УДК 651.926

# Использование CAT-программ для перевода технических текстов

**Костенко Евгений Владимирович [1], Лапшина Светлана Николаевна [2]**

[1, 2] Уральский федеральный университет имени первого Президента России
Б. Н. Ельцина, Екатеринбург, Россия

[1] shakilov@mail.ru

[2] s.n.lapshina@urfu.ru

**Аннотация.** В статье рассмотрены программные системы для автоматизированного перевода (CAT), включая системы памяти переводов (TM), как ключевой элемент перевода текстов. Представлена сравнительная аналитика функциональных возможностей выбранных CAT-систем в процессе перевода. Определены лексические и грамматические проблемы, возникающие при переводе английских терминов в текстах нефтегазовой индустрии.

**Ключевые слова:** специализированные словари, английский язык, терминология нефтегазовой отрасли, CAT-системы, память переводов

# Using CAT programs to translate technical texts

**Evgeniy V. Kostenko[1], Svetlana N. Lapshina[2]**

[1, 2] Ural Federal University named after the First President of Russia B. N. Yeltsin,
Ekaterinburg, Russia

[1] shakilov@mail.ru

[2] s.n.lapshina@urfu.ru

**Abstract.** The article discusses software systems for automated translation (CAT), including translation memory systems (TM), as a key element of text translation. A comparative analysis of the functionality of selected CAT systems in the translation

process is presented. The lexical and grammatical problems arising in the translation of English terms in the texts of the oil and gas industry are determined.

Modern automated translation programs (CAT programs - Computer Aided/Assisted Translation) make it possible to simplify the translation process by providing the opportunity to make a translation quickly and efficiently.

The purpose is to simplify the translation process of the same type of documents. In this automated process a translator performs his functions to the fullest extent. The only thing the machine does is memorize translated text fragments and use them in subsequent translations. This technology is called Translation Memory (TM).

A modern TM tool is a complex computer system that includes many programs with different functions that can be divided into two groups – basic and service. The Translator's Workbench program automatically enters translated sentences into the database (translator's memory). Usually, the length of the record is determined by dots. Then the completed translation is automatically added to the original sentence in the database. After finishing the work, the translator receives the translation he has performed, and two segmented parallel texts are sent to the database. The next translated text will be processed in the same way and entered into the database. At the beginning of the translation of each new sentence, the TM tool automatically checks the database for the presence of the same or similar sentence among those translated earlier.

Genre-thematic specialization is becoming a general trend in the modern translation business. For example, after the translation of the first document of the oil and gas industry, no more than half of the new text will be found in the second document. If the first document was translated using a TM tool, then only the different fragments will have to be translated into the second one. The great advantage of TM tools is also that they fully preserve the format of the original text when translating. Thus, the translator is freed from the need to keep track of the size, paragraph, etc.

Most TM tools exist in two versions: for installation on a personal computer and for network use. The TM tool, installed in the local network of the enterprise, turns all translation activities into a collective one.

The main advantages of using CAT programs based on TM tools: 1) guarantee that the document is fully translated; 2) ensuring uniformity of terminology (abbreviations, names, etc.); 3) the ability to work with documents of different formats, for example, with Adobe Reader (.pdf), Microsoft Excel (.xls) files; 4) increased productivity when working on a project; reduced time and labor costs; 5) the ability to save the translation database for further use when working on new projects; 6) the ability to choose the software shell of a particular developer that most fully meets your requirements for functionality and usability; 7) the ability to automatically verify correctness (most CAT tools have embedded quality control modules).

Difficulties and limitations encountered when using TM technology: 1) sometimes it may be difficult for a translator to track the context of the translated text segment; 2) if TM technology has not been used before, the necessary training course; 3) the software do not support all possible forms; 4) the programs are usually incompatible with each other; 5) the high cost of programs, especially network versions; 6) all software add-ons are also paid; 7) if a text segment was translated incorrectly, and the translator or editor missed it, then when if you use this database in the future, all similar segments will also be translated incorrectly; 8) it is necessary to constantly monitor the quality of TM databases, update and change their contents very often manually – otherwise their value is significantly reduced; 9) low efficiency when working with texts in which there are few repetitive segments or where there are no such segments at all.

As part of the current study, an analysis of CAT systems such as Trados, SmartCAT and MemoQ was carried out.

