Kaunas University of Technology

Faculty of Social Sciences, Arts and Humanities

Quality and Acceptability of Machine Translation in Spam Websites: a Linguistic and User-Centered Analysis

Master's Final Degree Project

Eglė Miltakienė

Project author

Prof. dr. Ramunė Kasperė

Supervisor

Kaunas, 2023

Kaunas University of Technology

Faculty of Social Sciences, Arts and Humanities

Quality and Acceptability of Machine Translation in Spam Websites: a Linguistic and User-Centered Analysis

Master's Final Degree Project

Translation and Localization of Technical Texts (6211NX031)

Eglė Miltakienė

Project author

Prof. dr. Ramunė Kasperė

Supervisor

Lect. dr. Brigita Dimavičienė

Reviewer

Kaunas, 2023

Kaunas University of Technology

Faculty of Social Sciences, Arts and Humanities

Eglė Miltakienė

Quality and Acceptability of Machine Translation in Spam Websites: a Linguistic and User-Centered Analysis

Declaration of Academic Integrity

I confirm the following:

1. I have prepared the final degree project independently and honestly without any violations of the copyrights or other rights of others, following the provisions of the Law on Copyrights and Related Rights of the Republic of Lithuania, the Regulations on the Management and Transfer of Intellectual Property of Kaunas University of Technology (hereinafter – University) and the ethical requirements stipulated by the Code of Academic Ethics of the University;

2. All the data and research results provided in the final degree project are correct and obtained legally; none of the parts of this project are plagiarised from any printed or electronic sources; all the quotations and references provided in the text of the final degree project are indicated in the list of references;

3. I have not paid anyone any monetary funds for the final degree project or the parts thereof unless required by the law;

4. I understand that in the case of any discovery of the fact of dishonesty or violation of any rights of others, the academic penalties will be imposed on me under the procedure applied at the University; I will be expelled from the University and my final degree project can be submitted to the Office of the Ombudsperson for Academic Ethics and Procedures in the examination of a possible violation of academic ethics.

Eglė Miltakienė

Confirmed electronically

Miltakienė, Eglė. Quality and Acceptability of Machine Translation in Spam Websites: a Linguistic and User-Centered Analysis. Master's Final Degree / supervisor prof. dr. Ramunė Kasperė; Faculty of Social Sciences, Arts and Humanities, Kaunas University of Technology.

Study field and area (study field group): Translation (N05), Humanities.

Keywords: machine translation, translation quality assessment, acceptability, linguistic error analysis, spam websites.

Kaunas, 2023. 54 p.

Summary

The study focuses on the prevalent issue of machine-translated spam websites which are rapidly proliferating in the search engines, offering information of dubious quality. These websites use machine translation software to create content in multiple languages, potentially leading to misinformation among users. The central problem is the users' lack of awareness about the machine-translated nature of the content and potential inaccuracies it may contain. Despite the growing abundance of such content, research on its linguistic quality and end-user acceptability is limited. The research aims to evaluate the linguistic quality through error analysis and end-user acceptability of machine-translated spam websites in the Lithuanian language, employing a combination of automated and human assessment tools for machine translation output quality.

The study aims to achieve these objectives:

1. to provide a review of the most recent literature on translation quality, the various methods of measuring it, and end-user acceptability.
2. to analyse the acceptability of machine-translated spam websites from the user's perspective.
3. to measure machine-translated spam websites text using automatic evaluation BLEU score.
4. to evaluate linguistic quality and identify errors present of machine-translated spam websites text following the Multidimensional Quality Metrics (MQM).

A comprehensive evaluation of machine-translated websites was conducted using a combination of qualitative, quantitative, and descriptive methods. The study employed three methodological strategies: a user-centered questionnaire to understand acceptability of machine-translated websites, the application of the BLEU score for an automated evaluation of machine translation quality, and human expert analysis using Multidimensional Quality Metrics by Lommel et al. (2014) to classify translation errors. Additionally, the acceptability of the end-user is of equal importance, as they are the intended recipients of the final translation product.

Based on the analysis of the data, several conclusions were reached. Translation quality evaluation studies acknowledge the lack of a universal evaluation method due to the subjectivity of the concept and the evolving translation technologies. Thus, incorporating various evaluation methods can lead to more comprehensive and accurate results in assessing translation quality. The evaluations conducted by end-users showed varying levels of acceptability regarding the quality of translation. Certain aspects, like clarity and informativeness, received positive feedback, while others like enjoyability, trustworthiness, and quality were questioned. Quality evaluations also varied across age groups. Moreover, machine translation quality varied across websites, as indicated by BLEU scores. Quality discrepancies could be attributed to text complexity, with better performance in familiar contexts. Yet, the BLEU score is limited in capturing all translation errors as it considers exact word matches only. Human expert evaluation using the MQM framework identified numerous translation

errors across various categories, with accuracy, linguistic conventions and terminology having the highest error rate. This underlines the need for thorough post-editing by professional translators to enhance the quality of machine translation.

This study integrates diverse methodologies, such as human linguistic analysis, social survey, and automatic language evaluation, to yield more holistic and objective insights. The study also emphasizes that this topic is equally important to both language researchers and end-users of the translation product. As the volume of machine-translated text grows with technological advances, there's a rising need to educate society about the technologies, their advantages, and risks tied to low-quality or inaccurate content, enabling informed decision-making. Notably, this study highlights the significance of both expert quality assessment and public involvement in evaluating translation quality.

The research conducted for this Master's thesis was presented at the student scientific conferences "SMILES 2022: Social Sciences, Arts and Humanities in Contemporary Society" and "The New Generation of Scientists" organized by the Lithuanian Council of Sciences in 2023. Additionally, a publication based on this research is planned.

Miltakienė, Eglė. Mašininio vertimo kokybės ir priimtinumo šlamšto svetainėse lingvistinė ir vartotojų požiūrio analizė. Magistro studijų baigiamasis projektas / vadovė prof. dr. Ramunė Kasperė; Kauno technologijos universitetas, Socialinių, humanitarinių mokslų ir menų fakultetas.

Studijų kryptis ir sritis (studijų krypčių grupė): Vertimas (N05), Humanitariniai mokslai.

Reikšminiai žodžiai: mašininis vertimas, vertimo kokybės vertinimas, priimtinumas, kalbos klaidų analizė, šlamšto svetainės.

Kaunas, 2023. 54 p.

Santrauka

Šiame tyrime nagrinėjamas vis dažniau pasitaikantis fenomenas – mašininiu būdu išverstos šlamšto svetainės, kurios, generuodamos abejotinos kokybės informaciją, gausėja paieškos sistemose. Kuriant šias svetaines yra pasitelkiama mašininio vertimo įrangą siekiant greitai sukurti turinį išverstą į kuo daugiau skirtingų kalbų. Turinys, kurį generuoja šlamšto svetainės, gali būti ne tik faktiškai netikslus, bet ir pasižymėti prasta lingvistine kokybe bei klaidų gausa. Vienas iš esminių iššūkių yra nepakankamas vartotojų informuotumas, susijęs ne vien su minėtų svetainių egzistavimu, bet ir su potencialiomis grėsmėmis, kurias gali sukelti prastos kokybės turinio priėmimas kaip faktiškai teisingas. Nepaisant to, kad šio pobūdžio turinys vis dažniau užima erdvę internete, moksliniai tyrimai, skirti mašininio vertimo lingvistinės kokybės bei priimtinumo galutiniam vartotojui nustatymui, yra nepakankami. Šio tyrimo tikslas yra išanalizuoti lingvistinę kokybę ir priimtinumą galutiniam vartotojui į lietuvių kalbą išverstų šlamšto svetainių, nagrinėjant klaidas, taikant tiek automatinio, tiek žmogaus mašininio vertimo kokybės vertinimo metodus.

Šiame darbe keliami tokie tyrimo uždaviniai:
1. atlikti išsamų naujausios literatūros, susijusios su vertimo kokybe, jos vertinimo metodikomis ir galutinio vartotojo priimtinumu, apžvalgą;
2. įvertinti mašininiu būdu išverstų šlamšto svetainių priimtinumą vartotojo perspektyvoje;
3. įvertinti mašininiu būdu išverstų šlamšto svetainių tekstus, taikant automatinio vertinimo BLEU metodiką;
4. nustatyti mašininiu būdu išverstų šlamšto svetainių teksto lingvistinę kokybę ir identifikuoti esamas klaidas, remiantis daugiamate kokybės metrika (angl. Multidimensional Quality Metrics, MQM).

Buvo atliktas nuodugnus mašininiu būdu išverstų šlamšto svetainių vertinimas, pritaikant kiekybinius, kokybinius ir aprašomuosius analizės metodus. Pasitelkta tyrimo metodika apimė tris strategijas: klausimyną skirtą galuniam vartotojui, kurio tikslas įvertinti mašininiu būdu išverstų šlamšto svetainių priimtinumą, BLEU balų pritaikymą mašininio vertimo kokybės automatizuotam vertinimui, bei žmogaus, remiantis Lommel et al. (2014) daugiamate kokybės metrika, analizę skirtą vertimo klaidoms identifikuoti.

Analizuojant surinktus duomenis, buvo suformuluoti keli esminiai apibendrinimai. Literatūroje pripažįstama, kad vertimo kokybės vertinimo srityje nėra universalaus vertinimo būdo, nes kokybės sąvoka yra subjektyvi, o ir vertimo technologijos nuolat vystosi. Todėl, siekiant gauti nuodugnesnius ir tikslesnius vertimo kokybės rezultatus, yra prasminga taikyti įvairius vertinimo metodus. Atliktas naudotojų vertinimas parodė, kad vertimo kokybė buvo priimama labai įvairiai. Kai kurie aspektai, tokie kaip teksto aiškumas ir informatyvumas, sulaukė teigiamų atsiliepimų, tačiau kiti aspektai, tokie kaip malonumas skaityti, patikimumas ir bendra kokybė, sukėlė naudotojams abejonių. Be to, įvairios

amžiaus grupės skirtingai įvertino mašininio vertimo kokybę. Surinkti BLEU balai rodo, kad mašininio vertimo kokybė skyrėsi tarp skirtingų svetainių. Tokius kokybės skirtumus galima paaiškinti teksto sudėtingumu, t. y. mašininis vertimas gali pasižymėti geresniais rezultatais, kai jis geriau supranta kontekstą ir terminologiją. Nepaisant to, BLEU metodika tik ribotai atspindi visus vertimo klaidų aspektus, nes ji orientuota į tiesioginius žodžių atitikimus sakiniuose. Atitinkamai žmogaus atliktas kokybės vertinimas, naudojant daugiamatės kokybės metriką (MQM), išryškino įvairių vertimo klaidų kategorijas, o daugiausia klaidų buvo pastebėta tikslumo, gramatikos ir terminologijos kategorijose. Tai rodo, kad norint optimizuoti mašininio vertimo kokybę, vis dar reikia profesionalių vertėjų ir redaktorių, kurie atliktų kruopštų teskto redagavimą.

Šiame tyrime derinamos įvairios metodikos, įskaitant lingvistinę analizę, socialinę apklausą ir automatinį vertimo kokybės vertinimą, siekiant gauti platesnes ir objektyvesnes įžvalgas. Taip pat akcentuojama, kad ši tema yra svarbi tiek kalbos tyrėjams, tiek galutiniams vertimo produkto nautotojams. Kalbos technologijoms progresuojant ir mašininio vertimo tekstų kiekiui internete augant, visuomenės švietimo poreikis apie tokias technologijas, jų privalumus ir rizikas, susijusias su nekokybišku ar netiksliu turiniu, taip pat didėja, tai svarbu, kad tokio turinio skaitytojas galėtų priimti informuotus sprendimus. Dar svarbu paminėti, kad šiame tyrime yra pabrėžiama tiek ekspertų kokybės vertinimo, tiek visuomenės dalyvavimo svarba vertinant vertimo kokybę.

Verta paminėti, kad šiame magistro darbe atliktas tyrimas buvo pristatytas mokslinėje studentų konferencijoje „SMILES 2022: Socialiniai, humanitariniai mokslai ir menai šiuolaikinėje visuomenėje“ ir taip pat buvo pristatytas 2023 metais organizuotoje Lietuvos mokslų tarybos „Naujoji mokslininkų karta“ studentų konferencijoje. Remiantis atlikto tyrimo rezultatais, planuojama parengti publikaciją.

Table of contents

List of figures

List of tables

Introduction

In the current landscape of rapid advancements in information technology, online content is disseminating at an unprecedented pace, has greater influence, and is more accelerated than ever. Notwithstanding its numerous advantages, the expansion of online content is a subject to debate. One issue is the lack of control over the amount and quality of online content available on the internet (Menczer, 2016). As an illustration, search engines have become a convenient way to find relevant information quickly and efficiently, but they may also present certain risks to society.

A matter of paramount concern is the rampant spread of machine-translated spam websites on search engines that offer dubious quality information. Machine-translated spam websites are websites that use machine translation software to create and distribute spam content in multiple languages. When a user performs a search query, the search engine generates a list of results and recommends websites that may not be in the original language of the entered keyword, but instead machine-translated version of a website. The information provided on these websites may be inaccurate, misleading, or harmful to the reader's perception, opinion or even health, depending on the topic.

The primary issue is that the users may lack the awareness that the text they are reading is machine-translated, let alone that the information provided may contain factual inconsistencies. Despite the abundance of machine-translated content available on the internet, research on its linguistic quality and end-user acceptability is limited. The central **problem** addressed by this study is whether the quality of machine-translated spam websites is acceptable and how users perceive it.

Several studies have indicated that the accuracy and quality of machine translation in low-resource languages like Lithuanian are not up to the mark (Petkevičiūtė & Tamulynas, 2011; Stankevičiūtė et al., 2017; Kasperavičienė et al., 2020). The accuracy of machine translation into Lithuanian from different languages has been found to be questionable, with potentially significant errors and inaccuracies in the output. Additionally, there is a lack of research regarding the end-user's perception and acceptability of machine-translated content (Castilho & O'Brien, 2017; Castilho, 2016). Within academic discourse, the term "acceptability" denotes the degree of usability, quality, and satisfaction achieved by a given text, which may be evaluated by means of survey-based approaches (Castilho, 2016).

The choice to investigate this topic was motivated by widespread distribution of machine-translated content online and its accessibility to society.

Therefore, **the aim** of the research is to evaluate the linguistic quality through error analysis and end-user acceptability of machine-translated spam websites in the Lithuanian language, employing a combination of automated and human assessment tools for machine translation output quality.

The study seeks to accomplish the following **objectives**:

1. to provide a review of the most recent literature on translation quality, the various methods of measuring it, and end-user acceptability.
5. to analyse the acceptability of machine-translated spam websites from the user's perspective.
6. to measure machine-translated spam websites text using automatic evaluation BLEU score.
7. to evaluate linguistic quality and identify errors present of machine-translated spam websites text following the Multidimensional Quality Metrics (MQM).

To ensure that the research is conducted with a thorough and rigorous approach, a combination of qualitative, quantitative, and descriptive methods are utilized to achieve the most comprehensive and robust results possible.

Within the current climate of swift content distribution on the internet, this study proposes that AI-generated and machine-translated content will experience a growth in the future. Hence, this study not only serves the purpose of evaluating the quality of machine-translated websites, , but also addresses the need to increase societal awareness regarding potential internet risks.

The research conducted for this Master's thesis was presented at the student scientific conferences "SMILES 2022: Social Sciences, Arts and Humanities in Contemporary Society" and "The New Generation of Scientists" organized by the Lithuanian Council of Sciences in 2023. Additionally, a publication based on this research is planned.

1. A Translation Studies Perspective on Evaluating Machine Translation Quality and Acceptability of Websites

The internet offers extensive access to a plethora of information on various subjects, including education, news, entertainment, personal growth, and beyond. Every day, an enormous number of users access the internet to consume content for diverse purposes. The quality and reliability of online available content is highly unpredictable. With the expanding use of AI-powered text creation technologies and tools, there is an expected flood of machine-generated content available online with no human input. This raises concerns about the potential quality of information that will be available in the future.

In addition to the enthusiasm surrounding neural machine translation, it is evident that the quality of machine translation has been improving and the texts generated are more fluent than ever before. These advancements are due to the use of neural networks and deep learning algorithms, which have considerably increased translation output accuracy. Machine translation is therefore becoming an increasingly popular and viable option for practitioners, students and ordinary users alike. It is clear that machine translation technologies are now automating a significant portion of the translator's work, causing the industry to shift towards post-editing. Also, there is a continual expansion of research in this field, leading to the emergence of novel machine translation paradigms. Today, machine translation has become a focal point of both translation studies, computational linguistics and artificial intelligence research. Nonetheless, the high level of quality achieved by machine translation is limited to only a few languages, neglecting the majority of low-resource languages, whose translation quality remains questionable (Costa-jussà, 2022).

In this study, the focus is on the analysis of machine-translated spam websites as an example of the mass production of content on the internet. The main concern is how machine-translated content is perceived and accepted by the end-users, particularly given the fact that users may engage with it unknowingly. The question is whether a typical user of a website can recognize if the text displayed on the website is machine-translated, and if they consider such content as trustworthy and acceptable. The current study employs multiple proven methods of translation evaluation, including automatic evaluation, human evaluation, and the survey method, to assess the quality of machine-translated websites from the user's perspective. When analyzing a subject such as spam websites, which are easily accessible to anyone, a multi-dimensional deabte emerges that takes into account the linguistic construction of the text, the user's perception of it, and the subjectivity involved in evaluation. Also, the resulting translated text is regarded as a technological product. Given the potential prevalence of such content in the future, the topic is particularly relevant, as human-written or human-translated text may become increasingly rare. The interdisciplinary nature of this subject, which encompasses translation and linguistics, technology, and social sciences, highlights the significance and relevance of this topic both for academia and to society. Thus, in the following sections, we will review some of the leading scientific discussions regarding these topics.

1.1. Machine-Translated Spam Websites

The development of machine translation technology has enabled website owners to extend their reach in other countries by creating multilingual websites. However, this practice has not always been used with positive intentions. Commercial websites aim to improve their rankings on search engine results pages, and as such, they may adopt various strategies, such as ethical methods like search engine optimization (Al-Kabi et al., 2012), or unethical and illegal techniques like website spam, which is considered a deceptive practice. Gyongyi et al. (2005) defined website spam as the act of publishing

hyperlinked websites on the internet to manipulate search engine rankings and improve their visibility on search results pages. Machine-translated spam websites use machine translation software to automatically translate content from one language to another, usually with the goal of creating low-quality, spam content for search engine optimization purposes. Typically, these websites are produced in large quantities by individuals or organizations commonly referred to as spammers, with the intention of tricking both search engines and internet users. Furthermore, Google's spam regulations[1] provide the following definitions for automatically generated spam content in search engines:

- text that includes search terms but lacks coherence;
- text that is machine-translated without human review;
- text that is produced by automated procedures without considering quality or user experience;
- text created through automatic synonymization, paraphrasing, or obfuscation;
- content that merges from other websites without providing sufficient value.

Aswani et al. (2021) found that users prefer to visit websites that are ranked higher and appear on the first page of search results. And despite the effort search engines put in to ensure quality search results, the web pages that appear on the first page of search results may not always be the most reliable or of the highest quality. Spam websites still frequently appear among the top search results for particular keywords. Consequently, users end up accessing these spam websites. Upon visiting such a webpage, a user may observe that the language employed on the page lacks logical coherence and is filled with grammatical errors. The user experience is typically poor and the website is full of odd advertisements and redirects. While it is reasonable to assume that an educated visitor to such a site will recognize that something is wrong and exit, this is not always the case.

The impact of machine-translated spam websites on society can be significant. First, these sites can be used to spread plagiarism, propaganda, or information that is misleading. For example, a machine-translated webpage, could be created to spread false information about medical conditions or medications, potentially resulting in harm to user with no medical knowledge, who may not be able to distinguish the misleading content. In this way, machine-translated spam websites also damage the reputation the overall trust in the internet.

In view of this, multiple studies have been conducted to examine an individual's ability to differentiate between machine-generated content and human-written content. While this paper specifically focuses on machine translation, research on automatically generated content is a very similar topic, exploring the impact of language technology, which is also relevant to this study. For instance, as machine-generated content becomes more prevalent in journalism, Clerwall (2014) found that readers have difficulty distinguishing between news articles written by humans and those generated by machines. Although the study specifically focused on online journalism, the implications of machine-generated content are far-reaching, as emerging technologies continue to expand its use. Consequently, the line between human-generated and machine-generated text may blur, affecting society as a whole.

Taken together, despite the considerable improvement in machine translation technology over the years, certain machine translation systems can now produce translations that are almost as accurate as those produced by human translators. Yet, machine translation is far from perfect, and the quality of the output may be influenced by factors such as language pair, text complexity, and software

[1] Spam policies for Google web search.
https://developers.google.com/search/docs/essentials/spam-policies?hl=en&visit_id=638020383807236347-3382164925&rd=1#spammy-automatically-generated-content

quality. For this reason, machine-translated content, in the spam websites referred earlier, is often of low quality and full of errors. Furthermore, these websites often contain irrelevant or misleading information, which can lead to a negative impact on user experience and even harm to society. It appears that the use of machine translation technology is not limited to ethical applications, as it can also be employed for unethical purposes such as generating plagiarized content, publishing misinformation, or misleading information. Unfortunately, even though search engines are implementing measures to combat website spam issue but the technology still has a long way to go. Therefore, machine-translated spam websites remain a major problem for search engines, and both to users. It appears that it is ultimately the responsibility of internet users to objectively evaluate the content of these websites and avoid accessing them.

In the context of this study, it is equally important to differentiate and evaluate what constitutes a high quality website from one that is low quality. Therefore, below we will examine the quality factors that influence how users perceive a website. Users visit websites for variety of reasons, however, their decision to stay and engage with the website, as well as trust the information presented on it, are largely dependent on the website's overall quality. The quality of a website is a critical factor in determining whether users will have a positive experience and continue to use the site. It is worth emphasizing that the website is intended for the end-user, who is the ultimate consumer of the product. Thus, it can be argued that it holds great significance to consider which evaluation criteria are pertinent for assessing the quality from the perspective of the end-user. Furthermore, as in many other domains of research, the quality of a website from an end-user standpoint can be highly subjective, and a website that appears to be of high quality to one reader may be perceived differently by another. As Herrera-Viedma et al. (2006) noted, subjectivity plays a role as each user's interaction with the content and degree of satisfaction with it, vary. Many researchers have attempted to identify and define indicators of a website quality. Although there is no single view, however many scientific studies suggest that the parameter of subjectivity should be considered when appraising the quality of a website (Thielsch & Hirschfeld, 2019; Herrera-Viedma et al., 2006; Koch & Hartmann, 2022). In this section, we will look at several cases from scientific literature.

Koch and Hartmann (2022) propose one viewpoint in their study, where they examine several other studies and identify key factors that influence website quality. The authors describe these factors as follows:

- content: refers to whether it answers the question of the visitor;
- design: which encompasses website layout, navigation and interactivity;
- and the overall impression: which denotes the coherence of all the individual website's elements.

In their study, Thielsch and Hirschfeld (2019) emphasize the significance of evaluating the individual facets of content that users experience. By drawing on prior online content studies, the authors developed a metric to measure user's subjective evaluation criteria of website content. This metric involves four main aspects of a user's content experience that reflects subjective content perception: clarity, likeability, informativeness, and credibility.

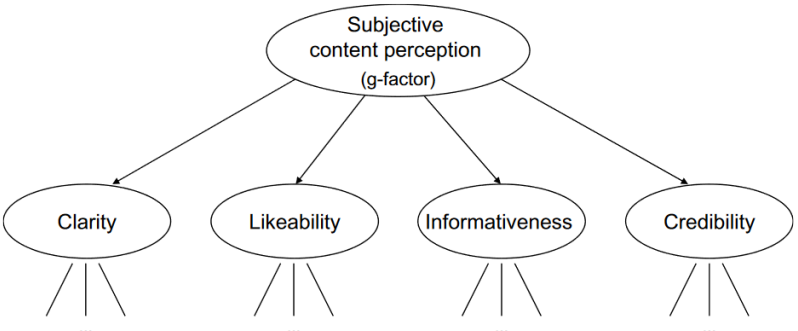


Fig. 1. The Web-structural CLIC's model (Thielsch & Hirschfeld, 2019)

According to Thielsch and Hirschfeld, the clarity criterion assesses the user's ability to understand the presented information, which is sometimes referred to as comprehensibility and ease of understanding in academic sources. Likeability measures both user engagement and emotional experience towards website content, and it is also studied as perceived attractiveness or entertainment. Informativeness represents the perceived amount of valuable and useful information on a website, while credibility assesses the trustworthiness of the provided information (2019). The authors state an important point that a high rating on a particular criterion does not necessarily indicate that the website's content is of remarkably good quality, but rather that the website user has a favorable perception of the content. This illustrates the importance of considering subjectivity when assessing website content quality.

The research performed by Kincl and Štrach (2012) indicates the importance of content when determining website quality and user satisfaction. The findings revealed that users give positive ratings to websites when they successfully accomplish the intended purpose they had in mind when visiting the site. Moreover, the study highlighted that content and navigation are the most influential factors in users' evaluation of website quality. These results emphasize the role of content and its function for website users.

Another noteworthy study, conducted by Kim and Niehm (2009), which explore various quality factors of websites that impact purchase intention. The study found that users evaluated the quality of website content based on its correctness, informativeness, timeliness, and relevancy. One notable finding was the correlation between the quality of information presented on a website, users' trust, and their intention to make purchases. Another observation was, users who perceived the website's content to be of high quality were more likely to consider the product prices to be fair. The scholarly literature emphasizes the necessity of understanding the factors that influence website users' perceptions. Kang and Kim (2006), for instance, propose that a user's website experience is primarily shaped by the quality of the website, encompassing the informativeness of the content, the entertaining aspect of the content, and the ease of navigation.

The evaluation of a website's quality requires consideration of several factors, but it is essential to recognize that users' educational level, age, technological proficiency, and a general understanding of how the website works can impact their judgment. In the contemporary digital era, online participation is a vital skill, akin to reading or writing. Despite this, many users lack basic knowledge and have limited comprehension of how internet operates. Consequently, a certain level of literacy is necessary to understand the digital medium's content. This perception could be called user website literacy and is another important aspect in the overall picture of quality from an end-user perspective.

Based on the literature review, it can be observed that various sources have slightly different criteria for determining website quality. However, there are some similarities such as fulfilling the purpose for which the visitor has come to the website; providing informative and reliable content; and presenting it clearly and comprehensibly. In this study, we will also use the quality factors identified by the authors mentioned above to assess how users perceive website quality.

1.2. Translation Quality Assessment

The concept of translation quality assessment (TQA) is a vital aspect of translation studies, relevant to researchers, professionals, practitioners, and students in the field. Because translation is a complex phenomenon where a text is re-contextualized from one language to another, influenced not only by linguistic and textual factors but also by a range of extra-linguistic conditions (House, 1997), unsurprisingly, there are various approaches in the literature to analyze the TQA notion, given the broad and multifaceted nature of this term. As a case in point, Gouadec (2010) states that quality in the field of translation encompasses both the end product, which is the translated text, and the transactional process, which is the service provided. It is evident that the scope of the discussion and literature related to the assessment of translation quality can be foreseen. Hence, it is noteworthy to mention that this review of literature will concentrate on the quality of translation as a product rather than a service. This section explores TQA from both scientific and practical perspectives, emphasizing the need to understand how to evaluate translation technologies output and their quality, particularly in light of the growing prevalence of automated text translation and production. As the distinction between human and machine-generated text blurs, adaptation to rapid change will become an industry-wide challenge.

