces (D.2.2, H.1.2, I.3.6); I.2.4 Artificial Intelligence: Knowledge Representation Formalisms and Methods (F.4.1)

Author Keywords

Metadata; Museum; Catalogue;

INTRODUCTION

Digital preservation is a significant focus for cultural heritage organisations. In particular, museums, libraries and archives are increasingly working with digital objects, resulting in a significant and growing literature on the challenges of digital preservation [17]. Long-term digital preservation refers to 'processes, strategies and tools used to store and access digital data for long periods of time', according to Factor et al. [11]. The time period in question is sufficiently long that technologies, formats, hardware, software and technical communities are likely to change.

There is, we are told, a risk of a 'digital dark age', in which data from the digital age is lost irrevocably [21]. Digital preservation attempts to mitigate such risks, a process which often involves some sort of maintenance: for example, in the case of digital objects, format shifting can resolve problems of format, hardware and software obsolescence.

In this position paper, we explore the issues which can and do contribute to a 'digital dark age', and propose an approach—identifying promising related work in human-computer interaction and information retrieval—to enhance usability of digital data in the future. Whilst we acknowledge that this alone cannot mean that an digital dark age—or similar lack of usability or access—is averted, it can begin to help researchers consider some of the issues in this area.

Our research interest is not solely in preservation of access to data across a large temporal gulf, since similar issues of accessibility and usability arise in contemporary contexts; for example, Burrows et al. [7] describe the benefits of 'actively develop[ing] a shared language' between specialist system designers/implementers and end-users. We have a keen interest in this area and hope to explore the adaptation of the methods in this paper to support such developments. However, cultural heritage catalogues, as large semi-structured datasets, offer an opportunity to evaluate the effects of cultural and contextual change over time.

In the following, we review some of the long-standing research which has contributed to our understanding of the usability and accessibility concerns presented by online/digital catalogue information. Given these problems, we then discuss in more detail how these and other related issues of digital preservation have an impact upon the cultural heritage sector. In particular, we focus upon the search and retrieval process relating to catalogues, as this is a core function of first- and second-generation cultural heritage organisation informatics as well as a key architectural component underlying implementation of additional functionality: education, study and enjoyment.

In doing so, this leads us to our proposal, which we believe will provide a new insight in detecting the early signs of possible digital dark ages in this sector. Specifically, we believe that since language is the 'currency' both in indexing catalogues and in search and retrieval behaviour, it makes sense to accommodate potential changes or differences in usage which may provide a barrier to usage for a proportion of those accessing it. By automatically considering linguistic differences, we can identify cases in which performance of mitigating maintenance actions may reduce impact of change, adding further information to support the active curation processes implemented by cultural heritage professionals. In particular, we

suggest that data drawn from social sensors and cultural media mining could usefully support such processes of revision.

Whilst the main body of this paper deals with digital infrastructure, we remark that objects and their surrounding data are accessed in a variety of physical and online contexts. Many of the issues described in this paper are relevant across online and physical contexts.

THE USABLE CATALOGUE

In the context of digital cultural heritage, interest in understanding the strengths and weaknesses of the online catalogue grew with the evolution of the Web. It is notable that the online library catalogue significantly pre-dates broader digital cultural heritage efforts. Consequentially the following literature review considers evidence drawn from study of the library catalogue as well as more recent studies focusing directly on cultural heritage.

Two influential papers by Christine Borgman and published a decade apart document the development of online catalogue usability between 1986 [3] and 1996 [4]. Conceptual aspects of system design were a major focus in the earlier paper: misunderstanding of system features, lack of use of advanced search techniques and difficulty in identifying appropriate subject headings ('headline labels' for relevant categories). Two key types of knowledge were identified [3]: knowledge of search syntax, semantics, structure and system, and knowledge of the conceptual aspects of search methodologies. In the later paper [4], Borgman refines the model, identifying

- conceptual knowledge—in which a person 'seeking knowledge or meaning [. . .] must formulate a query in terms of the content of information entities' [or proxies],

- semantic knowledge of query implementation—in this case, the semantics of the catalogue system

- technical skills which allow the user to navigate the computer interface and query syntax (syntactic knowledge).

Borgman noted that the capabilities of information retrieval far exceeded those exhibited by catalogue interfaces. Through literature review, Borgman discussed the effects of various factors relating to the catalogue user, such as personality, age and experience, on user search behaviour.

