

PROMT Opens the Door to MSDN Library for Russians

Case Study

The Case

The Russian branch of the Microsoft Corporation has decided to translate the information contained in the MSDN Library - the main resource for technical information on Microsoft software, - into Russian for Russian-speaking programmers who design software products based on Microsoft applications and technologies. Because of the complexity of topics and the great volume of data, the task required unique solutions. PROMT, the leader in the field of machine translation, has been invited to participate in a co-operative effort of investigation and achievement.

The Solution

Using their own, second-to-none, technology of online-translation, experts of PROMT could create the first translation service adapted to a specific information resource. The terminology analysis of one of the MSDN Library sections has been completed, structural features of the technical text have been taken into account and, for the first time, the technology of the associated memory is such that it is no longer necessary to translate anew each time the common terms and phrases of the technical literature. The originality of the translation method combined with unique solutions has made this joint project successful and has raised the bar of machine translation quality for all participants in the market.

MSDN Library is an extensive resource for technical information created by Microsoft Corporation to help experts, who are engaged in designing software products using Microsoft applications and development tools. Until recently, Microsoft provided developers with its technical documentation written primarily in English, although the number of foreign programmers working with MSDN Library is not less than the number of English-speaking users. The Russian branch of the company has decided to take care of the interests of Russian-speaking developers by making the information of the electronic library accessible to them in their native language. Experts of the PROMT, the world's leader in the field of machine translation technologies, have been involved in meeting this challenge. As a result, it was possible to create a unique translation service adapted to a specific information resource. While addressing this challenge, the experts of PROMT have found a new application for earlier created machine translation technologies, and also have worked out in great detail a number of unique solutions.

PROMT Ltd (Headquarters)
16 Birzhevaya Liniya
Vasilyevsky Island
Saint-Petersburg
199034 Russia

Tel.: +7-812-331 75 40
Fax: +7-812-327 44 83
E-Mail: info@e-prompt.com
Internet: www.e-prompt.com

PROMT (Moscow office)
6a Letnikovskaja Street
Moscow
115114 Russia

Tel.: +7-095-580 48 48
+7-095-509 35 94
E-Mail: moscow@prompt.ru
Internet: www.prompt.ru

PROMT GmbH
Eiffestr. 632
20537 Hamburg

Tel.: +49-40-219 01 140
Fax: +49-40-219 01 143
E-Mail: info@prompt.de
Internet: www.prompt.de

One of the reasons why Microsoft has been able to maintain absolute leadership in the market of computer software for many years is the availability of information on any of its technologies and solutions. This availability is not absolute. It is possible to give an opposite example in the notorious privacy surrounding OS Windows's codes, although this is most likely an exception to the rule. Another reason for Microsoft leadership is that it is interested in attracting the greatest degree of support in the development of software products that use Microsoft technologies and applications.

When working with software, questions sometimes arise for which answers cannot be found in "Help". Therefore, Microsoft experts have created an extensive electronic library of articles and materials that are devoted to the technical aspects of issues - MSDN (Microsoft Developer Network) Library - and have put it out on the Internet at www.msdn.com/library. Developers quickly evaluated its utility and "the electronic information bureau" has become very popular despite the fact that the information is available only in English.

It is necessary to note here, that the Microsoft Corporation has never previously undertaken any translation of materials contained in MSDN Library into other languages, not wishing to spend energy in constant updating of huge volumes of material and also believing that there is no programmer who is unfamiliar with the technical computer terminology, which is a priori in English, despite the fact that English is not the native language of the majority of professional developers of products that are based on the MS applications. In fact, the complicated and often changing technical information from MSDN Library is not clear to most Russian-speaking programmers.

If now, in our country, we would speak about deficiency, we would undoubtedly be referring to a shortage of competent paper technical dictionaries. It is well-known, that computer terms are also rare occurrences in the general dictionaries. As a result, Russian experts, who didn't possess a complete understanding of English, did not have full access to the information on the site <http://www.msdn.com/library>.

