

**What's on Wikipedia, and What's Not...?
Completeness of Information on the Online Collaborative Encyclopedia**

Cindy Royal, Ph.D.

Assistant Professor
Texas State University
School of Journalism and Mass Communication
102 Old Main
601 University Avenue
San Marcos, TX 78666
Ph. 512-245-3481
croyal@txstate.edu

Deepina Kapila

Graduate Student
Texas State University
School of Journalism and Mass Communication
102 Old Main
601 University Avenue
San Marcos, TX 78666
deepina@txstate.edu

January 24, 2007

What's on Wikipedia and What's Not...? Completeness of Information on the Online Collaborative Encyclopedia

Abstract

The World Wide Web continues to grow closer to achieving the vision of becoming the repository of all human knowledge. While improved search engines such as Google facilitate access of knowledge across the Web, some sites have increased in popularity and have attracted the attention of more Web users than others. Wikipedia is one such site that is becoming an important resource for news and information. It is an online information source that is increasingly used as the first, and sometimes only, stop for online encyclopedic information.

Much discussion has dealt with the accuracy of information on Wikipedia. While accuracy is important, that is not what this project is measuring. Using a method employed by Tankard and Royal (2005) to judge completeness of Web content, completeness of information on Wikipedia is assessed. What we found was that some topics were covered more comprehensively than others and that predictors of these biases included recency, importance, population, and financial wealth.

Introduction

The World Wide Web continues to grow closer to achieving the vision of becoming the repository of all human knowledge (Heylighten, 1995). While improved search engines such as Google facilitate access of knowledge across the Web, some sites have increased in popularity and have attracted the attention of more Web users than others. Wikipedia is one such site. It is an online information source that is increasingly used as the first, and sometimes only, stop for online encyclopedic information.

Wikipedia (www.wikipedia.org), deemed “the free encyclopedia,” was launched on the Web in 2001. (Wikipedia:About, 2007)) It was started by Jimmy Wales, formerly a futures trader in Chicago, as an open information source, allowing anyone with access to the Internet to post or edit content on the site. Wikipedia uses the wiki software format, which is a collaborative development environment. Established as a non-profit organization, Wikipedia currently receives over 38 million unique visitors per month and is ranked #13 on ComScore Media Metrix Top 50 Web Properties (Holiday Fever..., 2007). This open source project operates under the assumption that more writers and editors are better than fewer, and that the community will develop and monitor content in a manner that is improved over that of traditional information publishing.

Wikipedia is now the Web's third most popular news and information source, with more unique visitors than Yahoo News, MSNBC, AOL News, and CNN (Half of All U.S. Internet Users..., 2006). Wikipedia's English-language version doubled in size last year and now has over 1 million articles. By this measure, it is almost 12 times larger than the print version of the Encyclopaedia Britannica. It has over 100,000 contributors writing in 200 languages (The Wiki Principle, 2006).

What's on Wikipedia, and What's Not...? Completeness of Information on the Online Collaborative Encyclopedia

Wikipedia has become a popular site frequented by students, scholars, business people, family members, and government officials for finding information on a variety of topics. But, due to the open nature of contributions, much attention has been given to the level of accuracy of information on Wikipedia. Many feel that Wikipedia's policy of letting anyone create and edit content causes the information to be inaccurate, misleading, or generally incorrect, both purposefully and accidentally. Instances have occurred in which rumors and falsities have been planted on Wikipedia articles. For example, a Wikipedia entry was created that falsely implicated John Siegenthaler, Sr. in the Robert Kennedy assassination (Giles, 2005; Udell, 2004; Johnson, 2006). While the error was eventually corrected, it was not done so before being picked up by other information resources and seen by untold numbers of users. Still, the philosophy of the site is that with so many people looking at the content, in the long run, accuracy will prevail.

Wikipedia has sought to counter some of the criticisms by instituting measures designed to reduce the number of attacks on the credibility of information on the site. Volunteer administrators monitor content on the site, and can now block users from editing content on specific articles. Some articles are temporarily protected from editing, until the climate for the attack has died down. Others, like the article on George W. Bush, are semi-protected and open to editing only by people who had been registered on the site for at least four days. (Hafner, 2006). But according to Wales, Wikipedia's founder, this type of protection affects a tiny fraction of the 1.2 million entries on the English-language site. "Protection is a tool for quality control, but it hardly defines Wikipedia," Mr. Wales said. "What does define Wikipedia is the volunteer community and the open participation." (Hafner, 2006)

Some studies have actually refuted Wikipedia's position as a reliable information source. In a recent study comparing the accuracy of science entries, *Nature* reported that Wikipedia's

What's on Wikipedia, and What's Not...? Completeness of Information on the Online Collaborative Encyclopedia

level of accuracy is close to that of *Encyclopedia Britannica* (Giles, 2005). The scientific journal reported that, within 42 randomly selected general science articles, there were 162 mistakes in Wikipedia versus 123 for *Britannica*, with the errors in *Britannica* being oriented towards omissions rather than factual errors.

There is an indication that even librarians are finding value in the usage of Wikipedia (Miller, et al., 2006). Attention to popular culture items and usage of links, objective presentation of controversy, and up-to-date nature of information are all improvements over traditional encyclopedias. Others characteristics of Wikipedia are considered strengths. Consistent presentation of information and format of pages, organization of articles, and links to outside sources provide users with a site that is robust and efficient. The basic search field allows users to quickly find information on their desired topic or related subjects.

While accuracy of information is important, that is not what this project is measuring. There are other criteria in which an information source can and should be judged. Using a method employed by Tankard and Royal (What's on the Web..., 2005) to judge completeness of Web content, completeness of information on Wikipedia is assessed. With Wikipedia becoming a popular online information destination, it might possess some of the biases inherent in the Web at large. Are some topics covered more comprehensively than others and, if so, are there systematic predictors, like those found on the Web in general, that determine the amount of coverage?

Tankard and Royal (What's on the Web..., 2005) studied completeness of information on the Web by creating systematic lists of topics to search in two popular search engines. Searching on several dimensions including recency, importance of information, country population, and company revenue, they also investigated these dimensions as predictors of completeness of

What's on Wikipedia, and What's Not...?

Completeness of Information on the Online Collaborative Encyclopedia

information on the Web. In each dimension, Tankard and Royal found that there were systematic biases. Coverage for items that were current, deemed important, or held power measured by country size or company wealth produced more hits, and thus were covered more frequently on the Web. This study was unique in that it was the first to propose a method for assessing the completeness of information on the Web (Finding Out What's on the World Wide Web, 2005). It seems appropriate to modify these methods and to apply them to specific Web resources, particularly those that have become population information destinations. Wikipedia is a likely candidate for analysis in that its goal is to provide information created and accessible by all with an Internet connection, much like the Web itself.

Review of Literature

While there have been many articles questioning Wikipedia's accuracy, few communication studies have focused on Wikipedia. Lih (2004) studied news articles citing Wikipedia and analyzed the trends in using Wikipedia as a source.

Denning, et al. (2005) listed several risks inherent in the Wikipedia model: accuracy, motives, uncertain expertise, volatility, coverage, and sources. Of coverage, the authors said,

Voluntary contributions largely represent the interests and knowledge of a self-selected set of contributors. They are not part of a careful plan to organize human knowledge. Topics that interest the young and Internet-savvy are well covered, while events that happened "before the Web" may be covered inadequately or inaccurately, if at all. More is written about current news than about historical knowledge.

