DOI: 10.15290/lingdid.2025.29.18

Anna Włodarczyk-Czubak
Uniwersytet Radomski im. Kazimierza Pułaskiego
e-mail: a.wlodarczyk-czubak@uthrad.pl
ORCID: 0000-0003-0521-6559

HOW HAS MACHINE TRANSLATION CHANGED THE WORK OF A PROFESSIONAL TRANSLATOR?

ABSTRACT

The paper deals with the analysis of the problem of machine translation replacing translation by a professional translator. The analysed text was taken from texts translated for 2nd Logistic Supply Unit in Wrocław. The aim of the analysis is to check how machine translations (3 machine translation systems RBMT, SMT and NMT) are different from a translation by a professional translator (called Poprawne) by comparing the outputs of each of the systems and a human translator, as well as to decide if they still require proofreading or if they can be seen as a final product of the translation process. There are 2 research methods used: quantitative and comparative analysis of the 4 texts. The research results revealed that there are numerous errors and mistakes in the analysed machine translation texts and cannot be seen as a final product of the translation process.

Keywords: machine translation, AI, translation engines, comparative analysis

ABSTRAKT

JAK TŁUMACZENIE MASZYNOWE ZMIENIŁO PRACĘ PROFESJONALNYCH TŁUMACZY?

Niniejsza praca dotyczy analizy problemu zastępowania profesjonalnego tłumacza tłumaczeniem maszynowym. Analizowany tekst pochodzi z dokumentów zleconych do tłumaczenia dla II Wojskowego Oddziału Gospodarczego we Wrocławiu. Celem analizy jest sprawdzenie w jaki sposób tłumaczenie maszynowe (3 systemy RBMT, SMT oraz NMT) różni się od tłumaczenia wykonanego przez profesjonalistę (nazwane Poprawne) poprzez porównanie tłumaczeń wykonanych przez ww. systemy oraz przez tłumacza, jak również aby zdecydować czy wciąż potrzebna jest korekta tłumaczenia maszynowego, czy też

może ono stanowić produkt finalny procesu tłumaczeniowego. Zostały wykorzystane 2 metody badawcze: analiza ilościowa i jakościowa zamieszczonych 4 tekstów. Wyniki badania pokazały, że analizowane tłumaczenia maszynowe zawierają błędy i nie mogą być uznane za produkt finalny tłumaczenia.

Słowa kluczowe: tłumaczenie maszynowe, AI, silniki tłumaczeniowe, analiza porównawcza

1. Introduction

As it was said, "Any person who knows more than one language has the ability to explain a word or a sentence in what translators call 'the source language' (the language you translate from) by using an equivalent word or sentence in what they call 'the target language' (the language you translate into). This, in effect, is the beginning of translation"[1]. If we follow this quotation, it means that everybody can translate a text, but in reality it is a complicated process requiring specialised knowledge and abilities. That is why more advanced tools like machine systems have started to be introduced in the work of translators.

According to Brown et al., "[o]ur approach to machine translation is based on the idea that every sentence in one language is a possible translation of every sentence in another language, with some translations more probable than others"[2]. We can argue with this statement but if we look closer, it appears to be true. Machine translation has been developing since the 1950s with various results, for many years not having been precise for many. But with the appearance of neural machine translation it has been surprising how the development of machine translation has been accelerated. There is the necessity to check if machine translation can be used as a final product and can replace the professional translator, or if machine-translated texts still need proofreading before publication, by focusing on 3 main machine translations, i.e. RBMT (Rule Based Machine Translation), SMT (Statistical Machine Translation) and NMT (Neural Machine Translation). In this paper there is a presentation of how they work, giving the general characteristics, and then a comparison of all 3 machine translations and a translation by a professional translator. Finally, the quantitative and comparative (with the use of

[1] C. Picken (ed.), *The Translator`s handbook,* London 1989, p. 33.

[2] P.F. Brown, J. Cocke, S.A. Della Pietra, V.J. Della Pietra, F. Jelinek, J.D. Lafferty, R.L. Mercer, P.S. Roossin, *A Statistical Approach to Machine Translation*, "Computational Linguistics" 1990, 16 (2), pp. 79–85.

