Assessing machine translation quality on uterine cancer content: MQM-based comparative study

Amal Akram Ismail Muhanna
Teaching Assistant, Ahram Canadian University, Faculty of Languages and Translation, English Department, Giza, Egypt
Supervisors:
- **Prof. Nagwa Younis,** Professor of Linguistics, Department of English, Faculty of Education, Ain Shams University
- **Dr. Lobna Elhakim,** Lecturer of Translation, Department of English, Faculty of Al-Alsun, Ain Shams University

Abstract
This study evaluates the effectiveness of machine translation in conveying medical content on endometrial cancer from English to Arabic. Machine translation is essential for medical awareness, yet few studies evaluate its error-based performance in translating English medical articles on women's tumors into Arabic. This study assesses machine translation efficiency to enable readers and specialists to access medical knowledge. By comparing three translation engines—Yandex, Systran, and Microsoft—the study aims to determine which engine performs best in delivering accurate and reliable medical content. Additionally, the study examines the extent to which target readers can depend on machine translation to understand medical information accurately. By employing the Multidimensional Quality Metrics (MQM) framework, the study identifies accuracy and terminology errors. The CAT tool, SDL Trados, is used in the assessment process. Both qualitative and quantitative approaches are adopted: the qualitative approach identifies errors and assesses their impact on translation quality, while the quantitative approach calculates the frequency of each error type, assigns penalty points, and generates quality scores. The data, sourced from WebMD, covers content on endometrial cancer, a prevalent cancer among women. Although there has been great advancement in machine translation efficiency, machine translation is still inadequate in conveying precise medical content, particularly regarding terminology and accuracy. These two aspects are crucial in translating medical content. This study, in addition to other previous studies, highlights the inaccuracy of machine translation, needing further refinements in translation technologies when dealing with specialized domains like medicine.

Keywords: Machine Translation, Translation Quality Assessment, Multidimensional Quality Metric, SDL Trados, Systran, Yandex, Microsoft, Endometrial Cancer

المستخلص:

تهدف هذه الدراسة إلى تقييم كفاءة الترجمة الآلية في نقل النصوص الطبية، وبالأخص سرطان الرحم، من اللغة الإنجليزية إلى اللغة العربية. كما تهدف الدراسة إلى تحديد محرّك الترجمة الآلية الأفضل في نقل المحتوى الطبي بدقة وسلاسة عن طريق المقارنة بين ثلاثة محرّكات: Yandex، وSystran، وMicrosoft. وكذلك، تبحث الدراسة في مدى قدرة القُرّاء على الاعتماد على الترجمة الآلية لفهم المعلومات الطبية بدقة. لإجراء التقييم البشري للترجمة الآلية الناتجة عن المحرّكات الثلاثة، تعتمد الدراسة على الإطار Multidimentional Quality Metric (MQM-Core) لتحديد الأخطاء الخاصة بالمصطلحات والدقة (Terminology and Accuracy)، وهما فئتان من الإطار. استخدم الباحث إحدى أدوات الترجمة بمساعدة الحاسوب وهي SDL Trados في عملية التقييم البشري عن طريق تفعيل خاصية "تقييم جودة الترجمة" (Translation Quality Assessment). تعتمد الدراسة على النهجين النوعي والكمي، حيث يحدد النهج النوعي الأخطاء ويقيم تأثيرها على جودة الترجمة، بينما يحسب النهج الكمي تكرار كل خطأ، ويعين نقاط جزاء، ويحدد نقاط جودة الترجمة. أُجريت الدراسة على محتوى يخص سرطان بطانة الرحم، لكونه أحد أنواع السرطان الشائعة التي تصيب النساء، ومصدر المحتوى هو الموقع الطبي WebMD. تشير النتائج إلى أنه على الرغم من تقدم كفاءة الترجمة الآلية في نقل المحتوى، فإن الترجمة الآلية لا تزال غير قادرة على نقل المحتوى الطبي الدقيق من حيث المصطلحات والدقة.

الكلمات المفتاحية:

الترجمة الآلية، تقييم جودة الترجمة، التقييم البشري، سرطان بطانة الرحم، SDL Trados، أخطاء المصطلحات، أخطاء الدقة، Systran، Yandex، Microsoft.

Assessing machine translation quality on uterine cancer content: MQM-based comparative study

Amal Akram Ismail Muhanna
Teaching Assistant, Ahram Canadian University, Faculty of Languages and Translation, English Department, Giza, Egypt
Supervisors:
- **Prof. Nagwa Younis,** Professor of Linguistics, Department of English, Faculty of Education, Ain Shams University
- **Dr. Lobna Elhakim,** Lecturer of Translation, Department of English, Faculty of Al-Alsun, Ain Shams University

I. Introduction:

Due to globalization and expansion in industrial fields, the language barrier has become a significant problem. As a result, machine translation has played a major role in bridging this gap. Mondale et al. (2023) stress the indispensable role of machine translation. They emphasize that it influences both the social and economic values for not only being used for translation but also for various functionalities; for instance, it is used for text-to-speech, speech-to-speech, and phonetic input methods. However, language barriers still exist, despite the crucial role of the internet in bringing people closer. As a result, there has been a notable increase in research studies on machine translation. Stressing the fact that machine translation has become a field of interest, Mondal et al. (2023) reinforce that machine translation, a key domain within computational linguistics and artificial intelligence (AI), has gained significant attention due to its potential to overcome language barriers globally.

Despite playing a major role in translation, machine translation involves various challenges. Mondal et al (2023) state that languages that are morphologically or syntactically rich are challenging when machine translation is used due to difference in language systems between the source and target languages. Arabic and English are entirely two different languages, and translating from or into Arabic can be challenging. Moreover, Adly and Ansary (2010) stated that Arabic is a syntactically and morphologically rich language, highlighting the difficulty of Arabic grammar. Thus, this study hypothesizes that the three machine translation engines are ineffective in translating content of specialized domains, specifically medical content, from English into Arabic.

II. Significance of the Study:

Machine translation has become an indispensable tool in everyday life. Although various studies have tackled the role of machine

translation, few studies have focused on the error-based evaluation of machine translation engines while translating from English into Arabic in medical articles related to women's tumors. With the spread of various diseases, such as cancers, people have become passionate about being medically well-informed. The website used in this study compiles medically reviewed articles by specialists regarding diseases and many other topics. However, people face difficulty in reading the content since it is not in their native language. Thus, they can use machine translation engines to gain knowledge. This study helps both readers and specialists to identify the efficiency of machine translation in translating medical articles.

III. Objectives of the Study:

This study aims to prove the efficiency of machine translation in challenging fields such as medicine while maintaining accuracy and medical terminology in translating articles related to women-related tumors. It seeks to highlight the severity of the errors arising from certain machine translation engines and their impact on the quality of the translation. It also seeks to help both medical practitioners and common readers to determine which MT engine from those analyzed can be used in such a field. Another objective is to explore the MT system that outperforms the other systems in this field.

IV. Research Questions:

The researcher attempts to answer the following questions:

1. What are the most frequent types of errors resulting from the three machine translation engines?
2. Which machine translation engine outperforms the others in translating specialized medical content?
3. Which error types have the highest effect on the quality of translation?
4. Which subcategories have the greatest impact on the overall quality of translation?

V. V. Review of literature

The role of machine translation has prevailed in translating various texts, including medical texts. The application of machine translation to medical texts poses multiple challenges for being highly specialized. This review seeks to present the current state of research and advancements in the field of machine translation evaluation.

This study focuses on English and Arabic language pair, where the chosen medical articles are in English, and the translation of the three MT systems is generated in Arabic. In a study conducted by Habeeb (2013),

the quality of translating Arabic headlines into English using two engines, namely, Google and Babylon, is evaluated. In addition, the human translation of such headlines is tested. The model used in his study is Swan's model (1996) for investigating structure and its impact on conveying the original meaning. To assess the clarity, accuracy, and style of the machine-translated headlines, a questionnaire is given to professionals. The results indicate that Google outperforms Babylon in style. However, both systems show advancement compared to their performance in previous studies.

Given that Arabic and English are two completely different systems, this study focuses on comparing the Arabic output of three MT engines. Similarly, Adly and Ansary (2009) focused on evaluating English-Arabic translations by comparing the MT systems based on the Universal Network Language (UNL) against other systems. According to their study, the interlingual approach involves converting the source language into an independent language that can then be transformed into the target language. UNL is a new interlingua, defined as an artificial language that includes lexical, syntactical, and semantic elements similar to those in any other natural language. Then, this language can be automatically transformed into a natural language by computers.

The current study's translation direction is from English into Arabic and the evaluation is done according to MQM error typology on the Arabic texts conducted by the selected MT systems. Similar to the MQM metric, Federico et al. (2014) developed an error taxonomy for evaluating machine translation from English into Arabic, Chinese, and Russian. Such taxonomy includes a set of error classes: morphological errors, lexical choice, additions, omissions, casing and punctuation, reordering errors, and many other errors. Therefore, both MQM and Federico's (2014) taxonomy share common error categories.

Examining the efficiency of machine translation in translating content of specialized fields, Mohamed (2023) compared the output of Google Translate with human translation regarding FIDIC (the Red Book 1999). The researcher focused on four main aspects: passivization and modality at the syntactic level and collocations and doublets (two paralleled synonyms) at the lexical level. The results of the study clarified that human translation is more accurate and effective in conveying legal language, whereas Google Translate fails to convey the legal nuances accurately.

The medical field is always challenging in translation because it is constantly evolving. Thus, terminologies, concepts, and expressions within this field need to be advanced and highly specialized. Like Skianis

et al. (2020), the current study also focuses on evaluating medical data. The objective of their study was to evaluate the effectiveness and accuracy of machine translation systems in translating medical terminologies.

In their study, Skianis et al (2020) tested national translation techniques dealing with medical terms. According to Skianis et al. (2020), the corpora used in their study were medical terminologies. The experiments conducted by Skianis et al (2020) were on both neural and statistical machine translation methods, like this study. The data used for the experiment was a sample of WHO International Classification of Diseases (ICD) terminologies and International Classification of Functioning, Disability and Health (ICF) terminologies. The language pair was from English to French. They used MOSES as the SMT toolkit while Fairseq was used as the NMT tool. The results revealed the following. First, the traditional SMT showed great results compared to the human sample. However, the NMT had a great performance with longer sentences. Second, "pre-trained models like Fairseq on medical terminologies can improve the quality of the translation" (Skianis et al., 2020, p. 67). Two of the engines tested were NMT while the third was hybrid: NMT and SMT. Likewise, NMT and hybrid machine translation systems are tested in the present study.

