

## **An Assessment of English-Arabic Translation Using ChatGPT**

**Ahmad Al-Harahsheh, Suha Tayseer Abualkarasanneh,\*, Rafat Al Rousan**

Translation Department Yarmouk University, Jordan

**Received on: 23-7-2024**

**Accepted on: 17-12-2024**

### **Abstract**

This study aims to assess the quality of translating political texts from English to Arabic using the artificial intelligence tool ChatGPT. It adopts the updated version of the Multidimensional Quality Metrics (MQM) framework to evaluate the quality of translations produced by ChatGPT. The data includes 30 political articles selected from BBC World and CNN World in English and their translations into Arabic. The study identified several linguistic errors encountered by translators using ChatGPT, such as fluency and accuracy. Translations lacking these important aspects often lead to misunderstandings and confuse the users. Despite these issues, the ChatGPT tool demonstrated good translation performance in political news, with an overall translation quality rating of 93%.

**Keywords:** ChatGPT, Artificial Intelligence, Political Texts, English, Arabic.

### **1. Introduction**

Artificial Intelligence (AI) is “the science and engineering of making intelligent machines, especially intelligent computer programs” (McCarthy 2007, 2). It concerns “the computer's development to employ human thought processes like self-correction, adapting and learning” (Kok et al. 2009, 2). Therefore, AI is a field of computer science where machine stimulates human intelligence to perform certain tasks, which can be developed based on the orders provided to the machine by humans. It has been used widely in different fields of knowledge, such as manufacturing, teaching, and translation. On the other hand, Machine Translation (MT) is the process of producing automatic translation without human assistance. It excludes “computer-based translation tools which support translators by providing access to online dictionaries, remote terminology databanks, transmission and reception of texts, etc.” (Hutchins1995, 431). There are several MT programs, such as Google Translate, Bing Microsoft Translator, DeepL, Reverso Translation, and ChatGPT.

According to Guinness (2023), ChatGPT predicts words by understanding a prompt, so it can generate the appropriate answer for a question. By using natural language, ChatGPT learns internet data, providing users with AI-generated written responses or answer prompts. By training these models on large datasets, AI enables them to predict the next word in a sentence (Sabzalieva et al. 2023).

---

© 2025 JJMLL Publishers/Yarmouk University. All Rights Reserved,

\* Doi: <https://doi.org/10.47012/jjml.17.4.8>

\* Corresponding Author: [harahsheh77@yu.edu.jo](mailto:harahsheh77@yu.edu.jo)

ChatGPT is based on the (GPT) Generative Pre-trained Transformer architecture; it has been trained on large datasets and generates responses like humans when provided by prompts (Zheng et al. 2021; Pokorny 2023). ChatGPT encodes the user prompts apart from old conversations in history using transformer blocks; then to generate a response, it uses these encoded data or information using the sequence-to-sequence framework modeling (Verma, 2022). Sequence-to-sequence, a model proposed by Srivastava et al. (2014), maps one sequence to another using a neural network. It was applied for the first time on machine translation replacing the whole-sentence machine translation model (Zhao 2023).

ChatGPT can be used as a translator; it can translate any language with output fluency and accuracy, and it can handle complex languages. The user employs ChatGPT with selected texts in a specific language, which is explicitly prompted to translate it into another language. Then, ChatGPT generates the translation immediately (Khoshafah 2023). In recent years, the use of Artificial Intelligence (AI) in translation has increased. ChatGPT is not a type of MT, although ChatGpt and MT have some shared features but their functions are different. ChatGPT can generate interactions, write human-like texts based on certain prompts, and translate texts of various fields of knowledge. It can also comprehend the text's context or the colloquial or slang terms in a text. However, MT generates automatic translation of a certain text directly without taking care of the context. Despite advancements in machine translation technologies, there is still a gap in investigating the accuracy and fluency of English-Arabic translations, especially when it comes to political news.

Political language has subtle expressions whose meanings depend on the situation, and its language has lots of details, making it tricky. The primary purpose of this study is to assess the accuracy and fluency of ChatGPT English-Arabic translation of political news. By selecting the political field, which is complex, the study aims to assess the ability of ChatGPT to maintain clear and coherent content during the translation process. The study seeks to learn more about ChatGPT, its performance, and challenges in rendering political news. This study is keen to answer the following questions:

1. How accurate is ChatGPT in translating political news from English into Arabic?
2. How fluent is ChatGPT in translating political news from English into Arabic?

## **2. Literature Review**

Gao et al. (2023) use the multi-reference test sets that allow a variety of attainable translations. They indicate that it is important to use more comprehensive evaluation criteria to assess the quality of ChatGPT's translation.

The study concluded that providing ChatGPT with complex or unpleasant information will generate serious devaluing performance; however, correct information about prompt text led to allowing free rein in the performance of ChatGPT. In addition, they employ ChatGPT and a pre-edit scheme to achieve accuracy in translating attributive clauses from Japanese to Chinese using ChatGPT. They indicated that the accuracy of ChatGPT should be improved by using linguistic knowledge by linking the selection of translation modes in Japanese attributive clauses and the semantic functions of changed nouns. Upon using this linguistic knowledge, they built tailored prompts to facilitate MT and then tested ChatGPT to

validate the viability of their strategy. Their strategy improved the accuracy of the translation score by 30% and optimized the translation input.

In the same vein, Işım and Balcıoğlu (2023) assess the performance of the translation of AI for English – Turkish. They provided an assessment of ChatGPT in the translation performance and translation prompt of machine translation. The study selected 50 Turkish paragraphs from the educational field. They found that the performance of ChatGPT is competitive with education, and for the output, out of the 50 paragraphs, 34 were correctly translated without errors, and the rest detected errors. While ChatGPT can render simple texts appropriately, the translation should be controlled as there are some translation errors.

Siu et al. (2023) discuss the strengths, use cases, and limitations of both ChatGPT and GPT-4 and their impact on the assessment of translation context. Their study aims to promote an extensive knowledge of ChatGPT and GPT-4 roles in the translation industry and training discussions regarding their potential and implications. They presented cases that demonstrate how ChatGPT can be used to assist translation tasks, such as contextual clarification of expressions, cultural explanation of expressions, explanation of technical terminology, simplification of complex text, and more. Then, they continued with the strength of ChatGPT in the translation context. In this section, Siu et al. (2023) explore the notable benefits of ChatGPT, such as emphasizing cost, creativity, speed, and context awareness. Furthermore, the limitations and constraints of ChatGPT, GPT-4 have major imperfections, including the architecture of the prediction of the next word, limited knowledge, incorrect predictions, unreliability, limited Chinese language support besides the translation quality (Chinese), and more.

Wang (2023) aims to demonstrate how a new AI or Chatbot can be used to construct bilingual knowledge as well as to compare and contrast AI-generated and human-translated texts. Powered by ChatGPT, Wang (2023) presents the created bilingual dataset about the Canton Fair and then compares the human-translated and ChatGPT-generated multimodal texts involving invitation letters and visual pictures. The study discovered that in contrast to bottom-up translation practice, instruction, and study. AI can aid the translation from a macroscopic and three-dimensional knowledge framework to language usage at a microscopic level (top-down way). Furthermore, comparing AI- and human-produced translations can provide empirical concepts and perspectives for translating, writing, and post-editing documents in the new AI era.

Cheng et al. (2023) discover prospects for AI-driven translation technology using ChatGPT. They selected ten Chinese text samples based on the accuracy of concepts in the fields of economy, politics, and arts. Then, they used translation tools such as ChatGPT, Baidu Translate, and Google Translate to translate the selected samples into Chinese. The study showed that ChatGPT translation is more accurate and relevant for contextual understanding, while MT tools may induce ambiguous understandings. Furthermore, Cheng et al. (2023) indicated that due to globalization, AI translation tools, besides translating accurately, must have a profound understanding of cultural contexts, which leads to minimizing cultural disparities and easing cultural tensions.

