EURESCOM Project BabelWeb

Multilingual Web Sites: Best Practice Guidelines and Architectures

Elisabeth den Os, Lou Boves

While the importance of the World Wide Web as a source of information and a vehicle for doing business is rapidly increasing, so is the need to make web services multilingual. Since all major telecommunication network operators aspire a role as Internet Service Provider (ISP) or Application Service Provider (ASP), it is clear that they invest resources in research and development in the field of multilingual web services. Because of the multilingual character of the research companies try and organize the work in an international framework. Such a framework is provided by EURESCOM, the European Institute for Research and Strategic Studies in Telecommunications. EURESCOM is devoted to collaborative R&D in telecommunications. It works as a virtual company using the resources of the shareholders (i.e., the former national Telecommunication companies) to perform high-impact research projects. The project BabelWeb was a collaboration of five companies, viz. BT Research, Portugal Telecom Inovação, CSELT (today TILab, the research company of Telecom Italia), France Télécom R&D, and KPN Research, that acted as project coordinator.

In today's market the Telecommunication companies need knowledge of and technology for the development of multilingual web services for several reasons. All companies seeing their traditional home market develop into a multilingual and multicultural environment because of massive migration. Thus, they need to offer their own services in multiple languages. In addition, all Telecommunication companies are trying to acquire a position as ISP, developing and hosting web services for other companies that are active in a global market.

The importance of multilingual web sites

In the past the WWW was virtually monolingual English, if only because the large majority of the customers who had Internet access were living in the USA and the UK. However, recent studies show that already now the majority of the web surfers are natives of other languages than English; in the year 2003 about 70% of the web surfers will be non-English.

At the same time it appears that even today's web surfers, who have above average education level and knowledge of foreign languages strongly prefer web sites in their native language. Research by Forrester shows that customers stay twice as long on a site in their native language, and perhaps even more importantly, buy three times as much.

Present Situation

In BabelWeb we have found that the large majority of the multilingual web sites that are in operation have never been planned to be multilingual. Almost invariably web services are designed with a single language and a single locale in mind. If the need arises, the service is localized for other cultures and languages. However, it also appeared that localization of existing monolingual web services is very difficult, and very expensive as well. In any case, the construction of a multilingual web service is a multidisciplinary enterprise. Fig. 1 shows the most important actors, technologies and their roles.

Maintenance of these sites is even more complex, time consuming, error prone, and expensive. This is especially important, because web-based information tends to have a short life cycle, which requires frequent updates. Fortunately, for some classes of multilingual web services the top-level pages, which must be of the highest quality and therefore require extensive expert involvement for their localization, are the most stable part of the service. However, even the home page may contain natural language information that changes daily, such as 'hot news'.

In some services lower level pages, which tend to be consulted only by prospective customers with a real interest in the service offered through the site, can tolerate somewhat lower quality translation and localization.

The aim of BabelWeb

Since many web sites may need to be localized in the future, it should pay to consider architectures and design principles that support localization. The aim of the BabelWeb project was to develop best practice guidelines for the design, implementation and maintenance of multilingual web services. To that end we analyzed a number of possible architectures for their suitability for multilingual services, and made an inventory of language technology tools that can support the implementation and design of multilingual web services. To put these somewhat theoretical results to the test, we decided to implement a real multilingual web service, and use that to conduct usability experiments with a number of realistic users. In the remainder of this paper we discuss the design and implementation of the multilingual demonstrator and the results of the user experiments.

The demonstrator service intended to support scientists working in EURESCOM projects to prepare and plan meetings in one of the cities where the partners' laboratories are located (Torino for TILab, Paris and Lannion for France Télécom R&D, Porto for Portugal Telecom, and the Hague for KPN). The web sites were meant to provide general information about the companies and their research labs, and about the countries. More specific information was provided about the cities, the closest airports and other travel opportunities, hotels, restaurants, etc. Also a link to a live weather information service was included. The information was made available in the native langue of the countries and in English, with two exceptions: the information on KPN, the Netherlands and the Hague was only available in English, because no suitable tools for automatic translation from Dutch to any of the other languages was available. The information about BT and England was also only available in English. This was considered acceptable, because the target population (scientists active in European projects) were supposed to be able to read and understand at least basic English.