TQA is a multi-component process that evaluates various aspects of a translated text, including linguistic accuracy, grammatical correctness, clarity, coherence, and more, often using a combination of human and automated evaluation methods. This process is important in determining the success of a translation project, study, or a task. Yet, the most important aspect determining the quality of a translated text is how to measure it, and throughout history, numerous methods and attempts have been made to assess translation quality. The quality of translation has been a significant aspect of the translation process since the early literature on translation studies. Nida (1964) established the theoretical foundation for evaluating translation quality, stressing its significance and presenting a methodology based on three important factors: accuracy, fluency, and naturalness. According to Nida, a high-quality translation should accurately express the intended meaning of the source material and be clearly readable. This perspective to translation quality assessment, which emphasizes accuracy, clarity, and readability, appears appropriate for ensuring faithful transmission of the source message. However, there has subsequently been a debate among scholars that such a requirement is somewhat vague. Holmes (1988) stressed the necessity of a systematic approach to assessing translation quality, and his work on TQA emphasised the importance of evaluating translations using both objective and subjective criteria. It is undisputed that a systematic method to evaluation not only provides greater clarity but also sets specific criteria and precise goals for quality assessment. According to Williams (2004), there are two types of translation quality assessment (TQA) models: quantitative models that focus on micro-textual analysis, that attempt to numerically quantify the quality of translation based on the number of errors, and qualitative or general models that use macro-textual analysis and do not focus on the number of errors. House presented an alternative view for evaluating translation quality in 2010, which considers both linguistic and cultural factors. House stressed the importance of evaluating translation quality holistically, considering the full translation process, rather than just the final outcome. When considering the wide range of theoretical perspectives in translation studies, it

is also undeniable that certain aspects, such as cultural factors in the context of localization, hold significant importance. Apart from House's ideas, another academic, Pym, proposed parallel model. Pym (2010) proposed a contextual and dynamic approach to assessing translation quality that considers a broad range of factors that can impact the quality of a translation. Pym suggested that the quality of a translation cannot be assessed using a standard set of criteria, but rather is determined by the context in which the translation is produced and then used (2010). This assertion is indeed agreeable, as different contextual factors such as the intended audience and cultural factors can influence the criteria for assessing translation quality in that particular context. As a result, the notion of a quality translation may differ according to the circumstances. As the several academic disciplines of TQA advanced, the rise of machine translation appeared as a crucial development that impacted the entire domain of translation studies. Despite the fact that machine translation has been present for 70 years, it is without question that it has hastened the development of new assessment methodologies, and has resulted in a wider range of such, in addition to the existing theoretical debate on quality assessment. In the following, we will look more closely at the evaluation of machine translation and its quality.

In a general sense, translation quality, like any other qualitative evaluation subject, can be assessed using traditional academic methods such as quantitative or qualitative methodologies, or it can be evaluated using surveys or interviews (Esmail & Jaza'ei, 2015), while modern and innovative approaches employ tools such as eye tracking or cognitive methods (Ferreira, 2016; Kasperavičienė 2020). Another academic viewpoint in translation studies worth exploring is the assessment of translation quality based on the end-user's viewpoint. For instance, it can be concurred that the quality of the translation is important as it assists in detecting potential issues present in the translated text, such as errors, omissions, or cultural diversity. These issues might adversely impact the translation's effectiveness or its acceptance by the end-user. The user-centered approach is an alternative viewpoint that has received limited attention from researchers in the context of TQA. It emphasizes the end-user's perspective, as they are the intended audience for the final translated product. Further to the point made above, a non-expert may consider a text as of acceptable quality while an expert in the field may perceive it as low-quality. Moreover, in terms of evaluating the quality of a specific translation, the end-user may frequently have inadequate literacy skills or insufficient knowledge of the target language.

Notwithstanding the importance of assessing the quality of translation, researchers and practitioners do not share a consensus or a universal model to accurately evaluate the quality of translated texts. At both the micro and macro levels, methods and procedures in TQA tend to vary since the concept of quality can differ significantly among different individuals, groups, and contexts (Castilho, 2018). Therefore, a majority of researchers emphasize that a single model cannot be applied to all translation projects or situations (Drugan, 2013; Lommel, 2014; Doherty 2017). Taking into account the above arguments, the significance of TQA is undeniable and constitutes a substantial portion of all research in translation studies, and nevertheless it is essential to view translation quality assessment as an ongoing practice that must be regularly updated and improved, as the translation industry context continually changes and evolves along with translation practice and technology. According to the foregoing review of the literature, there is no universal criterion to evaluate the quality of a particular project or output. This implies that a combination of methods and tools can be used based on the context and purpose for which the translation is being assessed.

1.3. Machine Translation Evaluation

Machine translation, originally developed in the 1940s as computational systems for translating texts, has recently undergone a paradigm shift from statistical to neural systems, which has propelled the technology to new heights (Bahdanau et al., 2015; Wu et al., 2016). This shift has not only resulted in rapid improvements in translation quality, but it has also opened up opportunities for machine translation applications (Costa-jussà, 2022). Nowadays, machine translation plays a significant role in global communication, work, travel, education, information access, and other important domains.

Building on the previous discussions in the preceding chapters, it is worthwhile to delve deeper into the topic by focusing on the instruments used to measure particularly the quality of machine translation. One of the key areas of concern with machine translation tools, despite their many benefits, is the quality of the resulting translation output, which typically falls short in comparison to the quality achieved by professional human translators (Rivera-Trigueros, 2022). In order to improve the low quality of machine translation, it is required to evaluate its performance and make necessary improvements. For this reason, systematic machine translation performance tracking and measuring is crucial. Castilho (2016) lists more areas where evaluating machine translation systems is important: within the academic context, evaluation of machine translation serves as a tool to determine whether novel methodologies have resulted in quality enhancements; while in the business sector, evaluation plays a crucial role in demonstrating the quality of machine translation solutions; end-users of machine translation can also benefit from comparative assessments of translation quality when selecting an machine translation system. The increasing use of translation technologies, particularly machine translation, has led to the emergence of numerous metrics and standards of quality, many of which are implicit and differ in operational definition (Doherty, 2017). More to this, according to Kenny (2022), the quality of machine translation is heavily reliant on factors such as the specific translation software employed, the translation's context, and the end-users' expectations, all of which must be taken into account. This diversity of approaches and metrics makes it difficult for researchers to agree on how to accurately evaluate the quality of machine translation. However, among the many different ways to measure the quality of an machine translation system, the most widely accepted methods for machine translation evaluation are either through human evaluation or automated metrics. The subsequent sections will delve into the overview of the three main categories of machine translation evaluation and their distinctive elements.

1.3.1. Human Evaluation of Machine Translation Quality

Human expert evaluation is probably the most dependable method for assessing and quantifying the quality of machine translation and is commonly known as the benchmark or the golden standard (Läubli et al., 2020) of TQA. The assessment process involves bilingual professionals in translation and linguistics evaluating the translation output. It is not surprising that numerous measures have been developed by academics and industry professionals to assess the quality of machine translation based on the human expert evaluation.

The beginning of human assessment methods for evaluating the quality of machine translation can be attributed to the researcher John B. Carroll, with the first evaluations being conducted in approximately 1966. Carroll (1966) proposed a method for evaluating the quality of translations by comparing them to their source texts and assessing the degree of semantic and syntactic similarity. In his study, Carroll assessed the adequacy, accuracy, and overall quality of translations, among other factors that affect translation quality. Already then, Carroll observed that a key criterion for evaluating

a translation is its comprehensibility and readability, which should be comparable to the original language.

Later on, comparable methods of quality evaluation can be discovered in the academic literature, such as the White et al. (1994) framework. White et al. (1994) established a framework for machine translation evaluation that includes three essential criteria - adequacy, fluency, and comprehension - based on which the evaluators are directed to assign aggregate scores. These fundamental criteria were described by White as follows: fluency is the ease with which a translation can be read and understood; adequacy is the degree to which the translation accurately conveys the intended meaning of the original text; and comprehension is the concerned with how well the target audience can understand the translation (1994). The original stance of the White's study was to evaluate machine translation quality based on the potential productivity it offers to its users.

In machine translation evaluation, adequacy and fluency are commonly evaluated using Likert scales and the assessment is usually performed at the sentence or segment level, without considering the broader context, for example, adequacy evaluation requires some level of bilingual proficiency, while fluency can be assessed solely based on proficiency in the target language (Moorkens, 2018).

In addition to these primary metrics, secondary metrics such as readability, comprehensibility, usability, and acceptability are also commonly used to assess machine translation output, as pointed out by Castilho (2018). Subsequently the concepts of fluency and adequacy have been widely adopted and adapted in various machine translation evaluation studies.

Popović (2020) has recently presented another novel method for manual evaluation of machine translation, in which evaluators identify problematic translational segments (such as words, phrases, or sentences), rather than assigning scores or classifying errors. The method proposed by Popović does not solely rely on the quantification or classification of errors, but also allows for a more in-depth analysis of the annotated texts. During the evaluation, annotators differentiate between only two categories of issues: major and minor errors (2020), and generally the evaluation method mainly focuses on comprehensibility and adequacy as machine translation quality criteria.

Human manual evaluation has both advantages and disadvantages. One apparent benefit is that human evaluation can provide a more accurate and thorough assessment of translation quality, allowing for the identification of specific translation strengths and weaknesses. This is because human evaluators may consider factors like as context, idiomatic expressions, and cultural nuances that automatic measurements may struggle to capture. Human evaluation, on the other hand, has several drawbacks. For one, it can be a time-consuming and expensive procedure since it requires qualified evaluators and significant effort to evaluate large volumes of text. Furthermore, human evaluation is inherently subjective, and different evaluators may have different viewpoints on what makes a high-quality translation. However, despite these drawbacks, human evaluation remains an important method for measuring translation quality and improving machine translation systems.

1.3.2. Error Typologies

Another method of assessing the quality of machine translation is through error typologies and classification, where errors in the translation output are identified and categorized according to a predefined set of error types. Over time, scholars have introduced numerous error typologies or error classification methods to achieve the most accurate evaluation of translation quality. Examples of these include error typology proposed by Flanagan in 1994, the taxonomy developed by Correa in

2003, the categorization scheme presented by Vilar et al. in 2006, the error taxonomy proposed by Stymne et al. in 2012, and the error classification system called MQM (Multidimensional Quality Metrics) by Lommel et al. (2014) among others. The above classifications are based on the following principle: after identifying and categorizing errors in machine translation output using error typology, errors are typically counted and recorded to generate overall scores or metrics. These metrics can provide feedback to developers of machine translation systems, serving both to evaluate the performance of a single system or to compare different machine translation systems.
Typologies vary in the types of errors they identify and classify, as well as the criteria and methodology used for evaluation. Each typology has its own benefits as well as drawbacks, and it can be used in a variety of contexts and for a variety of purposes. Nonetheless, they are not without limitations. For instance, not all errors have the same weight or degree of impact on the overall meaning of the translation, and the significance of an error may vary depending on the context. Moreover, error taxonomies may overlook some types of errors or language-specific subtleties. However, this approach is considered to be a systematic way of assessing the quality of machine translation output. On the whole, error taxonomies are applicable to both manual and automatic assessment of machine translation quality. In the instance of manual evaluation, human evaluators employ error taxonomies to identify and classify errors in the machine translation text. For automated evaluation, error taxonomies can be applied to train machine learning models that automatically recognize and classify errors in the machine translation text, without human intervention.

1.3.3. Automatic Evaluation of Machine Translation Quality

Over the past couple of decades, the use of automated evaluation methods has gained widespread application among developers of machine translation systems due to their capacity to enable fast evaluation of machine translation systems and facilitate prompt implementation of enhancements to the system based on the evaluation results. Fully automated evaluation methods are based on comparing the result of machine translation with a reference translation, or in other words they evaluate the degree of similarity between reference and output translations. Thus, the employment of automatic measurements can considerably speed up the process of assessing the quality of machine translation. More to this, automatic machine translation evaluation measures are intended to be more objective than human evaluation or any alternative evaluation approach (Chauhan & Daniel, 2022). Another advantage of automated methods is their cost-effectiveness and ease of deployment, which makes them such a a popular tool for assessing the quality of machine translation. Subsequently, we will examine several of the most commonly used metrics.

The most distinguished metrics used in this field are WER (Su et al., 1992), BLEU (Papineni et al., 2002), NIST (Doddington, 2020), METEOR (Banerjee & Lavie, 2005), TER (Snover et al., 2006), ROUGE (Kilickaya et al., 2016).

- The Word Error Rate (WER) (Su et al., 1992) is a metric used for assessing the performance of machine translation system at the word-level. This method evaluates the quality of translated sentences by calculating the minimum number of word-level edits required to transform the machine translation output into the reference translation. WER is widely applied in machine translation research, particularly for evaluating speech-to-text translation systems (Chauhan & Daniel, 2022).
- Bilingual Evaluation Understudy (BLEU) (Papineni et al., 2002) metric operates by calculating the correspondences of n-grams in the machine translation output with respect to those in the reference translations, while also considering penalties for short sentences. To put it more precisely, this score measures how closely the machine translation output matches

the human translation, based on a comparison of contiguous sequences of words or so called n-grams in the two texts. Furthermore, it is worth mentioning that BLEU is the most widely used metric in the machine translation industry to assess translation quality.

- National Institute of Standards and Technology (NIST) (Doddington, 2002, March) evaluation metric shares some similarities with the BLEU metric.The NIST score is computed by taking the average of n-gram matches between machine translation and reference translations, and applying a reduced penalty for slight discrepancies in phrase lengths. NIST is one of the official metrics used in the annual Conference on Machine Translation (WMT), along with BLEU, METEOR, and TER, among others (Lavie, 2010).
- Metric for Evaluation of Translation with Explicit Ordering (METEOR) (Banerjee & Lavie, 2005) is an evaluation metric for machine translation that combines the unigram precision, recall, and alignment-based metrics to evaluate the similarity between the machine-generated output and the reference translations. The use of METEOR score is particularly prevalent in machine translation research and development as it allows for a holistic evaluation of the output quality, taking into account both fluency and adequacy of the generated translation, as stated in Banerjee & Lavie (2005) study.
- Translation Error Rate (TER) (Snover et al., 2006) measures the number of edits required to transform the machine translation translation into a human-generated reference translation. The score is calculated by dividing the total number of edits by the total number of words in the reference translation. The lower the TER score, the closer the machine-generated translation is to the human-generated reference translation. TER score is highly prevalent in machine translation evaluation due to its ability to account for various types of errors in translation,including insertions, deletions, and substitutions (Snover et al., 2006), providing a useful and informative measure of translation quality.
- Recall Oriented Understudy for the Gisting Evaluation (ROUGE) (Lin, 2004) score is a set of parameters used to compare the quality of machine-generated summaries with human-generated reference summaries. This score is a widely used metric in the field of automatic summarization evaluation, as it enables a comprehensive and objective assessment of the quality of machine translation output by taking into account the similarity between the generated summary and the human-generated reference summary at various levels of granularity.

Considering the previously mentioned information, there exist both advantages and disadvantages of using automatic methods for assessing the quality of machine translation that are noteworthy to mention. Automatic metrics offer several advantages in the evaluation of machine translation output, such as they provide an objective and repeatable method of evaluation; they are quick and efficient, enabling to assess vast amounts of machine translation output accurately; and they are cost-effective and language-independent, which makes them a valuable tool for assessing the quality of machine translation output in any language. Yet, there are certain downsides to using automatic measures to evaluate machine translation: they are limited in scope and may overlook issues related to grammar, syntax, or context. Further to this, there is a lack of consistency among different automatic metrics, which can make it difficult to compare results. Finally, automatic metrics require reference translations, which can be a limitation in cases where there is no reference translation available or the reference translation is of poor quality.

Generally, evaluating machine translation system quality is a complex and multifaceted matter as there is no singular correct answer regarding what constitutes a good translation. Even a single phrase or sentence can have multiple acceptable versions, making it difficult to determine the best

translation. However, automatic metrics are a helpful instrument to evaluate the quality of machine translation output; Alternatively, they could be employed in tandem with human evaluation and other approaches to give a more comprehensive and precise representation of machine translation quality.

1.3.4. Machine Translation Evaluation for Low Resource Languages: The Lithuanian Language Case

The subject of machine translation quality is more extensive than it may appear, and it encompasses various dimensions. Another matter that scholars frequently bring up is the quality of machine translation for languages with low resources. Low-resource languages refer to languages that lack sufficient linguistic resources, such as parallel corpora, lexicons, and grammar, which are essential for training and optimizing machine translation systems. In the preceding decades, most of the advancements in machine translation have been focused on high-resource languages, that is, languages with a large amount of training data (Costa-jussà et al., 2022), for instance English, French, German, or Russian. Consequently, individuals communicating in these languages are expected to benefit more from the evolution of machine translation technology than those speaking low resource languages such as Lithuanian, Basque, Maltese, or Bengali, which have limited training data (Costa-jussà et al., 2022). The primary concern is that the quality of machine translation for low resource languages has not yet reached the level of proficiency achieved for high resource languages. Lithuanian is considered a low resource language, and despite recent acceleration in the improvement of translation quality, it still falls below an acceptable standard. Several Lithuanian scholars have examined the matter of machine translation quality and the reality that the quality of machine translation in Lithuanian remains inadequate. Their research will be discussed in the following paragraphs.

The significance of machine translation efficacy for Lithuanian language has been the subject of various scholarly discussions. Among the early contributions is the study by Labutis (2005), who highlighted the importance of establishing a strong linguistic foundation for machine translation systems to produce acceptable quality of translation output. Following that, Daudaravičius (2006) tackled the challenging issue of machine translation quality assessment criteria and briefly described the distinct characteristics of machine translation systems. Rimkutė et al. (2008) and Cvilikaitė (2008) also reviewed the relevance of machine translation quality assessment and delved into the prevalent errors in machine translation systems and linguistic peculiarities of the output of machine translation. The research carried out by these researchers undoubtedly serve as a firm foundation for additional discussions among professionals and academics in the field. Daubarienė and Ziezytė (2013) subsequently conducted a study that involved a comparative analysis of two machine translation systems and aimed to review the current state of machine translation quality. The findings indicated that the translation quality from English to Lithuanian was of poor quality, with semantic issues being the primary cause, leaving it essentially unacceptable from the reader's standpoint. In a separate research, Ralys (2017) provided an overview of the development and advancement of machine translation tools, including a review of one of the newest neural network-based machine translation systems. The primary focus of the study was on evaluating the quality of machine translation using the BLEU score metric and exploring potential perspectives for the quality improvement. Petkevičiūtė and Tamulynas (2011) completed a significant study that focused on assessing the quality of English-Lithuanian machine translation. At that point, given the widespread use of machine translation systems, there was a growing need to assess the quality of such systems in a systematic manner and to develop an error classification framework tailored to the Lithuanian language. In their research, Petkevičiūtė and Tamulynas (2011) successfully achieved this objective by creating a comprehensive

error typology that encompassed three primary categories of errors: morphological, lexical, and systematic errors. The error classification developed by the authors not only effectively identified errors in their research but has also been adopted in subsequent studies by other researchers. To the best of our knowledge, no other classification have been developed exclusively for Lithuanian language. Recently, a few of research studies on the topic of machine translation quality for Lithuanian language have been published, including works by Stankevičiūtė et al. (2017), Kasperavičienė et al. (2020), and Povilaitienė and Kasperė (2022). These authors have concluded that the quality of machine translation systems for the Lithuanian language remains inadequate. Stankevičiūtė et al. (2017) analyzed machine translation output and indicated syntax and lexical errors, demonstrating the continued challenge of achieving acceptable machine translation quality for low-resource languages such as Lithuanian. Kasperavičienė et al. (2020) used an eye-tracking method to investigate the cognitive effort required to process different types of errors in machine translation. Meanwhile, Povilaitienė and Kasperė (2022) evaluated the quality of popular machine translation tools using both automatic BLEU score and MQM annotation methods, and also concluded that the issue of machine translation quality in low-resource languages remains unresolved, despite technological progress.

Based on a literature review of machine translation quality evaluation for Lithuanian language, it is noteworthy that the quality of translations in low-resource languages like Lithuanian is not yet at a satisfactory level compared to high-resource languages. Furthermore, an important point to highlight is that evaluating the quality of machine translation for low-resource languages is challenging due to the lack of suitable evaluation metrics and benchmarks. Therefore, the development of effective machine translation evaluation methods for low-resource languages is still an active research area in the machine translation field.

1.3.5. BLEU Evaluation Metrics

BLEU, meaning Bilingual Evaluation Understudy, is a automatic metric used to evaluate the efficacy of machine-translated text between languages. Moreover, even after two decades since its introduction by Papineni et al. (2002), BLEU metric is still considered as the benchmark for evaluating the quality of machine translation, moreover, it is the most frequently used metric in the industry. The reason why BLEU became popular was because it was one of the first metrics to demonstrate a significant correlation with human assessments of machine translation quality.

This metric based on the comparison of n-gram sentence units, often words, between the output generated by the machine translation system and one or more reference sentences (Kocmi et al., 2022). More specifically, translated sentences are evaluated by comparing them to a set of high quality reference translations. The frequency with which a word appears in the reference sentences is also counted (Chauhan & Daniel, 2022). Following that, the evaluations are aggregated throughout the corpus to generate an estimate of the translation's overall quality. While quality is assessed by the degree of similarity between the output generated by a machine translation and the output generated by a human. BLEU score uses a scale between 0 and 1, with the aim of measuring both its adequacy and fluency. For example, when the test segments score near to 1, it implies a higher level of agreement between the system's translations and the reference translations made by humans. When evaluating machine translation, the BLEU score is limited to the source sentences and the translations that are included in the test, and do not necessarily reflect the correctness of all possible translations. This limitation can result in a good translation being scored poorly, as the selected translation for each segment may not be the only valid option.

Notwithstanding the prevalence of this measurement, in their recent publication, Chauhan and Daniel (2022) identify the primary advantages and drawbacks of the BLEU score. BLEU metrics have several key advantages, including ease of implementation, simplicity and language neutrality. It is worth noting that these metrics can efficiently scale to large amounts of data and can quickly evaluate the quality of machine translation based on predefined criteria. Another significant advantage is that it is widely accepted as a metric of machine translation quality within the translation industry, thus making it simple to evaluate the quality of different systems. Conversely, BLEU drawbacks consist of the subsequent aspects: it is ineffective for morphologically complex languages, also fails to account for sentence structure and word order, and ignores the language's syntactical structure. (Dorr et al., 2011) further emphasise that BLEU scores are not considered accurate in evaluating the machine translation quality of individual sentences, indicating the rather limited applicability of automatic metrics. Therefore, these limitations contribute to a weak correlation between the BLEU metric and human judgement.

The application of BLEU metrics is not confined to translation, but extends to tasks like text summarization, where it assesses the quality of machine-generated summaries compared to those written by humans. Furthermore, it can also evaluate the quality of automatically generated image captions by comparing them to captions generated by humans, and the quality of automatic speech recognition transcriptions by comparing them to transcriptions written by humans.

Once again, the adoption of the BLEU metric has contributed to the rapid advancements of machine translation research, and it has established itself as the default automated evaluation metric within the field (Dorr et al., 2011). The BLEU metric serves as a benchmark for evaluating new metrics, and despite its limitations, it continues to be the prevailing automated metric for evaluating the quality of machine translation systems, primarily due to its simplicity and ease of use.

1.3.6. MQM Evaluation Framework

Based on the literature reviewed in this paper, it is apparent that various assessment methods or tools can be employed to evaluate machine translation quality. However, due to the high cost of human machine translation evaluation and the shallow nature of automated metrics, there was a growing demand for a more comprehensive and unified evaluation framework that addresses specific translation errors and issues (Lommel, 2018). Lommel et al. (2014) have addressed the limitations of previous machine translation quality assessment methods by developing a systematic and detailed evaluation framework known as Multidimensional Quality Metrics (MQM). This framework was created during the QTLaunchPad project, which was financed by the European Union, and was further developed as part of the QT21 project (Burchardt et al., 2021). MQM addresses the problem of translation quality evaluation's subjectivity and aims to provide a more comprehensive and accurate assessment of translation quality compared to commonly used metrics like BLEU, which focus on a single metric such as n-gram overlap. According to Lommel et al. (2014), the MQM framework is suitable for both quick evaluation and comprehensive assessment that aim to identify particular issues in a given text. Furthermore, users can customize the metric to their own objectives, allowing MQM to be applied in a variety of quality evaluation tasks, including both human and automatic assessment projects. The proposed method evaluates machine translation output using multiple dimensions or criteria that cover various aspects of translation quality, such as accuracy, fluency, terminology, or style. These criteria are language-neutral and applicable to any type of text or project. The advantage of this model is that evaluators can use MQM to rate the quality of translations according to each criterion, providing a more detailed and objective assessment of

translation quality, this provides an unbiased machine translation quality rating that can be used to determine whether a translation should be accepted, revised, or rejected (Burchardt et al., 2021).

It is noteworthy that MQM is subject to continuous improvement by its developers, and its most recent version 2.0 is presented here https://themqm.org. The website provides not only a history of the framework's development and explanation of error-related issues, but also detailed instructions on how to use MQM for evaluation purposes. Also, an earlier version of MQM, version 1.0, is accessible on this page here (Lommel et al., 2015a), however, the creators recommend the adoption of the most recent version. In the following paragraps we will further discuss the dimensions of the latest version 2.0 and how they are employed in the machine translation evaluation process.

In the translation assesment process, evaluators start by selecting a set of source texts and their corresponding translations. The translations are then evaluated using the MQM framework, with the translation quality rated for each dimension and sub-dimension. The latest MQM framework has a total of 7 core level error type dimensions, each of which consists of several sub-dimensions. These dimensions and sub-dimensions cover various aspects of translation quality, such as terminology, accuracy, linguistic conventions (fluency in previous version), style, locale conventions, audience appropriateness (verity in previous version), design and markup, and custom error types. The MQM core typology error types are explicated as follows[2]:

- Terminology errors occur either when a term employed in the source text differs from the normative standards of the domain or the organization, or when the term utilized in the target text does not correspond to the correct, normative equivalent found in the source text. Terminology also includes the following sub-dimensions: inconsistent with terminology resource, inconsistent use of terminology, wrong term.
- Accuracy errors arise when there is a discrepancy between the intended meaning of the source text and the message conveyed in the target text, brought about by altering, omitting, or adding information. The following sub-dimensions make up accuracy as well: mistranslation, over-translation, under-translation, addition, omission, do not translate, untranslated.
- Linguistic conventions (formerly known as fluency) pertain to errors concerning the grammatical correctness of the text, encompassing difficulties with grammar, punctuation, and overall mechanical accuracy. These sub-dimensions are also included in linguistic conventions: grammar, punctuation, spelling, unintelligible, encoding of the character.
- Style – these errors involve a discrepancy from organizational style guidelines or an unsuitable language style, despite the grammatical correctness of the text. Also, style has these sub-dimensions: organizational style, third-party style, inconsistent with external reference, register, awkwardness, unidiomatic style, inconsistency (Lommel, 2015a).
- Locale conventions errors arise when the translated text fails to meet to locale-specific content or formatting criteria. These sub-dimensions are also part of locale conventions dimension: number format, currency format, measurement format, time format, date format, address format, telephone format, short key (Lommel, 2015a).
- Audience appropriateness errors are caused by the use of inappropriate content in the translation output for the target location or target audience. This dimension also includes the sub-dimension culture-specific reference.
- Design and markup errors refer to issues associated with the physical layout of a translation output, such as the formatting and markup of characters, paragraphs, and user interface

[2] The comprehensive list of MQM dimension issues and the corresponding error types explained in this study can be found at https://themqm.org.

elements, among others. The following are the sub-dimensions: formatting, layout, markup, truncation/text expansion, missing text, link or cross reference (Lommel, 2015a).

Although having created a comprehensive framework with over 100 error types and their detailed explanations, the MQM developers recommend against relying solely on this catalogue, instead emphasize the customization of the framework to suit to unique assessment requirements (Lommel, 2018). In this way, the translation is evaluated at a more granular level, allowing for a more in-depth understanding of the output quality.

Furthermore, the framework allows translation quality evaluators to give numerical values to identified translation errors and to assign their severity levels. According to Lommel (2018) severity is defined by the characteristics of the error and its impact on the translation's efficiency. The greater the severity of an error, the more likely it will have a negative impact on the user experience. Independent of the error category, severity levels are allocated as Major, Minor, and Neutral (Lommel et al., 2015b). The authors of MQM define the default severity levels as follows:

- None: 0. This is for issues that require further attention but should not impact the evaluation of the overall translation. It serves as a flag for attention without imposing a penalty and can be used for minor and systematic errors that can be easily fixed. As no penalty is assigned for this level, it is not taken into account in the scoring.
- Minor: 1. These are errors that do not affect the usability or comprehension of the content. For example, errors like punctuation, can be considered minor as long as they can be easily corrected by the typical reader without affecting the overall functionality of the content. Resolving these errors is optional for content creators since they do not significantly impact the content's usability.
- Major: 10. Major issues in content refer to errors that affect its usability or comprehensibility, but do not render it completely unusable. For instance, a misspelled word can cause confusion and extra effort to comprehend, but does not make the content unusable. If the reader cannot correct the error and the intended meaning is unclear, it should be classified as a major issue.
- Critical: 100. Critical issues in content refer to errors that make it unsuitable for use, such as a severe grammatical error that alters the text's meaning. An error must be classified as critical if it prevents the reader from using the content as intended or provides false information that might be harmful.

