Evaluation of Instagram's Neural Machine Translation for Literary Texts: An MQM-Based Analysis

Altaf Fakih [a]
a.afakih97@gmail.com
School of Languages, Literacies and Translation
Universiti Sains Malaysia, Malaysia

Mozhgan Ghassemiazghandi [b]
mozhgan@usm.my
School of Languages, Literacies and Translation
Universiti Sains Malaysia, Malaysia

Abdul-Hafeed Fakih
a.hafeed1@gmail.com
Najran University, Saudi Arabia
&
Ibb University, Yemen

Manjet K. M. Singh
manjeet@usm.my
School of Languages, Literacies and Translation
Universiti Sains Malaysia, Malaysia

ABSTRACT

Addressing the global increase in social media users, platforms such as Instagram introduced automatic translation to broaden information dissemination and improve cross-cultural communication. Yet, the accuracy of these platforms' machine translation systems is still a concern. Therefore, this paper aims to explore the potential of Neural Machine Translation utilized by Instagram in producing high-quality translations. In doing so, this study attempts to scrutinize the reliability of Instagram's "See Translation" feature in the translation of literary texts from Arabic to English. A selection of auto-translated Instagram captions is analyzed through the identification, classification, and assignment of error types and penalty points, utilizing the MQM core typology. Subsequently, the Overall Quality Score of the error-based analysis is calculated automatically using the ContentQuo platform. Furthermore, the study investigates whether Instagram Neural Machine Translation can effectively convey the intended message within literary texts. From 30 purposively selected Instagram captions with literary content, the study found Instagram's machine translation lacking in 90% of cases, particularly in accuracy, fluency, and style. Among these, 61 errors were identified: 26 in fluency, 25 in accuracy, and 10 in style, adversely affecting the quality and failing to convey the original message. The findings suggest a need for enhanced algorithms and linguistic architecture in Neural Machine Translation systems to better recognize linguistic variants and text genres for more accurate and fluent translations.

Keywords: Literary Text Translation; Multidimensional Quality Metrics; Neural Machine Translation; Translation Quality Assessment

---

[a] *Main author*
[b] *Corresponding author*

eISSN: 2550-2131
ISSN: 1675-8021

INTRODUCTION

Instagram is a popular social networking platform where users share updates in the form of photo and/or audio-visual elements as posts. These posts are often accompanied by textual details, referred to as captions. According to Dixon's (2023) report on Statista, Instagram is currently the fourth most popular social media platform, with 1.28 billion active users as of January 2022, and is projected to reach 1.44 billion monthly active users by 2025. To address the growing global use of Instagram, Meta (formerly Facebook), the parent company of Instagram, is developing new Machine Translation (MT) innovations and MT systems to facilitate cross-lingual communication among users from different countries, aiming at improving the effectiveness of interactions among global users regardless of their language backgrounds. In this context, Meta has recently introduced a new innovative MT system, called the *No Language Left Behind (NLLB-200)*, built upon a single artificial intelligence-based model that uses neural networks which was claimed that they matched human's performance. Given the consistent developments in the MT field, regular evaluations of MT systems are highly required to monitor their quality and identify areas for improvement. To this aim, the current study seeks to closely examine the quality of the recent NMT system implemented in Instagram. Another gap that will be bridged in this study is the lack of the studies that address the quality of "See translation" feature of Instagram in translating literary texts in the Arabic context. As each language has its own unique characteristics, it is profitable to conduct evaluations on various languages and context that yields to revealing the language-related strong and weak aspects of each development, subsequently leaving a plenty of room for improvement.

The study's significance lies in its evaluation of machine translation, an essential area of research that helps improve the performance of existing MT systems and understand how they function (Dorr et al., 2011). Furthermore, Trigueros (2021) stressed that there is a need for more standardization for MT quality assessment and error analysis. Therefore, this study's findings can serve as a valuable reference for the translation technology field in general and for MT evaluation development, translation error-analysis methodology, and computational linguistics in particular. Additionally, this study will offer practical benefits by providing insights for Meta developers to improve the algorithms and linguistic architecture of their MT models, as well as grow awareness among Instagram users of the reliability level of the instant translations provided by the platform. Given the latest development of MT implemented in Instagram, this study aims to closely inspect the potential of "See Translation" feature in auto-generating adequate translations, thereby check on the concerns regarding the quality of such feature and highlighting where improvement is needed for Meta developers. To this end, this study seeks to achieve the following two objectives: a) To assess the quality of Instagram's Neural Machine Translation (NMT) system in translating literary texts by using an analytical error-based approach, which utilizes the MQM system that includes structured translation specifications, an error typology, and a scoring system integrated with the ContentQuo platform; b) To examine the NMT system's ability to convey the expressive function of literary texts, as defined by Nord's translation function theory.

eISSN: 2550-2131
ISSN: 1675-8021

LITERATURE REVIEW

MACHINE TRANSLATION OVERVIEW

Machine Translation is an interdisciplinary paradigm that involves different fields, including Natural Language Processing (NLP), which focuses on developing and optimizing computer-based translation systems (Ameur et al., 2020). It is also considered a branch of Computational Linguistics, which investigates the use of computer software to translate text from one natural language to another (Arnold et al., 1994; Sipayung et al., 2021). Due to its multidimensional nature, MT presents complexities different from those in Human Translation and is continuously evolving with the development of technology. MT is the process of translating text from one language into one or more other languages utilizing computer-based systems and tools, and may involve varying degrees of human intervention.

The increasing demand for translation, driven by economic globalization, exceeded human capacity to handle all translation tasks, leading to the introduction of automatic translation systems and resulting in significant changes in related fields. MT systems, according to their computational architecture (Chéragui, 2012), are classified into four approaches: Rule-based MT (RBMT) approach, Corpus-based MT (CBMT) approach, Hybrid MT approach, and Neural MT approach (Trigueros, 2022). The first approach was Rule-based MT, which used two linguistic sub-approaches: transfer and interlingua (Chéragui, 2012) that relied on monolingual and bilingual dictionaries, grammar and transfer rules for generating translations (Espana-Bonet & Costa-jussa, 2016; Castilho et al., 2017; Trigueros, 2022). Later on, Corpus-based MT was introduced as an alternative approach for MT in order to overcome the shortcomings of RBMT (Chéragui, 2012), and was the first approach of data-driven methods that used sophisticated algorithms and mathematical models to automatically learn the translation process from data (Ameur et al., 2020). CBMT used monolingual and bilingual corpora of parallel texts in the translating process (Hutchins, 1995). This approach was divided into two systems: Statistical MT system (SMT) and Example-based MT system (EBMT). The advantage of this system is that it requires less human effort for automatic training, along with its solid performance in terms of selection (Hutchins, 2007; Koehn, 2009; Trigueros, 2022). However, it sometimes outputs bad-quality translations that are ill-structured or grammatically incorrect attributed to the difficulty in reaching corpora of specific domains or language pairs (Habash et al., 2009; Espana-Bonet & Costa-jussa, 2016; Trigueros, 2022). Nevertheless, corpus-based systems dominated the field for a while as many MT developers adopted the approach to their MT systems, including Google Translate, Facebook, and Instagram. The hybrid MT approach combines both Rule-based MT and statistical MT systems, resulting in a solution that overcomes the deficiencies of each system and produces high-quality translations with a high level of precision (Thurmair, 2009; Hunsicker et al., 2012; Tambouratzis et al., 2014; Trigueros, 2022).

