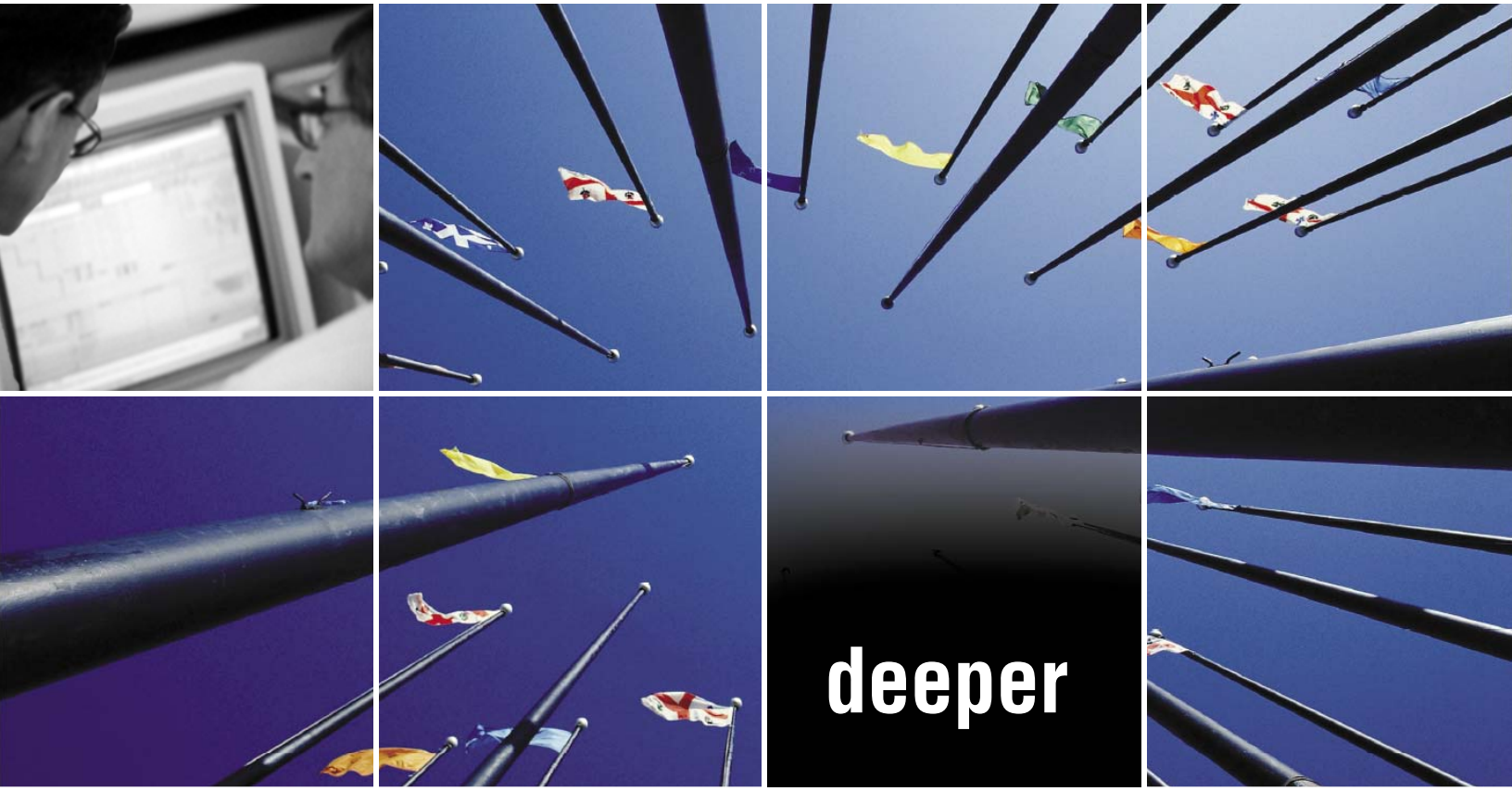
Building blocks, not stumbling blocks

Language strategies ease collaboration in a multilingual world



An IBM Institute for Business Value executive brief

The IBM Institute for Business Value develops fact-based strategic insights for senior business executives around critical industry-specific and cross-industry issues. This executive brief is based on an in-depth study created by the IBM Institute for Business Value. This research is a part of an ongoing commitment by IBM Business Consulting Services to provide analysis and viewpoints that help companies realize business value. You may contact the authors or send an e-mail to iibv@us.ibm.com for more information.