Wu et al. (2023) analyze the translation quality of ChatGPT for the language pairs Chinese-English in terms of the syntactic arrangement, consistency, and domain modification by comparing the traditional

MT and the official one. The study selected some Chinese words mentioned in the Chinese Government Work Report, including idioms, abbreviations, political jargon, and terms with a Chinese background. They found that ChatGPT illustrates considerable advantages in translating professional terminology or political jargon compared to Google Translate. In addition, when translating some Chinese idiomatic expressions, the translations generated by ChatGPT are equivalent to translations provided by experts or professionals. Furthermore, ChatGPT encounters difficulties in conveying implicit meanings regarding terms or words with specialized knowledge.

Farrell et al. (2023) have done two experiments. The first aimed to measure the translation quality of a single language combination (English to Italian) and a specific text genre by using the accessible version of DeepL, ChatGPT, and Google Translate through human assessment. The result was that DeepL had the best performance (279 points), followed by Google Translate (239 points), and finally ChatGPT (186 points). In the second experiment, the specific translation of "*bigram ci sono*" was chosen, meaning they are in Italian and made by ChatGPT for the MT marker. It has been analyzed, and its variations and frequency were assessed against human translation. ChatGPT (free version) equaled the average HT lexical variety levels in automatic monolingual machine translation post-editing (MTPE). ChatGPT failed to deliver a lexical variety after producing *bigram ci soon* four times out of four during the raw output generation. In addition, in subsequent experiments, ChatGPT surpassed the average HT lexical variety levels in automatic bilingual MTPE, demonstrating a higher level of variety and diversity than HT. Farrell (2023) added that even though ChatGPT demonstrated a higher level of variety and diversity, it also has a limitation, including the assessment focused only on the translation of "*bigram ci sono*," and a deeper examination of various translation features is required to comprehend ChatGPT's performance in contrast of HT.

In a recent work, Khoshafah (2023) evaluates the translation accuracy of various text genres of Arabic - English using ChatGPT 3.5. The study concluded that ChatGPT could be used as a translator for simple kinds of text; however, it endures with challenges that require intervention from a human, such as lacking proper phrasing (Arabic) even that is grammatically correct, the issue of word order, the omission of Arabic nouns but it cannot detect the Arabic diacritical marks which could be a struggle for AI. Moreover, Khoshafah (2023) added that translating cultural differences could be challenging. Although ChatGPT generally offers accurate translations, its constraints make it inappropriate for some texts, such as legal documents, research papers, medical reports, and works of literature.

Moneus et al. (2023) investigate the accuracy of legal translation provided by humans and AI translation and examine their strengths and weaknesses besides the practical approach. It has selected legal texts from different contracts and translated them by legal translators and AI translation tools. Computer programs, such as (Google Translate, Bing, Microsoft Translator, DeepL, Reverso, ChatGPT-4, etc.) can be faster and more effective than human translations. Still, they may need to consistently capture the intricacies and cultural nuances of the text as precisely as a human translator does. On the other hand, human translators convey the exact meaning and provide a deeper understanding of the cultural context of the text being translated, making them the best choice for translating text with high accuracy and cultural awareness. Consequently, the results show that in Arabic translations the human

translation was recorded (92.2) and (92.7) in English translation in contrast to the AI translation was recorded in Arabic translation (88.2) and (89.1) in English translation, which means that translation by humans is superior because of the legal background with keeping the legal effect which is not included in AI translations, which means if AI translation is fed and include with a legal dataset (legal text, etc.), this may lead to the replacement of human translation services.

Al-Mubassyir (2023) examines the accuracy of translating Arabic into Indonesian texts regarding the context using ChatGPT (GPT-3) and Google Translate. Following analysis, it is clear that Arabic translation using ChatGPT (GPT-3) and Google Translate can deliver a general comprehension of the topic efficiently and timely. However, ChatGPT surpasses Google Translate in terms of accuracy, since some translated words were found to be inconsistent in their usage context. The study indicates that MT can translate Arabic texts into Indonesian quickly and effectively, but translators are still required to improve the quality of the final translation. Translators can adjust and examine the translation to fulfill the requirements, notably in terms of word usage context.

Kadaoui et al. (2023) present an assessment of ChatGPT (GPT3 and GPT4) and Bard across ten varieties of Arabic. It covers Modern Standard Arabic, Classical Arabic, and several country-level dialectal variants. Besides, a study revolves around humans to examine the effectiveness of the recent model (Bard) in following human instructions during translation tasks. Manually, it has created a multi-dialect Arabic dataset for MT evaluation to assess the ability of LLMs to understand non-visual data. It was found that LLMs face particular challenges, especially in the rare categories present at the end of the most uncommon general data. However, it can perform well in some categories that have been considered. Furthermore, they do not include different varieties of various language types purported to exhibit good performance, like Arabic.

### **3. Methodology**

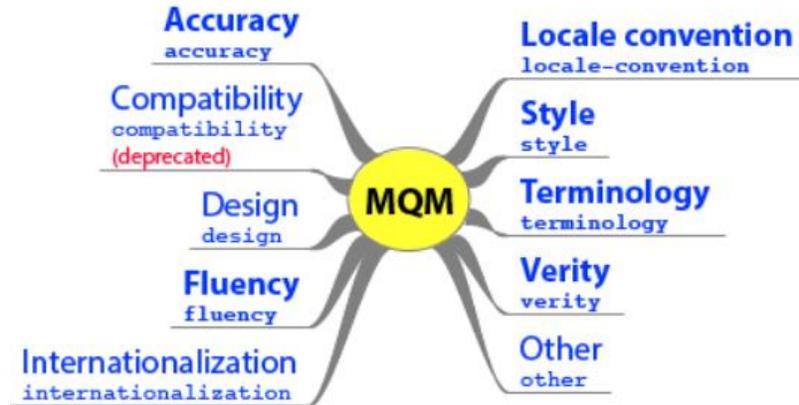
#### *3.1 Data Collection*

The sample of this study consists of thirty political news articles selected randomly from the major international news organizations: BBC (World) and CNN (World) for the period 7 October to 30 November 2023. The news was related to the Hamas-Israeli conflict. In this study, ChatGPT was used as the translation tool, and the system was explicitly prompted to translate them using the sentence “*Translate into Arabic.*” The study randomly selects this news for assessment to analyze the issues related to the fluency and accuracy of ChatGPT in rendering political news. The study mainly focuses on these two issues as they are primarily important in determining whether the translation is adequate or inadequate for the target readers.

#### *3.2 Data analysis*

ChatGPT 3.5 (free version) was asked to translate each news separately from English into Arabic. After that, the accuracy and fluency of the translated news were analyzed and assessed linguistically using the MQM (Multidimensional Quality Metric). A holistic approach was used to detect the translation issues or errors. MQM framework is describes and defines quality utilized in assessing

translated texts and identifying their particular issues (WayBackMachine 2016) (See figure 1). The MQM framework provides an entire set of criteria for evaluating the major, minor, and critical issues of the media content (WayBackMachine 2016). The MQM presents a comprehensive framework to assess the reliability and validity of translation. It assesses crucial criteria, such as accuracy, fluency, terminology, design, and style. Therefore, it offers a comprehensive understanding of the translated text.



**Figure 1:** A tree graph of the MQM issues

The default MQM severity model has three levels:

1. Minor issue (1 point) does not affect the understandability of the content (e.g., a space after a full stop).
2. Major issue (10 points) affects the understandability of the content but it does not make it impossible for the readers to understand (e.g., misspelling of a word).
3. Critical issue (100 points) makes the content inappropriate for use (e.g., errors prevent users to use the content properly, incorrect information, and grammatical errors that completely change the meaning).