Moreover, most recently, a new data-driven MT approach, called Neural Machine Translation (NMT) has been developed with a different mechanism. NMT is the latest technology in Artificial Intelligence (AI), which consists of a system that uses neural networks and works in building and training a single large neural network that reads a sentence and outputs correct translations (Bahdanau et al., 2014; Trigueros, 2022). This system is based on the encoder-decoder model in which the encoder reads the input and encodes it into a fixed length vector while the decoder produces the translation output from the encoder vector (Cho et al., 2014; Bahdanau et al., 2014; Trigueros, 2022). NMT represents the latest development of MT systems, which has become the dominant paradigm that is currently applied in machine translation field (Ragni &

Vieira, 2021; Trigueros, 2022). Moreover, Trigueros (2022) pointed out that the architecture of NMT is characterized by some advantageous properties that prior MT systems do not own. For instance, it uses fewer components and processing steps, and it requires less memory than SMT. Moreover, it allows the use of human and data resources more efficiently than RBMT (Cho et al., 2014; Bentivogli et al., 2016; Trigueros, 2022). Furthermore, the findings revealed that NMT output contained fewer overall errors compared to SMT at the accuracy and fluency levels (Wu et al., 2016; Castilho et al., 2017; Moorkens, 2018; Ragni & Vieira, 2021) since the neural networks can be trained to recognize patterns in data and deal with massive amount of language data with much ease, hence making NMT output more accurate (Das, 2018). Such characteristics have pushed Meta, along with many other major companies, such as Google, Systran, and Microsoft (Ameur et al., 2020; Trigueros, 2022) to shift from SMT and RBMT approaches to Neural MT approach.

NMT OF INSTAGRAM

In 2017, Meta, Instagram's parent company, announced its shift from phrase-based statistical machine translation to neural machine translation (Mannes, 2017), resulting in more accurate and fluent translations (Pino et al., 2017). In 2020, Meta introduced a new neural machine translation model, the multilingual machine translation (M2M-100), which automatically translates between any pair of 100 languages, including translation across 2,200 language pairs, without relying on English as an intermediary source. The M2M-100 model aims to improve translation quality for low-resource languages (Bhattacharyya, 2022). Additionally, Meta developed a single artificial intelligence-based model, the *No Language Left Behind (NLLB-200)*, which translates 200 languages, including those not adequately addressed by machine translation tools in Instagram. The NLLB-200 model aims to improve the quality of machine translations and facilitate communication worldwide. Meta evaluated the NLLB-200 model using automatic evaluation metric, the BLEU algorithm, which measures how closely machine translations match human translations and reported that it achieved BLEU scores that were 44% higher than any previous record (Meta, 2022). Therefore, this study investigates whether the new advanced model, NLLB-200, of Instagram MT can make any improvements in this respect.

MT QUALITY EVALUATION

Various studies have evaluated the quality of Instagram machine translation (MT) since its introduction. Fadilah (2017) identified three types of semantic errors in its output: referential, grammatical, and contextual. Grammatical and contextual errors were the most frequent, while the translation of dictionary meaning performed better. Mawarni et al. (2017) focused on cultural-specific terms (CSTs) and found a loss of meaning in the translations, which failed to transfer the expressive meaning to the target culture. In line with previous findings, MT succeeded in translating referential meaning but failed in translating pragmatic meaning.

Furthermore, Meilasari (2019) evaluated the accuracy of Instagram MT translations related to ecology and environment vocabulary. The study found that the MT was unreliable, with 40% of the translations being inaccurate and only 24% being accurate. Susanti (2018) analyzed Instagram MT translations and identified incorrect and missing words as the most frequent lexical errors. The study also found that the MT tended to use a word-for-word translation method, resulting in a lack of recognition of the text's context and failing to represent the authentic language. Other researchers have compared the quality of Instagram machine translation with human translation. Arvianti (2018) compared the performance of Instagram MT and human translators in translating formal and informal language. The study found that while Instagram MT produced good translations for formal language, it failed to translate texts written in an informal language. Human translators were better able to recognize particular languages and better understand context due to their more extensive vocabulary and context understanding. In addition, Instagram MT translations have been compared to output from other MT systems. Larassati et al. (2019) evaluated the output of neural machine translation utilized in Google Translate and Instagram and found that both systems had translation errors, with Instagram MT having more errors. The most frequent error types were terminology errors, syntax errors, and literalness, which were interrelated. Similarly, Pujakesuma (2022) found that Google Translate and Instagram MT made similar errors, such as mistranslation, and applied the same translation strategies, such as literal translation.

Moreover, some researchers have evaluated Instagram MT by exploring its translation strategies. Purwaningsih (2019) investigated the translation strategies employed by Instagram MT in translating culturally specific Indonesian items, particularly Banyumas Batik motifs. The study found that Instagram MT used three techniques, including literal translation, borrowing, and particularization, with borrowing being the dominant technique for translating cultural items. However, this led to a loss of the cultural sense. Purwaningsih (2019) recommended that Instagram developers enrich the MT with a more extensive contextual linguistics database to improve the quality of translation results.

The current study will focus on evaluating the output of the new innovative model "NLLB-200" recently implemented and claimed to produce more accurate translations than the prior MT models that were examined by the previous studies. Furthermore, the existing literature on Instagram NMT evaluation has apparently focused on the Indonesian-English language pair, leaving a research gap for Arabic language translation. Ameur et al. (2020) note that there are still many linguistic problems related to Arabic that require further investigation as they pose significant challenges to current Arabic MT systems. Therefore, this study aims to fill this research gaps by evaluating the translation quality of the new AI-powered MT system for Arabic captions.

TEXT TYPE

Nord (2005) proposed a tripartite model of the functions of linguistic signs inspired by Bühler's (1934) work, which includes four basic functions of communication in language: referential, expressive, operative, and phatic. The referential function focuses on the meaning or content referred to and represented in informative texts, such as scientific articles and news. The expressive function refers to the emotions and attitudes of the sender towards the referred object, thought, or idea, as often found in texts of high aesthetic value, such as literary works. The operative or appellative function is concerned with the direction of the text toward the addressee. The phatic function, attributed to Roman Jakobson, focuses on establishing communication between sender and receiver and attracting the attention of the receiver regarding certain things. The expressive function implied in literary texts is the focus of this study; it seeks to explore how Instagram NMT can deal with the unique sentence structures, cultural elements, and aesthetic features present in literary language that have fewer counterparts stored in the MT database. Additionally, the study aims to investigate the extent to which this system can convey the expressive function inherent in literary texts.

eISSN: 2550-2131
ISSN: 1675-8021

METHOD

RESEARCH DESIGN

The present study utilized a qualitative descriptive method and analyzed the written content (captions) taken from the @cairo_mockingbird Instagram account, a virtual platform for visual arts and literary writings. The account contains over 12,000 posts, each featuring a photo and a caption written in Arabic (as of last access on 24/3/2023). It is a community platform designed for displaying visual arts and literary writings. To achieve the objectives of the study, the selected data were analyzed by using a non-DEJ-based analytical evaluation method called the Multidimensional Quality Metrics (MQM) core typology.