Additionally, Mujjiga et al. (2019) examined a novel NMT model, compared with a regular expression-based one, to address the challenges of large-scale information retrieval in healthcare, particularly in clinical reports. Their study aims at solving the problem of efficient and accurate document retrieval in large healthcare systems by mapping documents into a concept space. NMT model significantly outperforms a regular expression-based concept mapper.

VI. Research Methods

1. Source of data

In this study, the source of data is a medical website called WebMD. WebMD is an American corporation that publishes trustworthy information regarding human health and well-being. It comprises experienced doctors, health experts, and journalists. This website aims at helping people receive accurate medical information on the internet. People use this website for various reasons, including but not limited to health research or medical information for healthcare and well-being. Doctors, similarly, rely on it to gather information about any topic since it is a medically reviewed site. This website is chosen since it provides

healthcare and medical information. Its content is easy, free, and accessible to all readers. The study investigates a common tumor, namely, endometrial cancer. This type of cancer is chosen as it is unique to women and affects numerous women, nowadays. Thus, women can gain knowledge by reading articles about such tumors. The study aims to determine whether women or readers in general, can rely on the output of machine translation to enrich their knowledge and receive health information either for avoidance, treatment, or other reasons.

2. Machine translation engines selected for application

In this study, three machine translation engines are tested: Microsoft Translator, Systran, and Yandex Translate. The three engines were chosen for their relevance to the study's focus on translating medical content from English into Arabic. Microsoft is a widely used neural machine translation (NMT) engine. Similar to Systran, it shows broad coverage of global languages, including Arabic. Systran is a hybrid engine including NMT that is well-suited for domain-specific tasks like medical translation. Yandex employs unique algorithms for dealing with linguistic challenges. The three engines provide an inclusive evaluation for machine translation performance across specialized content into Arabic.

2.1 Microsoft Translator

It is a cloud-based engine, a neural machine translation that is integrated into Microsoft Office (Dalibor, 2022). As mentioned by Microsoft Corporation (2023), the Microsoft Translator engine offers versatile applications across different domains.

This engine is chosen because it stands out among other MT engines. It is distinguished, according to Microsoft Corporation (2023), by its deep integration, allowing users to use it easily while using Office or Bing. Moreover, it is known for its accuracy as it continuously improves its models that can be used effectively with different domains (Microsoft Corporation, 2023). In this study, the main focus is examining the efficiency of Microsoft Translator in translating the medical content of articles on a website called WebMD. Accordingly, the researcher tests the quality of the translation produced by Microsoft Translator for personal use.

2.2 Systran

The second chosen engine is Systran. According to the Systran website (n.d.), it provides industry-specific translation models that allow users to translate specific jargon. It is one of the top-listed machine translation engines, as mentioned on the Systran official website (n.d.). It has two versions: free and pro. The free version allows translating texts

into more than 50 languages, including Arabic, the target language of this study, while the pro-version supports domain-specific translations, including the medical domain that is the focus of the current study (Neural Machine Translation by Systran, n.d.).

Systran also supports translating websites and unlimited texts as well as different document formats (Text Translation — SYSTRAN, n.d.). Like Microsoft Translator, it is accessible as it can be easily integrated into browsers such as Google Chrome, facilitating the translation of websites, as mentioned in their official website (SYSTRAN Browser Extensions — BROWSER EXTENSIONS Documentation, n.d.). Since this study examines the quality of translating articles on the WebMD website, Systran is a good MT engine to be chosen for the study.

2.3 Yandex Translate

The third engine examined is Yandex Translate. As mentioned on the Yandex Website (n.d.), Yandex Translate is one of the most used engines comprising more than 90 languages, including Arabic and English.

In Yandex Translate's official website, *About Machine Translation* (n.d.), the process of translation by Yandex is clarified. It is stated that when the user enters a text for translation in Yandex Translate, the input is sent to both the statistical and the neural systems. Before generating the final output to the user, this algorithm evaluates multiple factors, and it generates the best translation after this evaluation. Yandex Translate stands out as a no-cost engine that supports text, website, document, and image translation (About Machine Translation, n.d.).

After the translation is generated by the three MT engines mentioned above, a process called translation quality assessment (TQA) is conducted to evaluate the quality produced. This process is discussed below.

3. Multidimensional Quality Metric (MQM):

To evaluate the translation quality of the previously mentioned MT engines, the researcher adopts the Multidimensional Quality Metric (MQM) for evaluation. This metric is an evaluation metric that can be applied to both human and machine translation (MQM, n.d.). MQM, as an integral part of quality management systems, determines translation errors, categorizes them according to their error typology, and determines whether the translation conforms to translation standards (MQM, n.d.). Although the MQM framework does not directly **provide alternatives translation errors**, it provides an approach to **identify, categorize, and evaluate these errors**.

Lommel et al., (2014) stated that MQM is a model that comprises multiple metrics. It is also applicable to all language pairs. It is a translation quality assessment metric that can be tailored to specific translation projects and contexts (Lommel et al., 2014). In this study, MQM is used to identify terminology and accuracy errors, defined below.

3.1 MQM typology:

3.1.1 Terminology

This category occurs when a term does not fit the domain or the organizational standards, or when the term in the target text is not equivalent to the term of the source text (*The MQM CORE Typology*, n.d.). It is divided into three sub-categories:

a. Inconsistent use of terminology: It implies that the same concept is translated into different terms in the target text, leading to inconsistency.

b. Wrong term: It occurs when the term used leads to conceptual mismatch or is not suitable for the domain.

The first sub-category is not suitable for this study since there is no terminology resource provided for the evaluation process. Thus, the main focus is on the other two sub-categories.

3.1.2 Accuracy

This type of error occurs when the source text is inaccurately conveyed. It might happen by distorting, omitting, or adding to the message (*The MQM CORE Typology*, n.d.). This is a key category in the translation quality evaluation as it comprises major subcategories that can affect the quality. Burchardt (2013) defined accuracy as "Issues related to how well the content of the target text represents the content of the source" (p. 4). It encompasses seven subcategories mentioned below *(The MQM CORE Typology*, n.d):

a. Mistranslation: It occurs when the source content is misrepresented in the target text leading to miscomprehension.

b. Over-translation: It means that the target text is inappropriately less specific than the source text.

c. Under-translation: It implies that the target text is inappropriately less specific than the source text (content is generalized)

d. Addition: It means adding content that is not in the source text to the target text.

e. Omission: It is the deletion of content that is in the source text from the target text.

f. Untranslated: It denotes the non-translation of a segment or part in the source text that is intended to be translated.

3.2 MQM approaches

MQM Metric comprises two approaches, namely, analytic and holistic approaches. They differ according to the way the user adopts while analyzing the translation. This study adopts an analytic approach clarified below.

3.2.1 Analytic approach

Lommel et al. (2014) explain that the analytic approach is used to specify and quantify certain errors in translation; thus, it does not focus on the whole text, it rather focuses on certain parts for identifying certain errors. In the analytic approach, errors cannot be counted as errors if they comply with the specifications. For instance, if the specifications of a service manual do not consider stylistic issues important, they can be neglected in the error count.

The analytic approach does not only depend on specifying errors, but it also includes scorecards for evaluation. Burchardt, A. et al. (2020) stated that scorecards are computer systems that have been long used in the translation industry by evaluators to identify and categorize errors for producing a quality score. Scorecards are helpful for evaluation using the analytic evaluation method.

According to Burchardt et al., (2020), using scorecards, evaluators can assess translation quality, specify points of strength and weakness, and provide feedback . Scorecards help ensure consistency and objectivity in the evaluation process, allowing for more accurate and reliable assessments of translation quality. However, in this study, using SDL Trados for review by activating the TQA feature for assessment enables the user to evaluate the translation without scorecards, highlight the error category, determine its severity, and add feedback.

3.3 MQM severity levels

Lommel (2018) defined severity as "the nature of the error itself and its effects on the usability of the translation" (p.121). In other words, the target audience is affected by the extent of the error depending on its severity. Severity is applied to single errors, not the whole error categories. Severity levels are helpful for categorizing errors based on their impact on the translation quality.

As explained by Lommel (2018), MQM, by default, supports four severity levels: critical, major, minor, and null. First, critical errors may lead to destruction or injury (Kocmi & Federmann, 2023). Lommel (2018) stated that critical errors “can prevent a translation from fulfilling its purpose (e.g. by preventing the intended user from completing a task) and may have safety or legal implications” (p. 120). In other words,

critical errors can severely impact the accuracy of the translation leading to misunderstanding.

Second, major errors are the errors that "make the intended meaning of the text unclear in such a way that the intended user cannot recover the meaning from the text, but are unlikely to cause harm" (Lommel, 2018, p. 120). Unlike the severity of critical errors, major errors can noticeably impact the quality of the translation and may require correction to ensure the text is effectively delivered to the target readers. Third, minor errors "are those that do not impact usability. In most cases, the intended user will correct them and move on" (Lommel, 2018, p. 120). Although minor errors are insignificant compared to major and critical, they still affect the overall translation quality. Fourth, null errors are not considered errors, since "No penalties are applied at this severity" (Lommel, 2018, p.120). The default penalties are 100 points (critical), 10 points (major), 1 point (minor), and 0 points (null) (Kocmi & Federmann, 2023).

VII. Translation Quality Assessment (TQA):

Translation quality assessment (TQA) is defined by Petrova (2019) as the process of evaluating the quality of translation in terms of accuracy, fluency, and overall content. She added that it can be applied using certain tools for checking the quality of accuracy, consistency, style, and cultural appropriateness.

Moorkens et al. (2018) stressed that Language Service Providers use either a TQA error checking integrated into a CAT tool, a TQA standalone software, or an evaluation model. In this study, the research adopts SDL Trados for implementing the TQA process.

