

## RESEARCH ARTICLE

# Leveraging Large Language Models for Enhanced Back-Translation: Techniques and Applications

JI-BUM CHUNG<sup>1</sup> AND TAEHYUN KIM<sup>2</sup><sup>1</sup>Department of Civil, Urban, Earth, and Environmental Engineering, Ulsan National Institute of Science and Technology, Ulsan 44919, Republic of Korea<sup>2</sup>Division for Environmental Planning, Korea Environment Institute (KEI), Sejong 30147, Republic of Korea

Corresponding author: Taehyun Kim (kimth@kei.re.kr)

This work was supported in part by the Research Work “Comparative Study on Urban Planning Elements and Policy Acceptability to Reduce Residents’ Carbon Footprint—Targeting Korea and Germany” under Grant RR2024-15, in part by the “Development of Planning Support Technology based on Carbon Spatial Map” under Grant 2025-014(R), conducted by the Korea Environment Institute (KEI), in part by the Korean National Research Foundation under Grant RS-2025-00517091, and in part by the Korea Agency for Infrastructure Technology Advancement (KAIA) of the Ministry of Land, Infrastructure and Transport under Grant RS-2023-00242291.

**ABSTRACT** Cross-cultural studies are prevalent in academia, yet challenges arise in conducting objective research due to linguistic and cultural disparities. Rigorous international comparative research requires appropriate questionnaires that can be used in all countries, and translation becomes an extremely important process. Brislin’s back-translation method is widely recognized, but it usually requires many skilled bilingual translators and is both time-consuming and expensive. This study aims to overcome these limitations by using Large Language Model (LLM) AI technology. We utilized the Application Programming Interfaces (APIs) of well-known commercial LLM models such as ChatGPT3.5, ChatGPT4o, Google-Gemini, and Anthropic-Claude 3. The entire program was built using the Python programming language, and the user interface was built using the Streamlit library. This pilot study’s results confirm the feasibility of LLM-assisted back-translation, particularly for complex topics like carbon footprint reduction planning. This represents a significant advance over traditional back-translation methods, offering substantial time and cost savings.

**INDEX TERMS** Artificial intelligence, machine translation, cross-cultural communication.

## I. INTRODUCTION

People’s thoughts can vary greatly depending on their country or the culture they belong to. While understanding the significance of certain research findings may be challenging within a single country or cultural context, a more conclusive understanding can be attained through cross-cultural comparisons. For example, consider a survey that asks individuals’ level of agreement with climate change mitigation policies implemented by their governments. If researchers seek to derive policy implications from such survey results, they must compare their own nation’s stance with that of other countries. In other words, researchers can derive practical policy recommendations by comparing their country’s results with others, identifying areas for improvement.

Cross-cultural comparisons are particularly important in fields that address global issues such as climate change

mitigation and adaptation. Cross-cultural research enables researchers, policymakers, and communities to understand how different cultural perspectives, values, and practices influence climate perceptions, mitigation efforts, and adaptation strategies. It can provide valuable insights into knowledge sharing, policy-making, and global cooperation.

Hence, to ensure the significance of survey research among different countries, it is imperative to conduct cross-cultural comparisons of its results. Cross-cultural studies are common in academia, yet challenges arise in conducting objective research due to linguistic and cultural disparities. This is especially true for self-administered surveys conducted in countries with different languages. Rigorous international comparative research requires appropriate questionnaires (instruments) that can be used in all countries, and translation becomes an extremely important process in this regard [1].

Generally, simple, direct, and unidirectional translation cannot guarantee content equivalence between the original language and the translated language. The equivalence

The associate editor coordinating the review of this manuscript and approving it for publication was Abdel-Hamid Soliman<sup>1b</sup>.

between the original text and the translation can be measured from various perspectives, such as: semantic, idiomatic, experiential, and conceptual equivalence [1], [2], [3]. Various methodologies have been proposed to ensure equivalence between the original text and the translation, but the most well-known method is the back-translation methodology proposed by Brislin [4]. However, Brislin’s classic back-translation model needs multiple proficient bilingual translators and continuous collaboration among them, potentially incurring significant time and expense. Consequently, researchers have opted to utilize various forms of modified methodologies rather than employing this traditional approach [1], [2], [3], [5]. However, these modified methods still require a certain level of bilingual translators or experts and involve multiple rounds of consultation and collaboration.

This study aims to overcome these limitations by using large language model AI technology. Specifically, it aims to significantly enhance the efficiency of the current back-translation technique by utilizing commercially accessible Large Language Models (LLMs) to replace the expensive and laborious translation and back-translation processes.