The latest version of MQM, from October 2021, was developed to allow a harmonization with TAUS DQF (Dynamic Quality Framework) error typology, which resulted in creating a flexible subset to MQM. The MQM error typology contains eight high-level dimensions; seven dimensions are the core and the eighth is additional to provide a wide range of more detailed error types that can be used where implementers require greater granularity. The tree view format illustrated in Figure 1 below depicts the MQM-Core error typology. Each dimension consists of more specific error subtypes:

- Accuracy contains three subtypes: mistranslation, over-translation, under-translation, addition, omission, Do not translate (DNT), and untranslated. It involves the errors occurring when the target text does not accurately represent the propositional content of the source text, either by distorting, mistranslating, omitting, or adding to the original message.
- Fluency (or Linguistic Conventions) comprises four subcategories: grammar, punctuation, unintelligible, and character coding. It focuses on the errors related to the linguistic form of the text, including problems with grammaticality, orthography, and other mechanical correctness.
- Terminology category includes three subcategories: inconsistent with terminology resource, inconsistent use of terminology, and wrong term and it includes incorrect terms in the target text that are not equivalent of the corresponding term in the source text.
- Style includes organizational style, third-party style, inconsistent with external reference, register, awkward style, unidiomatic style, and inconsistent style. Style refers to the errors that are grammatically acceptable but are inappropriate because they deviate from inappropriate language style or organizational style guides.
- Audience appropriateness contains only one subcategory: cultural-specific reference. In this category, the errors arising from the use of content in the translation product that is invalid or inappropriate for the target audience are addressed.
- Locale conventions are the issues related to the locale-specific content (e.g., date/name format, calendar type, postal code, locale-specific punctuation, or national language standard) or formatting requirements for data elements.
- Design and markup include the issues related to the physical design (e.g., graphics and tables) or the layout of a translation product.
- Custom: Any other issue observed or suggested by the evaluator(s) can be added to this category.

This study employed a non-DEJ-based evaluation method, in which the judge (annotator) assesses the translation quality indirectly. Such evaluation methods are commonly used to evaluate the accuracy and fluency of both human and machine translation results and involve comparing either the source text with the target text or the target text with the translation reference (Chatzikoumi, 2020). The rationale behind the choice of MQM core typology over other existing analytical-based approach lies on two reasons. Firstly, this approach is based on a functional-oriented perspective that was formulated on Melby's (2002) paralleled work, which parallels Skopos theory and the translation brief, and Nord's extension of Skopos theory (1997), which is known as Functionalism in translation theory and practice. This mainly serves the fulfillment of the study's second objective that involves investigating the expressive function included in the MT translations. Secondly, MQM-core typology is characterized by its flexibility and usability (Lommel et al., 2013). That is, the framework can be adjusted in a way that it serves the purpose of the analysis and accounts for specific required needs. Besides, it should be noted that the MQM can be applicable to professional translations as well as to MT output, i.e., the metric is designed to evaluate the translation product, regardless of how the target text is generated.

MQM-core typology quality assessment metric includes three different stages for the evaluation. The first stage is the *Preliminary Stage* that is conducted before the evaluating process and includes three phases: Translation Specifications Evaluation, Evaluation Metric Design, and Data Collection. The second stage *Error Annotation* is where annotation and error analysis of the data is taken place, and lastly, the third stage *Automatic Calculation* includes the calculation of the Overall Quality Score of the analysis. Second and third stages were conducted in ContentQuo platform to assure more accurate results. Each evaluation stage is elaborated in details in next section.

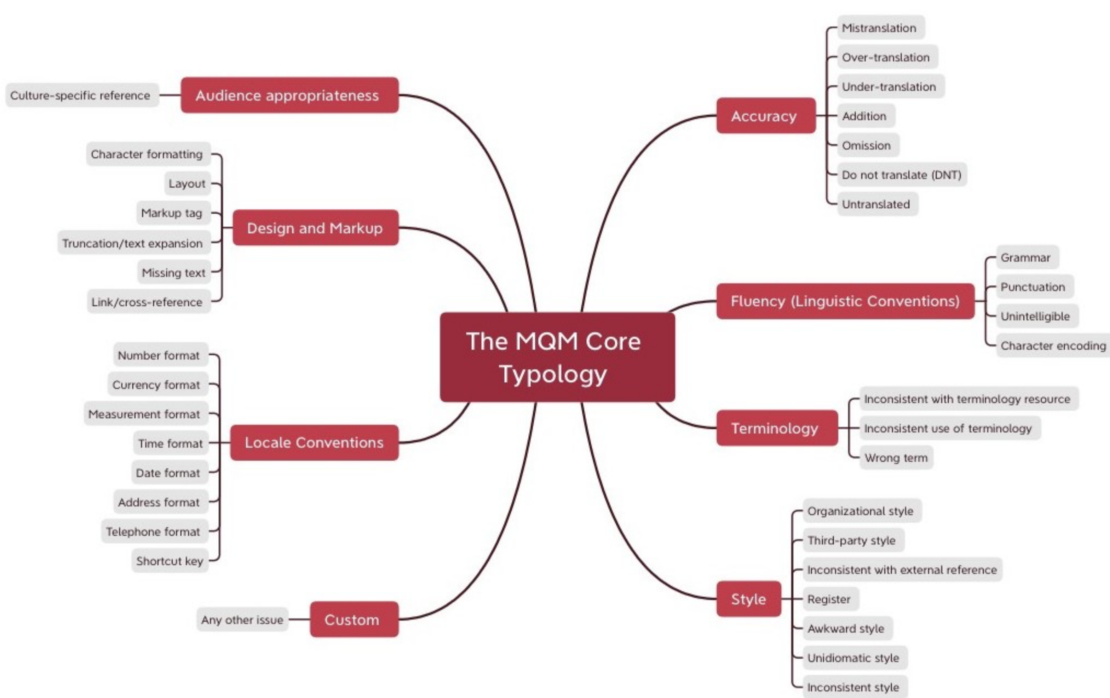

FIGURE 1. The MQM-Core Error Typology (http://www.themqm.info/ (Last access 7/3/2023)).

TRANSLATION QUALITY EVALUATION (TQE) STAGES

STAGE 1: PRELIMINARY STAGE
TRANSLATION SPECIFICATIONS

In this phase, we determined translation parameters that should be met adopted from the 2006 ATSM Standard Guide for Quality Assurance in Translation's structured translation specification framework (ASTM F2575-06). They include metadata of the text under evaluation and its original. This step is prerequisite as it works as a guideline for the evaluators or annotators to determine the translation quality parameters that the translated text should meet and upon which the text should be evaluated. The translation parameters, as shown in Table 1 and 2, and adopted in this paper, were selected to align with the objectives of this study.

TABLE 1. Source Content Information

Textual characteristics	
Source language	Arabic (Modern Standard Arabic and Egyptian Arabic)
Text type	Literary texts
Audience	Instagram users who are familiar with Arabic language and culture
Purpose	Expressive function: the text is intended to convey a particular message in the mind of an author in an artistic form.
Specialized language (Subject field)	The captions consist of sayings and texts quoted from novels and other literary sources.
Specialized language (Terminology)	The texts do not include specialized or complicated terminology, but rather everyday use vocabulary. Therefore, it does not require a specialized term base.
Volume	30 captions (376)
Complexity	Some captions are written in a straightforward form while some others in an artistic style.
Origin	The source texts are captions posted on @cairo_mockingbird Instagram account.

TABLE 2. Target Text Requirements

Target language	English
Audience	Instagram users who can understand English.
Purpose	Expressive function
Content Correspondence	The ST should be translated accurately and fluently.
Register	Texts written in Modern Standard Arabic should be translated into formal English while texts written in Egyptian Arabic should be translated into informal-colloquial English.
Format	Captions underneath a photo on Instagram
Style	Stylistics should be taken into consideration in translating the ST.