Once the errors have been aggregated and assigned severity levels, the scores for each dimension and sub-dimension are summed up to give an overall translation score. The score can be used to determine whether the translation is of good quality, requires only minor post-editing, or is of unacceptable quality.

This approach is built upon the most effective practices in translation quality assessment, making it a methodical and cohesive framework to evaluate the quality of machine translation (Lommel, 2018). Accordingly, it is unsurprising that many researchers in the translation field have successfully used MQM due to its multiple advantages (Klubička et al., 2017; Freitag et al., 2022; Rei et al., 2021). The MQM evaluation process is not only highly flexible but also can be adapted to different translation projects and scenarios. It can also be used with other metrics and evaluation methodologies to provide a more comprehensive evaluation of translation quality. As a result, an evaluation of the framework's characteristics suggests that its flexibility, granularity, and versatility make it one of the finest methods to evaluate the translation quality of a morphologically rich language like Lithuanian.

1.3.7. Human Acceptability and Methods to Measure it

After considering the various ideas expressed about TQA evaluation methods, a central point that emerges in machine translation paradigm is the degree to which the translation output matches the original message. As described earlier, we have outlined the perspectives regarding the assessment of translation quality by industry professionals and machine translation system developers. However, an equally important aspect is understanding the process of machine translation quality assessment from the perspective of end-user, who is the primary intended audience for machine translation product. The concept of acceptability has been a topic of discussion among scholars in the field for more than four decades, therefore there is a substantial body of scientific findings related to this subject. Subsequently, we will review the scientific literature on the notion of acceptability and the methodologies used to measure it.

Acceptability in translation studies refers to the degree to which a translated text is considered linguistically and stylistically adequate for its intended purpose and audience. In his work, Castilho and O'Brien (2017) elaborates on this concept by defining it as the degree of acceptance by the end-user of language, a text, or a product. Acceptability, together with fluency and adequacy, is one of the most fundamental dimensions of machine translation quality. While fluency refers to the degree to which the translated text reads fluently and has no grammatical errors, and adequacy refers to the degree to which the translation accurately conveys the meaning of the source text (Castilho et al., 2018), acceptability focuses on whether the translation is acceptable to the target audience in terms of textual norms. For example, if machine translation is intended for use in a professional medical context, it should use appropriate terminology, appropriate register and avoid the use of jargon. Similarly, if the translation is intended for a general audience, it should be clear, concise and easy to understand.

According to Beaugrande and Dressler (1981), the term of "acceptability" in communication refers to the recipient's attitude towards a message. Specifically, the recipient needs to perceive the structure of the text as a coherent and cohesive that serves a functional purpose. The authors explain that acceptability can also have a level of tolerance for minor errors as long as they do not impact overall text comprehension. Beaugrande and Dressler (1981) provide further elaboration on the concept. They argue that a number of linguistic and non-linguistic elements, such as the recipient's background knowledge, cultural norms, and social norms, influence acceptability. Furthermore, the authors make the interesting point that acceptability can function at various levels of linguistic analysis, including individual words, phrases, and sentences, as well as the general structure and coherence of the text. The authors also emphasize that acceptability is not an absolute construct that varies based on the recipient, with possible differences in the degree of acceptance towards the same text. Overall, Beaugrande and Dressler's concept of acceptability encompasses multiple factors that aid in comprehending how recipients evaluates and use texts in communication.

Chomsky (2014) also proposed a definition for acceptability, in which the term refers to utterances that are effortlessly understandable and natural, without requiring analytical examination. He emphasizes that in a given text, the sentences that are perceived as more acceptable are characterized by greater naturalness and comprehensibility. Chomsky notes that acceptability of sentences is an essential aspect of evaluation, as it reflects their practical suitability for communication. He proposes that acceptability is separate concept from grammaticality, which is an internal property of sentences, whereas acceptability is based on efficiency and reflects how easily a sentence can be understood in context. He also argues that acceptability is closely related to other aspects of language, such as the

overall coherence of the text, and communicative effectiveness, which reflects how well the text conveys the intended message to the target audience. This implies that Chomsky's ideas are significant because they highlight the fundamental elements that are crucial to the end-user of the text, including its level of comprehension, clarity of the central idea, and overall readability.

Van Slype conducted a comprehensive examination of methods used to evaluate machine translation and categorized evaluation criteria into two main groups: macro-evaluation and micro-evaluation (1979). Among the other macro-level evaluation methods, Van Slype proposed acceptability as the cognitive level assessment method. In his study Van Slype defined "acceptability" as the degree to which an end-user finds a translation acceptable. Furthermore, he claimed that surveying end-users is the most effective means of measuring acceptability. Although Van Slype's ideas were proposed more than 40 years ago, they have been used by numerous researchers and practitioners in the field, including Hutchins (2014); Läubli et al. (2018); Castilho et al. (2018); Kasperavičienė et al. (2020) among others.

Coughlin conducted a thorough investigation in 2003 to look at the correlation between automated measurements and human evaluations of machine translation quality. In order to assess the acceptance of machine translation, he used a scale of 1 to 4, with raters having to select the level of acceptability. The levels of evaluation suggested in the Coughlin (2003) scale were as follows:

- 4 = denotes an ideal translation that may not be flawless but is grammatically correct and conveys all the information with precision.
- 3 = is considered acceptable, though not perfect in terms of grammar and style. Nevertheless, it is comprehensible and conveys all the essential information accurately.
- 2 = is possibly acceptable, which means it may be understandable with sufficient context or time for interpretation. In this case, only some information is conveyed accurately.
- 1 = is considered unacceptable, because it is completely incomprehensible and gives little or no accurate information.

Coughlin's proposed the evaluation approach that is simple to apply, and evaluators are able to assess when the criteria are brief and clear.

Roturier (2006) conducted a study aimed at examining the effectiveness of implementing controlled language rules, and assessed their impact on the comprehensibility, usefulness, and acceptability of such documents. In this study, the term "acceptability" adopted by Roturier is based on the standard of textuality, which was initially introduced by Beaugrande and Dressier (1981). According to Roturier, this textuality standard relates to the reader's perception that the content they are reading is both intelligible and relevant to them. In other words, the text should be cohesive and coherent, and should serve a purpose for the reader. Furthermore, author notes that acceptability is not just about whether a text is relevant to the reader, but also how the reader perceives its textual features. This includes whether a reader finds the language and style of the text acceptable, or if he rejects it for some reason. Roturier's study specifically focused on the end-users, and evaluated acceptability through a survey in the form of a customer satisfaction questionnaire. The author reasoned that using a survey would assist reveal how effective the machine-translated document was to the user.

In his study on the acceptability of computer systems, Nielsen (1993) argues that acceptability is an important component of usability, as it encompasses not only the system's effectiveness but also the user's satisfaction and enjoyment. Nielsen suggests that acceptability can be assessed through various methods, such as user feedback, surveys, and observational studies. Later, Castilho initiated research (2016), using Nielsen's acceptability model, which also aligns with the De Beaugrande and Dressler's

(1981) notion of acceptability. Castilho's study focused on how the relevance of the text to the reader influences its acceptability. The research measured the acceptability of machine-translated content based on evaluation criteria of usability, satisfaction, and quality. In doing so, the study aimed to bridge the gap between the concepts of acceptability and usability. Castilho's overall argument is that a user is more likely to consider machine-translated content acceptable if it can still be used to complete a task despite its imperfections. On the other hand, if the translation is so poor that it cannot be used to accomplish the task, the user will find the content unacceptable. It is reasonable to support the author's viewpoint because, from a human perspective, a translation must meet to the end-user's requirements, and the standard for quality can vary depending on factors such as language knowledge or education. Additionally, it should be noted that even if a linguist finds a translation output unacceptable, it may still fulfill the needs of an ordinary user. Castilho's research has made a major contribution to existing research on analyzing the acceptability of machine translation from the user's perspective.

Kasperavičienė et al. (2020) conducted an experiment that was built upon Castilho's (2016) ideas, and the authors also believed that a text can be considered acceptable if it meets the requirements of a reader. The main idea of the study was to identify machine translation errors that typically cause problems for readers. During their experiment, Kasperavičienė et al. (2020) used an eye tracking methodology and a reader survey, which was suggested by Van Slype (1979). Their research evaluated the acceptability of the text by taking into account the criteria of satisfaction, usability, and quality. Undoubtedly, evaluating the acceptability of machine translation from an end-user perspective is important for obtaining a comprehensive understanding of translation errors. This approach provides insights not only from the experts' perspective but also from the viewpoint of end-users. Kasperavičienė et al. made a contribution to the field of machine translation research by analyzing errors and assessing quality acceptability from the perspective of the user, which also aligns with previous studies (2020).

Another study by Kasperė et al. (2021) also looked at the quality of machine translation and how it is perceived by the end-user. The authors discovery that quality is correlated to the education level is noteworthy. According to the findings of the study, respondents with lower education levels were more satisfied with machine translation than those with higher education levels. The results of the study indicate that people with higher levels of education tend to have better language skills and, therefore, expect higher quality translations. Consequently, if machine translations contain grammatical errors or lack of coherence, they are more likely to receive lower quality scores. This research is significant as it examines the perception and usability of machine translation by end-users. It is noteworthy that there are not many studies that focus on end-users' perspective.

As revealed by a brief review of the literature, many authors have explored the concept of "acceptability" in translation studies. These discussions have highlighted the fact that acceptability is a subjective and complex notion, which is difficult to measure. Typically, in research studies, individuals with expertise in translation, such as professional translators, editors, linguists, or students studying these fields, are the ones who evaluate the acceptability of a translated text (Castilho & Guerberof Arenas, 2018). This approach to quality, unfortunately, then represents only one part of the evaluation of the translation, and the opinion of the end-user is not taken into account. However, it should be mentioned that evaluation of the translation quality involves many subtleties, including language proficiency, education, cultural background, and other factors, that may lead to one person finding it acceptable while another may find it completely unacceptable. Though, the literature review shows several approaches on evaluating the quality of translation. Nonetheless, researchers

emphasize the significance of determining how the translation is perceived by the end-user who is the final recipient of the translation systems product. Evaluating and measuring the acceptability of machine translation from the end-user's perspective is equally important because the quality of a product is also determined by how well it meets their needs. In this study, acceptability is also treated as a concept of the extent to which a translation text is acceptable to the end-user. The measure used is a survey based on the Likert scale, as recommended by a number of the above-mentioned authors.

Taking everything into account, the literature overview highlights the prevalence of questionable content on the internet, particularly in the form of machine-translated spam websites. These websites not only contain potentially inaccurate or false information, but the use of machine translation can also compromise language quality, message transmission, and user experience. Addressing this issue requires evaluating the quality of machine-translated websites and understanding how users perceive them. The literature suggests that assessing machine translation output is a relevant topic, with researchers exploring various approaches including automated methods and human judgment. Additionally, considering the end-user's evaluation is important since they are the ultimate consumers of the translation product. Given the subjective nature of translation quality, combining multiple assessment methods can provide a more comprehensive evaluation while mitigating subjective biases. The subsequent empirical part of this study will employ three machine translation evaluation methods, which will be further discussed in the following section.

2. A User-Centered Qualitative Analysis of Machine-Translated Websites

The empirical section of this paper examines the quality of machine-translated websites, employing established industry-standard methods for assessing translation quality, and integrating end-user evaluation. In this chapter, the methods for collecting and analyzing data will be presented and report on the findings of the assessment of machine translation quality, which was carried out using both automatic and human translation quality assessment approaches.

2.1. Methodology

This part presents the methodology employed in this study to evaluate the quality of machine-translated websites, with a focus on qualitative analysis and a user-centered perspective. The approach combines descriptive, qualitative descriptive, and comparative methods, which are elaborated on in the following paragraphs. The descriptive literature review was conducted to compile and analyze the most relevant research on the assessment of machine translation quality, and to serve as a foundation for selecting the most suitable research instruments to effectively address the study aim. The qualitative descriptive method was employed to analyze and describe the data obtained from automatic and human translation error analysis as well as an end-user survey. Additionally, the intention is to compare the employed methods with one another. This multi-faceted approach will not only provide insights into the actual quality of website translation, but will also offer a comprehensive understanding of the factors that impact the user experience of machine-translated websites and the extent to which machine translation errors can be tolerated by end-users.

This study focuses on the analysis of machine-translated spam websites, which can be easily discovered through targeted keyword searches on search engines. The chosen websites (see Appendix 1) adhered to Google's spam regulations, as they were incoherent, machine-translated, and disregarded for text quality or user experience. The URLs of these websites suggested that they provided little value and consisted primarily of content merged from other sources. To conduct a comprehensive analysis, five machine-translated websites were selected for the study, covering topics in mechanics, popular science, medicine, psychology, and chemistry. These topics were intentionally chosen to reflect the broad range of subjects that can be exploited by the creators of such websites. The selected websites were subject to a multifaceted error analysis and user survey as part of the research methodology. The research consists of three primary components: an assessment of the quality of machine translation on the selected websites utilizing automatic scoring; an evaluation of translation errors by a human expert; and a user survey aimed at determining the acceptability of these websites. The methodology employed for each of these components will be elaborated in detail below.

The initial stage of the research involved questionnaire, aimed at evaluating the quality and acceptability of machine-translated websites through a user-centered perspective. In order to conduct the survey, the LimeSurvey tool was employed. Participants were asked to rate the quality of five distinct machine-translated websites that were presented in the Lithuanian language. The selected website excerpts showed the overall view of the website, rather than just the text alone, as this enables the user to assess the overall quality of the website. By including multiple components of the website allows users to make a more informed judgment on the quality of the machine-translated website, leading to a more accurate assessment. The excerpts of websites contained texts ranging from 66 to 135 words, which was considered enough to detect mistakes in grammar or terminology, and also to get an overall impression of the website. To evaluate the websites, respondents were asked to rate each website according to five criteria, including the clarity and likability of the text, the

informativeness of the website, the credibility of the information, and the accuracy and quality of the text. The evaluation criteria were established using the website quality criteria introduced by Thielsch and Hirschfeld (2019). Each criterion was rated on a 5-point Likert scale, ranging from "Strongly Agree" to "Strongly Disagree." Two additional questions were included to understand the respondents' educational background and age, which may also influence user preferences on the websites evaluated. The survey was open to native speakers of Lithuanian from 1 01/01/2023 to 28/02/2023. A total of 99 completed questionnaires were collected during this period. To increase the number of survey participants, the questionnaire was promoted on both the website of the Kaunas University of Technology, Faculty of Social Sciences, Arts and Humanities and on the Faculty's Facebook page. The survey was anonymous, and respondents were informed about the conditions for data collection and privacy. The gathered data were processed using MS Excel and IBM SPSS Statistics 27 software. The practical part of the study will provide a detailed discussion of the obtained results.

The second part of the research aimed to assess the quality of website texts using automatic machine translation metrics. To achieve this objective, the research employed the BLEU score, which is the most commonly used and recognized automated metric in the field of machine translation evaluation. This score provides a quantitative evaluation of translation output by comparing it to reference human translation. The BLEU score is a numerical rating ranging from 0 to 100, which evaluates how closely the machine-translated text matches a reference translation. A score of 0 indicates that there is no similarity between the machine-translated output and the reference translation, resulting in low quality, whereas a score of 100 indicates a perfect match with the reference translations, resulting in high quality. For example, score above 60 to 80 is generally considered very good, while a score of 100 is extremely rare and would indicate that the machine translation is identical to a human translation word for word. To put it simply, a higher BLEU score usually suggests that the quality of the translation is better. For this part of the study, the Interactive BLEU score evaluator available at Tilde was used, which can be accessed at https://www.letsmt.eu/Bleu.aspx. This tool helps compare the quality of translations produced by one or more machine translation systems. It also enables comparison between machine and human translations and analysis of BLEU scores obtained from translation systems. To perform the evaluation, three components were necessary: the webpage text in English, the same text version in Lithuanian, and a reference translation done by a human. Three different texts were chosen, totaling 2450 words, for the evaluation process. The selection of texts aimed to include representative content commonly found on spam websites and provide a comprehensive basis for evaluating translation quality. The evaluation process consisted of the following steps: initially, inputting the original English text, machine-translated text in Lithuanian, and the human reference translation into the Interactive BLEU score evaluator. Subsequently, generating the comparison results between machine and human translations, and finally, analyzing these results. The results of the BLEU score evaluation are presented and discussed in the practical part of the study.

The third part of the study involved examining machine translation errors through human expert analysis. The errors were classified using the Multidimensional Quality Metrics (MQM) error typology proposed by Lommel et al. (2014). The MQM error typology was chosen for this study because it is considered one of the most recent and extensive error classification systems in the TQA field. Moreover, it is highly adaptable and inclusive of various error types, making it suitable for assessing the translation quality of morphologically complex languages like Lithuanian. To maintain consistency in the evaluation process, excerpts from the same source texts, totaling 1060 words, were used for both the automatic and human expert evaluations. The translations were assessed using the

following dimensions from the MQM error typology: terminology, accuracy, linguistic conventions, style, and locale conventions. The dimensions audience appropriateness and design and markup were excluded as they are not relevant for this study.

The following is a comprehensive list of the error categories and subcategories that were selected:

Terminology
- Inconsistent with terminology resource
- Inconsistent use of terminology
- Wrong term

Accuracy
- Mistranslation
- Over-translation
- Under-translation
- Addition
- Omission
- Do not translate
- Untranslated

Linguistic conventions
- Grammar
- Punctuation
- Spelling
- Unintelligible
- Character encoding

Style
- Organizational style
- Third-party style
- Inconsistent with external reference
- Register
- Awkward style
- Unidiomatic style
- Inconsistent style

Locale conventions
- Number format
- Currency format
- Measurement format
- Time format
- Date format
- Address format
- Telephone format
- Short key

In addition, errors were categorized according to their severity level. The MQM setting, commonly used in the translation community, was used in this study. Under this setting, the severity levels are neutral = 0, minor = 1, major = 5, and critical = 25. A maximum of 100 points is allowed to pass the quality assessment. The TAUS error typology evaluation template, available at https://info.taus.net/dqf-mqf-error-typology-template-download, was used to manually annotate errors in the translations. After identifying and annotating all errors, the data was systematised and examined. The results will be thoroughly discussed in the practical part of this research.

In the final step of the research, the results obtained from the machine and human translation error analysis and end-user survey are compared. This approach aims to identify patterns and correlations between translation error rates and user satisfaction with machine translation. It will also help to determine the extent to which machine translation errors are tolerated by the end-users and how these errors affect the overall user experience. By using descriptive, qualitative, and comparative methods, this study aims to provide a holistic understanding of the factors that influence the quality of machine-translated websites and how users perceive machine translation errors.

2.2. End-user Acceptability of Machine-Translated Websites

This section will present and thoroughly discuss the results of the end-user survey. During the survey, the participants assessed five distinct machine-translated websites by rating five quality statements on a 5-point Likert scale that ranged from strongly agree to strongly disagree. The statements are as follows (refer to the provided questionnaire in Appendix 2):

- **Statement No. 1.** The text of the website is clear. (lt – Pateiktos interneto svetainės tekstas yra aiškus.)
- **Statement No. 2.** The text is enjoyable to read. (lt – Tekstą yra malonu skaityti.)
- **Statement No. 3**. The website is informative. (lt – Pateiktos interneto svetainės puslapis yra informatyvus.)
- **Statement No. 4.** I trust the information on this website. (lt – Pasitikiu svetainėje pateikta informacija.)
- **Statement No. 5.** The text is accurate and of high quality. (lt – Tekstas yra taisyklingas, jo kokybė puiki.)

To analyze the data in this research, IBM SPSS was used to calculate frequency and percentage counts. Additionally, a cross-tabulation table was employed to illustrate the relationship between categorical variables.

The survey also gathered demographic information from respondents, including age, which is anticipated to contribute to the overall understanding. The pie chart (Fig. 2) presented displays the age distribution of survey respondents. Using this data, the following conclusions can be drawn.

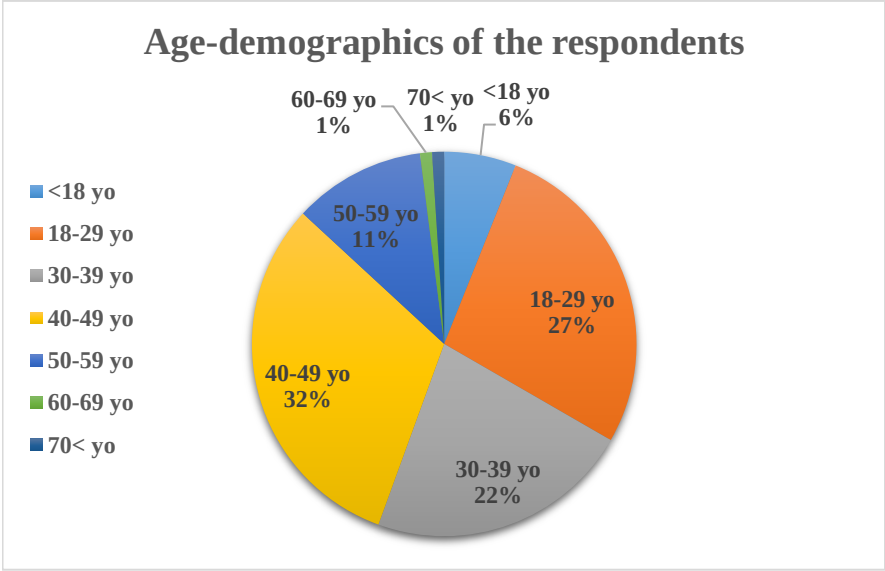


Fig. 2. Age-demographics of the respondents

The data on the age distribution of survey participants indicate a broad spectrum of age groups. This could offer valuable insights into how various age groups perceive the quality of machine-translated websites. The largest age group is the 40-49 years category, with 31 respondents (32%). This is followed by the 18-29 years and 30-39 years groups, with 27 respondents (27%) and 22 respondents (22%), respectively. The age groups with fewer respondents are the 50-59 years (11 respondents, 11%), < 18 years (6 respondents, 6%), 60-69 years (1 respondent, 1%), and 70 < years (1 respondent, 1%) categories. From the data presented, it is clear that that the age groups of 50-59 years, 60-69 years, and 70 years and older are underrepresented in the survey sample. Furthermore, this indicates that the results may not fully capture the opinions and experiences of these age groups, which could affect the interpretation and applicability of the survey findings. For instance, younger respondents might be more familiar with internet technologies and have different expectations for translation quality compared to older respondents.

The survey also obtained information on the educational background of respondents. The collected data is illustrated in the bar chart below (Fig. 3).

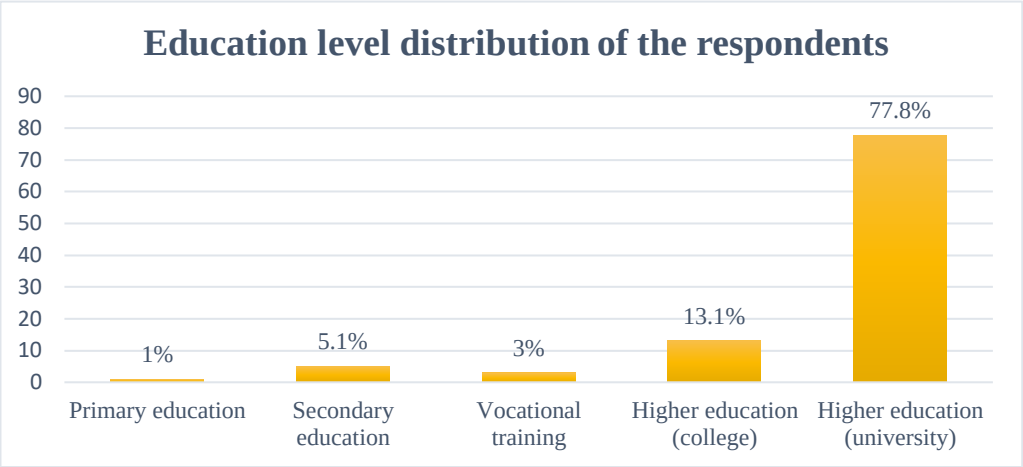


Fig. 3. Respondent education level distribution

The majority of the respondents (77.8%) have a university-level higher education. This implies that the survey sample is predominantly well-educated, with most respondents having at least a university degree. In contrast, there is a relatively small percentage of respondents with primary (1%), secondary (5.1%), and vocational (3%) education levels. The data also shows that 13.1% of respondents have a college-level higher education. As evident from the survey data, the majority of respondents have a higher education level, which may have an impact on the interpretation and generalisation of the survey results. This could mean that the opinions and perspectives gathered in the survey may not accurately reflect the broader population's views, as individuals with lower levels of education are underrepresented. The high proportion of respondents with higher education degrees might have an impact on the way they perceive and evaluate the quality of machine-translated websites. These respondents may have higher expectations for translation quality, greater linguistic knowledge, or more experience with professional translations, potentially influencing their evaluation of the machine-translated websites.

Table 1. Aggregated and separate statement counts

	Total		Statement No. 1		Statement No. 2		Statement No. 3		Statement No. 4		Statement No. 5	
	Responses		Responses		Responses		Responses		Responses		Responses	
	N	Percent	N	Percent	N	Percent	N	Percent	N	Percent	N	Percent
Strongly Agree	397	16.00%	95	19.20%	77	15.60%	83	16.80%	69	13.90%	73	14.70%
Agree	518	20.90%	124	25.10%	88	17.80%	112	22.60%	107	21.60%	87	17.60%
Neither Agree nor Disagree	664	26.80%	100	20.20%	123	24.80%	173	34.90%	148	29.90%	120	24.20%
Disagree	553	22.30%	118	23.80%	133	26.90%	78	15.80%	96	19.40%	128	25.90%
Strongly Disagree	343	13.90%	58	11.70%	74	14.90%	49	9.90%	75	15.20%	87	17.60%
Total	2475	100.00%	495	100.00%	495	100.00%	495	100.00%	495	100.00%	495	100.00%

The survey requested respondents to evaluate website quality using a 5-point Likert scale that ranged from strongly agree to strongly disagree. Table 1 presents the condensed data gathered from the survey, which will be discussed in further detail. This will be done by examining both the overall aggregated data and an analysis of seperate statements. Based on the aggregated data provided for the five websites (Fig. 3) and their machine translation quality, here are the findings.

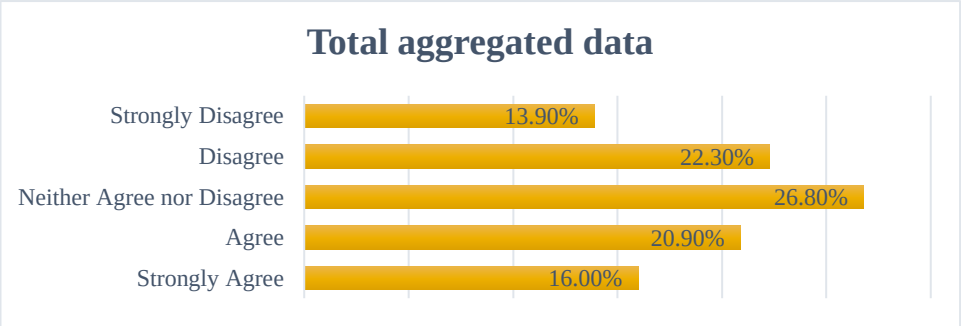


Fig. 4. Total aggregated data

A small proportion of respondents found the machine-translated websites to be of high quality, with 16.00% strongly agreeing that the text quality was good. Additionally, 20.90% agreed that the text quality was good, which indicates that 36.9% of the respondents had a positive perception of the machine translation quality. The largest group of respondents (26.80%) neither agreed nor disagreed with the quality of the translations, suggesting that they had a neutral perception. This might suggest that, that the translations were of average quality or that the respondents were unsure about the quality. A considerable proportion of respondents had a negative perception of the machine-translated websites' text quality, with 22.30% disagreeing and 13.90% strongly disagreeing. This means that 36.2% of respondents found the translations to be of poor or very poor quality. It's worth noting that the positive and negative perceptions are almost equal, with 36.9% having a positive perception and 36.2% having a negative perception.