To use the TQA feature for assessment, metrics for evaluation should be used. As clarified by Moorkens et al. (2018), evaluation can be achieved using evaluation models that are error-based. In these evaluation models, the evaluator detects these errors in the target text. Then, such errors are counted, categorized, and weighted according to the model used In this study, the adopted metric for evaluating the machine translation outputs is the Multidimensional Quality Metric (MQM) Metric. Two categories and their subcategories are chosen for the analysis, namely, terminology and accuracy errors.

VIII. Computer Assisted Translation Tools (CAT Tools):

Schoening (2014) defined CAT tools as "software programs that assist human translators in transferring the meaning of text between languages more quickly and efficiently" (Para.1). Additionally, Petrova (2019) highlighted that Computer Assisted Translation Tools (CAT

Tools) were developed for texts that have repetitive patterns such as contracts and technical texts.

CAT tools also enable evaluators to activate the TQA feature while activating certain metrics for evaluation. After that, reviewers can specify the error category/subcategory and its severity level. Moreover, using CAT tools, the reviewer can add a comment or even reject the translation. After the evaluation process, CAT tools generate a TQA report including all the details of the evaluation including the number of errors of each category and subcategory with their severity level.

There are multiple CAT tools, such as SDL Trados. Lommel (2018) stated that Trados is the most widely used CAT tool and stressed the role of Trados in translation evaluation as Trados offered to use MQM since 2016. In other words, the full MQM typology is used in the error-checking module (p.126). In this study, the researcher uses SDL Trados, one of the most commonly used CAT tools in the translation industry, to evaluate the output of three MT engines, namely, Systran, Yandex Translate, and Microsoft. The activated metric for TQA is the MQM metric. It is customized according to the medical corpus examined in this study; the metric used in SDL Trados for assessment is customized to fit the subcategories of accuracy and terminology.

IX. Evaluation method:

Manual evaluation is a method of MT evaluation that is mainly based on the human evaluator. Moorkens et al. (2018) stated that for human translation quality assessment to be performed, particularly the evaluation of machine translation outputs, evaluators assess by looking at adequacy (accuracy) and fluency. In addition, secondary measures are also used to evaluate the readability, comprehensibility, usability, and acceptability of translations. Adequacy ("also called accuracy or fidelity in certain studies devoted to MT Evaluation") measures the extent of transferring the meaning of the target text into the source language using the translation process (Moorkens et al., 2018).

Adequacy requires proficiency in both the source and target languages. Similarly, Guessoum and Zantout (2005) reinforced the importance of evaluating adequacy, i.e. the extent to which the meaning is conveyed from the source to the target language. For evaluating the output of MT, the role of bilingual human evaluators is essential to score the quality of the translation. Popovic (2018) also stressed the importance of the role of human evaluators.

X. Analysis

This part covers the accuracy and terminology errors in the three machine translation outputs, namely, Yandex, Systran, and Microsoft. For each subcategory, the error is detected with its severity level. Using SDL Trados for the translation quality assessment process, the three engines fail the assessment. The CAT tool detects the penalty points that determine the efficiency of each engine whether it passes the assessment or fails. The analysis below shows some examples of their analysis. In each example, the error related to the subcategory being discussed is both in bold and underlined. In addition, the errors that are only underlined are the errors that are related to other subcategories but affect the translation. Both types of errors are analyzed below each example.

1. Accuracy errors:

1.1 Addition errors:

Addition errors are the errors that occur by adding information in the translated text that is not mentioned in the original text. In some cases, addition errors can cause awkwardness and redundancy. This part shows some addition errors highlighted in bold. Errors related to other subcategories are underlined.

Example 1:

Source text (Segment 58):

Endometrioid cancer starts in the glands of your endometrium (uterine lining), and under the microscope it looks very similar to normal uterine lining (endometrium).

Target text (Yandex):

شبيهة ببطانة الرحم سرطان يبدأ في الغدد من **بطانة الرحم (بطانة الرحم)** ، تحت المجهر أنها تبدو مشابهة جدا **عادي بطانة الرحم (بطانة الرحم) .**

Target text (Systran):

يبدأ سرطان بطانة الرحم في غدد **بطانة الرحم (بطانة الرحم)**، وتحت المجهر يبدو مشابها جدا **لبطانة الرحم الطبيعية (بطانة الرحم).**

Target text (Microsoft):

يبدأ سرطان بطانة الرحم في غدد **بطانة الرحم (بطانة الرحم)**، وتحت المجهر يبدو مشابها جدا **لبطانة الرحم الطبيعية (بطانة الرحم).**

Suggested Translation:

يبدأ الورم الغدي نظير بطانة الرحم في غدد بطانة الرحم (وهو الغشاء المُبطّن للرحم)، حيث يشبه جدا الغشاء المُبطّن للرحم (بطانة الرحم) تحت المجهر.

In this example, the three engines implement addition errors by using the same translation for translating the medical term “endometrium” and its explanation “uterine lining”. Although the translation is correct, there is redundancy in Arabic. It can be translated as suggested above to both differentiate between them and clarify the

meaning of “endometrium”. It is a minor error as it does not change the intended meaning despite the repetition. Another minor addition error is detected in Yandex output because "عادي" is irrelevant to the context.

Furthermore, the three engines do not convey the accurate terminology for translating the medical term “Endometrioid cancer”. According to Hitti and Al-Khatib (2011), this term is translated as "نظير بطانـة الـرحم" (p.286). This is a critical wrong term error where the translation used is equivalent to another term, “endometrial cancer”. Using the same Arabic translation for different concepts can mislead the readers.

Yandex, compared to the other two engines, shows grammatical errors by inaccurately using "يبـدأ" with "شـبيهة", where the word "شـبيهة" in Arabic is a feminine noun that should be referred to in feminine as well, "تبدأ".

Example 2:

Source text (Segment 38):

Prognosis is an estimate of how the course of the disease will go for you.

Target text (Yandex):

التكهن هو تقدير مدى مسار المرض سوف تذهب إلى لك.

Target text (Systran):

التشخيص هو تقدير لمدى سير المرض بالنسبة لك.

Target text (Microsoft):

التشخيص هو تقدير لكيفية **سير مسار** المرض بالنسبة لك.

Suggested translation:

تشخيص مآل المرض هو تقدير مسار المرض بالنسبة لك.

The addition error is evident in the translation of Microsoft. There is no need to add "كيفيـة سـير" as the word “مسـار” shows the implied meaning of “how the course of the disease will be”. It is a minor error because the translation is understandable, although following the source structure leads to awkwardness.

A major terminology error is shown in both Systran and Microsoft outputs in translating “prognosis”. It is defined as "the likely course of a medical condition" (Oxford, n.d.). Thus, it can be translated as "تكهن/تشخيص مآل". The word "تشـخيص" is equivalent to “diagnosis” that is a different term and is defined as “the process of identifying a disease, condition, or injury from its signs and symptoms” (National Cancer Institute, n.d.). The meaning of both terms is different. By translating them similarly, the meaning is changed, causing confusion for readers. Thus, the direct equivalent of “prognosis” can be better translated into "تكهن" or "تشخيص مآل", not just "تشخيص".

Mistranslation errors are also determined in Yandex and Systran. In Yandex output, a critical mistranslation error of "سوف تذهب لك" is shown. The target text is meaningless compared to the meaning intended in the source text. It is literally translated, yet the meaning intended is that prognosis shows the course of the disease. It is critical since, by being meaningless and non-understandable, it affects the flow of the translation.

In Systran, the translation of "تقدير مدى سير" is not accurate. It is a major mistranslation error as it does not show the intended meaning, "the development of the disease". The word "مدى" means the limit of something (Omar, 2008).

Example 3:

Source text (Segment 27):

If left untreated, endometrial cancer can spread to the bladder or rectum, or it can spread to the vagina, fallopian tubes, ovaries, and more distant organs.

Target text (Yandex):

إذا تركت دون علاج, سرطان بطانة الرحم يمكن أن تنتشر إلى المثانة أو المستقيم أو **يمكن أن ينتشر** إلى المهبل أو قناة فالوب أو المبيضين و الأجهزة البعيدة.

Target text (Systran):

إذا ترك سرطان بطانة الرحم دون علاج، فقد ينتشر إلى المثانة أو المستقيم، أو **قد ينتشر إلى** المهبل وقناتي فالوب والمبيضين والأعضاء **الأكثر** بعدا.

Target text (Microsoft):

إذا ترك دون علاج ، يمكن أن ينتشر سرطان بطانة الرحم إلى المثانة أو المستقيم ، أو **يمكن أن ينتشر** إلى المهبل وقناتي فالوب والمبيضين والأعضاء البعيدة.

Suggested translation:

قد ينتشر سرطان بطانة الرحم، حال تُرك دون علاج، ليصيب المثانة أو المستقيم، أو ليصيب المهبل وقناتي فالوب والمبيضين وأعضاء أخرى أبعد.

The three outputs include a minor addition error by adding "قد ينتشر إلى". The spread of endometrium cancer can affect all the following organs. Thus, there is no need to follow the source text structure and repeat "قد ينتشر إلى" before the word "vagina". Another addition minor error is detected in adding "الأكثر" by Systran. An alternative structure can be conveyed as "أبعد", where the comparison is clearer rather than "أكثر بعدًا".

The three engines show addition errors. Furthermore, Yandex comprises a minor over-translation by translating "tubes" as "قناة". There are two tubes in a female's body where endometrium cancer may affect both and not just one. Thus, a dual structure should have been used as "قناتي فالوب". It is major since the meaning is changed, and this can be misleading for the reader by faulty medical information.

Yandex, moreover, includes a major wrong term error by translating "organs" as "أجهزة" not "أعضاء". In this segment, all the

mentioned items are “organs”, "أعضـاء", not "systems". There are also grammatical minor mistakes determined in Yandex output where it refers to endometrial cancer in the feminine form as "تُركـت-تنتشـر". In Arabic, the word “cancer” means "سـرطان", and it is a masculine noun that should be indicated in the masculine form as "تُرك-ينتشر".

1.2 Mistranslation

Mistranslation is a critical type of accuracy errors. Mistranslation errors can lead to severe consequences, particularly in medical contexts. These errors change the intended meaning of the source text, leading to misunderstanding, misdiagnoses, and potentially incorrect treatments. Such errors can be life-threatening for readers or patients. The three engines show mistranslation errors. Below are some examples of their analysis. Mistranslation errors are underlined and bold. The underlined texts are errors but related to other subcategories.

