

Internet for the Translator

Navigating language resources on the Web

by Christophe Declercq

The Internet resources concerning translation are extremely scattered. Like every Net domain, there is no standard guide or authoritative categorization, let alone a premium organization. There are only two kinds of links—the good and the bad. However, the Internet has sharpened the communicative skills of language users on the whole, specifically those of the translator. With unlimited access to the immense bank of information on the Web, it is easy for translators to feel overwhelmed. Work- and language-addicted as the translator is, the complex Web can be navigated endlessly with virtually no orientation whatsoever. However, there are a few beacons.

The Future is Translation

Not only has the Net become a platform of information and communication, it has grown into a tool in its own right. In 1997 two thirds of all Web sites were American and 80 percent of all of the information on the Web was in English. Today about half of all new Web sites are multi- or plurilingual. Search engines support queries in different languages and are sometimes coupled with a simple but fast online automatic-translation tool. Growing worldwide trade and expanding international communications have increased the demand for translation. To master this growing workload, translators need software technology that speeds up translation, saves expenses, and improves quality.

Interpreting Web Offerings for Translators

The Net is a fast way to access information on any subject and to obtain firm product information. There is, however, the issue of the authoritative weight. Can the information on the Web be trusted?

Often, translators are not experts in the domain in which they are working. They have to trust upon their resources in order to find as accurate terminology as possible. Whether or not translators are authoritative specialists in a specific field (and the specialists not being terminologists themselves), the Internet offers relief. It is one of the inherent capacities of the Internet to improve the distribution, accuracy, and quality of terminology—therefore bettering the translation process itself.

Coordination in Navigation

With the large number of Web sites promoting translation tools, it should be no surprise that the tools vary widely in cost, quality, and accuracy. The number of sites with overviews on translation software is even larger.

Three resource lists are particularly useful: Rivendell, the University of Mainz, and the EU DGXIII site.

Rivendell (<http://rivendell.com/~ric/resources/mt.html>) discusses some 50 programs briefly. Some systems take on just one language in one direction only, while

others translate to and from multiple languages. Via the Rivendell Language Resources Chat (in cooperation with Uni-verse, Inc.) you can chat about languages. You are able to select the source language you will be chatting in and the target language. The chat text should be instantly translated. Note that Rivendell repeatedly states that automatic translation is not perfect. However, you should comprehend the context of your chat communication through this service. Accurate spelling and grammar help ensure good translations, local idioms and abbreviations can lead to bad translations. Rivendell suggests rewriting your sentence if the translation is not understood.

Rivendell also has a vast list of links to online dictionaries. The list ranges from monodirectional and bilingual to bidirectional and multilingual. The new type of learning via the global Web is also taken into consideration—Rivendell lists online courses, ranked by language. Two basic references referred to are The Virtual Call Library (www.sussex.ac.uk/langc/call.html) and Languages for Travelers (www.travlang.com/languages). Rivendell also has a large store of printed language resources and multilingual books and tapes, all alphabetically ranked from Albanian to Zulu. The section on machine-translation software might be the best. The sole criterion on the basis of which MT tools were selected was whether or not the attested system is ca-

pable of translating whole sentences. To name but a few: Uni-verse, Promt, Systran, Globalink Power Translator, Targumatic, and Deluxe.

The University of Mainz (www.fask.uni-mainz.de/ze/term/cat-pas.html) has an extensive list of translation tools as well. Products are briefly described, along with vital information concerning addresses of distributors and prices. The list is compiled in Germany and therefore focuses on products available there. The range of fields is broad: machine translation, automatic translation, computer-aided translation, terminology management, translation memory, and electronic references and dictionaries. Though the list is somewhat dated (last update stems from August 1996), it most certainly has its merits: it is very useful as a who is what and what is where, a valuable feature in the chaotic world of translation software.

CAT Tools: Computer-Aided Translation Software

CAT tools are systems used to lighten the translator's work, ensure higher terminology accuracy, and increase speed. They incorporate tools such as translation memory, terminology-management tools, alignment schemes, taggers, and the like. For a good overview of CAT, check these sites: www.languagepartners.com/catintr.htm and www.who.int/pll/cat/index.html.

IBM distributes the TranslationManager product family via Praetorius (www.praetorius.com/ibm/version2.html). IBM TranslationManager and IBM Personal TranslationManager are software products that provide a set of translation, text processing, and dictionary aids to meet the needs of translators and other bilingual personnel engaged in translation, localization, and publication-related activities. The TranslationManager products are computer-assisted translation systems that automate repetitive tasks, freeing the translator to attend to the finer points of translation that require the judgement of an expert. The use of the products helps accelerate material output and save money in overall translation costs. Using the controls and mastering the required keystrokes is easy. The matching function is quite efficient, with a good hit rate for fuzzy matches, or the degree in which to-be-translated material resembles text from the TM database.

On www.atril.com, you can find Déjà Vu Database Maintenance Align. Déjà Vu is a CAT package with different tools that enable the translator to work both faster and more accurately. It is a memory database which stores all translated texts, and

Often, translators are not experts in the domain in which they are working. They have to trust upon their resources in order to find as accurate terminology as possible.

