Readability Assessment of Policies and Procedures of Social Networking Sites

Gabriele Meiselwitz

Department of Computer and Information Sciences,
Towson University
8000 York Road
Towson, MD, 21252, USA
gmeiselwitz@towson.edu

Abstract. Many internet users today are members of social network sites, building personal profiles and interacting with millions of users worldwide. These virtual environments are based on Web 2.0 technology and offer rich user interaction, personalized use of the environment, and the option for sophisticated user-created content. Some of these environments have developed into large communities with complex relationships within the community, which are covered by policies and procedures. Users accept these when they sign up with the site, and many find that these policies and procedures can be quite complex and difficult to read. A large number of participants in these environments are children or teenagers, making it even more important to ensure that all users fully understand what these policies and procedures entail. Even adult users often have trouble understanding and applying the policies and procedures, and in many cases users just accept the default when registering with the site. This paper addresses the readability of such statements and evaluates the comprehension difficulty of standard policies and procedures of selected social network sites. It concludes with a summary and suggestions for future research.

Keywords: Readability, Comprehension Difficulty, Policies, Procedures, Social Networking Sites.

1 Introduction

Social networking sites are very popular with teenagers and adults, and many internet users participate in more than one social network. The growth of social network users is part of a technological movement referred to as Web 2.0, which involves user participation, where users are actively producing and sharing information [9,11]. Expanding on the social connection in Web 2.0, the semantic web (sometimes referred to as Web 3.0) promotes structuring data on the web to enable better organization of web content and facilitate better information sharing and collaboration [11]. These trends of sharing personal and other information and organizing this information make it necessary to define relationships within and between online social communities, and most communities cover these in their online policies and procedures.

A.A. Ozok and P. Zaphiris (Eds.): OCSC/HCII 2013, LNCS 8029, pp. 67–75, 2013.

The majority of teens (82%) who are online participate in social networks, and over 50% of adults have accounts on social networks, with many internet users having more than one account or profile and/or participating in several social networks [8]. Currently, most social networking sites have default policies that users agree to when signing up for participation in such sites. Although most sites offer changing of certain elements in their default policies and procedures, users tend not to change the default settings, experiencing changing of the defaults as a burden, and even perceiving them as authoritative recommendations [12]. In addition, many users experience difficulty when reading these policies; they find them overly complex, difficult to understand, and have little knowledge about how and when sites may change their policies and how these changes will be communicated to the user [5,6,10]. Users frequently decide not to read policies because it takes considerable time and efforts to locate, read, and analyze them. Studies have shown that it took users an average of 35 minutes to locate and analyze privacy policies of e-commerce sites, and that it can take analysts up to several hours to examine each policy [10].

In higher education, many instructors already are using Facebook groups, wikis, or blogs; and current Learning Management Systems like Blackboard offer synchronization with student Facebook accounts. New learning paradigms such as connectivism, distributed cognition, and communal constructivism address a shift to community knowledge and learning, making it a necessary for educators to consider social networking integration into their instruction and in support of departmental and faculty collaboration [9].

This project evaluates terms of service policies of twenty social networking sites with possible application in higher education. Policy accessibility is assessed by evaluating how users can locate the policy (e,g, is there a link easily accessible) and how the policy is presented to the user (e.g. one or several pages, downloadable as pdf). Readability is assessed by using several instruments to compute grade level readability scores.

2 Research Method

Twenty social networking sites are evaluated regarding the accessibility and readability of their site specific published policies and procedures. Accessibility of a site's policy for this study is defined as the ease of accessing the policy. Users must be able to find and access the site easily to enable them to read the policy. Frequently, the policy is presented at the time a user creates an account at the site. However, many users elect not to read the policy at all at this time, read only parts of it, or give the text a brief scan, ready to continue with the process of account creation [10]. Moreover, it is preferable to not tie the policy to the sign up process and give users the option to read the policy before they decide to sign up for an account. Therefore, accessibility of the policy was evaluated by assessing the location and presentation of the policy. For example, is the policy easily available on the homepage, does the user have to go through several clicks to access the policy, what is the length of the policy, and what is the file format of the policy (html, pdf, or other).