Other studies have looked at Wikipedia's strength as a reference source. Bill Katz developed six fundamental evaluation criteria for reference work: purpose, authority, scope, audience, cost, and format (Wallace and Van Fleet, 2005). Wikipedia did not perform well on the brief analysis performed by Wallace and Van Fleet on these criteria. Value, however, was

What's on Wikipedia, and What's Not...? Completeness of Information on the Online Collaborative Encyclopedia

identified in the democratic and timely circumstances under which articles are created and revised. According to Bopp and Smith (2001), coverage in an encyclopedia reference source “should be even across all subjects,” although “it is important to note that some subjects, by their very nature, demand greater emphasis.” While Wikipedia boasts over 1 million articles, Wallace and Van Fleet expressed that volume of articles alone is not a useful indicator of scope.

Like the Tankard and Royal study, this project challenges the notion that the Web may be the repository of all human knowledge by assessing the coverage on one of its most popular information destinations, Wikipedia. By making systematic measurements of the amount of information on Wikipedia using the same dimensions, we attempt to identify factors that predict Wikipedia's completeness.

Borrowing methods from the Tankard and Royal study, this project measures the content of Wikipedia against various indexes or standards of completeness to identify and uncover potential inherent biases. Communication research provided direction in identifying predictor variables. Journalism scholars have often included completeness as one of the basic concepts of journalism. McQuail stated that completeness “is usually thought to be a precondition of proper understanding of news, and the media generally promise completeness in the sense of a full range of information about significant events of the day” (McQuail, 1992, p. 211).

In an early study of the completeness of newspaper coverage, Danielson and Adams (1961) examined coverage of the 1960 presidential election campaign. They developed a list of 1,033 campaign events and then drew a random sample of 42 events to be used as a checklist against which articles were judged.

Tankard and Showalter (1977), in their study of coverage of the 1972 Surgeon General's report on television violence, constructed an index of completeness by checking for presence or

What's on Wikipedia, and What's Not...? Completeness of Information on the Online Collaborative Encyclopedia

absence of “three elements that were judged necessary for full reader understanding.” While the present study does not focus on individual news stories, it borrows the technique of using a list of facts or concepts as an effective means of measuring completeness.

Research on news flow has identified a number of factors that influence the presence or absence of information. A related research approach—theoretical influences on mass media content—has identified five major categories of influence on news content: the individual journalist, media routines, the journalistic organization, extramedia sources, and ideology (Shoemaker & Reese, 1995).

Current information is the bedrock of journalistic reporting (Berkowitz, 1990; McMillin, 1996; Curtin & Rhodenbaugh, 1999). With regard to the Web, currency comes into play in another sense. Shoemaker & Reese (1995) identified the individual as a news influencer. Web users and content creators tend to be young, with strong ties to current popular culture. The contributors to Wikipedia are likely to mirror the demographics of the Web at large. This factor would tend to weight the content of the Web, and ostensibly Wikipedia, toward material that these individuals would be interested in—material of greater currency or recency.

Galtung & Ruge (1965), identified signal strength, or amplitude as another significant factor influencing the flow of news. This factor might also be thought of as the importance of information. When considering the probability of information being on Wikipedia, importance of the information is likely to be a useful predictor, with the more important items having the most attention paid to them.

Kariel and Rosenvall (1984) identified country population as an important predictor of international news flow. Countries with larger populations have more individuals to become the

What's on Wikipedia, and What's Not...?

Completeness of Information on the Online Collaborative Encyclopedia

focus of news coverage, hold greater political influence, and have more people who could potentially create and contribute to online content.

Shoemaker and Reese (1995, p. 190) suggested that capitalist-owned media content tends to favor those with economic power. In addition, corporations that are larger have more market impact, have larger budgets for advertising and public relations, and have influence on more people.

Research Questions

Borrowing from the Tankard and Royal study of completeness of information on the Web, the following research questions were developed as they related to Wikipedia:

1. Are there some systematic gaps or biases in the overall presentation of information made available on Wikipedia?
2. Is recency (or currency) a predictor of amount of information on Wikipedia?
3. Is importance of information a predictor of amount of information on Wikipedia?
4. Is population a predictor of amount of information about particular countries on Wikipedia?
5. Is economic power a predictor of amount of information about individual corporations on Wikipedia?

Method

Using the same predictors as Tankard and Royal, recency, importance, country population, and economic power, several systematic searches on Wikipedia were conducted. Lists were developed within each of the dimensions, the contents of which are described in the results section. Each term on the lists was searched using the Wikipedia search feature. A determination was made as to the main page of content for that term. In some cases, such as the

What's on Wikipedia, and What's Not...?

Completeness of Information on the Online Collaborative Encyclopedia

countries of the United Nations, the list of countries on the United Nations page was used to find the main article on a particular country. Each page was visited and the relevant content was highlighted. Wikipedia navigation and other superfluous links that were not related to the actual term being searched were not included in the selection. To capture the word count of items selected on a page, an extension of the Firefox Web browser, Word Count, was downloaded (<http://roachfiend.com/archives/2005/03/03/word-count/>). This extension counted the number of words in the selection by simply using the Ctrl key on the computer's keyboard. Word counts were captured in a spreadsheet for each dimension. Items were plotted on charts, first in ascending order, then by predictor variable. Items within dimensions were then compared and correlated with predictor variables. When possible, the same search terms that were used in the Tankard & Royal study were employed here.

All statistical analyses were conducted with Spearman (rank order) correlation coefficients because parametric statistics (such as the Pearson correlation coefficient) are inappropriate for L-shaped distributions (Bradley, 1982), which occurred with most of our data. The correlations represent relative, as opposed to absolute, relationships.

Results

Several variables were used to test the currency dimension. First, using the same method as Tankard and Royal, years were assessed. Wikipedia conveniently provided an article depicting the highlights of each year. Figure 1a depicts the word count of each article in ascending order, disregarding year. A backward L-shaped curve is evident. Figure 1b depicts the word count by year in chronological order, starting with 1900 and going through 2010. There is a clear progression of the length of each article with a dramatic increase occurring

What's on Wikipedia, and What's Not...?

Completeness of Information on the Online Collaborative Encyclopedia

starting in 2001. Years in the future, understandably, were shorter, given that there was not yet much to write about them. The average word count for the years since 2001 was 90% greater than the average for the entire preceding 100 years (4566 vs. 8692).

The chart in Figure 8 depicts correlations of dimensions variable with predictor variables. The Spearman correlation for Years was .79, indicating a very strong relationship of article word count to the recency of information.

What's on Wikipedia, and What's Not...?
Completeness of Information on the Online Collaborative Encyclopedia

