

Chapter III

Machine Translation as the Future of International Online Communication

Nicole St. Germaine-Madison
Texas Tech University, USA

ABSTRACT

American “ownership” of the World Wide Web is on the decline. In fact, two-thirds to three-fourths of Internet users are not native speakers of English. It thus does not require a leap of logic to imagine how many more non-U.S. users would buy from American e-commerce sites or even purchase American-manufactured products if they had access to translations of these e-commerce pages and technical documents in their native languages. The most oft-cited reason for not providing content in other languages, however, is the sheer cost involved of first hiring a translator to translate the original content and then keeping material updated. This chapter examines uses of machine translation as a mechanism for addressing these linguistic needs.

INTRODUCTION

The Internet used to be an American product. It was largely developed by researchers at MIT and various other universities as a means of connecting researchers and developers, and since its commercialization it has traditionally also been used as a vehicle for American information and American commerce. As a result, virtually all

of the content was, and is, written in English (Internet Society, n.d.).

American “ownership” of the World Wide Web is on the decline. According to the Online Computer Library Center’s annual Web Characterization Project, the most cited resource for World Wide Web statistics, while only 55% of sites of the World Wide Web were American in origin as of 2003, 72% of the total Web sites are

in English. This statistic has remained stable since the first Web Characterization Project report in 1999.

Paradoxically, two-thirds to three-fourths of Internet users are not native speakers of English (Cronin 2001, p. 16). Nor do the majority of these non-native speakers necessarily prefer to read Web sites in English: 50% of Web users in Europe, 75% of users in Latin America, and 80% of users in Japan prefer to read Web sites in their native language over English (2001, p. 16). According to some estimates, the willingness of a target audience to read a Web site increases by 300% if the site is in their native language, even if they are competent in the other language the site is presented in (Forsyth, 2005) This disconnect represents a large gap where industry and e-commerce can step in to increase their markets. In 2003, the non-U.S. portion of the e-commerce market would account for 46% of a US\$1.3 trillion market (2005, p. 16). It does not require a leap of logic to imagine how many more non-U.S. users would buy from American e-commerce sites or even purchase American-manufactured products if they had access to translations of these e-commerce pages and technical documents in their native languages.

Aside from e-commerce, other types of Web sites could also benefit from having a translation made available. Web sites representing international organizations and research applications would gain more members and more hits if these sites had available translations into other languages as well. Even the Society for Technical Communication (<http://www.stc.org>), an international organization, does not have content available in any language other than English.

The most oft-cited reason for not providing content in other languages is the sheer cost involved of first hiring a translator to translate the original content, and then keeping a subscription at a translation and localization firm in order to keep the site's information continually updated in all of the available languages. Most of these companies,

such as SDL International (<http://www.sdl.com>) and Translation Services USA, hesitate to even promise a delivery time that is faster than two to three weeks *or more*, depending on the size of the Web site, the number of languages, and the commonality of the desired target languages. Each of these factors can greatly inflate the cost and time needed to complete the translation.

So what is a company to do? According to Cronin (2001, p. 112), "The best scenario for updating Web content is simultaneously updating all languages. This might not be possible except with automation." However, automated translation (or machine translation) is still not a reliable method for obtaining a completely accurate translation.

The main objectives of this chapter are:

1. To explain the need for rapid, inexpensive translation for online content
2. To present different case studies for how companies are approaching the production of different language sites
3. To explore the differences of translating for the Web vs. translating for print, as well as to discuss the implications for translation
4. To explain machine translation: a brief history, how it works, and its reliability
5. To explore current uses of machine translation for the Web
6. To make specific recommendations for optimizing the use of machine translation for the Web

THE NEED FOR RAPID, INEXPENSIVE TRANSLATION FOR THE WEB

In 2005, it was estimated that around three-quarters of the approximately one billion people who were conducting business or communicating online were living outside of the U.S., and most of these people did not speak English (Schenker, n.d.). Taken with the staggering size of the profits

9 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/machine-translation-future-international-online/25562

Related Content

Psychological Correlates of Perfectionistic Self-Presentation Among Social Media Users

Anantha Ubaradka, Ayesha Fathima and Shreya Batra (2023). *International Journal of Cyber Behavior, Psychology and Learning* (pp. 1-13).

www.irma-international.org/article/psychological-correlates-of-perfectionistic-self-presentation-among-social-media-users/324089

Social Media Use and Job Performance: Moderating Roles of Workplace Factors

Peerayuth Charoensukmongkol (2015). *International Journal of Cyber Behavior, Psychology and Learning* (pp. 59-74).

www.irma-international.org/article/social-media-use-and-job-performance/135316

Internet vs. Matter: Differences in Students' Concept Development from Elementary through High School

Zheng Yan and Xiufeng Liu (2012). *International Journal of Cyber Behavior, Psychology and Learning* (pp. 60-72).

www.irma-international.org/article/internet-matter-differences-students-concept/75172

Using CMC in order to investigate the language system

Elke Hentschel (2010). *Handbook of Research on Discourse Behavior and Digital Communication: Language Structures and Social Interaction* (pp. 866-875).

www.irma-international.org/chapter/using-cmc-order-investigate-language/42823

The Evolution of E-learning Management Systems: An Ethical Approach

Nuno Sotero Alves da Silva, Gonalo Jorge Morais da Costa, Mary Prior and Simon Rogerson (2011). *International Journal of Cyber Ethics in Education* (pp. 12-24).

www.irma-international.org/article/evolution-learning-management-systems/56105