However, employees of the Microsoft Corporation's Russian office became interested in a problem. Due to the complexity of the subject and the great volume of data, the problem required unique solutions. Only one Russian company possessed the technology of online machine translation with which it was possible to achieve the desired results. For this reason, the Microsoft's Russian branch office has applied to the PROMT, the world leading machine translation technologies developer, with an offer to participate in the project.

It was decided to organize a translation service within the framework of project GotDotNet (www.gotdotnet.ru/), a site specially created by Microsoft Corporation for Russian programmers. Various help and teaching materials related to Microsoft products are available at this site, in addition to forums for discussion of hot questions concerning work with Microsoft products.

Though PROMT has provided translation systems and specialized dictionaries on computer topics for a long time, a decision was made to develop a special service for the English-Russian translation of the documentation in the MSDN Library. Nobody had undertaken such a task previously. Therefore, the work was performed in an atmosphere of inspired trail-blazing.

The first stage of the work consisted in the organization of a translation service on a basis of .NET service. The documentation on Microsoft product .NET Framework SDK has been chosen for translation from the majority of the library sections. Here it is worth noting, that it was impossible to complete the translation of all contents of MSDN library at once because of the enormous volume and heterogeneity of the information. Therefore, the section on the most modern subject Microsoft technologies .NET was chosen for the first stage.

When programmers of PROMT had built a framework of the future translation service, the task of tuning and optimization of service came to the fore. The professional tool for the analysis of texts and extraction of terminology, PROMT Terminology Manager, was therefore applied. With the help of this tool, the statistical analysis of terminology in the documentation from MSDN Library has been completed, the list of terminological candidates has been made and the dictionary of terms has been prepared. This terminological dictionary became the basic working tool of the translation service.

As is well known, the same word in the heading, in the text and in the list of program codes in technical articles can mean entirely different things. Hence, the method of translation depends on the structure of the HTML page. Technical optimization consisted of "teaching" the special program to recognize the structure of various page elements and to report this to the translation service.

During linguistic customization, PROMT experts used the Translation Memory (TM) database on the scale of the complete service for the first time. The Translation Memory (TM) database differs from usual dictionaries in that it contains sentences and even whole text fragments, named segments, rather than separate words and their translations. Thus, the TM base contains source text segments and their translation.

The TM base that has been created for the purpose of translating texts from MSDN Library contains more than 2000 segments - ready sentences and fragments of the text, used in the technical documentation. If an article requiring translation contains a fragment, that coincides 100 % with one that is stored in the Translation Memory, the service takes the text translation from the database, rather than translating it again. It provides guaranteed translation quality of standard phrases and expressions.

"The last part of service tuning was the most complicated" - recalls Alexander Serebryakov, the technical director of the PROMT company. "It can be said without exaggeration that the team, which worked at this stage, consisted of the best linguists in our country. We are really proud of the work done because, firstly, nobody has ever done anything similar before, and secondly, new opportunities for online-translation have been demonstrated."

The documentation translation service .NET Framework SDK received a great deal of positive responses from professional programmers, many of whom were surprised with the high quality of machine translation. With the successfully introduced technology as their foundation, PROMT experts also completed the translation of the Microsoft SQL Server documentation, intended for the creation and analysis of electronic databases. The high quality of the translation, as well as that produced for the .NET Framework SDK section, was the result of the adjusting the service according to the terminology and features of formatting that are peculiar to Microsoft SQL Server documents.

Dmitry Starostin, the representative of the Microsoft Corporation's Russian branch: "It is very satisfying to us that the idea of MSDN Library translation from the English language to the native users' language originated within the Russian Microsoft branch. Certainly, without the participation of PROMT, everything would still remain a dream. I want to emphasize that only innovative decisions could cause it to materialize, and the fact that the PROMT managed to develop them, shows the Russian experts to have great mental abilities and a high level of professionalism. Interest in the results of the experiment from other foreign Microsoft branches attests to its importance.