Three Layer Architecture

To a large extent the requirements to the architecture of a multilingual web service run parallel to those of services that need to be accessed from terminals with different capabilities (such as powerful desktop workstation on the one hand, and small mobile terminals on the other). The single most important architectural feature that is needed in these circumstances is a systematic separation of the presentation of the pages on the screen, the way the navigation through the options is implemented (also indicated as the business tier), and the data or contents that the service accesses for its information. Thus, we ended up with a three layer architecture, as shown in Fig. 2.

From bottom to top, the data layer consists of a relational database, that is used to store the factual data, but also the text files pertaining to the countries and the companies. For a multilingual web service the database must be structured in two parts. The first part is locale independent. It defines the overall structure and organization of the information. The locale independent part of the database specifies the format in which the locale specific information in the second part of the database must be provided. It also is a key factor in the navigation on the next higher level. The second part of the database is by necessity locale dependent. It contains pointers to country specific data and properly formatted country specific information. The information for each individual country can be present in multiple languages.

Since navigation is based on the locale independent view of the underlying database, essentially the same 'look and feel' and the same paths through the menus are provided for all relevant languages. As can be seen in Fig. 2, information that is retrieved from the database can be processed in two different ways. It can be forwarded to the presentation layer directly, or it can first be processed by one of the language tools. Two types of language tools were implemented and tested for their usability, viz. Machine Translation and Automatic Summarization.

The business layer is also responsible for the generation of the structure and contents of the pages that will eventually be displayed on the user's screen. In a multilingual service one should use a technology that generates web pages on demand, such as Active Server Pages from Microsoft, or PHP3 in the Unix world.

Web services can be fully managed (closed) or partly managed (open). In open web services pointers to web sites maintained by other organizations can be included. In the EURESCOM meeting support service

we included links to a live weather forecast service, and to external web services with information about cultural activities in the countries and cities. In a multilingual service the information retrieved from external web sites may need to be processed (translated, formatted) before it can be shown to the user. The need to translate and perhaps localize information from external web sites should be anticipated in the design and implementation of multilingual web services.

Language selection is one of the most important issues that needs attention in the design of a multilingual web service. In the BabelWeb demonstrator it was decided to allow language switch at all times. Thanks to the systematic separation of presentation, navigation and data/contents it is possible to obtain the page with the exact same information immediately, without the need to return to the home page, and find the way back to the page that one wants to consult in another language.

The presentation layer in a multilingual web service also makes special requirements. In the BabelWeb demonstrator language selection buttons are present in all pages. We opted for a combination of country flags and a linguistic indication of the languages (cf. Fig. 3). This seems to be the best way to help the user quickly identify the correct switch.
Much effort was spent on the design of the graphical presentation, to avoid seemingly trivial, but difficult problems due to the fact that expressions of the exact same content in different languages are likely to contain different numbers of characters, and therefore to require different amounts of real estate. In a multilingual web service one must absolutely avoid embedding text in what is essentially graphics. Technology that supports layered pictures is available. With this technology it is relatively easy to overlay text in different languages over a picture.

Language Technologies

In the BabelWeb demonstrator we investigated three major types of language technology tools, viz. Translation Memories (TM), Machine Translation (MT), and Automatic Summarization (AS). After an extensive analysis of the tools that were on the market in mid 2000 we decided to choose TRADOS as the TM, Systran Enterprise for MT, and Extractor for AS.
One of the most important reasons to select TRADOS (in addition to its own merits) was the fact that KPN had extensive information with the software, and had already built a sizable memory of texts related to the telecommunication domain. Building our own translation memory would have been beyond the scope of the BabelWeb project.
Systran Enterprise was selected because it was available for the largest number of language pairs, and with the shortest time to delivery. Extractor was selected as the tool for AS because it was available for two languages (English and French) in a form that was royalty free for scientific experiments.

In BabelWeb we subjected the MT and AS tools to two different kinds of tests, both equally important: the ease with which the tools can be integrated in a multilingual web service, and the usefulness of the tools from the point of view of the end user of the service. TRADOS was not fully integrated, but the usefulness of the TM built by KPN was evaluated for the domain of the demonstrator.

Integration of Language Tools

We managed to integrate Systran Enterprise and Extractor in the BabelWeb demonstrator, be it not without some difficulties. We made the experience that the vendors could improve the effectivity of their support for users who intend to deply the software in somewhat unconventional ways. For Systran we had to develop a work around, in order to be able to integrate it in an Active Server Pages environment. Extractor may be easier to integrate, but it comes without support for a proper presentation of the output. This necessitates in house development of software to supply appropriate mark-up for the plain ASCII output of the tool.