In the 1990s, key questions about cultural heritage were asked about both physical and online visitors: wants, needs and strategies for information access [9]. In general, a greater focus was placed on developing well-informed user models, seeking to understand patterns of use, visitor aims, information needs and search strategies.

Innovation in catalogue design continued into the 2000s. These included the practical implementation of 'next-generation' features [18] such as faceted navigation, attempts at correcting user error via methods such as spell-checking and automated format validation, enriched search metadata, ranking of search results and a greatly simplified interface that comprehended lessons learned from the simple, clear search interfaces offered by search engines such as Google [6]. The arrival of Web 2.0 gave rise to further innovations [34], many of which were focused on active, ongoing user participation, as well as the information discovery and retrieval model implemented in previous generations of catalogue software.

Interactive cultural heritage

The increasing focus of the 2000s on supporting user activity beyond catalogue search and discovery heralded further research projects and practical developments focusing on broader views of the role of a cultural heritage platform. Broader platform functionality is intended to support individual participant needs. Roles that participants may hold include that of student, self-directed learner, an individual planning a physical visit to the museum, subject expert and teacher [13].

Personalisation

Personalised access to museums, libraries and archives was explored during the 2000s by a large number of high-profile cultural heritage institutions [5]. The Rijksmuseum, for example, created a service called Cultural Heritage Information Presentation (CHIP) [32]. This made use of 'likes' and 'dislikes' expressed by the user; this 'non-obtrusive collection of user data' was designed to support the generation of personalised tours.

Under the name 'personal digital collections systems', functionality allowing users to select items from a museum's catalogue for a personalised collection was implemented by many museums, such as the Museum of Fine Arts in Boston, the National Museum of Australia and Tate Online [26]. Use of these systems was found to be popular with a subset of users [13], although often collections were abandoned shortly after creation [14], causing suggestions that the outcome of such systems is 'a landscape of "lost" personal museums' [26].

It is not clear whether the ephemeral nature of an individual's interaction with a cultural heritage system—creating and forgetting a list—should be seen to imply that the interaction is incomplete or superficial, as is often suggested in the literature. It is partially as a consequence of the short-lived nature of many visitor interactions, however, that many institutions have chosen to work with external services, often commercial in nature. These allow the institution to indirectly provide functionality such as social bookmarking or personal digital collection rather than providing the service themselves.

Introduction of a dependency on social websites introduces further risk of a 'digital dark age' [20], since it is not clear that either service model is sustainable. As such, information (such as course notes, expert or public comment and further annotation) may not be captured by the institution itself. Consequentially, such interactions may be both ephemeral and unlogged, unless action is taken to observe and document: an appropriate social sensor may be used to evaluate specific dimensions of social annotations or interactions, such as information shared, emotional responses and geographical localisation [27].

Co-visiting and shared spaces

In the examples shown above, the system is designed with the individual user in mind. Others, such as ARCHIE [31], were designed to support 'co-visiting'—interactive learning within the physical borders of the museum. ARCHIE followed the

contextual model of learning, which point to the contributions and influence of sociocultural, physical and personal context in interactions with objects and subsequent learning [31, 12]. The Sotto Voce project explored co-visiting and the creation of shared audio media spaces, focusing in part on shared listening to promote interaction [1]. Further projects consider location-aware mobile gaming as elements in cooperative learning [10].

The mission of the museum
These activities reflect what Lin [22], citing the International Council of Museums (ICOM), still describe in a vision statement as the mission of the museum:

> A museum is a non-profit, permanent institution in the service of society and its development, open to the public, which acquires, conserves, researches, communicates and exhibits the tangible and intangible heritage of humanity and its environment for the purposes of education, study and enjoyment.

This definition is a little puzzling, in that it states that a museum is 'open to the public' and identifies communication for education, study and enjoyment, but does not clearly identify any target audiences. Perhaps this simply reflects the span of ICOM's membership—given over 20,000 museums, it is to be expected that a variety of answers might be given to the question, 'who is this museum primarily intended to serve?'

In any case, this definition provides us with a series of items that a museum must support: the requirements of education, of study, and of enjoyment. The evidence suggests that provision of a service capable of providing all three of these is an ongoing process that is deeply dependent on a strong and current understanding of the visitor. In particular, we suggest that ongoing service provision in the museum context in particular requires careful monitoring and proactive response to changes in user behaviour, profile and context.

