

## **AN EXAMPLE OF MT USE BY THE U.S. GOVERNMENT**

Joel Ross

Foreign Broadcast Information Service  
POB 2604, Washington, DC, USA 20013  
jrmt@ix.netcom.com

For over 50 years, the Foreign Broadcast Information Service (FBIS) has been the primary collector of foreign open source information for the U.S. Government. To bring the latest foreign political, military, economic, and technical information to analysis, warning, and operations processes, FBIS monitors over 3,000 publications, radio and television stations, news agencies, Internet sources, and databases in 210 countries and 73 languages. FBIS makes available to U.S. Government offices and to academic institutions many products and services: For example, FBIS databases of translations are accessible from worldwide electronic information handling systems that function via the Internet or similar technology. Also, FBIS linguists produce foreign language terminology glossaries, which the National Technical Information Service (NTIS) hosts on its free FedWorld service on the Internet (<http://www.fedworld.gov/fbis>).

As with many government and private-sector offices, FBIS has had to deal with shrinking financial resources that have led to a smaller workforce. Complicating matters further, there has been an enormous burst of new foreign information sources, particularly over the Internet, and consumers have increased their requests for FBIS products based on these sources. FBIS has been trying to resolve the dilemma of having fewer translators process a greater demand for translations, and machine translation (MT) has emerged as part of the solution.

One of the main problems in implementing MT in the Government workplace is identifying the appropriate circumstances in which it can be applied. While MT cannot substitute for the finely tuned translations of experienced linguists, it can reach a quality high enough to replace various summarization tasks. The following is an example of how MT was inserted into a work process to increase production and reallocate valuable resources (figures are rough estimates):

An FBIS French linguist had been receiving 100 requests a month from consumers for full translations of French-language newspaper articles on the Internet. The linguist and her contractors provided those translations, of which only 10 were eventually used by the consumers in end products. As resources dwindled, the linguist could only supply 75 of the translations, while requests increased to 125/month. A French MT system was installed, and the linguist responded to requests with machine translation output. The consumers found that the MT documents were good enough to clearly indicate whether “human” translations were necessary. This resulted in only 20 requests for full translations. While the consumers still incorporated only half of those requests into final products, the FBIS linguist had her translation workload reduced from 125 to 20. Eventually, she even handed over her duties of identifying French Internet sites to a non-linguist, who used French MT to “gist” a site, enabling him to decide whether the linguist should take a closer look. This enabled the linguist to concentrate on what she does best: translate. Production and morale increased; quality was maintained; and no one lost one's job to a machine.

While the above example illustrates how MT was successfully integrated into one FBIS work process, there are a number of obstacles impeding expansion. While MT developers have made

enormous steps in improving access to their products, users still do not generally find MT products as easy to use as they would like. This is partly a result of users' unrealistic expectations, but it also is a consequence of developers' lack of appreciation for what users need and want. The success story above works only because the linguist had the patience to learn the system and the perseverance to deal continuously with it. Unfortunately, many users do not share these traits. The process from French Internet webpage to English text document involves over a dozen steps over three different applications — a route most FBIS linguists will refuse to try. Simpler methods are possible, but the translation quality inevitably suffers.

FBIS deals in many foreign languages, but most MT systems are not able to handle different language code-page formats. If the MT program is not equipped with identification and conversion tools, the onus is placed on the user to prepare the document, but the user usually does not know a file's code page. This situation is unacceptable if the developer is trying to create a user-friendly product.

Many Government offices are understaffed with over-worked employees. The last thing many workers want is a new software gizmo that they must learn and apply while the “regular work” piles up — even if the promise is that their jobs will become easier by using the new product. Developers must be aware of this attitude and find ways to encourage evaluation. If developers want their products to be utilized by U.S. Government organizations such as FBIS, they must:

- Provide user-friendly, intuitive interfaces that need little more than the “click of a button” to translate a document. For many users, anything beyond this is frightening.
- Account for various language code pages and fonts, providing font files and automatic conversion programs if necessary. Users will choose not to use the MT program rather than having to personally convert formats beforehand.
- Ensure that the program is easily installed. An automatic CD startup will be more welcome than a package of 10 floppy disks.
- Keep in touch with Government contacts. It is necessary to be aware of user needs and upcoming projects.
- Be proactive, not reactive. Chances of successfully bidding on a Government contract is more of a gamble than selling to the Government a product that already exists.

Developers have their own financial concerns as well, and there is a limit to the powers of technology. One way to address these problems is to join forces. Two MT companies separately building parsing engines and dictionaries for a language pair would have a more successful result if one company built the engine and the other concentrated on the dictionaries. We are not calling for less competition; we are proposing more cooperation. The Government wants the best possible solutions for its MT needs, and a healthy MT industry is necessary to fulfill such needs. FBIS MT users have been frustrated by finding the best parsing engine in the most difficult-to-use program, while the most technologically advanced interfaces seem to have the smallest dictionaries. It is hoped that one day the best of each system could be combined to create a machine that might easily lose to Garry Kasparov in chess, but would rival him in translating a Russian sentence into English.