2 criteria) analyses of 4 texts are made to decide which translation is the most accurate, and if machine translations can be seen as the final product of the translation process.

2. RBMT, SMT and NMT comparison

As we all know, translation is a process with the aim and purpose of leading to a target text. According to Hans J. Vermeer, the aim and the mode of the action is determined by a commission. Vermeer presents and describes skopos theory in his article "Skopos and Commission on Translational Action":

> The *skopos*, which is (or should be) defined in the commission, expands the possibilities of translation strategies, and releases the translator from the corset of an enforced – and hence often meaningless – literalness; and it incorporates and enlarges the accountability of the translator, in that his translation must function in such a way that the given goal is attained. This accountability in fact lies at the very heart of the theory: what we are talking about is no less than the ethos of the translator[3].

Vermeer postulates that every translation has its aim, and the basis of each process of translation is interpretation. What is more, the translation has to be functional and culturally-oriented. "It reflects a paradigm shift from the linguistic to the functionalist approach, which defines translation as a purposeful transcultural action"[4]. But how can machine translation be defined? Simply as the translation process from the source to the target language without the use of a person. According to Arnold et al., "Machine Translation started out with the hope and expectation that most of the work of translation could be handled by a system which contained all the information we find in a standard paper bilingual dictionary"[5]. It can be said that it all started in the 1950s with the appearance of RBMT. RBMT systems are built on linguistic rules created by human experts. These rules cover the grammar, syntax, and semantics of both the source and target languages. RBMT relies heavily on extensive bilingual dictionaries which contain not just word-to-word translations, but also information about parts of speech, verb conjugations, and other linguistic details. The source text is analysed linguistically

---

[3] H.J. Vermeer, *Skopos and Commission in Translational Action. The Translation Studies Reader*, L. Venuti, M. Baker (eds.), London–New York 2001, p. 237.

[4] Ibidem.

[5] D. Arnold, L. Balkan, S. Meijer, R.L. Humphreys, L. Sadler, *Machine Translation: An Introductory Guide*, Manchester 1994, p. 10.

and the analysed structure is transferred to the target language. RBMT systems require deep, language-specific knowledge to be programmed into them. They can struggle with ambiguous language, idiomatic expressions, and context-dependent translations. This can make RBMT systems less scalable compared to data-driven approaches. Many modern systems combine RBMT with statistical or neural approaches to leverage the strengths of each method. RBMT systems can be fine-tuned for specific domains by adding specialised vocabulary and rules. This makes them particularly useful in fields like legal or medical translation where precision is crucial.

In the 1990s there appeared SMT, or Statistical Machine Translation, which is a type of machine translation that relies on statistical models to translate a text from one language to another. SMT tools use large bilingual text corpora to learn translation patterns. They rely on statistical models to predict the most likely translation for a given input sentence. These tools use probabilistic models to determine the best translation. They calculate the probability of a translation being correct based on the frequency of word and phrase alignments in the training data. SMT often uses phrase-based translation, where the text is divided into phrases rather than individual words. SMT systems require extensive training on large datasets to improve accuracy. SMT tools can struggle with idiomatic expressions, rare words, and languages with complex morphology.

It can be said that the Neural Machine Translation (NMT) tool, which represents the current state-of-the-art in machine translation technology, appeared in the 2010s. However the first works on NMT started in the1980s-1990s. The biggest development of NMT can be observed after Microsoft released AI-powered online NMT, and then GPT-3 by Open AI. NMT uses artificial neural networks, specifically deep learning techniques, to learn and perform translations. It typically employs sequence-to-sequence models with attention mechanisms or transformer architectures. According to D. Bahdanau et al., "[u]nlike traditional statistical machine translation, neural machine translation aims at building a single neural network that can be jointly tuned to maximize the translation performance"[6]. The NMT system learns to map between languages during the training process and requires vast amounts of parallel text data for training. Advanced NMT models use techniques like sub-word tokenisation (e.g., Byte Pair Encoding) to handle rare or unknown words effectively. Transfer learning techniques allow leveraging pre-trained models for new language pairs or domains. Training and running NMT