Contents

Building blocks, not stumbling blocks: Language strategies ease collaboration in a multilingual world

Picture a world where technologies allow people to collaborate – in realtime or asynchronously – no matter where they are located. Information posted to internal databases is instantly accessible to a company's employees anywhere in the world. Information access is directly and readily available to customers as well. Company Internet sites allow customers to go online to peruse and order a company's products and services, and obtain many other types of information, such as answers to frequently asked questions and fixes to technical problems.

In many cases, this vision is close to becoming a reality. Internal databases and Internet sites have become easier to search and collaborative technologies are increasingly common on desktops. Organizations are undertaking new approaches to manage in-person and e-mail communication. But, a stumbling block looms large because the primary focus of many organizations has been either:

- To increase collaboration among people who *speak the same language*, or
- To make information widely accessible in *English-only*.

While in an ideal world, machine translation (MT) would translate seamlessly and instantly between languages, technology has not yet caught up with our needs. Thus, to understand how companies currently are addressing language translation related to knowledge management, the IBM Institute for Business Value conducted over twenty in-depth interviews with a cross-section of Fortune 500 companies and government agencies. The interviews explored how these organizations overcame language barriers by using a combination of technology and human solutions.

Knowledge sharing in a multilingual world

Although we do live in a multilingual world, English has become the dominant language within global organizations. At a large, European-based manufacturing company, for example, our study found that fluency in English was one important factor in hiring and promoting mid-level managers. However, some non-English-speaking managers were in place before these criteria were introduced. Even though English probably *will* remain the language of international business, it is highly unlikely that all people sharing information between countries will be able to speak English.

As one translation company has candidly pointed out, global organizations must remember: "The e in e-business does not stand for English."[1]

Of even greater importance to an organization is the fact that many of their *customers* may not speak English. Not surprisingly, most people prefer to view Web sites in their native languages.

Clearly, language can present a barrier to fast and efficient information exchange. This paper focuses on two areas of language-related hurdles encountered by the companies that IBM interviewed: the management of *active communication*, such as face-to-face communication and e-mail, and *stored content*, such as internal databases and Internet sites.

Analysis of the interview results revealed important points to consider when developing a language strategy to support knowledge sharing. Below, we discuss the translation needs for both active communication and stored content and illustrate:

- How individuals in some organizations have adapted to assess and handle translation needs
- The extent to which current technologies can be supportive.

In a later section, we explore the evolution of translation technology methods in greater detail.

Active communication

Active, face-to-face communication between two or more people can occur through direct interaction when everyone is in one place, or be simulated by using technologies such as videoconferencing. In addition, active communication can happen through various synchronous technologies, such as telephone and instant messaging, or asynchronous technologies such as e-mail, fax and electronic teamrooms.

Improving how human interpreters are used

One organization participating in the study had frequent face-to-face meetings between English-speaking executives and their Japanese colleagues. At first, they used an outside vendor to supply translators on an as-needed basis. Not surprisingly, translation time meant that a "one-hour meeting" often took nearly two hours. More unexpected, meeting success was largely dependent on the translator's effectiveness, and not on how well each team presented its case. After several meetings, it became apparent that the translators who had a greater understanding of the context of discussions were far more adept at conveying conversational nuances than those who were merely fluent in the two languages and translated word-for-word.

The executives quickly realized that too much was at stake to continue with this ad hoc method of hiring meeting translators from an outside agency. Their solution was to audition a number of translators and evaluate their skills. They hired those with the most business knowledge and trained them about the organization, its practices and products so that the translators gained the context to more accurately convey what was being said. Building translators into the fabric of the team allowed for more productive and successful meetings.

Face-to-face communication

Even in today's high-tech world, face-to-face communication remains the favored medium of transferring information between people. Despite cutbacks in travel expenditures over the last two years, enormous numbers of people still fly around the globe to see and talk with their colleagues. Face-to-face communication between people without a common language poses a specific challenge in that it is synchronous – it requires instant translation. There are also cost and availability considerations because translators are expensive and not always at hand, especially when conversations occur serendipitously.

From a technology perspective, there are various companies working on realtime speech-to-speech machine translation; as yet, there are no commercial products available. This is an extremely complex endeavor and is unlikely to be in everyday use for several years. In the meantime, we will be reliant on human translators of varying abilities (see Figure 1).