Example 1:

Source text (Segment 74):

Went through menopause late

Target text (Yandex):

ذهبت من خلال انقطاع الطمث في وقت متأخر

Target text (Systran):

مر بمرحلة انقطاع الطمث في وقت متأخر

Target text (Microsoft):

ذهب من خلال انقطاع الطمث في وقت متأخر

Suggested translation:

انقطاع الطمث في سِنّ متأخرة

This segment expresses one of the risk factors of endometrial cancer. The intended meaning of the original text is that one of the main causes of endometrial cancer is the occurrence of menopause for women after the age of fifty. None of the three engines conveys the same meaning. The phrase "ذهـب مـن خـلال" is meaningless in this context as it is literally translated into “to pass through”. This error is a major error for being misleading and incomprehensible. In Systran output, it is also a major error as "مـر بـ" means “to go through a phase and then it ends” (Masoud, 1992). However, menopause is a phase that does not end.

Addition minor error is detected in the three outputs. The addition of "فـي وقـت" is not correct as “late” refers to the age of women who can suffer from this type of cancer, not time. This error is minor as it is not significant and can be understood. Thus, it can be translated as suggested.

Example 2:

Source text (Segment 49):

Your doctor may call these “lower grade.”

Target text (Yandex):

طبيبك **قد نطلق على هذه** "الدرجة السفلى".

Target text (Systran):

وقد **يطلق عليك طبيبك** وصف "الدرجة الأدنى."

Target text (Microsoft):

قد **يطلق طبيبك على هذه** "الدرجة الأدنى."

Suggested translation:

قد يصنّف الطبيب هاتين الدرجتين من سرطان بطانة الرحم ضمن "درجة أقل حدة".

This segment refers to the one preceding it, “grades 1 and 2 endometrioid adenocarcinoma are considered type 1 endometrial cancer”. The determiner “these” refers to grades one and two that are considered type one of endometrial cancer. In Arabic, to transfer the same sense of dualism in the feminine form, “هاتين” can be used to refer to the two grades. None of the translation outputs accurately translates the intended meaning. Instead, "هذه" in Yandex and Microsoft is vague implying a minor addition error. In Systran output, it is a major mistranslation error as "يطلق عليك" is not the exact meaning where "lower grade" refers to type one, including grade one and two, not the patient herself.

"Lower grade" means less severe concerning the disease severity level. Thus, it is better translated as "أقل حدة". However, "الدرجة الأدنى" is correct despite being literally translated. In Yandex translation, "السفلى" is a major mistranslation error.

The word order in Yandex’s output is also confusing by beginning with the noun instead of the verb. This can be a minor style error by mirroring the English structure in the Arabic version. This structure is not natural in the Arabic syntax. Although it is understandable, it might impact the overall flow and naturalness of the translated text.

Example 3:

Source text (Segment 108):

And if you’ve just gotten diagnosed with endometrial cancer, you may have a better outlook than these numbers suggest.

Target text (Yandex):

و إذا كنت قد حصلت للتو تشخيص سرطان بطانة الرحم ، قد يكون أفضل من **التوقعات من هذه الأرقام تشير إلى**.

Target text (Systran):

وإذا شخصت إصابتك بسرطان بطانة الرحم للتو، فقد تكون نظرتك المستقبلية أفضل **مما توحي به** هذه الأرقام.

Target text (Microsoft):

وإذا كنت قد تم تشخيصك للتو بسرطان بطانة الرحم ، فقد يكون لديك نظرة أفضل **مما توحي به** هذه الأرقام.

Suggested translation:

إذا شُخصت المرأة للتو بسرطان بطانة الرحم، فإن مآل المرض قد تكون أفضل مما تشير إليه هذه الأرقام.

The word “outlook” means “the probable or expected condition or outcome of something” (Collins, n.d.). In Arabic, the word "مـآل" means "the outcome of something in the future” (Omar, 2008, p.141). Thus, it is better translated as "مـآل" rather than "توقعـات" since it only means predictions, not results. Similarly, "نظرتـك المسـتقبلية/نظرة" are not suitable for this medical context as these words mean "the way you predict or see something in the future". However, in such a context, the intended meaning is the disease prognosis. Therefore, this is a major terminology error in the three engines. In addition, in Microsoft, it can be an error of inconsistent use of terminology where it was previously translated as "توقعات" in segment 107.

The verb “suggest” is mistranslated by the three engines. In Yandex, the whole phrase is a major mistranslation error. The phrase " التوقعـات مـن هـذه الأرقـام تشـير إلـى" cannot be understood in the context. It is incorrectly ordered where it should be "التوقعات التي تشـير إليهـا هـذه الأرقـام". In both Microsoft and Systran, the verb "توحى بـه" means “to give a signal for someone to do something” (Omar, 2008, p. 2415). Thus, it can be better translated as "تشير إليه" numbers. According to Omar (2008), the verb "أشار إلى" means “to indicate something” (p. 1245).

In addition, the translation of “have gotten” as " حصـلت" is literally translated by Yandex. It should have been translated directly as "شُخصـت". It is a minor mistranslation error as, despite being awkward, it is understandable. There is also a minor omission error of a preposition "على" with "حصلت".

1.3 Over-translation

Over-translation errors arise when the content is rendered into the target language with additional specifications, not present in the original. Such errors can compromise accuracy by misleading readers, particularly in medical texts, where clarity and precision are crucial. In this article, minor over-translation errors are present in the three outputs. These errors often stem from using second-person pronouns in Arabic and sticking to the source text's structure, resulting in excessive specificity. As the article provides general information about endometrial cancer rather than

addressing a specific patient, generalization should be maintained in the translation. Example one illustrates this type of error.

Example 1:

Source text (Segment 98):

It depends on how far your cancer has spread when you're first diagnosed with it.

Target text (Yandex):

ذلك يعتمد على مدى إنتشار السرطان عندما كنت أول تشخيص ذلك.

Target text (Systran):

يعتمد ذلك على مدى انتشار السرطان عند تشخيص الإصابة به لأول مرة.

Target text (Microsoft):

يعتمد ذلك على مدى انتشار السرطان عند تشخيص **إصابتك** به لأول مرة.

Suggested translation:

يعتمد ذلك على مدى انتشار السرطان عند تشخيص الإصابة به لأول مرة.

Microsoft output comprises a minor over-translation error. The pronoun "كـ" attached as a suffix to "إصـابة" is not accurate. This segment discusses the factors of the survival rates of endometrial cancer in general for all people suffering from such disease. It does not focus on a specific person being diagnosed. On the other hand, this pronoun in Arabic is used to refer to the second-person pronoun "you". Contrariwise, Systran renders the translation accurately and concisely.

Yandex includes a major mistranslation error. The phrase "عندما كنت أول تشخيص ذلـك" is meaningless and does not transfer the intended meaning of "first diagnosed with it". The intended meaning is that when patients get diagnosed with endometrial cancer for the first time, their survival rates vary based on the extent of the cancer's spread.

Example 2:

Source text (Segment 32):

If you get diagnosed with endometrial cancer, your doctor might tell you it's type 1 or type 2.

Target text (Yandex):

إذا كنت تحصل على تشخيص سرطان بطانة الرحم, **طبيبك** قد اقول لكم انها من النوع 1 أو النوع 2.

Target text (Systran):

إذا شخصت إصابتك بسرطان بطانة الرحم، فقد **يخبرك طبيبك** بالنوع 1 أو النوع 2.

Target text (Microsoft):

إذا تم تشخيص إصابتك بسرطان بطانة الرحم، فقد **يخبرك طبيبك** أنه من النوع 1 أو النوع 2.

Suggested translation:

إذا شُخِصت المريضة بسرطان بطانة الرحم، فإن هذا النوع من السرطان يُصنّف على أنه من النوع الأول أو النوع الثاني.

Similar to example one of over-translation errors, the use of the second person pronoun "كـ" in "يخبرك/طبيبك" is not accurate. This segment introduces the two types of endometrial cancer for all women being diagnosed with endometrial cancer, not a certain woman whose doctor categorizes her cancer type. Therefore, adding the phrase "your doctor might tell you" when translated into Arabic, is unnecessary, as in the suggested translation. Otherwise, it can be used in the general form as "قد يُصنّفه الطبيب على أنه من النوع الأول أو الثاني". This is a minor error as the meaning is conveyed clearly.

Yandex, in addition to the over-translation error, includes a major mistranslation error. The phrase "قد اقول لكم انها" is meaningless and can mislead the readers. Such engine also comprises a minor wrong term error of translating "get diagnosed" literally as "تحصل على تشخيص" while adhering to the original structure. Adding "تحصل على" is unnecessary.

Over-translation errors are very few in the three engines. Most of them are minor errors by using inaccurate pronouns in the target content. Two examples are discussed to clarify this type of error in this article.

In Systran translation, the preposition "بـ" is not accurately used. The intended meaning of this segment is that the doctor, after detecting the tumor, categorizes it into type one or two. However, by adding the preposition "بـ" the meaning is changed so that the doctor explains the types for the patient, not diagnoses the type. This is a major omission error where the meaning is altered.

1.4 Under-translation

Under-translation errors can impact the faithfulness of the translation by failing to convey the meaning or detail mentioned in the source text. Such incomplete information transfer can mislead the readers where details are crucial in medical contexts. Under-translation errors can significantly impact the clarity and informativeness of the text.

Example 1:

Source text (Segment 62):

This rarer form of endometrial and uterine cancer starts in the lining of the uterus.

Target text (Yandex):

هذا أندر شكل **الرحم وعنق الرحم** يبدأ في بطانة الرحم.

Target text (Systran):

يبدأ هذا **الشكل النادر** من سرطان بطانة الرحم وسرطان الرحم في بطانة الرحم.

ISSN 1110-2721

Occasional Papers
Vol. 89: January (2025)

Target text (Microsoft):

يبدأ هذا **الشكل النادر** من سرطان بطانة الرحم والرحم في بطانة الرحم.