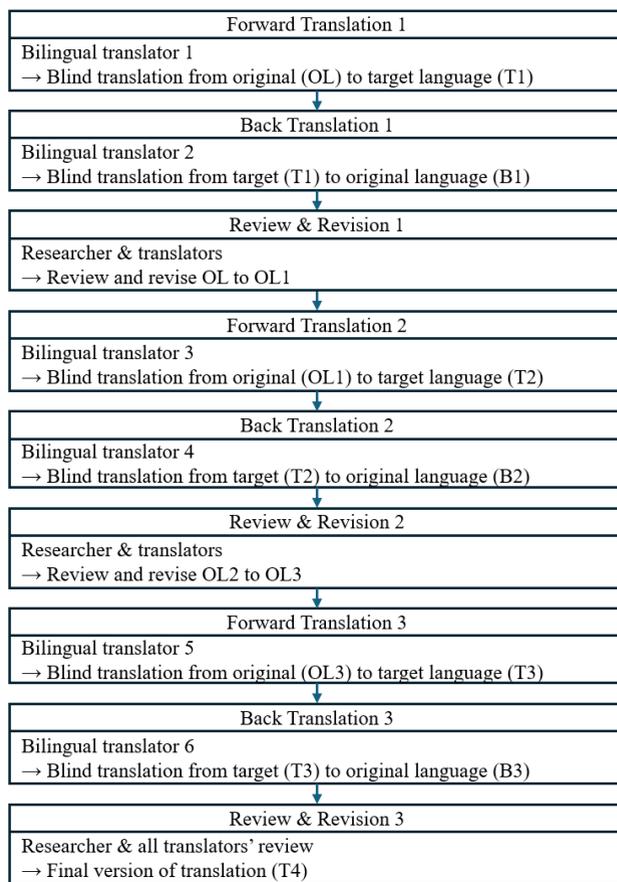
**II. BACKGROUND**

**A. CROSS-CULTURAL STUDY AND BACK-TRANSLATION**

Cross-cultural research involves comparing people from different cultures to understand how their different backgrounds affect their behavior [6]. The efficiency of cross-cultural research can be improved using established instruments with proven psychometric properties. However, to ensure validity, instruments must undergo cultural adaptation and translation; the full potential of cross-cultural research can only be realized when researchers employ instruments that are culturally appropriate [1]. Consequently, the translation process holds critical importance in cross-cultural studies.

Techniques for maintaining equivalence between original and translated measures include back-translation method, bilingual technique, committee approach, and pretest procedure [4]. The classic back-translation model proposed by Brislin [4] has been widely employed for instrument validation. It requires an iterative process involving repeated independent translation and back-translation by a team of translators [1]. The back-translation process can be described as follows:

The first bilingual translator (bilingual translator 1) translates an instrument from the original language (OL) to the target language (T1); a second bilingual translator (bilingual translator 2) independently back-translates the instrument from the target language (T1) to the original language (B1). Subsequently, the researcher and two translators compare the two versions of the instrument to ensure equivalence. When discrepancies are found, the team revises the translation to achieve a more appropriate form (review and revision 1). Then, another translator (bilingual translator



**FIGURE 1. Classical back-translation process.**

3) independently retranslates the revised original text (OL1) into the target language (T2). Another independent translator (bilingual translator 4) then back-translates the revised target language (T2) into the original language (B2). Once again, the researcher and translators compare the two versions and revise it to achieve a more proper form (OL3). These iterative procedures continue until a team of researchers and bilingual translators reaches a consensus that the two versions of the instruments are identical and devoid of any errors in meaning. This iterative process can be summarized in Figure 1.

The Brislin’s classic translation model can be enhanced by incorporating additional bilingual translators with diverse backgrounds to capture the nuances of the language more effectively. For example, Beaton et al. [3] employed two bilingual translators to simultaneously translate a medical questionnaire, wherein one translator possessed knowledge of medical or clinical concepts, while the other did not. The inexperienced translator was more likely to capture subtle differences in the original text that might not match the interpretation of the expert translator. They were less influenced by academic goals and instead provided a translation that used the language of the intended audience, often showing possible unclear parts in the original questionnaire.

Although, these back-translation procedures can be the most appropriate way of translation in cross-cultural studies, there are challenges in finding a sufficient number of bilingual translators and the process can be time-consuming and expensive [1], [2], [5].

