



The Technology Game

An industry insider's review of the latest and greatest in translation technology.

by Bert Esselink

Exactly a year ago, in December 1999, I wrote an article for *Language International* entitled "The End of Translation As We Know It." The piece described views on possible future developments in the language industry and translation tools world. It pictured how localization models were changing from projects to flows, i.e. from shrink-wrap software products to highly dynamic Web sites. It also described how translation-memory engines could be moving closer to the source to obtain a multilingual content database, not just bilingual memory. The article closed with some theories on how translators could be working in the future, linked to a particular publisher and receiving numerous small updates to translate on a daily basis and with improved turnaround.

In the year 2000, we have already seen many of these theories become reality, with software applications such as Idiom's World-Server and translation job portals such as Aquarius becoming increasingly popular. Also, many new buzzwords and much new jargon can be found in industry publications and forums, such as translation workflow automation, ASP models, translator portals, multilingual content management, XML localization portals, etc.

For all the people who have just come to grips with the concept of translation-memory systems, and for those who have lost track of what these buzzwords and new translation technologies are all about, the following article may shed some light.

Below I will summarize and categorize some of the latest developments in translation technology, without the goal of selecting "the best" technology in existence today. Each business model,

each file format, and even each personal preference determines which tool or technology is the best choice. Where for virtually all software products and services comparative information is readily available—for example the pros and cons of using Netscape versus using Internet Explorer—most players in the localization industry still need to figure out by themselves which technology is most suitable for which product or business situation. Up-to-date comparative analyses of translation technologies are virtually nonexistent.

Translation Aid

Software that helps translators do their work more efficiently and quickly has been in (professional) existence for approximately 10 years. Before the Internet also started impacting the translation and localization industry, translation technologies could be categorized in the following two groups:

- Computer Aided Translation tools
- Machine Translation tools

Computer Aided Translation (CAT) tools include technologies such as translation memory, terminology management, software UI localization, and glossary management. The core function of CAT tools was—and still is—to help translators do their work efficiently and maximize reuse of previously translated material. CAT tools are designed to work on fairly static, project-based material, such as a set of HTML or FrameMaker files. CAT tools

Each business model, each file format, and even each personal preference determines which tool or technology is the best choice. Where for virtually all software products and services comparative information is readily available—for example the pros and cons of using Netscape versus using Internet Explorer—most players in the localization industry still need to figure out by themselves which technology is most suitable for which product or business situation. Up-to-date comparative analyses of translation technologies are virtually nonexistent.

typically contain filters that convert a limited set of file formats to the format supported by the tool.

Leading suppliers of CAT translation-memory technology are Trados, Star, and SDL. Apart from the translation-memory desktop products for translators, most translation-memory providers are also offering more advanced solutions that include a combination of consulting, tools customization, and translation-workflow management. Star, for example, has developed a multilingual content management and workflow system that automatically imports and pretranslates files, then forwards the files to the translators and reviewers, and exports the translations for delivery to the client.

Another important type of CAT tool are software user-interface localization tools such as Corel Catalyst and Passolo, that are mainly intended as software UI resource editors with translation-memory functionality. Most of these tools offer limited support for programming environments, for example only Microsoft MFC. Although new languages such as Java are widely being used in Web-based applications, tools developers have been slow to catch up.

Machine Translation (MT) tools have the (often highly undoable) task of replacing human translations by producing translations that only need minor post-editing by a human translator prior to publication. The main reason for the existence of MT tools was to speed up the human, manual translation process. Where MT has been highly successful in the automotive and aerospace industries, after thorough preparation in the authoring process and terminology setup, the recent hype of MT systems has mainly been triggered by the machine translation “gisting” engines offered for free on the Web, such as the AltaVista Babelfish (www.babelfish.altavista.com) translation service.

In software and Web localization, there is serious interest in machine translation to bring translation costs down, but in most cases the content type and authoring process are not suitable for professional application of machine translation systems. Leading suppliers of MT tools are L&H, Systran, and Logos.

Translation Management

Translation management refers to technology that combines computer-aided translation with various online automation and reporting methods that are mostly Web-based. These types of technology are not only designed to help translators do their work more efficiently, but mainly to assist publishers and translation suppliers in managing the increasing complexity of translating large volumes of frequently updated information.

Translation management software typically contains the following core components:

Content Management—automating the process of managing content in multiple languages and detecting changes in the source language. Multilingual management applications such as Idiom’s WorldServer often integrate tightly with content-management systems such as Vignette or Broadvision.

