Visualizing Human Factors and Ergonomics Publications: Word clouds and Word networks

John D. Lee
Department of Industrial and Systems Engineering
University of Wisconsin—Madison

Objective: This paper explores the strengths and limits of word clouds and word networks in visualizing the field of Human Factors and Ergonomics (HFE). **Background:** Large volumes of textual data present unique analysis and visualization challenges. Visualization techniques must balance the need to make graphics engaging, accessible, and informative. **Method:** The analysis considers 11,911 abstracts from two decades of papers published in 11 HFE journals. A common graphical representation of text data—word clouds—is compared with word networks. **Results:** Word clouds and word networks both highlight important terms of HFE, but word networks provide additional information regarding the sematic structure of the field. Restyling what has become known as the "mullet of the internet" could make a valuable contribution to analyzing textual data. **Conclusion:** Like other visualization techniques that convey complex relationships, word networks offer more insight than word clouds, but at the cost of being less accessible than simple word clouds. **Application:** Word networks offer a promising alternative to word clouds by providing a more complete view of textual data.

INTRODUCTION

Human Factors and Ergonomics (HFE) is a broad field with contributions ranging from applied psychology to biomechanics. Broadly concerned with designing systems to make technology work for people, HFE addresses many aspects of technology and human behavior. Consequently, thousands of papers from many journals contribute to the field. This volume of papers can bewilder students trying to understand the field. The many papers can also make it difficult for experienced researchers to discern links between apparently disparate, but actually related research. Understanding the increasing volume of publications has become urgent as the pace of technology change accelerates.

Journals archive the field and its constituent topics, but the thousands of articles make a close reading of the field challenging. Text analysis and associated visualization tools may offer an effective way for people to quickly grasp and explore the field.

Designing visualizations of large volumes of textual data presents the same challenge faced by researchers in understanding and communicating any research results. This challenge can be viewed as managing tradeoffs among six design objectives (Gelman & Unwin, 2013):

- Provide an overview
- Convey scale and complexity
- Facilitate exploration
- Communicate a message
- Tell a story
- Attract attention and engage

As Gelman and Unwin (2013) report, Wordle and the word clouds it produces received an honorable mention as one of the best data visualization projects of 2008, despite its substantial limitations. Word clouds provide a simple overview of complex data in an engaging manner; however, they simply highlight frequent terms. This article compares word clouds to word networks, which can convey a coherent story and facilitate more productive exploration of textual data.

METHOD

Eleven prominent HFE journals indexed in the ISI subject category of ergonomics were included in the analysis. The analysis included abstracts from 1991 through 2010. Table 1 summarizes the sample of abstracts and journals considered in this analysis.

Table 1. Journals considered in the analysis.

Journal	Abstracts
Accident Analysis and Prevention	2213
Applied Ergonomics	1313
Ergonomics	2243
Human Factors and Ergonomics in Manufacturing	324
Human Factors	1133
International Journal of Human Computer Interaction	442
International Journal of Human Computer Studies	1116
Industrial Ergonomics	1412
Interacting with Computers	638
Safety Research	743
Le Travail Humain	334
Total	**11,911**

Data Reduction

The title and abstract of each paper were combined and the resulting text was prepared in a manner consistent with standard text analysis procedures using the tm package in the statistical program R (Feinerer, 2013; R Core Development Team, 2013). All words were converted to lower case; all numbers and punctuation were removed, as were words one or two letters long. Stop words—extremely common words, such as "of" and "the"—were also removed. Headings in structured abstracts were also removed, as were place names.

Misspellings and British English spellings were converted to standard American English spellings. After these transformations, the text was stemmed and then the stems were completed to produce more easily understood terms. For example, "driving" stems to "driv" and "driv" completes to "driver." Stem and stem completion eliminate word endings so that words with a common stem are treated as a single term (Meyer, Hornik, & Feinerer, 2008).