Usefulness of MT

The usefulness of MT was tested in a scenario where scientists were asked to imagine that they were going to attend a meeting in Lannion, and that they wanted to stay an extra day for touristic reasons. To prepare their activity for the day they were presented with a text about fishing traditions in Brittany. The text fitted on one screen, and it was available in the French original and two different translations into English, one produced fully automatically by Systran, and one produced by a human translator. Subjects in the experiment were scientists working in the laboratories of the partners in BabelWeb, viz. 6 Dutch, 6 Italian, 6 English and 4 Portuguese, who had had at best some high school French.

All subjects followed the same procedure. They first had to read the English MT version of the text, and then answer nine factual questions about its contents. Next the were asked to read the human translation of the text, and subsequently score four Likert scales related to the usefulness, the trustworthiness, the comprehensibility and the ease of reading of the text.

The performance of the subjects on the factual questions was related to the scores of 6 native French subjects, who read the French original of the text. It appeared that the non-French subjects do worse on all questions, so we are left with the conclusion that some information was lost due to the translation process. Interestingly, the non-English subjects did better than the native English, who complained about the unusual vocabulary and syntax, not only in the MT version of the text, but also in the translation produced by the human translator.

As for the Likert scales, the majority of the subjects found MT a useful way of presenting the information. However, they also found the presentation difficult to read and to comprehend. The majority of the subjects expressed doubts regarding the trustworthiness of the information.

Usefulness of AS

The procedure used to assess the usefulness of Automatic Summarization was similar to the test of MT: 24 subjects were presented with a text on France Télécom R&D, that again was available in three versions, the French original and two English versions, one produced by human and one by machine translation.

The three versions of the text were processed by Extractor. It appeared that the output with the French original did not reflect the contents of the full text in any conceivable way. This made it impossible to test summarization of the original document followed by translation of the summary. This experience shows that the technology is not yet fully mature, at least not for French.

The summaries produced for the two English versions of the text differed substantially, although both did reflect the major gist of the original information. Of the 24 subjects 12 read the summary of the MT version, and another 12 the summary of the human translation. After reading the summary they were asked to answer two factual questions, and to rate the same four Likert scales as used to assess MT. The answers to the factual questions were related to the responses of 6 native French subjects, who read the original, full, version of the text in French.

The results of the experiment show that the accuracy of the answers to the factual questions on the basis of the automatically generated summaries is slightly worse than the performance of the French reference subjects. It also appeared that the accuracy with the MT version was slightly lower than the performance of the subjects who read the summary of the human translation.

The scores on the Likert scales were quite different between the summaries derived from human or machine translated texts. On average, the assessment of the summarized human translation was positive in all respects, while the corresponding scores for the MT based summary were slightly negative. We hypothesize that the different assessment is due to the fact that the sentences in the MT derived summary are more difficult to read. Moreover, the lack of discourse coherence, which is characteristic for all automatic summarization tools that extract key phrases from the original text, is probably more distractive in the MT than in the human translation version.

Usefulness of Translation Memories

The assessment of the usefulness of a TM in a multilingual web service had to follow different routes, simply because the target audience is not the end user population, but rather the translators and web site managers. In assessing the usefulness of a tool such as a TM two issues are at stake, viz. the integration of the tool in the daily work flow, and potential cost savings due to the tool.

The best way to integrate a TM in the workflow of the implementation and maintenance of a multilingual web service is not yet fully clear. Decisions must be made about a number of issues. Is it feasible to supply software that automatically detects changes in documents in a web service, so that the procedure to estimate the translation costs due to those changes can be started automatically? Who makes the decision that a change justifies the involvement of human translators? What is the position of the human translators? Are they employed by the company that owns or hosts the web service? Or are they employed by a service bureau, or are they perhaps self-employed? Will one accept fully automatic translation, if the changes in the texts are only minor, and adequate translations are present in the memory? Can one accept such an automatic translation for all pages in a service, or only for a subset of the pages at the lower levels? In BabelWeb we have only been able to make an inventory of these questions. Answers that are always true, irrespective of the type of web service and the type of company that owns the service are obviously impossible.

