Are Online Privacy Policies Readable?

M. Sumeeth  
University of Alberta, Canada

R.I. Singh  
University of Alberta, Canada

J. Miller  
University of Alberta, Canada

ABSTRACT

This paper examines the question of are on-line privacy policies understandable to the users of the Internet? This examination is undertaken by collecting privacy policies from the most popular sites on the Internet, and analyzing their readability using a number of readability measures. The study finds that the results are consistent regardless of the readability measure utilized. The authors also compare their findings with the results from previous studies. The authors conclude that, on average, privacy policies are becoming more readable. However, these policies are still beyond the capability of a large section of Internet users, and roughly 20% of the policies require an educational level approaching a post-graduate degree to support comprehension.

1. INTRODUCTION

Human readable privacy policies are widely used in websites as they allow users to interpret privacy policies without machine intervention. An important factor in these policies is their readability. The Federal Trade Commission (FTC) describes a privacy policy to be a comprehensive description of: a domain’s collection of user-related information, located on a website that may be accessed by clicking on a hyperlink (Federal Trade Commission, 1998).

The Graham Leach Bliley Act (GLBA), states that policies must be “clear and conspicuous”. GLBA’s privacy requirements state “organizations or institutions should post a notice that is reasonably understandable and designed to call attention to the nature and significance of the information in the notice” (Anton, 2004). Hence,
privacy policies are examined for their coherence, readability and information they convey to users. Klare defines readability as “the ease of understanding or comprehension due to the style of writing.” (Klare, 1963). This definition focuses on the writing style of privacy policies. Similar emphasis on writing style and clarity is mentioned by Hargis et al. (1998). The SMOG readability formula defines readability as: “the degree to which a given class of people find certain reading matter compelling and comprehensible.” (McLaughlin, 1969). This definition stresses the interaction between the text and a class of readers of known characteristics such as reading skill, prior knowledge, and motivation. Perhaps, Dale and Chall provided the most comprehensive definition of readability:

The sum total (including all the interactions) of all those elements within a given piece of printed material that affect the success a group of readers have with it. The success is the extent to which they understand it, read it at optimal speed, and find it interesting. (Dale, 1949).

The focus of most readability definitions is on clarity, ease-of-readability, and level of comprehension. However, the question of are current privacy policies published on websites meeting these requirements has received limited attention. These policies are usually long and unstructured documents that are difficult for users to read and understand. Lack of clarity and understanding of privacy policies has resulted in several privacy-related complaints. Hence, we urgently need more research, which initially assesses the readability of these policies; and if they are found it be deficient, research to improve their readability.

The remainder of this paper seeks to address this first research question and is structured as follows: Section 2 covers some legal issues regarding privacy policies. Section 3 briefly outlines the requirements of a readability policy. Section 4 introduces the approaches used to estimate readability in this study. Section 5 reviews previous empirical research on this topic; and Section 6 introduces the study’s experimental design and analyses the principally results from our study. Section 7 provides a comparison of the results from our study (Section 6) against previous results (Section 5). And Section 8 concludes the paper.

2. PRIVACY POLICIES: THEIR REQUIREMENTS AND RESTRICTIONS

An overview of privacy policies would seem to be the first order of business at hand. To start with, are privacy policies likely to be read? And then there are also questions about legalities, user characteristics and privacy policies. Readability is an issue to be addressed as well as the contents of an ideal privacy policy. These issues and other will be explored further in this section.

Regarding how probable is it that a privacy statement gets read; is easy to bring forth anecdotal evidence when it comes to speculating how likely privacy statements are to be read. Far more difficult is to find empirical evidence on the subject. In a study of 2,468 adult US internet users, it was found that reading rates are linked to several factors. One factor is the concern for privacy. Another is positive ideas about the understandability of the notice. Related to this are higher degrees of trust in the privacy notice. (Milne & Culin, 2004). The reader is referred to Milne and Culin (2004) for further reading.

A literature review was conducted to find out the implications of a privacy policy being legally and contractually binding. A contract can contain almost anything; the question is if it is enforceable. And therein lies the starting point of the literature search, namely what laws apply to privacy in which countries. Generally speaking there are two levels of law governing privacy: legal jurisdictional laws and sectoral laws. Legal jurisdictional laws relate to the “laws of the land” and sectoral laws “govern