

Words and Emotional Work: Classification Theory's Constructs Useful for the Analysis of Social Media Data in Terms of Gender, Race, and Sexuality

Joseph T. Tennis¹

¹University of Washington

Abstract

Presents constructs from classification theory and relates them to the study of hashtags and other forms of tags in social media data. Argues these constructs are useful to the study of the intersectionality of race, gender, and sexuality. Closes with an introduction to an historical case study from Amazon.com.

Keywords: Indexing; Classification Schemes; Semantics; Ethics

Citation: Tennis, J. T. (2015). Words and Emotional Work: Classification Theory's Constructs Useful for the Analysis of Social Media Data in Terms of Gender, Race, and Sexuality

Copyright: Copyright is held by the authors.

Contact: jtennis@uw.edu

1 Words and Work – Emotional and Otherwise

Classification theorists are interested in the semantics and structure of what we call *indexing languages*. These languages are used to identify characteristics of documents; once identified the documents can be placed into sets that share characteristics by a searcher. Indexing languages are either designed or emerge from tagging practices. The latter is most common in social media and results in folksonomies, tag clouds,

In the context of microblogging, social networking sites, and comment systems we see hashtags used to tag tweets, images, videos, wall posts, etc. From a classification theory perspective hashtags carry similar traits to terms used in other indexing languages. Because of that, it seems reasonable to consider what might be useful in the analysis of social media tags that has proven useful in classification theory. To that end, this paper starts that comparison. Drawing on three major areas of classification theory work: semantics, structure, and ethics. In these, and at their intersection, we can explore how words and emotional work manifest in indexing languages employed by social media.

2 Semantics, Structure, and Ethics in Classification Theory

In classification theory we are concerned with words, the relationship that obtains between words, and the resources those words represent. For the purposes of this paper we can talk about these as the semantics, structure, and textual aura of indexing languages.

Semantics in indexing languages is defined by three things: its extension, intension, and textual aura. This is divergent from common linguistic debates of semantics where pragmatics, network theory, and cognitive science prevail. The extension of a word is the number of other words it covers. This is a hierarchical notion. The word cat might have an extension of Maine Coon, Jaguar, and Ocelot. We understand the meaning of the cat in this context because of what it contains. The intension of a word are the characteristics it contains. So cat might be made up of biological characteristics of been a chordate, or it could be a made up of cultural characteristics of being a domestic pet or a agile predator. Finally the textual aura is the texts associated with a particular word in the system. In the case of a library catalogue it would be all the resources (books, archival materials, etc.) that have cat (or more likely written in the form cats) associated with it. We would understand the meaning of cat in this particular system based on the resources assigned to that word.

Semantics, as so defined, opens up hashtags to a particular form of analysis, especially in relation to gender, race, and sexuality. We can take a hashtag like #transphobia and look at a particular system and ask: what hashtags are broader or narrower in this system? What seem to be the characteristics of #transphobia as a term in this system? And what resources are assigned the hashtag #transphobia? The concept of intersctionality surfaces immediately in upon asking these questions. Is #transphobia linked to broad concepts of social justice? Is #transphobia nowhere linked to #feminism? What are the nature of resources linked to #transphobia and #blacklivesmatter?

Structure is another concern for classification theory that may also be a concern for analysis of social media. Structure is the nature of the *binding* between words in the system. Is this binding loose or tightly constrained? Is the hierarchy deep or shallow? While structure was originally conceived with an

eye to helping the user of information systems, strict hierarchical structure has been linked to privilege and other forms of patriarchal power structures (Olson, 2001; Olson 2007; cf., Tennis, 2012). That is, because we impose a structure on our words in an information system we make clear our priorities and those priorities reinforce the privileged understanding of the world. This means the information system reproduces and reinforces this systemic privilege. Further, infrastructure studies, which lays bare the invisible structures of information systems (writ large) grows out of the investigation of classification system structures and indexing language structures (Star, 1998; Bowker and Star, 1999). In the context of social media and gender, race, and sexuality, infrastructure is residual, installed over a large space, and only visible on breakdown. So we might ask: what words are in social media systems? What are not? What terminology that is present is harmful, which is emancipatory? What is easily changed and what is not? Is the panel of choices now available in Facebook a structure that escapes criticism because it offers a range of gender expressions? Is social media a space to problematize race because of flexibility or does the long tail of tags serve to hide variation in expression?

