

LEARNING NEURAL MACHINE TRANSLATION TOOLS AS A COMPONENT OF TRAINING PROFESSIONAL TRANSLATORS

ВИВЧЕННЯ ЗАСОБІВ НЕЙРОННОГО МАШИННОГО ПЕРЕКЛАДУ ЯК СКЛАДОВА ПІДГОТОВКИ ПРОФЕСІЙНИХ ПЕРЕКЛАДАЧІВ

The article deals with solutions to problems of multilingual communication using the capabilities of neural machine translation. Some examples of work being done by companies and organizations using machine translation tools are given. The need for an active response of educational institutions to transformations that require new skills and qualifications is substantiated. Attention is drawn to the content of training future specialists in the field of foreign languages. It is noted that some educational institutions are beginning to train specialists in the field of correcting and editing machine translation. The necessity of harmonization of language learning and assessment is confirmed. The opportunities to improve the use of modern technologies in the digital era are considered. Furthermore, some peculiarities of the use of publicly available free online translation services are indicated. The advent and spread of artificial neural networks provided the emergence of neural machine translation (NMT). It is pointed out that the functioning of NMT-systems is accompanied by its constant training which occurs by processing huge collections of existing translations into different language pairs. In addition, online translation services reserve the right to dispose of the content at their own discretion. The information stored on the server of the online translator can be used as a source of data for third-party organizations involved, for instance, in marketing or espionage. This feature of online translation should be taken into consideration by both ordinary users and professional translators who work with free public services.

It is concluded that translation is an important aspect in the workflow of documentation processing in all the organizations. The use of machine translation provides opportunities to improve the speed and quality of services as well as achieving significant cost savings. Therefore, the professional training of future specialists in the use of information and communication technologies, in particular, the qualified use of machine translation, should already be envisaged in curricula today. Coordination of the requirements and expectations of students and employers is a necessary condition for the further development of the system of higher education.

Key words: *foreign language, NMT-systems, editing, correcting, educational curricula, multilingual communication, language services, linguistic technologies.*

У статті йдеться про вирішення проблем багатомовного спілкування з використанням можливостей нейронного машинного перекладу. Наведено деякі приклади роботи, яку виконують компанії та організації з використанням засобів машинного перекладу. Обґрунтовується необхідність активного реагування вищих навчальних закладів на вимоги ринку праці, якими передбачається потреба у нових навичках та кваліфікації. Звертається увага на зміст підготовки майбутніх спеціалістів у галузі іноземних мов. Вказується на появу такого нового напрямку підготовки фахівців з іноземних мов, як коректування та редагування машинного перекладу. Розглядаються можливості покращення використання сучасних технологій у цифрову епоху з урахуванням існуючих прикладів їх застосування на практиці. Вказується на неявні особливості використання загальнодоступних безкоштовних онлайн-сервісів перекладу. Констатується, що у деяких випадках служби онлайн-перекладу залишають за собою право розпоряджатися контентом на свій розсуд. Робиться акцент на принципах функціонування систем нейронного машинного перекладу (NMT). Вважається встановленим фактом, що застосування штучного інтелекту та машинного навчання призвели до експоненційного зростання якості перекладу. Використання технологій NMT на сьогодні вимагає, щоб машина для перекладу отримувала парні речення для порівняння для задіяних мов перекладу. Зазначається, що машинний переклад за певних умов здатний підвищити швидкість та якість послуг перекладачів, а також сприяє значній економії коштів. Привертється увага до того, що натепер професійна підготовка майбутніх спеціалістів в галузі інформаційно-комунікаційних технологій, зокрема, кваліфікованого використання машинного перекладу, повинна знайти належне місце у навчальних програмах. Висловлюється спостереження, що інструменти машинного перекладу можуть сприяти реалізації концепції відкритої освіти із забезпеченням вибору освітніх стратегій. Зроблено висновок, що узгодження вимог і очікувань студентів і роботодавців є необхідною умовою подальшого розвитку системи вищої освіти.

Ключові слова: *іноземна мова, NMT-системи, редагування, виправлення, навчальні програми, багатомовне спілкування, мовні послуги, лінгвістичні технології.*

УДК 378.016, 81.33
DOI <https://doi.org/10.32843/2663-6085/2022/47.36>

Morentsova A.V.,
Lecturer at the Technical English
Department № 1
National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”

Introduction. The development of modern society, its economic, political and cultural life is impossible without high-tech and information-rich communication.

The incessant information flow and communication revolution increases the intensity of communication process, promotes the formation and dynamic development of cultural, social, public, professional, and other spheres. Of course, communication

processes are not limited to language groups. The use of machine translation capabilities helps to solve multilingual communication problems.