## B. MACHINE TRANSLATION AND LARGE LANGUAGE MODEL

Computer algorithms can be employed for more efficient translation of natural language into another language, a process known as machine translation (MT). With regard to methodologies, MT approaches can be broadly categorized into rule-based and corpus-based method [7]. Initially, rule-based methodologies were prevalent in the field. Rule-based translation approaches utilized bilingual dictionaries and manually constructed rules to facilitate the translation process between two languages [7]. However, the inherent complexity of languages and the labor-intensive process of rule creation pose limitations in terms of scalability and adaptability [8].

With the advent of bilingual corpora, corpus-based methodologies gained prominence in the post-2000 era [7]. Corpus-based methodologies encompass three main approaches: example-based machine translation (EBMT), statistical machine translation (SMT), and neural machine translation (NMT). EBMT uses a set of texts in two languages (bilingual corpora) as its main source of information when translating. This approach can be viewed as a form of machine learning that relies on previous instances for knowledge acquisition. SMT utilizes large corpora of bilingual data to decide the most probable translation for a given input, leveraging statistical relationships between original texts and their corresponding human translations. Although both EBMT and SMT can generate more comprehensible translations than RBMT, their potential is limited by the difficulty of constructing appropriate examples or a bilingual corpus [7], [8].

With significant advancements in deep learning technology, researchers have begun to incorporate it into MT. NMT employs an artificial neural network to forecast the probability of a word sequence, typically modeling whole sentences within a unified model. The fundamental principle of NMT is to encode the source language into a dense semantic representation, and subsequently generate the translation by employing an attention mechanism [8]. NMT is the predominant methodology employed today, and prominent companies such as Google and DeepL utilize this technique. Under certain conditions, it can generate translations that are comparable to human translations.

The recent innovation of LLMs has led to increased interest in using them for machine translation. Rather than adjusting a pre-trained language model for the translation task, it is possible to directly prompt extensive generative models to translate a sentence into the desired language. In 2023, . . . . Hendy et al. [9] conducted a comprehensive test and evaluation of this approach on GPT 3.5. Their findings showed that

GPT systems could produce highly fluent and competitive translation outputs, particularly for high-resource language translations, even in the zero-shot setting. Furthermore, . . . . Yan et al. [10] demonstrated that GPT-4's performance was comparable to that of junior translators in terms of the total number of errors made, but it lagged behind senior translators.

Based on these results, we utilized up-to-date LLM technology as bilingual translators for more efficient back translation.

## III. METHOD

In this study, while using the traditional back-translation method structure, we tried to overcome the difficulty of employing multiple independent bilingual translators and improve the efficiency of the entire translation process by utilizing LLM technology. We utilized the Application Programming Interfaces (APIs) of well-known commercial LLM models such as ChatGPT3.5, ChatGPT4o, Google-Gemini, and Anthropic-Claude3. Furthermore, we also made use of the API of DeepL which is based on NMT technology to compare functional efficiency.

The entire program was built using the Python programming language, and the user interface was built using the Streamlit (<https://streamlit.io/>) library. Streamlit is an open-source Python framework for data scientists and AI/ML engineers to deliver interactive data apps. The step-by-step process of the program for translating English text into Korean and back-translating it is as follows.

First, the English source text (OL) to be translated is initially received by the researcher through the input window (Figure 2 Input). Second, the English text (OL) is translated into Korean using the LLM APIs. In this study, a total of five APIs were used, which is the same concept as using five independent bilingual translators in the classical back-translation process of Brislin [6] (Figure 2, Forward translation). The APIs of ChatGPT-3.5 [11], ChatGPT-4o [12], Google-Gemini [13], Anthropic-Claude 3 (Claude 3 Sonnet) [14], and DeepL [15] were used to obtain Korean translations T1, T2, T3, T4, and T5, which were then displayed in the output window. This process is similar to the approach of simultaneously utilizing multiple independent interpreters [3]. The command entered into the LLM API is presented in Korean and is as follows: “다음 문장을 한국어로 번역해줘 한국인들이 자주 쓰는 표현을 사용해서 쉽게 이해할 수 있게 하고 되도록 원문과 같은 문장 형태를 가지도록 만들어주고 번역 결과만 보여줘” which has a meaning of “Translate the following sentence into Korean, using common expressions that Koreans frequently use so that it can be easily understood, and try to maintain the same sentence structure as the original text as much as possible. Show only the translation result.”