Following this, it is important to explore how the survey respondents' answers were distributed among the different quality statements.

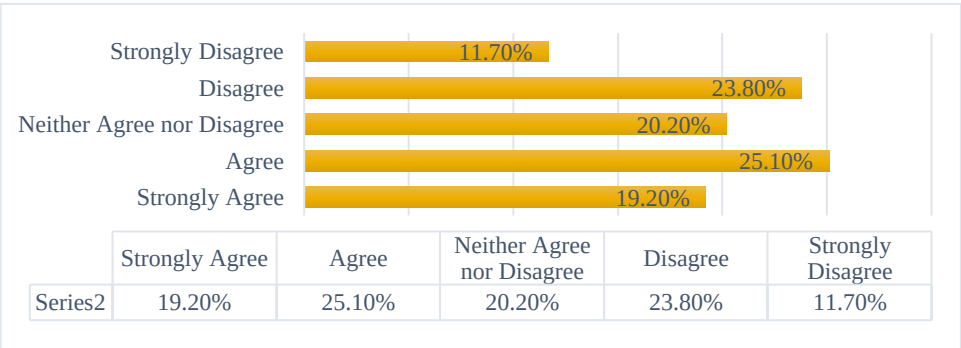


Fig. 5. The breakdown of responses for statement No. 1: The text of the website is clear

Regarding the first statement concerning text clarity, Figure 5 presents the distribution of participants' responses. A notable proportion of respondents found the machine-translated website text to be clear, with 19.20% strongly agreeing and 25.10% agreeing. This may be an indication, that 44.3% of the respondents had a positive perception of the text clarity. A considerable group of respondents 20.20% neither agreed nor disagreed with the clarity of the translations, indicating a neutral perception. This could be a sign that, the translations had average clarity or that the respondents were uncertain about the clarity. A significant proportion of respondents had a negative perception of the text clarity, with 23.80% disagreeing and 11.70% strongly disagreeing. This means that 35.5% of respondents found the text to be unclear or very unclear. The survey results show that the positive and negative

perceptions of text clarity are relatively balanced, with 44.3% having a positive perception and 35.5% having a negative perception. The neutral group might indicate that the translations have average clarity or that respondents are unsure about the clarity, which could warrant further investigation.

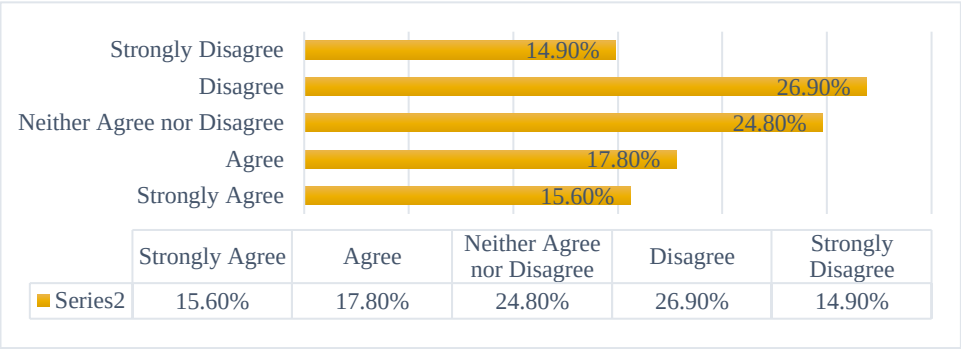


Fig. 6. The breakdown of responses for statement No. 2: The text is enjoyable to read

In the analysis of the second statement, "The text is enjoyable to read" (Fig. 6), a diverse response pattern can be noticed. A minority of respondents found the machine-translated website text to be enjoyable, with 15.6% strongly agreeing and 17.8% agreeing, for a total of 33.4% who considered the content was enjoyable. Meanwhile, 24.8% held a neutral position, neither agreeing nor disagreeing, signaling that they had a neutral stance towards the translations. A considerable number of respondents, 26.9% and 14.9%, disagreed and strongly disagreed, respectively, amounting to 41.8% with negative perceptions of the text's enjoyability. This implies that 41.8% of the participants perceived the text as either not enjoyable or highly unenjoyable to read. The data reveals a higher percentage of negative 41.8% than positive 33.4% perceptions regarding text enjoyability.

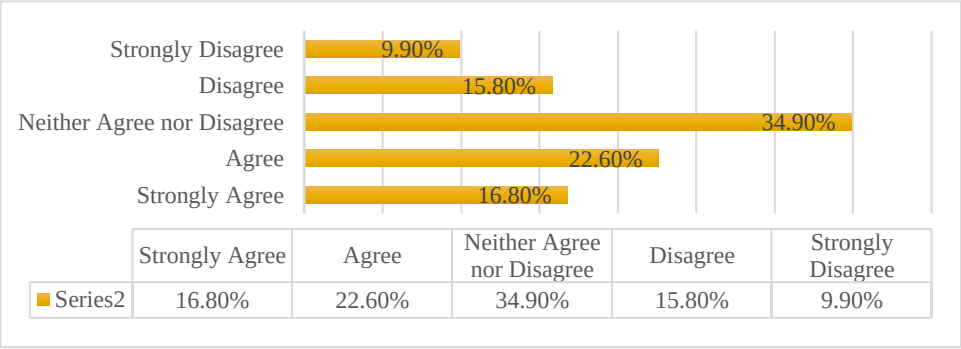


Fig. 7. The breakdown of responses for statement No. 3: The website is informative

Survey respondents provided ratings for the third statement, "The website is informative" (Fig. 7). A significant percentage of respondents found the machine-translated websites to be informative, with 16.8% strongly agreeing and 22.6% agreeing, totaling 39.4% who held a positive view of the website's informativeness. Meanwhile, 34.9% of respondents held a neutral position, neither agreeing nor disagreeing with the website's informativeness. A sizable group of respondents (15.8% disagreeing and 9.9% strongly disagreeing) had a negative impression of the website's informativeness, accounting for 25.7% of respondents. This means that 25.7% of respondents considered the website to be either uninformative or very uninformative. The survey outcomes reveal that positive views on website informativeness exceeded negative ones, with 39.4% holding a positive opinion and 25.7% a negative one.

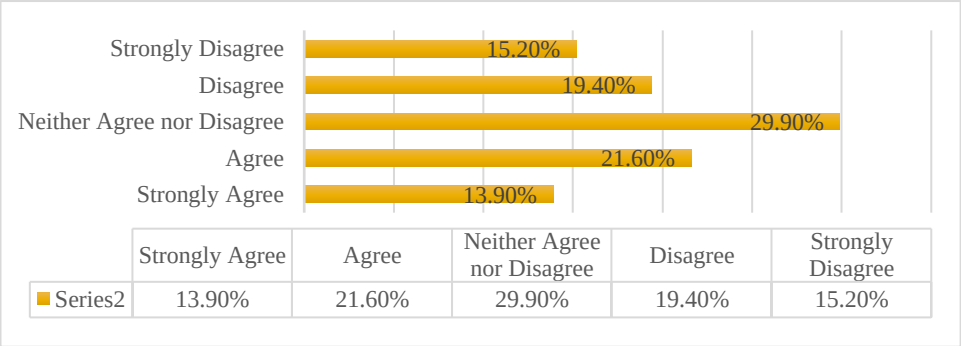


Fig. 8. The breakdown of responses for statement No. 4: I trust the information on this website

Figure 8 presents the distribution of responses for the fourth statement "I trust the information on this website". A portion of respondents trusted the content on machine-translated webpages, with 13.9% strongly agreeing and 21.6% agreeing, for a total of 35.5% respondents assuming that the content was trustworthy. Moreover, 29.9% of respondents maintained a neutral stance, neither agreeing nor disagreeing with the information's trustworthiness. A sizable proportion of participants 34.6% had a negative view of the information's trustworthiness (19.4% disagreed and 15.2% strongly disagreed). The survey findings reveal that positive and negative perceptions of trustworthiness are relatively evenly distributed, with 35.5% expressing a positive view and 34.6% expressing a negative one.

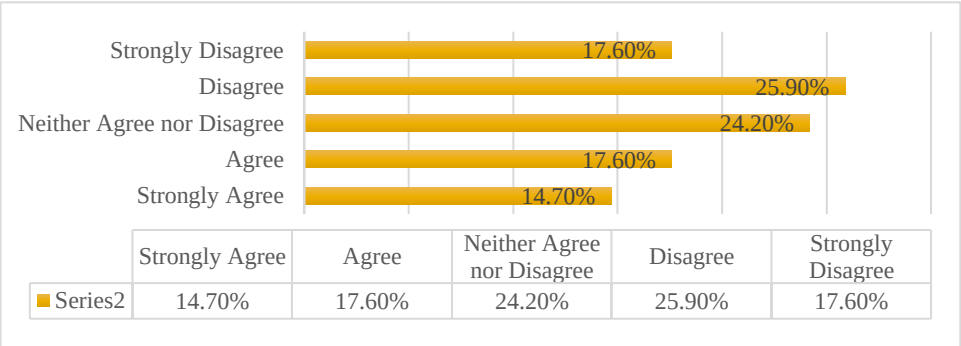


Fig. 9. The breakdown of responses for statement No. 5: The text is accurate and of high quality

The distribution of responses for the fifth and final statement "The text is accurate and of high quality" is shown in Figure 9. The machine-translated website text appeared accurate and of high-quality for 32.3% of individuals, with 14.7% strongly agreeing and 17.6% agreeing. Furthermore, 24.2% of respondents indicated a neutral opinion, neither agreeing nor disagreeing with the accuracy and quality of the translations. The text's accuracy and quality were evaluated negatively by a significant number of respondents (25.9% disagreeing and 17.6% strongly disagreeing), accounting altogether 43.5% of respondents. According to the results, a greater proportion of respondents held negative views regarding the accuracy and quality of the text, as compared to those who held positive views. Specifically, 43.5% of the respondents expressed a negative opinion, while 32.3% held a positive one.

An interesting aspect of the data is the connection between responses and age, which is illustrated in Figure 10. According to five quality statements sorted by age groups, the following conclusions can be drawn.

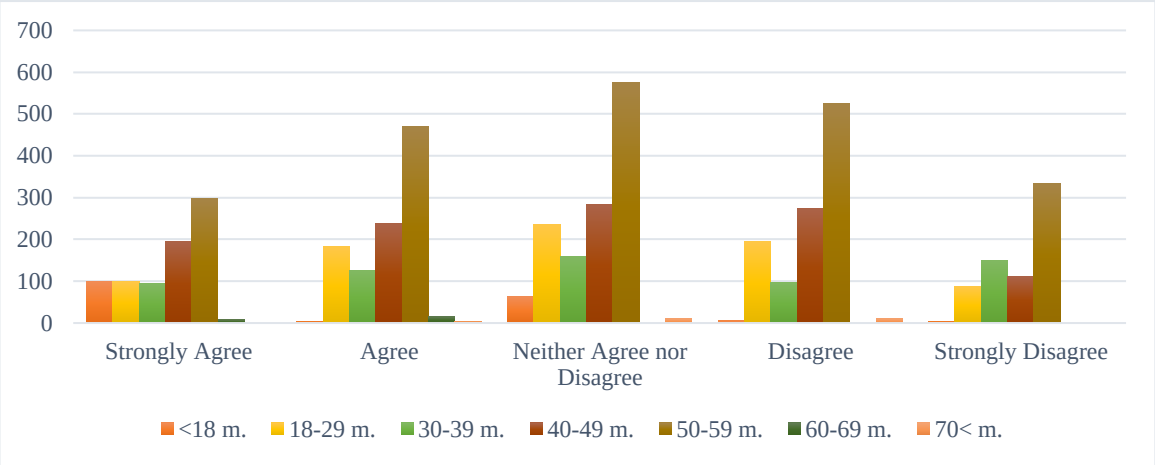


Fig. 10. Breakdown of responses by age group

- Age group <18 years: The majority (57.1%) of respondents in this age group strongly agree with the quality parameters, suggesting a high level of satisfaction with the machine-translated websites.
- Age group 18-29 years: The distribution of opinions regarding the quality of the translations is more even in this age group, with 35.2% agreeing or strongly agreeing and 35.4% disagreeing or strongly disagreeing. The largest group (29.5%) is neutral, neither agreeing nor disagreeing with the quality parameters.
- Age group 30-39 years: In this age group, 35% of respondents agree or strongly agree that the texts are of high quality, while 39.4% disagree or strongly disagree with the quality. 25.6% of respondents are neutral, neither agreeing nor disagreeing.
- Age group 40-49 years: Respondents in this age group have a relatively balanced distribution of opinions regarding the quality, with 39.1% agreeing or strongly agreeing and 35.1% disagreeing or strongly disagreeing. The largest group (25.7%) is neutral.
- Age group 50-59 years: In this age group, 34.9% of respondents agree or strongly agree that the texts are of high quality, while 38.9% disagree or strongly disagree that they are reading a quality website. 26.1% of respondents are neutral.
- Age group 60-69 years: The majority (64%) of respondents in this age group agree with the quality parameters, and 36% strongly agree. No respondents in this age group disagree or strongly disagree with the quality parameters.
- Age group 70+ years: In this age group, 12% agree, 44% neither agree nor disagree, and 44% disagree with the quality of the texts. Participants in this age group did not provide any strong agreements or disagreements with the quality parameters.

Overall, younger respondents (<18 years) have the most positive perception of the machine-translated websites, with a majority strongly agreeing with the quality parameters. The age groups 18-29, 30-39, 40-49, and 50-59 years show a more balanced distribution of opinions, with a significant proportion of respondents having positive, neutral, or negative perceptions of the quality parameters. The age group 60-69 years has a predominantly positive perception of the machine-translated websites, with no negative perceptions. The age group 70+ years has a more uniform distribution of positions, with no respondents strongly agreeing or strongly disagreeing with the quality parameters.

In summary of the data gathered from the entire survey and its responses distribution, the following conclusions can be made. In general, these results indicate that while there are some positive perceptions of machine-translated website quality, there is a significant portion of respondents who had a negative experience. Overall, the end-user survey data reveals that the acceptability of machine-translated English-to-Lithuanian spam websites from the users' perspective is mixed. While some aspects, such as clarity and informativeness of the text, received relatively positive feedback, other aspects like enjoyability, trustworthiness, and accuracy and quality raises questions. The even distribution of survey responses implies that participants may have reservations about the translation quality, resulting in average ratings, or that they are uncertain about the quality itself. Nevertheless, this finding may require further investigation. Also, the data suggests that younger and older age groups (specifically <18 years and 60-69 years) have a relatively positive perception of the quality of machine-translated websites, whereas the remaining age groups demonstrate a more proportional opinions. For the age groups with more balanced views it would be helpful to identify the specific aspects of the translations that lead to negative or positive opinions, but this would require even more in-depth investigation.

2.3. An Analysis of BLEU Scores for Machine-Translated Websites

In the subsequent part of the research, the objective was to examine and evaluate the translation quality of the machine-translated websites. To do this, an automated evaluation method, the BLEU score, was employed to assess the texts of the chosen three websites. The BLEU scores provide a quantitative measure of translation quality, with higher scores indicating better translations. It is essential to remember that a higher BLEU score indicates better translation quality, specifically in terms of agreement between machine and human translations. A score between 60-80 is considered very good, while a score of 100 indicates a perfect match between human and machine translations. Three websites were assessed: the first website featured a text length of 140 characters, the second website contained a text length of 1200 characters, and the final third website comprised a text length of 1224 characters. The following are their BLEU scores.

Table 2. BLEU score for website No. 1

BLEU:	7.15			
Precision x brevity:	8.01 x 89.22			
Type	1-gram	2-gram	3-gram	4-gram
Individual	48.99	12.59	4.38	1.53
Cumulative	43.71	22.16	12.42	7.15

Website No. 1 has a low BLEU score of 7.15, indicating poor translation quality (Table 2). The precision x brevity score of 8.01 x 89.22 suggests that the translation is not very accurate and tends to be slightly shorter than the reference translation. Individual and cumulative n-gram scores for Website No. 1 are low across all n-grams (1-gram to 4-gram), indicating that the translation struggles at both the word and linguistic levels. The overall low BLEU score suggests that the machine-translated content on this website is likely to be difficult to understand and may require thorough editing or re-translation to be understandable to the reader.

Table 3. BLEU score for website No. 2

BLEU score:	48.74			
Precision x brevity:	51.26 x 95.09			
Type	1-gram	2-gram	3-gram	4-gram
Individual	74.29	55.08	44.96	37.52
Cumulative	70.64	60.82	54.08	48.74

Website No. 2 has a relatively high BLEU score of 48.74, suggesting a good translation quality (Table 3). The score of 51.26 x 95.09 for precision and brevity indicates that the translation is more accurate and closely matches the length of the reference translation. Also, for Website No. 2, the individual and cumulative n-gram scores are consistently higher than those of the other two websites, indicating better translation quality regarding both word-level and language structure. In general, the translation quality of Website No. 2 is better than that of the other two websites. As a result, the content of Website No. 2 is more likely to be understandable and useful to the target audience without the need for major editing.

Table 4. BLEU score for website No. 3

BLEU score:	37.21			
Precision x brevity:	37.21 x 100.00			
Type	1-gram	2-gram	3-gram	4-gram
Individual	63.92	41.38	30.65	23.64
Cumulative	63.92	51.43	43.28	37.21

Finally, the average BLEU score of the third site was 37.21, indicating that the quality of the translation is fair (Table 4). Also, the translation is quite accurate and in line with the length of the reference translation, as shown by the precision x brevity score of 37.21 x 100.00. Website's individual and overall n-gram scores are average across all n-grams, indicating that the quality of words and sentence structures is acceptable. The overall BLEU results imply that the content of this website should be acceptable and clear to the reader.

Based on the analysis of the BLEU scores, Website No. 2 demonstrates the highest translation quality among the three machine-translated websites. Website No. 3 has a moderate translation quality, while Website Nr. 1 shows a poor translation quality with a low BLEU score and n-gram scores. This indicates that the translation quality is varied among the different websites, with some being of low quality and others being more understandable and usable. Noteworthy, that variations in quality scores can be attributed to the complexity of the text. Machine translation performs better in cases where it has been trained on and recognizes the specific context, resulting in higher quality translations. However, texts with more specialized terms or nuanced meanings that require indirect translation pose greater challenges for machine translation systems. It's also important to point out, that the BLEU score has its limitations and should be interpreted in context with other evaluation metrics, for example, human judgment to assess the overall translation quality. Therefore, the final part of the study will concentrate on assessing the quality of spam website texts through the perspective of a professional translator.

2.4. An Analysis of Translation Errors for Machine-Translated Websites Using MQM

The final part of the research concentrates on the evaluation of errors by human expert in machine-translated websites. It is considered that human judgement, although subjective, is clearly superior to automatic error assessment. This is because human evaluation considers errors more comprehensively and acknowledges that there may be several variants of a correct translation. Automatic evaluation, on the other hand, unfortunately only takes into account the overlap between machine translation and reference translation. With this in mind, it will be interesting to discuss not only the errors identified, but also the correlations between human evaluation and automatic evaluation. In order to evaluate the quality of the machine-translated spam websites, Multidimensional Quality Metrics (MQM) was employed. The evaluation tool is widely recognized, covers a broad range of error types, and can be easily used in various translation assessment tasks. The current analysis was conducted using the latest version of the MQM issue types.

This analysis aims to provide valuable insights into the translation quality of the websites examined and to pinpoint any areas where the machine translation could be improved. In this study, a sample of 1,060 words of text from machine-translated websites was analyzed, with a focus on six error dimensions and their subtypes, including accuracy, linguistic convention, terminology, style, locale conventions, audience appropriateness and design. For each dimension, the identified errors were categorized according to their severity: neutral, minor, major, and critical. The severity penalty points for each error were calculated based on the MQM error penalty guidelines. Later, the total number of errors and the corresponding penalty points were aggregated for each dimension, providing an overall evaluation of the translation quality (the analyzed segments, refer to Appendix 3). Table 5 presents the complete data collected from the error annotation.

Table 5. Translation error distribution and count

Error Types	**Error Severity Levels**					
	Critical	**Major**	**Minor**	**Neutral**	**Total**	**Severity Points**
Accuracy	**9**	**19**	**23**	**3**	**54**	**343**
Addition		1			1	
Mistranslation	9	17	17	1	44	
Under-translation		1	6	2	9	
Linguistic conventions			**15**	**22**	**37**	**15**
Grammar			12	9	21	
Punctuation			2	9	11	
Spelling			1	4	5	
Locale convention			**1**		**1**	**1**
Measurement format			1		1	
Style			**5**	**1**	**6**	**5**
Awkward			5	1	6	
Terminology	**10**	**9**	**9**		**28**	**304**
Inconsistent use of terminology	10	9	9		28	
Grand Total	19	28	53	26	126	668

The results indicate that the translation fails to meet acceptable quality standards, with a total of 126 errors and a severity penalty points of 668. It is important to note that according to the developers of MQM, a high-quality translation should not exceed 100 penalty points, therefore the data obtained suggest that machine-translated spam websites text has significant room for improvement in terms of translation quality. Next, we will discuss each error dimension in detail and highlight the most notable examples within each category.

Accuracy errors

In the accuracy dimension, errors are related to how well the translation reflects the original text. This dimension has subcategories such as mistranslation, over-translation, under-translation, addition, omission, do not translate, and untranslated. A total of 54 errors were identified in the accuracy dimension, with 3 being neutral, 23 minor, 19 major, and 9 critical errors (see Fig. 11). The severity penalty points for this category amounted to 343. The distribution of errors based on their subcategories is as follows:

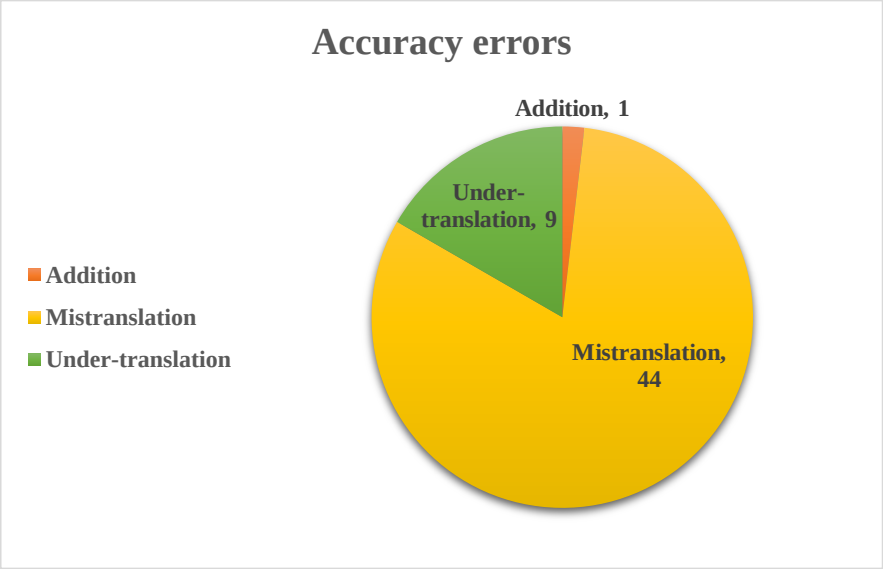


Fig. 11. Distribution of accuracy errors

The accuracy category had a relatively high number of errors, indicating that there are many issues with machine-translated text. Clearly, it is still challenging for machine translation technologies to convey the original text's meaning accurately. Among the subcategories within this dimension, mistranslation had the highest number of errors, suggesting that the machine-translated text did not accurately convey the meaning of the original text. Out of the total errors found, 54 errors were related to mistranslation subcategory. There were hardly any other types of errors found, except for the under-translation sub-category, which had 9 errors and 1 addition error. We will now examine a few examples.

Example 1

Source:	The Steerable Rock System (SRS) is the auger boring market's first steerable head designed to navigate not only solid rock **but difficult fractured rock conditions as well**.
Target:	Steerable Rock System (SRS) yra pirmoji sraigtinių gręžinių rinkoje valdoma galvutė, **skirta važiuoti** ne tik kietomis uolienomis, **bet ir sudėtingomis skilimo sąlygomis**.

Reference translation:	Valdomoji uolienų gręžimo sistema (angl. Steerable Rock System, SRS) – tai pirmoji tokio tipo sistema sraigtinių grąžtų rinkoje. Valdomasis gręžimo antgalis **sukurtas valdyti gręžimą** ne tik kietose uolienose, **bet dirbti ir sudėtingomis sąlygomis, kuomet uolienos skilinėja**.

In the following example, we can see how machine translation can misrepresent the meaning of the original text. This is often caused by literal translation which fails to capture the intended meaning. For instance, the phrase "designed to navigate" has been translated as "skirta važiuoti" when the intended meaning was related to controlling the drilling process. Additionally, the portion of the sentence "fractured rock conditions as well" has been translated as "bet ir sudėtingomis skilimo sąlygomis", when a more accurate translation would be "bet dirbti ir sudėtingomis sąlygomis, kuomet uolienos skilinėja". Such an error is considered to be a major error because it undoubtedly fails to convey the meaning and misleads the reader.

Example 2

Source:	For example, Archimedes built a steam gun **Heron of Alexandria** used the energy of steam to open the doors of ancient temples.
Target:	Pavyzdžiui, Archimedas pastatė garo pistoletą **Aleksandrijos garnys** panaudojo garų energiją senovinių šventyklų durims atidaryti.
Reference translation:	Pavyzdžiui, Archimedas sukūrė garais varomą patranką, kurios garų energiją **Heronas iš Aleksandrijos** panaudojo atidaryti senovinių šventyklų durims.

In the example 2, the machine translation inaccurately translated the pronoun "Heron of Alexandria" as "Aleksandrijos garnys", leading to confusion in the target text. The accurate translation of the pronoun should be "Heronas iš Aleksandrijos".

Example 3 below in the accuracy dimension will pertain to the sub-category of under-translation. In this sub-category, errors occur in the cases when the target text is less specific than the source text.

Example 3

Source:	Engineered to operate in rock up to 25,000 psi, the SRS **allows operators to keep bores online and grade even in the toughest ground conditions for an on-grade bore**.
Target:	Sukurtas veikti uolienose iki 25,000 4 psi, SRS **leidžia operatoriams išlaikyti gręžinius tinkle ir greiderį net sunkiausiomis žemės sąlygomis**.
Reference translation:	Valdomoji uolienų gręžimo sistema (sukurta darbui uolienose iki 1724 ba slėgio) **leidžia gręžimo operatoriams matyti gręžiamą ertmę ekrane net gręžiant žemę sunkiausiomis sąlygomis**.

In this case, we see an under-translation error where the machine translation fails to capture the complete meaning of the original text. It is a literal translation that lacks context and may result in ambiguity or confusion. In comparison, human translation is more thorough and provides a better understanding of the text.

The relatively large proportion of major and critical errors in the accuracy category is particularly concerning, as these errors can severely impact the comprehensibility and usability of the translated text. A translation of this kind on websites is expected to result in a text that is hard to comprehend. Furthermore, it can lead to readers losing trust in the website and the information presented there. It is also worth noting that the severity penalty points for the accuracy category (343) account for more than half of the total severity penalty points (668) across all error categories. This shows that this error dimension is significant for the overall qualitative result, indicating that the machine translation technology is still imperfect, particularly in the English-Lithuanian language pair.

Linguistic convention errors

The second most frequent type of errors found in the analysis pertained to linguistic conventions. These errors concerned the appropriateness and correctness of language usage, including issues with grammar, spelling, punctuation, and overall mechanical accuracy. Linguistic conventions errors were less prevalent in the text compared to accuracy errors, with a total of 37 errors found. Most of these errors were considered neutral (22), followed by minor errors (15), and there were no major or critical errors, also 15 severity penalty points were attributed to this category.