Readability of the policy assesses the degree to which a user understands and comprehends the content of policy. Several standardized tests are available to assess the

readability of text; the most often used is the Flesch Reading Ease Score (FRES) [3]. The FRES computes a final reading score considering average sentence length and average number of syllables per word. Longer words and sentences are more difficult to read and produce a lower reading score. A higher score means an easier reading level [3]. FRES has been used for decades ubiquitously and now is often bundled into word processing software; it is used by the Department of Defense, educational institutions, and several states require that some of their legal documents are written at a particular reading level according to FRES [3,5]. The final score as computed with the FRES can then be translated into a Flesch Grade Level which maps the FRES score to a U.S. grade school level, making it easier to associate the readability level of a certain text with a grade school education level [3,5].

Although the Flesch Reading Ease Score is the most widely used method, other tests are available to assess the reading level. The Gunning-Fog and SMOG indices use a similar method as the FRES, and also compute a final score of grade level. FRES uses total syllables; FOG and SMOG are based on the use of complex words (words with three or more syllables). The Coleman-Liau (CL) index also produces a final score of grade level, however, it relies on numbers of characters rather than the syllable/word approach [3].

For the purpose of this study, multiple methods are used to evaluate the policies of social networking sites. Text will be evaluated using the Flesch grade level, Gunning-Fog, SMOG, and Coleman-Liau assessment to compute an average of the reading level.

One may want to assume that the majority of high school students entering college reads at a 12-th grade reading level, however, research indicates a different trend. The average reading level of adults in the U.S. is at a 7th grade reading level, and research shows that the average reading performance of 12th- grade high school students has been steadily declining [4]. Although female students outperform male students and the percentage of students with parents who graduated from college increased, reading levels have fallen since 1992 (actually since 1985) across all groups [4, 13]. Studies also show that textbooks used by college freshmen are often several grade levels above their reading ability; sometimes over 50% of students have a reading level of several grades below that of the text they are using [2, 7].

Sites for this project were selected using the Wikipedia/Alexa Top 500 social network compilation [14]. The following selection criteria have been applied: sites have to focus mainly on the U.S., must have applicability in higher education, and sites with a higher number of registered users were selected over sites with a lower number of registered users.

The program used to calculate the readability scores of social network site policies in this project is Text-Statistics, a public domain program providing several methods to compute readability scores. The tool was selected because it is available free of charge, calculates several scores, and is able to process HTML files (it will strip all html code from the text file) [1].

3 Results

Results include the evaluation of twenty sites, and evaluate the terms of service policies of the site. Results are discussed by a) evaluating the Flesch Reading Ease Score (FRES)

and the Flesch Grade Level (FGL), b) evaluating the average of FGL, FOG, SMOG, and CL and c) comparing results achieved with FGL and FOG, SMOG, and CL.

3.1 Flesch Reading Ease Score and Flesch Grade Level

Table 1 displays the sample of social networking sites including URL, the total number of words, total number of pages, Flesch Reading Ease Score, and Flesch Grade Level. Since all documents were online in HTML format, the number of pages was calculated by allocating 500 words per single spaced page. The total numbers of pages reanges from 3-27, with an average of 9.5 pages per policy.

Table 1. Overview of Sites including URL and FRES and FGL

Site	URL	Words	Pages /500wds	FRES	FGL
Facebook	www.facebook.com	4583	9.17	55.9	8.7
Twitter	www.twitter.com	3504	7.01	42	13.9
Google+	plus.google.com	1699	3.40	53.7	10.5
Pinterest	www.pinterest.com	2290	4.58	49.3	10.6
Mylife	www.mylife.com	6643	13.29	49.4	10.1
Friendster	www.friendster.com	2969	5.94	31.5	16.4
Secondlife	www.secondlife.com	11600	23.20	39.3	14.5
Flickr	www.flickr.com	5656	11.31	47.1	10.2
Youtube	www.youtube.com	3828	7.66	43.7	12.9
Tumblr	www.tumblr.com	5202	10.40	41.4	13.7
Wikipedia	www.wikipedia.org	5817	11.63	38.5	13.9
Xanga (Blog)	www.xanga.com	2733	5.47	31.3	15.7
Blogger	www.blogger.com	1699	3.40	53.7	10.5
Livejournal	www.livejournal.com	5489	10.98	35.2	13.7
Blogspot	www.blogspot.com	1699	3.40	53.7	10.5
Edublogs	www.edublogs.org	2463	4.93	39.8	12.8
LinkedIn	www.linkedin.com	6424	12.85	35.1	13.5
Xing	www.xing.com	3728	7.46	48.7	10.6
Ziggs	www.ziggs.com	3273	6.55	35.9	15.9
Ning	www.ning.com	13654	27.31	37.4	15
Average		**4747.65**	**9.50**	**43.13**	**12.68**
Standard Deviation		**3140.60**	**6.28**	**7.82**	**2.29**