Extraction—somehow translatable information must be extracted from a wide variety of file formats, in particular all variants of HTML and XML.

Workflow—the translatable files must be routed through a predefined process, e.g., hand-off to translator, translation, editing, approval, proofreading, etc. Tools that offer this functionality include SDL’s WebFlow and Lionbridge’s LionTrack, which has now been integrated with the Lionbridge Globalization Platform.

CAT—translation memory and terminology management. See previous section.

Reporting—online reporting of the number of words, the number of matches with existing translation memories, and the status of translation jobs. Many localization vendors have created client portals where, depending on access rights, customers can log in to check the status of their projects and answer to queries.

The company that has been given most publicity for automating translation workflows is Uniscape.com. Previously focused on software internationalization services and products, Uniscape re-invented themselves in 1999 to become an Application Service Provider (ASP) of translation workflow and management solutions. Put simply, this means that Uniscape is running the CAT tools mentioned above on their servers, and have automated the steps in a typical translation process.

These systems are all based on an ASP model, which means the core software required to manage these complex translation jobs is not sold and distributed, but hosted centrally and licensed on a subscription basis. Some of these systems still require a thin client download by the user, but the servers provide most of the processing power. True ASP models with no client-based software are no realistic option until all translators using the service have a fixed Internet connection.

Uniscape's success has impelled Trados, the leading translation-memory tools provider, to launch their Trados.net portal, which will be a hosted solution that "manages the enterprise document and translation process via a robust workflow engine" according to the press release. This development is very much in line with Microsoft.net, Microsoft's new ASP and application-hosting model. Another company that recently entered the ASP arena is Alpnnet, with their recently announced Globelix translation tools suite.

These systems are all based on an ASP model, which means the core software required to manage these complex translation jobs is not sold and distributed, but hosted centrally and licensed on a subscription basis. Some of these systems still require a thin client download by the user, but the servers provide most of the processing power. True ASP models with no client-based software are no realistic option until all translators using the service have a fixed Internet connection.

Some of these ASP companies have also integrated multilingual content-management solutions to their product range. For example, Uniscape recently announced a new tool called Global Content Manager, which automatically initiates the localization of new content added to a Web site.

Examples of suppliers that have focused on multilingual content management solutions are Idiom and GlobalSight. These companies do not provide ASP services, but are essentially software integrators. Their technology manages both the translation process and the multilingual content on Web sites. Idiom, for example, sells a product called WorldServer, which companies can install in their existing Web architecture to help them create and manage a multilingual Web presence.

Future Developments

Undoubtedly, traditional translation-memory tool providers will continue to move into the area of hosted services and applications. And current translation-management technology providers will continue to enhance the CAT functionality in their products. Eventually, publishers and translators will be able to pick from a limited set of all-encompassing translation-management and multilingual content-management solutions that directly plug into common Web architectures and publishing processes. These solutions will be provided both by localization service providers and tools developers. The question for publishers whether to pick

technology developed by a service provider or an (independent) tools developer will remain.

Choosing a technology developed by a service provider will in most cases mean that publishers also need to outsource translation and other localization activities to this provider. This approach has both pros and cons. The pros are that the service provider has a translator base experienced with the tools so quick project initiation is guaranteed, and service providers have had the chance to test the technology internally on projects. The cons are obviously that publishers cannot easily switch service providers when needed. Choosing an independent tool developer will allow publishers to use the translation service provider of their choice, although with the ongoing consolidation in the localization industry this could suddenly change as well.

For translators, the recent translation-technology developments are essentially good news. A likely scenario will be that, depending on the customer, downloading a free, thin translation client will be the only necessary technology investment. With ASP translation models, all pretranslation, (fuzzy) matching, and terminology processing is done by the server, not the client machine used by the translator.

In conclusion, the best translation technology solution for the future will be a combination of CAT, multilingual content management, workflow, and reporting features. All processing will take place on a server machine and translators will only be required to download a free translation client to do their work.

Bert Esselink is a regular contributor to Language International. The new edition of his highly regarded Practical Guide to Software Localization is available on amazon.com

For translators, the recent translation-technology developments are essentially good news. A likely scenario will be that, depending on the customer, downloading a free, thin translation client will be the only necessary technology investment. With ASP translation models, all pretranslation, (fuzzy) matching, and terminology processing is done by the server, not on the client machine used by the translator.