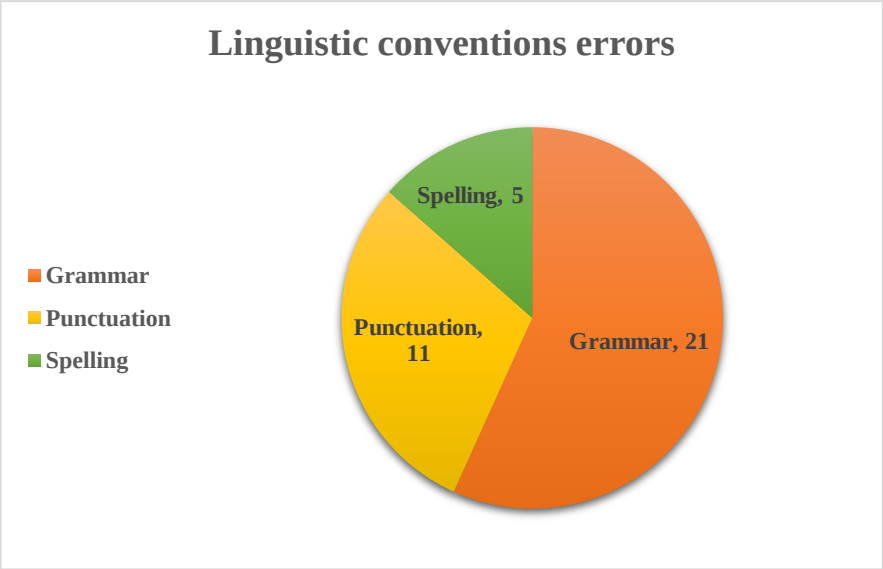


Fig. 12. Distribution of linguistic conventions errors

Figure 12 presented above shows the breakdown of errors. The errors found in the linguistic conventions dimension were primarily related to grammatical errors, punctuation, and a few cases of misspelling. In terms of the distribution of errors in the linguistic conventions dimension, there were a total of 21 grammatical errors, 11 punctuation errors, and 5 spelling errors. Grammatical errors in this dimension pertained to the incorrect use of number or case agreement in the sentence and also issues with sentence structure. Punctuation errors were mainly related to the incorrect use of long dash punctuation, which is a minor but frequent error in machine translation. Spelling errors, on the other hand, were related to inaccuracies in spelling numbers and percentages. However, regarding severity penalty points, as there were no major or critical errors found in this category, it suggests that these errors may not have a significant impact the understanding of the text, but may instead affect its overall quality and readability. We will examine some typical examples below.

Example 4

Source:	Valerian (tablets) – **instructions** for use, reviews, analogues
Target:	Valerijonas (tabletės) – naudojimo **instrukcijos**, apžvalgos, analogia
Reference translation:	Valerijono tabletėmis naudojimo **instrukcija**, vertinimas, sudėtis

Example 4 demonstrates a number agreement error, where the machine translation literally translates a word’s meaning. The word "instructions" in the source text is plural, but the machine translation also renders it as plural, whereas the more appropriate translation would be the singular "instrukcija" in this context. It is important to note that this error is minor, meaning that the meaning of the text

remains unchanged, and it is not likely to mislead the reader. However, it may affect the overall quality of the text's style.

Example 5

Source:	In rhizomes and roots, the amount of essential oil reaches 2%.
Target:	**Šakniastiebiuose ir šaknyse eterinio aliejaus kiekis** siekia 2%.
Reference translation:	**Eterinio aliejaus kiekis šakniastiebiuose ir šaknyse** siekia 2 %.

Example 5 shows a syntactic error in the machine translation where the word order in the sentence is not followed correctly. The reference translation, on the other hand, chooses a different word order in the sentence to highlight the most important part of it. Although this is an error, it has been classified as a neutral error, meaning it does not count as a penalty since it is more related to the translator's style or approach.

Example 6

Source:	Composition 1 tablet: **active substance**:
Target:	1 tabletės sudėtis: **veiklioji medžiaga**:
Reference translation:	Vienos tabletės sudėtis ir **veikliosios medžiagos**:

A minor translation error was identified in example 6, specifically related to the incorrect translation of agjective and noun number. The phrase "active substance" should have been translated as "veikliosios medžiagos" to accurately represent that the medicine consists of multiple substances. However, the machine translation rendered it as "veiklioji medžiaga", incorrectly using the singular form, as in the original English text.

Example 7

Source:	- age up to 3 years;
Target:	- amžius iki **3** metų;
Reference translation:	- amžius iki **3 - jų** metų;

Example 7 demonstrates a translation mistake involving a misspelling of numbers. In Lithuanian, where Arabic numerals are used to denote a sequence, the word endings can be added after a hyphen, especially in a sentence. This error is considered a neutral error, representing a choice in translation style, and therefore no penalty points are assigned.

Overall, it is important to note that severity penalty points for the linguistic conventions category (15 points) are lower compared to the accuracy and terminology categories. This finding means that while language correctness is an essential translation quality aspect, the machine-translated websites had relatively minor errors in this category. However, the translation accuracy and terminology errors were more problematic in the texts. Nevertheless, addressing linguistic conventions issues remains important to improve the overall quality of the translation.

Terminology

The category with the third-highest number of errors was related to terminology. Terminology errors are related to issues where the translation of a term does not align with the terminology standards in a particular field, or when a term is used incorrectly in the target text compared to recognized sources of terminology. In the terminology error category, a total of 28 errors were found, including 9 minor, 9 major, and 10 critical errors. While errors in this category resulted in a severity penalty score of 304. Comparing the penalty scores of the terminology category to those of the linguistic conventions

and style categories, it is noteworthy that the former has a higher score. This suggests that errors related to terminology are more problematic and significant in machine-translated spam websites content. The significant number of major and critical errors found suggests that errors in terminology translation can have a significant impact on both the quality of the text and the reader's ability to understand it. In the terminology category, all errors were annotated under the subcategory of inconsistent with terminology resource, which means that a term does not match the required term usage specified by a specified terminology resource. We shall examine some examples of terminology errors.

Example 8

Source:	The main components of the oil: **bornyl isovalerate**, **isovalerianic acid**, borneol, **camphene**, α-pinene, limonene, etc.
Target:	Pagrindiniai aliejaus komponentai: **borniloizovalatas**, **izovalerianinė rūgštis**, borneolis, **kafelenas**, α-pinenas, limonenas ir kt.
Reference translation:	Pagrindinės aliejaus sudedamosios dalys: **bornilizovaleratas**, **izovalerijono rūgštis**, borneolis, **kamfenas**, α-pinenas, limonenas ir kt.

In this example, we see multiple terminology errors in a single sentence produced by the machine translation. The Ministry of Health of The Republic of Lithuania refers to "bornyl isovalerate" as "bornilysovalerate", while the Lithuanian Term Bank translates "camphene" as "kamfenas" and the Lithuanian State Medicines Control Agency translates "isovalerianic acid" as "izovalerijono rūgštis". These errors are considered critical as accurate translation of medical terminology is crucial for ensuring the safety and well-being of the reader. It is important to translate such terms correctly in order to provide accurate information.

Example 9

Source:	**Low Height Pancake Hydraulic Cylinder** Used In Auger Boring Steerable Rock System
Target:	**Mažo aukščio blynų hidraulinis cilindras**, naudojamas sraigtinio gręžimo valdomoje uolienų sistemoje
Reference translation:	**Žemo profilio hidraulinis stūmimo cilindras** skirtas valdomajai sraigtinei uolienų gręžimo sistemai

Example 9 highlights a critical error is noted in the machine translation where the term "Low Height Pancake Hydraulic Cylinder" has been mistranslated as "Mažo aukščio blynų hidraulinis cilindras", leading to confusion and misunderstanding for the reader. The correct translation, "Žemo profilio hidraulinis stūmimo cilindras", accurately reflects the product's functionality and should have been used. This error not only misrepresents the product but also leads to a lack of understanding of the whole text.

Example 10

Source:	Because the Papin **steam engine** like most subsequent projects, they were called steam-atmospheric machines.
Target:	Kadangi Papin **garo mašina** kaip ir dauguma vėlesnių projektų, jie buvo vadinami garo-atmosferinėmis mašinomis.
Reference translation:	Papino garbei, **garo variklis**, kaip ir dauguma vėlesnių projektų, buvo vadinami garo-atmosferinėmis mašinomis.

In example 10, there is a minor mistake in terminology where the machine translation translates "steam engine" as "garo mašina", while the Lithuanian Term Bank translates it as "garo variklis".

This mistake is classified as minor since it does not make the text difficult to understand, but it is stil important for the accuracy of the translation.

The examples we have discussed, along with all the terminology errors found, indicate that machine translation had difficulty recognizing and translating standardized terms. This is because such terms may not be included in their corpora and are not recognized by the machine translation systems. In a general sense to maintain the credibility and professionalism of translated text, it is important to use proper terminology. However, machine translation still has room for improvement in accurately recognizing and translating standardized terms, and the human ability to search for and identify equivalents of such terms remains essential.

Style errors

In comparison to the other categories, this dimension had a relatively small number of errors. It is significant to recall that style errors are errors in a text that are grammatically correct but unsuitable because they don't match the style of language or don't follow the style guide. The style of a text varies depending on its domain, function, and author's competence. A certain style may be appropriate for one type of text, but not for another. For instance, a fiction text may focus on emotional and visual language, while a scientific or administrative text may use more formal and straightforward language.

Following this, errors related to style were not as common as errors in other categories, as only 5 errors were found. The style errors were categorized as 1 neutral, 5 minor, resulting in a total of 5 severity penalty points, with no critical or major errors found. For the annotation of errors, the style error dimension had several subcategories including organizational style, third-party style, inconsistent with external reference, register, awkward style, unidiomatic style, and inconsistent style. However, all the errors identified belonged to the subcategory of awkward style, which refers to style that contains excessive words or overly complex sentence structures, often caused by inappropriate translation of the source text style in the target text. We will now examine some examples of awkward style identified in machine translation.

Example 11

Source:	**Before you buy Valerian in tablets, we recommend that you familiarize yourself with the proposed instructions**, which describe the methods of application, dosage, and the concomitant effect of this drug on the human body.
Target:	**Prieš pirkdami valerijoną tabletėmis, rekomenduojame susipažinti su siūlomomis instrukcijomis**, kuriose aprašomi šio vaisto vartojimo būdai, dozavimas ir lydimasis poveikis žmogaus organizmui.
Reference translation:	**Prieš įsigyjant valerijono tabletes, rekomenduojame susipažinti su vartojimo instrukcija**, kurioje aprašomi vartojimo būdai, dozavimas ir poveikis žmogaus organizmui.

Example 11 illustrates a case of awkward style, where the machine translation translates the sentence literally, resulting in a minor error with minimal penalty points. Although this may not affect the reader's understanding of the text, the stylistic quality of the translation is hardly satisfactory. For instance, the machine translation translates the phrase "Before you buy Valerian in tablets, we recommend that you familiarize yourself with the proposed instructions" as "Prieš pirkdami valerijoną tabletėmis, rekomenduojame susipažinti su siūlomomis instrukcijomis". The quality of the text is reduced by the use of the participle and the plural in this case. However, a more acceptable translation is "Prieš įsigyjant valerijono tabletes, rekomenduojame susipažinti su vartojimo instrukcija", which is much more fluent in terms of style.

Example 12

Source:	**Valerian has a calming effect on the nervous system**, which improves cardiac function.
Target:	**Valerijonas ramina nervų sistemą**, todėl gerėja širdies veikla.
Reference translation:	**Valerijonas pasižymi raminančiu poveikiu nervų sistemai** bei gerina širdies darbą.

Example 12 shows another instance of awkward style in machine translation, where the translated text literally states that valerian calms the nervous system "Valerijonas ramina nervų sistemą", whereas the original text meant that it has a "calming effect" on it. The preferred translation is "pasižymi raminančiu poveikiu nervų sistemai" which is a more commonly used phrase in the context of medications. While this is a minor error and the meaning can still be understood, it is not an ideal translation.

Example 13

Source:	**Valerian is a unique sedative** used to alleviate stress.
Target:	**Valerijonas yra unikali raminamoji priemonė**, naudojama stresui malšinti.
Reference translation:	**Valerijonas yra išskirtinis raminamasis vaistas**, naudojamas streso malšinimui.

Example 13 demonstrates a minor style error in the given context. The machine translation uses "unikali raminamoji priemonė" to describe a medication, while the translator chooses "išskirtinis raminamasis vaistas", which is a more appropriate choice. "Priemonė" is too broad term to accurately describe the medication in this context.

In general, although the number of errors in the style category is lower than in other categories, it is still important to correct these errors to ensure that the translated text is consistent and professional in tone. After analyzing the machine-translated websites, it was found that there were a relatively low number of errors in the Style category. While these errors may not greatly affect the overall quality of the text, it is important to note that machine translation is still not perfect and errors can make it difficult to understand the intended meaning.

Locale Convention errors

Locale convention errors refer to issues in translation that do not comply with the content or data formatting standards of a particular locale, such as errors in number, currency, measurement, time, date, address, telephone or shortcut formats. Locale convention errors were relatively rare in this study, with just 1minor error found. There were no critical errors in this category, and the severity penalty points totaled to 1. Here is an example of translation error related to locale conventions.

Example 14

Source:	Engineered to operate in rock up to **25,000 psi**, the SRS allows operators to keep bores online and grade even in the toughest ground conditions for an on-grade bore.
Target:	Sukurtas veikti uolienose iki **25,000** 4 **psi**, SRS leidžia operatoriams išlaikyti gręžinius tinkle ir greiderį net sunkiausiomis žemės sąlygomis.
Reference translation:	Valdomoji uolienų gręžimo sistema (sukurta darbui uolienose iki **1724 ba** slėgio) leidžia gręžimo operatoriams matyti gręžiamą ertmę ekrane net gręžiant žemę sunkiausiomis sąlygomis.

This instance demonstrates a locale convention error in machine translation where the unit of measurement, "25,000 psi," is translated literally despite the fact that pressure is typically measured in bars in Lithuanian according to the Lithuanian Term Bank. Therefore, a conversion of the unit of

measurement is necessary for accurate translation into Lithuanian, where "25,000 psi" is equivalent to "1724 bar."

On the whole, locale convention errors may not have a major impact on the quality of the machine-translated text, but they do suggest that human involvement in post-editing is still required. Machine translation is not advanced enough to account for all the specific nuances of different locales, particularly in languages like Lithuanian where resources are limited.

In summary, the analysis of 1060-word machine-translated websites revealed 126 errors according to the MQM framework, with a severity penalty of 668, indicating poor overall translation quality into Lithuanian. The number of penalty points indicates poor translation quality, as a high-quality translation should not accumulate 100 penalty points. The accuracy and terminology categories had the most significant issues, indicating that the machine translation system had difficulty conveying the intended meaning and following to domain-specific terminology. Other categories, such as linguistic conventions, style, and locale conventions, had fewer errors and lower severity penalty points. No design or verity errors were found in the sample text. However, even in categories with fewer errors, the identified errors were still significant enough to require editing by a professional translator before using the translation as an end product.

After analyzing the results of the three methods used, we can summarize the findings and offer some insighst. This study used three different methods to analyze the translation quality of machine-translated websites, including a human expert analysis using the MQM framework to identify errors, a survey to evaluate user perception of quality and acceptability, and the BLEU score for automatic TQA evaluation. All three methods were employed in a systematic manner to comprehensively evaluate the translation quality and to investigate its perception by the end-user. Each method employed in the research has its own strengths, providing distinct perspectives and valuable information regarding the translation quality of machine-translated websites. The MQM error annotation method has several advantages, such as providing a thorough analysis of various error categories and identifying specific problematic areas like accuracy, linguistic conventions, terminology, and style. Additionally, it allows for the quantification of errors using severity penalty points, making it easier to compare different translation issues. The end-user survey method collects feedback from the actual users of the translated content to gain insight into what is important to them and their perspectives on the quality of the translation. While, the BLEU score evaluation provides an automated and objective approach to assess the quality of translations. It measures both precision and brevity, ensuring translations are both accurate and fluent. This method is cost-effective and efficient, especially when evaluating large amounts of text. By using a combination of these methods, the study offers a more complete insight into the quality of translations produced in machine-translated websites.

Returning to the study’s findings, an end-user survey revealed mixed perceptions of the websites’ translation quality. The majority of respondents (26.8%) had a neutral perception, while 20.9% agreed that the translations were of good quality, and 16% strongly agreed they were of high quality. On the other hand, 22.3% disagreed that the translations were of good quality, and 13.9% strongly disagreed. The age and education distribution of the respondents were diverse, with most having higher education (77.8%). The data implies that the translation quality may be mediocre, or participants could be uncertain about the quality. Although certain elements like clarity and informativeness were favorably rated, other factors such as enjoyability, trustworthiness, and accuracy raised concerns. The survey outcomes did not definitively determine if the participants were satisfied or dissatisfied with

the machine-translated websites' quality. There are several possible interpretations of this result: it is possible that the quality of the websites was neither important nor relevant to the respondents, or that the respondents did not have enough information, or were not sure about the topic, to express a strong opinion. An alternative explanation could be that the Likert scale fails to capture the respondent's complex opinion, leading them to occasionally select a neutral response. A notable finding is that a significant number of highly educated respondents in the survey, which may not reflect the views of the general population. Additionally, there were varying opinions among different age groups, with younger and older respondents (<18 years and 60-69 years) showing a more positive perception of the translation quality. This suggests that different age groups may have different standards for quality. Future research could explore the opinions of respondents from diverse educational backgrounds and age ranges.

In the second method, BLEU scores were used to assess the translation quality of three different websites. The scores varied, with Website No. 1 having a low score of 7.15, while Websites No. 2 and No. 3 had higher scores of 48.74 and 37.21, respectively. This indicates that the translation quality varied among the different websites, with some being of poor quality and others being more understandable and functional. Nonetheless, it is important to recognize that the BLEU score has constraints in assessing subtleties and meaning within a text, making it more appropriate for evaluating large datasets. Moreover, the length of texts used in this study may also limit the accuracy of the BLEU scores obtained; therefore, longer texts may be needed for more precise results.`

Lastly, with the MQM error annotation method a total of 126 errors in different categories were identified, with a severity penalty points of 668. The accuracy category had the highest number of errors (54), followed by linguistic conventions (37) and terminology (28). These results support the conclusion that the quality of machine translation into Lithuanian language is insufficient and post-editing is still required. The high number of errors in the accuracy and terminology categories indicate that the end-user is unlikely to receive a high-quality final product.

In summary, the relationship among the outcomes of the three methods reveals that the translation quality of machine-translated websites is not uniformly high. The MQM error annotation findings and the diverse opinions from the end-user survey imply that enhancements and post-editing are needed in multiple aspects of the translations. Additionally, the inconsistent BLEU scores suggest that translation quality may vary according to the particular website.

Conclusions

In this section, the study's main results are presented in relation to the research objectives and questions; also their value and contribution to the field of translation studies are discussed. The present study aimed to evaluate both the linguistic quality by analyzing errors and end-user acceptability of machine-translated spam websites in the Lithuanian language. This was accomplished by employing a combination of automated and human assessment tools for machine translation output quality. After analyzing the data, the following conclusions can be made in response to the objectives raised in the study:

1. Translation quality assessment studies recognize that there is no universal method for evaluating translation quality due to the subjective nature of the concept and the continuously evolving landscape of translation technologies, which present constant challenges. In this regard, different methods are employed to evaluate the quality of translation, but automatic and human expert assessments are commonly used in the field. The BLEU score is widely used for automatic evaluation, while the MQM framework is considered a flexible and comprehensive method for human expert evaluation. Assessing the quality of websites content is equally important in this study. From an end-user perspective, it is important to consider if the website meets the user's objective, if the content is reliable and informative, and if the content is clear and understandable. In addition to automatic and human expert evaluation, it is also essential to consider the opinion of the end-user. Their perspective is important to determining whether the translation product meets their expectations. Thus, incorporating various evaluation methods can lead to more comprehensive and accurate results in assessing translation quality.
2. The initial stage of the study investigated the perceptions of end-users regarding the quality of translated content on different websites. The findings revealed a range of opinions on the acceptability of the translations, highlighting some uncertainty and concerns among the respondents. While certain aspects of the translated content, such as clarity and informativeness, received relatively positive feedback, other aspects such as enjoyability, trustworthiness, accuracy, and quality raised questions among the participants. The findings also revealed variations in quality evaluations across different age groups, with both younger and older participants tending to rate the quality more positively. Overall, the absence of consensus among evaluators regarding the quality of translation indicates that it is a subjective measure of the assessment, underscoring the importance of public education on the subject.
3. The study found variations in machine translation quality across different websites as evaluated by the BLEU score. Inconsistency in quality scores can be attributed to the text's complexity, with machine translation performing better in familiar contexts. However, specialized terms and nuanced meanings that necessitate indirect translation present challenges for machine translation systems. However, it is important to note that the BLEU score only considers exact word matches and does not account for possible correct translations that may differ from the reference translation. This evaluation method, while widely used in the industry, has limitations that fail to capture the full range of translation errors.
4. The use of the MQM framework in evaluating the translation quality of selected spam websites proved to be an effective and systematic approach. The human expert's evaluation identified a significant number of translation errors across various categories, with the accuracy category having the highest number of errors. This highlights the importance of thorough editing by a professional translator to ensure the quality of the translated content before it is used as a final product. In accordance with the findings of other researchers, the results obtained show that the machine translation quality is not up to the mark and that the translated product cannot not be yet

acceptable. Therefore, it needs to be post-edited to improve its terminology, linguistic aspects, and text structure.

The focus of this study was to investigate the quality of machine-translated spam websites that are widely available on the internet. Since these websites are accessible to end-users who consume content daily, the study aimed to understand how these users perceive the quality of such translations. The study conducted in three parts revealed that the quality of machine-translated websites was not acceptable and required editing to make them accurate and understandable. The end-user perception of the texts was mixed and uncertain.

It is necessary to point out that linguistic research incorporating a variety of methods, such as expert linguistic analysis, social surveys, and language technology evaluation, provides a more comprehensive and diverse understanding of the topic, leading to more objective and insightful results. To bring it all together, this study involves multiple layers, including language, technology, and society perception, which adds value and depth to the findings. The study also emphasizes that this topic is equally important to both language researchers and end-users of the translation product. As language models continue to advance, the amount of machine-translated and machine-generated text will likely increase. This study's findings might suggest a need to educate the society on available technologies, their benefits, and the risks associated with low-quality or factually inaccurate content. This could help users recognize such content and make informed decisions. Lastly, it is important not only for experts to assess the quality of translated content, but also to involve the society in evaluating it.

List of references

1. Al-Kabi, M., Wahsheh, H., Alsmadi, I., Al-Shawakfa, E., Wahbeh, A., & Al-Hmoud, A. (2012). Content-based analysis to detect Arabic web spam. *Journal of Information Science*, 38(3), 284-296. [viewed 28 March 2023]. Retrieved from https://doi.org/10.1177/0165551512439173
2. Aswani, R., Ghrera, S. P., Chandra, S., & Kar, A. K. (2021). A hybrid evolutionary approach for identifying spam websites for search engine marketing. *Evolutionary Intelligence*, 14(4), 1803-1815. [crossref] [viewed 28 March 2023]. Retrieved from https://doi.org/10.1007/s12065-020-00461-1
3. Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. arXiv:1409.0473. [viewed 28 March 2023]. Retrieved from https://arxiv.org/abs/1409.0473
4. Banerjee, S., & Lavie, A. (2005, June). METEOR: An automatic metric for MT evaluation with improved correlation with human judgments. *In Proceedings of the acl workshop on intrinsic and extrinsic evaluation measures for machine translation and/or summarization* (pp. 65-72). [viewed 25 March 2023]. Retrieved from https://aclanthology.org/W05-0909.pdf
5. Burchardt, A., Lommel, A., & Macketanz, V. (2021). A new deal for translation quality. *Universal Access in the Information Society*, 20, 701-715. [viewed 2 April 2023]. Retrieved from https://doi.org/10.1007/s10209-020-00736-5
6. Carroll, J. B. (1966). An experiment in evaluating the quality of translations. Mech. Transl. Comput. *Linguistics*, 9(3-4), 55-66. [viewed 2 April 2023]. Retrieved from https://aclanthology.org/www.mt-archive.info/MT-1966-Carroll.pdf
7. Castells, M. (2014). The impact of the internet on society: a global perspective. *Change*, 19, 127-148. [viewed 2 April 2023]. Retrieved from http://aasa.ut.ee/augsburg/literature/CASTELLS_BBVA-OpenMind-book-Change-19-key-essays-on-how-internet-is-changing-our-lives-Technology-Internet-Innovation.pdf
8. Castilho, S. (2016). *Measuring acceptability of machine translated enterprise content: Doctoral dissertation*. Dublin City University. [viewed 21 March 2023]. Retrieved from https://doras.dcu.ie/21342/
9. Castilho, S., & Guerberof Arenas, A. (2018). Reading comprehension of machine translation output: what makes for a better read?. European Association for Machine Translation. [viewed 5 April 2023]. Retrieved from http://hdl.handle.net/10045/76032
10. Castilho, S., Doherty, S., Gaspari, F., & Moorkens, J. (2018). Approaches to human and machine translation quality assessment. *Translation quality assessment: From principles to practice*, 9-38. [viewed 21 March 2023]. Retrieved from https://doi.org/10.1007/978-3-319-91241-7_2
11. Chauhan, S., & Daniel, P. (2022). A comprehensive survey on various fully automatic machine translation evaluation metrics. *Neural Processing Letters*, 1-55. [viewed 21 March 2023]. Retrieved from https://doi.org/10.1007/s11063-022-10835-4
12. Chomsky, N. (2014). *Aspects of the Theory of Syntax* (Vol. 11). MIT press. [viewed 21 March 2023]. Retrieved from http://www.colinphillips.net/wp-content/uploads/2015/09/chomsky1965-ch1.pdf
13. Clerwall, C. (2014). Enter the robot journalist: Users' perceptions of automated content. *Journalism practice*, 8(5), 519-531. [viewed 21 March 2023]. Retrieved from https://doi.org/10.1080/17512786.2014.883116

14. Correa, N. (2003). A fine-grained evaluation Framework for machine translation system development. *In Proceedings of Machine Translation Summit* IX: Papers. [viewed 2 April 2023]. Retrieved from https://aclanthology.org/2003.mtsummit-papers.7/
15. Costa-jussà, M. R., Cross, J., Çelebi, O., Elbayad, M., Heafield, K., Heffernan, K., ... & Wang, J. (2022). No language left behind: Scaling human-centered machine translation. arXiv preprint arXiv:2207.04672. [viewed 2 April 2023]. Retrieved from https://arxiv.org/abs/2207.04672
16. Coughlin, D. (2003). Correlating automated and human assessments of machine translation quality. *In Proceedings of Machine Translation Summit IX*: Papers. [viewed 2 April 2023]. Retrieved from https://aclanthology.org/2003.mtsummit-papers.9/
17. Cvilikaitė, J. (2008). Leksinės mašininio vertimo klaidos: beekvivalenčių žodžių vertimas. *Filologija*, (13), 27-38. [viewed 2 April 2023]. Retrieved from https://etalpykla.lituanistika.lt/object/LT-LDB-0001:J.04~2008~1367167919772/
18. Daubarienė, A., & Ziezytė, G. (2013). Machine translation: translated texts in terms of standards of textuality. *Kalbų studijos*, (22), 55-61. [viewed 2 April 2023]. Retrieved from https://doi.org/10.5755/j01.sal.0.22.4301
19. Daudaravičius, V. (2006). Pradžia į begalybę. Mašininis vertimas ir lietuvių kalba. *Darbai ir dienos*, (45), 7-18. [viewed 21 March 2023]. Retrieved from https://doi.org/10.7220/2335-8769.45.1
20. De Beaugrande, R. A., & Dressler, W. U. (1981). *Introduction to text linguistics* (Vol. 1). London: longman. [viewed 21 March 2023]. Retrieved from http://library.lgaki.info:404/2017/De%20Beaugrande%20R.-A..pdf
21. Doddington, G. (2002, March). Automatic evaluation of machine translation quality using n-gram co-occurrence statistics. *In Proceedings of the second international conference on Human Language Technology Research* (pp. 138-145). [viewed 24 March 2023]. Retrieved from https://doi.org/10.3115/1289189.1289273
22. Doherty, S. (2017). Issues in human and automatic translation quality assessment. *In Human issues in translation technology* (pp. 149-166). Routledge. [viewed 22 March 2023]. Retrieved from https://doi.org/10.4324/9781315648934
23. Dorr, B., Snover, M., & Madnani, N. (2011). *Part 5: machine translation evaluation*. Handb. Nat. Lang. Process. Mach. Transl. DARPA Glob. Auton. Lang. Exploit, 936. [viewed 28 March 2023]. Retrieved from https://www.cs.cmu.edu/~alavie/papers/GALE-book-Ch5.pdf
24. Drugan, J. (2013). *Quality in professional translation: Assessment and improvement* (Vol. 9). A&C Black. [viewed 28 March 2023]. Retrieved from https://doi.org/10.5040/9781472542014
25. Esmail, F., & Jaza'ei, M. (2015). A Translation Quality Assessment of Two English Translations of Nazim Hikmet's Poetry. *Translation Journal*. [viewed 24 March 2023]. Retrieved from https://translationjournal.net/January-2015/a-translation-quality-assessment-of-two-english-translations-of-naz-m-hikmet-s-poetry.html
26. Ferreira, A., Schwieter, J. W., Gottardo, A., & Jones, J. (2016). Cognitive effort in direct and inverse translation performance: Insight from eye-tracking technology. *Cadernos de tradução*, 36, 60-80. [viewed 2 April 2023]. Retrieved from https://doi.org/10.5007/2175-7968.2016v36n3p60
27. Flanagan, M. (1994). Error Classification for MT Evaluation. *In AMTA*. [viewed 24 March 2023]. Retrieved from https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=c6230c506d17d337a9e2fc15afe4f1729e5bfe34