If we attentively address the needs of Microsoft products users in our countries, we can try to provide our experts with the same easy approach to information that their English-speaking colleagues have, develop loyalty to the company name and expand the growth of our products. As a matter of fact, our expectations have already begun to come true. We have experienced record growth in the number of visitors to the site www.gotdotnet.ru following the introduction of the PROMT translation service. The activity of users in the professional sector has increased. It is interesting to them, they exchange opinions about the new service at the forum, and most responses are positive, which is rare for an experimental undertaking".

At last we shall give a quotation from the forum: "Of course, there is rather crude text in the machine output for the moment, but the site and service are excellent. It is clear that developers of the machine translation technology are applying a great deal of energy and diligence to modification of the technology! It means that this site shows advantages to (a) technologies in general and (b) Microsoft technologies. Even the most inveterate skeptics, after having walked through the site, will come to the conclusion that machine translation is important."

The promotion of new technologies and solutions is a common feature of leading companies, such as Microsoft and PROMT. The MSDN Library translation project shows how a search for solutions to specific questions can stimulate the development of unique methods. Their successful application always brings additional results, in addition to attaining the original goal. Consequently, by working on a specific problem of translation of one information resource, the PROMT has raised the bar of machine translation quality as a whole.

About Microsoft

Founded in 1975, Microsoft (Nasdaq "MSFT") is the worldwide leader in software, services and solutions that help people and businesses realize their full potential.

Note to editors: If you are interested in viewing additional information on Microsoft, please visit the Microsoft Web page at

<http://www.microsoft.com/presspass> on Microsoft's corporate information pages. Web links, telephone numbers and titles were correct at time of publication, but may since have changed. For additional assistance, journalists and analysts may contact Microsoft's Rapid Response Team or other appropriate contacts listed at <http://www.microsoft.com/presspass/contactpr.asp>.

About PROMT

PROMT (www.e-promt.com) is a world leading provider of machine translation technologies. The company was established in 1991 and its expertise spans over 13 years of machine translation technologies development. Combining the innovation and technological perfection, PROMT has developed the revolutionary machine translation (MT) technology and became the technological leader in MT industry. PROMT core technology powers customized translation solutions that bring together in-depth linguistic knowledge and cutting-edge computer technologies.

Nowadays PROMT provides the full range of translation solutions: machine translation systems and services, dictionaries, translation memory systems, data mining systems for all kinds of business, professionals and Internet/Intranet. Our policy of continuous innovation in software and service delivery has achieved market share growth in all these sectors. New translation software product line @prompt consists of a wide range of sophisticated software products and solutions that leverage PROMT 13 plus years of enhancements. @prompt translation software is rated as the most sophisticated on the market and ensures the highest possible quality and efficiency of translation processes.

Our customers appreciate our commitment to the continued research and investment in the fields of linguistics and Natural Language Processing: we have customers spanning consumer, SME and enterprise markets, including global industry leaders Xerox, SAP, Lukoil, General Motors, Volvo Car International, Deloitte&Touch and others.

PROMT mission is to enable global business by building new generation MT systems that provide high-quality translation.

PROMT Ltd (Headquarters)
16 Birzhevaya Liniya
Vasilyevsky Island
Saint-Petersburg
199034 Russia

Tel.: +7-812-331 75 40
Fax: +7-812-327 44 83
E-Mail: info@e-promt.com
Internet: www.e-promt.com

PROMT (Moscow office)
6a Letnikovskaja Street
Moscow
115114 Russia

Tel.: +7-095-580 48 48
+7-095-509 35 94
E-Mail: moscow@promt.ru
Internet: www.promt.ru

PROMT GmbH
Eiffestr. 632
20537 Hamburg

Tel.: +49-40-219 01 140
Fax: +49-40-219 01 143
E-Mail: info@promt.de
Internet: www.promt.de