EVALUATION METRIC DESIGN

A metric is a measurement with a specific purpose (Lommel & K. Melby, 2018). Due to the scope of this study, researchers did not include all the dimensions appeared in Figure 1 and designed a specific metric for evaluation that served the aim and objectives of the study. To verify the metric of the evaluation, three dimensions were selected: accuracy, fluency, and style, as per the MQM core typology and its subsets, as shown in Figure 2. The main goal of an MT system is to automatically translate text while preserving its meaning and style, ensuring that the output is as linguistically fluent as possible (Ameur et al., 2020). Accordingly, the evaluation focused on three aspects: accuracy (adequacy), which considers the semantic and pragmatic equivalence of lexis between the source and target texts; fluency, which refers to the linguistic conventions of the target language and naturalness (Chatzikoumi, 2020); and style, which measures the extent to which the translated text uses appropriate language to convey the message effectively. Thus, the evaluation encompassed the lexical, syntactic, semantic, pragmatic, and stylistic aspects of the translated texts. The errors extracted from the TT were measured according to the following Error Severity Levels:

1. Minor errors, which do not affect the comprehension of meaning but affect the fluency (Weight: 1)
2. Major errors, which make TT difficult to understand, yet the general message is conveyed. (Weight: 5)
3. Critical errors, which change the meaning of ST and make it incomprehensible or distorted (Weight: 10)

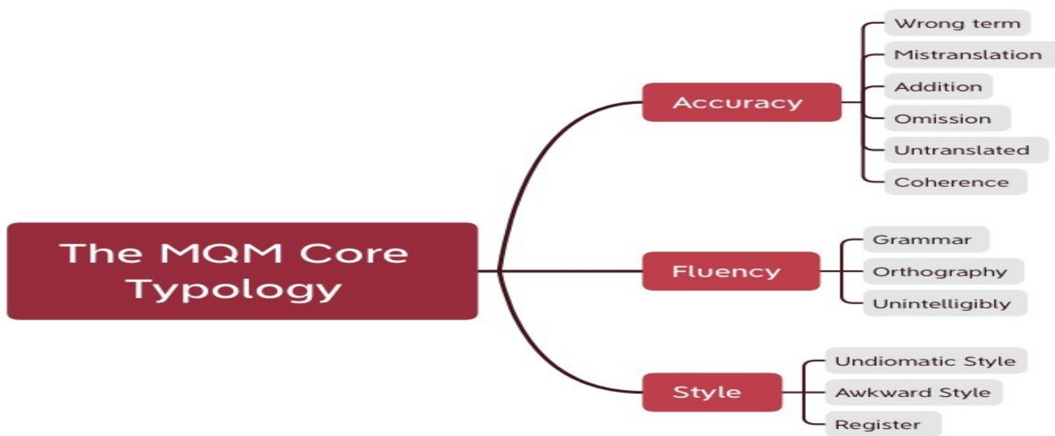

FIGURE 2. A Metric Designed for the based on the MQM Framework Evaluation in the Present Study

DATA COLLECTION

The data collection process included three phases: (i) selecting the source of the data, namely, an Instagram account, (ii) selecting the data (captions), and (iii) collecting the selected data. Data source selection phase is determined by the following criteria:

- The source material should contain the data that are necessary to answer the stated research objectives.
- The data included in the source should be within the scope of the study.
- The source should be a verified account with a substantial number of followers.
- Data included in the account should be in form of captions (texts) and not audio-visual elements.

The researchers selected the @cairo_mockingbird Instagram account as it met the data source selection criteria. The account shares a variety of literary writings daily, providing ample samples for the evaluation and contributing to answering the research questions. All captions are written in Arabic, ensuring the data remains within the study's scope. Additionally, the account is verified with 826 thousand followers (as of last access on 24/3/2023). Finally, the account often uploads literary writings illustrated in a photo with the text replicated in a caption below the photo, making the data easily accessible.

In the study, there are two types of data: first, the original captions written in Arabic, referred to as "Source Text" (ST), and secondly, the English machine translations that is referred to as "Instagram Machine Translation" (IMT). The researchers added another additional data that includes human translations, referred to as "HT", for the captions as a reference for the reader. The two types of data were collected manually by the researchers using a purposive sampling. Firstly, the researchers read intently all the captions posted on the @cairo_mockingbird Instagram account. Secondly, 30 captions were selected purposively, ranging from short to medium-length sentences (total of 376 words) written in Modern Standard Arabic (MSA) and Egyptian Dialect in the form of a poetic language. Thirdly, the researchers collected the translated results manually after tapping on "انظر للترجمة" or "See Translation" feature set beneath the selected captions, which instantly translates the captions into English, the language set in their personal Instagram application. Finally, source texts (the captions) and target texts (their translations) were collected and divided into segments. Each segment pair, containing corresponding content (a source text and target text), is termed a translation unit (TU).

STAGE 2: ERROR ANNOTATION

In this stage, the annotation was conducted semi-automatically using the harmonized MQM-Core Typology and DQF error typology, integrated with ContentQuo platform. The annotators (two experienced translators along with a skilled linguist) examined the translated text against the source text based on the agreed translation specifications, and analytically annotated errors, which involved identifying, classifying, and assigning error type and penalty points, in accordance with the designed metric.

STAGE 3: AUTOMATIC CALCULATION

At this stage, the Overall Quality Score was calculated automatically by ContentQuo according to the selected scoring model using the following formula: QualityScore = 100 - 100 * (ErrorPoints / Wordcount), then compared to the Threshold Value (100%) to assign a pass/fail rating.

ANALYSIS AND RESULTS

This section shows the results of the analytical evaluation conducted on the Instagram NMT translation of 30 captions selected from @cairo_mockingbird account using ContentQuo platform. As illustrated in Figure 3, Instagram NMT failed at translating 90% of the data from three different aspects: accuracy, fluency, and style. 61 errors were found in the selected data, classified as follows: 26 errors in Fluency, 25 errors in Accuracy, and 10 errors in Style. The severity of these errors ranged from minor to critical, as depicted in Figure 4 and 5. In accuracy, Quality Score was the lowest because the errors found in this category were critical and seriously affected the content message as only 39.9% of the content was translated correctly. Fluency category came in the second lowest Quality Score as the fluency-related errors was less severe on the original content message than the accuracy-related errors. Lastly, as inferred from the style issues that they slightly affected the meaning of the sentences in which they were found. Each category has been explained in more detail below.