Figure 1a
Years - Ascending Order

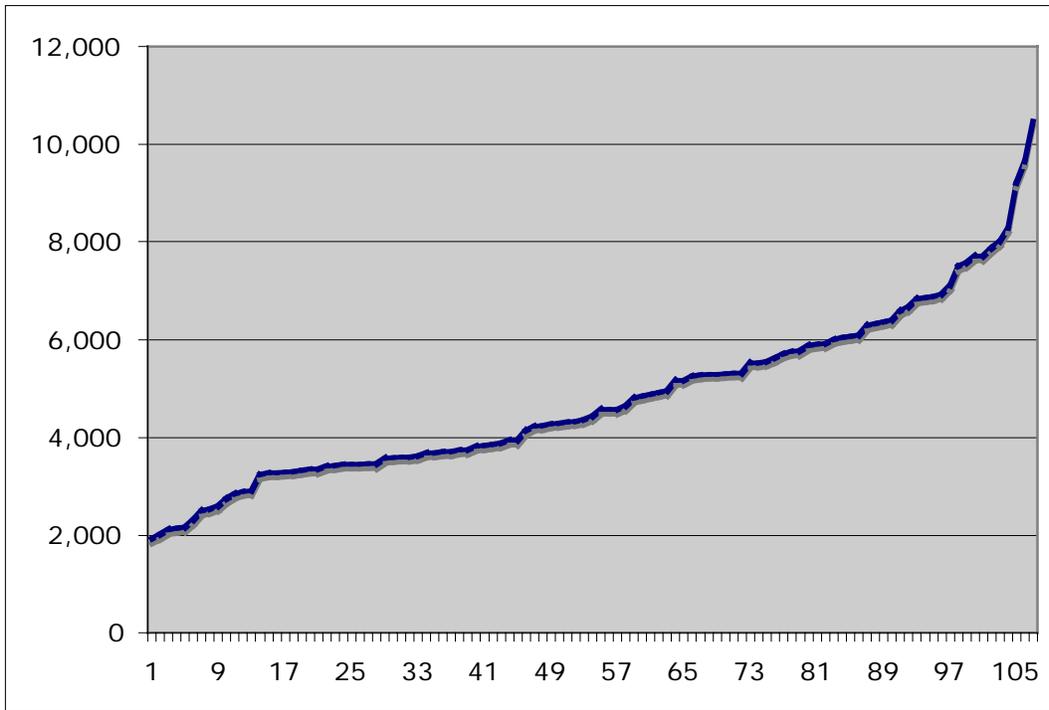
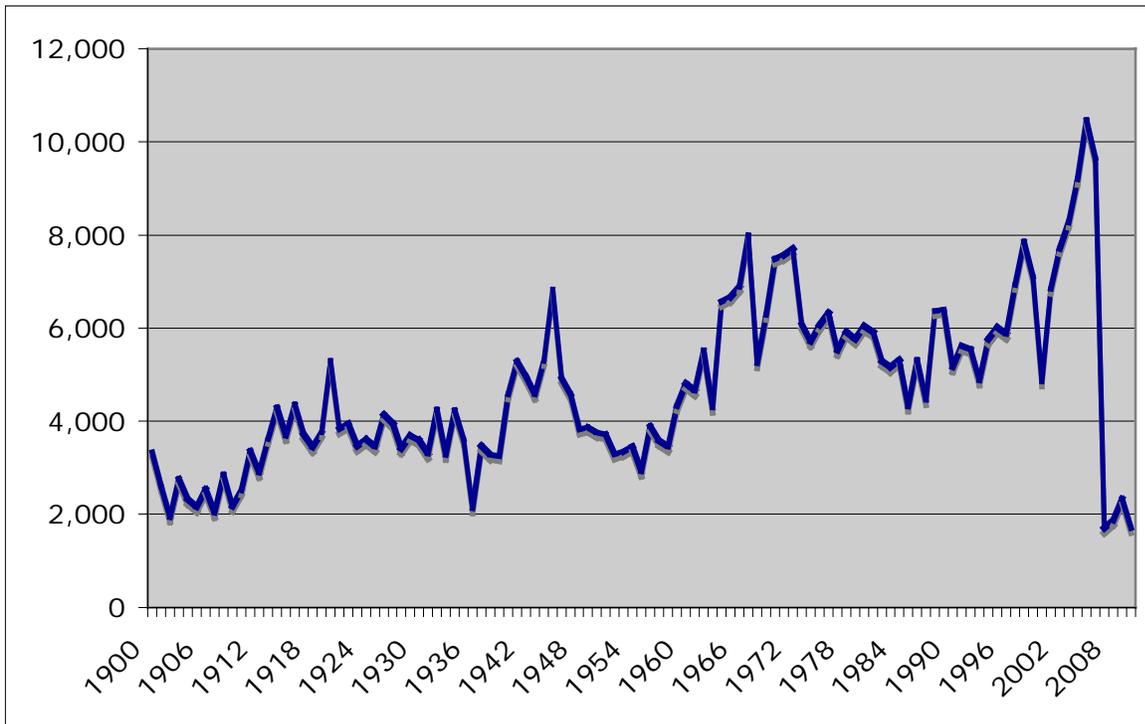


Figure 1b
Years - Chronological Order



What's on Wikipedia, and What's Not...?
Completeness of Information on the Online Collaborative Encyclopedia

Figure 2a shows the word count for articles on Wikipedia for the Academy Award winning films in ascending order. This list was not searched in the Tankard & Royal study, as it was difficult for them to identify only Web sites associated with films with common names, such as *Wings* or *Rebecca*. However, this was made easier on Wikipedia, with each film having a specific article associated with it. Another backward L-shaped distribution is displayed. With few exceptions, such as *Gone with the Wind* (1939) and *Casablanca* (1943) the analysis in Figure 2b plotted by year (1928-2005) shows a progression favoring more current films. This demonstrates that while recency is an important predictor, some films transcend time and are deemed important for other reasons, and thus have a strong share of coverage on Wikipedia. The average word count for the films since 2001 was 80% higher than the average word count for the time prior to 2000 (3190 vs. 1771). These last five years accounted for 11% of the total word count for the 78 years of the award. The Spearman correlation for films over years was .49 (see Figure 8), but that increased to .62 simply by removing the two outliers mentioned above. This indicates a strong relationship between word count and time for films.

What's on Wikipedia, and What's Not...?
Completeness of Information on the Online Collaborative Encyclopedia