Analysis of recent researches and publications. The problems related to translation by means of modern machine translation systems are of interest to ordinary users as well as scientists and professionals (N. Panasenkov, V. Kotenko, I. Kirichenko, N. Sigacheva, O. Kuzmin, E. Kartseva, G. Gurov,

N. Kochetkova, O. Revina, V. Zhovtyak, V. Samusevych etc.). The view that the results of machine translation can be used for familiarizing with the content, provided that the text is used as signal information and does not require careful editing, is shared by most experts. It can be stated that today the application of new successful developments in the field of artificial intelligence helps to quickly obtain translation conveying the meaning almost without errors and taking into account the style of language. The use of machine translation systems is widespread among individual users, as well as among companies, educational establishments and institutions. For instance, according to CSA Research, 66% of consumers in 29 countries use machine translation when shopping online [1]. As an example of the use of machine translation systems by organizations and enterprises, we can cite the Design Bureau «Pivdenne» (Dnipro) [2]. In order to ensure prompt translation in terms of expanding international cooperation in the bureau a new algorithm for the implementation of machine translation systems was developed and implemented, based in particular on the programs PROMT, Pragma, PROMT for Trados [2]. It is worth noting that the company uses corporate solutions for these products, as well as that the final editing of machine translation is carried out by translators with the involvement of highly qualified engineers [2]. The possibility of free use in the training of future professionals of all the basic features of Google services, in particular, platforms and applications for learning foreign languages, according to L. Cherchata, is an advantage in education [3].

Due to the rapid spread and growing popularity of the variety of information and communication services offered online, there is an increasing need for multilingual communication. Numerous online translation services are useful, which are always ready to quickly offer a fairly accurate translation. The availability and dissemination of public online translation services provide the opportunity to conveniently receive the necessary translation around the clock, but there are some features that should not be ignored, especially when it comes to translation that may contain confidential information. It is interesting to investigate how the problems of adequacy and confidentiality of translation are solved when using machine translation programs.

The aim of the article is to study some features of the experience of machine translation by companies and organizations, in particular, the risk of leakage of confidential data in the case of public online translation services, and substantiate the need to train machine translation specialists in foreign languages. A set of theoretical methods were used, in particular: analysis, generalization and systematization.

Results. The problem of leaking confidential data in the process of using public online translators has

arisen recently. To some extent, this is due to the emergence and proliferation of artificial neural networks on which neural machine translation (NMT) is based. The previously widely used Statistical Machine Translation (SMT) system used simple verbal statistical analysis, which was gradually modified to probabilistic analysis of set phrases and expressions. Modern neural network translation (NMT) is taught with bilingual corpora, i.e. collections of texts in a certain language in written or oral form in electronic version [2]. The difference is that in the process of learning NMT operates not by individual phrases, but the whole sentences. NMT is constantly learning by working on huge collections of existing translations into different language pairs. The first attempt to offer neural network translation to the general public was the online translation service Google Translate, which included neural translation of nine languages in 2016. By the way, from 2009 to 2019, Google ran the Translator Toolkit service. This service made it possible to receive machine translation in a web interface. Due to this, a large array of data on translation and manual corrections of translators for 10 years was collected. In the future, this data was used to train algorithms to improve the quality of translation. The online translator DeepL (Germany) started using neural networks in early 2020. Thus, it is stated that NMT is a fairly new technology, the intricacies of which are not yet sufficiently disclosed to ordinary users. Of great interest are the issues of machine translation of information materials and documents belonging to organizations and companies, which, moreover, may contain confidential information, and the need for training specialists in the sphere of information and communication technologies, in particular, highly qualified developers of machine translation services [4, 5].

The experience of such a global international organization as the United Nations, which aims to maintain and strengthen peace, international security and promote cooperation between all countries, is interesting to learn. The UN declares the principle: "Multilingualism in the United Nations system is synonymous with respect for all languages and equal treatment of all languages recognized as official languages in every organization" [3]. The importance of multilingualism and its absolute necessity for communication is emphasized as well as the need to develop and improve the "policy of attracting new translators and interpreters" and retain qualified language professionals [3].

Written translation of documents of the meeting bodies and normative documentation is performed by the language services of the specialized agencies of the UN system. The translation process is subject to revision in order to increase the economic feasibility of documentation processing efficiency and cost savings. To this end, the Department of General Assembly and Conference Management introduced (2014) the eLUNa electronic translation program [6].

The acronym eLUNa means "electronic languages of the United Nations". Developed in-house specifically for the translation of UN documents, this program is an online computer translation tool with typical translation platform functionality. The system provides instant access to previously translated documents, special terms databases and machine translation systems.

