

# On social media and document theory: a exploratory and conceptual study

---

## **General outline of the study**

There is a growing body of research that accentuates the importance of employing a socially, culturally, and linguistically oriented analytic framework when studying information and knowledge (Duguid & Brown, 2000; Talja et al, 1999; Hjørland & Albrechtsen, 1995; Frohmann, 2004a; Hjørland, 2011). A wide variety of contextually oriented approaches have been employed in the study of equally diverse subject domains, including scholarly communication (Francke, 2010), the production of scientific knowledge (Knorr Cetina, 2000; Bazerman, 1985), professional information seeking (Sundin & Johannisson, 2007), and information retrieval (Hjørland, 2002). There are, however, few conceptual and empirical inquiries in LIS that manifest such a contextual credo in the study of the relationship between informational aspects of social media and the conditions of documentation, archiving, and cultural heritage.

The goal of the present study is to contribute to contextually oriented, social media research by discussing how blogs about three-dimensional virtual worlds, from a document theory perspective, can be conceptualized as documentation of the virtual worlds to which they pertain. The document theoretical view of blogs is further analyzed from a joint LIS and postmodern archival science perspective, where the archival notion of record is used to explore virtual world blogs as a potential information source in archiving efforts directed towards virtual worlds. An empirical case study of World of Warcraft blogs (described below) is used as an example of how document theory and the concept of records in postmodern archival science can provide a fruitful social constructivist analytic framework in this particular area of research. In order to crystallize the above stated research aims, the essential theoretical underpinnings of the inquiry will be briefly presented in the following section.

## **A short introduction to document theory, the concept of record, and the theory of remediation**

Document theory – pioneered by Otlet (1868-1944) and Briet (1894-1989); modified and reintroduced into LIS research by scholars such as Frohmann (2004b), Francke (2008), and Windfeld Lund (2009, 2010) – is a contextually cognizant approach that accentuates the material, and consequently the contextual and processual, nature of the document (Windfeld Lund, 2009, 2010). What should be considered a document is a topic that has been discussed extensively in both historical and contemporary academia (cf. Buckland, 1997, 1998; Francke, 2005; Frohmann, 2009). However, strong reasons to define the concept of document using a heuristic approach has been put forward by, among others, Francke (2005) and Frohmann (2009). For the purpose of this study, Winfeld Lunds definition of documents as “any results of human efforts to tell, instruct, demonstrate, teach or produce a play [...] by using some means in some ways“ will be employed (Winfeld Lund, 2010, p. 743). In this view, it is plausible to conceptualize blogs as documents.

From a document theory perspective, blogs are thus essentially cultural, social, and historical constructs that are shaped by the situated document practices that create them (Francke, 2008; Windfeld Lund, 2009, 2010). Coupled with a view of information as a fundamentally social

phenomenon, mediated through people, artefacts, and technologies (Francke, 2008), the study of blogs, from a document theory point of view, becomes a reasonable way of investigating how virtual world blogs remediate and represent the virtual worlds which they are about. The theory of remediation will be used as a supplementary theoretical tool to understand and analyze the relation between the virtual world and blogs about it, and how one represents, remediates, and reforms, the other (Bolter & Grusin, 1999).

The attempt to understand and explore the social nature of documents – a central tenet of document theory – resonate with the notion of record, a fundamental concept in archival science as well as one of the most basic units of the archive. In traditional archival thought the record is a neutral object that provides authentic information about the context of its creation (Thomassen, 2001). However, postmodern archival science asserts that the record is socially constructed and, furthermore, accentuates the need to develop new archival frameworks and practices on the basis of this viewpoint (Cook, 2001; Trace, 2002).

## **Case study – methods and materials**

The document-centered theoretical framework described above was operationalized in a case study of how the landscape alternations of the original continents of World of Warcraft's game-world – brought about by the release of the expansion pack Cataclysm in December 2011 – was represented in 156 World of Warcraft blogs. How were the drastic changes to the familiar surroundings of the game-world, hitherto unchanged since the release of World of Warcraft in 2004, represented in the blogs? Which representations of the landscape metamorphosis were frequent in the World of Warcraft blogosphere, and which were uncommon? How were the representations interrelated?

The case study was guided by the following research questions: (i) what are the characteristics of and variations in the remediation and representation of the virtual world World of Warcraft in the studied World of Warcraft-related blogs? (ii) How does research-based knowledge of the remediation and representation of World of Warcraft in the studied blogs inform a working approach to how virtual worlds could be archived?

The World of Warcraft blogs that form the empirical basis of the present case study was found using the blog search engines Technorati and Blogpulse, and the blogrolls of the blogs identified this way. All of the blogs were written in English and active at the time of data collection. Relevant blog posts were identified using the blogs' built-in search tools and folksonomic tags, and selected for analysis using purposive sampling (Wildemuth, 2009). The data was analyzed using co-word analysis, a content analysis-method designed to measure the frequency and co-occurrence of linguistic units (words, pairs of words, phrases) in a body of text (Courtial, 1994; Callon, et al, 1983; He, 1999). Multidimensional scaling (MDS) was used to visualize the interrelations between the indexed terms in a two-dimensional network maps suitable for qualitative interpretation (Holmberg, et al., 2009). The analysis was performed using BibExcel, an infometric computer software tool written especially for this purpose (Persson, 2006). Co-word analysis was chosen as the method of data analysis because it allows for efficient identification, quantification, and interlinkage of the dominant modes of linguistic expressions – representations – in a large set of texts, as well as an easily accessible way to synthesize and present these results in a graphical representation well-suited for qualitative analysis.