The final vocabulary of the corpus was defined by generating the trigrams of the terms, which consisted of all the one, two, and three-word combination of terms across the 11,991 abstracts, resulting in 1,775,613 unique terms, where a term could be a three-word combination of words, such as "subjective mental workload". These terms were filtered using the term frequency inverse document frequency (TFID) metric, which removed terms that occurred very infrequently or those that occured so frequently that they did not discriminate between documents. Terms with a TFID score in the bottom 30th percentile, as well as those that occurred in less than 10 abstracts were removed. Removing these terms left some abstracts with very few terms and so those with less than 15 terms were removed, leaving 11,108 abstracts described by 10,036 terms.

The statistical technique of topic modeling identified topics defined by groups of terms that tend to co-occur across documents (Blei & Lafferty, 2009). The joint probabilities of terms across these topics define the links in the word network. The network analysis and associated graphs were produced by igraph (Csardi & Nepusz, 2006). The word clouds were produced by wordcloud (Fellows, 2013).

RESULTS

Figure 1 shows a word cloud of the most prominent terms and Figure 2 shows the same terms in the form of a word network.

Figure 1. A word cloud of HFE terms arranged and oriented in a random fashion, with their size proportional to frequency.

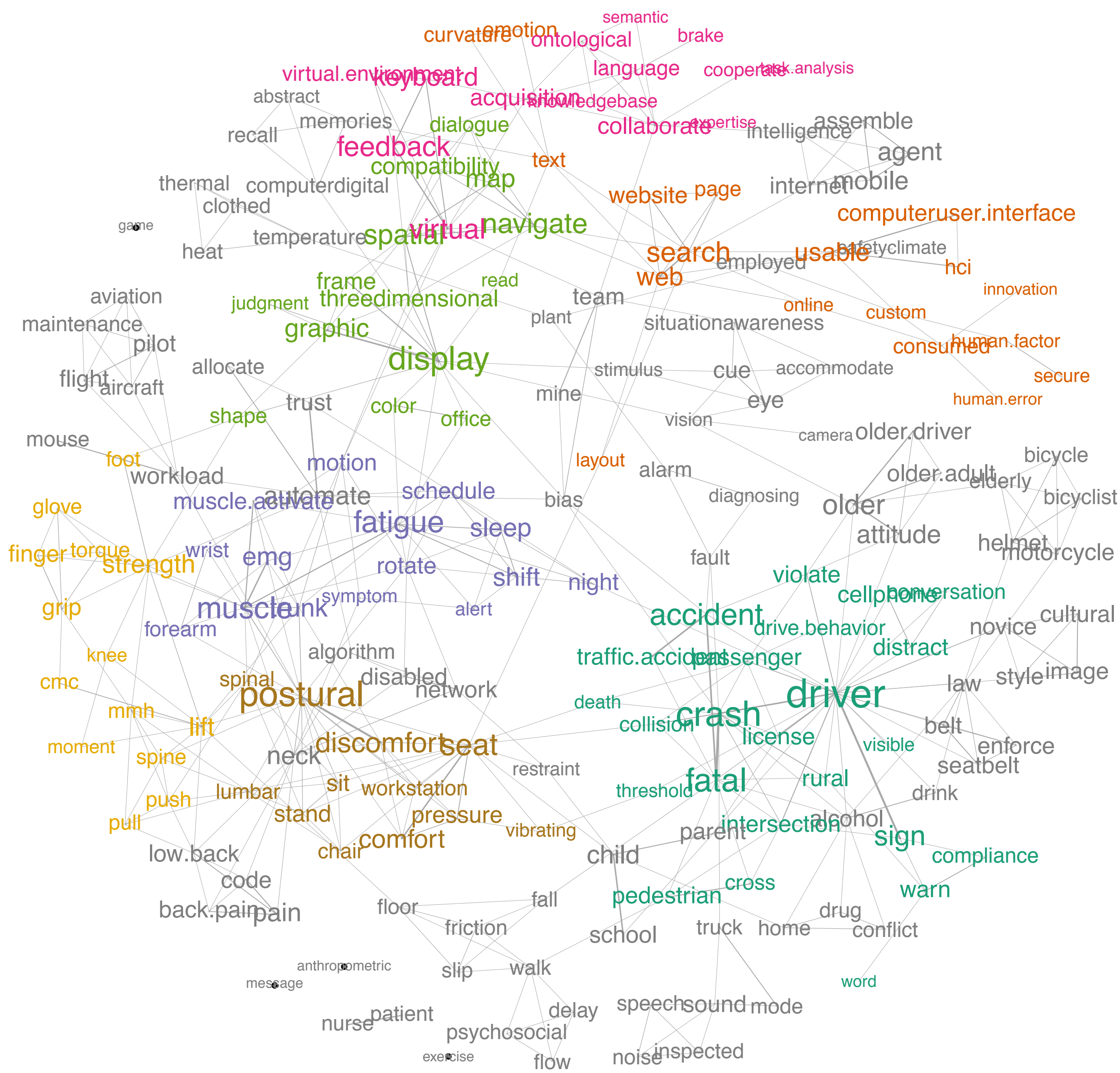

Figure 2. A word network of HFE terms arranged in a manner that approximates a multidimensional scaling solution, with their size proportional to frequency and links connecting terms that have a high joint probability.

Figure 1 provides a limited view of the central terms of HFE: it is simply a random array of terms scaled by their frequency. It highlights the most common terms by size and provides an engaging visual puzzle, but gives no other insight into the underlying papers.

Figure 2 shows a word network that also highlights the most common terms by size. Unlike a word cloud, the terms are linked according to the joint probability of the terms. The links also constrain the layout, so that the terms' position approximates a multidimensional scaling solution (Fruchterman & Reingold, 1991).

Eight colors highlight groups of terms identified by the walktrap community identification algorithm (Pons & Latapy, 2005). The lower right shows terms associated with crashes and driving safety, the lower left shows terms associated with physical ergonomics, whereas the upper right shows terms associated with displays and computer-mediated work. Condensing over 11,000 publications into 200 terms provides a very limited, low-dimensional representation of the HFE field. However, by selecting topically important terms and showing the interconnections between these terms provides a surprisingly useful overview of the field.