Extending the use of the eLUNa system the UN Secretariat made it possible to speed up the translation of information materials during the COVID-19 pandemic. In particular, this system is used in the International Maritime Organization (IMO), the World Meteorological Organization (WMO), the World Health Organization (WHO) and UNESCO [6]

The Assemblies of the World Intellectual Property Organization (WIPO) have approved of a new approach that provides a more diverse, digitally indexed and video-searchable session, supplemented by text transcripts in six official languages that are automatically compiled using artificial intelligence. The introduction of the latest advances in digital technology aims to help increase the speed and quality of language services while saving money. The organization's management believes that this approach will facilitate resource-intensive preparation of reports. The document also cites the views of language experts, who note that the proper level of quality cannot be achieved with the help of automated transcripts. The World Intellectual Property Organization has launched an experiment to implement neural machine translation outside the UN system in collaboration with the translation services of the Organization for Economic Cooperation and Development. According to the evaluation of this experimental project, «the accuracy of neural machine translation can reach an average of 38-40 percent» [6]. Thus, the successful use of machine translation tools should be expected to increase productivity by 25-50%. Under such conditions, translators would perform the work of automated translation editors when performing written translations. The use of the eLUNa system with direct access to terminology and other reference materials should reduce the period of search for reference information by translators. According to the results of the experiment, the UN Secretariat proposed to move translation services to a new scheme, which would look like self-editing of texts, instead of the previously used two-step process, which first translates and then edits.

The complete automation of search work provided by the eLUNa system has led to the elimination of positions of reference language services. The report notes that, unfortunately, the ICT services of many organizations are not involved in the development and use of computer translation tools. The inspector of the commission that prepared the report emphasizes the need to review the integration of linguistic technologies with

a view to their widespread use. The report states that the integration of pre-automated machine translation into the translation process has led to a high demand for relevant training and argues that in the near future organizations will seek more proofreaders and editors than translators. In view of this, UN organizations are trying to «attract and train potential linguists from the time when they are still students of linguistic faculties» [4]. In compliance with modern requirements the educational institutions with which the UN cooperates have introduced courses in addition to the standard curricula in machine translation, editing of texts translated by computers and the use of computer translation tools. Students first learn the profession of a translator, and then acquire the skills of a machine translation editor. The report of the commission states that «editing and proofreading of automated translation have already become the most popular training courses at the University of Geneva» [6].

As follows from the above, the corporate segment, i.e. large companies and government agencies are intensively implementing machine translation for their own needs, while saving resources and money. For companies that use public services and do not allocate funds for an automatic translation service that works on the local network, there is a risk that the information intended for translation will be widely available. Large corporations and organizations use corporate solutions located in the company's internal network, which ensure the secrecy of confidential information. However, the COVID-19 pandemic has forced them to work from home, so it is very likely that some workers will not hesitate to use public translation services. It should be noted, that free online translators collect and store translated information. This approach provides continuous training of the artificial neural network, and, as a consequence, the improvement of neural network translation. But at the same time, a huge amount of information is created that can be used for advertising or marketing purposes.

Conclusions. It is obvious that neural automatic translation systems give a better result than the previous translation based on a statistical approach. Modern possibilities of neural machine translation will be increasingly used to complement and expand the work of professional translators [5]. Due to the fact that neural network technologies provide an opportunity to quickly get a translation that is almost ready for publication, there is a risk that due to carelessness or negligence of the user (translator), or lack of corporate translation service the leak of confidential information may occur. Today, many companies and government agencies use corporate software products for multilingual translation, hosted on the corporate intranet. This approach is justified. After all, in addition to information security, the creation of terminological databases of the company facilitates the convenient work with forms and documentation.

However, it should be noted that although machine translation technologies are constantly evolving and becoming more sophisticated, the responsibility for the end result of translation rests with the individual. Hence, the necessity of cooperation and partnership between educational institutions and companies, taking into account their requirements and expectations.

REFERENCES:

1. DePalma D. A., O'Mara P. D. Can't Read, Won't Buy – B2C. *CSA Research*: веб-сайт. URL: <https://insights.csa-research.com/reportaction/305013126/Marketing> (дата звернення: 20.05.2022).
2. Котенко В. В. Перспективы развития нейронного машинного перевода в контексте концепции открытого образования. *Ученые записки университета им. П.Ф. Лесгафта*. 2020. № 4 (182). С. 225-230.
3. Черчата Л. Використання сервісів Google у навчанні іноземної мови студентів неспеціальних факультетів. *Молодь і ринок*. 2021. № 4 (190). С. 51-56. DOI: 10.24919/2308-4634.2021.236340
4. Morentsova A. V. Enhancement of efficiency of teaching reading scientific and technical texts at higher technical educational institutions. *Актуальные научные исследования в современном мире*. 2020. № 5-7 (61). С. 36-40.
5. Моренцова А., Варянюк Т. Робота з машинним перекладом як складова змісту підготовки фахівців з іноземних мов. *Молодь і ринок*. Дрогобич. 2022. № 1 (199). С. 166-170. DOI: 10.24919/2308-4634.2022.252192
6. Lozinskiy N. (Ed.). Multilingualism in the United Nations system. JIU/REP/2020/6. *The Joint Inspection Unit*: веб-сайт. URL: https://www.unjui.org/sites/www.unjui.org/files/jiu_rep_2020_6_english.pdf (дата звернення: 20.05.2022).