Elizabeth Milonas (emilonas@citytech.cuny.edu)
Long Island University, Brookville NY USA

Wittgenstein and Web Facets

Abstract: The World Wide Web has grown exponentially in the last few years. The popularity of Web search engines has also grown in a similar manner. The task of a Web search engine is to provide the Web searcher with accurate and targeted information from the plethora of information available on the Web. This is a daunting task that requires the careful usage of language to ensure accuracy. As a result, the importance of the usage and meaning of language in the Web domain has become the focus of recent research. In this paper, the author will explore Wittgenstein's later philosophy of language as it applies to the language used in the search result pages of a Web search engine in an effort to broaden the understanding of language usage within this domain.

1.0 Words and the Web

In recent years the growth of the World Wide Web has prompted a refocusing on the importance of the usage and meaning of words, specifically in the search process where words play an integral part in the retrieval and exploration of information. In the Web domain, Web search engines are the main tools used to retrieve information. The retrieval process is accomplished by the use of a "search term(s)," a word or group of words typed into the search box of the particular search engine. The ineffective use of words can provide either too much irrelevant information or too little information (Chekuri, Goldwasser, Raghavan, and Upfal 1997). The effective use of words yields information that is targeted specifically to the initial search request.

Within the result pages of a Web search engine, words play a vital role in expanding the understanding of the concept presented by the search term. Web search engines generate words that appear in the search result pages typically under the "Related Search" section. These words, often called Web facets, represent different aspects of the search term and are frequently used to explore further the search topic. Web facets are evident in the search result pages of all the search engines found on the Web. The purpose of this paper is to broaden the understanding of language usage within the search result pages of Web search engines. Wittgenstein's philosophy of language, presented in his later work *Philosophical Investigations* (1953), will be explored as it applies to the language used to represent Web facets in the search result pages of a Web search engine. The findings of a case study in which the language used to denote Web facets was examined for evidence of Wittgenstein's later theories of forms of life, language games, and family resemblances will be discussed.

2.0 Wittgenstein's later philosophy of language

Wittgenstein in his *Philosophical Investigations* (1953) posits that the meaning of a word is independent of the existence of the object that it represents (p. 28e). Meaning is multi-dimensional and is based on the use of the word (p. 4e-5e) that can be as diverse as the use of each tool in a toolbox (p. 6e). Meaning in language emerges from ordinary usage of language, that is, from the day-to-day discourse (Blair 2003). A word can have several meanings depending

on its role in different "forms of life" (2003). "Forms of life" are the activities and practices in which individuals participate (2003). Within these "forms of life," patterns of word usage (language games) develop that relate to the specific form of life (2003). For example, depending on the form of life, the word "pitch" can mean the slope of a roof, the intonation of a voice, a specific action in a baseball game, a salesperson's convincing dialogue (2003). Groups of words or phrases also adhere to the forms of life theory. For example, the phrase "five slabs" can mean the command "Bring me five slabs!" or it may be used to indicate a numeric count, "There are five slabs" (Wittgenstein 1953, p. 10e). In addition, words that look different can have similar meaning if they are used within the same form of life. For example, the word "slab" and the word "block" are two different words that share similar meaning when used in relation to the activity of building or construction (p. 8e). Understanding the meaning of a word is a two-step process. First one must be able to use the word correctly in the activity (form of life) in which it is normally used, and second, one must know how to use the word (the language game) (Blair 2003).

The meaning of words changes as the forms of life shift and change (Andersen and Christensen 1999) and therefore words are not defined by one true meaning, rather they are grouped together into families (Mazzocchi and Tiberi 2009). As members of a family, they possess family resemblance but they do not share a set of traits or characteristics that are common to all the members of the family (2009). Within families, words form a complex network of interweaving and overlapping relationships (Wittgenstein 1953, p 32e). For example, chess, tennis, baseball, solitaire and the like, are all members of the family "games." However, the concept of "games" is difficult to define (Taivalsaari 1997). Some games involve amusement, some involve competition while others do not, some involve luck, some involve skill and some involve both luck and skill (1997). Some games involve one player such as solitaire, while other games involve thousands and thousands of players such as the lottery or horserace betting (1997). "Games" do not share any common defining characteristics; instead they share a family resemblance; that is, there is no single collection of properties shared by all the members of the family (1997). Wittgenstein believed that language itself is intricate and complicated (1953, p. 8e), bewitching its user (p. 47e) into relying on the object or the "picture" for meaning and blinding him from the real meaning, that which is based on use. The "essence of the matter" is not in "focus" (p. 48e) and the user of the language may think that he is "tracing the outline of the thing's nature over and over again" when in actuality, he is "merely tracing round the frame through which [he] look[s] at it" (p. 48e). True meaning can only be derived from use in a particular form of life.