The following formula shows how to calculate the penalties (P) (WayBackMachine, 2016)

$$P = \frac{\left( \text{Issues}_{\text{minor}} + \text{Issues}_{\text{major}} \times \text{SeverityMultiplier}_{\text{major}} + \text{Issues}_{\text{critical}} \times \text{SeverityMultiplier}_{\text{critical}} \right)}{\text{Word count}}$$

**Figure 2:** (Severity formula)

The following formula shows how to generate the Quality Score (TQ) after calculating the penalties to shows the overall rating of quality:

$$\text{Quality Score (TQ)} = 100 - \text{Penalties for the target content (TP)} + \text{Penalties for the source content (SP)}$$

$$\text{TQ} = 100 - \text{TP} + \text{SP}$$

#### 4. Findings and Discussion

After the texts were translated into Arabic by ChatGPT, they were analyzed and discussed based on Multidimensional Quality Metrics (MQM). The following categories of Multidimensional Quality

## An Assessment of English-Arabic Translation Using ChatGPT

Metrics (MQM) were applied to the data analysis: Verity, style, terminology, design, fluency, and accuracy.

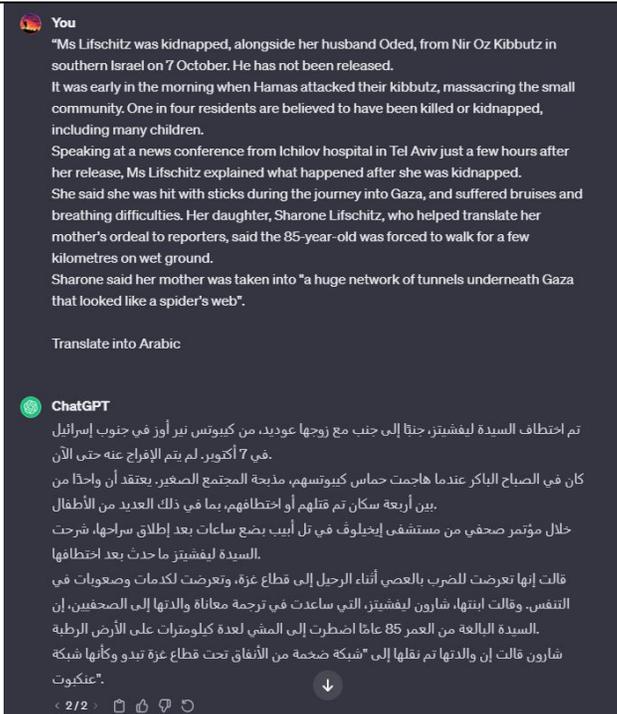
### Example (1)

ChatGPT Translation	Calculated Penalties & Quality Scores	
<div style="background-color: #f0f0f0; padding: 5px; border: 1px solid #ccc;"> <p><b>You</b></p> <p>"DJ Rosenthal, who served as director for counterterrorism at the US National Security Council during the Obama administration, wrote that "Israel must ensure that its pursuit of its objectives against Hamas is conducted with the utmost care to minimize civilian casualties. While it is true that Hamas uses civilians as human shields to create a significant tactical complication for Israel's mission, Hamas' inhumanity does not form a basis upon which Israel can depart from its obligations to minimize civilian casualties. To fail to do so is to risk undermining Israel's security interests in the region, Western support and its legitimacy."</p> <p>"Take, for example, the Israel Defense Forces strike on Gaza's Jabalya refugee camp on Tuesday. While the IDF said that it was targeting high-level Hamas leadership, which might make the strikes legal under international law, Israel must not ignore the realpolitik implications."</p> <p>Translate into Arabic</p> </div> <div style="background-color: #f0f0f0; padding: 5px; border: 1px solid #ccc; margin-top: 5px;"> <p><b>ChatGPT</b></p> <p>الذي شغل منصب مدير مكتب مكافحة الإرهاب في مجلس الأمانة القومي، الأمريكي خلال إدارة أوباما، كتب: "يجب على إسرائيل أن تضمن أن سعيها لتحقيق أهدافها ضد حماس يتم بأقصى درجات العذر لتقليل الخسائر المدنية. بينما صحيح أن حماس تستخدم المدنيين كدروع بشرية لإعاقة مهمة إسرائيل بشكل تكتيكي، إلا أن وحشية حماس لا تشكل أساساً يمكن لإسرائيل أن تتخذ على أساسه تحليلاً عن التزاماتها لتقليل الخسائر المدنية. فعل ذلك يعرض إسرائيل لخطر تقويض مصالحتها الأمنية في المنطقة والدعم العربي وشرعيتها"</p> <p>لنأخذ، على سبيل المثال، ضربات القوات الدفاع الإسرائيلية على مخيم جنابا للاجئين في غزة يوم الثلاثاء. بينما قالت القوات الدفاع الإسرائيلية إنها تستهدف قيادة حماس على مستوى عالٍ، مما قد يجعل الضربات قانونية وفقاً للقانون الدولي، يجب على إسرائيل ألا تتجاهل آثار الواقعية السياسية</p> </div> <p>Source: (CNN, 5/Nov/2023)</p>	<p>Total word of text= 125  (Penalties= issues minor × Severity Multiplier minor + issues major × Severity Multiplier major) / word count)  <math>P = (5*1+1*10*0)/125</math>  <math>= 15/125</math>  <math>= 0.12</math>  (100 – Penalties for the target content + Penalties for the source content = Quality Score)  <math>100 - (0.12* 100) + 0 =</math>  <math>100 - 12 + 0 = 88</math>  Quality Score of the text is 88%</p>	
<b>ST/TT</b>	<b>Errors that have been detected</b>	<b>Dimensions/ Penalties</b>
ST (a)	... DJ Rosenthal, who served as director for counterterrorism at the US National Security Council during the Obama administration ...	<b>Major:</b> Terminology P= 0.09
TT	الذي شغل منصب مدير مكتب مكافحة الإرهاب في مجلس الأمانة القومي، الأمريكي خلال إدارة أوباما، كتب: " يجب على إسرائيل أن ... Rosenthal	
ST (b)	...Take, for example, the Israel Defense Forces strike on Gaza's Jabalya refugee camp on Tuesday...	<b>Minor:</b> <i>Terminology:</i> (domain-terminology) P= 0.06
TT	لنأخذ، على سبيل المثال، <u>ضربات</u> القوات الدفاع الاسرائيلية على مخيم جنابا للاجئين في غزة يوم الثلاثاء .	
ST (c)	The whole text	Minor: Design (Text alignment) P= 0.008
TT	كامل النص"	

In example (1 a), ChatGPT did not transliterate the name "DJ Rosenthal"; it kept it as it was. In addition, it renders the "US National Security Council" as "مجلس الأمانة القومي الأمريكي" in Arabic, which is wrong. ChatGPT provides inadequate translation for the term "Security" as "الأمانة" literally means "honesty" in Arabic; it should be translated as "الأمن" "security" to provide the correct meaning. In this sentence, terminology has a penalty rate of (0.09). In (1B), the phrase "the Israeli Defense Forces strike" is rendered inadequately in Arabic as ضربات القوات الدفاع الاسرائيلية. It should be rendered as قصف قوات

الدفاع الاسرائيلية and there is no need for the definite article "ال(the). This minor issue may have a penalty of (0.06). Finally, it is noted that ChatGPT lacks design dimension with penalties of (0.008) during the translation process, like overall layout and text alignment, which affected the process of reading the text. The translation generated by ChatGPT quality was nearly 88% in quality.