Assessing the expected cost savings that can be obtained through the use of a TM is not straightforward either. We conducted an informal experiment in KPN Research, where the translation department has extensive experience with TRADOS to support in house translations between Dutch and English. One of the results of this experience was a sizable TM for texts related to the telecommunication business. Intuitively, one would expect that a TM should yield lower translation costs because of the frequent, but small updates in the type web texts that one sees in a wide range of services.
For the experiment we selected two texts from each of three domains, viz. news messages that appear on the KPN web site, promotional texts taken from the web sites of competing telecommunication service providers, and pages containing information about theatre programs in the Hague.
Somewhat to our surprise it appeared that the TM coverage for all three domains was relatively low. This is despite the fact that the majority of the texts relate to the telecommunication business. The experiment was not big enough to allow far-reaching conclusions, but it seems wise to expect that TMs in the maintenance of multilingual web sites will yield most of the cost savings with texts that are subject to frequent minor updates.

Conclusion

In conclusion we can say that BabelWeb has produced best practice guidelines for the design, implementation and maintenance of multilingual web services. It has been shown that multilinguality must be taken into account from the very beginning of the design process. An architecture must be selected that supports systematic separation of the presentation, navigation and data/contents layers.
Language tools, such as Machine Translation, Translation memories and Automatic Summerization, can certainly be used to advantage, but at this moment the integration of these tools in a web service requires substantial expertise in the translation industry and in software development. End users find fully automatic MT of factual texts in languages that they do not speak useful, be it that reading and comprehending that information is considered rather difficult. However, native speakers of English were also critical about the quality of human translations from original french texts. Apparently, the translation quality provided by the best MT systems is approaching the quality that can be expected from the type of 'quick and dirty' human translation that is affordable in multilingual web services.
BabelWeb could only scratch the surface of multiingual web services. Much research remains to be done, especially with respect to the integration of language tools in the workflow.

Acknowledgement

BT. We also want to acknowledge the support of EURESCOM.

Fig.1 Design, production and maintenance of multilingual web services is a multidisciplinary enterprise

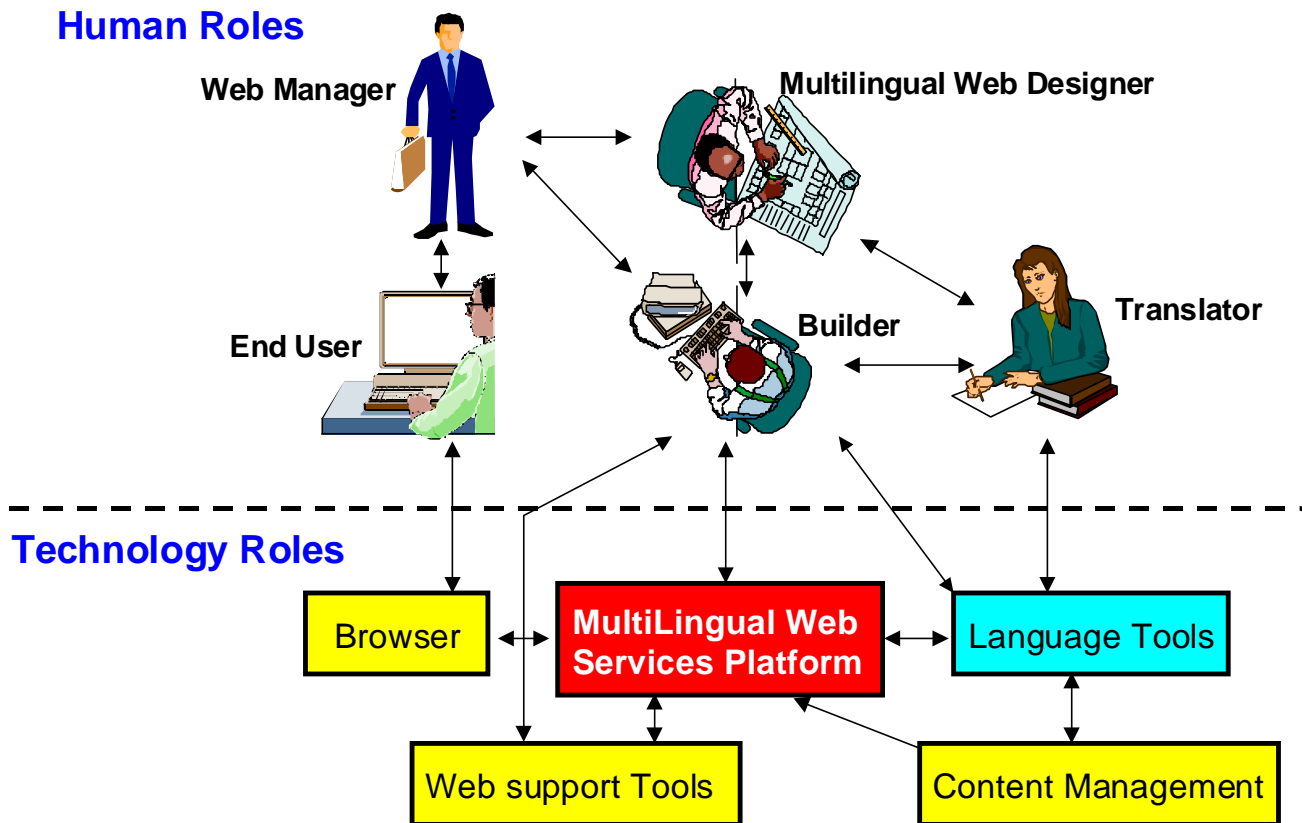# Actors involved in ML web site development