Because classification theory is focused on semantics and structure, and because the kinds of questions listed above pertain to catalogues, indexing and abstracting databases, and other tools used for retrieval and sense-making we have been preoccupied with the ethical dimension of our work. If the problems of naming and the power to position concepts in relation to one another is in the hands of the classificationist, that person must act with intention (Mai 2013). Further, while questions of gender, race, and sexuality materialize as conceptual questions, they are also fundamentally ethical questions about how people treat people. Even if we mediate behavior through online presence, and delay that interaction through asynchronous techniques of logging on, writing, reading, and logging off, we are still acting with hashtags and with resources in a social space.

To date the ethical dimension of classification theory has been cataloguing the problems that surface. A vocabulary for talking about these issues has surfaced, as has a cross-cultural investigation into issues within different borders and contexts (Pinho and Guimerães, 2012). What has surfaced is a tension for revision versus preservation of semantics and structure in systems (Tennis 2012). It is marked by questions like: what should change? Who would change it? What remains as a marker of a particular time and context? What benefit is gained from such preservation?

Wikipedia, where titles are URIs and where categories are assigned and removed on some articles at a rapid pace is a good example of social media's theatre of this issue. In order to faithfully represent what one editor might see as the reality of the situation, there may be an argument over whether to change the title of a page after a person transitions from male to female, from man to woman. The category system may be viewed as a place to highlight what would be otherwise hidden in the article.

3 Contributing toward a Vocabulary of Words and Work in the Intersectionality of Gender, Race, and Sexuality in Social Media

This paper offers one view on how we might talk about intersectionality of gender, race, and sexuality in social media. In this way it can be considered a methodology paper. We can draw from classification theory to talk about hashtags and their semantics, structure, as well as the ethics of their deployment in information systems. We can highlight where gender, race, and sexuality intersect and where they do not. From this we can account for the words used in social media systems, and gain some insight into the emotional work these words are doing.

4 Case Study: An Ann Coulter Book on Amazon.com 2007-2010

In an effort to demonstrate the value of this kind of analysis in social media will present case studies of book tagging on Amazon.com. For the sake of time I only present one case. The Amazon.com tagging case spans three years and is screen captures of tags associated with the Ann Coulter book, *If Democrats Had Any Brains They'd Be Republicans*. I will use the constructs listed above to analyze how gender, race, and sexuality play out in the *social* space. I will also use this case study to show where this kind of analysis can inform classification theory, in turn.

5 References

Bower, G., and Star, S. L. (2001). *Sorting Things Out: Classification and Its Consequences*. (MIT).

Mai, J-E. Ethics, values, and morality in contemporary library classifications. *Knowledge Organization*, 40 (3): 242-253. 2013.

Olson, H. A. (2001). Patriarchal Structures of Subject Access and Subversive *Techniques* for Change. In *The Canadian Journal of Information and Library Science* 26(2/3): 1-29.

Olson, H. A. (2007). How We Construct Subjects: A Feminist Analysis. In *Library Trends* 56(2): 509-541.

Pinho, F. A. ; Guimarães, J.A.C. . Male homosexuality in Brazilian indexing languages: some ethical questions. *Knowledge Organization*, v. 39, p. 363-369, 2012.

Star, S. L. (1998). Ethnography of InfrastructurePreview the documentView in a new window. *American Behavioral Scientist* 43(3): 377-391.

Tennis, J. T. (2012). "A Convenient Verisimilitude or Oppressive Internalization? Characterizing the Ethical Arguments surrounding Hierarchical Structures in Knowledge Organization Systems." In *Knowledge Organization*. 39(5): 394-397.