**Example (2)**

ChatGPT Translation		Calculated Penalties & Quality Scores
 <p>Source: (BBC, 24/Oct/2023)</p>		<p>Total word of text= 135                      (Penalties= issues minor × Severity Multiplier minor + issues major × Severity Multiplier major) / word count)  <math>P = (7*1+0*0)/135</math>  <math>= 7/135</math>  <math>= 0.05</math>  <math>100 - (0.05 * 100 )+ 0 = \text{Quality Score}</math>  <math>100 - 5+ 0 = 95</math></p> <p>Quality Score of the text is <u>95%</u></p>
ST/TT	Issues that have been detected	IssueType/ Penalties
ST (A)	... It was early in the morning when Hamas attacked their kibbutz, massacring the small community...	Minor: Terminology (Culture-specific reference) P= 0.09
TT	كان في الصباح الباكر عندما هاجمت حماس كيبوتسهم، مذبحه المجتمع الصغير	
ST (B)	... Speaking at a news conference from Ichilov hospital in Tel Aviv just a few hours after her release, Ms. Lifschitz explained what happened after she was kidnapped ...	Major: Accuracy, Fluency P= 0.18
TT	وخلال مؤتمر صحفي من مستشفى إيلوف في تل أبيب بضع ساعات بعد إطلاق سراحها شرحت السيدة ليفشيتس ما حدث بعد اختطافها.	
ST (C)	... Sharone said her mother was taken into "a huge network of tunnels underneath Gaza that looked like a spider's web..."	Minor: Fluency (word order) P= 0.05
TT	شارون قالت إن والدتها تم نقلها إلى "شبكة ضخمة من الأنفاق تحت قطاع غزة تشبه شبكة العنكبوت	
ST (D)	The whole text	Minor: Design (Text alignment) P= 0.007
TT	كامل النص"	

## An Assessment of English-Arabic Translation Using ChatGPT

In (2A), In the sentence “*It was early in the morning when Hamas attacked their kibbutz, massacring the small community.*” The term “kibbutz” is an Israeli culture-specific term that means “a type of farm or factory where a group of people live together and share all the work, decisions, and income” (Oxford Dictionary Online). The appropriate translation for this term in Arabic is "مستوطنة". In (2B), the following sentence has a major issue in translation:

وخلال مؤتمر صحفي من مستشفى إيكيلوف في تل أبيب بضع ساعات بعد إطلاق سراحها شرت السيدة ليفشيتس ما حدث بعد اختطافها.

This sentence has accuracy and fluency issues as ChatGPT mistranslates it. ChatGPT provides an inappropriate translation for this sentence, and the translation lacks coherence to the text. The reason for that is word-for-word translation. It should be rendered as: بعد ساعات قليلة من خروجها من مستشفى إيكيلوف أوضحت السيدة ليفشيتس في مؤتمر صحفي ما حدث لها بعد اختطافها. These issues have penalties of (0.18). In (2 c), there is a minor issue of word order, which relates to fluency with a penalty of (0.05). In the Arabic translation, the verb شارون comes after the subject قالت, which is unacceptable in Arabic as the verb must proceed with the subject. In (2 d), there is a minor design issue with a penalty of (0.007). The translation quality achieved was impressively high, reaching 95%.

### Example(3)

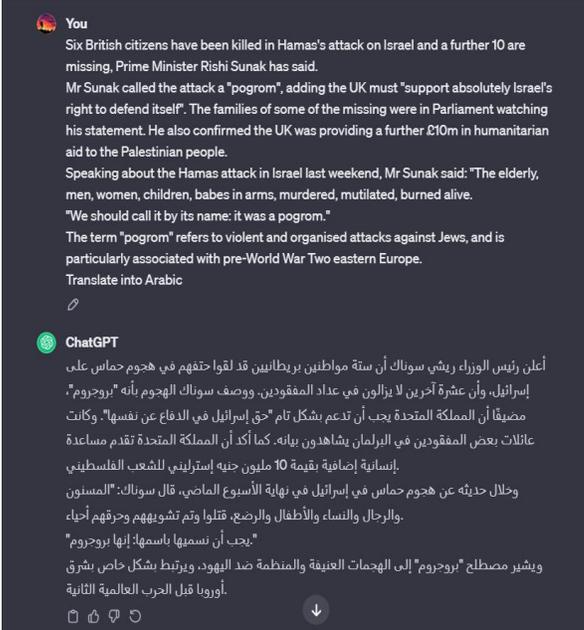
ChatGPT Translation		Calculated Penalties & Quality Scores
<p><b>You</b></p> <p>Three of Wales' police forces said they had not seen a notable increase in reported hate crime, but Supt Tim Morgan of South Wales Police said the service was "actively reaching out" to offer support.</p> <p>The Muslim Council of Wales' secretary general Dr Abdul-Azim Ahmed said it was reassuring that Welsh police had not seen a rise in hate crime, but many Muslims were "hurting" and "afraid."</p> <p>"People are concerned about the impact of campaigning for Palestine. They're concerned as well about the reception and sentiment others hold.</p> <p>"It's such a polarised time, especially some of the rhetoric coming from authorities."</p> <p>The Welsh government said: "The first minister has strongly condemned the appalling attacks carried out by Hamas and the horrific escalation in violence against Israel."</p> <p>Translate into Arabic</p> <p><b>ChatGPT</b></p> <p>قالت ثلاثة من قوات الشرطة في ويلز أنها لم تلاحظ زيادة ملحوظة في الجرائم التي تم الإبلاغ عنها كجرائم الكراهية، ولكن قال المفتش تيم مورغان من شرطة جنوب ويلز إن الخدمة "تسعى نشطاً" لتقديم الدعم. وقال الأمين العام لمجلس المسلمين في ويلز الدكتور عبد العظيم أحمد إنه من المطمئن أن الشرطة الويلزية لم تشهد زيادة في جرائم الكراهية، ولكن العديد من المسلمين "يعانون" و"يشعرون بالخوف".</p> <p>الناس قلقون بشأن تأثير الدعوة من أجل فلسطين، إنهم يشعرون أيضاً بالقلق بشأن الترحيب، والمشاعر التي يحملها الآخرون.</p> <p>"إنه وقت متقطع، خاصة بعض الخطابات الواردة من السلطات"</p> <p>وقالت حكومة ويلز: "قد أدان رئيس الوزراء بشدة الهجمات الرهيبة التي نفذتها حماس والتصعيد المرع في العنف ضد إسرائيل"</p>		<p>Total word of text= 111</p> <p>(Penalties= issues minor × Severity Multiplier minor + issues major × Severity Multiplier major) / word count)</p> $P = (4*1+0*0)/111$ $= 4/111$ $= 0.03$ <p>100 – (0.03 * 100 )+ 0 = Quality Score</p> <p>100 – 3 + 0 = 97</p> <p>Quality Score of the text is <u>97%</u></p>
Source: (BBC, 22/Oct/2023)		
ST/T	Issues that have been detected	Issue Type/ Penalties
TS (a)	Three of Wales' police forces said they had not seen a notable increase in reported hate crime, but Supt Tim Morgan of South Wales Police said the service was "actively reaching out" to offer support	Minor: Accuracy (mistranslation) P= 0.02

TT	قالت ثلاثة من قوات الشرطة في ويلز انها لم تلاحظ زيادة ملحوظة في الجرائم التي تم الإبلاغ عنها كجرائم الكراهية، ولكن قال المفتش تيم مورغان من شرطة جنوب ويلز إن الخدمة "تسعى نشطاً" لتقديم الدعم	
ST (b)	...People are concerned about the impact of campaigning for Palestine. They're concerned as well about the reception and sentiment others hold...	Minor:Terminology (term-inconsistency) P= 0.05
TT	"الناس قلقون بشأن تأثير الدعوة من أجل فلسطين. إنهم يشعرون أيضا بالقلق بشأن الترحيب والمشاعر التي يحملها الآخرون.	
ST (c)	"It's such a polarised time, especially some of the rhetoric coming from authorities."	Minor:Terminology (Inconsistent use of terminology) P= 0.11
TT	"إنه وقت متقطع، خاصة بعض الخطابات الواردة من السلطات".	
ST	The whole text	Minor: Design (Text alignment) P=0.009
TT	كامل النص"	

In (3 a), there is a major issue of accuracy as the phrase “actively reaching out” is inadequately translated as "تسعى نشطاً" as it should "dimension with penalties of (0.02) identified constitute to the subcategory of "Mistranslation” sentence “تسعى نشطاً” could be translated more accurately as “تسعى” to capture the meaning of “actively reaching out”.