Human Roles

Web Manager

Multilingual Web Designer

End User

Builder

Translator

Technology Roles

Browser

MultiLingual Web Services Platform

Language Tools

Web support Tools

Content Management

Fig 2. Three layer architecture of a multilingual web service.

Presentation Tier

Service User Browser

Business Tier

Web Site ASP Based

Internet

MT Engine

Summarisation Tool

Data Tier

Data Repository

Operator's Data
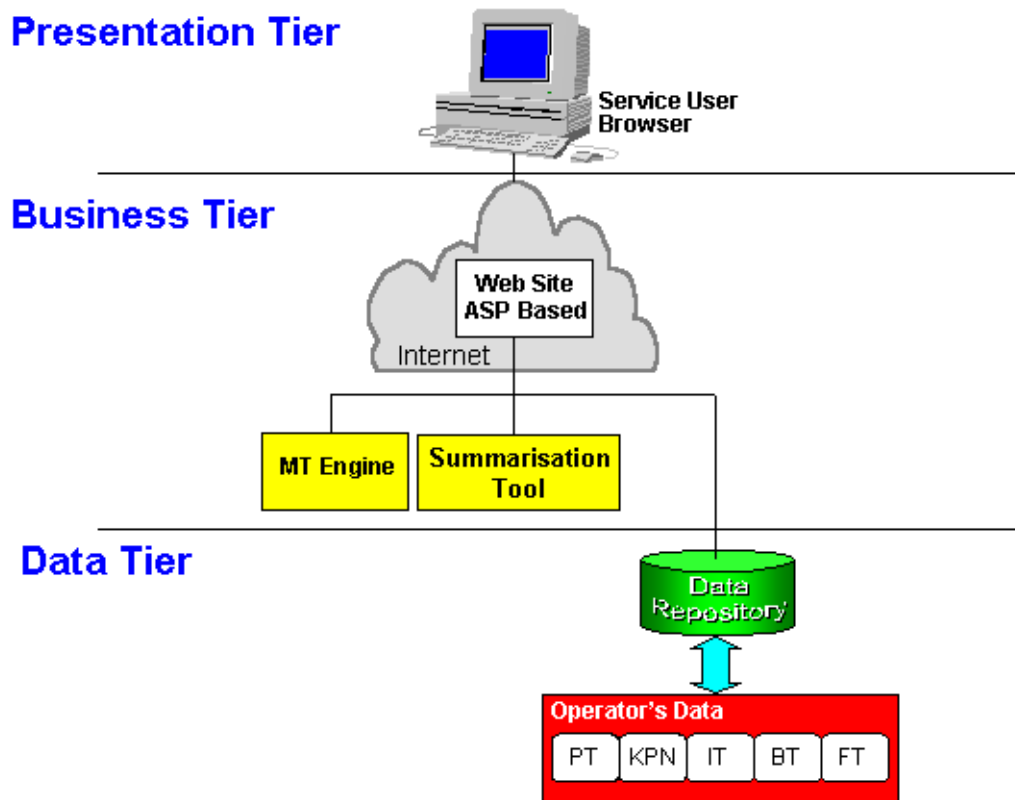
| PT | KPN | IT | BT | FT |

Fig. 3 Screen shot of the home page of the EURESCOM Meeting Support Service