For the social networking sites evaluated, the average FGL was 12.68 (SD=2.29). The most difficult Terms of Service had a FGL of 15.9, roughly equivalent to someone with an undergraduate college degree. The most readable policy required a reading grade level of 8.7, comparable to the reading level of a high school freshman.

All sites have their policies available on the main page, and have direct links to "Terms of Service" and "Privacy". Facebook, however, has a sophisticated structure for policies with several links to the main policy page, and some policies are linked to other policy pages. In addition, Facebook also uses a different naming convention than the other evaluated 19 sites; on the Facebook Site, Terms of Service are referred to as "Terms and Rights and Responsibilities", and Privacy policies are referred to as "Data Use Policy", and include several sections and links.

3.2 Reading Grade Level Using Average of Several Instruments

Table 2 summarized the grade level scores using FGL, FOG, SMOG, and CL. Using several instruments, the average of all surveyed sites is 13.02, representing the reading ability of a second year college student. The site with the highest average reading grade level is Friendster, with an average reading grade level of 15.20, the site with the most readable policy again is Facebook with a reading grade level of 10.23, representing the reading level of a college sophomore. Gunning-Fog produces the highest scores, with some scores ranging up to 18, which is comparable to a post graduate college education. This may possibly be due to the fact that the Gunning-Fog index does not take into account that not all multi-syllabic words are difficult, and thus may result in a higher score of difficulty.

Table 2. Average of Sites using FGL, FOG, SMOG, and Coleman-Liau

Site	FGL	FOG	SMOG	Coleman-Liau	AVG FGL/ FOG/SMOG/CL
Facebook	8.7	10.1	8.9	11.7	9.85
Twitter	13.9	16	12	13	13.73
Google+	10.5	12.7	10.1	13	11.58
Pinterest	10.6	11.7	10.3	13	11.40
Mylife	10.1	11.9	9.7	13.5	11.30
Friendster	16.4	18.4	13.6	13.6	15.50
Secondlife	14.5	16.1	12.3	12.9	13.95
Flickr	10.2	11.8	9.5	14	11.38
Youtube	12.9	14.3	11.4	13.2	12.95
Tumblr	13.7	16	12.5	13.8	14.00
Wikipedia	13.9	16.4	12.5	13.2	14.00
Xanga (Blog)	15.7	17.2	13.1	13.8	14.95
Blogger	10.5	12.7	10.1	13	11.58

Table 2. (*Continued*)

Livejournal	13.7	15.1	12.3	14.9	14.00
Blogspot	10.5	12.7	10.1	13	11.58
Edublogs	12.8	15.2	12.2	14.2	13.60
LinkedIn	13.5	14.7	12.7	14.5	13.85
Xing	10.6	12.9	10.1	12.5	11.53
Ziggs	15.9	18	13.6	13	15.13
Ning	15	16.8	12.6	13.9	14.58
Average	**12.68**	**14.54**	**11.48**	**13.39**	**13.02**
Standard Deviation	**2.29**	**2.35**	**1.47**	**0.73**	**1.60**

3.3 Categories

Figure 1 summarizes the readability status of the terms of service policies of all evaluated social networking sites. The figure illustrates the distribution of site readability across grade levels. Evaluation shows that most policies (55%) require users to read on a college level, 40% of evaluated sites require users to read on a high school 11th/12th grade reading level. Only one of the sites has a 9th/10th grade reading level, and no sites have a reading level below high school. It should also be considered that this assumes the best case that a graduating high school senior entering college actually reads on a 12th grade reading level. Unfortunately, this only applies to some students.