In the provided context, the term “الدعوة” has terminology dimension “minor”, and it falls under the category of inconsistent use with penalties of (0.05). Regarding the term “campaigning”, it could be translated as “حملة”, which refers to an invitation to specific campaigning. Besides, the term “الترحيب” may not capture the intended meaning; in this context “reception” refer to the reception of the campaign more than welcoming. Also, the use of "متقطع" to translate "polarised" in this context can be categorized as a minor issue for the Terminology dimension of inconsistent use with penalties of (0.11). The term “polarised” could be suitably translated into “متقسم” in this context to convey the meaning where public opinions are divided between conflicting views. It is essential to mention that the text has the same Design dimension issue with penalties of (0.009). To conclude, the translation approached a high level of quality with 97%.

**Example (4)**

ChatGPT Translation		Calculated Penalties & Quality Scores
 <p><b>You</b> Six British citizens have been killed in Hamas's attack on Israel and a further 10 are missing, Prime Minister Rishi Sunak has said. Mr Sunak called the attack a "pogrom", adding the UK must "support absolutely Israel's right to defend itself". The families of some of the missing were in Parliament watching his statement. He also confirmed the UK was providing a further £10m in humanitarian aid to the Palestinian people. Speaking about the Hamas attack in Israel last weekend, Mr Sunak said: "The elderly, men, women, children, babes in arms, murdered, mutilated, burned alive. "We should call it by its name: it was a pogrom." The term "pogrom" refers to violent and organised attacks against Jews, and is particularly associated with pre-World War Two eastern Europe. Translate into Arabic</p> <p><b>ChatGPT</b> أعلن رئيس الوزراء ريشي سوناك أن ستة مواطنين بريطانيين قد لقوا حتفهم في هجوم حماس على إسرائيل، وأن عشرة آخرين لا يزالون في عداد المفقودين. ووصف سوناك الهجوم بأنه "بروجروم"، مضيفاً أن المملكة المتحدة يجب أن تدعم بشكل تام "حق إسرائيل في الدفاع عن نفسها". وكانت عائلات بعض المفقودين في البرلمان يشاهدون بيانه. كما أكد أن المملكة المتحدة تقدم مساعدة إنسانية إضافية بقيمة 10 مليون جنيه إسترليني للشعب الفلسطيني وخلال حديثه عن هجوم حماس في إسرائيل في نهاية الأسبوع الماضي، قال سوناك: "المسنون والرجال والنساء والأطفال والرضع، قتلوا وتم تشويههم وحرقهم أحياء". يجب أن نسميها باسمها: إنها بروجروم". ويشير مصطلح "بروجروم" إلى الهجمات العنيفة والمنظمة ضد اليهود، ويرتبط بشكل خاص بشرق أوروبا قبل الحرب العالمية الثانية.</p>		<p>Total word of text= 116 (Penalties= issues minor × Severity Multiplier minor + issues major × Severity Multiplier major) / word count) <math>P = (5*1+0*0)/116</math> <math>= 5/116</math> <math>= 0.04</math> <math>100 - (0.02 * 100) + 0 = \text{Quality Score}</math> <math>100 - 4 + 0 = 96</math></p> <p>Quality Score of the text is <u>96%</u></p>
Source: (BBC, 16/Oct/2023)		
ST/TT	Errors that have been detected	Issue Type/ Penalties
ST( A)	.. Mr. Sunak called the attack a " <u>pogrom</u> ", adding the UK must "support absolutely Israel's right to defend itself."	Minor: Fluency (spelling) and Terminology (Inconsistent use) P= 0.1
TT	وصف سوناك الهجوم بأنه 'بروجروم'. مضيفاً أن المملكة المتحدة يجب ان تدعم بشكل تام "حق اسرائيل في الدفاع عن نفسها".	
ST (b)	The whole text	Minor: Design (Text alignment) P= 0.008
TT	"كامل النص"	

In example (4A), the word "بروجروم" means "the organized killing of large numbers of people, because of their race or religion (originally the killing of Jewish people in Russia) (Oxford Dictionary Online). It is repeated three times and was transliterated rather than translated into Arabic. It should be translated as "مذبحة منظمة" in order to clear for the Arab audience. This minor issue of fluency has a penalty of (0.06). It should be mentioned that the text also has major issue of Design dimension like other text mentioned in another example. In this example the penalties of Design were (0.008). In conclusion, ChatGPT's translation was remarkably high reaching 96%.

**Example (5)**

Translation Generated by ChatGPT		Calculated Penalties & Quality Scores
<p><b>You</b></p> <p>"' Hamas acquires its weapons through smuggling or local construction and receives some military support from Iran," the CIA's World Factbook says.</p> <p>While the Israeli and US governments have yet to find any direct role by Iran in last weekend's raids, experts say the Islamic Republic has long been Hamas' main military supporter, smuggling weapons into the enclave through clandestine cross-border tunnels or boats that have escaped the Mediterranean blockade.</p> <p>"Hamas' tunnel infrastructure is still massive despite Israel and Egypt regularly degrading it," said Bilal Saab, senior fellow and director of the Defense and Security Program at the Middle East Institute (MEI) in Washington."</p> <p>translate into Arabic</p> <p><b>ChatGPT</b></p> <p>حصلت حماس على " (CIA) حسب "كتاب حقائق العالم" الخاص بوكالة الاستخبارات المركزية المركزية "أسلحتها من خلال التهريب أو الإنتاج المحلي وتلقى بعض الدعم العسكري من إيران</p> <p>بينما لم تجد الحكومتان الإسرائيلية والأمريكية دوراً مباشراً لإيران في هجمات العطللة الأخيرة، يقول الخبراء إن الجمهورية الإسلامية كانت داعمة رئيسية لحماس لفترة طويلة، حيث تهرب الأسلحة إلى القطاع من خلال أنفاق سرية عبر الحدود أو قوارب تفلت من الحصار البحري في البحر الأبيض المتوسط.</p> <p>في (MEI) قال بلال سعد، كبير الباحثين ومدير برنامج الدفاع والأمان في معهد الشرق الأوسط واشنطن: "بنية أنفاق حماس لا تزال ضخمة على الرغم من جهود إسرائيل ومصر المستمرة في تدهورها بشكل منتظم".</p> <p>Source: (CNN, 12/Oct/2023)</p>		<p>Total word of text= 103 (Penalties= issues minor × Severity Multiplier minor + issues major × Severity Multiplier major) / word count) P = (5*1+0*0)/103 = 5/103 = 0.04 100 – (0.04 * 100 )+ 0 = Quality Score 100 – 4 + 0 = 96 Quality Score of the text is <u>96%</u></p>
ST/TT	Issues that have been detected	Issue Type/ Penalties
ST (a)	... Hamas acquires its weapons through smuggling or local construction and receives some military support from Iran," the CIA's World Factbook says...	<b>Critical:</b> Fluency (word order and coherence), and Accuracy (untranslated) P= 0.08
TT	حصلت حماس: (CIA) حسب "كتاب حقائق العالم" الخاص بوكالة الاستخبارات المركزية المركزية "أسلحتها من خلال التهريب أو الإنتاج المحلي وتلقى بعض الدعم العسكري من إيران."	
ST (b)	While the Israeli and US governments have yet to find any direct role by Iran in last weekend's raids...	<b>Major:</b> Accuracy (mistranslation) P= 0.07
TT	بينما لم تجد الحكومتان الإسرائيلية والأمريكية دوراً مباشراً لإيران في هجمات العطللة الأخيرة،	
ST (c)	... Hamas' tunnel infrastructure is still massive despite Israel and Egypt regularly degrading it said Bilal Saab, senior fellow and director of the Defense and Security Program at the Middle East Institute (MEI) in Washington...	Minor: Fluency (spelling and word order) P= 0.12
TT	في (MEI) قال بلال سعد كبير الباحثين ومدير برنامج الدفاع والأمان في معهد الشرق الأوسط واشنطن: "بنية أنفاق حماس لا تزال ضخمة على الرغم من جهود إسرائيل ومصر المستمرة في تدهورها بشكل منتظم"	