Suggested translation:

يبدأ هذا النوع من سرطان بطانة الرحم والرحم في الغشاء المبطن للرحم، وهو أندر من الورم الغدي نظير بطانة الرحم.

The three engines fail to convey the comparative form between this type of endometrial cancer, known as “uterine papillary serous carcinoma”, and the other type mentioned in segment 58, known as “endometrioid”. Yandex output uses the superlative form rather than the comparative. Both Systran and Microsoft do not show the comparative structure, they just describe this type as rare. This is a major error because the meaning is changed.

A critical wrong term error is detected in Yandex output where the term “endometrial cancer”, known as "سرطان بطانة الرحم", is translated as "عنق الرحم" whose equivalent is “cervical cancer”. These are different types of cancer where cervical cancer is the “cancer that starts in the cells of the cervix”, not the endometrium (National Cancer Institute, 2023). Inaccuracy in translating medical terms can cause severe implications.

In addition, the word “cancer” is omitted in Yandex output, leading to an omission critical error. By deleting this term, the word "form" does not describe cancer, but the organ, "الرحم وعنق الرحم". In contrast, the intended meaning is the description of the cancer type, known as “uterine papillary serous carcinoma”.

A minor wrong term error is determined in the three outputs by translating “form” as "شكل". It is better translated as "نوع" as this segment defines a type of endometrial cancer which is “uterine papillary serous carcinoma”, and not its shape.

Example 2:

Source text (Segment 48):

For instance, grades 1 and 2 endometrioid adenocarcinoma are considered type 1 endometrial cancer.

Target text (Yandex):

على سبيل المثال, الصفوف 1 و 2 شبيهة ببطانة الرحم غدية تعتبر نوع 1 سرطان بطانة الرحم.

Target text (Systran):

على سبيل المثال، يعد الصنفان 1 و 2 من **سرطان الغدد الصماء البطاني الرحمي** النوع 1 من سرطان بطانة الرحم.

Target text (Microsoft):

على سبيل المثال ، تعتبر الدرجات 1 و 2 من **سرطان بطانة الرحم الغدي** من النوع 1 من سرطان بطانة الرحم.

Suggested translation:

فمثلا، تُصنّف الدرجتين 1 و2 من الورم الغدي نظير بطانة الرحم على أنها النوع الأول لسرطان بطانة الرحم.

The translation of the term "endometrioid" as "سـرطان بطانـة الرحم/البطـاني الـرحم" is considered both a wrong term and an under-translation critical error. According to the American Cancer Society (2019), it "can be divided into different types based on how the cells look under the microscope". Endometroid is a type of endometrial cancer (American Cancer Society, 2019. By translating it as "سـرطان بطانـة الـرحم", the term is incorrect and generalized, misleading and confusing to the readers. Therefore, it should be translated as suggested. Hitti (2016) translated the term "endometrioid" as "شـبيه بطانـة الـرحم" (p.286). The term "adenocarcinoma" is defined as "cancer that forms in the glandular tissue, which lines certain internal organs as well as makes and releases substances in the body" (National Cancer Institute, n.d.) and is translated as "ورم غدي" (Peter MacCallum Cancer Centre, n.d., p.1).

The translation of "grade" in both Yandex and Systran is a wrong term. Grade, in this context, refers to the severity levels of the cancer, not class, as translated by Systran, leading to a minor error. Yandex's translation of this term is a major wrong term error as it is translated as "classes". Although the translation is clear, a grammatical error is clear in using the plural, not the dual form in Arabic for the two grades 1 and 2.

Example 3:

Source text (Segment 44):

Grading Endometrial Cancer

Target text (Yandex):

الدرجات سرطان بطانة الرحم

Target text (Systran):

تصنيف سرطان بطانة الرحم

Target text (Microsoft):

تصنيف سرطان بطانة الرحم

Suggested translation:

تصنيف مراحل سرطان بطانة الرحم

The word "تصـنيف" is defined as "the classification of something into types" (Masoud, 1992, p.502). According to Collins (n.d.), the word "grading" is defined as "the act of classifying something on a scale by quality, rank, size, or progression, etc.". The translations of the three engines show the concept of classifying; however, in this context, the word grading refers to the severity levels of endometrial cancer. In the section following this segment, the word "grade" is defined as the number

that specifies the extent of the cancer cells being abnormal under the microscope. Thus, in this context, “grading” means the progression stages of endometrial cancer. None of the three outputs shows that such grading means the classification of endometrial cancer’s stages. This is a minor under-translation error by Systran and Microsoft. Yandex shows a major under-translation error by its failure to show that "الدرجات" refers to the cancer progression, i.e., “grading”, of such type of cancer.

Under-translation errors are few in this article. However, they impact the overall quality of the translated content. These errors arise by generalizing the original content in the target text where details can be missed, leading to a lack of accuracy and fidelity.

1.5 Omission

Omission errors occur when essential content of the original text is omitted in the translated content. The missing information can lead to significant changes in the intended meaning, causing misinterpretation. In medical content, omissions can severely affect the accuracy of the translation by omitting medical information.

Example 1:

Source text (Segment 109):

That’s because treatments get better over time, and these numbers are based on women who got diagnosed and treated at least 5 years earlier.

Target text (Yandex):

ذلك لأن العلاجات الحصول على أفضل مع مرور الوقت, و هذه الأرقام تستند إلى النساء اللواتي تم تشخيص الحالة وعلاجها **5 سنوات على الأقل** في وقت سابق.

Target text (Systran):

وذلك لأن العلاجات تتحسن بمرور الوقت، وهذه الأرقام تستند إلى النساء اللواتي تم تشخيصهن وعلاجهن **قبل 5 سنوات على الأقل.**

Target text (Microsoft):

ذلك لأن العلاجات تتحسن بمرور الوقت ، وتستند هذه الأرقام إلى النساء اللائي تم تشخيصهن وعلاجهن **قبل 5 سنوات على الأقل.**

Suggested translation:

قد تكون التوقعات أفضل لأن طرق العلاج تتحسّن بمرور الوقت، ولأن هذه الأرقام، أي نسب معدلات البقاء، تستند إلى نساء شُخصن وتلقين العلاج منذ خمس سنوات على الأقل من الوقت الحالي.

Although the phrase "من الوقت الحالي" is not mentioned in the source text, it should be added to clarify that the phrase "five years" refers to the time when women got diagnosed earlier than the current time. This segment suggests that as time passes, treatment improves as such survival rates refer to women diagnosed from five years. The three engines fail to convey the exact meaning of the phrase “5 years earlier”. It is a major omission error sincethe reader cannot understand the meaning of "قبل خمس

"سنوات". In addition, a minor omission error is represented in not adding "معدلات البقاء" to the target text. The word “numbers” in this context refers to the endometrial cancer survival rates. For more clarification, the phrase "نسب معدلات البقاء" can be added.

Example 2:

Source text (Segment 20):

Cancer can affect the uterus, the hollow, pear-shaped organ where a baby grows.

Target text (Yandex):

السرطان يمكن أن تؤثر على الرحم، **مجوف** على شكل كمثرى الجهاز حيث الطفل ينمو.

Target text (Systran):

يمكن أن يؤثر السرطان على الرحم، وهو العضو المجوف الذي يشبه الكمثرى حيث ينمو الطفل.

Target text (Microsoft):

يمكن أن يؤثر السرطان على الرحم ، العضو المجوف على شكل كمثرى حيث ينمو الطفل.

Suggested translation:

هذا النوع من السرطان يصيب الرحم، وهو العضو المُجوّف الذي يشبه الكمثرى حيث ينمو الجنين.

Among the three engines, Yandex includes a major omission error. It sticks to the source structure which leads to a major loss of meaning. The phrase "the hollow" defines the shape of the uterus. Thus, in Arabic, the pronoun "وهو” should be added to refer to the uterus. In addition, the noun "مجوف" should be definite with "ال" in Arabic as the uterus is a unique organ in women's bodies.

The three engines show major mistranslation errors. The modal verb “can” in this segment is not used for possibility, but for truth; according to Collins (n.d.), “can” is defined as “mentioning a quality or fact about something". The intended meaning of this segment is that this cancer affects the uterus. In contrast, the three outputs show the possibility of the uterus being affected.

Moreover, the three engines produce a minor wrong term error by translating “baby” as "الطفل". According to Omar (2008), the word "طِفل" is defined as “a young child whose age is from birth till maturation” (p.1405). This term can be translated as "جنين" since Omar (2008) defined it as “the child in the uterus” (p.408).

Yandex also shows a major wrong term for translating “organ” as "جهاز". “Organ” means a part of the human body that performs a particular function. Similarly, "عضو" is “a part of a human body” (Omar, 2008, p. 1513). However, the word "جهاز" is defined as “a group of organs in the body that perform certain functions, such as the respiratory or digestive systems” (Omar, 2008, p. 412).

Example 3:

Source text (Segment 72):

Postmenopausal women have a high risk for endometrial cancer if they:

Target text (Yandex):

النساء بعد سن اليأس لديهم ارتفاع **خطر** سرطان بطانة الرحم إذا:

Target text (Systran):

النساء بعد انقطاع الطمث معرضات لخطر كبير للإصابة بسرطان بطانة الرحم إذا كن:

Target text (Microsoft):"

النساء بعد انقطاع الطمث لديهن خطر كبير للإصابة بسرطان بطانة الرحم إذا:

Suggested translation:

يتصاعد خطر الإصابة بسرطان بطانة الرحم بعد انقطاع الطمث إذا كانت المرأة:

The word "للإصـابة" has a significant role in showing the relation between the endometrial cancer and the danger. By omitting this word, the meaning is severely affected. Yandex's translation shows that the tumor is the danger that increases. However, the meaning of the original text is that the risk of having cancer increases. Thus, the danger that increases is the chance of suffering from it, not the tumor itself. This is a major omission error committed by the Yandex engine.

In addition, this segment refers to women, and the pronoun referring to them shall be in the feminine form in Arabic. Yandex, on the contrary, uses "لـديهم". It should be "لـديهن" or simply as suggested above. There is also a minor inconsistent use of terminology detected in Yandex's output. In other segments, including 74 and 86, the term "menopause" is translated as "انقطاع الطمث", not "سن اليأس".