Figure 2a
Films - Ascending Order

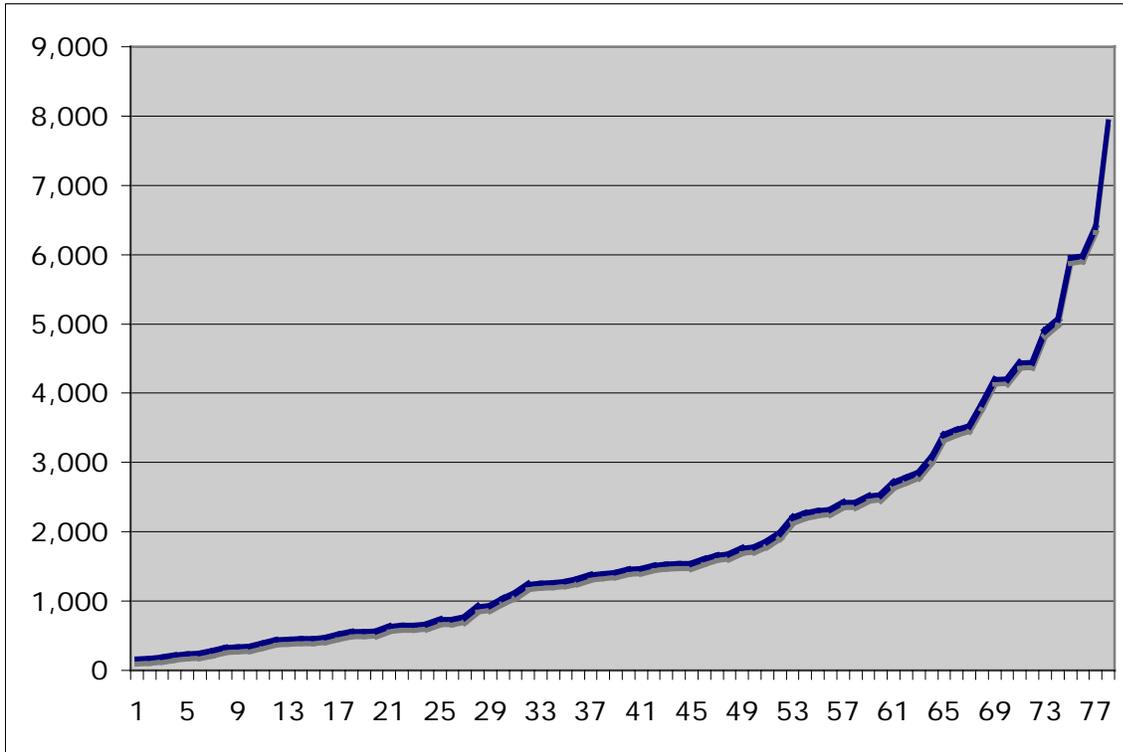
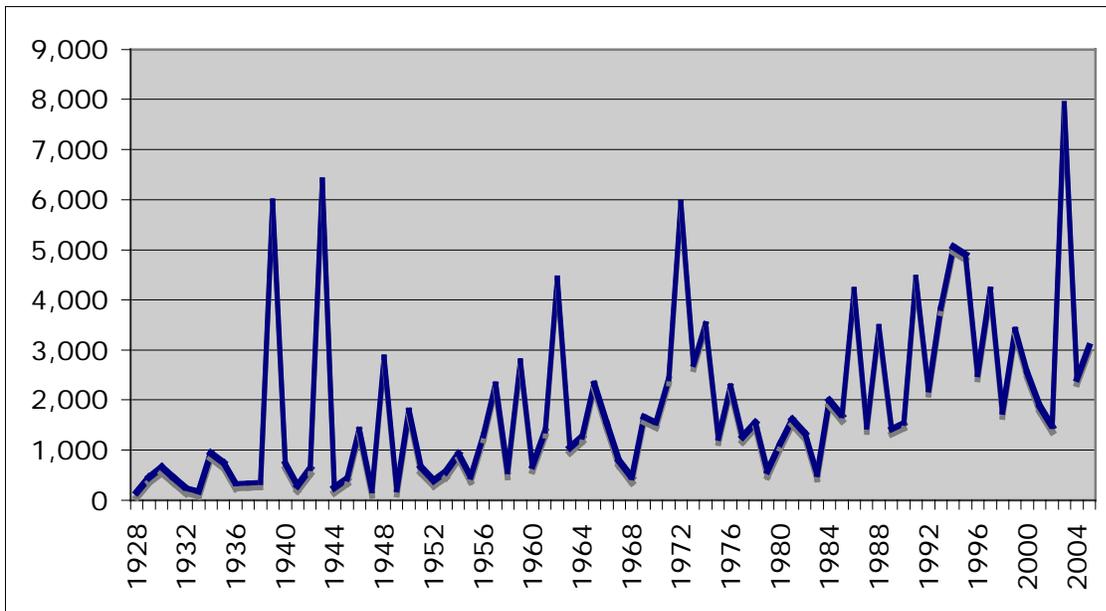


Figure 2b
Films - By Year



What's on Wikipedia, and What's Not...?
Completeness of Information on the Online Collaborative Encyclopedia

Figures 3a and 3b show another test of recency not performed in the Tankard & Royal study by looking at *Time Magazine's* Person of the Year. Some years that did not include an individual were discarded (for example, in 2002, "The Whistleblowers"). Figure 3a shows a backward L-shaped distribution when disregarding time, although not as steep as some of the others experienced in this analysis. The progression appears evenly distributed, only slightly skewed to the upper half of the distribution (the median was 93% of the average). But, Figure 3b shows a more random pattern than those experienced with Year and Film. The Spearman correlation (see Figure 8) for recency was close to 0, thus indicating no relationship with time. This indicates that while a bias is evidenced in the consistently upward progression of Figure 3a, the bias is not due to recency in regard to Person of the Year, but perhaps to some other measure of importance.

Figure 3a
Time's Person of the Year - Ascending Order

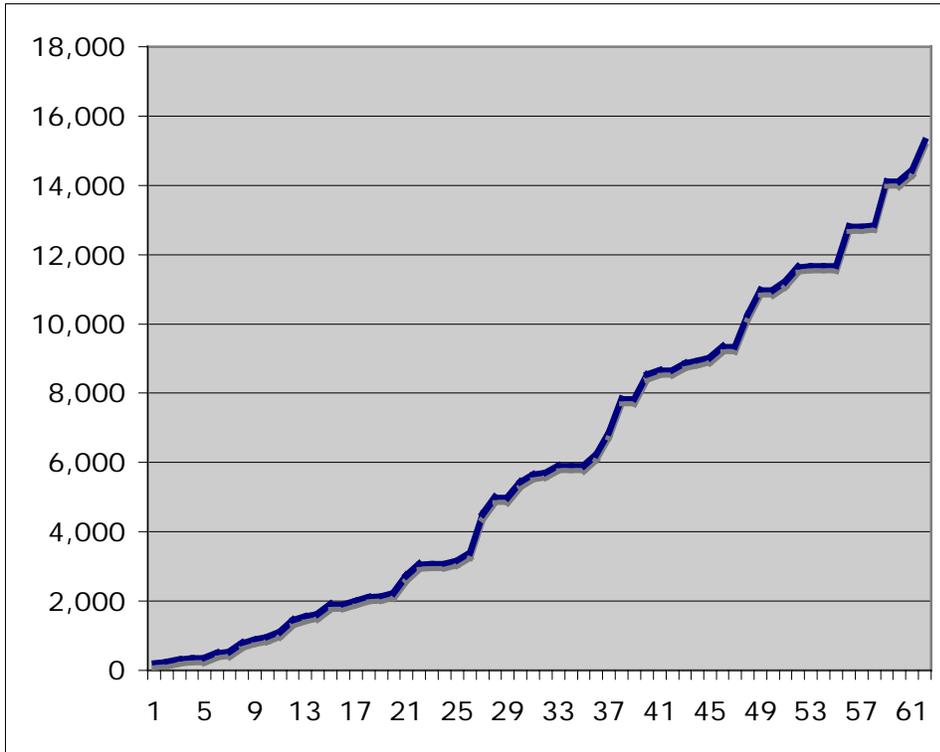
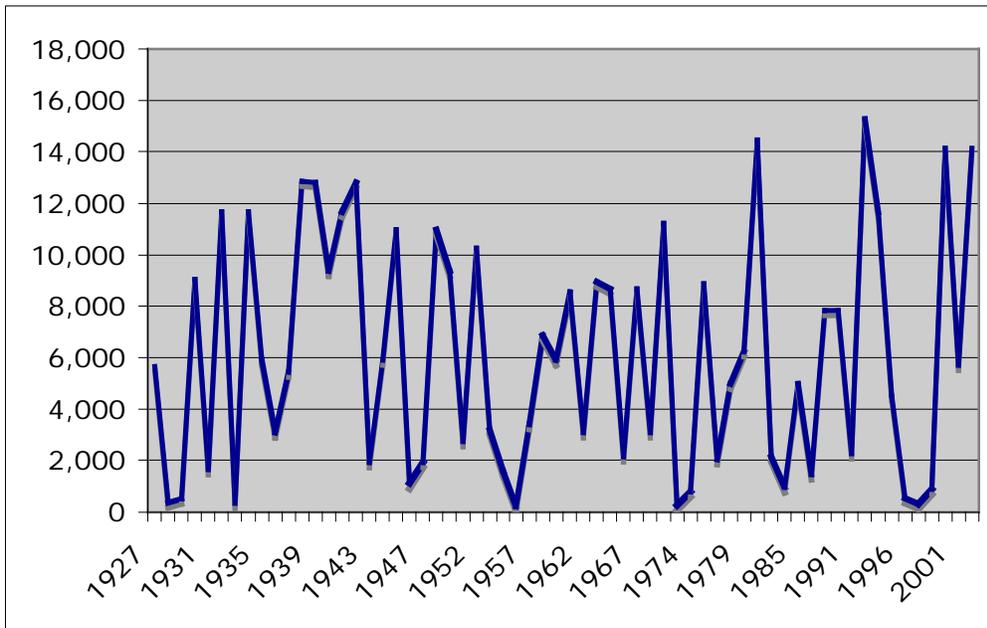