[6] D. Bahdanau, K. Cho, Y. Bengio, *Neural Machine Translation by Jointly Learning to Align and Translate* (paper presented at 3rd International Conference on Learning Representations), ICLR 2015, San Diego 2015, p. 1.

models often require significant computational resources, including GPUs or TPUs. As neural network architectures and training techniques evolve, NMT systems continue to improve in quality and efficiency. Innovations like transformer models have led to significant advancements in recent years. While early NMT models struggled with long sentences, modern architectures like transformers have largely overcome this limitation, and "numerous efforts have since continued to push the boundaries of recurrent language models and encoder-decoder architectures and to rely entirely on an attention mechanism to draw global dependencies between input and output"[7]. NMT may occasionally *hallucinate,* which means producing outputs that are unrelated to the source text.

This comparison highlights the evolution of machine translation technologies, with NMT currently leading in terms of translation quality and adaptability, while RBMT and SMT have their own niche strengths and, as Arnold et al. asserted, 'Scientifically, MT is interesting, because it is an obvious application and testing ground for many ideas in Computer Science, Artificial Intelligence, and Linguistics, and some of the most important developments in these fields have begun in MT'[8].

3. Quantitative and comparative analysis of translations made by RBMT, SMT and NMT machine tools and by a professional translator.

The following is a comparison between translations made by a professional translator called *Poprawne* and by machine translation services: Systran (RBMT), Google (SMT), and DeepL (NMT) to indicate differences between translations, and how the systems translate the original text.

The original text in Polish:

> PODPISANIE POROZUMIENIA O WSPÓŁPRACY Z NARODOWYM INSTYTUTEM DZIEDZICTWA. 9 LUTEGO 2024 R. DYREKTOR MIĘDZYNARODOWEGO OŚRODKA SZKOLENIA I BADAŃ NAD DZIEDZICTWEM KULTURY W ZAGROŻENIU PŁK PAWEŁ NAWARYCZ PODPISAŁ Z DYREKTOR NARODOWEGO INSTYTUTU DZIEDZICTWA DR HAB. KATARZYNĄ ZALASIŃSKĄ POROZUMIENIE O WSPÓŁPRACY W OBSZARZE OCHRONY DÓBR KULTURY.
>
> Międzynarodowy Ośrodek Szkolenia i Badań nad Dziedzictwem Kultury w Zagrożeniu oraz Narodowy Instytut Dziedzictwa, realizując swoje cele i zadania statutowe, podpisali porozumienie o współpracy w obszarze ochrony dóbr kultury. Przedmiotem

[7] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, Ł. Kaiser, I. Polosukhin, *Attention is All You Need. Advances in Neural Information Processing Systems*, New York 2017, p. 6001.

[8] Ibidem, p. 5.

współpracy będzie w szczególności realizacja wspólnych działań edukacyjnych i szkoleniowych oraz wymiana informacji i doświadczeń w zakresie ochrony, poszanowania i opieki nad dziedzictwem materialnym na wypadek konfliktu zbrojnego i sytuacji kryzysowych, a także podejmowanie wspólnych działań na rzecz budowy, kształtowania i umacniania tożsamości narodowej poprzez prezentowanie oraz popularyzację materialnego i niematerialnego wojskowego dziedzictwa kulturowego[9].

The choice of the original text was generally obvious, because it is said that machine translation systems have difficulties in dealing with proper names and specialised vocabulary, and the text in Polish presents the required complexity. The given text is in Polish, and discusses the signing of a cooperation agreement between the International Centre for Training and Research on Cultural Heritage in Danger and the National Heritage Institute. The main topics covered are: the directors of the two institutions, Colonel Paweł Nawarycz and dr. hab. Katarzyna Zalasińska, signed a cooperation agreement in the field of cultural heritage protection. The cooperation will focus on joint educational and training activities, as well as the exchange of information and experience in the protection, respect, and care of tangible heritage in the event of armed conflict and crisis situations. Specific terminology and phrases that might be challenging for translation include: Międzynarodowy Ośrodek Szkolenia i Badań nad Dziedzictwem Kultury w Zagrożeniu (International Centre for Training and Research on Cultural Heritage in Danger), Narodowy Instytut Dziedzictwa (National Heritage Institute), dobra kultury (cultural property), dziedzictwo materialne (tangible heritage), dziedzictwo niematerialne (intangible heritage), and tożsamość narodowa (national identity). The text is written in a formal, legal-administrative style, with the use of specific terminology related to cultural heritage and institutional cooperation.