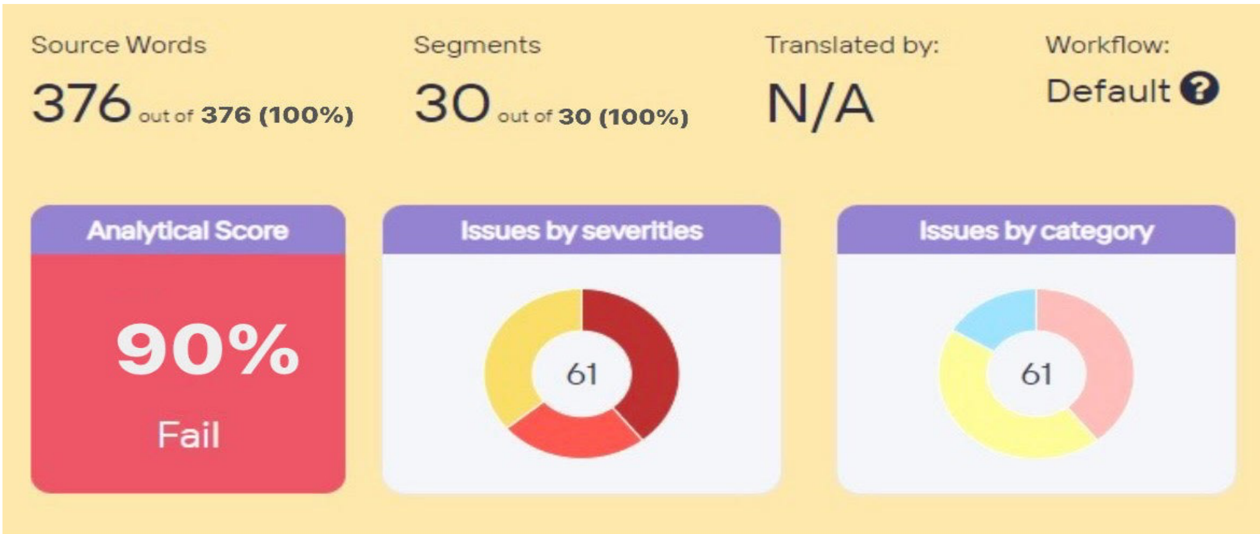

FIGURE 3. The Overall Quality Score at ContentQuo

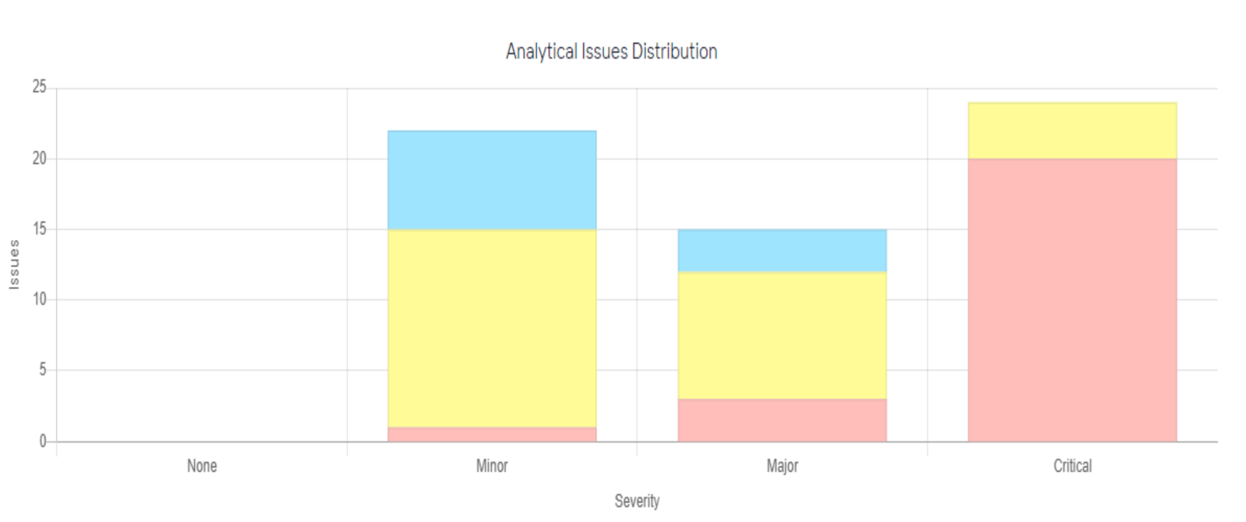

FIGURE 4. Issue Severity Levels

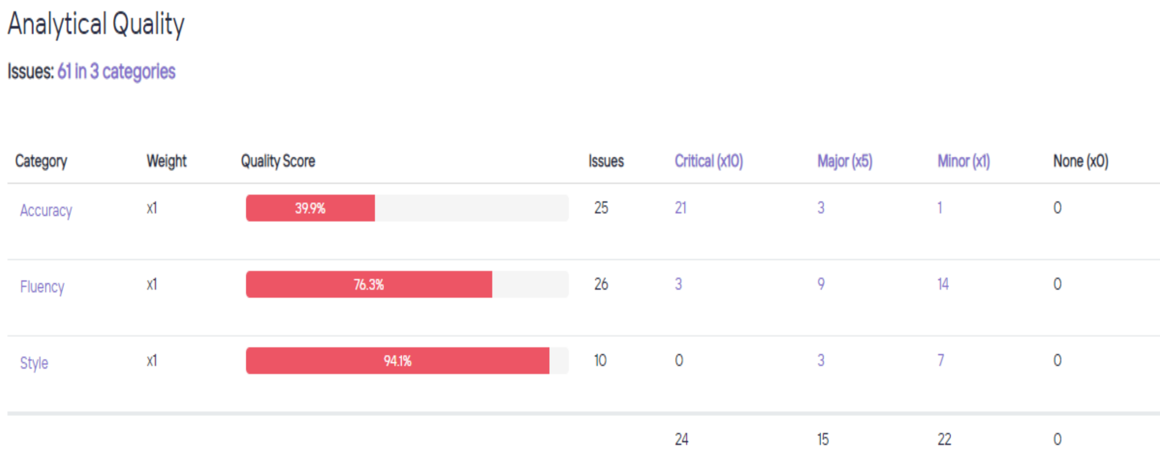

FIGURE 5. Quality Score of Each Category

ACCURACY

Accuracy categories concerned how the MT system recognized the meaning of the source text and reproduced it in the target text. Based on the results of the TQE, Instagram MT produced plenty of errors under this category. For instance, Instagram MT system was unable to recognize the exact meaning of the ST term within the context, thereby failing to find an appropriate equivalent. As shown in Table 3, The TM was unable to recognize the exact meaning of the polysemous word (بحث search), so it incorrectly chose 'research' as the equivalent in the TL.

TABLE 3. Example 1 of Inaccuracy

TU	ST	IMT	HT
1	الأمان جميل جداً، أظنه الشعور الوحيد الذي يستحق عناء البحث.	Security is so beautiful, I think it's the only feeling worth the effort to research.	Security is so beautiful. I think it is the only feeling worth seeking out.

Moreover, in literary texts, authors sometimes represent the message they want to express as a figure of speech, such as a metaphor. Instagram MT struggled with understanding and translating the metaphors in the source texts. In Table 4, the MT system's word-for-word translation of the caption distorted the intended meaning of the metaphor. The vehicles (سماء Sky) and (أرض Earth) that carry the meaning of the topic (God) and (People) indicates that God is above the sky and people are down on the earth. The MT translation failed to convey the ground relationship implied, resulting not only distorting the meaning but also the aesthetic value of the literary device, i.e., metaphor.

TABLE 4. Example 2 of Inaccuracy

TU	ST	IMT	HT
2	نَلتَمِسُ بالسماء ما تَرفضُ الأرضُ أن تمنحهُ لنا.	We touch heaven what the earth refuses to give us.	We ask God what people refuse to grant us.

eISSN: 2550-2131
ISSN: 1675-8021

Furthermore, within the same caption, it was found that Instagram MT system tended to misread the captions even though they were written in a direct sentence structure and partially or fully vocalized. Unlike English, Arabic language is characterized by having no letters to represent the vowel sounds. Instead, the Arabic writing system uses small signs that are added above or below the letters as vowel sounds called diacritics and the presence of such diacritical signs is known as "Vocalization" (Ameur et al., 2020). Vocalization clarifies the way of reading words and their exact meanings which indeed helps in solving lexical, semantic, and pragmatic ambiguities in translation. The MT system in the present study failed to misread these signs, hence reproducing wrong equivalents. In Table 4, the verb (نَلْتَمِسُ), meaning (ask for) was mistranslated into (touch نَتَلَمَسَ). It can be concluded that the MT system still cannot decide the exact equivalent for a word with or without vocalization.

TABLE 5. Example 3 of Inaccuracy

TU	ST	IMT	HT
3	نجيب محفوظ	Najeeb is safe	- Naguib Mahfouz
4	د. جاسم المطوع	Jasim the volunteer - د	- Dr. Jasem Al-Mutawa

Additionally, one of the most frequent translation errors that Instagram MT produced was the mistranslation of proper names. As demonstrated in Table 5, the MT system transliterated the first names while it translated literally the surnames. This type of proper nouns falls under "Adjective Constituent" noun-compound class; it consists of noun and adjective which are connected with each (Bounhas & Slimani, 2009; Omar & Al-Tashi, 2018). This is a common issue occurring when translating from Arabic into English because the Arabic language lacks a unified system or rules in writing named entities in, such as capitalization. Additionally, the rich lexical variations and highly inflected nature of Arabic further complicate this issue.

