

Localizing the Bull way

By Andrew Joscelyne

French computer company Bull has developed an ambitious localization program to help it stay competitive in the international arena. Survival tactics?

One of the greatest stimuli for language industry applications during the 1980s was the mass adoptance of product localization within the computer industry. Translation companies sprung up to address this lucrative market, vying for contracts by offering one-stop package deals for multiple target languages, the fast turnaround required in the notoriously ephemeral software world, and ostensibly supplying the technical expertise and infrastructure which computer companies might reasonably expect. With the translation companies arrived various forms of computer-aided translation software to enhance the process. Within a few short years, localization has evolved from a casual afterthought to be a highly automated, vital marketing weapon.

An internationalization strategy

One company that learned the localization lesson the hard way was the French computer giant Groupe Bull, still mopping up the debts incurred by an insufficiently pro-active product internationalization strategy in the mid-1980s. Today, however, Bull can boast an ambitious and robust localization process, automated to the highest possible degree. It is the result of a key decision taken by Bull CEO Francis Lorentz in 1989. Realizing that his firm was dropping behind market-leaders in getting products on to world markets, he assigned top priority to the creation of an interdepartmental team to study ways of internationalizing Bull's major products. This would include designing standard facilities for localizing such parameters as date, time, currency, sort order, telecom standards, etc, as well as streamlining the translation of the user-interface and documentation into foreign languages.

The result was the Bull ILO group (*l'Internationalisation et Localisation de l'Offre Bull*) led by Olivier Surun, whose job was to oversee the implementation of four new basic principles in product design. First, every new product had to be internationalized to the degree specified in the overall marketing strategy. Second, English had to be the development language, for both products and documentation (an inheritance from Bull's Honeywell days as well as an implicit recognition of the preeminence of English in the computing sector). Third, "Corporate Strategic" products had to be systematically translated into French, German, Italian, Spanish and Dutch. Fourth, any other localization effort had to show a return on investment.

Self-perpetrating

As Olivier Surun says today, "this interdepartmental approach to solving a problem, stretching from the design stage to the marketing department, comprises a large number of facets and can in fact be a good recipe for failure. Luckily we have succeeded, with the end result that the project itself will dissolve and the system we have installed will take on an autonomous existence of its own." Since special emphasis has been placed on cost-effectiveness, the group has had to identify key stages in the process at which robust systems could be deployed with minimum danger. This inevitably meant drawing upon proven technology and well-tried process control strategies rather than doing any proprietary linguistic engineering on a grand scale.

For the localization of a software package, the task divides naturally into the user-interface (including help screens)

on the one hand and the accompanying user documentation on the other. For the user-interface, Bull filters the source material through swim (Software user Interface Manager), a package developed by Keyword, in Montpellier, France, to handle large batches of screens. To ensure consistency in corporate terminology, Bull has constructed a multilingual database of standardized Bull interface terms. The system is based on SPIRIT, a multilingual database interface manager from the French company Systex, to search the database for the correct target version of a given text string. In addition, all local parameters for dates, etc, are drawn from a parameters database. Before finalizing the product, Bull sends the screens and messages to the local distributor or subsidiary for validation.

There is always Systran

As for the user documentation, Bull has decided to utilize the services of what Surun calls “the least bad MT system on the market,” namely Systran. This system currently offers an English to multiple target language pair batch translation service that meets Bull’s corporate strategic product requirements, even though the German target module is judged below-standard at present. However, to optimize output quality, Bull pre-edit source texts by using a form of controlled English, as do a number of other large-scale Systran users.

Using a Windows version of the MAX text editor developed by US firm Smart Communications, a document pulled out of the documentation database can be batch-scanned for sentence length and complexity, then edited accordingly. This reduces ambiguity, thereby lowering translation costs in the long run. Since Bull also integrates a number of third party software packages from Informix, Microsoft, and others into its range of hardware platforms, the documentation accompanying these products is also passed through the MAX editor to standardize translation input, and hence output.

Interestingly, the MAX editor is also used interactively as a writing aid by Bull technical writers. The system signals style transgressions, with the inevitable result that, in due time, the writers learn to generate controlled text directly, speeding up the process. While MAX does not provide a wide-ranging critique of how technical information as a whole should be presented – it is strictly speaking a pre-editor for Systran – it is useful for determining the readiness of a text for MT. Once the MAX writing “error” rate is brought down to three percent, the text is considered acceptable for Systran..

Once the document in question is fully edited following MAX’s advice, it is sent to the Systran computer a few dozen kilometers away and returned translated minutes later. Lexical resources for the translation process are fed by Systran’s Customer Specific Dictionary facility, using Bull’s standardized terminology database mentioned above. The automatic translation stage is now reaching cruising speeds. The average quality rate of seventy percent is deemed adequate for ILO’s purposes.

Three million words a month

With the current volume running at one million words per month, Olivier Surun predicts an increase to over three million words per month by the end of the year. He calculates the average cost at ECU0.042 per word (7ECU/page) for the translation alone. Bull ILO reckons the total cost of localizing a page, including translation, restructuring the entire document, post-editing, at US\$25 (ECU20), about fifty- five percent of the price previously paid to translation companies. According to Surun, ILO has resulted in localization which is three times quicker and twice as cheap as the previous non-engineered approach: the ILO mission has, to a degree, been accomplished.

Now that the ILO process seems under control, Bull can begin to address medium-term needs in the localization process. Among them is the question of source language production. Although English is the language of development, about seventy percent of Bull’s developers are French-speaking. The group is therefore looking at ways of developing text production tools that are sensitive to French. As Olivier Surun says, “the process will only reach full profitability when we can substitute controlled French – or any other language – as input language.” French source text would also be better suited to the current version of Systran, whose French-to-English quality is higher than the converse.

As for the translation stage itself, Surun emphasizes that industrialization of the localization process required

qualified tooling. However, the group in no way feels “dependent” on a given MT platform, and potential interoperability has been an underlying design feature, so that ILO can adapt to improved MT systems and additional language pairs.

COPYRIGHT © 1992 BY LANGUAGE INDUSTRY MONITOR