The principle of operation of the Trados system is completely based on the TM concept, that is, it assumes the detection of fragments in the translated text, the translations of which are already in the program database. Untranslated fragments are transmitted to the translator for manual processing or a machine translation system is used

to translate them. The translator can select the translated fragments and add new pairs of parallel texts in two languages to the database, thus expanding it [5].

Another automatic translation system is SmartCAT, which is a cloud platform and cannot be installed on a computer. SmartCAT works using a browser, which greatly facilitates the process of using it, because it frees the translator from having to download the software installation in compliance with a number of technical requirements. SmartCAT can be attributed to automated translation systems, globalization management systems, and even freelancer exchanges [6].

SmartCAT supports various input formats, including text documents, presentations, spreadsheets, scanned documents and images (providing the latter with a paid service for recognizing text from photos and images), HTML pages, resource files, standardized bilingual formats and others. The system supports SDL Trados packages, which allows you to use SmartCAT to work on projects originally intended for execution in the Trados environment, as well as create source packages that can be uploaded to Trados.

Another AI tool of automated translation is MemoQ, which makes up an automated translation software package. MemoQ integrates with translation memory, terminology, machine translation, and reference information management in desktop, client/server, and web application environments.

It is important that documents can be imported between Trados and MemoQ systems.

The MemoQ system is almost as good as the functionality of Trados Studio, while it is easier to learn and more intuitive to use, which is more suitable for use by student translators.

When comparing programs, the following conclusions can be drawn:

1. All programs show high translation efficiency,

2. Trados Studio has the greatest functionality, but other programs can cover most of the translation needs.,

3. SmartCAT and MemoQ have more intuitive interfaces compared to Trados Studio.

There are some difficulties of translating texts in the oil and gas industry, it can be noted that there is no necessary equivalent of the term in the translation language, or the selection of such an equivalent itself may present some difficulty. For example, "classification of oil and gas reserves" where difficulties in translating terminology can be explained by a partial discrepancy between the concepts used in Russian and English. The international terminological unit "reserves" does not act as a complete analogue of this term in Russian which is closer to the term "stocks".

Another difficulty may be presented by complex terms consisting of several words that can be translated by a single word or a terminological phrase. So, the terminological combination of the CAT - system "well construction project" can be translated in different ways: "a well drilling project", for example. However, a "well construction project" would be more correctly translated as "well design". Translation is also complicated by the need to understand which word is determinative and which is definable. Such phrases usually include those in which the connection between words is formed grammatically, for example with the help of prepositions. The term "development", which is often translated in the context of field development: "plan of development", "develop reserves" in the geological aspect means "reservoir development": co-mingled development – joint development of reservoirs. However, in combination with well or drilling, this terminological unit begins to be translated differently: development drilling – production drilling, development well – production well. The lack of new terms in dictionaries is also a problem. For the correct translation of such terms, special dictionaries should be used with the possibility of considering the same term in different branches of science [3].

For the correct translation of terms by the CAT system, it is necessary to establish a connection between the term and the text, as well as to correctly understand the situation in which these terms are used. Scientists believe that before starting to translate special texts, the translator should obtain all the necessary knowledge about the field to which the source text belongs in order to be able to choose an appropriate equivalent or resort to translation transformations in cases where such an equivalent is not available in the translation language. In cases where there are no equivalents of the term, it is

recommended to resort to the following transformations: 1) descriptive translation – this method is the most convenient, but it can be too cumbersome: "evolution apparatus" – "прибор для выделения и измерения объема вещества" (a device for isolating and measuring the volume of matter); 2) translation using the genitive case and the preposition of: "dynamic holdup of column" – "динамическая задержка колонны); 3) transcription in cases where the term is in the original language and the translation is consonant with: fitting – "фитинг"; fretting – "фретинг", gas oil – "газойль"; 4) transliteration – involves recreating the graphic form using the letters of the translation language: regenerator – "регенератор"; barrel – "баррель"; 5) translation using prepositions: sampling and analyzing unit – "установка для отбора и анализа проб". When translating terms by the CAT system, there is a possibility of encountering the problem of ambiguity of a particular term. To prevent such problems, it is recommended to first comprehend the text, understand its essence, and then try to choose the meanings of the term that will fit the context. In some cases, the CAT system may not understand the multicomponence of the term. In order for the translation to be performed correctly, it is necessary to check the translation of the concept separately and together with the word that is used in conjunction with this term. Among other difficulties of translation by the CAT system of oil and gas texts, there is a discrepancy in the requirements for all possible abbreviations, as well as a difference in translation standards adopted in a particular country. The abbreviations themselves may also differ depending on the country, as a result of which they may be incomprehensible even to a specialist in the field to which the text refers [1]. In this case, it is very difficult to avoid semantic errors, distortion or inaccuracy of translation. For example, M&R (maintenance and repairs) – maintenance and repair in short in Russian looks like TO (technical service). The analysis of the scientific and professional array of the language of the oil and gas industry showed that the texts contain a large number of special terms. Due to the fact that some lexical units are not typical, for example, of the Russian language, their translation and correct interpretation present translation difficulties.