Omission errors in this article are not many. However, their impact can be major or critical as shown above. The accuracy of the translated text can be seriously affected by omitting essential information. Omission errors can also arise by not adding information crucial for the target readers to understand in the target text even if such content is not mentioned in the original text.

1.6 Untranslated

Untranslated errors are evident when the original content intended to be translated is kept in the source language. These errors can impact the overall accuracy of the translation where the missing and untranslated content can be crucial for readability and comprehension.

Example 1:

Source text (Segment 4):

Written by WebMD Editorial Contributors

Target text (Yandex):

كتبه **WebMD** التحرير المساهمين

Target text (Systran):

كتب WebMD **Editorial Contributors**

Target text (Microsoft):

كتبه المساهمون التحريريون في WebMD

Suggested translation:

بقلم فريق التحرير بموقع ويب مِد(WebMD):

This segment refers to the people who wrote the article, i.e., the editorial contributors. The word "editorial" is defined by Cambridge (n.d.) as the person who writes, changes, or corrects work to be published". Systran fails to translate the whole term, leading to a major untranslated error which hinders the readability of the target text. Although Yandex translated the term, it shows a major mistranslation error. Yandex's output fails to convey the exact meaning leading to confusion and lack of understanding. Microsoft represents the best translation among the other three engines.

Despite not being an error, the Arabic transliterated term for "WebMD" should be added for more clarity as "ويب مِد". In addition, the word "موقع" can be added before the website's name to define and clarify it.

Example 2:

Source text (Segment 66):

Uterine carcinosarcoma

Target text (Yandex):

الرحم carcinosarcoma

Target text (Systran):

سرطان الرحم:

Target text (Microsoft):

ساركوما الرحم:

Suggested translation:

ساركوما سرطان الرحم

A major untranslated error is detected in Yandex's translation where the word "carcinosarcoma" is not translated into Arabic. Using a medical term in its original English language hinders the reader's understandability. The word (sarcoma), defined as "A type of cancer that begins in bone or soft tissue, including cartilage, fat, muscle, blood vessels and other connective tissue." (Peter MacCallum Cancer Centre, n.d.), is omitted, and "carcino" is a term that refers to the word cancer. According to the World Health Organization (n.d.), this term is translated as "سركومة سرطانية".

In both Microsoft and Systran, the term "Uterine carcinosarcoma" is generalized. Thus, it is a major under-translation error. By omitting the word "sarcoma", the meaning is both generalized and changed.

Example 3:

Source text (Segment 57):

The most common type of endometrial cancer is a kind of adenocarcinoma (cancer that forms in glandular tissue) called endometroid cancer.

Target text (Yandex):

النوع الأكثر شيوعا من سرطان بطانة الرحم هو نوع من غدية (السرطان التي تشكل في الأنسجة الغدية) دعا **endometroid** السرطان.

Target text (Systran):

النوع الأكثر شيوعا من سرطان بطانة الرحم هو نوع من السرطان الغدي (السرطان الذي يتشكل في الأنسجة الغدية) يسمى سرطان بطانة الرحم.

Target text (Microsoft):

النوع الأكثر شيوعا من سرطان بطانة الرحم هو نوع من السرطان الغدي (السرطان الذي يتشكل في الأنسجة الغدية) يسمى سرطان بطانة الرحم.

Suggested translation:

النوع الأكثر شيوعًا من سرطان بطانة الرحم هو نوع من الورم الغدي (الورم الذي يتكوّن في الأنسجة الغدّية) ويسمى سرطان شبيه بطانة الرحم.

The term "endometroid" is not translated by the Yandex engine. This is a major untranslated error because using a medical English term in an Arabic text leads to a lack of readability and comprehensibility. Hitti (2016) translated the term “endometrioid” as “شبيه بطانة الرحم” (p.286); so this term should have been conveyed in Arabic.

The same term is inaccurately translated. In a medical context, inaccuracy of medical terminology can lead to crucial consequences. This example shows a critical wrong term error while translating “endometroid” by both Systran and Microsoft. Translating this term into "endometrial cancer" where both tumors are not the same may lead to incorrect diagnosis, significant misunderstanding, and inappropriate prognosis. This is a critical error, based on the MQM metric, as serious errors significantly impact understanding and lead to incorrect interpretation.

A major omission error is evident in Yandex’s output. The term “adenocarcinoma” is defined as “cancer that forms in the glandular tissue, which lines certain internal organs and makes and releases substances in the body” (National Cancer Institute, n.d.) and is translated as "ورم غدي" (Peter MacCallum Cancer Centre, n.d., p.1). The term "ورم/سرطان" is omitted, impacting the meaning and understanding. The word “type” refers to “adenocarcinoma", which is a type of cancer. The meaning is altered in Yandex's translation where the "type" refers to "غدية" meaning "glandular”.

Untranslated errors are not detected in Microsoft output. Systran also shows only one untranslated major error. On the contrary, Yandex is the engine that shows the highest number of untranslated errors. Although untranslated errors can be less significant than other error subcategories such as terminology or mistranslation, some of them can show major or even critical impact based on their influence on the overall translation quality and understandability.

2. Terminology errors

2.1 Inconsistent use of terminology

Inconsistent use of terminology errors is the use of more than one term to translate the same concept. This type of error results in a lack of consistency that hinders comprehension and confuses the readers.

Example 1:

Source text (Segment 113):

Taking birth control lowers the risk, but first talk with a doctor about possible pros and cons. Being healthy, eating well, and watching your weight may help lower the risk.

Target text (Yandex):

اتخاذ النسل يخفض **خطر**, ولكن أولا التحدث مع الطبيب حول الايجابيات والسلبيات المحتملة. كونها صحية, تناول الطعام بشكل جيد, و مشاهدة الوزن قد يساعد على خفض **خطر.**

Target text (Systran):

يقلل إستخدام وسائل منع الحمل من خطر الإصابة، ولكن تحدث أولا مع الطبيب حول الإيجابيات والسلبيات المحتملة. قد يساعد التمتع بصحة جيدة وتناول الطعام بشكل جيد ومراقبة وزنك على تقليل **المخاطر**.

Target text (Microsoft):

يقلل تناول وسائل منع الحمل من **المخاطر** ، ولكن تحدث أولا مع الطبيب حول الإيجابيات والسلبيات المحتملة. قد يساعد التمتع بصحة جيدة وتناول الطعام بشكل جيد ومراقبة وزنك في تقليل **المخاطر.**

Suggested translation:

وسائل تحديد النسل تقلل من خطر الإصابة، لكن يُنصح بالتحدث مع الطبيب حول مزاياها وعيوبها، كما يقلل الحفاظ على الصحة وتناول الطعام الصحي ومتابعة الوزن من خطر الإصابة.

The three engines fail to keep consistency in the terminology used. The translation of the term "risk" is both a wrong term and an inconsistent use of terminology error. It is a wrong-term error where it refers to the risk factors of having endometrial cancer. It is also an inconsistent use of terminology error because it is translated as "خطـر الإصـابة" by Systran and Microsoft in segments 72, and as "الخطـر" in segment 69. Based on the context, it should be translated into "خطـر الإصـابة" as this part tackles the steps that women can follow to prevent endometrial cancer. It is a major wrong-term error as using "المخـاطر" can be misunderstood to mean that

the woman suffers from this cancer, but these steps reduce its risks. In contrast, such steps reduce the possibility of having cancer. In addition, it is a minor inconsistent use of terminology because such inconsistencies reflect the same meaning of the term itself.

A minor mistranslation error is clear in the three outputs. The translation of “eating well” in this context means eating healthy food. It is translated literally like the source text, leading to awkwardness in its meaning. Although its meaning is vague, it can be understood.

The verb "تحـدث" should refer to women, since these recommendations are for women to avoid endometrial cancer. The form of this verb should be in the feminine form like "تحدثي". It is a minor error.

Yandex translation shows mistranslation errors that impact the intended meaning. The translation of "taking birth control" means “the use of medicines, devices, or surgery to prevent pregnancy” (MedicinePlus, 2024). The use of "اتخـاذ النسـل", meaning “to take birth”, is both meaningless and severely hinders comprehension. It is a critical mistranslation error. Another major mistranslation error is "being healthy"; it is literally translated as "كونهـا صـحية". In this context, it means that when women’s health is good, this can help in avoiding cancer. However, Yandex’s translation is incomplete where the pronoun "ـها" does not refer to any noun. This leads to confusion and misinterpretation.

The term “watch weight” means to “lose weight or to not gain weight" (Merriam-Webster, n.d.). However, Yandex translates it literally as “view" which changes the meaning and shows vagueness in the context. It is thus a major wrong-term error.

All the inconsistent use of terminology errors detected are analyzed in the examples above, despite being analyzed in other subcategories. For example, the error of translating "menopause" as "سـن اليـأس" and "انقطـاع الطمث" in example three, omission section, is analyzed in detail above.

This article contains very few inconsistent use of terminology errors. However, they have an impact on the overall quality of the translated text; these errors confuse readers by translating the same term into various concepts in the target text.

2.2 Wrong term

Wrong-term errors are the most significant types of errors in this study. This study evaluates the translation quality of a medical text where accurate and precise terminology is crucial in this field. Using wrong medical terminology can bring about severe consequences, including misdiagnosis, misinterpretation, prognosis, and other implications.

Example 1:

Source text (Segment 101):

Regional means the disease has spread to nearby areas or lymph nodes.

Target text (Yandex):

الإقليمية يعني المرض قد انتشر إلى المناطق المجاورة أو الغدد الليمفاوية.

Target text (Systran):

الوسائل الإقليمية:انتشار المرض إلى المناطق القريبة أو العقد اللمفية.

Target text (Microsoft):

الإقليمية تعني أن المرض قد انتشر إلى المناطق المجاورة أو الغدد الليمفاوية.

Suggested translation:

السرطان الناحي هو انتشار المرض إلى المناطق أو الغدد اليمفاوية المجاورة.

This segment explains one of the stages of endometrial cancer. In anatomy, the term “regional” is known as "النـاحي", according to Khayat (2006, p.1303). None of the three engines translate the medical term correctly. This error implies a severe misunderstanding where the readers cannot grasp the content or even the definition of this cancer stage. For this reason, this is a critical wrong-term error in the three engines.