Figure 1. Comparison of translation techniques for face-to-face communication.

Technique	Strengths	Constraints
• Use human interpreters ad hoc	• Costs incurred on as needed basis	• Word-for-word translation doesn't capture all nuances of conversation
• Integrate human interpreters into team	• Can understand business context to convey nuances accurately	• With rising importance of international markets, especially in East Asia, human interpreters must have the right knowledge and skills to fit the situation • Higher costs to permanently add translators to team

Source: IBM Institute for Business Value.

Speeding up the process to translate e-mail via a human focal point

In one organization, there was an important business need for e-mail correspondence between English-speaking executives and their Japanese-speaking colleagues. Initially, e-mails were sent in English and the Japanese recipient would then seek out a colleague to translate them. The process would then be reversed with responses sent back in English. Unfortunately, the transfer of information was slow and dependent on the availability of multilingual people in the Tokyo office.

Over time, however, when a Japanese executive identified a trusted, multilingual person, the process was streamlined by having e-mails from English-speaking executives sent directly to that Tokyo employee. This transition occurred gradually, with the consent of all involved. Each of the trusted, multilingual people in the Tokyo office would answer e-mails that they felt able to handle, but would translate and refer those deemed more critical to their non-English-speaking executive colleagues. After trusted relationships had formed and everyone had clearly defined roles and responsibilities, information flowed significantly faster.

E-mail communication

Successful collaboration via e-mail may also need to address language barriers. Because of e-mail's asynchronous nature, people can seek help on an *ad hoc* basis to translate messages they receive in a different language. However, it is not always possible to find a colleague to assist in a timely manner. And, without the visual clues and voice intonations apparent in face-to-face communication, e-mails can easily be misunderstood – even more so among people who do not share the same language.

The need for e-mail translation suggests a strong need for a technological solution. At the moment, people with at least a minimal understanding of a language are finding that simple desktop translation tools can be very helpful, but human translation is often still needed to clarify meaning (see Figure 2).

Figure 2. Comparison of translation techniques for e-mail communication.

Technique	Strengths	Constraints
• Locate colleague for ad hoc translation	• Low cost and flexible	• Not always possible to locate help when needed • Information transfer slows down
• Use trusted, multilingual employees as e-mail translation focal point	• Established, trusted relationships can significantly improve speed of information flow	• People's abilities and availability vary
• Desktop translation tools	• Relatively cheap to install on PC • Provide rudimentary translation between many pairs of languages	• Require user to have a minimal understanding of source language • Often produce inaccurate results due to inability to understand context behind language, requiring human intervention to clarify

Source: IBM Institute for Business Value.

Stored content

As with active communication, it is important to have a strategy for dealing with the language aspects of stored content. As more organizations are codifying what they know and storing it in internal databases, making that knowledge available throughout the organization becomes increasingly necessary. Similarly, from a customer's perspective, retail outlets and sales representatives are no longer the major source of information about an organization's products and services – instead, people now consult company Web sites for information, as well as to make purchases.

Managing local language postings with a central knowledge coordinator

One global software firm allowed documents to be posted in their internal databases in any language, as long as there was a summary in English. In their various regional offices, one or more knowledge coordinators (KCs) were available to help with localized translation of documents that were relevant to local projects. For example, in the Tokyo office, the KC:

- Translated a weekly listing of new database documents from English into Japanese for distribution throughout the office. Upon request, the KC would translate the summary document and, where necessary, the whole document.
- For translation of a document that was not in English, contacted the document's author or counterparts in other offices to ask if they could translate it into English (or Japanese).
- Supported document submission by writing the required English language summary for anyone wanting to post to the global database.
- Played a proactive, knowledge harvesting role by inquiring whether project leaders within her office had any documents that should be posted.

Internal databases

The primary language challenge with regard to internal databases is one of *complexity,* largely due to the volume of stored information and the number of possible languages into which it might need translation. It is unlikely, due to cost and time constraints, that an organization will translate every database document into every language spoken by its employees. Rather than implementing such a comprehensive solution that is both time consuming and expensive, many organizations are looking for cost-effective, interim solutions. Several organizations participating in the study were establishing specific processes to make documents in their knowledge repositories available to non-English-speaking employees.