Figure 3b
Time's Person of the Year - By Year



What's on Wikipedia, and What's Not...? Completeness of Information on the Online Collaborative Encyclopedia

Another search that was added that was not performed in the Tankard & Royal study was to consider musical artists over time. An artist holding the #1 song on the *Billboard* Top 100 for the first week in February of each year since 1940 was selected. Figure 4a depicts the word count of the main Wikipedia article associated with that artist in ascending order for each of the selected artists, again depicting the backward L-shaped distribution. Figure 4b shows each artist by year. While the pattern in the graph appears to indicate a random distribution, the Spearman correlation with time was .30 (See Figure 8). By eliminating just two outliers (Bing Crosby – 1945 and the Beatles – 1964), the correlation increases to .40. The average word count for the artists since 1990 was 32% higher than for the years from 1940-1989 (3332 vs. 2511). Similar to the trends found in film, it shows that while the recency relationship is strong, some artists transcend time and receive more coverage on Wikipedia than would be indicated by their currency.

Figure 4a
Artists with #1 Songs on Billboard – Ascending Order

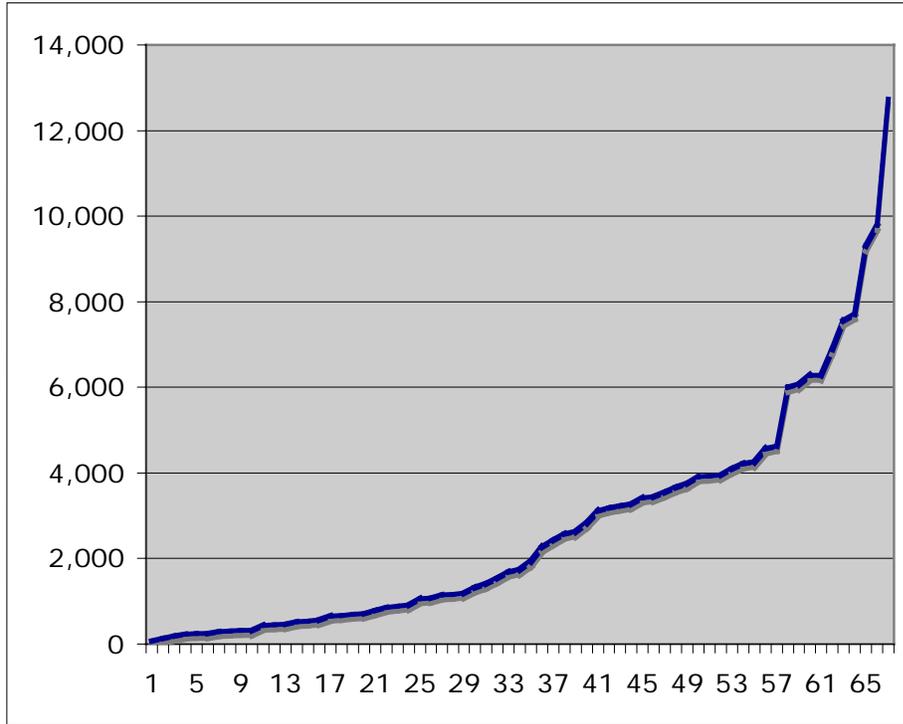
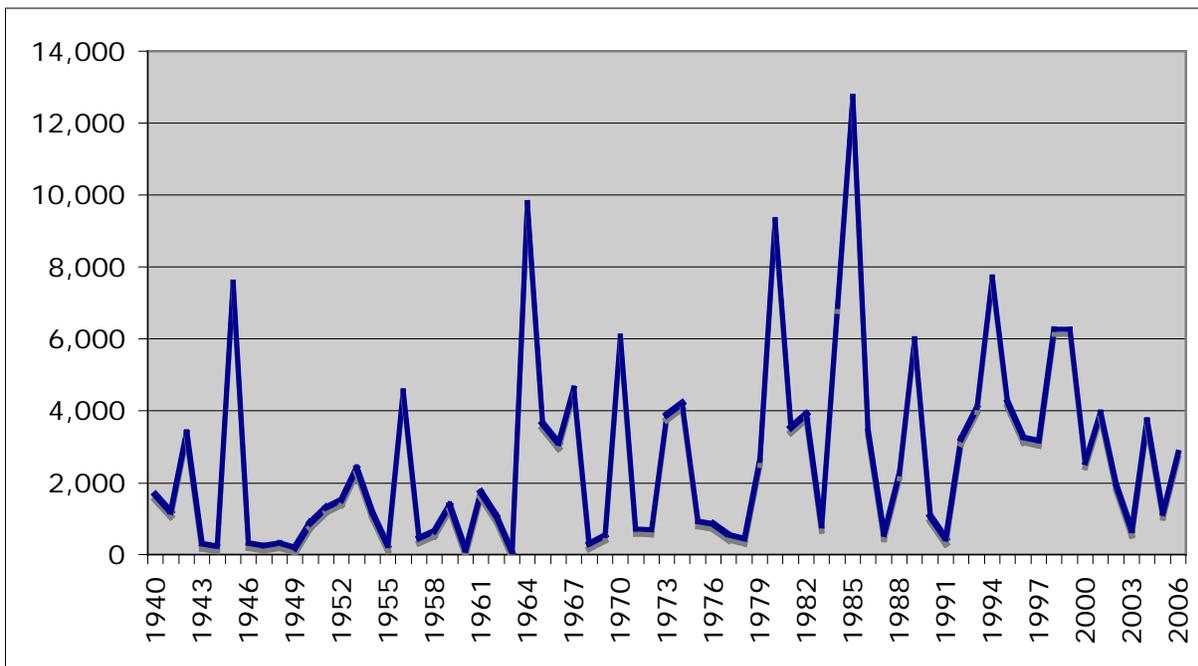


Figure 4b
Artists with #1 Songs on Billboard – By Year



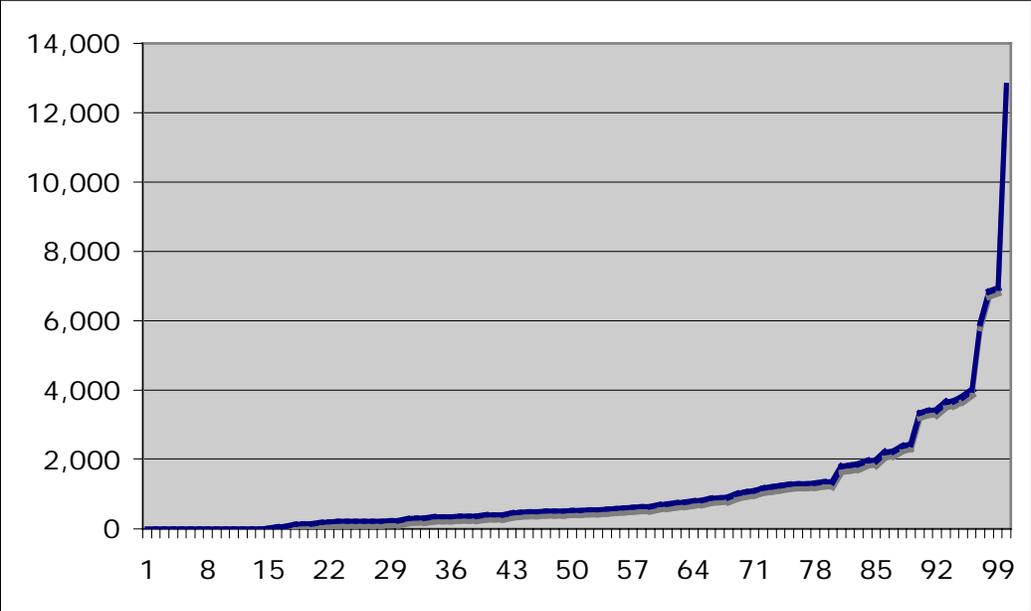
What's on Wikipedia, and What's Not...?
Completeness of Information on the Online Collaborative Encyclopedia