In (5 a), there is a critical issue of dimension with a penalty of (0.08) in the translation of the following sentence, "Hamas acquires its weapons through smuggling or local construction and receives some military support from Iran," the CIA's World Factbook says". The lack of coherence and word order leads to misunderstanding for Arab audience. It is hard for the reader to understand how Hamas acquires its weapons. Therefore, the correct rendition should be وفقاً لكتاب الحقائق الصادر عن المخابرات

## An Assessment of English-Arabic Translation Using ChatGPT

المركزية الأمريكية "حصلت حماس على أسلحتها من خلال التهريب أو التصنيع المحلي وتلقت بعض الدعم العسكري من إيران"

In addition, in (5b), the phrase “*in last weekend’s raids*” was translated as “هجمات العطلة الأخيرة”. This word-for-word translation lacks coherence, leading to ambiguity in the target text. This major issue has a penalty of (0.07). It should be translated as “هجمات نهاية الأسبوع الماضي”. Moreover, in (5c), there is a minor issue with the acronym (MEI) as it should be omitted because it was translated in full as “معهد الشرق الأوسط”. Another minor issue in (5c) is the misspelling of the name “Saab” as “سعد”, this issue has a penalty of (0.17). Although there were many issues in this example, the translation reached 96%.

### Example (6)

ChatGPT Translation		Calculated Penalties & Quality Scores
<p><b>You</b></p> <p>"I am still frightened," farm worker Nattaphong Duangchan told CNN, now back home in Thailand after the October 7 attacks.</p> <p>Wounded by shrapnel, he hid for two days on the farm where he worked along the Gaza border, before being rescued by Israeli security forces. Return is not an option for him, he said.</p> <p>"Nothing is left there, and I am too afraid to go back."</p> <p>But workers like him are sorely needed. On Israel's farms, "the vast majority nowadays of the kind of the people who are doing the grunt work, you know, the heavy work for low wages are Thai and have been since the early 90s," Matan Kaminer, an anthropologist studying Thai labor in Israel, told CNN.</p> <p>translate into Arabic</p> <p><b>ChatGPT</b></p> <p>وهو الآن عائد إلى بلاده في CNN، ما زلت خائفاً، قال عامل الزراعة نتافونغ دوانجتشان لشبكة تايلاند بعد هجمات 7 أكتوبر.</p> <p>جرحه شظايا القذائف، واختبأ لمدة يومين في المزرعة التي كان يعمل فيها على طول حدود غزة، قبل أن يتم إيقاده من قبل قوات الأمن الإسرائيلية. وقال إن العودة ليست خياراً بالنسبة له.</p> <p>"لم يتبقى شيء هناك، وأنا خائف جداً من العودة".</p> <p>ولكن العمال مثلهم مطلوبون بشدة، في مزارع إسرائيل، "غالبية الأشخاص الذين يقومون بالعمل الشاق بأجور منخفضة في الوقت الحالي هم تايلانديون وكانوا كذلك منذ أوائل التسعينات"، كما قال CNN، ماتان كامينر، عالم الأنثروبولوجيا الذي يدرس العمالة التايلاندية في إسرائيل لشبكة</p>		<p>Total word of text= 102 (Penalties= issues minor × Severity Multiplier minor + issues major × Severity Multiplier major) / word count)</p> $P = (4*1+0*0)/102$ $= 4/102$ $= 0.03$ $100 - (0.03 * 100) + 0 = 97\%$ <p>Quality Score of the text is 97%</p>
Source: (CNN, 26/Nov/2023)		
ST/TT	Issues that have been detected	Issue Type/ Penalties
ST (a)	... I am still frightened," farm worker Nattaphong Duangchan told CNN, now back home in Thailand after the October 7 attacks...	Major: Fluency (word order and coherence & cohesion) P= 0.04
TT	وهو الآن عائد إلى بلاده في CNN، ما زلت خائفاً، قال عامل الزراعة نتافونغ دوانجتشان لشبكة تايلاند بعد هجمات 7 أكتوبر	
ST (b)	... Wounded by shrapnel, he hid for two days on the farm where he worked along the Gaza border, ...	Major: Accuracy (mistranslation) P= 0.06
TT	جرحه شظايا القذائف، واختبأ لمدة يومين في المزرعة التي كان يعمل فيها على طول حدود غزة	
ST (c)	... Matan Kaminer, an anthropologist studying Thai labor in Israel, told CNN...	Minor: Fluency (word order) P= 0.28
TT	كما قال CNN، ماتان كامينر، عالم الأنثروبولوجيا الذي يدرس العمالة التايلاندية في إسرائيل لشبكة	

In (6 a), there is a major issue of fluency (i.e., word order cohesion and coherence) with a penalty of (0.04). The majority of words are confused in the translation, which results in misunderstanding by the readers. The translation should be:

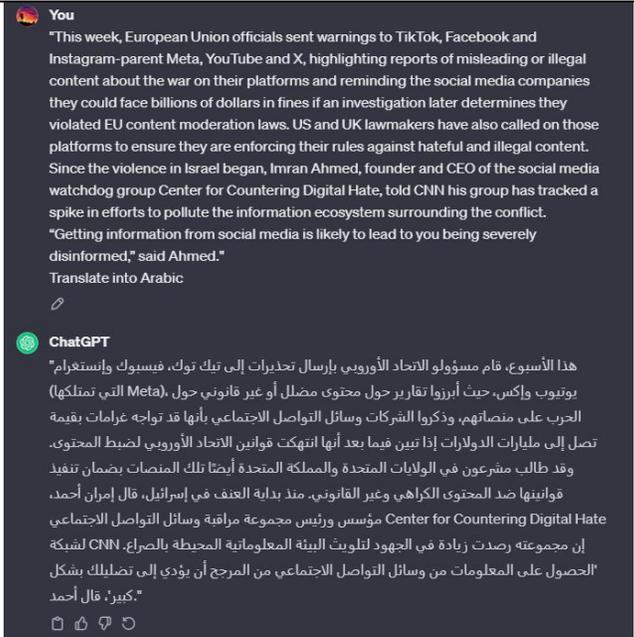
قال عامل الزراعة نتافونغ دوانجشتان لشبكة CNN "مازلت خائفا"، وهو الآن عائد الى بلاده في تايلند بعد هجمات 7 أكتوبر"

Moreover, in (6b), there is also a major issue of accuracy, which is a mistranslation of the phrase “Wounded by shrapnel” as it was translated as “جرحه شظايا القذائف” with a penalty of (0.06). This is a word-for-word translation, and it is inadequate as there is a personification for the word “shrapnel”, and this leads to a misunderstanding of the meaning. Therefore, it should be translated "أصيب بشظايا". In addition, in (6c), there is a minor issue of fluency (word order) as this may confuse the reader about who said this statement: CNN or Matan Kaminer. This error has a penalty of 0.28. It should be translated as:

أخبر ماتان كامينر ، عالم الأنثروبولوجيا الذي يدرس العمالة التايلاندية في اسرائيل , شبكة CNN

At last, the quality of ChatGPT’s translation was 97%.