3.0 The role of facets in the Web domain

The applications and benefits of facets and facet-classification have been widely explored in relation to the organization and retrieval of information within the Web domain. Research has been generated in the area of facet-enhanced design and organization of corporate and Library and Information Science (LIS) websites (Franklin 2003, Mills 2004, Broughton 2006, La Barre 2006, Uddin 2007, Uddin and Janecek 2007, Capra et al. 2007, Crystal 2007, and Wilson and Mc Schraefel 2008). Research has also been generated on the effects of implementing the concepts of facets and faceted classification methods in the information retrieval processes within the corporate and LIS websites (Broughton 2001, Yee et al. 2003, Kules, Kustanowitz, and

Shneiderman 2006, Hong 2006, Gnoli and Mei 2006, Broughton and Slavic 2007, and Capra et al. 2007).

The design and organization of information on a website is a difficult task (Uddin 2007), however, it is a task that warrants careful attention. Website information that is well organized and easily accessible helps maximize the website search process. In recent years, designers and information architects have been incorporating faceted classification methods (La Barre 2006) into their website designs in an effort to improve website organization and accessibility. The incorporation of faceted classification methods has been implemented through the physical display of "building blocks" (Hjørland 2008) or "facet elements" (Bates 1988). These facets are often displayed in a drop-down, windows-based method that has been proven effective for displaying information given that Web searchers are already familiar with drop-down techniques (Broughton 2006) from using windows tools. From these drop-down displays, Web users choose a combination of facets that help them navigate through website information. Research has shown that the utilization of facets can be effective (Franklin 2003 and La Barre 2006) in improving this navigation process and ultimately improving the query process (Crystal 2007, Uddin and Janecek 2007, Wilson and Mc Schraefel 2008 and Milonas 2010). One important aspect of facet-enhanced website design is the incorporation of facets that are valuable to the user. This is accomplished by the careful study of user's varied search approaches (Capra et al. 2007, and Uddin and Janecek 2007). Websites that benefit the most from facet-enhanced design are those that display information that is not easily placed into rigid (Franklin 2003) hierarchies. Research conducted in this area has shown that the use of facets in these type of websites can facilitate the organization of information into logical concepts allowing the Web user to cover a greater segment of the information (Mills 2004) found on the website.

In the area of information retrieval, the implementation of facets and faceted classification methods is beneficial to Web searchers especially to those who are unfamiliar (Kules, Kustanowitz, and Shneiderman 2006, Yee et al. 2003, and Capra et al. 2007) with the search topic. Faceted classification methods facilitate the search and retrieval process (Milonas 2010) by presenting the Web searcher with categorized overviews (Kules, Kustanowitz, and Shneiderman 2006) of the search results. These overviews enable the Web searcher, who is unfamiliar with the topic, to easily browse through the categories. Research has proven that facet-enhanced searches are a powerful means of retrieving complex and multidimensional (Broughton 2001, Hong 2006, and Broughton and Slavic 2007) context since they allow Web information to be accessed from different points of access (Broughton and Slavic 2007). There is consensus that this method is effective in accessing Web information (Broughton and Slavic 2007, Hong 2006, and Gnoli and Mei 2006) that has multi-dimensional properties. Within the search result pages of Web search engines, Web facets are used to expand the Web searchers knowledge of the search topic. These Web facets, often displayed within the "Related Searches" section of the search result pages, are similar to traditional facets in that they represent various aspects of the search topic. However, unlike traditional facets, Web facets do not represent modes of controlled vocabulary. Relatively few studies have been conducted on the effects and benefits of Web facets.

4.0 Methodology

A case study was conducted in which Web facets displayed as a result of a search using the term "wine" in the Google search engine were examined for evidence of Wittgenstein's later philosophy of language (search conducted on 11/07/10).

4.1 Procedure
4.1.1 Displaying Web facets in the Google search result pages
Web facets are not immediately visible to the Web searcher once the search term has been executed in the Google search engine. A list of Web facets can be made visible if the "Related Search" option is invoked by first choosing the option labeled "More Search Tools" that appears on the left side of the search result page and then by choosing the "Related Searches" option. Once this option is chosen, a list of Web facets related to the search term appears beneath the Google search box. Only the Web facets that appeared on the first page of the search results were examined.