To measure comprehensiveness of information, we used the same random sample employed in the Tankard & Royal study of 100 topics from the *Micropaedia* of the *Encyclopaedia Britannica*. Figure 6 shows the word count of each term's main page on Wikipedia. Once again, a backward L-shaped distribution emerged. Of the 100 items, 14 were not represented at all on Wikipedia. Fifteen of the terms had articles with a word count of 2000 or more. The average word count for those 15 terms was 5 times that of the average word count for the other items on the list with Wikipedia articles.

A Spearman correlation was used to compare inches of content in the *Micropaedia* of the *Encyclopaedia Britannica* with word count on Wikipedia. This correlation was calculated at .26, indicating some relationship with the importance placed on information in the traditional encyclopedia with that in Wikipedia (See Figure 8). In some cases, the articles on Wikipedia indicated that the content had been derived from a print encyclopedia source. There was no time dimension or other predictor variable with which to compare for encyclopedia terms.

What's on Wikipedia, and What's Not...?
Completeness of Information on the Online Collaborative Encyclopedia

Figure 5
Encyclopedia Terms – Ascending Order



What's on Wikipedia, and What's Not...? Completeness of Information on the Online Collaborative Encyclopedia

Figure 6a shows the word count for the main Wikipedia article by country in the United Nations in ascending order. Articles were analyzed for the 192 countries of the United Nations. Once again, a backward L-shaped distribution emerged. The distribution is fairly even, with a sharp increase experienced for the top 22 countries. Figure 6b shows a gradual upward distribution when charted in order by population (higher number indicates higher population). Spearman correlation for countries with population was .55, indicating that the larger countries were more represented on Wikipedia in terms of word count per article (see Figure 8). The top 10% of countries by population accounted for 15% of the total word count for country articles and the average word count for the top 10% of countries was 63% higher than those on the rest of the list.

Figure 6a
Countries in UN – Ascending Order

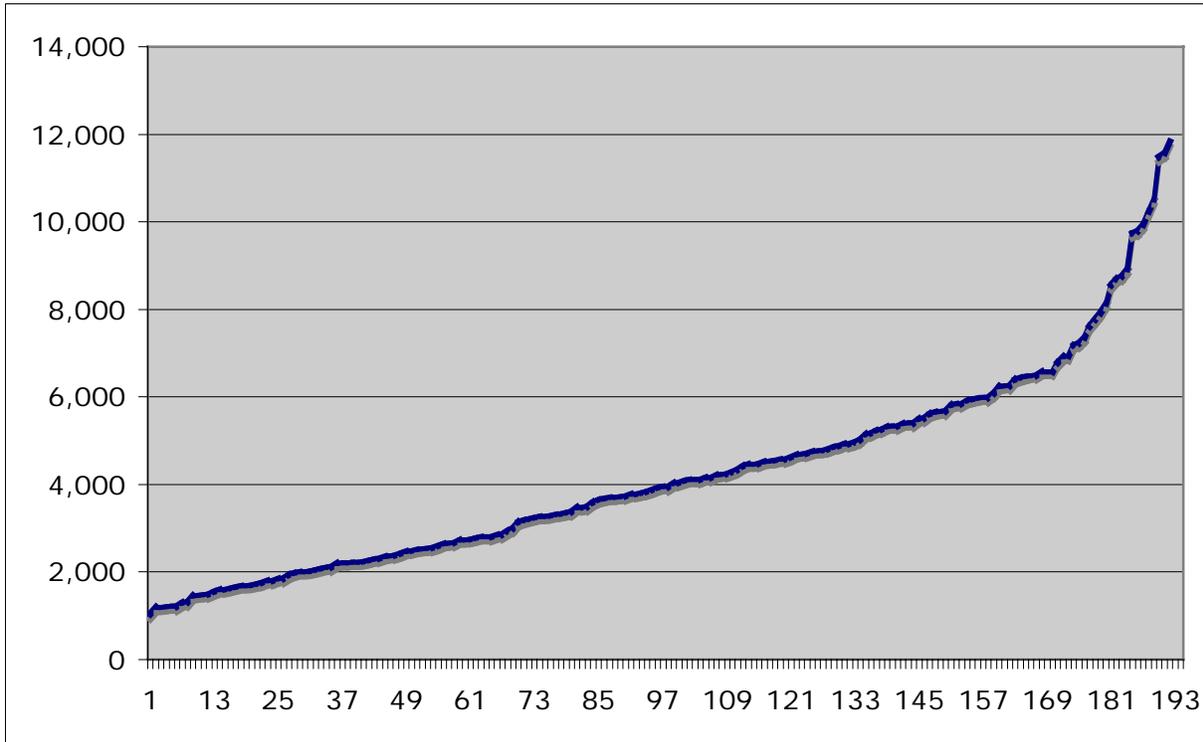
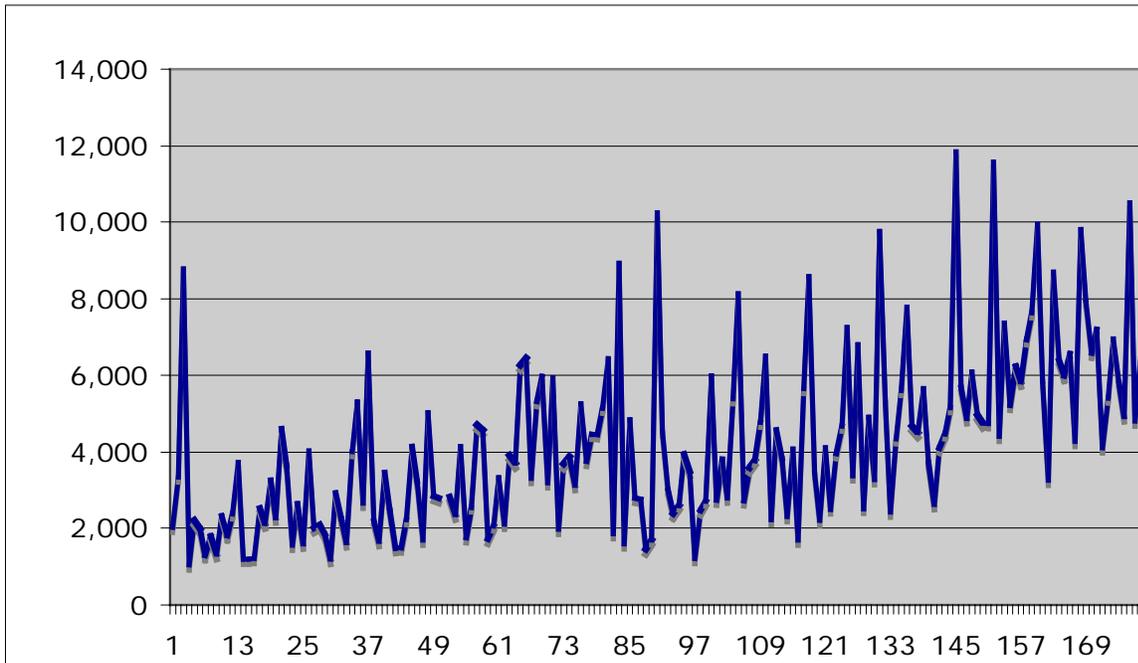


