

Corporate Deployment of Machine Translation: Case Study

One of the first major corporate deployments of machine translation shows that a wide range of application benefits can be obtained in a multinational environment, if realistic expectations are set regarding the quality of the output.

Core Topic

Emerging Trends and Technologies: The Future of Technology, Business and Society

Key Issue

Which emerging and embryonic technologies should early adopters be examining for competitive advantage?

Machine translation automatically translates text in one language (for example, English) into another language (for example, French or Japanese). The technology has been available for decades, but because of the inability of computers to understand the context and meaning of human language, the quality of translations created by a machine translation system is significantly lower than translations created by professional human translators. However, there are many situations in which the "imperfect but fast" approach of machine translation can bring worthwhile benefits, as PricewaterhouseCoopers (PwC) discovered in its global deployment of machine translation.

Problem: The impetus for PwC to investigate machine translation was that senior management in the multinational enterprise found it suboptimal to communicate in English when it was not their native language. As part of a broader knowledge management initiative, PwC set up a working group to investigate language issues. Because of technology constraints, the group focused on addressing the written language problem, as automated translation of speech is beyond the state of the art.

Objective: The objective of the initial investigation was to determine the potential benefits, costs and user acceptance of various language technologies, including machine translation. This was more difficult than quantifying the benefits of related technologies, such as translation memory (a tool for translators that stores bodies of pretranslated text), because machine translation was to be targeted as part of the desktop, more like a search engine.

Approach: Having identified machine translation as one of the key language technologies that warranted investigation, the working group tested multiple machine translation systems online. Three vendors set up small pilots as an Internet service

Gartner

Entire contents © 2002 Gartner, Inc. All rights reserved. Reproduction of this publication in any form without prior written permission is forbidden. The information contained herein has been obtained from sources believed to be reliable. Gartner disclaims all warranties as to the accuracy, completeness or adequacy of such information. Gartner shall have no liability for errors, omissions or inadequacies in the information contained herein or for interpretations thereof. The reader assumes sole responsibility for the selection of these materials to achieve its intended results. The opinions expressed herein are subject to change without notice.

and entered into pricing discussions. PwC translators tested the accuracy and readability of the translations. Based on these activities, PwC selected Systran, because it offered state-of-the-art translations and was willing to defer payment until the system was up and running (compared with other vendors that asked for multimillion-dollar payments upfront).

PwC and the vendor spent the next three months in further testing and in setting up the authentication and access to the machine translation as part of PwC's globally deployed intranet KnowledgeCurve. KnowledgeCurve incorporated knowledge repositories and general information supported by desktop tools such as search, translation and personalization. Speed of deployment was helped because PwC had processes in place for integrating content and external applications, such as those from Dow Jones. Additionally, the availability of glossaries from different translation teams was useful in customizing the translation engine in the early rollout phase.

Operational access to the machine translation system was established on an application service provider (ASP) basis hosted at the Systran site, seamlessly available via KnowledgeCurve or any other network-based application (for example, local implementations of KnowledgeCurve). Once the user registered, the user ID was passed on to Systran with appropriate levels of encryption, providing the user with instant access to the translation service.

Based on the success of an initial deployment in Spain, other countries were added one at a time by customizing the system for each new language and adding authenticated access for the users in that country. PwC paid for the service under a global agreement, with individual countries paying for the cost of local customization (that is, adding PwC-specific and department-specific vocabulary to the generic translation dictionaries). List price for a machine translation desktop software license is approximately \$1,000, whereas a corporate service, including setup costs, ranges from \$13,500 per annum for 100 users with five language pairs to \$77,200 per annum for unlimited users and five language pairs.

Results: As with most deployments of machine translation, PwC found the main benefit for users to be understanding the gist of documents in a language they do not speak well. The following uses have been identified to date.

- A major use is translating the results of Web searches of internal or external sites. Finding the relevant content and Web pages prior to translation is a separate and ongoing issue. PwC's selected search engine does not yet support cross-lingual search (as is the case with most major search

engines), and smaller vendors with good cross-lingual search were difficult to scale for global needs.

- Management personnel use the system to get the main ideas from text that they don't understand (for example, internal documents, meeting minutes, client documents, e-mail from overseas clients) to decide whether to have human translation. Frequently, opportunities can be identified by homing in on document types where a significant amount of human translation is already performed, but where understanding the gist is enough for some portion of the documentation (for example, in a time-critical application such as an investment opportunity, where the time delays in professionally translating all relevant documentation would cause missed deadlines).
- Use of the system has allowed specialists to be assigned to project teams where they may not be skilled in the team's language. The machine translation system is used to translate agendas, minutes and project documentation. What helps make this application workable is knowing the context of the documentation (that is, understanding what the project is about, knowing the subject matter and knowing the nature of the document).
- Non-English speakers may use the system to create a document in English, if they do not know the relevant terminology. This is used only for internally targeted documents that are not worth translating professionally. Users become familiar with how to tailor their original text to help the system work better (for example, by using simple, unambiguous language and sentence constructions).

In PwC's case, members of the translation department do not use the machine translation system, because they prefer to use traditional translation tools. The system is not appropriate for legal documents (for example, judgments or statutes) or text that may have legal ramifications. It works well with highly technical documents, but not where the language is abstract or philosophical. Longer term, PwC aims to integrate machine translation with its e-mail system so that users can see the two languages side by side.

Critical Success Factors/Lessons Learned: Factors that helped this project become successful included:

- Gaining buy-in from the professional translators to help evaluate and customize the system, even if they are not the target users
- Making the business case that this is a knowledge management tool targeted at individual productivity, rather than for quantifiable cost savings

- Setting realistic expectations as to the constraints of the technology; the logon screen has a "health" warning telling users what the system can do
- Creating a user interface that allows users to access translation functionality from any application or intranet site without leaving their documents
- Establishing a global license that makes it straightforward to add new country sites
- Making the system customizable for each country (and even individual, if desired). If more budget were available, PwC would have performed more customization to specific businesses, as this makes a big difference in the quality of the translation.

Bottom Line: Enterprises with global operations should examine how machine translation can improve access to internal documentation and international Web and intranet sites, as well as enhancing collaboration between employees in different countries. The technology is not ready for creating documents for external consumption and is not necessarily a benefit for professional translators.