Another option is making documents available in multiple languages, which adds complexity and is a requirement for some governments and international organizations. For example, in the case of the Canadian Government, all official documents created in either English or French must be sent to the official translation group for translation into the other language.

Document authors or their colleagues perform translation

Another organization we interviewed was rolling out a new asset management system, but not all of their employees were fluent in English. So, their system allows the individual posting a document to choose from one or more key words from six categories in their local language – key words that are also stored in English. Allowing authors to contribute documents in their local languages was expected to significantly increase the number of postings.

People searching for documents could search the internal database using key words in English or their own languages. For documents in a language they could not read, database links let them check whether they had a language in common with the author or any of his colleagues – if so, they could ask any qualified colleague for a translation.

When developing a language strategy for internal databases, there are two key points to consider. First, determining the percentage of people in the organization who only read – or would prefer to read – a language other than English will help define your translation needs. For example, if 95 percent of your organization has good English comprehension, then an English-only database would suffice. Alternatively, if 50 percent spoke English and the other half spoke up to six other languages, looking at English key word or summary solutions would be worthwhile. Second, whether your organization has a highly centralized internal database or many localized databases will determine how much content is likely to be shared between countries and hence, the need for a translation strategy.

For most organizations, the intermediate solution of posting documents in the local language, while providing some way of making them accessible to others, may be the most fruitful (see Figure 3). In cases of multiple languages, the key word solution may be the better option. In cases of one dominant language, it may be most effective to have local translators assisting people who do not speak that language.

Figure 3. Comparison of translation techniques for internal content.

Technique	Strengths	Constraints
• Post all documents in one language	• Lower costs • Quicker to search documents	• Some employees will not be able to read documents • Some employees will be less likely to post documents
• Post all documents in more than one language	• More employees will be able to read documents	• Need to decide which languages to use for posting • Higher costs • Need for translation infrastructure
• Post documents in local language with summary or key words in English	• Medium costs • Translation can be done locally on an as-needed basis	• Need local translators to create English summaries • Need some level of English to understand key words or summaries
• Post documents in local language	• Lower costs	• May lead to silo effect by language • Many people will not be able to read documents

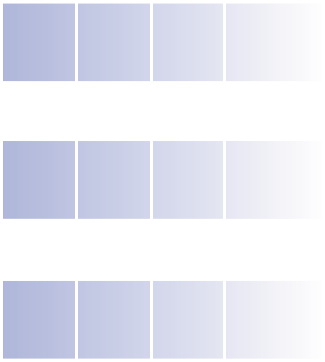
Source: IBM Institute for Business Value.

Internet sites

When devising a strategy for your organization's Internet site, language considerations are similar, but more complex than for internal databases. Though the languages spoken by potential customers could include any of the myriad world languages, cost constraints make it unfeasible for an organization to translate every Web page into every language – rather, the question is largely *which* pages to translate into *which* languages. Translating too much would waste time and money, but translating too little is liable to hinder initial customer contact or alienate customers and lose their business.

Our research suggests that organizations have four options regarding translation of their Internet sites:

- *Create mirror-image sites in multiple languages* – The Canadian government site is one of the most comprehensive examples of this approach: users can toggle between French and English on any page.[2] Obviously, the creation and maintenance costs increase with the number of mirror-image pages. What is practical for the Canadian government working in two languages is likely to be unworkable for a global organization serving customers in thirty or more different languages.
- *Translate some of the content into more than one language* – This partial mirror-image approach allows users to toggle between different languages. An example of this is the U.K. government's site, which is partly translated into Welsh to meet legal requirements.[3]
- *Create country-specific Web sites with the same look and feel, but with different content* – For example, the IBM home pages for Brazil (in Portuguese)[4] and the United Kingdom (in English)[5] have the same layout and many sections are the same, but different content and links are featured. This flexibility allows each country to focus its message on the products and services that are most relevant for that particular market.
- *Maintain country-specific sites* – Johnson & Johnson, a globally distributed company, has completely different Web sites for different countries.[6] This scope allows each country office to create aesthetically pleasing Web pages for its local market, to optimize those customer relationships.