28. Freitag, M., Grangier, D., Tan, Q., & Liang, B. (2022). High quality rather than high model probability: Minimum Bayes risk decoding with neural metrics. *Transactions of the Association for Computational Linguistics*, 10, 811-825. [viewed 21 March 2023]. Retrieved from https://doi.org/10.1162/tacl_a_00491
29. Gyongyi, Z., & Garcia-Molina, H. (2005). Web spam taxonomy. *In first international workshop on adversarial information retrieval on the web* (AIRWeb 2005).
30. Google search central. *Spam policies for Google web search*. [viewed 25 March 2023]. Retrieved from https://developers.google.com/search/docs/essentials/spam-policies?hl=en&visit_id=638020383807236347-3382164925&rd=1#spammy-automatically-generated-content
31. Gouadec, D. (2010). *Quality in translation*. Handbook of translation studies, 1, 270-275. [viewed 21 March 2023]. Retrieved from https://doi.org/10.1075/hts.1.qua1
32. Herrera-Viedma, E., Pasi, G., Lopez-Herrera, A. G., & Porcel, C. (2006). Evaluating the information quality of web sites: A methodology based on fuzzy computing with words. *Journal of the American Society for Information Science and Technology*, 57(4), 538-549. [viewed 25 March 2023]. Retrieved from https://doi.org/10.1002/asi.20308
33. Holmes, J. (1988). *Translated*! Papers on Literary and Translation Studies. With an Introduction by Raymond van den Broek. [viewed 2 April 2023]. Retrieved from https://doi.org/10.1163/9789004486669
34. House, J. (1997). *Translation quality assessment: A model revisited*. Gunter Narr Verlag. [viewed 2 April 2023]. Retrieved from https://doi.org/10.4324/9781315752839-11
35. House, J. (2001). Translation quality assessment: Linguistic description versus social evaluation. *Meta*, 46(2), 243-257. [viewed 2 April 2023]. Retrieved from https://doi.org/10.7202/003141ar
36. Hutchins, W. J. (2014). *History of research and applications*. The Routledge encyclopedia of translation technology, 120. [viewed 2 April 2023]. Retrieved from https://doi.org/10.4324/9780203359792-38
37. Kang, Y. S., & Kim, Y. J. (2006). Do visitors' interest level and perceived quantity of web page content matter in shaping the attitude toward a web site?. *Decision support systems*, 42(2), 1187-1202. [viewed 24 March 2023]. Retrieved from https://doi.org/10.1016/j.dss.2005.10.004
38. Kasperavičienė, R., Motiejūnienė, J., & Patašienė, I. (2020). Quality assessment of machine translation output: cognitive evaluation approach in an eye tracking experiment. *Texto Livre: Linguagem e Tecnologia*, 13(2), 271-285. [viewed 5 April 2023]. Retrieved from https://doi.org/10.35699/1983-3652.2020.24399
39. Kasperė, R., Horbačauskienė, J., Motiejūnienė, J., Liubinienė, V., Patašienė, I., & Patašius, M. (2021). Towards sustainable use of machine translation: usability and perceived quality from the end-user perspective. *Sustainability*, 13(23), 13430. [viewed 5 April 2023]. Retrieved from https://doi.org/10.3390/su132313430
40. Kenny, D. (Ed.). (2022). Machine translation for everyone. Empowering users in the age of artificial intelligence. *Quaderns* IEE, 2(1), 148–152. [viewed 5 April 2023]. Retrieved from https://doi.org/10.5565/rev/quadernsiee.49
41. Kim, H., & Niehm, L. S. (2009). The impact of website quality on information quality, value, and loyalty intentions in apparel retailing. *Journal of interactive marketing*, 23(3), 221-233. [viewed 19 March 2023]. Retrieved from https://doi.org/10.1016/j.intmar.2009.04.009

42. Kincl, T., & Štrach, P. (2012). Measuring website quality: asymmetric effect of user satisfaction. *Behaviour & Information Technology*, 31(7), 647-657. . [viewed 19 March 2023]. Retrieved from https://doi.org/10.1080/0144929x.2010.526150
43. Klubička, F., Toral, A., & Sánchez-Cartagena, V. M. (2017). Fine-grained human evaluation of neural versus phrase-based machine translation. *The Prague Bulletin of Mathematical Linguistics*, 108(1), 121. [viewed 18 March 2023]. Retrieved from https://doi.org/10.1515/pralin-2017-0014
44. Koch, C., & Hartmann, M. (2022). The Impact of a Company Website and Its Perceived Quality on the Buying Intention in B2B-Settings. *In Advances in Digital Marketing and eCommerce* (pp. 9-18). Springer, Cham. [viewed 18 March 2023]. Retrieved from https://doi.org/10.1007/978-3-031-05728-1_2
45. Kocmi, T., Macháček, D., & Bojar, O. (2022). The Reality of Multi-Lingual Machine Translation. arXiv preprint [viewed 18 March 2023]. Retrieved from https://arxiv.org/abs/2202.12814
46. Labutis, V. (2005). Išaugusi vertėjų paklausa–nauji pavojai lietuvių kalbai. *Bendrinė kalba* (iki 2014 metų–Kalbos kultūra), (78), 205-209. [viewed 15 March 2023]. Retrieved from https://www.ceeol.com/search/article-detail?id=93919
47. Läubli, S., Castilho, S., Neubig, G., Sennrich, R., Shen, Q., & Toral, A. (2020). A set of recommendations for assessing human–machine parity in language translation. *Journal of Artificial Intelligence Research*, 67, 653-672 [viewed 15 March 2023]. Retrieved from https://doi.org/10.1613/jair.1.11371
48. Läubli, S., Sennrich, R., & Volk, M. (2018). Has machine translation achieved human parity? a case for document-level evaluation. [viewed 22 March 2023]. Retrieved from https://arxiv.org/abs/1808.07048
49. Lavie, A. (2010). Evaluating the output of machine translation systems. *In Proceedings of the 9th Conference of the Association for Machine Translation in the Americas: Tutorials*. [viewed 19 March 2023]. Retrieved from https://aclanthology.org/2010.amta-tutorials.4.pdf
50. Lin, C. Y. (2004, July). Rouge: A package for automatic evaluation of summaries. *In Text summarization branches out* (pp. 74-81). [viewed 19 March 2023]. Retrieved from https://aclanthology.org/W04-1013.pdf
51. Lommel, A. (2018). Metrics for translation quality assessment: a case for standardising error typologies. *Translation quality assessment: From principles to practice*, 109-127. [viewed 29 March 2023]. Retrieved from https://doi.org/10.1007/978-3-319-91241-7_6
52. Lommel, A., Burchardt, A., & Uszkoreit, H. (2015). Multidimensional Quality Metrics (MQM) Definition. (2015b). [viewed 29 March 2023]. Retrieved from https://web.archive.org/web/20160425060914/http://www.qt21.eu/mqm definition/definition-2015-12-30.html#default-values
53. Lommel, A., Uszkoreit, H., & Burchardt, A. (2014). Multidimensional quality metrics (MQM): A framework for declaring and describing translation quality metrics. *Tradumàtica*, (12), 0455-463. [viewed 29 March 2023]. Retrieved from https://doi.org/10.5565/rev/tradumatica.77
54. Lommel, Ariel, Aljoscha Burchardt, Attila Görög, Hans Uszkoreit, and Alan K. Melby. Multidimensional Quality Metrics (MQM) Issue Types. (2015a). [viewed 18 March 2023]. Retrieved from https://web.archive.org/web/20160505053317/http://www.qt21.eu/mqm-definition/issues-list-2015-12-30.html
55. MQM (Multidimensional Quality Metrics). *The place to go to learn about MQM*. [viewed 18 March 2023]. Retrieved from https://themqm.org

56. Nida, E. A. (1964). *Toward a science of translation*. Leiden: Brill. [viewed 18 March 2023]. Retrieved from https://doi.org/10.1163/9789004495746
57. Papineni, K., Roukos, S., Ward, T., & Zhu, W. J. (2002, July). Bleu: a method for automatic evaluation of machine translation. *In Proceedings of the 40th annual meeting of the Association for Computational Linguistics* (pp. 311-318). [viewed 18 March 2023]. Retrieved from https://aclanthology.org/P02-1040.pdf
58. Petkevičiūtė, I., & Tamulynas, B. (2011). Kompiuterinis vertimas į lietuvių kalbą: alternatyvos ir jų lingvistinis vertinimas. *Kalbų studijos*, (18), 38-45. [viewed 18 March 2023]. Retrieved from https://etalpykla.lituanistika.lt/object/LT-LDB-0001:J.04~2011~1367174892606
59. Pym, A. (2017). *Exploring translation theories*. Routledge. [viewed 22 March 2023]. Retrieved from https://doi.org/10.4324/9781315857633
60. Popović, M. (2020, December). Informative manual evaluation of machine translation output. *In Proceedings of the 28th International Conference on Computational Linguistics* (pp. 5059-5069). [viewed 22 March 2023]. Retrieved from https://doi.org/10.18653/v1/2020.coling-main.444
61. Povilaitienė, M., & Kasperė, R. (2022). MACHINE TRANSLATION FOR POST-EDITING PRACTICES. *Scientific Journal of National Pedagogical Dragomanov University. Series 9. Current Trends in Language Development*, (24), 47-62. [viewed 18 March 2023]. Retrieved from https://doi.org/10.31392/npu-nc.series9.2022.24.04
62. Ralys, D. A. (2017). Mašininis vertimas lietuvių kalbai. *Bendrinė kalba* (iki 2014 metų–Kalbos kultūra), (90), 1-20. [viewed 18 March 2023]. Retrieved from https://www.ceeol.com/search/article-detail?id=605540
63. Rei, R., Farinha, A. C., Zerva, C., van Stigt, D., Stewart, C., Ramos, P., ... & Lavie, A. (2021, November). Are references really needed? unbabel-IST 2021 submission for the metrics shared task. *In Proceedings of the Sixth Conference on Machine Translation* (pp. 1030-1040). [viewed 18 March 2023]. Retrieved from https://aclanthology.org/2021.wmt-1.111/
64. Rimkutė, E., & Kovalevskaitė, J. (2008). Kalbos pusfabrikačiai–sunkus maistas mašininio vertimo sistemos skrandžiui. *Gimtoji kalba*, 3, 3-11. [viewed 18 March 2023]. Retrieved from http://donelaitis.vdu.lt/lkk/pdf/pusfab.pdf
65. Rivera-Trigueros, I. (2022). Machine translation systems and quality assessment: a systematic review. *Language Resources and Evaluation*, 56(2), 593-619. [viewed 18 March 2023]. Retrieved from https://doi.org/10.1007/s10579-021-09537-5
66. Roturier, J. (2006). *An investigation into the impact of controlled English rules on the comprehensibility, usefulness and acceptability of machine-translated technical documentation for French and German users*. Doctoral dissertation. Dublin City University. [viewed 18 March 2023]. Retrieved from https://doras.dcu.ie/18190/
67. Snover, M., Dorr, B., Schwartz, R., Micciulla, L., & Makhoul, J. (2006). A study of translation edit rate with targeted human annotation. *In Proceedings of the 7th Conference of the Association for Machine Translation in the Americas: Technical Papers (*pp. 223-231). [viewed 29 March 2023]. Retrieved from https://aclanthology.org/2006.amta-papers.25/
68. Stankevičiūtė, G., Kasperavičienė, R., & Horbačauskienė, J. (2017). Issues in Machine Translation: A case of mobile apps in the Lithuanian and English language pair. *International Journal on Language, Literature and Culture in Education*, 4(1), 75-88. [viewed 19 March 2023]. Retrieved from https://etalpykla.lituanistikadb.lt/object/LT-LDB-0001:J.04~2017~1550500353375/

69. Stymne, S., & Ahrenberg, L. (2012, April). On the practice of error analysis for machine translation evaluation. *In LREC* (pp. 1785-1790). [viewed 19 March 2023]. Retrieved from https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=311f739c5bb50a783c7f111cc8838bd1ce0c3fa7
70. Su, K. Y., Wu, M. W., & Chang, J. S. (1992). A new quantitative quality measure for machine translation systems. *In COLING 1992 Volume 2: The 14th International Conference on Computational Linguistics*. [viewed 19 March 2023]. Retrieved from https://doi.org/10.3115/992133.992137
71. Sulubacak, U., Caglayan, O., Grönroos, S. A., Rouhe, A., Elliott, D., Specia, L., & Tiedemann, J. (2020). Multimodal machine translation through visuals and speech. *Machine Translation*, 34, 97-147. [viewed 22 March 2023]. Retrieved from https://doi.org/10.1007/s10590-020-09250-0
72. Thielsch, M. T., & Hirschfeld, G. (2019). Facets of website content. *Human–Computer Interaction*, 34(4), 279-327. [viewed 19 March 2023]. Retrieved from https://doi.org/10.1080/07370024.2017.1421954
73. Van Slype, G. (1979). Critical Study of Methods for Evaluating the Quality of Machine Translation (Final Report), Prepared for the Commission of the European Communities. [viewed 25 March 2023]. Retrieved from http://aei.pitt.edu/39751/1/A4102.pdf
74. Vilar, D., Xu, J., Luis Fernando, D. H., & Ney, H. (2006, May). Error Analysis of Statistical Machine Translation Output. *In LREC* (pp. 697-702). [viewed 25 March 2023]. Retrieved from https://www.cs.brandeis.edu/~marc/misc/proceedings/lrec-2006/pdf/413_pdf.pdf
75. White, J. S., O'Connell, T. A., & O'Mara, F. E. (1994). The ARPA MT evaluation methodologies: evolution, lessons, and future approaches. *In Proceedings of the First Conference of the Association for Machine Translation in the Americas*. [viewed 19 March 2023]. Retrieved from https://aclanthology.org/1994.amta-1.25.pdf
76. Williams, M. (2004). *Translation quality assessment: An argumentation-centred approach*. University of Ottawa Press. [viewed 25 March 2023]. Retrieved from https://doi.org/10.7202/013278ar
77. Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., ... & Dean, J. (2016). Google's neural machine translation system: Bridging the gap between human and machine translation. [viewed 25 March 2023]. Retrieved from https://arxiv.org/abs/1609.08144

Appendices

Appendix 1. List of evaluated spam websites

Website No. 1

Website No. 2

Asmenybės psichologija Psichikos sutrikimai Depresinės būsenos Fobijos ir manija

HOUSEPSYCH.COM - APIE PSICHOLOGIJĄ

NAMAI | PSICHOLOGIJA | PSICHIATRIJA | KONTAKTINIAI DUOMENYS | SVETAINĖS SCHEMA

Antraštės

- Nervų sistemos ligos
- Depresinės būsenos
- Vaiko psichologija
- Vaikų psichiatrija
- Konflikto psichologija
- Mąstymo psichologija
- Psichologinės sąvokos
- Psichikos sutrikimai
- Asmenybės psichologija
- Santykių psichologija
- Bendravimo psichologija
- Psichologinė pagalba
- Psichoterapija
- Savivertė
- Fobijos ir manija
- Mūsų skaitytojų istorijos
- Vaistai

Populiarios fobijos

Hemofobija ar kraujo baimė yra nekontroliuojama baimė sunkios panikos lygyje ...

Sociofobija yra neracionali baimė, nekontroliuojama bet kurios visuomenės baimė ...

Ergofobija yra darbo baimė, bet kokių tikslingų veiksmų

Valerijonas (tabletės): vartojimo instrukcijos

Prieš pirkdami valerijoną tabletėmis, rekomenduojame susipažinti su siūlomomis instrukcijomis, kuriose aprašomi šio vaisto vartojimo būdai, dozavimas ir lydimasis poveikis žmogaus organizmui. Reikėtų nepamiršti, kad tik gydytojas gali profesionaliai susieti paciento problemą ir apsiribojimą valerijono vartojimu, todėl nereikėtų savarankiškai gydytis.

Svetainėje „Psichologija ir psichiatrija“ pateikiama visa reikalinga informacija šiais klausimais: aprašyta išleidimo forma ir forma, vartojimo instrukcijos ir indikacijos vartoti valerijoną, kontraindikacijos, rekomenduojama dozė, populiarūs analogai, gydytojų ir pacientų apžvalgos apie šio vaisto vartojimą.

Valerijonas yra unikali raminamoji priemonė, naudojama stresui malšinti. Valerijonas ramina nervų sistemą, todėl gerėja širdies veikla. Valerijonas sugeba pašalinti stipriausio širdies plakimo priepuolius. Taip pat ši priemonė yra efektyvi sprendžiant virškinimo trakto problemą, jei ji yra psichosomatinio pobūdžio.

Website No. 3

Website No. 4

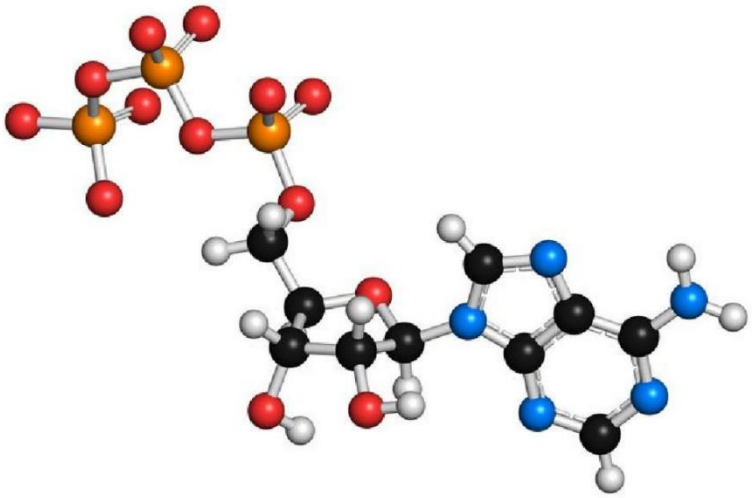
Website No. 5

EFerrit

Mokslas › Chemija

Fosforilinimas ir kaip tai veikia

by Anne Marie Helmenstine, Ph.D.



Oksidacinis, gliukozės ir baltymo fosforilinimas

Fosforilinimas Apibrėžimas

Fosforilinimas yra fosforilo grupės (PO_3^-) cheminis pridėjimas prie **organinės molekulės** . Fosforilo grupės pašalinimas vadinamas defosforiliavimu. Tiek fosforilinimą, tiek defosforilinimą vykdo **fermentai** (pvz., Kinazės, fosfotransferazės). Fosforilinimas yra svarbus biochemijos ir molekulinės biologijos srityse, nes tai pagrindinė baltymų ir fermentų funkcija, cukraus metabolizmas, energijos kaupimas ir išleidimas.

Fosforilinimo tikslai

Fosforilinimas atlieka svarbų reguliavimo vaidmenį ląstelėse. Jo funkcijos apima:

- Svarbu glikolizei
- Naudojamas baltymų ir baltymų sąveikai
- Naudojamas baltymų degradacijai
- Reguliuojamas fermentų slopinimas
- Išlaiko homeostazę reguliuojant chemines reakcijas, reikalingas energijai

Appendix 2. Sample questions of end-user survey

Pateiktos interneto svetainės tekstas yra aiškus.

*

Pasirinkite vieną iš atsakymų:
Pasirinkite **tik vieną** iš pateiktų variantų:

- ◯ 1 – Visiškai sutinku
- ◯ 2 – Sutinku
- ◯ 3 – Nei sutinku, nei nesutinku
- ◯ 4 – Nesutinku
- ◯ 5 – Visiškai nesutinku

Tekstą yra malonu skaityti.

*

Pasirinkite vieną iš atsakymų:
Pasirinkite **tik vieną** iš pateiktų variantų:

- ◯ 1 – Visiškai sutinku
- ◯ 2 – Sutinku
- ◯ 3 – Nei sutinku, nei nesutinku
- ◯ 4 – Nesutinku
- ◯ 5 – Visiškai nesutinku

Pateiktos interneto svetainės puslapis yra informatyvus.

*

Pasirinkite vieną iš atsakymų:
Pasirinkite **tik vieną** iš pateiktų variantų:

- ◯ 1 – Visiškai sutinku
- ◯ 2 – Sutinku
- ◯ 3 – Nei sutinku, nei nesutinku
- ◯ 4 – Nesutinku
- ◯ 5 – Visiškai nesutinku

Pasitikiu svetainėje pateikta informacija.

*

Pasirinkite vieną iš atsakymų:
Pasirinkite **tik vieną** iš pateiktų variantų:

- ◯ 1 – Visiškai sutinku
- ◯ 2 – Sutinku
- ◯ 3 – Nei sutinku, nei nesutinku
- ◯ 4 – Nesutinku
- ◯ 5 – Visiškai nesutinku

Tekstas yra taisyklingas, jo kokybė puiki.

*

Pasirinkite vieną iš atsakymų:
Pasirinkite **tik vieną** iš pateiktų variantų:

- ◯ 1 – Visiškai sutinku
- ◯ 2 – Sutinku
- ◯ 3 – Nei sutinku, nei nesutinku
- ◯ 4 – Nesutinku
- ◯ 5 – Visiškai nesutinku

Appendix 3. Evaluation of translation segments using MQM error typology

Source	Target	Suggested Target	Error Category	Error Subcategory	Severity
Low Height Pancake Hydraulic Cylinder Used In Auger Boring Steerable Rock System	**Mažo aukščio blynų hidraulinis cilindras, naudojamas sraigtinio gręžimo valdomoje uolienų sistemoje**	Žemo profilio hidraulinis stūmimo cilindras skirtas valdomajai sraigtinei uolienų gręžimo sistemai	Accuracy	Mistranslation	Critical
Low Height Pancake Hydraulic Cylinder Used In Auger Boring Steerable Rock System	**Mažo aukščio blynų hidraulinis cilindras**, naudojamas sraigtinio gręžimo valdomoje uolienų sistemoje	Žemo profilio hidraulinis stūmimo cilindras skirtas valdomajai sraigtinei uolienų gręžimo sistemai	Terminology	Inconsistent use of terminology	Critical
50-ton low height pancake hydraulic cylinders used in auger boring steerable rock system	**50 tonų žemo aukščio blynų hidrauliniai cilindrai, naudojami sraigtinio gręžimo valdomų uolienų sistemoje**.	50 tonų žemo profilio hidrauliniai stūmimo cilindrai naudojami gręžimo įrenginiuose su valdomąja sraigtine uolienų gręžimo sistema.	Accuracy	Mistranslation	Critical
50-ton low height pancake hydraulic cylinders used in auger boring steerable rock system	50 tonų **žemo aukščio blynų hidrauliniai cilindrai**, naudojami sraigtinio gręžimo valdomų uolienų sistemoje.	50 tonų žemo profilio hidrauliniai stūmimo cilindrai naudojami gręžimo įrenginiuose su valdomąja sraigtine uolienų gręžimo sistema.	Terminology	Inconsistent use of terminology	Critical
50-ton low height pancake hydraulic cylinders used in auger boring steerable rock system	50 tonų žemo aukščio blynų hidrauliniai cilindrai, naudojami **sraigtinio gręžimo valdomų uolienų sistemoje**.	50 tonų žemo profilio hidrauliniai stūmimo cilindrai naudojami gręžimo įrenginiuose su valdomąja sraigtine uolienų gręžimo sistema.	Terminology	Inconsistent use of terminology	Critical
Steerable head designed for boring through sand or soft rock with real-time steering adjustments made out of the pit with a hydraulic control system.	**Vairavimo galvutė**, skirta gręžti smėlį ar minkštą uolieną, realiuoju laiku reguliuojant vairavimą iš duobės su hidrauline valdymo sistema.	Hidraulinės valdymo sistemos pagalba, smėlio ar minkštų uolienų gręžimui sukurtą antgalį galima valdyti realiu laiku, jam esant gręžinyje.	Terminology	Inconsistent use of terminology	Major
Steerable head designed for boring through sand or soft rock with real-time steering adjustments made out of the pit with a hydraulic control system.	Vairavimo galvutė, skirta gręžti smėlį ar minkštą uolieną, realiuoju laiku **reguliuojant vairavimą iš duobės su hidrauline valdymo sistema**.	Hidraulinės valdymo sistemos pagalba, smėlio ar minkštų uolienų gręžimui sukurtą antgalį galima valdyti realiu laiku, jam esant gręžinyje.	Accuracy	Mistranslation	Major
The Steerable Rock System (SRS) is the auger boring market's first steerable head designed to navigate not only solid rock but difficult fractured rock conditions as well.	Steerable Rock System (SRS) yra pirmoji sraigtinių **gręžinių** rinkoje valdoma galvutė, skirta važiuoti ne tik kietomis uolienomis, bet ir sudėtingomis skilimo sąlygomis.	Valdomoji uolienų gręžimo sistema (angl. Steerable Rock System, SRS) – tai pirmoji tokio tipo sistema sraigtinių **grąžtų** rinkoje. Valdomasis gręžimo antgalis sukurtas valdyti gręžimą ne tik kietose uolienose, bet dirbti ir sudėtingomis sąlygomis, kuomet uolienos skilinėja.	Terminology	Inconsistent use of terminology	Major
The Steerable Rock System (SRS) is the auger boring market's first steerable head designed to navigate not only solid rock but difficult fractured rock conditions as well.	Steerable Rock System (SRS) **yra pirmoji sraigtinių gręžinių rinkoje valdoma galvutė, skirta važiuoti ne tik kietomis uolienomis, bet ir sudėtingomis skilimo sąlygomis**.	Valdomoji uolienų gręžimo sistema (angl. Steerable Rock System, SRS) – tai pirmoji tokio tipo sistema sraigtinių grąžtų rinkoje. Valdomasis gręžimo antgalis sukurtas valdyti gręžimą ne tik kietose uolienose, bet dirbti ir sudėtingomis sąlygomis,	Accuracy	Mistranslation	Major

		kuomet uolienos skilinėja.			
Engineered to operate in rock up to 25,000 psi, the SRS allows operators to keep bores online and grade even in the toughest ground conditions for an on-grade bore.	Sukurtas veikti uolienose iki **25,000 4 psi**, SRS leidžia operatoriams išlaikyti gręžinius tinkle ir greiderį net sunkiausiomis žemės sąlygomis.	Valdomoji uolienų gręžimo sistema (sukurta darbui uolienose iki 1724 ba slėgio) leidžia gręžimo operatoriams matyti gręžiamą ertmę ekrane net gręžiant žemę sunkiausiomis sąlygomis.	Locale convention	Measurement format	Minor
Engineered to operate in rock up to 25,000 psi, the SRS allows operators to keep bores online and grade even in the toughest ground conditions for an on-grade bore.	Sukurtas veikti uolienose iki 25,000 4 psi, SRS leidžia operatoriams išlaikyti gręžinius tinkle ir greiderį net sunkiausiomis žemės sąlygomis.	Valdomoji uolienų gręžimo sistema (sukurta darbui uolienose iki 1724 ba slėgio) **leidžia gręžimo operatoriams matyti gręžiamą ertmę ekrane** net gręžiant žemę sunkiausiomis sąlygomis.	Accuracy	Under-translation	Major
Engineered to operate in rock up to 25,000 psi, the SRS allows operators to keep bores online and grade even in the toughest ground conditions for an on-grade bore.	Sukurtas veikti uolienose iki 25,000 **4** psi, SRS leidžia operatoriams išlaikyti gręžinius tinkle ir greiderį net sunkiausiomis žemės sąlygomis.	Valdomoji uolienų gręžimo sistema (sukurta darbui uolienose iki 1724 ba slėgio) leidžia gręžimo operatoriams matyti gręžiamą ertmę ekrane net gręžiant žemę sunkiausiomis sąlygomis.	Accuracy	Addition	Major
On the drilling head, there is peace's of small round pipe, that's where the cylinders sit.	Ant **gręžimo galvutės yra mažo apvalaus vamzdžio ramstis**, ten yra cilindrai.	Ant vamzdžio, juosiančio gręžimo antgalį, išdėstyti cilindrai.	Accuracy	Mistranslation	Minor
There are 4 pcs 50-ton pancake hydraulic cylinders on each side(top, bottom, right, and left) of the steerable head for a total of 16pcs, it works when it's drilling horizontally underground we use the cylinders to navigate the drilling head.	Kiekvienoje valdomos galvutės pusėje (viršuje, apačioje, dešinėje ir kairėje) yra 50 vnt 16 tonų blynų hidrauliniai cilindrai, iš viso XNUMX vnt., jis veikia gręžiant horizontaliai po žeme, cilindrus naudojame gręžimo galvutei naršyti.	Kiekvienoje valdomojo gręžimo antgalio dalyje (viršuje, apačioje, kairėje ir dešinėje) yra **po 4 vnt.** 50 tonų hidraulinių stūmimo cilindrų, viso, 16 cilindrų, kurie gręžiant po žeme horizontalia kryptimi naudojami gręžimo antgalio valdymui.	Accuracy	Mistranslation	Major
There are 4 pcs 50-ton pancake hydraulic cylinders on each side(top, bottom, right, and left) of the steerable head for a total of 16pcs, it works when it's drilling horizontally underground we use the cylinders to navigate the drilling head.	Kiekvienoje valdomos galvutės pusėje (viršuje, apačioje, dešinėje ir kairėje) yra 50 vnt 16 tonų blynų hidrauliniai cilindrai, iš viso XNUMX vnt., jis veikia gręžiant horizontaliai po žeme, cilindrus naudojame gręžimo galvutei naršyti.	Kiekvienoje **valdomojo** gręžimo antgalio dalyje (viršuje, apačioje, kairėje ir dešinėje) yra po 4 vnt. 50 tonų hidraulinių stūmimo cilindrų, viso, 16 cilindrų, kurie gręžiant po žeme horizontalia kryptimi naudojami gręžimo antgalio valdymui.	Fluency	Grammar	Minor
There are 4 pcs 50-ton pancake hydraulic cylinders on each side(top, bottom, right, and left) of the steerable head for a total of 16pcs, it works when it's drilling horizontally underground we use the cylinders to navigate the drilling head.	Kiekvienoje valdomos galvutės pusėje (viršuje, apačioje, dešinėje ir kairėje) yra 50 **vnt** 16 tonų blynų hidrauliniai cilindrai, iš viso XNUMX vnt., jis veikia gręžiant horizontaliai po žeme, cilindrus naudojame gręžimo galvutei naršyti.	Kiekvienoje valdomojo gręžimo antgalio dalyje (viršuje, apačioje, kairėje ir dešinėje) yra po 4 vnt. 50 tonų hidraulinių stūmimo cilindrų, viso, 16 cilindrų, kurie gręžiant po žeme horizontalia kryptimi naudojami gręžimo antgalio valdymui.	Fluency	Punctuation	Minor
There are 4 pcs 50-ton pancake hydraulic cylinders on each side(top, bottom, right, and left) of the steerable head for a total of 16pcs, it works when it's drilling horizontally underground we use the cylinders to navigate the drilling head.	Kiekvienoje valdomos galvutės pusėje (viršuje, apačioje, dešinėje ir kairėje) yra 50 vnt 16 tonų **blynų hidrauliniai cilindrai**, iš viso XNUMX vnt., jis veikia gręžiant horizontaliai po žeme, cilindrus naudojame gręžimo galvutei naršyti.	Kiekvienoje valdomojo gręžimo antgalio dalyje (viršuje, apačioje, kairėje ir dešinėje) yra po 4 vnt. 50 tonų hidraulinių stūmimo cilindrų, viso, 16 cilindrų, kurie gręžiant po žeme horizontalia kryptimi naudojami gręžimo antgalio valdymui.	Terminology	Inconsistent use of terminology	Major