A mistranslation error is also pinpointed in the three engines by translating “nearby areas or lymph nodes”. The word “nearby” should describe both the “lymph nodes” and the “areas”. However, in the three outputs, the word "مجاورة/قريبـة" only describes “areas”. According to the National Cancer Institute (n.d.), regional cancer is defined as the type of cancer that has spread to nearby tissues, organs, or lymph nodes.

Example 2:

Source text (Segment 83):

Have a family history of endometrial, colorectal, or breast cancer

Target text (Yandex):

لديهم تاريخ عائلي من بطانة الرحم, **القولون** أو سرطان الثدي

Target text (Systran):

وجود تاريخ عائلي للإصابة بسرطان بطانة الرحم أو **القولون** أو الثدي

Target text (Microsoft):

لديك تاريخ عائلي للإصابة بسرطان بطانة الرحم أو القولون والمستقيم أو الثدي

Suggested translation:

وجود تاريخ عائلي للإصابة بسرطان بطانة الرحم، أو سرطان القولون والمستقيم، أو سرطان الثدي.

Critical wrong-term error is evident in both Yandex and Systran outputs. The translation of the medical term “colorectal” is "قُولـوني مُسـتقيمي", as found in Khayat (2006, p.296). This type of cancer is defined by the American Cancer Society (2024) as the type of cancer that starts in

the colon or the rectum. By translating the term into "قولـون", the meaning is changed where it may mislead the reader to think that this is another type of cancer that affects the colon only.

In addition, Yandex includes a minor omission error where "الإصابة بسـرطان" should be mentioned before "بطانـة الـرحم". Although this omission error might hinder the flow of the translation, the reader can understand the missing text from the context. It is also a minor error because the word "سـرطان" is mentioned before "الثـدي" to determine the breast cancer, which can contribute to clarification for the missing part. In addition, "لـديهم" is incorrect as the pronoun that should be used is "ـهن" to show the grammatical feminine structure that refers to women who have a family history.

Example 3:

Source text (Segment 36):

Sometimes it develops from a condition called atypical hyperplasia, which is when too many cells grow in the endometrium.

Target text (Yandex):

في بعض الأحيان يتطور من حالة تسمى **فرط تنسج شاذة** ، وهو عند الكثير من الخلايا تنمو في بطانة الرحم.

Target text (Systran):

وفي بعض الأحيان، يحدث نتيجة حالة تسمى فرط التنسج اللانمطي، وهي الحالة التي ينمو فيها عدد كبير جدا من الخلايا في بطانة الرحم.

Target text (Microsoft):

في بعض الأحيان يتطور من حالة تسمى **تضخم غير نمطي** ، وهو عندما تنمو الكثير من الخلايا في بطانة الرحم.

Suggested translation:

في بعض الأحيان ينشأ سرطان بطانة الرحم من حالة طبية تُسمى فرط التنسج اللانمطـي، وتعنـى نمو عدد كبير من الخلايا في بطانة الرحم.

There is a major wrong term error produced by both Yandex and Microsoft while translating the medical term "atypical hyperplasia". The Arabic equivalent for this term is provided by Khayat (2006) as " فرط تنسج بطانة الـرحم" (p.692). The word "atypical" is translated by the World Health Organization (n.d.) as "لا نمطـي". This condition is defined by the National Cancer Institute (n.d.), as "a benign (not cancer) condition in which cells look abnormal under a microscope and are increased in number". Therefore, the word "تضـخم" is not accurate. Similarly, "شـاذة" as an equivalent for "atypical" is awkward.

Yandex shows a minor mistranslation error by adding "عنـد" as an equivalent for "when". This leads to ambiguity in the flow of the translation, yet the segment can still be interpreted.

Wrong-term errors can seriously affect the translation quality, particularly in a highly specialized field like medicine. The three engines show critical and major wrong-term errors that can bring about grave consequences.

XI. Discussion

The current study investigates the competency of three machine translation engines in translating medical content. The chosen data for the analysis is related to endometrial cancer for being one of the most common types of cancer nowadays. The three engines under investigation are Yandex, Systran, and Microsoft. The medical field is a highly specialized field that both medical terminology and content must be delivered precisely and accurately. This study focuses on evaluating the accuracy and terminology errors arising from machine translation and their impact on the overall quality.

The three engines, despite being different in their efficiency, fail to convey accurate terminology and meaning. Based on the word count (1079 words), the allowed penalty points for the translation to pass the translation quality assessment is 338 errors. None of the three translations are below the threshold of the penalty points allowed. First, the Microsoft engine showed1025/338. Second, Systran produced 1057/338. Third, Yandex's penalty points were 2657/338.

The three engines showed critical errors related to both accuracy and terminology. The translation, for example, of "regional" in the three outputs is categorized as a critical wrong-term error. It is translated as "الإقليمية", which means "regional" related to areas, by the three engines. However, it should be translated as "الناحي", that means "regional cancer", as this term refers to a type of cancer. Mistranslation critical errors are seen in the three outputs. For instance, the translation of "risk" should be "خطر الإصابة" which means the risk of having cancer. However, it is translated as "خطر/مخاطر" which changed the meaning that can be interpreted as the risk of spread or development of the cancer.

The occurrences of wrong-term errors are significant in the three engines. Yandex shows the highest number of occurrences: 32 wrong-term errors, divided into 22 major, 4 minor, and 6 critical errors. Microsoft also shows a significant number of wrong terms errors: 22 errors, divided into 7 minor, 9 major, and 7 critical. The least, yet evident, number of wrong term errors are created by Systran. There are 20 wrong term errors in Systran output: 7 major, 7 minor and 6 critical.

Terminology precision is crucial in translating medical content. Serious consequences can result from terminology errors in translating

medical texts, including misdiagnosis, misinterpretation, confusion, or risks. In addition, terminology errors hinder the development of research and treatment protocols by distorting clinical data.

The second subcategory in the three engines is the mistranslation errors. In Yandex, 52 mistranslation errors are detected: 31 major, 9 minor, and 12 critical Systran shows 17 errors: 9 minor, 7 major, and 1 critical. Similarly, Microsoft produces 17 errors. However, 11 of them are minor, 5 are major and 1 is critical.

Mistranslation is the failure to convey the intended meaning of the original text, thus failure to translate accurately. Yandex is the engine that shows the highest number of mistranslation errors that entirely change the meaning. For instance, the translation of "This kind usually doesn't form, grow, or spread very quickly" should refer to a type of cancer that neither grows nor forms nor spreads. Yandex translated this segment as "هذا النـوع عادة لا شـكل النمـو، أو تنتشـر بسـرعة كبيـرة". This is a major error where it should be translated as "هذا النوع من السرطان لا يتشكّل أو ينمو أو ينتشر بسـرعة". However, "شـكل النمـو" is a mistranslation error where this describes the shape of the growth.

The three engines show major and critical errors that can mislead the readers. Misinterpretations in highly specific fields lead to misunderstandings of medical information or instructions that can cause adverse impacts on dosage, treatment, or procedures.

The other subcategories are categorized and analyzed in this study with their penalty points and severity levels. However, the most evident and impactful subcategories in the three engines are the wrong term and mistranslation errors. In the three figures below, the translation quality assessment overview is represented by SDL Trados. These charts highlight the severity levels and the number of occurrences for each subcategory of both terminology and accuracy dimensions.

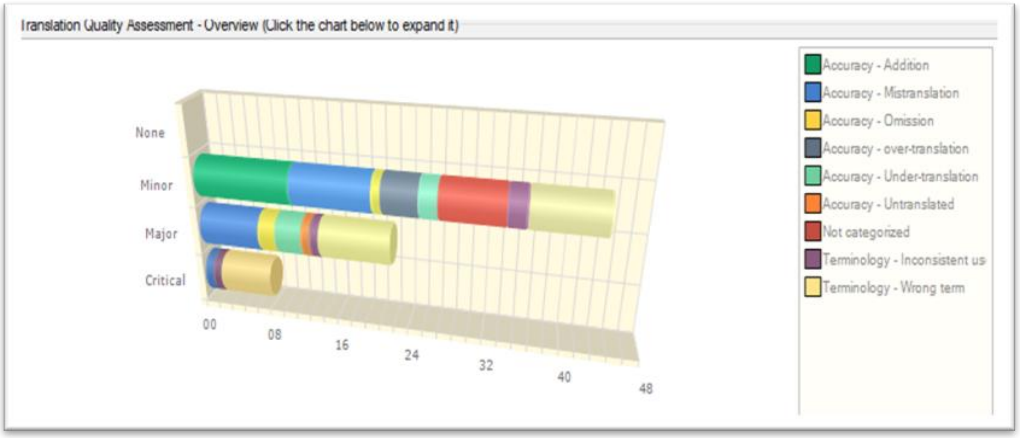

Fig.(1): Translation Quality Assessment Overview (Systran)

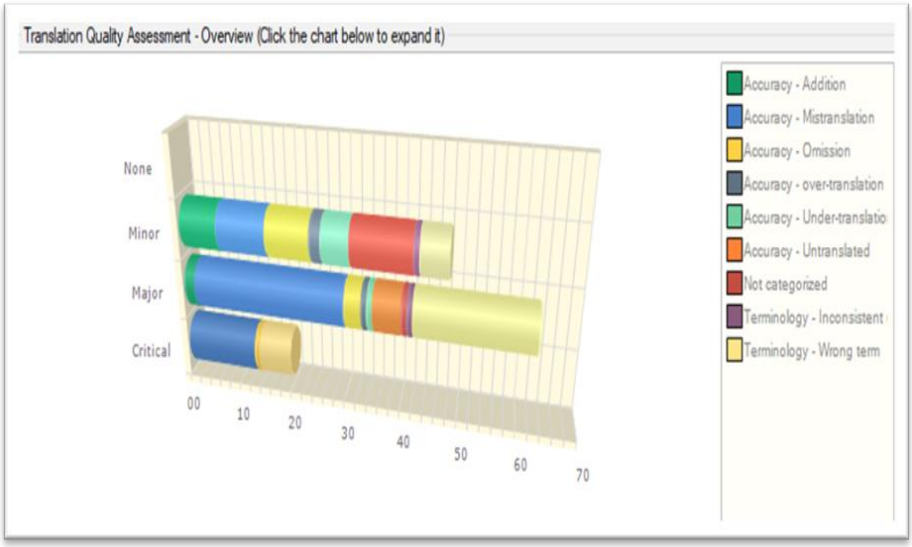

Fig. (2): Translation Quality Assessment Overview (Yandex)