4.1.2 Executing the search term "wine"
The search term "wine" was randomly chosen and typed into the Google search box and the search results page was displayed. Upon choosing the "Related Searches" option, nineteen Web facets appeared directly below the Google search box. These nineteen Web facets were examined for this case study.   The Web facets were; types of wine, wine 101, red wine, wine download, online wine, wine enthusiast, wine reviews, wine ubuntu, wine pairing, wine gifts, alcohol, beer, bottle, liquor, merlot, champagne, vineyard, cabernet, pinot noir.

Upon close examination of these Web facets, common topic areas emerged under which the Web facets could be grouped. The topic areas were loosely based on Ranganathan's five fundamental categories (PMEST) of personality, matter, energy or action, space and time (Ranganathan 1951, p. 101). Table 1 below displays the common topic areas and the corresponding Web facets:

Search term: Wine
Type (role, color, type, taste, brand):
Oenology: type of wine, red wine, merlot, champagne, cabernet, pinot noir, liquor
Computer program: wine ubuntu
Space (location including; country, region, , or storage location): vineyard, bottle, wine download
Education (classes, instruction, recipes, literature): wine 101, wine reviews
Social (culture, organizations, clubs, people, relationships): wine enthusiast, wine pairing
Retail (buy, sell, products, stores): wine gifts, online wine
Elements (ingredients, components): alcohol
Other (related): beer

Table 1: Common topic areas and corresponding Web facets

The "type" topic area has two sub-topic areas: "oenology" and "computer programming." The Web facets listed under these sub-topic areas describe the term "wine" in reference to the role it plays, the color, kind, taste and brand. The Web facets listed under the topic area "space" describe the term "wine" in reference to location, for example, the country, region, and the storage location. The Web facets listed under the topic area "education" describe the term "wine" in reference to formal classes, instructions, recipes and literature. The Web facets listed under the "social" topic area describe the term "wine" in reference to culture, social organizations such as clubs, people and other relationships concerning the term. The Web facets listed under the topic area "retail" describe the term "wine" in reference to the processes of buying and selling as well as retail stores and related products. The Web facets listed under the topic area "elements" describe the term "wine" in reference to its ingredients or components. The "other" topic area includes only one Web facet whose meaning is ambiguous and as a result cannot be grouped under any of the categories identified.

5.0 Analysis and discussion
5.1 Family resemblances and forms of life

The Web facets returned as a result of the Google search using the search term "wine" lack a single common set of characteristics. This is evident of all the Web facets returned both within and across the topic areas: oenology, computer programming, location, education, social or community, retail, elements or ingredients and other. These Web facets possess unique sets of characteristics but lack a shared common set of characteristics (Talvassari 1997). Like members of a family, they show some family resemblances however; there is no single collection of traits that can be said to be common to all. In addition, each topic area and sub-topic area under which the nineteen Web facets were grouped can be considered a form of life. These forms of life include: oenology, computer

programming, education, location, social or community, retail, ingredients and other. The Web facets displayed draw their meanings as a result of their participation in these particular forms of life.

6.0 Conclusion

The findings of this case study seem to indicate that the Web facets displayed as a result of the Google search using the search term "wine" exhibit aspects of Wittgenstein's later philosophy of language. This is evident in the discovery of forms of life and family resemblances. The nineteen Web facets appear to enhance the meaning of the initial search term by providing different vantage points (forms of life) expanding the Web searcher's understanding of the concepts presented by the search term at a specific point in time. As the forms of life change over time, it can be assumed that new Web facets will be presented embodying meaning relevant to the new forms of life and reflecting new language games. Further research is needed to explore these initial finding. In the future, a longitudinal study of multiple Web search engines and varying search terms will be conducted to investigate further the application of Wittgenstein's later philosophy of language to Web facets and ultimately broadening our understanding of Web facet utilization within this domain.

Acknowledgements

References

Andersen, Jack and Frank Sejer Christensen. 1999. Wittgenstein and indexing theory. Proceedings of the 10th ASIG SIG/CR Classification Research Workshop. Washington D.C. October 31, 1999.

Bates, Marcia. 1988. How to use controlled vocabularies more effectively in online searching. Online 12: 45-56.

Blair, David C. 2003. Information retrieval and the philosophy of language. *Annual Review of Information Science and Technology* 37:3-50

Broughton, Vanda. 2001. Faceted classification as a basis for knowledge organization in a digital environment; the Bliss Bibliographic Classification as a model for vocabulary management and the creation of multi-dimensional knowledge structures. New review of hypermedia and multimedia 7:67-102.

_____ . 2006. The need of a faceted classification as the basis of all methods of information retrieval. ASLIB proceedings: New information perspectives 58:49-72.