It is not yet clear in what time span issues associated with the types of change process monitored by 'long term' digital preservation become significant to each of these activities. A body of literature exists on website design and associated metrics for evaluating enjoyment [22, 23]. A broader literature on factors associated with enjoyment in human-computer interaction, such as flow [15], suggests a link between enjoyment and level of challenge (a factor associated with individual level of certain forms of preexisting knowledge). Education and study may be viewed as relating to preexisting formal knowledge, to lived experience, to exposure to contexts, peers and experts with shared areas of interests and to learning opportunities [19].

We suggest that the effect of ageing on each key aspect of the museum's mission is a useful area of study, in that the frequency of intervention and hence the sustainability of any institution's services is greatly influenced by this factor. In the remainder of this position paper, however, we will not consider these services further. Rather, we will, by reference to relevant theory, consider the effect of ageing on one relatively straightforward element of the system: the museum catalogue and its accompanying index.

THE AGEING OF CULTURAL COLLECTIONS
We are familiar with the marks of time. On stone, we expect to see erosion where water once flowed or generations of visitors have walked. On skin, liver-spots and crows' feet. Some of these marks are informationally rich physical clues to the experiences lived by other visitors in past years, such as a well worn passage in a book, or the scuffed floor where furniture once stood. A great deal of recent research on digital preservation has focused on forms of erosion that affect digital objects themselves, such as lack of compatibility with modern software [11]. The structural conveniences that accompany these objects—the texts and interlinks that provide context and render the material searchable and accessible—are generally engineered for a shorter timescale. What are the effects of age on the tertiary indexes and metadata structures that accompany those objects and make them actionable resources, such as catalogues of digital heritage material?

There are relatively few studies of the effects of age on catalogues over a significant period of time. There are various reasons for this, notably the fact that online catalogues are of themselves a fairly recent phenomenon. The opportunity to observe issues that occur over the long term, as in long-term data preservation scenarios, has therefore been limited. Confounding factors may also be identified. As we see above, catalogue interfaces have typically undergone significant amounts of re-engineering as new technologies and standards are developed and gain in popularity. Where funding exists, cultural heritage organisations will often look to state-of-the-art research, implementation guidelines and even upcoming trends in interface design to overhaul aspects of their online and public presence, meaning that the presentation of information is likely to change rapidly and often, even if the information itself is edited relatively infrequently. This adds a confounding aspect to any longitudinal study. Waterfield's review [33] of the development of art cataloguing demonstrates the significance that pragmatic, idealistic and nationalistic concerns took in the nascent field, discussing the complex interplay between these factors.

We must therefore look for evidence from indirect sources, such as theoretical and experimental research, which may inform our hypotheses about digital preservation issues on the humble museum catalogue.

Understanding the museum catalogue
The vast majority of museums hold some form of catalogue: in the simplest and most essential form, this is simply a list of objects. Lists usually include several pieces of information about the object, such as a title, date of entry into the catalogue, the origin of the object or the name of the object's creators or contributors and descriptive information about the object. Such information is generally described as 'metadata'. The art and technology of information description has developed over hundreds of years and remains in flux today, where several complementary approaches may be applied.

Manually contributed metadata
Traditionally, catalogue information is manually generated by expert cataloguers chosen from subject experts in the field.

3

It typically contains elements drawn from a controlled vocabulary (a taxonomy or list of subject headings) as well as free-text elements that may contain any choice of string input by the user.

As this process is manual and involves expert input, cataloguing is an expensive process. In some of its more expansive forms, cataloguing work is broadly viewed as prohibitively expensive. From time to time, efforts have been made to reduce cost by involving non-expert contributors: in the Great Depression, for example, non-experts were hired via the Welfare To Work platform to contribute to extensive cataloguing of historical records [2].

In more recent years, the annotation platforms offered on the Social Web offered further support for non-expert annotations [30], although opinions differ on the utility of the outcome, which supports serendipity over structured search [8].

Automated metadata
A further development is that of automated metadata generation, also known as automated indexing. The field of automated indexing has developed since the 1970s [29] to encompass areas such as image and multimedia resource indexing. A wide variety of methods are used to extract specific types of information, ranging from extraction of textual features and video captions to extraction of image features. Because the features extracted are seldom directly comparable to the types of metadata created in a traditional catalogue record, there is often a semi-supervised training process necessary to 'translate' findings to compatible catalogue terms.

The search and retrieval process
In the most basic search process, the user simply formulates a search query by providing a number of textual key terms. By interrogating the index of objects, making use of catalogue data, the service is able to identify and return matches. This process draws heavily on the user's ability to generate appropriate search key terms, ordinarily with little prompting from the interface. S/he is likely to be supported by second-generation site features such as search query processing via a thesaurus built into the interface and automated error identification/correction: such thesauri are not automatically updated and require ongoing work to maintain.

