Quality and Acceptance of Crowdsourced Translation of Web Content

Ajax Persaud, University of Ottawa, Ottawa, Canada
Steven O’Brien, University of Ottawa, Ottawa, Canada

ABSTRACT
Organizations make extensive use of websites to communicate with people. Often, visitors to their sites speak many different languages and expect that they will be served in their native language. Translation of web content is a major challenge for many organizations because of high costs and frequent changes in the content. Currently, organizations rely on professional translators or machines to translate their content. The challenge is that professional translations is costly and too slow while machine translations do not produce high quality or accurate translations even though they may be faster and less expensive. Crowdsourcing has emerged as a technique with many applications. The purpose of this research is to test whether crowdsourcing can produce equivalent or better quality translations than professional or machine translators. A crowdsourcing study was undertaken and the results indicate that the quality of crowdsourced translations was equivalent to professional translations and far better than machine translations. The research and managerial implications are discussed.

KEYWORDS
Crowd Diversity, Machine Translators, Professional Translators, Task Repetition, Web 2.0

INTRODUCTION
Crowdsourcing takes a job normally performed by a designated person and having it done by a large, undefined, and dispersed number of participants (Howe, 2008). Participants are recruited online to perform defined, sub-autonomous tasks (Behrend, Sharek, Meade, & Wiebe, 2011). Crowdsourcing leverages Web 2.0 tools (O’Reilly, 2007) to foster intellectual cooperation among human collectives in order to create, innovate and invent (Grasso & Convertino, 2012).

Crowdsourcing has been used in a variety of contexts including ideation generation (Schweitzer, Buchinger, Gassmann, & Obrist, 2012), citizen journalism (Tilley & Cokley, 2008), disaster and emergency management (Dailey & Starbird, 2014), scientific collaboration (Jirotka, Lee, & Olson, 2013), health care management (Adams, 2014), and linguistics (Munro et al., 2010). The study by Eagle (2009) is an early attempt to use crowdsourcing to translate single words and terminologies that have specific meanings among particular groups (e.g., engineers, accountants). Ledlie, Odero, Minkov, Kiss, and Polifroni (2010) used crowdsourcing for speech recognition translation i.e. converting speech to text. The use of crowdsourcing to translate website content that is longer, more descriptive, open to many interpretations and targeted to the general public rather than specific groups is a qualitatively more complex and difficult undertaking that has not yet been done but needed (Snow, O’Connor, Jurafsky, & Ng, 2008).

Translation of website content is important for several reasons. Today, users turn to the web as their first source to find information on virtually any topic, product or service. Therefore, it is important for organizations, businesses, and governments to make their content accessible to users in their native
languages. Hutchins (2001) contends that people prefer to read content in their first language, even with some errors, than struggle to understand a website in a foreign language. DePalma, Sargent, and Beninatto (2006) found that language quality is a major factor influencing consumers’ online purchase decisions or online interactions with organizations. Consequently, it seems that organizations serving people that speak different languages should offer web content into multiple languages.

There is also an important social challenge pertaining to the digital divide associated with language barriers. Although the digital divide is often used to describe whether or not people have access to the Internet or information and communication technologies (ICTs), Selwyn (2004) posits that it can also be seen as a practical embodiment of the wider theme of social inclusion or digital exclusion. Selwyn (2004) argues that access to the Internet does not guarantee that users can effectively access every available website and online resource. Similarly, Van Dijk and Hacker (2003) argue that access to a technology is useless without the necessary skills, knowledge and support to use it effectively. According to Selwyn (2004), ‘meaningful’ use of the web requires that users are able to engage with the content. Moreover, content is only useful if it is relevant to the user and those who are isolated or marginalized face a digital divide (Selwyn, 2004). Thus, not serving users in their first language is a form of digital divide that needs bridging.

Translating website content is a formidable challenge because of the enormous amount of content on websites, which tend to change often. Translating web content using professional translators is slow and costly. On the other hand, machine translation, while faster and less expensive, produces inferior quality translations (Bowker, 2008). These challenges combined with the growing use of crowdsourcing motivate the question posed in this study: Can crowdsourcing produce equivalent or higher quality translations compared to those produced by professional or machine translators?

At a practical level, if crowdsourcing produces high-quality translations quickly, reliably, and cost-effectively, then organizations would have yet another option to perform this important task. At a theoretical level, this study provides evidence on the validity of the principles of crowdsourcing and illustrates how this new technology may effectively alter the processes and human interactions involved in language translation.

The remainder of this article presents a review of the relevant literature, the research questions, hypotheses, and framework, the methodology, the analysis and results, and the discussion and conclusion.

RELATED WORK

Crowdsourcing belongs to a new class of collective intelligence tools that facilitate computer supported cooperative work (Grasso & Convertino, 2012). These new generation of collective intelligence tools, often described as Web 2.0 technologies (O’Reilly, 2007), have the ability to change the way people work and how work is organized (Grasso & Convertino, 2012). Essentially, the Internet and these Web 2.0 tools enable crowds of users from anywhere in the world to work collectively by leveraging their social and professional networks to create and share knowledge and undertake complex tasks (Grasso & Convertino, 2012). Crowdsourcing takes a task that is normally done by designated professionals or experts and having it performed by a large, undefined group of individuals (Howe, 2008). The tasks are usually broken down in numerous sub-tasks and each individual within the crowd performs various sub-tasks (Brabham, 2008). Payments for work performed vary from no monetary reward or very small amounts to large amounts, which are paid by large corporations (Hossain & Kauranen, 2015).

In a recent review of the literature, Hossain and Kauranen (2015) reported that scholarly research on crowdsourcing is in the early stages but growing despite its widespread use among practitioners.
Related Content

Script Identification of Camera Based Bilingual Document Images Using SFTA Features

Adherence: A Behavioral Economists’ View
[www.igi-global.com/article/adherence/166569?camid=4v1a](www.igi-global.com/article/adherence/166569?camid=4v1a)

Challenges for Libraries in the Information Age
[www.igi-global.com/chapter/challenges-libraries-information-age/55829?camid=4v1a](www.igi-global.com/chapter/challenges-libraries-information-age/55829?camid=4v1a)
Reflections on U-PriSM 2: The Second Workshop on Usable Privacy and Security for Mobile Devices
International Journal of Mobile Human Computer Interaction (pp. 73-78).
www.igi-global.com/article/reflections-on-u-prism-2/112032?camid=4v1a