There are 4 translations made by a professional translator and 3 by machine systems, which will be analysed in the next step.

Poprawne translation (by a professional translator):

SIGNING OF A COOPERATION AGREEMENT WITH THE NATIONAL INSTITUTE OF CULTURAL HERITAGE OF POLAND (POLISH: NARODOWY INSTYTUT DZIEDZICTWA NID). ON 9 FEBRUARY 2024 THE DIRECTOR OF THE INTERNATIONAL CENTRE FOR TRAINING AND RESEARCH ON CULTURAL HERITAGE IN DANGER COL. PAWEŁ NAWARYCZ SIGNED AN AGREEMENT WITH THE DIRECTOR OF THE NATIONAL INSTITUTE OF CULTURAL HERITAGE OF POLAND KATARZYNA ZALASIŃSKA, DR HAB, A COOPERATION AGREEMENT IN THE FIELD OF CULTURAL HERITAGE PROTECTION.

[9] *Podpisanie porozumienia o współpracy z Narodowym Instytutem Dziedzictwa*, [online], https://mosib-ndkwz.wp.mil.pl/aktualnosci/podpisanie-porozumienia-o-wspolpracy-z-narodowym-instytutem-dziedzictwa/ [accessed: 28.01.2025].

The International Centre for Training and Research on Cultural Heritage in Danger and the National Institute of Cultural Heritage of Poland, in pursuit of their statutory objectives and tasks, have signed a cooperation agreement in the field of cultural heritage protection. In particular, the implementation of joint educational and training activities and the exchange of information and experience in the field of protection, respect and care of tangible heritage in the event of armed conflict and crisis situations, as well as undertaking joint activities for the construction, formation and strengthening of national identity through the presentation and popularisation of tangible and intangible military cultural heritage will be the subject of cooperation. (author's own resources)[10].

Systran translation (by RBMT):

SIGNINNG A COOPERATION AGREEMENT WITH THE NATIONAL HERITAGE INSTITUTE. 9 FEBRUARY 2024 DIRECTOR OF THE INTERNATIONAL CENTRE FOR TRAINING AND RESEARCH THE HERITAGE OF CULTURE IN DANGER WAS SIGNED BY COLONEL PAWEŁ NAWARYCZ WITH THE NATIONAL INSTITUTE OF HERITAGE DR. HAB. KATARZYNA ZALASIŃSKA COOPERATION AREA OF PROTECTION OF CULTURAL PROPERTY.

International Centre for Training and Research on Cultural Heritage in Danger and the National The Institute of Heritage, pursuing its objectives and statutory tasks, signed a cooperation agreement in the area of protection of cultural property. The cooperation will focus in particular on the implementation of joint education and training activities and the exchange of information and experience on protection, respect and protection of material heritage in the event of an armed conflict and crisis situations, as well as undertaking joint activities to build, shape and strengthening national identity through the presentation and popularisation of material intangible military cultural heritage[11].

Google translation (by SMT):

SIGNING OF A COOPERATION AGREEMENT WITH THE NATIONAL HERITAGE BOARD. ON 9 FEBRUARY 2024, THE DIRECTOR OF THE INTERNATIONAL CENTRE FOR TRAINING AND RESEARCH ON CULTURAL HERITAGE IN DANGER COL. PAWEŁ NAWARYCZ SIGNED WITH THE DIRECTOR OF THE NATIONAL INSTITUTE OF HERITAGE DR HAB. KATARZYNA ZALASIŃSKA, THE AGREEMENT ON COOPERATION IN THE FIELD OF PROTECTION OF CULTURAL GOODS.