The problem of ageing indexes and potential mitigation

Studies in recognition and term generation
A contemporary text is written with an audience in mind, according to the author's perception of the strategies that will most effectively transmit his or her point. The same is true of catalogue entries, particularly elements that serve an interpretive or contextual purpose, such as descriptive text.

A subtle effect of passing time is an ongoing change in the style of written and spoken speech. Terminology, in particular, undergoes ongoing processes of change, with terms falling out of favour and being replaced by alternatives. A similar effect occurs between speech communities, which may prefer different terminology to others. Despite this change in preference, low-frequency lexical items are typically recognised long after falling out of regular use. However, low-frequency words are not well recalled [24].

Thus, as an index ages unmodified, the search process is compromised by difficulty in generating the necessary search terminology. The search process becomes slower and more error-laden, and the accessibility of the collection consequentially reduces.

Catalogue and index are not immutable: they are frequently updated for particular purposes, ranging from addition of new items to correction of existing entries for various purposes. A high-profile recent review of Rijksmuseum metadata, for example, saw the alteration of over a hundred items containing racially-charged terminology [28]. User complaints have been identified as a factor in the decision to review the metadata.

Mitigation
Taking these things together, we use them as our motivation and basis for addressing digital preservation within cultural heritage institutions. As we have just noted, words come and go from fashion: In some cases historical or outmoded terms may cause offence (such as that of the Rijksmuseum), whereas in others, they may cause a breakdown in understanding or ability to retrieve certain information.

As with many other aspects of ensuring a good user experience to the user of an online catalogue, the solution involves a strongly data-driven and frequently updated set of user profiles. As we have discussed earlier in this article, the 'user' of a museum is not clearly defined: reuse and interest may spring from many individuals or groups for a wide variety of purposes, from the sheer joy of collection to the requirements of academia. It is therefore necessary to take a broad statistical view rather than (or alongside) a detailed focus on individuals.

As an institution, it is preferable to schedule preventative maintenance in good time than to discover issues that have significantly inconvenienced large numbers of users. Thus, it is proposed that ongoing and non-invasive methods of data collection are used to identify cases in which catalogue (or, more frequently, thesaural) maintenance may become necessary, such that the proposed changes may be reviewed and implemented as part of the routine of data preservation.

For this purpose, a set of social [25] and corresponding sensors operating on cultural and political data (e.g., see [35]) is proposed that, in combination, represent a Museum Observatory. Such an observatory allows the museum, in an automated manner, to look out at the constellations of human activity and discourse that surround it and to identify and react to ongoing processes of change. For example, such an observatory might process data relating to news reports referencing the cultural institution, in order to get a sense of the context in which it operates and the discussions surrounding it at a formal level. In addition, social media may provide indications of the concerns of the general public (or sections of the general public). Inclusion of more structured data such as the catalogue search terms (or those in relation to the website), will also allow a more focused perspective on users of the cultural institution's resources.

The incorporation of a wide range of relevant resources which form the immediate ecosystem of the museum, enable a broad and varied view of its context. This observatory could then be integrated into the museum catalogue, for example, by generating relevant terms for automated (or semi-automated) annotation of objects, or for identifying and generating new relationships within the catalogue (e.g., between objects or between different indexing terms, or between objects and indexing terms). Although itself an actor in the ecosystem, the museum is one among many. We argue that to remain accessible to the visitor, the museum must retain an active connection with the general public and along with its wider context. Such an observatory may act as a virtual mirror on an institutional level [16]. It would, as the poet Robert Burns once put it, 'the giftie gie us/To see oursels as ithers see us!'

Such infrastructure, although cheap by comparison with the human-led processes which it can inform, is far from free: we remark, however, that such observatories do not need to exist in large numbers. In the case of larger institutions, they will most likely have the expertise and resources to create, and curate, such and observatory. However smaller institutions may need to form consortia or create links with the larger institutions; as with significant manual infrastructural efforts such as the development of openly accessible thesauri or taxonomies, the work is better shared than individually replicated.