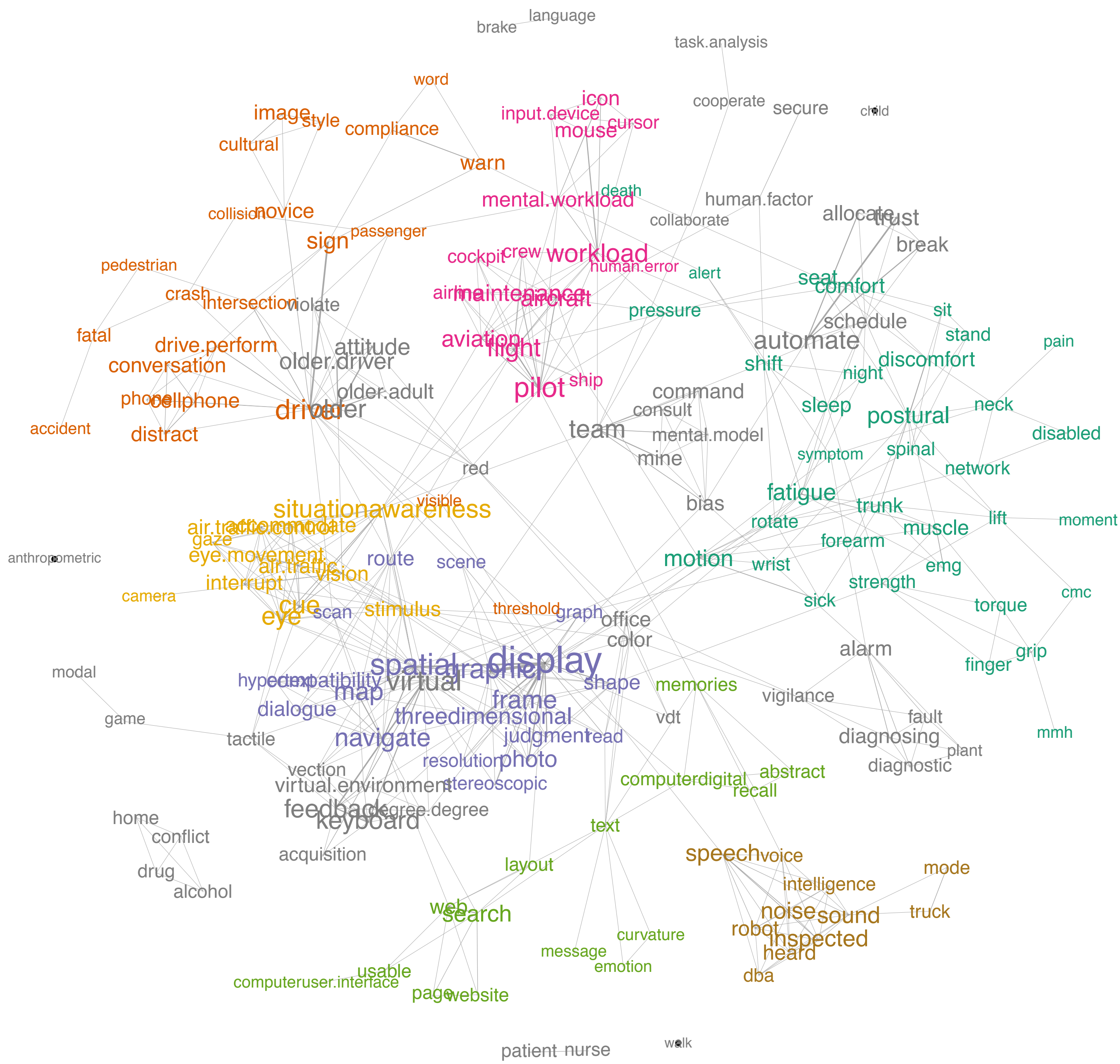

Figure 3. A word network based on the probability of topics in the journal *Human Factors*.

Figure 3 shows a word network that focuses on those terms occurring in abstracts from the journal *Human Factors*. The difference between this and the previous word network highlights the difference in focus of the journal relative to the other ten journals. Most notably, *Human Factors* places a greater emphasis on display design, situation awareness, workload, and aviation-related research.

Unlike a word cloud, the word network makes these differences more apparent because it clusters related terms, shows links, and highlights prominent communities of terms. In a word cloud, the random orientation and placement makes it difficult to identify which of the 200 words have stayed the same and which have changed.

DISCUSSION

Following Gelman and Unwin's critique of word clouds (Gelman & Unwin, 2013), the comparison of word clouds and word networks should not be viewed as one being superior to the other, but instead they represent different design tradeoffs. The simplicity of the word cloud provides an overview of a set of documents quickly. The random orientation and unusual, puzzle-like appearance is eye-catching and engaging. Also the layout algorithm prevents overlap of words, and creates an aesthetically pleasing visual puzzle.

The word networks offer a much richer view of the text. The highlighted communities and links between terms of a word network lend themselves to exploration and storytelling, two important design objectives (Gelman & Unwin, 2013). The approximation to multidimensional scaling of the Fruchterman-Reingold network layout also provides broad overview of the dimensions underlying the papers. These benefits come at a cost. The links between topics contribute to visual noise that is absent in word clouds. The layout algorithm does not avoid overlap of the text and so the spacing might not be as aesthetically pleasing as a word cloud, but overlaps do highlight closely related terms.

The word networks described in this paper are also more complex to construct than word clouds. A web-based application (wordle.net) makes word clouds very easy to construct. Word networks require a multistage analysis process that involves filtering and transforming the text, building statistical models of the latent topics underling the text, and then calculating the joint probabilities between terms (Blei & Lafferty, 2009). Plotting the word network requires layout (Fruchterman & Reingold, 1991), and community detection algorithms, such as the walktrap algorithm (Pons & Latapy, 2005). Fortunately, these tools are easily accessible through the R statistical analysis platform (R Core Development Team, 2013).