**Example (7)**

Translation Generated by ChatGPT		Calculated Penalties & Quality Scores
 <p>Source: (CNN, 16/Oct/2023)</p>		<p>Total word of text= 125                      Penalties= issues minor × Severity Multiplier minor + issues major × Severity Multiplier major)/ wordcount)  <math>P = (5*1+0*0)/125</math>  <math>= 5/125</math>  <math>= 0.04</math>  <math>100 - (0.04 * 100) + 0 = \text{Quality Score}</math>  <math>100 - 4 + 0 = 96</math>                      Quality Score of the text is <u>96%</u></p>
ST/TT	Issues that have been detected	Issue Type/ Penalties
ST (a)	...This week, European Union officials sent warnings to TikTok, Facebook and Instagram-parent Meta, YouTube and X, highlighting reports of misleading or illegal content about the war on their platforms...	<b>Minor: Fluency (Coherence and word order)</b> P= 0.03
TT	هذا الأسبوع، قام مسؤولو الاتحاد الأوروبي بإرسال تحذيرات إلى تيك توك، فيس بوك وانستغرام	

An Assessment of English-Arabic Translation Using ChatGPT

	فيسبوك وانستغرام ويوتيوب وإكس، حيث ابرزوا تقارير حول محتوى مضلل أو غير قانوني حول، (Meta التي تمتلكها) الحرب على منصاتهم	
ST (b)	...Imran Ahmed, founder and CEO of the social media watchdog group Center for Countering Digital Hate, told CNN his group has tracked a spike in efforts to pollute the information ecosystem surrounding the conflict...	Major: Accuracy (mistranslation, untranslated) , Fluency (word order), P= 0.07
TT	قال عمران أحمد، مؤسس ورئيس مجموعة مراقبة وسائل التواصل الاجتماعي إن مجموعته رصدت زيادة في الجهود لتلويث البيئة المعلوماتية المحيطة بالصراع. CNN لشبكة	

In (7 a), there is a minor issue of fluency (word order and coherence) with a penalty of 0.03. The expression “ Facebook and Instagram-parent Meta” was split and translated separately as انستغرام فيسبوك و... (Meta التي تملكها). So, the translation lacks coherence. Therefore, the translation should be as Meta. In addition, in (7b), there is a major issue of accuracy. ChatGPT did not translate the expression “Center for Countering Digital Hate” into Arabic. In addition, there is a minor issue in the transliteration of proper names into Arabic, i.e., Imran should be transliterated in عمران not عمران. This issue has a penalty of 0.07. The translation should be as follows:

وقال عمران أحمد، المؤسس والرئيس التنفيذي لمجموعة مراقبة وسائل التواصل الاجتماعي، مركز مكافحة الكراهية الرقمية، لشبكة CNN، إن مجموعته لاحظت ارتفاعاً كبيراً في الجهود المبذولة لتلويث النظام البيئي المعلوماتي المحيط بالنزاع. Therefore, the translation quality was 96%.

Example (8)

Translation Generated by ChatGPT	Calculated Penalties & Quality Scores
<p><b>You</b></p> <p>"Both Hamas and Israel have committed war crimes in the past month, the United Nations human rights chief said, as thousands of Palestinians fled south amid Israel's intensifying offensive against the Islamist militant group.</p> <p>"The atrocities perpetrated by Palestinian armed groups on October 7 were heinous, brutal and shocking, they were war crimes – as is the continued holding of hostages," said Volker Türk, the UN High Commissioner for Human Rights.</p> <p>"The collective punishment by Israel of Palestinian civilians amounts also to a war crime, as does the unlawful forcible evacuation of civilians," he added.</p> <p>Israel declared war on Hamas on October 7 after militants killed 1,400 people in Israel and kidnapped about 240 others, according to the Israel Defense Forces (IDF)."</p> <p>Translate into Arabic</p> <p><b>ChatGPT</b></p> <p>الرئيس التنفيذي لحقوق الإنسان في الأمم المتحدة أكد أن كلا من حماس وإسرائيل ارتكبتا جرائم حرب خلال الشهر الماضي. اعتبر أن الفظائع التي ارتكبتها الفصائل المسلحة الفلسطينية في 7 أكتوبر كانت جرائم حرب، وكذلك استمرار احتجاز الرهائن. وأشار إلى أن العقوبة الجماعية التي تفرضها إسرائيل على المدنيين الفلسطينيين تُعد أيضاً جريمة حرب، فضلاً عن الإغلاء القسري غير القانوني للمدنيين. أعلنت إسرائيل الحرب على حماس في 7 أكتوبر بعد قتل المقاتلين نحو 1400 شخص في (IDF) إسرائيل واحتطاف نحو 240 آخرين وفقاً للقوات الدفاع الإسرائيلية.</p>	<p>Total word of text= 84</p> <p>Penalties= issues minor × Severity Multiplier minor + issues major × Severity Multiplier major) / wordcount)</p> <p><math>P = (5*1+2*10*0)/84</math></p> <p><math>= 25/84</math></p> <p><math>= 0.29</math></p> <p><math>100 - (0.29 * 100) + 0 =</math></p> <p>Quality Score</p> <p><math>100 - 29 + 0 = 71</math></p> <p>Quality Score of the text is <u>71%</u></p>

Source: (CNN, 09/Nov/2023)		
ST/TT	Issues that have been detected	Issue Type/ Penalties
ST (a)	... Both Hamas and Israel have committed war crimes in the past month, the United Nations human chief said, <u>as thousands of Palestinians fled south amid Israel's intensifying offensive against the Islamist militant group...</u>	Critical: Accuracy (omission) P= 0.52
TT	الرئيس التنفيذي لحقوق الإنسان في الأمم المتحدة أكد أن كلا من حماس وإسرائيل ارتكبتا جرائم حرب خلال الشهر الماضي،	
ST (b)	The atrocities perpetrated by Palestinian armed groups on October 7 were heinous, brutal and shocking, they were war crimes – as is the continued holding of hostages,” <u>said Volker Türk, the UN High Commissioner for Human Rights...</u>	<i>Critical:</i> Accuracy (omission) P= 0.58
TT	اعتبر ان الفظائع التي ارتكبتها الفصائل المسلحة الفلسطينية في 7 أكتوبر، كانت جرائم حرب، وكذلك احتجاز الرهائن.	
ST (c)	Israel declared war on Hamas on October 7 after militants killed 1,400 people in Israel and kidnapped about 240 others, according to the Israel Defense Forces (IDF)	Minor: Fluency (addition) and word order) P= 0.03
TT	اعلنت اسرائيل الحرب على حماس في 7 أكتوبر بعد قتل المقاتلين لنحو 1400 شخص في "IDF إسرائيل واختطاف نحو 240 آخرين وفقاً (IDF) للقوات الدفاع الإسرائيلية".	

In (8a), there is a critical issue of accuracy, which is the omission of a whole sentence “as thousands of Palestinians fled south amid Israel’s intensifying offensive against the Islamist militant group...”. This omission leads to misunderstanding and lack of coherence in the target text as it causes ambiguity for the target readers. This issue has a penalty of 0.52. The translation should be:

إن كلا من حماس وإسرائيل ارتكبتا جرائم حرب في الشهر الماضي، مع فرار آلاف الفلسطينيين جنوباً وسط الهجوم الإسرائيلي المكثف ضد الجماعة الإسلامية المسلحة.

In (8b), there is also a critical issue of accuracy which is the omission of the clause “said Volker Türk, the UN High Commissioner for Human Rights”, and “heinous, brutal and shocking”. This issue has a penalty of 0.58. This omission again leads to misunderstanding. Therefore, the translation should be as follows:

وقال فولكر تورك، مفوض الأمم المتحدة السامي لحقوق الإنسان: "إن الفظائع التي ارتكبتها الجماعات الفلسطينية المسلحة في 7 أكتوبر/تشرين الأول كانت بشعة ووحشية وصادمة، وكانت بمثابة جرائم حرب - كما هو الحال مع استمرار "احتجاز الرهائن".