FINDINGS
In an initial study, quantitative methods have been explored to support direct comparison with modern corpora drawn from the same topic area and corresponding repair processes. Material drawn from social sensors is considered as a source for data relating to impact and affect, as well as an entry point into relevant online discourse. It is hoped that this work will be reported in detail in an upcoming publication. This study, however, has given us confidence that text analytic methods based on the distributional hypothesis, twinned with topic-comparable corpora from distinct temporal or social contexts, can be used both to describe the gulf that lies between the two corpora, and to identify strategies that enable the construction of knowledge structures that facilitate the bridging of the gap.

CONCLUSION
Digital preservation of cultural heritage often focuses on digital objects or digital proxies of physical objects, however museums require significant surrounding infrastructure to fulfil their mission which includes both the core function of search and discovery of information, as well as supporting visitors in education, learning and enjoyment.

In this paper, we have begun to explore the impact of language change in accessing catalogue information. Given that catalogue function is a key element of cultural heritage sites, in aggregate such effects may cause a significant decline in accessibility of information. One proposal to potentially mitigate this might be the use of social sensors and cultural media mining. In this paper, we have proposed that the collection of this kind of information from open data sources could be used to provide a landscape from which to understand and interpret catalogue information. By automatically and periodically collecting this information, in a manner that takes into consideration ethical concerns, it becomes possible for the catalogue/index to model—and therefore take account of—changes in common understanding or usage of language, with a view to cultural context, including specific search terms. This therefore may go some way towards supporting pro-active maintenance of museum infrastructure as it is currently supported for digital object formats. We hope that by doing so, this may begin to address the problems of catalogue accessibility with relation to averting a digital dark age. Such an approach may also contribute to supporting the increasingly extensive objectives of cultural organisations in supporting education, study and enjoyment.

We hope to explore the use of similar linguistically inspired approaches across heterogeneous groups for practical purposes. One such is the support of system development by identifying problematic or specialised jargon. A second is to provide support for detailed analysis of specialised cataloguing approaches and practices, which is of use in the normalisation of catalogue metadata, and in the sociological study of practices in a specialised field.

REFERENCES
1. Paul M. Aoki, Rebecca E. Grinter, Amy Hurst, Margaret H. Szymanski, James D. Thornton, and Allison Woodruff. 2002. Sotto voce: exploring the interplay of conversation and mobile audio spaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 431–438.

2. Baltimore City Archives. 2014. Transcribing and Inventorying the Records of Baltimore City, 1905–1940. (2014). `http://baltimorecityhistory.net/research-at-the-baltimore-city-archives/transcribing-and-inventorying-the-records-of-baltimore-city-1905-1940/`

3. Christine L. Borgman. 1986. Why are online catalogs hard to use? Lessons learned from information-retrieval studies. *Journal of the American society for information science* 37, 6 (1986), 387–400.

4. Christine L. Borgman. 1996. Why are online catalogs still hard to use? *JASIS* 47, 7 (1996), 493–503.

5. Christine L. Borgman. 2003. Personal digital libraries: creating individual spaces for innovation. In *NSF Post-digital library futures workshop*.

6. Marshall Breeding. 2007. Next-gen library catalogs. *Library Technology Reports* (July/August 2007), 10–13.

7. Alison Burrows, Rachael Gooberman-Hill, and David Coyle. 2015. Shared Language and the Design of Home Healthcare Technology. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*.

8. Sebastian Chan. 2007. Tagging and searching: serendipity and museum collection databases. (2007).

9. Daniel Cunliffe, Efmorphia Kritou, and Douglas Tudhope. 2001. Usability evaluation for museum web sites. *Museum Management and Curatorship* 19, 3 (2001), 229–252.

10. Riccardo Dini, Fabio Paternò, and Carmen Santoro. 2007. An environment to support multi-user interaction and cooperation for improving museum visits through games. In *Proceedings of the 9th international conference on Human computer interaction with mobile devices and services*. ACM, 515–521.

11. Michael Factor, Ealan Henis, Dalit Naor, Simona Rabinovici-Cohen, Petra Reshef, Shahar Ronen, Giovanni Michetti, and Maria Guercio. 2009. Authenticity and Provenance in Long Term Digital Preservation: Modeling and Implementation in Preservation Aware Storage. In *First Workshop on on Theory and Practice of Provenance (TAPP'09)*. USENIX Association, Berkeley, CA, USA, 6:1–6:10.

12. John H. Falk and Lynn D. Dierking. 2000. *Learning from museums: Visitor experiences and the making of meaning*. Altamira Press.