The design tradeoffs with word clouds and word networks parallel those associated with graphical representations that confront HFE researchers and practitioners more frequently. Choosing between a simple bar chart that shows only the mean, rather than a more complicated representation that conveys the mean, confidence intervals, and other elements of the distribution is a common dilemma. Unfortunately the simpler, less informative, option is the frequent choice. Previous analysis of HFE publications has focused on citation patterns or relatively simple content analysis. One such analysis considered publications in *Human Factors* between 1970 and 2000 and found that the human factors community is highly stratified with few authors publishing many papers and few papers receiving many citations (Lee, Cassano-Pinché, & Vicente, 2005). A content analysis of *Human Factors and Ergonomics Society* (HFES) meetings from 1959-1972 considered publication year, author affiliation, authors' country of residence, and research topic (Meister, 1995). More recent content analyses considered 3324 HFES conference proceedings articles (Hitt, 1998), followed by 511 *Human Factors* journal articles from 1988-1997 (Zavod & Hitt, 2000). None of these analyses considered the detailed content of the abstracts. Neglect of textual data might be masking important information (Ghazizadeh, McDonald, & Lee, 2014).

Text analysis enables analysis of large volumes of qualitative data that might otherwise go unanalyzed. A recent analysis considered 509 million messages from 2.4 million individuals from across the globe to show circadian and seasonal variation in mood (Golder & Macy, 2011). In HFE, text analysis identified lags in task-relevant topic transitions underlying team coordination (Gorman & Foltz, 2003; Gorman, Martin, & Dunbar, 2013). In sum, text analysis is a promising analysis approach that complements qualitative analysis techniques. Such text analysis could benefit from visualizations that provide deeper insight than word clouds.

ACKNOWLEDGMENTS

Thanks to Madeline Gibson for compiling paper database and to the members of CSL for comments on this paper.

REFERENCES

Blei, D. M., & Lafferty, J. D. (2009). Topic models. In A. N. Srivastava & M. Sahami (Eds.), *Text Mining: Classification, clustering, and applications*. CRC Press.

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems*, (CX.18), 1–9.

Feinerer, I. (2013). Introduction to the tm Package: Text Mining in R. *//cran. R-Project. Org/web/packages/tm/vignettes/tm.*, 1–8.

Fellows, I. (2013). wordcloud. http://cran.r-project.org/package=wordcloud.

Fruchterman, T., & Reingold, E. (1991). Graph drawing by force-directed placement. *Software: Practice and Experience*, *21*(11), 1129–1164.

Gelman, A., & Unwin, A. (2013). Infovis and Statistical Graphics: Different Goals, Different Looks. *Journal of Computational and Graphical Statistics*, *22*(1), 2–28. doi:10.1080/10618600.2012.761137

Ghazizadeh, M., McDonald, A. D., & Lee, J. D. (2014). Text mining to decipher free-response consumer complaints: Insights from the NHTSA vehicle owner's complaint database. *Human Factors*.

Golder, S. A., & Macy, M. W. (2011). Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science*, *333*(6051), 1878–1881. doi:10.1126/science.1202775

Gorman, J., & Foltz, P. (2003). Evaluation of Latent Semantic Analysis-based measures of team communications content. In *Proceedings of the Human Factors and Ergonomics Society annual meeting* (Vol. 47, pp. 424–428). SAGE Publications.

Gorman, J., Martin, M., & Dunbar, T. (2013). Analysis of Semantic Content and Its Relation to Team Neurophysiology during Submarine Crew Training. In *Foundations of Augmented Cognition* (pp. 143–152). Springer.

Hitt, J. M. (1998). 10-year (1988-1997) review of the publishing trends in the proceedings of the Human Factors and Ergonomics Society. In *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting* (pp. 679–682). Santa Monica, CA: HFES.

Lee, J. D., Cassano-Pinché, A., & Vicente, K. J. (2005). Bibliometric analysis of Human Factors (1970-2000): A quantitative description of scientific impact. *Human Factors*, *47*(4), 753–766.

Meister, D. (1995). Human factors—the early years. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 39, pp. 478–480). SAGE Publications.

Meyer, D., Hornik, K., & Feinerer, I. (2008). Text mining infrastructure in R. *Journal of Statistical Software*, *25*(5), 2–55.

Pons, P., & Latapy, M. (2005). Computing communities in large networks using random walks. *Journal of Graph Algorithms and Applications*, *10*, 191–218. doi:10.1007/11569596

R Core Development Team. (2013). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.

Zavod, M., & Hitt, J. M. (2000). Summary of the Publishing Trends of the Journal of Human Factors from 1988-1997. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *44*(33), 6–108–6–111.