FLUENCY

Fluency error categories include errors related to the linguistic well-formedness of the translated text, including morphology, syntax, orthography, and sentence readability. The evaluation results showed that fluency errors were the most frequent errors produced by Instagram MT. These errors range from minor ones affecting only the TT's fluency, to major errors that make the text hard to understand but convey the general message, and critical errors that distort the meaning and make the TT unintelligible.

One of the root causes that led to the fluency errors was the flexible word order of Arabic language. Unlike English language that has only one rigid SVO word order, Arabic has a flexible sentence structure that can occur in multiple orders, such as SVO, VSO, OVS, etc. These flexibility poses several problems when translating from Arabic into English. MT systems, built on fixed encoding and decoding mechanisms and algorithms, often get confused by the multiple sentence structure (i.e., word order) that Arabic language can take, particularly in literary texts. Therefore, these MT systems fail to produce the Arabic text into the TT. For example, as shown in Table 6, Instagram MT mistranslated the Arabic sentence that has an OVS word order, resulting in an unintelligible output.

TABLE 6. Example 1 of Non-Fluent Translation

TU	ST	IMT	HT
5	على دفء العائلة تتكئ البيوت.	Warmth of family leaning homes.	A home rests on the warmth of a family.

Another problem that was observed during the TQE of Instagram MT translations was that they lacked pronoun-antecedent agreement. In English, the pronoun and its antecedent (the word to which a pronoun refers) must agree in number, person, and gender. The MT system translated each caption segment separately. The pronoun (i.e., it) in the target text in Table 7, for instance, contradicts with its antecedent (years) in number. The MT system read and translated the two sentences independently, out of the context, resulting in incohesive translations.

TABLE 7. Example 2 of Non-Fluent Translation

TU	ST	IMT	HT
6	لأعوام تغير الكثير.. أنها تبدل تضاريس الجبال، فكيف لا تبدل شخصيتك؟ - أحمد خالد توفيق	The years changed a lot… It changes the mountains, how can it not change your character? - Ahmed Khaled Tawfiq	Years make a lot of changes. They change the terrains of mountains, let alone your character? -Ahmed Khaled Tawfik

Furthermore, another frequent issue was errors related to orthography as shown in some samples, which involve the target language's conventions of writing, such as norms of spelling, hyphenation, capitalization, word breaks, emphasis, and punctuation. These errors might not be critical, but they negatively affect the readability of the translations. It was noticed that Instagram MT tended to imitate the ST writing conventions which resulted in poorly written translations. This strategy might be usable in languages that have similar writing norms, but in our case, the source language and the target language have completely different orthographic systems, it led to considerable issues, such as small letters at the beginning of a sentence and capital letters in the middle of a sentence, and a lack of proper punctuation marks, among other things.

STYLE

Literary writings, as expressive texts, highly value the form of texts. Stylistics played a significant role in the evaluation Several stylistic errors were found in Instagram MT output. The MT system used basic translation strategies, such as literal and word-for-word translations with all types of texts, be it informative, expressive, or persuasive. While literal translation could work in informative texts that focus only on the content, in the literary texts that also value the form, it was a root cause of generating translations that lacked the aesthetic values and had awkward sentence structures, as illustrated in Table 8.

TABLE 8. Example 1 of Stylistic Errors

TU	ST	IMT	HT
7	لستُ أفهمُ من معنى الحبِ إلا أن الرُوح قد اهتدت إلى شيء من سِر الإنسانية في إنسانٍ جميل. - مصطفى الرافعي	I do not understand the meaning of love, except that the soul has been guided to something of the secret of humanity in a beautiful human being. - Mostafa Al-Rafay	The only thing I can understand about the meaning of love is that the soul has found a secret of humanity in a beautiful human being. - Mostafa Al-Rafe'ie

eISSN: 2550-2131
ISSN: 1675-8021

Idiomatic expression translation can be problematic, especially in machine translation that most of times these expressions end up being translated literally. It occurs when there are linguistic or cultural gaps between the SL and TL. However, Instagram MT failed to translate expressions that had one-to-one direct equivalent, producing unidiomatic style in the TT. This is clearly demonstrated in Table 9, where Instagram MT translated the ST literally, despite the existence of a direct equivalent in English.

TABLE 9. Example 2 of Stylistic Errors

TU	ST	IMT	HT
8	ما تزرعه اليوم تحصدهُ غداً..	What you plant today you will harvest tomorrow.	You reap what you sow.

Another issue commonly found in the output of Instagram MT was the lack of conformity to the register of the ST. The translations seemed to have informal style by using colloquial terms and contractions to reproduce the formality of the ST that represented in in using Modern Standard Arabic. This issue is demonstrated in Table 10.

TABLE 10. Example 3 of Stylistic Error

TU	ST	IMT	HT
9	مش معنى إن حد شايل الشيلة كويس يبقى الشيلة مش تقيلة!	It doesn't mean that someone is carrying the burden well Then the burden is not heavy!	Just because someone else is carrying the burden well, it doesn't mean the burden is not heavy!

Despite the above-mentioned weaknesses in Instagram MT system, the system has shown improvement in some other aspect. It was able to properly translate texts written in the Egyptian dialect. As shown in Table 10, the MT system managed to recognize the colloquial words (شايل), (الشيلة), and (كويس), and translated them into their proper equivalent terms in English (is carrying), (the burden), and (well).

DISCUSSION

This small-scale exploration questioned whether Instagram MT is capable of producing accurate and fluent translations that well maintain the intended message implied within the literary texts to the target language. The results of the evaluation revealed that the MT system produced several translation errors, covering different linguistic aspects including lexis, syntax, semantics, pragmatics, orthography, and stylistics, which hindered the process of transferring the accurate meaning of the source texts in fluent well-structured translations which definitely go against the translation specifications, Table 1 and 2, that were set by the researchers before the evaluation. These results contradict the concept of translation quality as defined by Koby et al. (2014), who defined translation quality as reproducing accurate and fluent translation results for the target audience that can serve the original purpose and comply with all other specifications negotiated between the requester and provider, while considering the needs of the end-users. Consequently, Instagram's NMT system is not capable of producing translations that are well-structured, properly convey the intended message, and preserve the aesthetic value of the literary texts. Despite significant improvements made on Instagram's NMT, e.g., the ability to recognize dialectical element as shown in Table 10, these linguistic aspects, such as accuracy, fluency, and adherence to stylistic conventions, remain challenging.

Issues related to accuracy included producing wrong equivalents in the TL either because the MT could not recognize the exact meaning of the term in the context as in Table 3 and 5, misread vocalization as in Table 4. This is in line with Susanti (2018), who questioned the reliability of Instagram MT translations by exploring the lexical errors and found out that MT is prone to generate mistranslated words, incorrect translation, and unknown words. Likewise, Cahyani et al. (2021) stressed that Instagram MT produced inappropriate translation because it used improper procedures by choosing the lexis in the target language literally through a term that has several synonyms which have different meanings in each use without considering the overall context of the caption. Additionally, literal texts usually include literary devices, such as metaphor, that consider a unique structure of language that implies contextual and cultural nuances. As we can see in Table 4 that Instagram MT apparently does not have the flexibility to deal with unusual forms of texts that do not have direct parallel structures in its database, and it fails to recognize the contextual and cultural knowledge implied in the inputs as it only uses literal translation to produce the dictionary meaning of the linguistic units (Purwaningsih, 2019; Omar & Gomaa, 2020). this issue. In a similar vein to the findings of Purwaningsih et al. (2019) and Meilasari (2019), Instagram NMT is still unable to recognize proper nouns, especially those come under "Adjective Constituent" noun-compound class as illustrated in Table 5. Due to complexities of the morphological differences of Arabic language, it possesses numerous types of noun compound types and the extraction of Arabic noun compounds is one of the challenging tasks in machine translation (MT) where in Arabic the words do not have capital or small letters which causes semantic ambiguity exactly as what happened when Instagram NMT attempted to identify and translate the names of well-known Arab writers (Naguib Mahfouz and Jasem Al-Mutawa) in Table 5. Though such proper nouns are well-known and frequently occur together, the MT system fails to recognize the context in which they occur and identify them as names not just nouns or adjectives. This issue can be attributed to two root causes: the lack of capitalization and available resources of Arabic noun compound lexicon (Omar & Al-Tashi, 2018). To overcome this limitation, we need to extract those compound nouns to process it further as well as improvement should also include the Named Entity Recognition (NER) and Part-of-Speech tagging tasks implemented in the MT. NER and POS tagging are responsible for identifying, determining, and classifying proper names in a text which can help compensate the absence of capital letters in the Arabic language that considers the main difficulty to achieving high performance in automated translations (Alkhatib & Shaalan, 2018).