Figure 6b
Countries of the UN - Ordered By Population



What's on Wikipedia, and What's Not...? Completeness of Information on the Online Collaborative Encyclopedia

Figure 7a shows the word count for a random selection of 86 Fortune 1000 companies in ascending order. This chart shows the backward L-shaped distribution with a sharp increase for 10% of the companies. Another 10% of the companies did not have Wikipedia entries. Figure 7b shows the companies ranked by revenue (higher number indicates higher revenue). The chart shows a distribution trending toward increased word count for companies with the highest revenue. The Spearman correlation for word count of these articles with company revenue was .49. The top 10% of the companies by revenue accounted for 30% of the total word count for articles about companies.

What's on Wikipedia, and What's Not...?
Completeness of Information on the Online Collaborative Encyclopedia

Figure 7a
Fortune 1000 Companies – Ascending Order

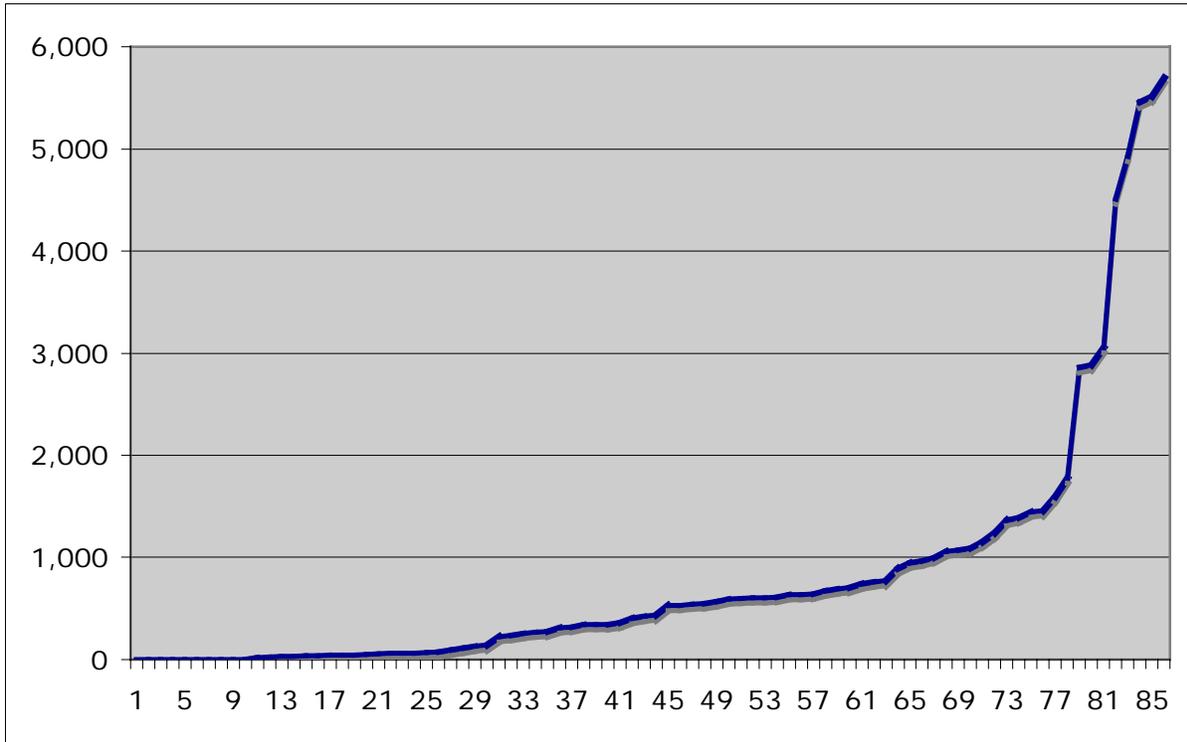


Figure 7b
Fortune 1000 Companies - By Revenue

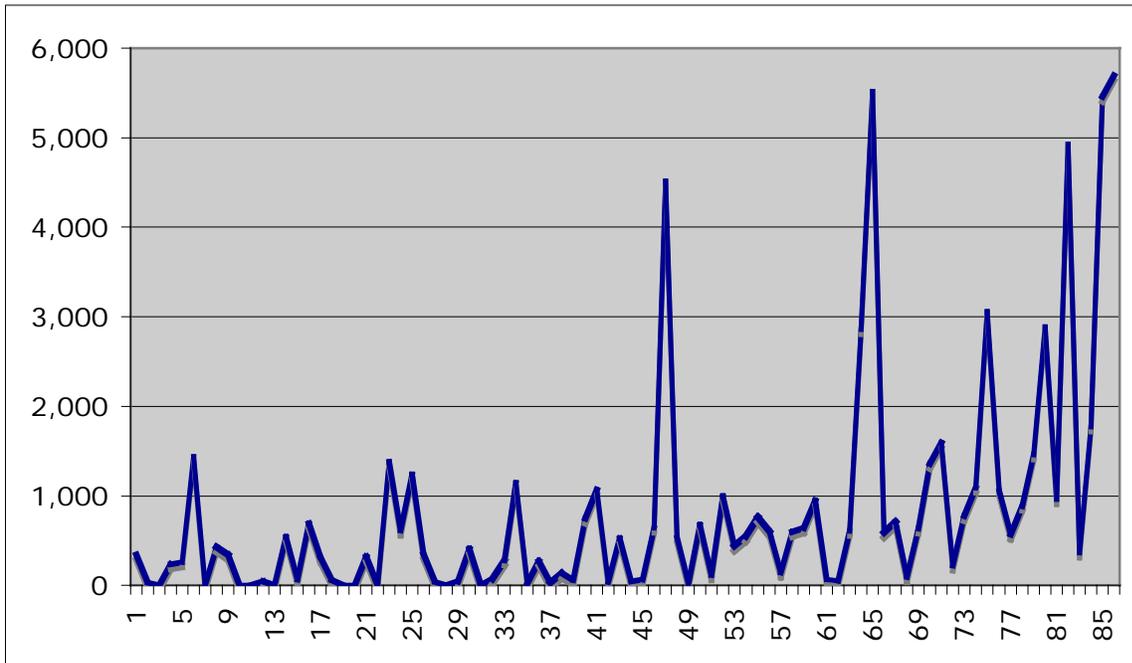


Figure 8
Spearman Correlation (Rank Order) with Predictor Variables

Dimension Variable	Predictor Variable	Correlation
Year	Time	.79
Academy Award Winning Films	Time	.47
Time's Person of the Year	Time	.00
Artist w/ #1 Song	Time	.30
Encyclopedia Term	Column Inches of Encyclopedia	.26
Country	2005 Projected Population	.55
Company	2005 Revenue	.49

Discussion

In each of the searches performed for the dimensions, a bias was evident. And, when considering the predictor variables, except when measuring *Time's Person of the Year*, a strong correlation was experienced. Within the currency or recency dimension, by looking at Year, Films, and Musical Artist, the more current topics were the most covered. While the Person of the Year category showed some bias, it was not closely correlated with time. This is perhaps due to the nature of the decision as being more editorial and reflecting opinions of the staff of Time. While they often select people who are historically important, over time, according to our study, they choose people with staying power or continued importance only about half the time.