The International Centre for Training and Research on Cultural Heritage in Danger and the National The Institute of Heritage, implementing its statutory goals and tasks, signed a cooperation agreement in the area of protection of cultural goods. The subject of cooperation will in particular be the implementation of joint education and training activities and the exchange of information and experience in the field of protection, respect and care for tangible heritage in the event of armed conflict and crisis situations,

[10] Author`s own resources.

[11] www.systransoft.com/lp/online-translation [online], [accessed: 28.01.2025].

as well as to take joint action to build, shape and develop a strengthening national identity through the presentation and popularisation of material and intangible military cultural heritage[12].

DeepL translation (NMT):

SIGNING OF COOPERATION AGREEMENT WITH NATIONAL HERITAGE INSTITUTE2024 02 09_01 FEBRUARY 9, 2024. DIRECTOR OF THE INTERNATIONAL CENTER FOR TRAINING AND RESEARCH ONCULTURAL HERITAGE IN DANGER COL. PAWEŁ NAWARYCZ SIGNED WITH THE DIRECTOR OF THE NATIONAL HERITAGE INSTITUTE DR. HAB. KATARZYNA ZALASIŃSKA AN AGREEMENT ON COOPERATION IN THE AREA OF PROTECTION OF CULTURAL PROPERTY.

The International Center for Training and Research on Cultural Heritage in Danger and the National Heritage Institute, pursuing their statutory goals and tasks, signed an agreement on cooperation in the area of protection of cultural property. The subject of cooperation will be, in particular, the implementation of joint educational and training activities, as well as exchange of information and experience in the field of protection, respect and care of material heritage in case of armed conflict and crisis situations, as well as undertaking joint activities to build, shape and strengthening of national identity through the presentation and popularisation of tangible and intangible military cultural heritage[13].

The following is the quantitative analysis of the translated texts (Poprawne, Systran, Google, DeepL) with a detailed comparison of the translations. The analysis was made with the use of Python. The code reads all translated documents, calculates texts metrics, identifies common words and computes the similarity between translations using TF (term frequency) -IDF (inverse term frequency) and cosine similarity.

First, basic metrics. The code to calculate basic metrics is: # Function to calculate basic metrics def calculate_metrics(text):

```
words = word_tokenize(text.lower())
return {
'word_count': len(words),
'unique_words': len(set(words)),
'avg_word_length': sum(len(word) for word in words) / len(words),
'sentence_count': text.count('.') + text.count('!') + text.count('?') }
```

(see Table 1):

[12] https://translate.google.pl [online], [accessed: 28.01.2025].

[13] www.deepl.com/pl/translator [online], [accessed: 28.01.2025].

Table 1. Basic metrics of translations

Type of translation	Word count	Unique words	Avg word length	Sentence count
Original	138	88	6.731884058	5
Poprawne	197	85	5.451766497	4
Systran	163	78	5.6257668712	5
Google	180	82	5.3	5
DeepL	173	81	5.4393063584	7

Source: compiled by the author.

Word count: Poprawne has the highest word count (197), while the original Polish text has the lowest (138). This suggests that Poprawne might be more explanatory or use more words to convey the same meaning.

Unique words: the original Polish text has the highest number of unique words (88), while Systran has the lowest (78). This could indicate that the original text is more diverse in vocabulary.

Average word length: the original Polish text has the highest average word length (6.73), while Google has the lowest (5.3). This difference is likely due to the nature of the Polish Language, which often has longer words compared to English.

Sentence count: DeepL has the highest sentence count (7), while Poprawne has the lowest (4). This suggests that DeepL might have broken down some complex sentences into simpler ones.

Lexical diversity: The original Polish text has the highest lexical diversity (0.638), while Poprawne has the lowest (0.431). This indicates that the original text uses a more varied vocabulary relative to its length.

Common words: all English translations show similar patterns in their most common words, with *of, the,* and *and* being the most frequent. This is typical for English texts. The original Polish text shows a different pattern due to the language difference.