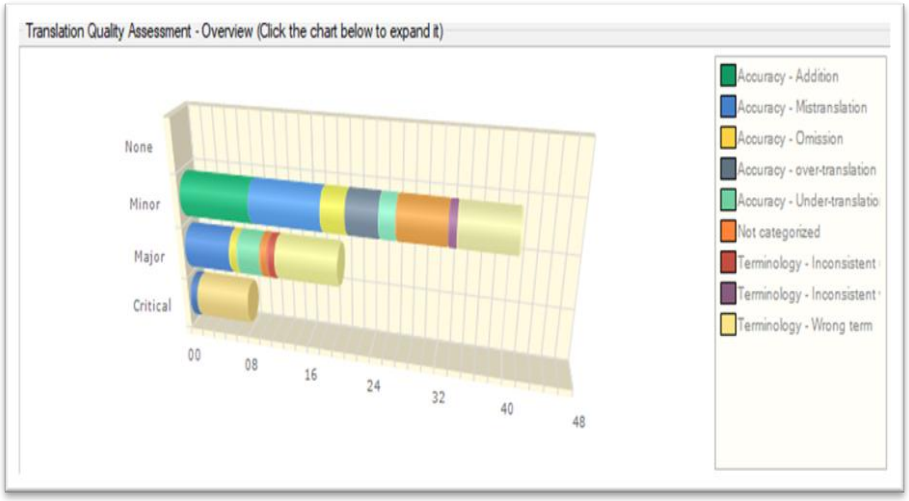

Fig. (3): Translation Quality Assessment Overview (Microsoft)

Based on the three charts above, the highest occurrences of major errors produced by the three outputs are mistranslation and wrong term errors. The other subcategories' errors are clear in the three translation outputs, but they are less than the other two subcategories.

XII. Conclusion

This study examines the competence of machine translation engines in translating medical texts; particularly endometrial cancer content. The three engines under examination areYandex, Systran, and Microsoft. The researcher compares the three outputs' performance based on the MQM Metric, where accuracy and terminology errors are investigated.

Despite the major role of machine translation in multiple fields and aspects of life, it still has limitations in rendering medical terminology and content. The three engines display major and critical terminology errors that have a severe impact on the overall quality of the

translated content. Similarly, mistranslation errors are significant in the three outputs. However, Yandex is the least efficient engine to convey the intended meaning. There are 28 major, 9 minor, and 13 critical mistranslation errors in the Yandex output. Both Systran and Microsoft show 17 errors. Microsoft comprises 11 minor, 5 major, and 1 critical error, while Systran includes 9 minor, 7 major, and 1 critical error.

Regarding the number of occurrences of terminology errors, Systran and Microsoft are close: Systran shows 20 and Microsoft shows 22 errors. Yandex includes the highest terminology error occurrences amounting to 32 errors.

As mentioned above, Microsoft shows the least penalty points (1025), followed by Systran (1057) and finally Yandex (2657). Despite failing in the translation quality assessment, Microsoft outperforms the other two engines. Thus, Microsoft translation can be promising in translating medical documents with human involvement for post-editing.

Collaboration of human translation and machine translation can result in more accurate and beneficial medical translation. By addressing the machine translation limitations pinpointed in this study, machine translation engines may have an insightful impact on translating medical content.

These insights not only contribute to medical translation but also have practical implications for machine translation performance in any highly specialized field. Further research can explore more recommendations to enhance the role of machine translation for researchers, readers, experts, or readers, in general, to have easy access to any content written in a foreign language.

References

Adly, N., & Al Ansary, S. (2010). Evaluation of Arabic machine translation system based on the universal networking language. *Natural Language Processing and Information Systems*, 243-257. https://doi.org/10.1007/978-3-642-12550-8_20

Burchardt, A. (2013). Multidimensional quality metrics: a flexible system for assessing translation quality. *Proceedings of Translating and the Computer 35*, London, UK. Aslib. https://aclanthology.org/2013.tc-1.6

Burchardt, A., Lommel, A., & Macketanz, V. (2020). A new deal for translation quality. *Universal Access in the Information Society*, *20*(4), 701-715. https://doi.org/10.1007/s10209-020-00736-5

Dalibor. (2022). Machine Translation: What It Is, and How It Works. Phrase. Available from: https://phrase.com/blog/posts/machine-translation/

Federico, M., Negri, M., Bentivogli, L., & Turchi, M. (2014). Assessing the Impact of Translation Errors on Machine Translation Quality with Mixed-effects Models. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*,1643–1653, Doha, Qatar. *Association for Computational Linguistics*. https://doi.org/10.3115/v1/d14-1172

Guessoum, A., & Zantout, R. (2004). A methodology for evaluating Arabic machine translation systems. *Machine Translation*, *18*(4), 299–335. https://doi.org/10.1007/s10590-005-2412-3

Habeeb, L. (2013). An evaluation of machine translation quality of news headlines. Faculty of languages and linguistics, University of Malaya.

Hitti, Y. & Al-Khatib, A. (2011). Hitti's New Medical Dictionary. Librairie du Liban, Beirut.

https://www.cancer.gov/types/cervical

Khayat, M. (2006). The Unified Medical Dictionary. Librairie du Liban, Lebanon.

Kocmi, T. & Federmann, C. (2023). Large Language Models Are State-of-the-Art Evaluators of Translation Quality. *Proceedings of the 24th Annual Conference of the European Association for Machine Translation*, 193–203, Tampere, Finland. European Association for Machine Translation.

Lommel, A. (2018). Metrics for Translation Quality Assessment: A Case for Standardising Error Typologies. *Machine translation*, 109–127. https://doi.org/10.1007/978-3-319-91241-7_6

Lommel, A., Uszkoreit, H., & Burchardt, A. (2014). Multidimensional Quality Metrics (MQM): a framework for declaring and describing translation quality metrics. *Revista Tradumàtica*, *12*, 455–463. https://doi.org/10.5565/rev/tradumatica.77

Mas'oud, G. (1992). Mu'jam Al-Ra'id [The lexicon of Al-Ra'id]. Dar El-'elm Lel Malayin, Beirut.

Mohamed, A. H. (2023). Google Translate: Lexical and Syntactic Problems of FIDIC Translation into Arabic. *CDELT Occasional Papers in the Development of English Education*, *83*(1), 33–53. https://doi.org/10.21608/opde.2023.325327

Mondal, S., Zhang, H., Kabir, H., Ni, K., & Dai, H. (2023). Machine translation and its evaluation: A study. *Artificial Intelligence Review*, *56*(9), 10137–10226. https://doi.org/10.1007/s10462-023-10423-5

Moorkens, J., Castilho, S., Gaspari, F., & Doherty, S. (2018). Translation Quality Assessment. Springer.

Omar, A. (2008). Mu'jam Al-lugha Al-'rabiyya Al-mu'aṣirah [The lexicon of the Contemporary Arabic]. 'Alam Al-kotob, Cairo.

Peter MacCallum Cancer Centre. (n.d.). https://www.petermac.org/

Petrova, V. (2019). Translation quality assessment tools and processes in relation to CAT tools. *Proceedings of the Second Workshop Human-Informed Translation and Interpreting Technology associated with RANLP 2019*, 89-97. https://doi.org/10.26615/issn.2683-0078.2019_011

Popović, M. (2018). Error classification and analysis for machine translation quality assessment. *Machine Translation: Technologies and Applications*, 129-158. https://doi.org/10.1007/978-3-319-91241-7_7

Schoening, S. (2024). CAT Tools: Exploring Computer-Assisted Translation. *Phrase*. https://phrase.com/blog/posts/cat-tools/

Skianis, K., Briand, Y., & Desgrippes, F. (2020). Evaluation of machine translation methods applied to medical terminologies. *Proceedings of the 11th International Workshop on Health Text Mining and Information Analysis.* https://doi.org/10.18653/v1/2020.louhi-1.7

Websites

American Cancer Society. (n.d.). What is endometrial cancer? *American Cancer Society*. Retrieved October 13, 2024, from https://www.cancer.org/cancer/types/endometrial-cancer/about/what-is-endometrial-cancer.html

Cambridge Dictionary. (n.d.). Retrieved from https://dictionary.cambridge.org/

Collins English Dictionary. (n.d.). Retrieved from https://www.collinsdictionary.com/

MedlinePlus. (2024). Birth control. *National Institutes of Health (NIH)*. Retrieved November 5, 2024 from https://medlineplus.gov/birthcontrol.html

Merriam-Webster. (n.d.). Retrieved from https://www.merriam-webster.com/

Microsoft Corporation. (2023). *Microsoft Translator Languages - Microsoft Translator*. Microsoft Translator for Consumers. https://www.microsoft.com/en-us/translator/languages/

Mujjiga, S., Krishna, V., Chakravarthi, K., & J, V. (2019). Identifying Semantics in Clinical Reports Using Neural Machine Translation. *Proceedings of the AAAI Conference on Artificial Intelligence*, *33*(01), 9552–9557. https://doi.org/10.1609/aaai.v33i01.33019552

Multidimensional Quality Metrics. (n.d.). *Multidimensional Quality Metrics (MQM)*. Retrieved September 5, 2024, from https://themqm.org/

National Cancer Institute. (n.d.). *NCI Dictionary of Cancer Terms.* Retrieved June 2, 2024, from https://www.cancer.gov/publications/dictionaries/cancer-terms/def

Neural machine translation by Systran. (n.d.). Systran by ChapsVision. https://www.systransoft.com/systran/translation-technology/neural-machine-translation-nmt/

Oxford University Press. (n.d.). *Oxford online dictionary*, from https://www.lexico.com

World Health Organization, Regional Office for the Eastern Mediterranean. (n.d.). *Unified Medical Dictionary (UMD)*. Retrieved June 8, 2024, from https://umd.emro.who.int/whodictionary

Yandex. (n.d.). How machine translation works. *Yandex*. Retrieved May 3, 2024, from https://yandex.com/dev/translate/doc/en/concepts/how-works-machine-translation