When looking at a random selection of encyclopedia terms, bias was also inherent. Most of the items that we searched had some information about them on Wikipedia, indicating broad coverage. But, it was clear that the more common or popular terms had the most detailed coverage. Coverage in Wikipedia was loosely correlated with the inches dedicated to the topics in one traditional encyclopedia, indicating the strength of the agenda established by these publications.

What's on Wikipedia, and What's Not...? Completeness of Information on the Online Collaborative Encyclopedia

In terms of country population, biases toward larger countries were found and were positively correlated with country size. This indicates that the democratic nature of Wikipedia on its own cannot counteract the effects of the magnitude of people that are available to participate.

And, in regard to Fortune 1000 companies, those with larger revenue streams and resources are more likely to have greater coverage on Wikipedia. This points to the strength of financial power in circumventing any type of democratizing feature of an online space.

Conclusion

In some ways, this was a more straightforward study than the one performed by Tankard & Royal. In their study, they had difficulty in determining whether certain searches were capturing all the information on a topic while not including irrelevant information. For example, search for years in a search engine can provide references to the numbers rather than the years. They attempted to alleviate this problem by searching for the word “year” before the numerical year and putting quotation marks around that text string. This did not capture hits regarding years that were not preceded by the word “year”. Some searches were difficult to perform if the topics were not presented consistently, as in the encyclopedia terms.

Shariatmadari (2006) identified characteristics of Wikipedia that make this case as well. Wikipedia is specifically intended as a work of reference while using a search engine is not. A search engine's purpose is to identify various sites as opposed to finding immediate context. Shariatmadari also indicated some coverage issues with Wikipedia, finding content more on popular culture and science fiction than history.

What's on Wikipedia, and What's Not...? Completeness of Information on the Online Collaborative Encyclopedia

In general, the searches on Wikipedia revealed individual articles on each topic, making it easier to identify the relevancy of it to the search item. And, Wikipedia conveniently provided an article stub for each year. A stub is an empty article that is ready to receive content. While this approach does not measure all the content on Wikipedia related to a particular year, it does provide one indicator of the amount of coverage and attention given to a year. Additional searches that were not done in the Tankard & Royal study were performed on the recency or currency dimension to help improve this area, including *Time Magazine* Person of the Year, Academy Award Winning Films by Year, and Artists having #1 Songs by Year.

Length of the individual article was all that was included in the Word Count for each topic. One feature of the Web that is also a feature of Wikipedia is the usage of links. Most articles included links to other articles that enhanced or augmented the content of a particular stub. Often these links are tangents, describing other people or events mentioned in the article. Trying to capture the word count of associated links would have made for an unwieldy study.

Information on Wikipedia is extremely volatile and dynamic. Articles can change dramatically over time. This study was performed during November 2006 and each search within a variable was performed on the same day during the same time period, to improve the comparison of that information. This project merely captures the presence of information in the timeframe under analysis. Some of the biases uncovered may subside or change over time. So, while this study uncovered important biases in information being presented on Wikipedia, it will be important to continue research in the area of measuring both accuracy as well as completeness of information on online sites that are becoming important information resources, particularly those taking advantage of the democratic and open source features of the technology.

What's on Wikipedia, and What's Not...?
Completeness of Information on the Online Collaborative Encyclopedia

Bibliography

- Berkowitz, D. (1990). Refining the gatekeeping metaphor for local television news. *Journal of Broadcasting & Electronic Media*, 34, p. 55-69.
- Bopp, R.E. & Smith, L.C. (2001), Reference and Information Services: An Introduction, 3rd ed., Englewood, Colorado, Libraries Unlimited, p. 436.
- Bradley, J. V. (1982). The insidious L-shaped distribution. *Bulletin of the Psychonomic Society*, 20, p. 85-88.
- Curtin, P. A., & Rhodenbaugh, E. (1999). It's not easy being green: Building the news media agenda on the environment. *Association for Education in Journalism and Mass Communication*. New Orleans, LA.
- Danielson, W. A., & Adams, J. B. (1961). Completeness of coverage of the 1960 campaign. *Journalism Quarterly*, 38, p. 441-452.
- Denning, P., Horning, J., Parnas, D., & Weinstein, L. (2005, December). Wikipedia Risks. *Communications of the ACM*.
- Giles, J. (2005, December). Internet Encyclopaedias Go Head to Head. *Nature*.
- Hafner, K. (2006, June 17). Growing Wikipedia Revises Its 'Anyone Can Edit' Policy." *New York Times*.
- Half of All U.S. Internet Users Visited News Sites in June 2006 (2006, August 7), ComScore Media Metrix Press Release, <http://www.comscore.com/press/release.asp?press=971>.
- Heylighten, F. (1995). From World-Wide Web to Super-Brain. *Principia Cybernetica Web*. Retrieved April 22, 2002, from <http://pespmc1.vub.ac.be/SUPBRAIN.html>
- Holiday Fever Drives Traffic to Shopping Sites in December (2007, January 16., ComScore Media Metrix Press Release, <http://www.comscore.com/press/release.asp?id=1177>.
- Johnson, G. (2006, January 3). The Nitpicking of the Masses vs. the Authority of the Experts. *New York Times*.
- Kariel, H. G., & Rosenvall, L. A. (1984). Factors influencing international news flow. *Journalism Quarterly*, 61, 509-516.
- Lih, A. (2004, April 16-17). Wikipedia as Participatory Journalism: Reliable Sources? Metrics for evaluating collaborative media as a news resource. *5th International Symposium on Online Journalism*. University of Texas at Austin.
- McMillin, D .C. (1996). Roles journalists play: An examination of journalists' roles as manifested in samples of their best work. *Association for Education in Journalism and Mass Communication*. Anaheim, CA.

What's on Wikipedia, and What's Not...?

Completeness of Information on the Online Collaborative Encyclopedia

McQuail, D. (1992). *Media performance: Mass communication and the public interest*. London: Sage.

Miller, B.X., Helicher, K., & Berry, T. (2006, April 1). I Want My Wikipedia. *Library Journal*.

Shariatmadari, D. (2006, Jul/Aug). Is A Million Articles Proof of Authentic Information? *Intermedia*, 34, 3, p. 17.

Shoemaker, P., & Reese, S. (1995). *Mediating the message: Theories of influences on mass media content*. Reading, MA: Addison-Wesley.

Tankard, J. W. & Royal, C. L. (2005), Finding Out What's On the World Wide Web, *Communication Impact: Designing Research that Matters*, Susanna Hornig Priest, Ph.D., Editor, p. 253-264.

Tankard, J. W. & Royal, C. L. (2005, Fall), What's on the Web and What's Not, *Social Science Computer Review*.

Tankard, J. W., & Showalter, S. S. (1977). Press coverage of the 1972 report on television and social behavior. *Journalism Quarterly*, 54, p. 293-298.

The Wiki Principle. (2006, April 22). *The Economist*, 379, p. 14-15.

Udell, J. (2004, January 9). *Wikipedia's Future*. Retrieved from <http://www.infoworld.com>

Voss, J. (2005). Measuring Wikipedia. In *Proceedings International Conference of the International Society for Scientometrics and Informetrics*. Stockholm (Sweden).

Wallace, D. & Van Fleet, C. (2005). The Democratization of Information? Wikipedia as a Reference Resource. *Reference and User Services Quarterly*, 45.

Wikipedia:About (2007), <http://en.wikipedia.org/wiki/Wikipedia:About>, accessed 1/23/2007.