13. Silvia Filippini Fantoni. 2006. Web-based solutions: Save it for later. (2006). `http://www.artsprofessional.co.uk/magazine/article/web-based-solutions-save-it-for-later`

14. Silvia Filippini Fantoni and J. P. Bowen. 2007. Bookmarking in museums: extending the museum experience beyond the visit. In *Museums and the Web 2007*, J. Trant and D Bearman (Eds.). Archives and Museum Informatics.

15. Jawaid A. Ghani and Satish P. Deshpande. 1994. Task characteristics and the experience of optimal flow in human-computer interaction. *The Journal of psychology* 128, 4 (1994), 381–391.

16. Peter A. Gloor, Daniel Oster, Ornit Raz, Alex Pentland, and Detlef Schoder. 2010. The virtual mirror: Reflecting on the social and psychological self to increase organizational creativity. *International Studies of Management & Organization* 40, 2 (2010), 74–94.

17. Margaret Hedstrom. 1997. Digital preservation: a time bomb for digital libraries. *Computers and the Humanities* 31, 3 (1997), 189–202.

18. Charles Hildreth. 1987. Beyond Boolean; designing the next generation of online catalogues. *Library Trends* (Spring 1987), 647–67.

19. Mizuko Ito, Kris Gutierrez, Sonia Livingstone, Bill Penuel, Jean Rhodes, Katie Salen, Juliet Schor, Julian Sefton-Green, and S Craig Watkins. 2013. *Connected learning: An agenda for research and design*. Digital Media and Learning Research Hub.

20. Stuart Jeffrey. 2012. A new Digital Dark Age? Collaborative web tools, social media and long-term preservation. *World Archaeology* 44, 4 (2012), 553–570.

21. Terry Kuny. 1998. The digital dark ages? Challenges in the preservation of electronic information. *International preservation news* 17 (1998), 8–13.

22. Aleck C. H. Lin and Shirley D. Gregor. 2006. Designing Websites for Learning and Enjoyment: A study of museum experiences. *The International Review of Research in Open and Distributed Learning* 7, 3 (2006).

23. Aleck C. H. Lin, Shirley D. Gregor, and Michael Ewing. 2008. Developing a scale to measure the enjoyment of web experiences. *Journal of Interactive Marketing* 22, 4 (2008), 40–57.

24. Lynn J. Lohnas and Michael J. Kahana. 2013. Parametric effects of word frequency in memory for mixed frequency lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 39, 6 (2013), 1943–1946.

25. Lev Manovich. 2011. Trending: the promises and the challenges of big social data. *Debates in the digital humanities* (2011), 460–475.

26. Paul F. Marty. 2011. My lost museum: User expectations and motivations for creating personal digital collections on museum websites. *Library & information science research* 33, 3 (2011), 211–219.

27. Bernd Resch, Anja Summa, Günther Sagl, Peter Zeile, and Jan-Philipp Exner. 2015. Urban emotions — Geo-semantic emotion extraction from technical sensors, human sensors and crowdsourced data. In *Progress in Location-Based Services 2014*. Springer, 199–212.

28. Nina Siegal. 2015. Rijksmuseum removing racially charged terms from artworks, titles and descriptions. *New York Times* (2015). `http://nyti.ms/1SQeoEX`

29. Mary Elizabeth Stevens. 1970. Automatic Indexing: A State-of-the-Art Report. (1970).

30. Jennifer Trant. 2009. Studying social tagging and folksonomy: A review and framework. *Journal of Digital Information* 10, 1 (2009).

31. Heleen Van Loon, Kris Gabriëls, Daniël Teunkens, Karel Robert, Kris Luyten, K. Coninx, and E. Manshoven. 2006. Designing for interaction: socially-aware museum handheld guides. *NODEM 06-Digital Interpretation in Cultural Heritage, Art and Science* (2006).

32. Yiwen Wang, Lora M. Aroyo, Natalia Stash, and Lloyd Rutledge. 2007. Interactive user modeling for personalized access to museum collections: The rijksmuseum case study. In *User Modeling 2007*. Springer, 385–389.

33. Giles Waterfield. 2000. The Origins of the Early Picture Gallery Catalogue in Europe, and its Manifestation in Victorian Britain. *Art in Museums* (2000), 42–73.

34. Katie Wilson. 2007. OPAC 2.0: Next generation online library catalogues ride the Web 2.0 wave! *Online Currents* 21, 10 (2007), 406.

35. Rong Zeng and Patricia M. Greenfield. 2015. Cultural evolution over the last 40 years in China: Using the Google Ngram Viewer to study implications of social and political change for cultural values. *International Journal of Psychology* 50, 1 (2015), 47–55.