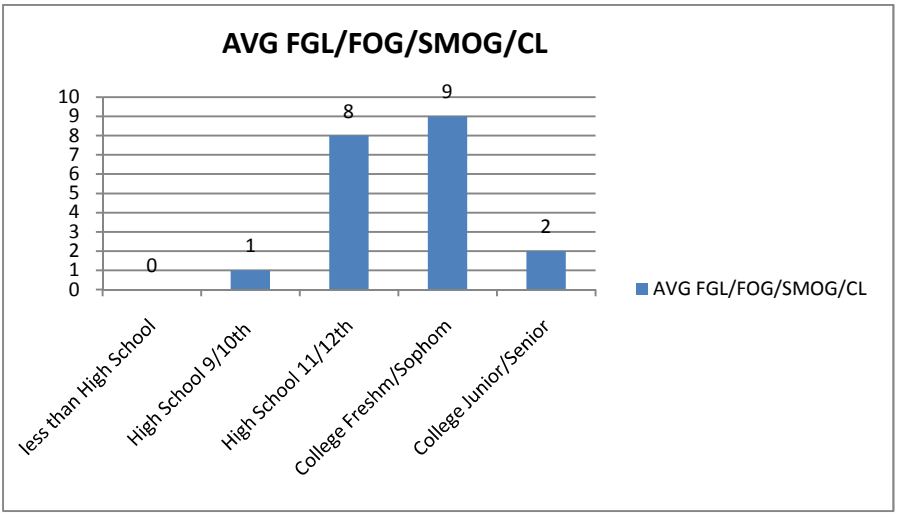

Fig. 1. Categories using FGL, FOG, SMOG, and CL Average

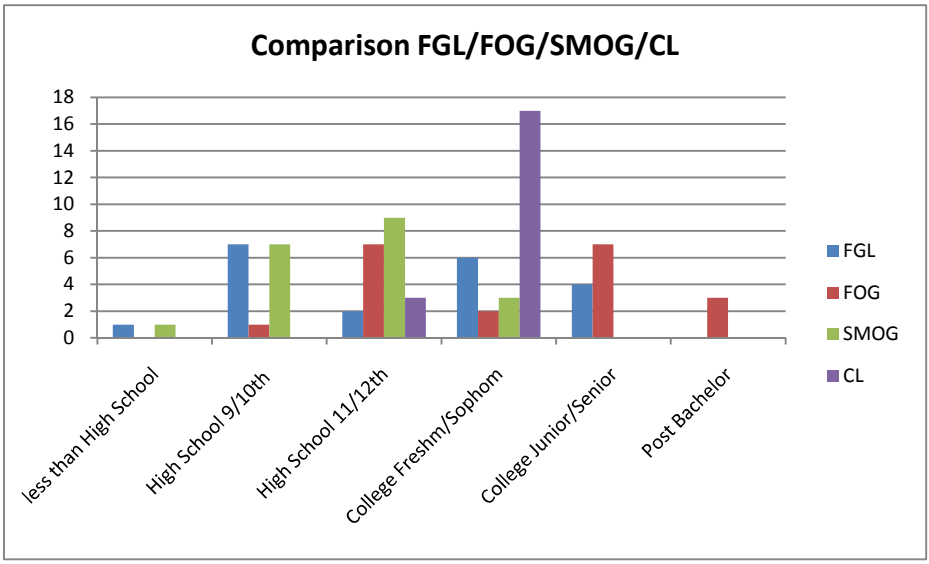

Fig. 2. Categories comparing FGL, FOG, SMOG, and CL

Figure 2 compares all of the used evaluation instruments, the Flesch grade level (FGL), Gunning-Fog (FOG), SMOG, and Coleman-Liau (CL). Comparing all four instruments shows that the majority of sites require an 11th/12th high school grade and college freshman/sophomore reading level. In contrast to FGL, FOG, and SMOG, who all show a somewhat even distribution across high school and college reading levels, the Coleman-Liau index shows a spike at the college freshman/sophomore level, in fact, 85% of all evaluated sites are in this category. Coleman-Liau relies on the total number of characters rather than the syllable/word approach, which may cause this different distribution.