The trade-offs among these approaches are detailed in Figure 4. There are several key points to consider when developing Internet language and content strategies. First, knowing the percentage of your customers who would prefer to read a language other than English will help determine your translation needs. Clearly, the number of supported languages will impact translation costs and requirements. Second, the frequency and extent to which content must be updated also affect translation costs and requirements. Third, the degree of your organization's centralization will govern the ease with which each geographic region's Web pages can be created and updated. For example, in a highly centralized company, it can take time to get approval for any changes to a Web page – although once permission is given, the change might be implemented fairly readily across each region. Alternatively, in a more decentralized company, it is easier to make changes to a particular region's Web site and probably more difficult to replicate these changes across all regions. Finally, organizations must also conform to any applicable legal requirements of individual countries.

Figure 4. Comparison of translation techniques for external content.

Technique	Strengths	Constraints
• Fully translated mirror image	• Accessible to all customers • Consistent message	• Higher costs • Added time for translation and updates • All changes have to be made more than once
• Selective translation	• Updates can be done quicker	• Some items are not accessible to some customers • Message is not consistent
• Tailored with same look and feel	• Consistent message • Flexibility of template allows for addition of new language markets • Local updates can be done quickly	• Corporate updates have to be done in all languages • Many links may be limited to one language • High cost for translation of all sites
• Tailored with different look and feel	• Greater flexibility in making site aesthetically pleasing to customer • Potential increased revenue • All updates can be done quickly	• Higher costs • Inconsistent message • New markets require new templates

Source: IBM Institute for Business Value.

Machine translation technology debuts, but high hopes are soon dashed

On January 7, 1954, a team of IBM researchers and Georgetown University professors met at IBM's midtown Manhattan office to demonstrate their new "$500,000 super-calculator" developed to translate Russian into English. This first-of-its-kind public demonstration of MT was enthusiastically received. True, only one language pair was available (English and Russian), the vocabulary was limited (250 words) and the test examples were simple (all straightforward, declarative sentences). Still, *The New York Times* and *The Christian Science Monitor* both carried prominent articles about this innovation.[7] A Georgetown professor, Dr. Leon Dostert, claimed that in "five, perhaps three years hence, interlingual, meaningful conversation by electronic process in important functional areas of several languages may well be an accomplished fact."[8]

Though this demonstration was impressive, the creators' initial predictions were quite far off. MT did not live up to the hype that surrounded it. In the next few years, follow-on articles – in both popular and technical trade publications – commented on the technology's apparent limitations. However, unrealistic expectations had been raised by Dostert's optimism about the impending, superlative translation systems that did not materialize. Disappointment led to the Automatic Language Processing Advisory Committee (ALPAC) report of 1966, which sent the clear message that MT was hopeless.[9] As a result, large-scale funding of MT research in the United States came to a virtual standstill for the next twenty years.

Beyond human translation: The impact of machine translation

While many of the previous examples highlight effective ways to use human translators, companies are also seeking technical solutions for their growing translation needs. As part of our research, we investigated the state of MT – technologies that translate static or dynamic content. We were interested in how MT has evolved, what technologies are available, how they are being used and what factors to consider in selecting a translation technology.

The resurgence of MT's appeal

Technological advances and economic changes are responsible for a resurgence of interest in machine translation. First, translation technology has improved since it first appeared in the mid-50s, (though not to the extent that Dostert prophesized in 1954 – see sidebar). Second, as noted earlier, the need for translation in business has grown and continues to increase significantly. Third, this burgeoning need has swamped the capacity of human translators. Finally, machine technology has been repurposed to meet the new needs.

Improved technology: Progress, not perfection

Several advances have played an important role in the technology's rebirth. Now, machine technology can process more languages, more specialized vocabularies and more complex sentence structure. In the past few decades, all of the world's alphabets have become candidates for translation software. Support for a variety of fields of expertise has also expanded; for example, specialist dictionaries in law, medicine and information technology are now available for translation in many language pairs. Another significant change is the syntax that technology can process. The first machine translation could only "understand" simple declarative

sentences in the form subject-verb-object. Today's technology is not hampered by such limitations. Questions, compound sentences and complex grammatical phrases are all within the capabilities of modern translation software.

Growing demand challenges human translators

Given increased globalization and Internet access, more and different languages need to be translated. Originally, the U.S. inventors of MT expected Russian to be the most important language since they envisioned the technology being used by the international affairs and intelligence fields. But currently, business people speaking Japanese or Spanish could find themselves needing translation services to communicate with colleagues, suppliers or customers speaking Cantonese or Polish.