## Findings and contributions

The present inquiry, consisting of a case study and a first attempt to establish a document theory-centered analytical framework for social media research, yields both conceptual contributions and empirical findings. Conceptually, the study contributes to the discussion of how virtual world blogs – employing document theory and a postmodern archival science view of the concept of records – can be conceptualized as documentation of the virtual worlds to which they pertain. Furthermore, the study stresses the importance of future research that continues to explore how informational dimensions of social media relate to questions of documentation, archiving, and cultural heritage. Additionally, it would be of great interest to investigate how the abstract, research-based, insights of this kind can be operationalized in order to develop tools and frameworks to be used in the archiving of virtual worlds.

Empirically, the results of the case study of World of Warcraft blogs indicate that there are distinct regularities in the frequencies and relationships of the concepts used to represent the changes made to the World of Warcraft game-world. These results support earlier observations (Yeo, 2007, 2008, forthcoming) of the highly heterogeneous documentary qualities of informal forms of communication such as blogs. The range of topics, sentiments, features, activities, and events represented in the studied World of Warcraft blogs was very unevenly distributed. The present study is a part of a doctoral research effort; in future inquiries I will seek to determine the information practices of blog creation, with the aim of illuminating the underlying social factors that directly influence and shape the creation of blogs, and explore how these factors manifest themselves in their construction, in order to further this area of study.

## References

- Bazerman, C. (1985). Physicists reading physics: schema-laden purposes and purpose-laden schema. *Written Communication*, 2(1):3-23
- Bolter, J.D. & Grusin, R. (2000), *Remediation: understanding new media*, MIT Press, Cambridge, MA
- Buckland, M. (1997). What is a “document”? *Journal of the American Society for Information Science*, 48(9):804-809
- Buckland, M. (1998). What is a “digital document”? *Document Numérique*, 2(2):221-30
- Callon, M., Courtial, J.-P., Turner, W.A. & Bauin, S. (1983). From translations to problematic networks: an introduction to co-word analysis. *Social Science Information*, 22(191)
- Cook, T. (2001). Archival science and postmodernism: new formulations for old concepts. *Archival Science*, 1(1):3-24
- Courtial, J.P. (1994). A co-word analysis of scientometrics. *Scientometrics*, 31(3):251-60
- Duguid, P. & Brown, J.S. (2000). *The social life of information*. Boston, Mass.: Harvard Business School
- Francke, H. (2005). What’s in a name? Contextualizing the document concept. *Literary and Linguistic Computing*, 20(1):61-69
- Francke, H. (2008). (Re)creations of scholarly journals: document and information architecture in open access journals. Borås: Valfrid. Diss.

- Frohmann, B.P. (2004a). *Deflating information: from science studies to documentation*. Toronto: University of Toronto Press
- Frohmann, B. (2004b). Documentation redux: prolegomenon to (another) philosophy of information. *Library Trends*, 52(3):387-407
- Frohmann, B. (2009). Revisiting “What is a document?”. *Journal of Documentation*, 65(2):291-303
- He, Q. (1999). Knowledge discovery through co-word analysis. *Library Trends*, 48(1): 133-159
- Hjørland, B. & Albrechtsen, H. (1995). Toward a new horizon in information science: domain-analysis. *Journal of the American society for Information Science and Technology*, 46(6):400-425
- Hjørland, B. (2002). Epistemology and the socio-cognitive perspective in information science. *Journal of the American society for Information Science and Technology*, 53(4):257-270
- Hjørland, B. (2009). Concept theory. *Journal of the American society for Information Science and Technology*, 60(8):1519-1536
- Hjørland, B. (2010). The foundation of the concept of relevance. *Journal of the American society for Information Science and Technology*, 61(2):217-237
- Hjørland, B. (2011). The importance of theories of knowledge: indexing and information retrieval as an example. *Journal of the American society for Information Science and Technology*, 62(1):72-77
- Holmberg, K., Huvila I., Kronqvist-Berg, M. & Widén-Wulff, G. (2009). What is Library 2.0?. *Journal of Documentation*, 65(4):668-681
- Johannisson, J. & Sundin, O. (2007). Putting discourse to work: information practices and the professional project of nurses. *Library Quarterly*, 77(2):199-218
- Kjellberg, S. (2010). *Scholarly blogs: scholarly communication and knowledge production in the blogosphere*. Lunds Universitet. Diss.
- Knorr Cetina, K.D. (1999). *Epistemic cultures: how the sciences make knowledge*. Cambridge, Mass.: Harvard University Press
- Persson, O. (2006). BibExcel – a bibliometric toolbox. Retrieved 23/02/2011 at [www.umu.se/inforsk/Bibexcel/](http://www.umu.se/inforsk/Bibexcel/)
- Talja, S., Keso, H. & Pietiläinen, T. (1999). The production of context in information seeking research: a metatheoretical view. *Information Processing and Management*, 35:751-763
- Thomassen, T. (2001). A first introduction to archival science. *Archival Science*, 1:373-385
- Trace, C.B. (2002). What is recorded is never simply “what happened”: record keeping in modern organizational culture. *Archival Science*, 2:137-159
- Wildemuth, B.M. & Cao L.L. (2009). Sampling for intensive studies, in: Wildemuth, B.M. (2009). *Applications of social research methods to questions in information and library science*, Westport, Conn.: Libraries Unlimited

Winfeld Lund, N. (2009). Document theory. *Annual Review of Information Science and Technology*, 43:399-432

Windfeld Lund, N. (2010). Document, text and medium: concepts, theories and disciplines. *Journal of Documentation*, 66(5):734-749

Yeo, G. (2007). Concepts of Record (1): evidence, information, and persistent representations. *American Archivist*, 70(2):315-343

Yeo, G. (2008). Concepts of Record (2): prototypes and boundary objects. *American Archivist*, 71(1):118-143

Yeo, G. (forthcoming). Rising to the level of a record? Some thoughts on records and documents. *Records Management Journal*