There are 4 pcs 50-ton pancake hydraulic cylinders on each side(top, bottom, right, and left) of the steerable head for a total of 16pcs, it works when it's drilling horizontally underground we use the cylinders to navigate the drilling head.	Kiekvienoje valdomos galvutės pusėje (viršuje, apačioje, dešinėje ir kairėje) yra 50 vnt 16 tonų blynų hidrauliniai cilindrai, iš viso **XNUMX vnt**., jis veikia gręžiant horizontaliai po žeme, cilindrus naudojame gręžimo galvutei naršyti.	Kiekvienoje valdomojo gręžimo antgalio dalyje (viršuje, apačioje, kairėje ir dešinėje) yra po 4 vnt. 50 tonų hidraulinių stūmimo cilindrų, viso, 16 cilindrų, kurie gręžiant po žeme horizontalia kryptimi naudojami gręžimo antgalio valdymui.	Accuracy	Mistranslation	Major
There are 4 pcs 50-ton pancake hydraulic cylinders on each side(top, bottom, right, and left) of the steerable head for a total of 16pcs, it works when it's drilling horizontally underground we use the cylinders to navigate the drilling head.	Kiekvienoje valdomos galvutės pusėje (viršuje, apačioje, dešinėje ir kairėje) yra 50 vnt 16 tonų blynų hidrauliniai cilindrai, iš viso XNUMX vnt., jis veikia gręžiant horizontaliai po žeme, **cilindrus naudojame gręžimo galvutei naršyti.**	Kiekvienoje valdomojo gręžimo antgalio dalyje (viršuje, apačioje, kairėje ir dešinėje) yra po 4 vnt. 50 tonų hidraulinių stūmimo cilindrų, viso, 16 cilindrų, kurie gręžiant po žeme horizontalia kryptimi naudojami gręžimo antgalio valdymui.	Accuracy	Mistranslation	Critical
Valerian (tablets) - instructions for use, reviews, analogues	**Valerijonas (tabletės) - naudojimo instrukcijos, apžvalgos, analogia**	Valerijono tablečių naudojimo instrukcija, vertinimas, sudėtis	Accuracy	Mistranslation	Major
Valerian (tablets) - instructions for use, reviews, analogues	Valerijonas (tabletės) - naudojimo instrukcijos, apžvalgos, analogia	Valerijono tabletėmis naudojimo instruk**cija**, vertinimas, sudėtis	Fluency	Grammar	Minor
Valerian (tablets) - instructions for use, reviews, analogues	Valerijonas (tabletės) - naudojimo instrukcijos, apžvalgos, **analogia**	Valerijono tabletėmis naudojimo instrukcija, vertinimas, sudėtis	Terminology	Inconsistent use of terminology	Major
Before you buy Valerian in tablets, we recommend that you familiarize yourself with the proposed instructions, which describe the methods of application, dosage, and the concomitant effect of this drug on the human body.	Prieš pirkdami valerijoną tabletėmis, rekomenduojame susipažinti su siūlomomis instrukcijomis, kuriose aprašomi šio vaisto vartojimo būdai, dozavimas ir lydimasis poveikis žmogaus organizmui.	Prieš įsigyjant valerijono tabletes, rekomenduojame susipažinti su vartojimo instrukcija, kurioje aprašomi vartojimo būdai, dozavimas ir poveikis žmogaus organizmui.	Style	Awkward	Minor
It should be borne in mind that only a doctor can professionally correlate the patient's problem and the limitation to taking Valerian, therefore self-medication should not be carried away.	Reikėtų nepamiršti, kad tik gydytojas gali **profesionaliai susieti paciento problemą ir apsiribojimą valerijono vartojimu**, todėl nereikėtų savarankiškai gydytis.	Prisiminkite, kad vaisto vartojimo trukmę, atsižvelgęs į jūsų sveikatos būklę, gali nustatyti tik gydytojas, tad vaisto vartoti savarankiškai nederėtų.	Accuracy	Mistranslation	Minor
The site "Psychology and Psychiatry" offers all the necessary information on the following issues: the composition and form of release, instructions for use and indications for taking Valerian, contraindications, recommended dosage, popular analogues, reviews of doctors and patients about taking this drug are described.	Svetainėje „Psichologija ir psichiatrija" pateikiama visa reikalinga informacija šiais klausimais: aprašyta išleidimo forma ir forma, vartojimo instrukcijos ir indikacijos vartoti valerijoną, kontraindikacijos, rekomenduojama dozė, populiarūs analogai, gydytojų ir pacientų apžvalgos apie šio vaisto vartojimą.	Daugiau informacijos apie vaisto sudėtį ir veikliąsias medžiagas, atsargumo priemones, vaisto tinkamumą ar netinkamumą, rekomenduotiną dozavimą, dažniausiai sutinkamus vaisto pakaitalus, vartojimo instrukcijos, gydytojų ir pacientų atsiliepimai pateikiami tinklalapyje: „Psichologija ir psichiatrija".	Style	Awkward	Minor
The site "Psychology and Psychiatry" offers all the necessary information on the following issues: the composition and form of release, instructions for use and indications for taking Valerian, contraindications, recommended dosage,	Svetainėje „Psichologija ir psichiatrija" pateikiama visa reikalinga informacija šiais klausimais: aprašyta išleidimo forma ir forma, vartojimo instrukcijos ir indikacijos vartoti valerijoną, kontraindikacijos,	Daugiau informacijos apie vaisto sudėtį ir veikliąsias medžiagas, atsargumo priemones, vaisto tinkamumą ar netinkamumą, rekomenduotiną dozavimą, dažniausiai sutinkamus vaisto pakaitalus, vartojimo	Accuracy	Mistranslation	Major

popular analogues, reviews of doctors and patients about taking this drug are described.	rekomenduojama dozė, populiarūs analogai, gydytojų ir pacientų apžvalgos apie šio vaisto vartojimą.	instrukcijos, gydytojų ir pacientų atsiliepimai pateikiami tinklalapyje: „Psichologija ir psichiatrija“.			
Valerian is a unique sedative used to alleviate stress.	Valerijonas yra unikali raminamoji priemonė, naudojama stresui malšinti.	Valerijonas yra išskirtinis raminamasis vaistas, naudojamas streso malšinimui.	Style	Awkward	Minor
Valerian has a calming effect on the nervous system, which improves cardiac function.	Valerijonas ramina nervų sistemą, **todėl gerėja širdies veikla.**	Valerijonas pasižymi raminančiu poveikiu nervų sistemai bei gerina širdies darbą.	Fluency	Grammar	Neutral
Valerian has a calming effect on the nervous system, which improves cardiac function.	Valerijonas ramina nervų sistemą, todėl gerėja širdies veikla.	Valerijonas pasižymi raminančiu poveikiu nervų sistemai bei gerina širdies darbą.	Style	Awkward	Minor
Valerian is able to remove attacks of the strongest heartbeat.	Valerijonas sugeba pašalinti **stipriausio širdies plakimo** priepuolius.	Vaisto vartojimo metu, gali išnykti stiprūs širdies plakimo priepuoliai.	Fluency	Grammar	Neutral
Valerian is able to remove attacks of the strongest heartbeat.	**Valerijonas sugeba pašalinti** stipriausio širdies plakimo priepuolius.	Vaisto vartojimo metu, gali išnykti stiprūs širdies plakimo priepuoliai.	Accuracy	Mistranslation	Minor
Also, this tool is effective in solving the problem with the gastrointestinal tract, if it is of a psychosomatic nature.	Taip pat ši priemonė yra efektyvi **sprendžiant virškinimo trakto problemą**, jei ji yra psichosomatinio pobūdžio.	Be to, priemonė veiksminga sprendžiant psichosomatinio pobūdžio virškinamojo trakto problemas.	Fluency	Grammar	Minor
Also, this tool is effective in solving the problem with the gastrointestinal tract, if it is of a psychosomatic nature.	Taip pat ši priemonė yra efektyvi sprendžiant virškinimo trakto problemą, jei ji yra psichosomatinio pobūdžio.	Be to, priemonė veiksminga sprendžiant psichosomatinio pobūdžio virškinamojo trakto problemas.	Style	Awkward	Minor
Pharmacological group: sedatives.	Farmakologinė grupė: raminamieji.	Farmakologinė grupė: raminamieji preparatai.	Accuracy	Under-translation	Minor
Active ingredient: Valerian medicinal rhizome with roots (Valerianae officinalis rhizomata cum radicibus).	Veiklioji medžiaga: valerijono šakniastiebis su šaknimis (Valerianae officinalis rhizomata cum radicibus).	Veiklioji medžiaga: vaistinių valerijonų (lot. Valerianae officinalis rhizomata cum radicibus) šaknų sausasis ekstraktas.	Terminology	Inconsistent use of terminology	Minor
Valerian - composition and form of release	Valerijonas - išsiskyrimo sudėtis ir forma	Valerijono sudėtis ir tablečių išvaizda	Terminology	Inconsistent use of terminology	Critical
Valerian - composition and form of release	Valerijonas - išsiskyrimo sudėtis ir forma	Valerijono sudėtis ir tablečių išvaizda	Fluency	Punctuation	Neutral
Dosage form - film-coated tablets: biconvex round, yellow with a greenish tint; the core is cross-sectional from light to gray-brown in color with a greenish tint (in a cardboard bundle of 1–5 contour cell packs of 10, 25 or 50 pieces, or 1 can, a bottle or polymer container of 30, 40, 50 or 100 pieces .).	Dozavimo forma - plėvele dengtos tabletės: abipus išgaubtos, apvalios, geltonos, žalsvo atspalvio; šerdis yra skerspjūvio nuo šviesiai iki pilkai rudos spalvos su žalsvu atspalviu (kartoniniame ryšulyje, kuriame yra 1–5 kontūro elementų pakuotės po 10, 25 arba 50 vienetų, arba 1 skardinėje, buteliuke ar polimerinėje talpykloje yra 30, 40, 50 arba 100 vienetų) .).	Vaisto išvaizda: plėvele dengtos tabletės yra apvalios, išgaubtos, geltonos spalvos su žaliu atspalviu. Tabletės spalva po apvalkalu gali būti nuo šviesiai iki tamsiai pilkai rudos spalvos su žalsvu atspalviu. Pakuotė: kartoninėje dėžutėje yra 1 – 5 lapeliai po 10, 25 ar 50 tablečių, stikliniame ar plastikiniame buteliuke – 30, 40, 50 arba 100 tablečių.	Accuracy	Mistranslation	Major
Dosage form - film-coated tablets: biconvex round, yellow with a greenish tint; the core is cross-sectional from light to gray-brown in color	Dozavimo forma - plėvele dengtos tabletės: abipus išgaubtos, apvalios, geltonos, žalsvo atspalvio; šerdis yra skerspjūvio nuo šviesiai	Vaisto išvaizda: plėvele dengtos tabletės yra apvalios, išgaubtos, geltonos spalvos su žaliu atspalviu. Tabletės spalva po apvalkalu gali būti nuo	Terminology	Inconsistent use of terminology	Major

with a greenish tint (in a cardboard bundle of 1–5 contour cell packs of 10, 25 or 50 pieces, or 1 can, a bottle or polymer container of 30, 40, 50 or 100 pieces .).	iki pilkai rudos spalvos su žalsvu atspalviu (**kartoniniame ryšulyje**, kuriame yra 1–5 kontūro elementų pakuotės po 10, 25 arba 50 vienetų, arba 1 skardinėje, buteliuke ar polimerinėje talpykloje yra 30, 40, 50 arba 100 vienetų) .).	šviesiai iki tamsiai pilkai rudos spalvos su žalsvu atspalviu. Pakuotė: kartoninėje dėžutėje yra 1 – 5 lapeliai po 10, 25 ar 50 tablečių, stikliniame ar plastikiniame buteliuke – 30, 40, 50 arba 100 tablečių.			
Dosage form - film-coated tablets: biconvex round, yellow with a greenish tint; the core is cross-sectional from light to gray-brown in color with a greenish tint (in a cardboard bundle of 1–5 contour cell packs of 10, 25 or 50 pieces, or 1 can, a bottle or polymer container of 30, 40, 50 or 100 pieces .).	Dozavimo forma - plėvele dengtos tabletės: abipus išgaubtos, apvalios, geltonos, žalsvo atspalvio; šerdis yra skerspjūvio nuo šviesiai iki pilkai rudos spalvos su žalsvu atspalviu (kartoniniame ryšulyje, kuriame yra 1–5 kontūro elementų pakuotės po 10, 25 arba 50 vienetų, arba 1 **skardinėje**, buteliuke ar polimerinėje talpykloje yra 30, 40, 50 arba 100 vienetų) .).	Vaisto išvaizda: plėvele dengtos tabletės yra apvalios, išgaubtos, geltonos spalvos su žaliu atspalviu. Tabletės spalva po apvalkalu gali būti nuo šviesiai iki tamsiai pilkai rudos spalvos su žalsvu atspalviu. Pakuotė: kartoninėje dėžutėje yra 1 – 5 lapeliai po 10, 25 ar 50 tablečių, stikliniame ar plastikiniame buteliuke – 30, 40, 50 arba 100 tablečių.	Terminology	Inconsistent use of terminology	Major
Dosage form - film-coated tablets: biconvex round, yellow with a greenish tint; the core is cross-sectional from light to gray-brown in color with a greenish tint (in a cardboard bundle of 1–5 contour cell packs of 10, 25 or 50 pieces, or 1 can, a bottle or polymer container of 30, 40, 50 or 100 pieces .).	Dozavimo forma - plėvele dengtos tabletės: abipus išgaubtos, apvalios, geltonos, žalsvo atspalvio; šerdis yra skerspjūvio nuo šviesiai iki pilkai rudos spalvos su žalsvu atspalviu (kartoniniame ryšulyje, kuriame yra 1–5 kontūro elementų pakuotės po 10, 25 arba 50 vienetų, arba 1 skardinėje, buteliuke ar **polimerinėje talpykloje** yra 30, 40, 50 arba 100 vienetų) .).	Vaisto išvaizda: plėvele dengtos tabletės yra apvalios, išgaubtos, geltonos spalvos su žaliu atspalviu. Tabletės spalva po apvalkalu gali būti nuo šviesiai iki tamsiai pilkai rudos spalvos su žalsvu atspalviu. Pakuotė: kartoninėje dėžutėje yra 1 – 5 lapeliai po 10, 25 ar 50 tablečių, stikliniame ar plastikiniame buteliuke – 30, 40, 50 arba 100 tablečių.	Terminology	Inconsistent use of terminology	Major
Composition 1 tablet: active substance:	1 tabletės sudėtis: veiklioji medžiaga:	Vienos tabletės sudėtis ir veikliosios medžiagos:	Fluency	Punctuation	Neutral
Composition 1 tablet: active substance:	1 tabletės sudėtis**:** veiklioji medžiaga:	Vienos tabletės sudėtis ir veikliosios medžiagos:	Fluency	Grammar	Minor
- thick valerian extract - 20 mg (the sum of carboxylic acid esters in terms of valerianic acid ethyl ester - 1.8%);	- storas valerijono ekstraktas - 20 mg (karboksirūgšties esterių suma, skaičiuojant nuo valerijono rūgšties etilo esterio, - 1,8%);	- tirštas valerijonų ekstraktas - 20 mg (karboksirūgšties esterių suma valerijono rūgšties etilo esterio atžvilgiu - 1,8 %);	Terminology	Inconsistent use of terminology	Critical
- thick valerian extract - 20 mg (the sum of carboxylic acid esters in terms of valerianic acid ethyl ester - 1.8%);	- storas valerijono ekstraktas - 20 mg (karboksirūgšties esterių suma, skaičiuojant nuo valerijono rūgšties etilo esterio, - 1,8**%**);	- tirštas valerijonų ekstraktas - 20 mg (karboksirūgšties esterių suma valerijono rūgšties etilo esterio atžvilgiu - 1,8 %);	Fluency	Spelling	Neutral
- thick valerian extract - 20 mg (the sum of carboxylic acid esters in terms of valerianic acid ethyl ester - 1.8%);	- storas valerijono ekstraktas - 20 mg (karboksirūgšties esterių suma, skaičiuojant nuo valerijono rūgšties etilo esterio, - 1,8%);	- tirštas valerijonų ekstraktas - 20 mg (karboksirūgšties esterių suma valerijono rūgšties etilo esterio atžvilgiu - 1,8 %);	Fluency	Punctuation	Neutral
- thick valerian extract - 20 mg (the sum of carboxylic acid esters in terms of valerianic acid ethyl ester - 1.8%);	- storas valerijono ekstraktas - 20 mg (karboksirūgšties esterių suma, skaičiuojant nuo valerijono rūgšties etilo esterio**,** - 1,8%);	- tirštas valerijonų ekstraktas - 20 mg (karboksirūgšties esterių suma valerijono rūgšties etilo esterio atžvilgiu - 1,8 %);	Fluency	Punctuation	Neutral
- auxiliary components: lactose monohydrate - 43.11 mg; magnesium hydroxycarbonate - 20	- **pagalbiniai komponentai**: laktozės monohidratas - 43,11 mg; magnio	- pagalbinės medžiagos: 143,11 mg laktozės monohidrato, 20 mg magnio	Accuracy	Mistranslation	Minor

mg; sucrose - 5 mg; potato starch - 10.84 mg; polysorbate 80 - 0.05 mg; calcium stearate monohydrate - 1 mg;	hidroksikarbonatas - 20 mg; sacharozė - 5 mg; bulvių krakmolas - 10,84 mg; polisorbatas 80 - 0,05 mg; kalcio stearato monohidratas - 1 mg;	hidroksikarbonato, 5 mg sacharozės, 10,84 mg bulvių krakmolo, 0,05 mg polisorbato 80, 1 mg kalcio stearato monohidrato;			
- auxiliary components: lactose monohydrate - 43.11 mg; magnesium hydroxycarbonate - 20 mg; sucrose - 5 mg; potato starch - 10.84 mg; polysorbate 80 - 0.05 mg; calcium stearate monohydrate - 1 mg;	- pagalbiniai komponentai: laktozės monohidratas - 43,11 mg; magnio hidroksikarbonatas - 20 mg; sacharozė - 5 mg; bulvių krakmolas - 10,84 mg; polisorbatas 80 - 0,05 mg; kalcio stearato monohidratas - 1 mg;	- pagalbinės medžiagos: 143,11 mg laktozės monohidrato, 20 mg magnio hidroksikarbonato, 5 mg sacharozės, 10,84 mg bulvių krakmolo, 0,05 mg polisorbato 80, 1 mg kalcio stearato monohidrato;	Fluency	Grammar	Minor
- auxiliary components: lactose monohydrate - 43.11 mg; magnesium hydroxycarbonate - 20 mg; sucrose - 5 mg; potato starch - 10.84 mg; polysorbate 80 - 0.05 mg; calcium stearate monohydrate - 1 mg;	- pagalbiniai komponentai: laktozės monohidratas - 43,11 mg; magnio hidroksikarbonatas - 20 mg; sacharozė - 5 mg; bulvių krakmolas - 10,84 mg; polisorbatas 80 - 0,05 mg; kalcio stearato monohidratas - 1 mg;	- pagalbinės medžiagos: 143,11 mg laktozės monohidrato, 20 mg magnio hidroksikarbonato, 5 mg sacharozės, 10,84 mg bulvių krakmolo, 0,05 mg polisorbato 80, 1 mg kalcio stearato monohidrato;	Fluency	Punctuation	Neutral
- shell: Opadry II yellow - 7 mg.	- apvalkalas: Opadry II geltonasis - 7 mg.	- apvalkalas: 7 mg Opadry II geltonasis.	Fluency	Grammar	Neutral
Pharmacological action of Valerian	Valerijono farmakologinis poveikis	Farmakologinis valerijono poveikis	Fluency	Grammar	Neutral
The therapeutic effect of Valerian is due to the complex of substances contained in it, and above all, essential oil and alkaloids.	Terapinį valerijono poveikį lemia jame esančių medžiagų kompleksas, o svarbiausia - eterinis aliejus ir alkaloidai.	**Gydomąjį valerijono poveikį** lemia jame esančių medžiagų kompleksas, ypač, eterinis aliejus ir alkaloidai.	Accuracy	Mistranslation	Critical
The therapeutic effect of Valerian is due to the complex of substances contained in it, and above all, essential oil and alkaloids.	Terapinį valerijono poveikį lemia jame esančių medžiagų kompleksas, o svarbiausia - eterinis aliejus ir alkaloidai.	Gydomąjį valerijono poveikį lemia jame esančių medžiagų kompleksas, ypač, eterinis aliejus ir alkaloidai.	Fluency	Punctuation	Neutral
In rhizomes and roots, the amount of essential oil reaches 2%.	Šakniastiebiuose ir šaknyse eterinio aliejaus kiekis siekia 2%.	Eterinio aliejaus kiekis šakniastiebiuose ir šaknyse siekia 2 %.	Fluency	Grammar	Neutral
In rhizomes and roots, the amount of essential oil reaches 2%.	Šakniastiebiuose ir šaknyse eterinio aliejaus kiekis siekia **2%**.	Eterinio aliejaus kiekis šakniastiebiuose ir šaknyse siekia 2 %.	Fluency	Spelling	Neutral
The main components of the oil: bornyl isovalerate, isovalerianic acid, borneol, camphene, α-pinene, limonene, etc.	Pagrindiniai aliejaus komponentai: borniloizovalatas, izovalerianinė rūgštis, borneolis, kafelenas, α-pinenas, limonenas ir kt.	Pagrindinės aliejaus **sudedamosios dalys**: bornilo izovaleratas, izovaleriano rūgštis, borneolis, kamfenas, α-pinenas, limonenas ir kt.	Accuracy	Mistranslation	Minor
The main components of the oil: bornyl isovalerate, isovalerianic acid, borneol, camphene, α-pinene, limonene, etc.	Pagrindiniai aliejaus komponentai: borniloizovalatas, izovalerianinė rūgštis, borneolis, kafelenas, α-pinenas, limonenas ir kt.	Pagrindinės aliejaus sudedamosios dalys: **bornilo izovaleratas**, izovaleriano rūgštis, borneolis, kamfenas, α-pinenas, limonenas ir kt.	Terminology	Inconsistent use of terminology	Critical
The main components of the oil: bornyl isovalerate, isovalerianic acid, borneol, camphene, α-pinene, limonene, etc.	Pagrindiniai aliejaus komponentai: borniloizovalatas, izovalerianinė rūgštis, borneolis, kafelenas, α-pinenas, limonenas ir kt.	Pagrindinės aliejaus sudedamosios dalys: bornilo izovaleratas, **izovaleriano rūgštis**, borneolis, kamfenas, α-pinenas, limonenas ir kt.	Terminology	Inconsistent use of terminology	Critical
The main components of the oil: bornyl isovalerate, isovalerianic acid, borneol, camphene, α-pinene, limonene, etc.	Pagrindiniai aliejaus komponentai: borniloizovalatas, izovalerianinė rūgštis, borneolis, kafelenas, α-pinenas, limonenas ir kt.	Pagrindinės aliejaus sudedamosios dalys: bornilo izovaleratas, izovaleriano rūgštis, borneolis, **kamfenas**, α-pinenas, limonenas ir kt.	Terminology	Inconsistent use of terminology	Critical