Here, the lexical diversity is analysed (see Table 2) (the code: # Calculate lexical diversity (unique words / total words).

```
df_metrics['lexical_diversity'] = df_metrics['unique_words'] / df_metrics['word_
count']
print("\
Lexical Diversity:")
print(df_metrics['lexical_diversity']):
```

Table 2. Lexical diversity in all translations

Type of translation	Lexical diversity
Original	0.6376811594
Poprawne	0.4314720812
Systran	0.4785276074
Google	0.4555555556
DeepL	0.4682080925

Source: compiled by the author.

Most common words: the items in brackets are individual words or punctuation marks found in the texts. The number following each word indicates how many times that word appears in the text.

- Original (Polish text): (i, 9), (w, 6), (,, 5), (współpracy, 4), (dziedzictwa, 4), (kultury,4), (., 4), (o, 3), (nad, 3), (dziedzictwem, 3)
- Poprawne: (of, 20), (the, 16), (and, 12), (heritage, 9), (cultural, 8), (in, 8), (,, 8), (cooperation, 4), (agreement, 4), (national, 4)
- Systran: (the, 13), (of, 13), (and, 11), (heritage, 7), (in, 6), (,, 5), (cooperation, 4), (national, 4), (protection, 4)
- Google: (the, 17), (of, 15), (and, 12), (heritage, 7), (,, 7), (in, 7), (cultural, 5), (cooperation, 4), (national, 4), (on, 4)
- DeepL: (of, 15, (the, 11), (and, 11), (,, 9), (heritage, 7), (in, 7), (cooperation, 4), (national, 4), (cultural, 4).

The frequency of common words like *the*, *of*, and *and i*ndicates that these translations maintain a similar structure to the original text, focusing on key concepts like *heritage*, *cooperation***,** and *national*. The presence of specific terms like *heritage* and *cooperation* across all translations suggests that these concepts were effectively translated. However, the frequency of function words *the* and *of* can also indicate the overall complexity and structure of the sentences in the translations.

Next, the similarity matrix between translations (see Table 3) (the code: # Calculate similarity between translations).

```
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.metrics.pairwise import cosine_similarity
vectorizer = TfidfVectorizer()
tfidf_matrix = vectorizer.fit_transform([poprawne, systran, google, deepl])
similarity_matrix = cosine_similarity(tfidf_matrix)
```

```
translation_names = ['POPRAWNE', 'SYSTRAN', 'Google', 'DeepL']
df_similarity = pd.DataFrame(similarity_matrix, index=translation_names, columns=translation_names):
```

Table 3. Similarity matrix between translations

	Poprawne	Systran	Google	DeepL
Poprawne	1	0.898187762	0.9235321434	0.9114227064
Systran	0.898187762	1	0.9260299352	0.9433779515
Google	0.9235321434	0.9260299352	1	0.917070105
DeepL	0.9114227064	0.9433779515	0.917070105	1
DeepL	173	81	5.4393063584	7

Source: compiled by the author.

All translations show high similarity to one another (above 0.98), with Systran and DeepL being the most similar (0.9433), and Poprawne and Systran being the least similar (0.8981). The results obtained of similarities and differences between all 4 translations show the basic metrics like word count, unique words, lexical diversity and similarity matrix. However, those parameters do not give an exhaustive answer to the question about choosing the best translation and if the professional translator can be replaced by machine translation.

That is why a comparative analysis of these 4 translations is needed to decide which one can be seen as the best, which as the worst, and if it is possible to use machine translation as the final product of translation.

The comparative analysis is conducted in 2 criteria: fidelity to the source, and fluency and style.

1. Fidelity to the source. This criterion evaluates how accurately the translation reflects the original text, including the preservation of names, terms, and essential details:

   Poprawne: this translation excels because it maintains the full official names of the institutions and all relevant details from the source text like "NATIONAL INSTITUTE OF CULTURAL HERITAGE OF POLAND (NARODOWY INSTYTUT DZIEDZICTWA NID)". This accuracy is crucial for formal documents.

   DeepL translation also performs well, closely mirroring the original text and maintaining the integrity of the names and details.

Google translation has a slight alteration in the name of one of the institutions, which can lead to confusion or misrepresentation, i.e. "National Heritage Board" instead of "Institute".