Whether or not translators are authoritative specialists in a specific field (and the specialists not being terminologists themselves), the Internet offers relief. It is one of the inherent capacities of the Internet to improve the distribution, accuracy, and quality of terminology—therefore bettering the translation process itself.

contains a term tool called Termwatch. A fully integrated terminology system, Déjà Vu also comprises the File Alignment Wizard which aligns all the translated text with the purpose of feeding the text into the translation memory. Atril's Déjà Vu is based on Microsoft Access.

From the Trados Web site you can download test versions of some of their products and access free updates. Trados offers translation-memory tools, alignment tools, and terminology tools. Their translation memory product Workbench (www.trados.com/workbench/index.html) memorizes all your previously translated material. You have to align them first by means of the Winalign, a visual alignment tool that resembles Database Maintenance Align but which also has a stronger set of possibilities. The Trados TagEditor edits tagged file formats of specialized translation, thus enhancing the Workbench.

The most popular Trados product is Multiterm, which is built into several end systems. Multiterm Plus is the actual terminology-management software which stores an unlimited number of terms in up to 20 languages. In Multiterm, the user defines the fields he/she needs. There is a demo version that can be downloaded and used for term banks of up to 99 items per demo installation. With MultiTerm Web Interface, terminology can be accessed and

distributed via the Internet/intranet. Multiterm Dictionary enables you to publish terminological data electronically.

Online Translation of Text Material

Online automatic translation has not yet proved to be fully successful, but it has its merits as a tool for producing draft translations. The most renowned online-translation services are those offered by Systran (www.systransoft.com), for its alliance with Alta Vista, a popular search engine. Other translation services can be found at www.worldblaze.com and www.tranexp.com.

Translations can also be word-by-word. Word-by-word translation translates each word or phrase that it understands, but does not take grammar into account. Word-by-word translators are not as effective as machine translators, but they still can be useful as a translation aid. Babylon (www.babylon.com) offers a free downloadable dictionary in many languages (see also www.download.com, and then go education>language).

A fun MT program is the Tolken97 v3.2 (http://lrbcg.softseek.com/Education_and_Science/Languages/Review_14498_index.html), a fast and easy-to-use language translator, text editor, dictionary, and vocabulary test. It is downloadable as a whole and works for 30 days. The program has been awarded five Stars by ZDNet. Tolken97 is aimed at English-speaking users all over the world. You can have very fast translations into and from the following languages: English, Spanish, German, Danish, Norwegian, and Swedish (though not all pairs are available). The databases used for words and phrases can be edited, without any limitations other than a limit of using a maximum of 35,000 items per language in each database. This limit will be removed when a user license is purchased. Tolken97 is amusing in that it also performs some playful translations, such as turning words backwards, translating to and from "backslang," etc.

Fantasy

On the whole the quality of translation software depends on many factors such as the translation program, the type of translation, the grammar of the document to be translated, the use of a specialized dictionary or glossary, and so on. The quality of the engine of the translation program is usually the most important factor. Generally, draft-quality translations can be

(continued on page 40) ◀

expected: the result can be understood, but will need editing for professional use.

One can easily conclude that many translation problems and localization difficulties can be overcome by implementing an integrated system of translation-memory and terminology-management tools, supported by a background of machine translation. Translation tools, like computer software, may boost the market, but the final touch in translation services will remain human. There will always be the typical human characteristics that a computer will never possess: empathy, emotion, the need to go to war, imagination, etc. However, the exponentially complex power of new technologies can make us dream of greater possibilities.

Some Sites of Interest

A short reference should be made to organizations and products in language engineering: www2.echo.lu/mlis/en/direct/services/0288s.html.

Michelle de Raaij's site is always worth a visit (<http://people.a2000.nl/mderaaij/irt/index.html>); listings are extensive and user-friendly. The site is extremely useful for those in need of fast, accurate online Russian translations.

www.angelfire.com is a site where you can find an online version of the Prompt MT system. You can only type in a couple of

Aquarius.net is not only a valuable resource, but also one of the largest interactive databases of translators and interpreters on the Net. Via The Language Hub (www.cetrodfft.com/langhub/transbil.htm) you can obtain bilingual dictionaries and glossaries from an extensive database (from African native to Zulu).

sentences at a time, but then again there are no limits to using the service.

Another frequently visited site is the Human Languages Page (www.june29.com/hlp), where you can find useful information concerning languages and literature, schools and institutions, linguistic resources, products and services, and organizations—a true navigation resource.

Aquarius.net is not only a valuable resource, but also one of the largest interactive databases of translators and interpreters on the Net. Via The Language Hub (www.cetrodfft.com/langhub/transbil.htm) you can obtain bilingual dictionaries and glossaries from an extensive database (from African native to Zulu). The University of Amsterdam offers its language knowledge on www.pscw.uva.nl/sociosite/language.html. Last but not least: several online terminological databases and resources are available at www.infoterm.or.at/webis/webis.htm.

Christophe Declercq is a Dutch-English-Russian translator working at the KVH in Antwerp. As a research assistant, he is assigned to a project on knowledge-based systems, CAT, MT, TM, NLP, and corpus linguistics. Email him at cdapen@hotmail.com

PROMT MT System