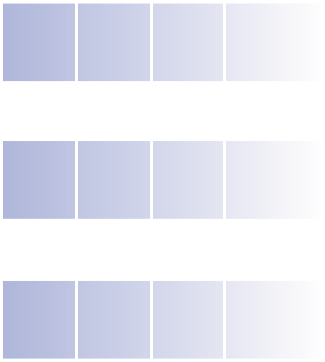
Further compounding demand, human translators must not only be fluent in the source and target languages, but also conversant with any particular industry terminology. As new fields develop (for example, biotechnology) and business complexity increases (such as strategic alliances), translators are being challenged by – and not necessarily keeping pace with – constant changes in business and technical vocabularies.

Thus, human translators are falling behind. Clearly, if MT can be adapted to global Internet time, it may be the tool that can scale human translation abilities to meet the burgeoning need.

Computer-aided translation is born

In our high-tech world, the obvious answer is to build technology that can accurately and quickly translate one language into another. Unfortunately, despite recent improvements, translation technology is not very accurate when used as a stand-alone solution because a literal, word-for-word translation does not automatically result in meaningful communication. For attempting perfect clarity, especially in the case of sensitive material, it is still necessary to use human translators, which can be slow and expensive.

Since human translation cannot keep pace with demand, a solution came with the recognition that not everything has to be translated perfectly. Sometimes, e-mail messages or online chat can make do with a rough translation. More to the point, it did not need to be an either/or choice. MT could be repurposed to aid a human translator. It was in this space, between using solely technology and only employing a person, that a new area developed: *Computer-aided translation (CAT).*



Trade-offs among translation options: Speed versus accuracy

Translation methods can be broken into three general types. Rather than discrete categories, the three are arrayed along a spectrum from translation completely by machine to translation entirely by a human translator. CAT combines varying amounts of machine and human translation. In general, the trade-off between speed versus accuracy is the primary differentiator among the three types (see Figure 5).

Figure 5. A comparison of the three types of translation.

Translation method	Trade-off: Speed vs. accuracy	Comments
Machine translation	Fastest, least accurate	Purely automatic word-for-word translation from source language to target language
Computer-aided translation	In between; speed and accuracy depend on the user and the technology	Stores and reuses a store of stock phrases and expressions previously vetted by human translators. As people add more words to the repository and learn when to use stock phrases, translation becomes quicker and more accurate
Human translation	Slowest, most accurate	Especially important when a document must appear to have been written in the target language by someone who is familiar with the idioms and semantics of that language and has the appropriate subject matter expertise

Source: IBM Institute for Business Value.

Defining a language translation strategy

With a clearer understanding of specific business needs and the translation options that could support them, organizations can begin to build language strategy solutions that are best suited to their own environments. Although the factors of accuracy, speed and cost may be weighed differently in different situations, companies can start to assess possible approaches by thinking through some key questions (see Figure 6).

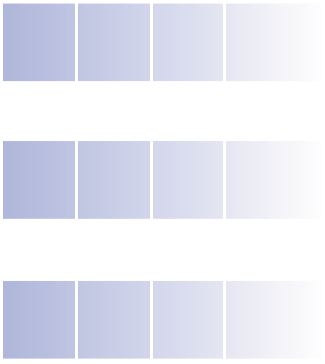


Figure 6. Key questions for evaluating language strategy solutions.

Accuracy requirements
• Who is the intended audience and do they share pertinent experience?
• How complex is the communication?
• What are the consequences of being misunderstood?
• How permanent is the communication expected to be?
Speed requirements
• How fast must translation be?
• What are the expected frequency and extent of updates?
• For internal databases, how centralized is the information?
Cost considerations
• When buying translation services, will there will be additional costs – such as formatting or testing hyperlinks – beyond the typical per-word translation costs?
• When buying licenses to run translation software packages internally, how much will it cost to buy licenses or the right to install the necessary number of copies of the software? What additional costs may be incurred by the need to purchase related services, such as integration of the software into existing systems?

Source: IBM Institute for Business Value.

Evaluating the context and risk associated with communications

Of accuracy, speed and cost, articulating the accuracy requirements for a translation task is often hardest because there are two intertwined dimensions of the communication to analyze: context and risk.