_____ , and Aida Slavic. 2007. Building a faceted classification for the humanities: Principles and procedures. Journal of documentation 63:727-54.

Capra, Robert, Gary Marchionini, Jung Sun Oh, Fred Stutzman, and Yan Zhang. 2007. Effects of structure and interaction style on distinct search tasks. Joint conference on digital libraries –JCDL (Vancouver: ACM and IEEE) 442-51.

Crystal, Abe. 2007. Facets are fundamental: Rethinking information architecture framework. Technical communication 54:16-26.

Franklin, Rosemary Aud. 2003. Re-inventing subject access for the semantic web. Online information review 27:94-101.

Gnoli, Claudio, and Hong Mei. 2006. Freely faceted classification for Web-based information retrieval. New review of hypermedia and multimedia 12:63-81.

Hjørland, Birger. 2008. Facet, facet analysis and the facet-analytic paradigm in knowledge organization (KO). Lifeboat for knowledge organization. Available at

Hong, Mei. 2006. Potential usage of faceted classification in Internet "information retrieval". Interdisciplinary information sciences 12: 43-51.

Kules, Bill, and Ben Shneiderman. 2005. Using meaningful and stable categories to support exploratory web search: Two formative studies. http://cgis.cs.umd.edu/localphp/hcil/tech-reports-search.php?number=2005-31 (Accessed December 5, 2009)

_____ . 2008. Users can change their web search tactics: Design guidelines for categorized overviews. Information processing and management 44:463-84.

Kules, Bill, Jack Kustanowitz, and Ben Shneiderman. 2006. Categorizing web search results into meaningful and stable categories using fast-feature techniques. Joint conference on digital libraries –JCDL (North Carolina: ACM and IEEE).

La Barre, Kathryn. 2006. The use of faceted analytic-synthetic theory as revealed in the practice of website construction and design. Submitted in partial fulfillment of the requirements of the degree Doctor of Philosophy in the School of Library and Information Science, Indiana University.

Mills, Jack. 2004. Faceted classification and logical division in information retrieval. Library trends 52:541-70.

Uddin, Mohammad Nasir. 2007. The implementation of faceted classification in web site searching and browsing. Online information review 31:218-33.

Uddin, Mohammad Nasir, and Paul Janecek. 2007. Performance and usability testing of multidimensional taxonomy in web site search and navigation. Performance measurement and metrics 8:18-33.

Wilson, Max L., and Mc Schraefel. 2008. A Longitudinal study of exploratory and keyword search. Joint conference on digital libraries –JCDL (Pittsburgh: ACM and IEEE) 52-5.

Yee, Ka-Ping, Kirsten Swearingen, Kevin Li, and Marti Hearst. 2003. Faceted metadata for image search and browsing. In proceedings of the 2003 conference on human factors in computing system - CHI '03 (New York: ACM) 401 – 8.

Chekuri, C., Goldwasser, M. H., Raghavan, P., and Upfale, E. 1997. Web search using automatic classification. *In Proceedings of the Sixth International Conference on the World Wide Web* (Santa Clara: California) April.

La Barre, Kathryn. 2006. The use of faceted analytic-synthetic theory as revealed in the practice of website construction and design. Submitted in partial fulfillment of the requirements of the degree Doctor of Philosophy in the School of Library and Information Science, Indiana University.

_____. 2008. Discovery and access systems for website and cultural heritage sites, reconsidering the practical application of facets. *Advances in Knowledge Organization.* 11:105-10.

Mazzocchi, Fulvio and Melissa Tiberi. 2009. Knowledge Organization in the Philosophical domain: Dealing with polysemy in the thesaurus building. *Knowledge Organization* 36: 103-12.

Milonas, Elizabeth. 2010. The use of facets in Web search engines. *In paradigms and conceptual systems in knowledge organization* (February 23 - 26), pp. 349-355.

Park, Ok nam. 2008. Opening ontology design: A study of the implications of Knowledge Organization for ontology design. *Knowledge Organization* 35:209-21.

Ranganathan, S.R. 1951. Colon Classification and its approach to documentation. In Jesse H. Shera, and Margaret E. Egan. *Bibliographic organization.* Chicago: The University Press of Chicago.

_____. 1962. Elements of library classification. India: Asia Publishing House.

Svenonius, Elaine. 2004. The epistemological foundations of knowledge representation. Library Trends 52:571-87.

Wittgenstein, Ludwig. 1953. *Philosophical Investigations*. Translated by G. E. B. Anscombe. Oxford: Blackwell Publishers Inc.