Issues under Fluency category, including producing unintelligible output, incohesive translations and orthography-related errors, can be attributed to the significant linguistic differences between the two languages. Arabic and English belong to different language families, respectively, and have distinct grammar rules, morphology, semantics, pragmatics, and writing conventions. These differences make it difficult for MT systems, resulting in inadequate translations. Morphologically rich languages like Arabic pose to accurately recognize and effectively bridge these linguistic and cultural gaps even more significant challenges for MT systems. This flexibility in word order within Arabic, for example, makes it difficult for translation systems to make accurate choices, negatively impacting the quality of translations (Ameur et al., 2020; Omar & Gomaa, 2020).

Literary texts pose a greater challenge for machine translation systems due to their unique style and use of figurative language, special diction, and language enhancers that carry implied meanings beyond words and sentences. Arvianti (2018) pointed out that compared to human translators, MT systems have a limited vocabulary and struggle with context understanding, making it difficult for them to recognize specialized language. Omar and Gomaa (2020) explored the challenges of applying MT systems to literary translation and concluded that despite the occurrence of errors, the usefulness of MT systems should not be underestimated. It is clear from the data shown that some of the resultant problems are brought about by translating the texts literally using basic translation strategies, such as word-for-word translations, as shown in Table 8 and 9, without considering the contextual, cultural, and aesthetic references which are essential in literary texts, hence resulting in a loss of the expressive sense, and in some cases, the whole meaning gets distorted.

Meilasari (2019) concluded that Instagram translation machine is not a reliable feature for the target language reader who wants to understand certain language or cultural-related terms in the source language because the MT only produces the translation product literally based on what is provided by the source text and has no ability to analyze and restructure the text that is being translated. As a matter of fact, MT technology has been introduced and implemented to social networking platforms as a response to the rising demand for multilingual content. It can be considered as a vehicle for accessibility as it provides a means for across nation communication to take place in a way it bridges between languages and cultures in which users lack proficiency. However, translation errors generated from MT systems can have negative impacts on the end-users' experience because, as the findings of this study shown, such errors affect the readability either by distorting the fluency, accuracy or the style of the original content.

CONCLUSION

This paper evaluates the output of Instagram's automatic translation of Arabic literary writings. The evaluation involves an analysis of translated texts, the identification, classification, and assignment of error types, along with the allocation of penalty points using the MQM core typology. The study explores the ability of Instagram MT to convey the implied message in literary texts. The findings indicate that Instagram MT fails to successfully translate 90% of the data at three levels: accuracy, fluency, and style. Specifically, the selected data exhibited 61 errors, comprising 26 in fluency, 25 in accuracy, and 10 in style. These errors significantly affect the quality of translations, thereby impeding the transfer of the intended message embedded within the source texts. The evaluation results reveal multiple translation errors. These errors negatively impact the translations' accuracy, fluency, and style, hindering the conveyance of the intended message of the source texts.

LIMITATIONS AND FURTHER RESEARCH

This study examines the quality of Instagram's AI-based machine translation for translating literary Arabic texts into English. Further research can investigate other aspects of the Arabic context and compare the linguistic needs of Arabic with those of other languages in MT systems. As AI continues evolving, further evaluations are necessary to assess MT applications and text genres in different language pairs for the purpose of further exploring this promising technology and enhancing its output for assuring more sustainable circulation of information worldwide and enriching end-user experience.

ACKNOWLEDGEMENT

We sincerely thank ContentQuo, a translation quality management SaaS, for their indispensable support in providing localization services and CAT tools to evaluate raw MT and PEMT output using Error Annotation and Rating Scale.

REFERENCES

Alkhatib, M., & Shaalan, K. (2018). The key challenges for Arabic machine translation. *Intelligent Natural Language Processing: Trends and Applications*, 139-156.

Ameur, M. S. H., Meziane, F., & Guessoum, A. (2020). Arabic machine translation: A survey of the latest trends and challenges. *Computer Science Review*, *38*, 100305. https://doi.org/10.1016/j.cosrev.2020.100305.

Arnold, D., Balkan, L., Meijer, S., Humphreys, R., & Sadler, L. (1994). *Machine translation: An introductory guide*. London: Blackwell.

Arvianti, G. F. (2018). Human translation versus machine translation of Instagram's captions: Who is the best? In *English Language and Literature International Conference (ELLiC) Proceedings* (Vol.2, pp.531-536).

Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*.

Bentivogli, L., Bisazza, A., Cettolo, M., & Federico, M. (2016). Neural versus phrase-based machine translation quality: a case study. *In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (pp. 257–267). arXiv preprint arXiv:1608.04631*.

Bhattacharyya, S. (2022). *Meta's machine translation journey*. Analytics India Magazine. Retrieved March22,2023, from https://analyticsindiamag.com/metas-machine-translation journey/#text=Meta%20used%20neural%20machine%20translation,training%20time%20from%2024%20hours.

Bounhas, I., & Slimani, Y. (2009). A hybrid approach for Arabic multi-word term extraction. In *2009 International Conference on Natural Language Processing and Knowledge Engineering* (pp. 1-8). IEEE.

Cahyani, N. L. D. (2022). Derivational affixes found in the caption of selected posts of @bawabali_official account on Instagram (Doctoral dissertation, Universitas Mahasaraswati Denpasar).

Castilho, S., Moorkens, J., Gaspari, F., Calixto, I., Tinsley, J., & Way, A. (2017a). Is neural machine translation the new state of the art? *The Prague Bulletin of Mathematical Linguistics*, *108*, 109–120.

Castilho, S., Moorkens, J., Gaspari, F., Sennrich, R., Sosoni, V., Georgakopoulou, P., ... & Gialama, M. (2017b). A comparative quality evaluation of PBSMT and NMT using professional translators. In *Proceedings of Machine Translation Summit XVI: Research Track* (pp. 116-131).

Chatzikoumi, E. (2020). How to evaluate machine translation: A review of automated and human metrics. *Natural Language Engineering*, *26*(2),137-161.

Chéragui, M. A. (2012). Theoretical overview of machine translation. *ICWIT*, 160-169.

Das, A. K. (2018). Translation and artificial intelligence: Where are we heading. *International Journal of Translation*, *30*(1),72-101.

Cho, K., Van Merriënboer, B., Bahdanau, D., & Bengio, Y. (2014). On the properties of neural machine translation: Encoder-decoder approaches. In *Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation (SSST-8). arXiv preprint arXiv:1409.1259*.

Dorr B., Snover M. & Madnani N. (2011). Chapter 5.1 introduction. In Olive J., McCary J. and Christianson C. (eds), *Handbook of Natural Language Processing and Machine Translation. DARPA Global Autonomous Language Exploitation*. New York: Springer, (pp. 801–803).