Conclusion. Therefore, it is quite obvious that using automation tools such as CAT programs, it is possible to increase the efficiency of the translation process of oil and gas texts, reduce costs and time for performing various types of work. The use of programs with support for storage memory technology has long been a standard in the oil and gas industry and the need to use CAT programs in their work should be perceived as a familiar thing. The prospect of further scientific research is the development of a glossary of oil and gas terms.

### Список источников

1. Коняева Л. А. О некоторых трудностях научно-технического перевода // Перевод и сопоставительная лингвистика. 2015. № 11. С. 50-54,

2. Слинкина Е. А., Краснова Е. В. Перевод англоязычных терминов нефтегазовой тематики // Молодой исследователь Дона. 2017. № 5 (8). С. 131-135.

3. Шурыгин Н. А. Филологические термины межсистемного функционирования // Филологические науки. 1998. № 5-6. С. 47-56.

4. CAT программа – что это такое и как начать делать переводы в 2 раза быстрее [Электронный ресурс]. URL: https://turbotehsnab.ru/raznoe-2/sistemy-cat-cat-sistemy-i-budushhee-perevoda-sistemnyj-blok.html (дата обращения: 12.04.2024),

5. SDL Trados [Электронный ресурс]. URL: https://www.sdltrados.com (дата обращения: 13.04.2024),

6. Smartcat [Электронный ресурс]. URL: https://ru.smartcat.com (дата обращения: 14.04.2024),

7. memoQ [Электронный ресурс]. URL: https://www.memoq.com (дата обращения: 14.04.2024).

### References

1. Konyaeva L.A. O nekotoryh trudnostyah nauchno-tehnicheskogo perevoda // Perevod I sopostavitelnaya lingvistika. 2015. № 11. P. 50-54,

2. Slinkina E.A., Krasnova E.V. Perevod angloyazychnyh terminov neftegazovoy tematiki // Molodoy issledovatel Dona. 2017. № 5 (8). P. 131-135.

3. Shurygin N.A. Philologicheskie terminy mezhsistemnogo funkcionirovaniya //

Philologicheskiye nauki. 1998. № 5-6. P. 47-56.

4.  CAT programma – chto eto takoe I kak nachat delat perevody v 2 raza bystreet [Electronic resource]. URL: https://turbotehsnab.ru/raznoe-2/sistemy-cat-cat-sistemy-i-budushhee-perevoda-sistemnyj-blok.html // (accessed: 12.04.2024),

5.  SDL Trados [Electronic resource]. URL: https://www.sdltrados.com (accessed: 13.04.2024),

6.  Smartcat [Electronic resource]. URL: https://ru.smartcat.com (accessed: 14.04.2024),

7.  memoQ [Electronic resource]. URL: https://www.memoq.com (accessed: 14.04.2024).

### Информация об авторах

**Костенко Евгений Владимирович** – студент Уральского гуманитарного института Уральского федерального университета (Екатеринбург, Россия), E-mail: shakilov@mail.ru, ORCID: https://orcid.org/0009-0003-2665-9882

**Лапшина Светлана Николаевна** – кандидат технических наук, доцент кафедры иностранных языков и перевода Уральского гуманитарного института УрФУ (Екатеринбург, Россия), E-mail: s.n.lapshina@urfu.ru, ORCID: https://orcid.org/0000-0001-6537-882X

### Information about the authors

**Kostenko Evgeniy V.** – student. Ural Federal University named after the First President of Russia B.N.Yeltsin, Ekaterinburg, Russia, E-mail: shakilov@mail.ru, ORCID: https://orcid.org/0009-0003-2665-9882

**Lapshina Svetlana N.** - PhD in Technical Science, Associate Professor, Departments of Foreign Languages and Translation, Ural Federal University named after the first President of Russia B.N. Yeltsin, Ekaterinburg, Russia, e-mail: s.n.lapshina@urfu.ru, ORCID: https://orcid.org/0000-0001-6537-882X

УДК 81.26:651.926:004.9