Context: Who is the intended audience? How complex is the communication?
The two primary context considerations focus on understanding both the audience and the communication's complexity. Fundamental audience questions include:

- Is the audience composed of internal people or individuals outside of the company? Some communication that is strictly for internal consumption may not have to be translated with 100 percent accuracy, while communication for an external audience may have to be more polished and thus require some "human touch."
- Does everyone have the same background, perform the same function or belong to the same team? People with pertinent shared experience often do not have to make all of their assumptions explicit since there may be an unstated, shared understanding of context. If everyone is acquainted with the subject and its jargon and assumptions, then a less precise translation may suffice.

There are several dimensions of complexity to consider. Something that may normally be straightforward can be complex because of a very granular level of detail. Other types of complexity are at the level of language, vocabulary or topic – for example, a more dense or arcane document may require more careful and precise translation.

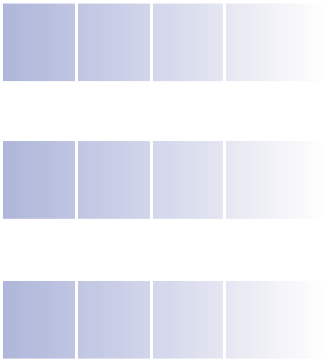
Risk: How grave are the consequences of being misunderstood? How permanent is the communication?
The two primary risk considerations include assessing the consequences of being misunderstood and taking into account the communication's permanence. It is crucial to identify the risk associated with being misunderstood. If there is no great potential fallout from being misinterpreted – as with an internal department meeting agenda that can be relatively painless (except, perhaps, for embarrassment) to correct in realtime – then accuracy can probably be traded for speed or cost. However, when the consequences of being misunderstood are serious, precision is required – such as a response to a libel suit.

Permanence dictates that higher risk is associated with a communication that will be published and cited long after it is written, such as an investment prospectus or research report. On the other hand, for something with a shorter shelf life – such as a routine meeting agenda – accuracy may not be as important.

Start now by making incremental changes

As collaboration and knowledge sharing needs continue to expand globally, organizations must face internal and external language barriers head on. There is no perfect solution to language translation challenges, but there are a number of ways to integrate existing technology with human translation to increase the flow of information across language barriers.

Since MT alone cannot yet suffice, it is essential to understand how organizations can support successful communication among speakers of different languages. Our study found that companies do not have to implement all-encompassing, costly solutions, but can undertake effective interim solutions. This awareness can help companies address their near-term translation needs by making incremental changes now that can also facilitate cost-effective solutions for the long term.



About the authors

Lisa Abrams is the Learning Lead for Communications Sector Learning and Knowledge in IBM. She can be reached at *labrams@us.ibm.com*.

Andrew Parker, formerly a Senior Consultant at the IBM Institute for Business Value, is a doctoral student at Stanford University.

Contributors

Eric Lesser is an Associate Partner and Team Leader at the IBM Institute for Business Value. Contact Eric at *elesser@us.ibm.com*.

Judith Quillard is Manager of Member Programs at the IBM Institute for Business Value. Send e-mail to Judith at *quillard@us.ibm.com*.

About IBM Business Consulting Services

With consultants and professional staff in more than 160 countries globally, IBM Business Consulting Services is the world's largest consulting services organization. IBM Business Consulting Services provides clients with business process and industry expertise, a deep understanding of technology solutions that address specific industry issues, and the ability to design, build and run those solutions in a way that delivers bottom-line business value.

References

1 Systran Information and Translation Technologies. Company press release. Accessed September 8, 2003. http://www.systransoft.com/About/News/pr000100.htm

2 Government of Canada. http://canada.gc.ca. Accessed on November 4, 2003.

3 Auditor General for Wales. http://www.agw.wales.gov.uk/index.htm. Accessed on November 4, 2003.

4 IBM Brasil. http://www.ibm.com/br. Accessed on November 4, 2003.

5 IBM United Kingdom. http://www.ibm.com/uk. Accessed on November 4, 2003.

6 Johnson & Johnson. Our Company Web Sites. http://www.jnj.com/our_company/company_websites/index.htm. Accessed on November 4, 2003.

7 European Association for Machine Translation. *MT News International* May 8, 1994, pp.15-18; *The New York Times*, January 8, 1954, page A1; *The Christian Science Monitor,* January 11, 1954.

8 Ibid.

9 Hutchins, John. *ALPAC: the (in)famous report*. Accessed September 8, 2003. http://ourworld.compuserve.com/homepages/WJHutchins/Alpac.htm

IBM Global Services
Route 100
Somers, NY 10589
U.S.A.

Produced in the United States of America
11-03

G510-3546-01