España-Bonet, C., & Costa-jussà, M. R. (2016). Hybrid machine translation overview. *Hybrid Approaches to Machine Translation*, 1-24.

Fadilah, E. (2017). Semantic Errors Analysis of Instagram Machine Translation from Indonesian to English. Published thesis, Syarif Hidayatullah State Islamic University of Jakarta, Indonesia.

Habash, N., Dorr, B., & Monz, C. (2009). Symbolic-to-statistical hybridization: extending generation-heavy machine translation. *Machine Translation*, *23*, 23-63.

Hunsicker, S., Chen, Y., & Federmann, C. (2012, June). Machine learning for hybrid machine translation. In *Proceedings of the Seventh Workshop on Statistical Machine Translation*, 312-316.

Hutchins, W. J. (1995). Machine translation: A brief history. In *Concise History of the Language Sciences* (pp. 431-445). Pergamon.

Hutchins, J. (2007). Machine translation: A concise history. *Computer Aided Translation: Theory and Practice*, *13*(29-70), 11.

Koby G.S., Fields P., Hague D., Lommel A. & Melby A. (2014). Defining translation quality. *Tradumàtica* 12,413–420.

Koehn, P. (2009). *Statistical machine translation*. Cambridge University Press.

Larassati, A., Setyaningsih, N., Nugroho, R. A., Suryaningtyas, V. W., Cahyono, S. P., & Pamelasari, S. D. (2019). Google vs. Instagram machine translation: multilingual application program interface errors in translating procedure text genre. In *2019 International Seminar on Application for Technology of Information and Communication (iSemantic)* (pp. 554-558). IEEE.

Lommel, A., & Melby, A. K. (2018). Tutorial: MQM-DQF: A good marriage (Translation quality for the 21st Century). In *Proceedings of the 13th Conference of the Association for Machine Translation in the Americas (Volume 2: User Track)*.

Mawarni, B., Pambudi, B. D., & Ghasani, B. I. (2017). The problem of cultural untranslatability found in the English translation of Jokowi's Instagram posts. In *UNNES International Conference on ELTLT* (pp. 104-108).

Meilasari, P. (2019). When Instagram translation machine translates ecology terms: Accurate or not? In *The 7th Library Studied Conference* (p. 129).

Melby, A. K. (2002). The mentions of equivalence in translation. *Meta*, *35*(1),207-213.

Meta AI. (2022). *200 languages within a single AI model: A breakthrough in high-quality machine translation*. Meta AI. Retrieved March 22, 2023, from https://ai.facebook.com/blog/nllb-200-high-quality-machine-translation/

Mannes, J. (2017, August 3). *Facebook finishes its move to neural machine Translationt*. TechCrunch. Retrieved March 22, 2023, from https://techcrunch.com/2017/08/03/facebook-finishes-its-move-to-neural-machine-translation/?guccounter=1

Moorkens, J., Toral, A., Castilho, S., & Way, A. (2018). Translators' perceptions of literary post-editing using statistical and neural machine translation. *Translation Spaces*, *7*(2),240-262.

Nord, C. (2005). *Text analysis in translation: Theory, methodology, and didactic application of a model for translation-oriented text analysis*. New York: Rodopi.

Nord, C. (1997). Functionalist approaches explained. *Manchester, UK: St. Jerome Publishing*.

Omar, N., & Al-Tashi, Q. (2018). Arabic nested noun compound extraction based on linguistic features and statistical measures. *GEMA Online® Journal of Language Studies*, *18*(2).

Omar, A., & Gomaa, Y. (2020). The machine translation of literature: Implications for translation pedagogy. *International Journal of Emerging Technologies in Learning (IJET)*, *15*(11),228-235.

Pino, J. M., Sidorov, A., & Ayan, N. F. (2017). *Transitioning to neural machine translation*. Tech at Meta. Retrieved March 22, 2023, from https://tech.facebook.com/artificial-intelligence/2017/8/transitioning-entirely-to-neural-machine-translation

Pujakesuma, G. A. (2022). *The Performance of Instagram's auto-translate and google translate in translating house of highlight's Instagram captions* [Thesis, Yogyakarta:Sanata Dharma University]. http://repository.usd.ac.id/id/eprint/42510

Purwaningsih, D. R., Sholikhah, I. M., & Wardani, E. (2019). Revealing translation techniques applied in the translation of Batik Motif names in see Instagram. *Celt: A Journal of Culture, English Language Teaching & Literature*, *19*(2), 287-301.

Ragni, V., & Nunes Vieira, L. (2022). What has changed with neural machine translation? A critical review of human factors. *Perspectives*, *30*(1),137-158.

Sabtan, Y. M. N., Hussein, M. S. M., Ethelb, H., & Omar, A. (2021). An evaluation of the accuracy of the machine translation systems of social media language. *International Journal of Advanced Computer Science and Applications*, *12*(7), 406-415. DOI: 10.14569/IJACSA.2021.0120746

S. Dixon. (2023). *Instagram users worldwide 2025*. Statista. Retrieved March 22, 2023, from https://www.statista.com/statistics/183585/instagram-number-of-global-users/

Sipayung, K. T., Sianturi, N. M., Arta, I. M. D., Rohayati, Y., & Indah, D. (2021). Comparison of Translation Techniques by Google Translate and U-Dictionary: How Differently Does Both Machine Translation Tools Perform in Translating? *Elsya: Journal of English Language Studies*, *3*(3),236-245.

Susanti, E. (2018). *Lexical errors produced by Instagram machine translation* [Doctoral dissertation, Universitas Islam Negeri Maulana Malik Ibrahim].

Stymne, S., & Ahrenberg, L. (2012, May). On the practice of error analysis for machine translation evaluation. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)* (pp. 1785-1790).

Tambouratzis, G. (2014, April). Comparing CRF and template-matching in phrasing tasks within a Hybrid MT system. In *Proceedings of the 3rd Workshop on Hybrid Approaches to Machine Translation (HyTra)* (pp. 7-14).

Thurmair, G. (2009). Comparing different architectures of hybrid machine translation systems. In *Proceedings of Machine Translation Summit XII: Posters*.

Widiastuti, N. M. A. (2021). Translation procedures of English phrasal verbs into Indonesian on Instagram captions. In *International Seminar on Austronesian Languages and Literature IX (399-408). Denpasar: Udayana University*.

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. *arXiv preprint arXiv:1609.08144*.

ABOUT THE AUTHORS

Altaf A. Fakih is a Ph.D. candidate at Universiti Sains Malaysia (USM) where she also obtained her M.A. in Translation and Interpreting studies. Her research areas of interest include Machine Translation, Artificial Intelligence, Audiovisual Translation, and Translation Quality Assessment. She is also a certified English-Arabic translator.

Mozhgan Ghassemiazghandi is a Senior Lecturer at the School of Languages, Literacies, and Translation at Universiti Sains Malaysia. She holds a Ph.D. in Translation Studies. Her research interests are in translation technology, machine translation, and audiovisual translation. Additionally, Mozhgan is an experienced freelance translator and subtitler, with more than a decade of experience in the field.

Abdul-Hafeed Fakih is a Professor of Linguistics at the Department of English, Najran University, Saudi Arabia (and formerly at Ibb University, Yemen). He published several papers in indexed journals. He is a member of different editorial boards of indexed journals.

Manjet K. M. Singh is an Associate Professor at Universiti Sains Malaysia. She holds a Ph.D. in language studies and has been attached to School of Languages, Literacies and Translation for the past 27 years. Manjet's interests are broad ranging and include sociolinguistics, language teaching and learning, academic literacy(ies), discourse, and multilingualism.