4 Conclusion

This project evaluated the readability of policies of social networking sites. Twenty social networking sites were evaluated, and the evaluation points out that most social networking sites assume a reading grade level beyond a high school reading level. Over half of all sites (51% of the average scores) require a college level reading ability. The average reading level of an adult in the U.S. is a 7th grade reading level, and although the college student population can be assumed to read at a higher level than 7th grade, reading levels of 12th grade high school students have steadily declined over the past years, and many high school seniors entering college read at a level below 12th grade. In addition, considering how many students sign up for social networks before they get to college, the grade level score is clearly beyond the reading abilities of many students at the time they sign up for a user account.

The evaluation concentrated on the main Policy page, the Terms of Services page. Unfortunately, if the main page is requiring a college reading level, it may be

suspected that other sites such as privacy policies or copyright policies are at a similar reading level. If a user encounters a high reading level on the main policy page and is unable to fully process the contents, he/she may get discouraged and not even pursue reading other policies or procedures pages. It should also be considered that some policies are of substantial length, page length assumed 500 words per page (appr. one page single spaced), and ranges between 3 and 27 pages, with an average of 9.5 pages per policy, and reading and processing this amount of material may be quite a task for users to tackle.

Considering that all policies of social networking sites are online and in HTML format, the computer science HCI community could significantly contribute to improve the situation of lengthy policies with high reading grade level requirements. The user interface could possibly consider implementations allowing users to configure the presentation form, such as audio or visual image support on demand; or allow users to decide if they want to see the complete policy or view smaller parts of the policy at a time.

Policy creators may also consider software assisting policy makers in developing, writing, and posting policies for the World Wide Web. Online generators could be made available to assist inexperienced site owners in creating meaningful, readable policies. This project was a pilot study to assess readability of social networking site policies, it is planned to continue research on this project by evaluating additional policies on the site, as well as increasing the number of sites evaluated. It is also planned to conduct a more detailed statistical analysis to evaluate the difference between instruments.

References

1. Child, D.: Text-Statistics (2012), `https://github.com/DaveChild/Text-Statistics`
2. Cline, T.: Readability of Community College Textbooks and the Reading Ability of the Students Who Use Them. Journal of Literacy Research 5(2), 110–118 (1972)
3. DuBay, W.H.: The principles of readability. Impact Information, 1–76 (2004)
4. Grigg, W., Donahue, P., Dion, G.: The Nation's Report Card: 12th-Grade Reading and Mathematics. National Center for Education Statistics (2005), `http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2007468`
5. Jensen, C., Potts, C.: Privacy Policies as Decision-Making Tools: An Evaluation of Online Privacy Notices. In: Proceedings of Computer Human Interaction (CHI). ACM (2004)
6. Kienle, H., Lober, A., Mueller, H.: Policy and Legal Challenges of Virtual Worlds and Social Network Sites. In: Proceedings of IEEE Conference of Requirements Engineering and Law (2008)
7. Kurzmann, M.: The Reading Ability of College Freshmen Compared to the Readability of Their Textbooks. Reading Improvement (1974)
8. Lenhart, A., Purcell, K., Smith, A., Zuckuhr, K.: Social Media & Mobile Internet Use Among Teens and Young Adults. PewResearch Center (2010), `http://pewinternet.org/Reports/201/Social-Media-and-Young-Adults.aspx`

9. Meiselwitz, G., Lazar, J.: Accessibility of Registration Mechanisms in Social Networking Sites. In: Stephanidis, C., et al. (eds.) 2009 Human Computer Interaction Conference Proceedings. ACM (2009)
10. Proctor, R., Ali, M., Vu, K.: Examining Usability of Web Privacy Policies. International Journal of Human-Computer Interaction 24(3), 307–328 (2008)
11. Thompson, J.: Is Education 1.0 ready for Web 2.0 Students? Journal of Online Education 3(4) (2007)
12. Toch, E., Sadeh, N., Hong, J.: Generating Default Privacy Policies for Online Social Networks. In: Proceedings of Computer Human Interaction (CHI). ACM (2010)
13. U.S. Department of Education. Adult Literacy in America. National Center for Education Statistics (2002), http://nces.ed.gov/pubs93/93275.pdf
14. Wikipedia: List of Social Networking Sites, http://en.wikipedia.org/wiki/List_of_social_networking_websites (retrieved online December 22, 2012)