The drug has a sedative effect, which manifests itself rather slowly, but stably.	Vaistas turi raminamąjį poveikį, kuris pasireiškia gana lėtai, bet stabiliai.	**Vaistas pasižymi raminamuoju poveikiu**, kuris pasireiškia gana lėtai, **tačiau yra pastovus.**	Accuracy	Mistranslation	Minor
Valerian has a multilateral effect on the body, inhibits the central nervous system, lowers its excitability, and facilitates the onset of natural sleep.	Valerijonas daro daugiašalį poveikį kūnui, slopina centrinę nervų sistemą, mažina jo jaudrumą ir palengvina natūralaus miego pradžią.	Valerijonas turi daugialypį poveikį organizmui, ramina centrinę nervų sistemą, mažina jos jaudrumą, palengvina natūralų miegą.	Accuracy	Mistranslation	Minor
Valerian also has a weak antispasmodic effect, is able to relieve mild cramps.	Valerijonas taip pat pasižymi silpnu antispazminiu poveikiu, geba palengvinti lengvus mėšlungį.	Valerijonas taip pat turi silpną antispazminį poveikį, gali palengvinti nestiprų mėšlungį.	Fluency	Grammar	Minor
Valerian also has a weak antispasmodic effect, is able to relieve mild cramps.	Valerijonas taip pat pasižymi silpnu antispazminiu poveikiu, geba palengvinti lengvus mėšlungį.	Valerijonas taip pat turi silpną antispazminį poveikį, gali palengvinti nestiprų mėšlungį.	Accuracy	Mistranslation	Minor
The muscle tissue of the organs of the urinary and digestive systems after taking Valerian relaxes, since the drug enhances the secretory activity of the gastrointestinal tract.	Šlapimo ir virškinimo sistemos organų raumeninis audinys išgėrus valerijono, atsipalaiduoja, nes vaistas sustiprina virškinimo trakto sekrecinį aktyvumą.	Vaistas sustiprina virškinamojo trakto sekrecinę veiklą, todėl išgėrus valerijono, atsipalaiduoja šlapimo ir virškinamojo trakto sistemos raumenų audiniai.	Accuracy	Mistranslation	Major
This drug dilates the coronary vessels, slows the heart rate, has a choleretic effect.	Šis vaistas plečia vainikines **kraujagysles**, lėtina širdies ritmą, turi choleretinį poveikį.	Šis vaistas plečia vainikines arterijas, lėtina širdies susitraukimų dažnį, turi choleretinį poveikį.	Terminology	Inconsistent use of terminology	Minor
This drug dilates the coronary vessels, slows the heart rate, has a choleretic effect.	Šis vaistas plečia vainikines kraujagysles, **lėtina širdies ritmą**, turi choleretinį poveikį.	Šis vaistas plečia vainikines arterijas, lėtina širdies susitraukimų dažnį, turi choleretinį poveikį.	Style	Awkward	Neutral
The therapeutic effect of Valerian is noted with prolonged course of treatment.	**Terapinis** valerijono poveikis pastebimas pailginus gydymo kursą.	Gydomasis valerijono poveikis pastebimas vaistą vartojant ilgą laiko tarpą.	Accuracy	Mistranslation	Major
Indications for use Valerian	Vartojimo indikacijos Valerijonas	Valerijono vartojimo indikacijos	Fluency	Grammar	Minor
Indications for use Valerian	Vartojimo indikacijos **V**alerijonas	Valerijono vartojimo indikacijos	Fluency	Spelling	Minor
Valerian drugs are used in many countries as a sedative for nervous agitation, insomnia, neurosis of the cardiovascular system, accompanied by spasm of the coronary vessels, tachycardia, as well as asthma, epilepsy, spasm of the gastrointestinal tract, migraine and for the treatment of neurodermatitis.	Valerijono vaistai daugelyje šalių naudojami kaip raminamasis poveikis nervų sujaudinimui, nemigai, širdies ir kraujagyslių sistemos neurozei, kartu su vainikinių kraujagyslių spazmais, tachikardija, taip pat astma, epilepsija, virškinimo trakto spazmais, migrena ir neurodermito gydymui.	Daugelyje šalių valerijonas vartojamas, kaip raminamasis vaistas padedantis nuo nervinio susijaudinimo, nemigos, širdies neurozės bei vainikinių arterijų spazmų, tachikardijos, taip pat astmos, epilepsijos, virškinamojo trakto spazmų, migrenos bei nervinio dermatito gydymui.	Accuracy	Mistranslation	Minor
Valerian drugs are used in many countries as a sedative for nervous agitation, insomnia, neurosis of the cardiovascular system, accompanied by spasm of the coronary vessels, tachycardia, as well as asthma, epilepsy, spasm of the gastrointestinal	Valerijono vaistai daugelyje šalių naudojami kaip raminamasis poveikis nervų sujaudinimui, nemigai, širdies ir kraujagyslių sistemos neurozei, kartu su vainikinių kraujagyslių spazmais, tachikardija, taip pat astma, epilepsija,	Daugelyje šalių valerijonas vartojamas, kaip raminamasis vaistas padedantis nuo nervinio susijaudinimo, nemigos, širdies neurozės bei vainikinių arterijų spazmų, tachikardijos, taip pat astmos, epilepsijos, virškinamojo trakto spazmų, migrenos	Terminology	Inconsistent use of terminology	Minor

tract, migraine and for the treatment of neurodermatitis.	virškinimo trakto spazmais, migrena ir **neurodermito** gydymui.	bei **nervinio dermatito** gydymui.			
Sometimes valerian is used in combination with bromine preparations, cardiac and other sedatives.	Kartais valerijonas vartojamas kartu su bromo preparatais, širdies ir kitais raminančiaisiais.	Kartais valerijonas vartojamas kartu su bromo preparatais, vaistais širdžiai ir kitais raminamaisiais vaistais.	Accuracy	Under-translation	Minor
Valerian is used in the form of decoctions, tinctures, infusions, powders and extracts.	Valerijonas naudojamas nuovirų, tinktūrų, užpilų, miltelių ir ekstraktų pavidalu.	Valerijonas naudojamas nuovirų, tinktūrų, užpilų, miltelių ir ekstraktų pavidalu.			
The work of Professor V.I. Ishchenko showed that tablets from whole finely ground vegetable raw materials are 2.5 times more effective than tablets with valerian rhizome extract.	Profesoriaus V. I. Iščenko **darbas** parodė, kad tabletės iš nesmulkintų smulkiai supjaustytų augalinių žaliavų yra 2,5 karto efektyvesnės nei tabletės su valerijono šakniastiebių ekstraktu.	Profesoriaus V.I. Iščenkos darbai parodė, kad tabletės iš nesmulkintų augalinių žaliavų yra 2,5 karto efektyvesnės nei tabletės su valerijono šakniastiebių ekstraktu.	Fluency	Grammar	Minor
Valerian - contraindications and side effects	Valerijonas - kontraindikacijos ir šalutinis poveikis	Valerijono kontraindikacijos ir šalutinis poveikis	Fluency	Grammar	Minor
Valerian - contraindications and side effects	Valerijonas - kontraindikacijos ir šalutinis poveikis	Valerijono kontraindikacijos ir šalutinis poveikis	Fluency	Punctuation	Neutral
Most patients are well tolerated by treatment with Valerian drugs, but with hypertension, the opposite sedative effect is stimulating, as well as sleep disturbance with heavy dreams.	Gydymas valerijono vaistais daugumą pacientų gerai toleruoja, tačiau esant hipertenzijai, stimuliuoja priešingas raminamasis poveikis, taip pat miego sutrikimas esant sunkiems sapnams.	Dauguma pacientų gydymą vaistiniu valerijonu toleruoja gerai, tačiau sergant hipertenzija, pasireiškia priešingas raminamasis poveikis, taip pat miego sutrikimai ir neramūs sapnai.	Accuracy	Mistranslation	Major
Absolute contraindications to Valerian therapy are:	Absoliučios valerijono terapijos kontraindikacijos yra:	Kontraindikacijos gydymui valerijonu yra:	Accuracy	Mistranslation	Major
- congenital fructose intolerance, sucrose / isomaltase deficiency, glucose-galactose malabsorption, lactose intolerance, lactase deficiency;	- įgimtas fruktozės netoleravimas, sacharozės / izomaltazės trūkumas, gliukozės ir galaktozės malabsorbcija, laktozės netoleravimas, laktazės trūkumas;	- įgimtas fruktozės netoleravimas, sacharozės-izomaltazės trūkumas, gliukozės ir galaktozės malabsorbcija, laktozės netoleravimas, laktazės trūkumas;	Fluency	Punctuation	Neutral
- I trimester of pregnancy;	- I nėštumo trimestras;	- I nėštumo trimestras;			
- age up to 3 years;	- amžius iki 3 metų;	- amžius iki 3 - jų metų;	Fluency	Spelling	Neutral
- individual intolerance to the components of the drug.	- individualus netoleravimas vaisto sudedamosioms dalims.	- asmeninis vaisto sudedamųjų dalių netoleravimas.	Accuracy	Mistranslation	Minor
Relative contraindications; (Valerian extract tablets are prescribed under medical supervision): chronic enterocolitis; II – III trimester of pregnancy; lactation period.	Santykinės kontraindikacijos; (Valerijono ekstrakto tabletės skiriamos prižiūrint gydytojui): lėtinis enterokolitas; II - III nėštumo trimestras; žindymo laikotarpis.	Santykinės kontraindikacijos (valerijono ekstrakto tabletes skiriant gydytojui prižiūrint): II-III nėštumo trimestras; žindymo laikotarpis.	Accuracy	Mistranslation	Neutral
Relative contraindications; (Valerian extract tablets are prescribed under medical supervision): chronic enterocolitis; II – III trimester of pregnancy; lactation period.	Santykinės kontraindikacijos**;** (Valerijono ekstrakto tabletės skiriamos prižiūrint gydytojui): lėtinis enterokolitas; II - III nėštumo trimestras; žindymo laikotarpis.	Santykinės kontraindikacijos (valerijono ekstrakto tabletes skiriant gydytojui prižiūrint): II-III nėštumo trimestras; žindymo laikotarpis.	Fluency	Punctuation	Neutral

Elderly patients should take Valerian with caution, since the drug is able to increase blood coagulation, while maintaining the risk of stroke or heart attack.	Vyresnio amžiaus pacientai valerijoną turėtų vartoti atsargiai, nes vaistas gali padidinti kraujo krešėjimą, išlaikant insulto ar **širdies smūgio** riziką.	Vyresnio amžiaus pacientai valerijoną turėtų vartoti atsargiai, nes vaistas gali paskatinti kraujo krešėjimą ir padidinti insulto ar infarkto riziką.	Terminology	Inconsistent use of terminology	Minor
Elderly patients should take Valerian with caution, since the drug is able to increase blood coagulation, while maintaining the risk of stroke or heart attack.	Vyresnio amžiaus pacientai valerijoną turėtų vartoti atsargiai, nes vaistas gali **padidinti** kraujo krešėjimą, **išlaikant** insulto ar širdies smūgio riziką.	Vyresnio amžiaus pacientai valerijoną turėtų vartoti atsargiai, nes vaistas gali paskatinti kraujo krešėjimą ir padidinti insulto ar infarkto riziką.	Accuracy	Mistranslation	Minor
It is also important to consider the following adverse reactions observed after prolonged use of Valerian drug: depressed state, weakness, lethargy, drowsiness, decreased performance, and occasionally - hypersensitivity reactions and constipation.	Taip pat svarbu atsižvelgti į šias nepageidaujamas reakcijas, pastebėtas po ilgalaikio valerijono vaisto vartojimo: prislėgta būsena, silpnumas, letargija, mieguistumas, sumažėjęs darbingumas, retkarčiais - padidėjusio jautrumo reakcijos ir vidurių užkietėjimas.	Taip pat svarbu atsižvelgti į šias nepageidaujamas reakcijas, pastebėtas ilgalaikio valerijono vartojimo metu: depresinė būklė, silpnumas, vangumas, mieguistumas, sumažėjęs darbingumas, o kartais ir padidėjusi jautrumo reakcija bei vidurių užkietėjimas.	Fluency	Grammar	Neutral
It is also important to consider the following adverse reactions observed after prolonged use of Valerian drug: depressed state, weakness, lethargy, drowsiness, decreased performance, and occasionally - hypersensitivity reactions and constipation.	Taip pat svarbu atsižvelgti į šias nepageidaujamas reakcijas, pastebėtas po ilgalaikio valerijono vaisto vartojimo: **prislėgta būsena**, silpnumas, letargija, mieguistumas, sumažėjęs darbingumas, retkarčiais - padidėjusio jautrumo reakcijos ir vidurių užkietėjimas.	Taip pat svarbu atsižvelgti į šias nepageidaujamas reakcijas, pastebėtas ilgalaikio valerijono vartojimo metu: depresinė būklė, silpnumas, vangumas, mieguistumas, sumažėjęs darbingumas, o kartais ir padidėjusi jautrumo reakcija bei vidurių užkietėjimas.	Terminology	Inconsistent use of terminology	Minor
It is also important to consider the following adverse reactions observed after prolonged use of Valerian drug: depressed state, weakness, lethargy, drowsiness, decreased performance, and occasionally - hypersensitivity reactions and constipation.	Taip pat svarbu atsižvelgti į šias nepageidaujamas reakcijas, pastebėtas po ilgalaikio valerijono vaisto vartojimo: prislėgta būsena, silpnumas, **letargija**, mieguistumas, sumažėjęs darbingumas, retkarčiais - padidėjusio jautrumo reakcijos ir vidurių užkietėjimas.	Taip pat svarbu atsižvelgti į šias nepageidaujamas reakcijas, pastebėtas ilgalaikio valerijono vartojimo metu: depresinė būklė, silpnumas, vangumas, mieguistumas, sumažėjęs darbingumas, o kartais ir padidėjusi jautrumo reakcija bei vidurių užkietėjimas.	Terminology	Inconsistent use of terminology	Minor
It is also important to consider the following adverse reactions observed after prolonged use of Valerian drug: depressed state, weakness, lethargy, drowsiness, decreased performance, and occasionally - hypersensitivity reactions and constipation.	Taip pat svarbu atsižvelgti į šias nepageidaujamas reakcijas, pastebėtas po ilgalaikio valerijono vaisto vartojimo: prislėgta būsena, silpnumas, letargija, mieguistumas, sumažėjęs darbingumas, retkarčiais - padidėjusio jautrumo reakcijos ir vidurių užkietėjimas.	Taip pat svarbu atsižvelgti į šias nepageidaujamas reakcijas, pastebėtas ilgalaikio valerijono vartojimo metu: depresinė būklė, silpnumas, vangumas, mieguistumas, sumažėjęs darbingumas, o kartais ir padidėjusi jautrumo reakci**ja** bei vidurių užkietėjimas.	Fluency	Grammar	Neutral
In some cases, dizziness, lowering body temperature, irritation of the stomach, weakening of the concentration of attention are possible.	Kai kuriais atvejais galimas galvos svaigimas, kūno temperatūros sumažėjimas, skrandžio **sudirginimas**, **susilpnėjusi** dėmesio koncentracija.	Kai kuriais atvejais galimas galvos svaigimas, kūno temperatūros sumažėjimas, skrandžio dirglumas, sumažėjusi dėmesio koncentracija.	Accuracy	Under-translation	Neutral

The steam engine was built by whom. The history of the creation of steam engines	Kas **pastatė** garo mašiną. Garo mašinų **kūrimo** istorija	Kas sukūrė garo variklį? Garo variklių sukūrimo istorija	Accuracy	Under-translation	Minor
The steam engine was built by whom. The history of the creation of steam engines	Kas pastatė **garo mašiną**. Garo mašinų kūrimo istorija	Kas sukūrė garo variklį? Garo variklių sukūrimo istorija	Terminology	Inconsistent use of terminology	Major
Interest in water vapor, as an affordable source of energy, appeared along with the first scientific knowledge of the ancients.	Susidomėjimas vandens garais, kaip prieinamu energijos šaltiniu, atsirado kartu su pirmosiomis senovės mokslo žiniomis.	Susidomėjimas vandens garais, kaip prieinamu energijos šaltiniu, atsirado kartu su pirmosiomis senovės mokslo žiniomis.			
People have been trying to tame this energy for three millennia.	Žmonės šią energiją bandė prisijaukinti tris tūkstantmečius.	Žmonės šią energiją bandė prisijaukinti tris tūkstantmečius.			
What are the main stages of this path?	Kokie pagrindiniai šio **kelio** etapai?	Kokie pagrindiniai šios kelionės etapai?	Terminology	Inconsistent use of terminology	Minor
Whose reflections and projects have taught mankind to extract the maximum benefit from it?	Kieno apmąstymai ir projektai išmokė žmoniją iš to išgauti maksimalią naudą?	Kieno apmąstymai ir projektai išmokė žmoniją iš to išgauti maksimalią naudą?			
Prerequisites for the emergence of steam engines	Prielaidos garo mašinoms atsirasti	Garo variklių atsiradimo prielaidos	Accuracy	Mistranslation	Minor
The need for mechanisms that can facilitate labor-intensive processes has always existed.	Mechanizmų, galinčių palengvinti darbui imlius procesus, poreikis egzistavo visada.	Poreikis palengvinti darbui imlius procesus naudojant mechanizmus, egzistavo visada.	Fluency	Grammar	Minor
Until about the middle of the 18th century, windmills and water wheels were used for this purpose.	Maždaug iki XVIII amžiaus vidurio tam buvo naudojami vėjo malūnai ir vandens ratai.	Maždaug iki XVIII amžiaus vidurio tam buvo naudojami vėjo malūnai ir vandens ratai.			
The possibility of using wind energy directly depends on the vagaries of the weather.	Galimybė panaudoti vėjo energiją tiesiogiai priklauso nuo oro sąlygų.	Galimybė panaudoti vėjo energiją tiesiogiai priklausė nuo oro sąlygų.			
And to use water wheels, factories had to be built along the banks of rivers, which is not always convenient and expedient.	O norint panaudoti vandens ratus, gamyklas reikėjo statyti prie upių krantų, o tai ne visada patogu ir tikslinga.	Norint panaudoti vandens ratus, gamyklos turėjo būti pastatytos prie upių krantų, o tai ne visada patogu ir tinkama.	Fluency	Grammar	Neutral
And the effectiveness of both was extremely low.	Ir abiejų efektyvumas buvo itin žemas.	Be to, abiejų efektyvumas buvo itin žemas.	Accuracy	Under-translation	Neutral
Essentially needed new engine, easily managed and devoid of these shortcomings.	Iš esmės reikia naujas variklis, lengvai valdomas ir neturi šių trūkumų.	Iš esmės, buvo reikalingas naujas variklis: lengvai valdomas ir neturintis šių trūkumų.	Fluency	Grammar	Minor
The history of the invention and improvement of steam engines	Garo **mašinų** išradimo ir tobulinimo istorija	Garo variklių išradimo ir tobulinimo istorija	Terminology	Inconsistent use of terminology	Major
The creation of a steam engine is the result of much thought, success and failure of the hopes of many scientists.	Garo variklio sukūrimas yra daugelio mokslininkų daug minčių, sėkmės ir nesėkmės rezultatas.	Garo variklio sukūrimas yra daugelio mokslininkų galybės minčių, sėkmės ir neišsipildžiusių vilčių rezultatas.	Accuracy	Under-translation	Minor
The beginning of the way	**Kelio** pradžia	Kelionės pradžia	Terminology	Inconsistent use of terminology	Minor
The first, single projects were only interesting curiosities.	Pirmieji pavieniai projektai buvo tik įdomūs kuriozai.	Pirmieji pavieniai projektai buvo tik įdomios keistenybės.	Accuracy	Mistranslation	Minor

For example, Archimedes built a steam gun Heron of Alexandria used the energy of steam to open the doors of ancient temples.	Pavyzdžiui, Archimedas pastatė garo pistoletą Aleksandrijos garnys panaudojo garų energiją senovinių šventyklų durims atidaryti.	Pavyzdžiui, Archimedas sukūrė **garais varomą patranką**, kurios garų energiją Heronas iš Aleksandrijos panaudojo atidaryti senovinių šventyklų durims.	Terminology	Inconsistent use of terminology	Critical
For example, Archimedes built a steam gun Heron of Alexandria used the energy of steam to open the doors of ancient temples.	Pavyzdžiui, Archimedas pastatė garo pistoletą Aleksandrijos garnys panaudojo garų energiją senovinių šventyklų durims atidaryti.	Pavyzdžiui, Archimedas sukūrė garais varomą patranką, kurios garų energiją **Heronas** iš Aleksandrijos panaudojo atidaryti senovinių šventyklų durims.	Accuracy	Mistranslation	Critical
For example, Archimedes built a steam gun Heron of Alexandria used the energy of steam to open the doors of ancient temples.	Pavyzdžiui, Archimedas pastatė garo pistoletą Aleksandrijos garnys panaudojo garų energiją senovinių šventyklų durims atidaryti.	Pavyzdžiui, Archimedas sukūrė garais varomą patranką, kurios garų energiją Heronas iš Aleksandrijos panaudojo atidaryti senovinių šventyklų durims.	Accuracy	Mistranslation	Critical
And researchers find notes on the practical application of steam energy to actuate other mechanisms in the works Leonardo da Vinci.	Tyrėjai randa pastabų apie praktinį garo energijos pritaikymą kitiems mechanizmams įjungti Leonardas da Vinčis.	Be to, mokslininkai Leonardo da Vinčio darbuose aptinka užrašų apie praktinį garo energijos panaudojimą kitų mechanizmų paleidimui.	Accuracy	Mistranslation	Major
Consider the most significant projects on this topic.	Apsvarstykite svarbiausius projektus šia tema.	Apsvarstykime žymiausius sumanymus šia tema.	Accuracy	Mistranslation	Minor
In the 16th century, the Arab engineer Tagi al Din developed a design for a primitive steam turbine.	XVI amžiuje arabų inžinierius Tagi al Dinas sukūrė primityvios garo turbinos projektą.	XVI amžiuje arabų inžinierius Tagi al Dinas (org. Tagi al Din) sukūrė primityvios garo turbinos modelį.	Accuracy	Mistranslation	Minor
However, it did not receive practical application due to the strong dispersion of the steam jet supplied to the turbine wheel blades.	Tačiau jis nebuvo praktiškai pritaikytas dėl stiprios garų srovės, tiekiamos į turbinos ratų mentes, sklaidos.	Tačiau jis negalėjo būti praktiškai pritaikomas dėl stipraus garų srauto, skleidžiamo į turbinos veleną.	Accuracy	Mistranslation	Critical
However, it did not receive practical application due to the strong dispersion of the steam jet supplied to the turbine wheel blades.	Tačiau jis nebuvo praktiškai pritaikytas dėl stiprios garų srovės, tiekiamos į **turbinos ratų mentes**, sklaidos.	Tačiau jis negalėjo būti praktiškai pritaikomas dėl stipraus garų srauto, skleidžiamo į turbinos veleną.	Terminology	Inconsistent use of terminology	Critical
Fast forward to medieval France. The physicist and talented inventor Denis Papin, after many unsuccessful projects, stops at the following design: a vertical cylinder was filled with water, over which a piston was installed.	**Greitai pirmyn į viduramžių Prancūziją**. Fizikas ir talentingas išradėjas Denisas Papinas po daugelio nesėkmingų projektų sustoja ties tokiu dizainu: vertikalus cilindras buvo pripildytas vandens, virš kurio buvo sumontuotas stūmoklis.	Persikelkime laiku į viduramžių Prancūziją. Talentingas išradėjas, fizikas – Denisas Papinas (org. Denis Papin) po daugelio nesėkmingų projektų apsistojo ties šiuo modelio dizainu: vandens pripildytas vertikalus cilindras, virš kurio buvo sumontuotas stūmoklis.	Accuracy	Mistranslation	Major
Fast forward to medieval France. The physicist and talented inventor Denis Papin, after many unsuccessful projects, stops at the following design: a vertical cylinder was filled with water, over which a piston was installed.	Greitai pirmyn į viduramžių Prancūziją. **Fizikas ir talentingas išradėjas Denisas Papinas po daugelio nesėkmingų projektų sustoja ties tokiu dizainu: vertikalus cilindras buvo pripildytas vandens, virš kurio buvo sumontuotas stūmoklis.**	Persikelkime laiku į viduramžių Prancūziją. Talentingas išradėjas, fizikas – Denisas Papinas (org. Denis Papin) po daugelio nesėkmingų projektų apsistojo ties šiuo modelio dizainu: vandens pripildytas vertikalus cilindras, virš kurio buvo sumontuotas stūmoklis.	Accuracy	Mistranslation	Minor

The cylinder was heated, the water boiled and evaporated.	Cilindras buvo kaitinamas, vanduo užvirinamas ir išgarinamas.	Cilindras buvo kaitinamas, tad vanduo virė ir garavo	Accuracy	Under-translation	Minor
The expanding steam lifted the piston.	Besiplečiantys garai pakėlė stūmoklį.	Besiplečiantys garai kėlė stūmoklį į viršų.	Fluency	Grammar	Minor
He was fixed in top point rise and waited for the cylinder to cool and the steam to condense.	Jis buvo užfiksuotas viršutinis taškas pakilti ir laukti, kol cilindras atvės ir garai kondensuosis.	Jis buvo pritvirtintas aukščiausiame taške ir turėjo pradėti veikti atvėsus cilindrui bei susikondensavus garams.	Accuracy	Mistranslation	Critical
After the steam condensed, a vacuum was formed in the cylinder.	Susikondensavus garams, cilindre susidarė vakuumas.	Susikondensavus garams, cilindre susidarė vakuumas.			
The piston, freed from fastening, rushed into vacuum under the action of atmospheric pressure.	Stūmoklis, atsilaisvinęs nuo tvirtinimo, veikiamas atmosferos slėgio pateko į vakuumą.	Veikiamas atmosferos slėgio, atsilaisvinęs nuo tvirtinimo stūmoklis buvo įtraukiamas į vakuumą	Accuracy	Mistranslation	Major
It was this fall of the piston that was supposed to be used as a working stroke.	Būtent šis stūmoklio kritimas turėjo būti naudojamas kaip darbinis eiga.	Būtent šis stūmoklio įtraukimas turėjo būti varomąja jėga.	Accuracy	Mistranslation	Major
So, the useful stroke of the piston was caused by the formation of a vacuum due to the condensation of steam and external (atmospheric) pressure.	Taigi naudingą stūmoklio eigą lėmė vakuumo susidarymas dėl garų kondensacijos ir išorinio (atmosferos) slėgio.	Taigi rezultatyviam stūmoklio judėjimui įtaką turėjo garų kondensacijos ir išorinio atmosferos slėgio dėka susidaręs vakuumas.	Accuracy	Under-translation	Minor
Because the Papin steam engine like most subsequent projects, they were called steam-atmospheric machines.	Kadangi Papin **garo mašina** kaip ir dauguma vėlesnių projektų, jie buvo vadinami garo-atmosferinėmis mašinomis.	Papino garbei, garo variklis, kaip ir dauguma vėlesnių projektų, buvo vadinami garo-atmosferinėmis mašinomis.	Terminology	Inconsistent use of terminology	Minor
Because the Papin steam engine like most subsequent projects, they were called steam-atmospheric machines.	Kadangi Papin garo mašina kaip ir dauguma vėlesnių projektų, jie buvo vadinami garo-atmosferinėmis mašinomis.	Papino garbei, garo variklis, kaip ir dauguma vėlesnių projektų, buvo vadinami garo-atmosferinėmis mašinomis.	Accuracy	Mistranslation	Major
Because the Papin steam engine like most subsequent projects, they were called steam-atmospheric machines.	Kadangi Papin garo mašina kaip ir dauguma vėlesnių projektų, jie buvo vadinami garo-atmosferinėmis mašinomis.	Papino garbei, garo variklis, kaip ir dauguma vėlesnių projektų, buvo vadinami garo-atmosferinėmis mašinomis.	Fluency	Punctuation	Minor
This design was very significant disadvantage - the repeatability of the cycle was not provided.	Šis dizainas buvo labai reikšmingas trūkumas - ciklo pakartojamumas nebuvo numatytas.	Tačiau šis modelis turėjo ženklų trūkumą – nebuvo galima užtikrinti veiksmo pasikartojimo.	Accuracy	Mistranslation	Major
This design was very significant disadvantage - the repeatability of the cycle was not provided.	Šis dizainas buvo labai reikšmingas trūkumas - ciklo pakartojamumas nebuvo numatytas.	Tačiau šis modelis turėjo ženklų trūkumą – nebuvo galima užtikrinti veiksmo pasikartojimo.	Fluency	Grammar	Neutral
Denis comes up with the idea of getting steam not in a cylinder, but separately in a steam boiler.	Denisas sugalvoja garą gauti ne cilindre, o atskirai garų katile.	Tad Denisui kilo mintis garus išgauti ne cilindre, o atskirame garų katile.	Accuracy	Mistranslation	Minor
Denis Papin entered the history of the creation of steam engines as an inventor of a very important detail- steam boiler.	Denisas Papinas įėjo į garo variklių kūrimo istoriją kaip labai išradėjas svarbi detalė- garo katilas.	Į garo variklių kūrimo istoriją Denisas Papinas įėjo, kaip labai svarbaus elemento (garo katilo) išradėjas.	Accuracy	Mistranslation	Critical