In addition, there is a minor issue of fluency as there is an inappropriate additional information in the translation, leading to incoherence. The words IDF and إسرائيل are redundant in the translation, and they must be omitted. This error has a penalty of 0.03. The translation should be:

أعلنت إسرائيل الحرب على حماس في 7 أكتوبر/تشرين الأول بعد أن قتل المسلحون 1400 شخص في إسرائيل واختطفوا حوالي 240 آخرين، وفقا للجيش الإسرائيلي.

The quality of the translation of this example was 71%.

## 5. Conclusions

The study found that ChatGPT's translation quality for political news from English into Arabic is 93%. Most errors were related to accuracy, fluency, terminology, and design. To illustrate, there were problems with coherence, cohesion, word order, grammar, mistranslation, untranslated parts, addition, omission, and the transliteration of proper nouns. Some of these issues were critical, major, or minor. Some of them affect the fluency and accuracy of the translation and lead to misunderstanding of the text. In addition, the translation provided by ChatGPT is to be reviewed by human translators to avoid any possible errors or inadequacies. ChatGPT should be trained on Arabic language structures and the alignment of the Arabic texts. Although ChatGPT saves time, convenience, and speed in performance, users should revise the translation provided. Therefore, the users of ChatGPT should revise the translation to ensure that it is adequate for the target readers. The study's recommendations also extend to the technology field, urging developers to continually improve and develop machine translation algorithms for ChatGPT in Arabic, as improving translation models can improve the accuracy and reliability of the translation. Finally, the MQM framework functions as a helpful model in analyzing the translation errors made by ChatGPT.

## تقييم الترجمة من اللغة الإنجليزية إلى اللغة العربية باستخدام ChatGPT

أحمد الحراشنة، سهى تيسير أبو الكراسنة، رأفت الروسان  
قسم الترجمة جامعة اليرموك، الأردن

### الملخص

تهدف هذه الدراسة إلى تقييم جودة ترجمة النصوص السياسية من الإنجليزية إلى العربية باستخدام أداة الذكاء الاصطناعي ChatGPT. وتعتمد الدراسة على الإصدار المحدث من إطار مقاييس الجودة متعددة الأبعاد (MQM) لتقييم جودة الترجمة التي تنتجها ChatGPT. وشملت عينة الدراسة 30 مقالاً سياسياً مختاراً من قناتي CNN و BBC World باللغة الإنجليزية وترجمتها إلى اللغة العربية. وتوصلت الدراسة إلى أن ثمة عدة أخطاء لغوية للترجمات باستخدام ChatGPT، مثل السلاسة والدقة. وغالباً ما تسبب الترجمات التي تفتقر إلى هذه العناصر المهمة سوء فهم المستخدمين وإرباكهم، وعلى الرغم من هذه المشكلات، فقد أظهرت أداة ChatGPT أداءً جيداً في ترجمة الأخبار السياسية. حيث بلغت جودة الترجمة الإجمالية لـ 93% ChatGPT.

الكلمات المفتاحية: الذكاء الاصطناعي، النصوص السياسية، الإنجليزية، العربية.

## References

- Al-Mubassyr, Muhammad. 2023. Accuracy in translating Arabic texts through automated applications: Google Translate and ChatGPT as examples. Surabaya, Indonesia: In Proceedings of International Conference on Islamic Civilization and Humanities.
- Yuxin, Cheng, Wang Ruochen, Chen Jiawei, Chao Yijun, Aliye Maimaitili, and Zhang Haoruo .2023. *Context-Based AI Translation from a Globalization Perspective: A Case Study of ChatGPT*. Shanghai, China: Sino-US English Teaching.
- Gao, Yuan, Ruili Wang, and Feng Hou .2023. *How to Design Translation Prompts for ChatGPT: An Empirical Study*. <https://arxiv.org/pdf/2304.02182>
- Hutchins, W. John .1995. Machine translation: A brief history. In *Concise history of the language sciences: from the Sumerians to the cognitivists*, edited by E.F.K. Koerner and R.E. Asher, 431-445. Pergamon Press: Oxford.
- Işim, Çağrı and Yavuz Selim Balcıoğlu. 2023. *ChatGPT: Performance of Translate*. Paper presented at the Conference of 3rd International ACHARAKA Congress on Humanities and Social Sciences at Acharaca, 11th to 13th March 2023, Turkey.
- Kadaoui, Karima, Samar M. Magdy, Abdul Waheed, Md Tawkat Islam Khondaker, Ahmed Oumar El-Shangiti, El Moatez Billah Nagoudi, Muhammad Abdul-Mageed. 2023. *TARJAMAT: Evaluation of Bard and ChatGPT on Machine Translation of Ten Arabic Varieties*. Singapore: arXiv preprint arXiv:2308.03051 [Computation and Language; Machine Learning].
- Khoshafah, Faten .2023. *ChatGPT for Arabic-English Translation: Evaluating the Accuracy*. Research Square. Ministry of Education, Yemen.
- Kok, Joost N., Egbert J.W. Boers, Walter A. Kosters, Peter van der Putten, and Mannes Poel .2009. *Artificial intelligence: definition, trends, techniques, and cases*. In Knowledge for sustainable development: an insight into the Encyclopedia of life support systems, 1 edited by UNESCO Publishing, 1095-1107. ELOSS Publishers.
- McCarthy, John. 2007. *What is artificial intelligence?* California, United States: Computer Science Department, Stanford University.
- Moneus, Ahmed Mohammed and Yousef Sahari .2023. *Artificial Intelligence and Human Translation: A Contrastive Study Based on Legal Texts*. Bisha, Saudi Arabia.
- Multidimensional Quality Metrics (MQM) Issue Types. 2016. WayBackMachine. <https://web.archive.org/web/20160505053317/http://www.qt21.eu/mqm-definition/issues-list-2015-12-30.html#term-inconsistency>
- Pokorny, Elizabeth. 2023. ChatGPT vs translation services: A comparative study. Retrieved from <https://www.weglot.com/blog/chatgpt-translation#:~:text=With%20ChatGPT%20being%20a%20form,or%20speech%2C%20into%20different%20languages>
- Sabzalieva, Emma and Arianna Valentini 2023. *ChatGPT and artificial intelligence in higher education: quick start guide*. Paris, France. UNESCO IESALC.

- Siu, Sai Cheong. 2023. *ChatGPT and GPT-4 for Professional Translators: Exploring the Potential of Large Language Models in Translation*. Hong Kong, China: SSRN 4448091.
- Srivastava, Nitish, Hinton, Geoffrey E., Krizhevsky, Alex, Sutskever, Ilya, and Salakhutdinov, Ruslan. 2014. *Dropout: a simple way to prevent Neural Networks from overfitting*. Ontario, Canada: Journal of Machine Learning Research- JMLR.
- Verma. Nilesh. 2022. An Overview of ChatGPT: How the Chatbot Model Works and Its Capabilities and Limitations. <https://medium.com/@techynilesh/an-overview-of-chatgpt-how-the-chatbot-model-works-and-its-capabilities-and-limitations-857d970c4b2a>
- Wang, Junchao. 2023. *A ChatGPT-driven Study of Translating and Post-editing Chinese Enterprise Publicity Materials*. Guangdong, China: Modern Languages, Literatures, and Linguistics.
- Wu, Jinrui. 2023. *A Comparative Analysis of Chinese-English Translation Quality Based on ChatGPT: A Case Study of Chinese Characteristic Words*. Xi'an, China: Journal of Social Science Humanities and Literature.
- Zhao, Yuming .2023. *The State-of-art Applications of NLP: Evidence from ChatGPT*. Tianjin, China: Highlights in Science, Engineering and Technology.
- Zheng, Xianrui, Chao Zhang, and Philip C. Woodland .2021. *Adapting GPT, GPT-2 and BERT language models for speech recognition*. Cambridge, United States: IEEE Automatic Speech Recognition and Understanding Workshop (ASRU).