Systran translation suffers from inconsistencies in terminology and may misrepresent some details, which lowers its fidelity to the source, i.e. containing typographical and grammatical errors (e.g. signinng) or 'The heritage of culture in danger".

2. Fluency and style. Th is criterion assesses how natural and polished the translation reads in the target language, concerning grammar, syntax, and overall style.

Poprawne can be noted for its fluency and expert crafting, making it read smoothly and professionally, which is essential for formal documents.

DeepL translation is also quite fluent, with natural phrasing. However, it may have minor formatting issues around the date that detract slightly from its overall presentation.

Google translation, while generally clear, lacks the formal style and precision of the top translations, making it feel less polished.

Systran contains several awkward phrases and grammatical errors, which significantly impact its readability and professionalism. It is less reliable due to its poor quality.

It can be said that Poprawne translation combines high accuracy with a polished presentation, making it the most reliable choice for formal contexts. The second best translation is DeepL, which is strong in natural phrasing and consistency, though it falls short of the professional Polish, then comes Google translation, which is generally clear, but the misnaming of an institution affects its reliability and credibility. The worst one is Systran translation because of the presence of typographical and grammatical errors, along with inconsistent terminology, which makes it the least reliable option.

4. Conclusion

By comparing the frequency of words across different translations, one can assess how each translation system handles the text. I.e. if one translation has a significantly higher frequency of a particular word, it may indicate a different approach to phrasing or sentence structure. The analysis provides insight into the linguistic

choices made by each translation system and how closely they align with the original Polish text in terms of word usage and structure.

Based on this quantitative analysis, it is difficult to definitely say which translation is the best or worst as each has its strengths and weaknesses. Poprawne uses more words, which might make it more explanatory but less concise. Systran has the lowest number of unique words, which might indicate less vocabulary richness. Google has the shortest average word length, which could make it more readable but possibly less precise. DeepL has the highest sentence count, which might make it easier to read but could alter the original text structure. The high similarity between translations (> 0.89) suggests that they all convey similar information. The choice of the best translation might depend on specific needs, such as conciseness (favouring Systran), readability (favouring Google or DeepL), or explanatory power (favouring Poprawne). However, as can be seen, all 3 machine translations contain errors, mainly stylistic, grammatical and inconsistency in terminology.

For a definite answer on quality, a comparative analysis is necessary to assess accuracy and nuance in translation. As the results of this analysis show, machine translations still require proofreading by a professional translator to be seen as the final product of the process of translation.

REFERENCES

Arnold D., Balkan L., Meijer S., Humphreys R.L., Sadler L., *Machine Translation: An Introductory Guide*, Manchester 1994.

Bahdanau D., Cho K., Bengio Y., *Neural Machine Translation by Jointly Learning to Align and Translate* (paper presented at 3rd International Conference on Learning Representations), ICLR 2015, San Diego 2015.

Brown P.F., Cocke J., Della Pietra V.J., Jelinek F., Lafferty J.D., Mercer R.L., Roossin P.S., *A Statistical Approach to Machine Translation*, "Computational Linguistics" 1990, 16 (2), pp. 79–85

Deepl, [online], www.deepl.com/pl/translator [accessed: 28.01.2025].

Picken C. (ed.), *The Translator's handbook*, London 1989.

Podpisanie porozumienia o współpracy z Narodowym Instytutem Dziedzictwa, [online], https://mosibndkwz.wp.mil.pl/aktualnosci/podpisanie-porozumienia-o-wspolpracy-z-narodowym-instytutem-dziedzictwa [accessed: 28.01.2025].

Systransoft, [online], www.systransoft.com/lp/online-translation/ [accessed: 28.01.2025].

Translate Google, [online], https://translate.google.pl [accessed: 28.01.2025].

Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A.N., Kaiser Ł., Polosukhin I., *Attention is All You Need. Advances in Neural Information Processing Systems*, New York 2017.

Vermeer H.J., *Skopos and Commission in Translational Action. The Translation Studies Reader*, L. Venuti, M. Baker (eds.), London–New York 2001.