Third, back-translation is performed on T1, T2, T3, T4, and T5 (Figure 2, Back translation). This back-translation (B1, B2, B3, B4, and B5) is also conducted using independent LLM APIs, and the results are displayed. The command entered into the LLM API was presented in English

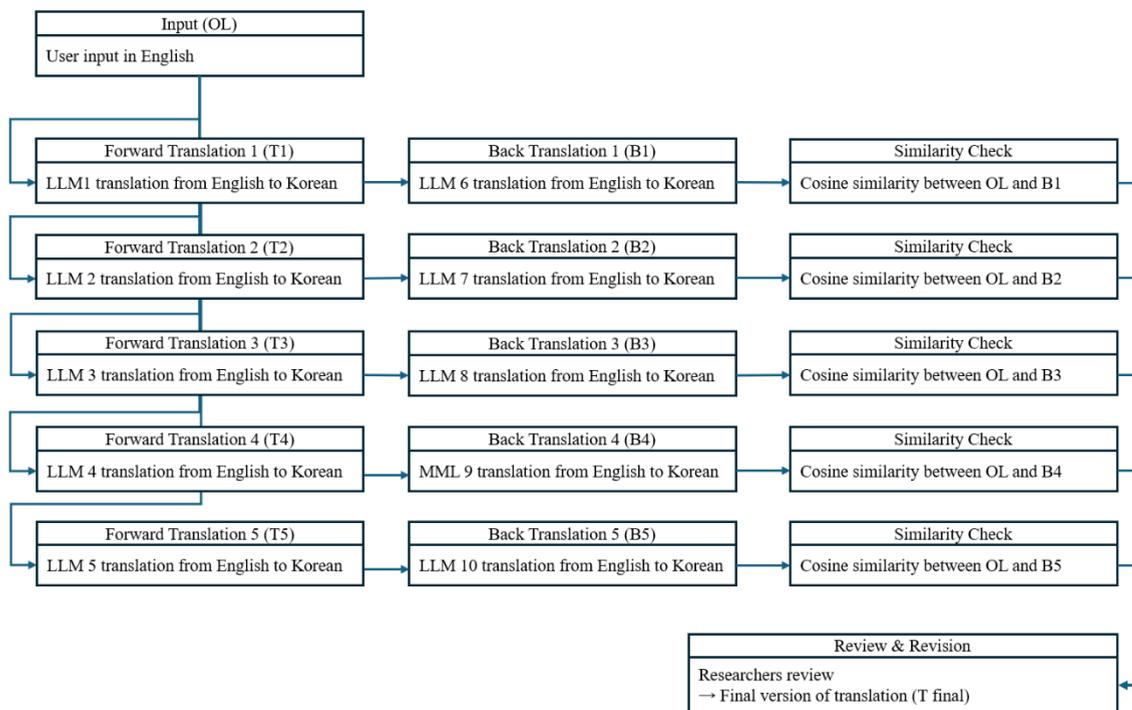


FIGURE 2. Process diagram of translation using LLMs.

and is as follows: “Please translate the following sentence into English, using common American expressions for easy understanding and maintaining the original sentence structure as much as possible. Show only the translated sentence.”

Fourth, the similarities between the original text (OL) and each back-translated text (B1, B2, B3, B4, and B5) are analyzed and outputted (Figure 2, Similarity check). Sentence similarity is computed using the cosine similarity algorithm. In LLM, each word is assigned to different coordinates, and a document is represented by a vector of the number of occurrences of each word within the document. Cosine similarity provides a useful measure of the similarity between two documents in terms of their subject matter [16]. Cosine similarity is the cosine of the angle between the vectors. For instance, two proportional vectors (the meanings of two documents are identical) have a cosine similarity of 1, while two opposite vectors have a similarity of -1.

Subsequently, the researchers examine the translation, back-translation results, and cosine similarity results to select the most proper translation text (Figure 2, Review and revision). This text is then revised to obtain the final translated text (T final). The aforementioned methods are summarized in Figure 2, and Figure 3 illustrates the user interface screen.

Along with the English-Korean translation program, an English-German translation program has also been developed. This is to conduct comparative research on

well-established original English survey instruments in Korea and Germany.

To verify the accuracy of the back translation program, the final translation results were evaluated by native speakers from both Germany and Korea, including both professionals and laypersons. They assessed the linguistic and conceptual precision of the translated statements.

#### IV. PILOT STUDY

##### A. PURPOSE OF PILOT STUDY

Cross-cultural research is essential for addressing global issues such as climate change and carbon footprint (CF) reduction. Successful implementation of such research requires developing tools that can be reliably understood and applied across diverse languages and cultural contexts, thereby ensuring cross-cultural consistency. For this pilot study, we selected Korea and Germany, two countries with distinct linguistic and cultural backgrounds and robust policy efforts to mitigate climate change.

Linguistically, Korean and German differ significantly in vocabulary, grammar, and cultural context. Their vocabularies are mostly different, except for words borrowed from English or other languages. German sentences usually follow a Subject-Verb-Object order, while Korean sentences use a Subject-Object-Verb order, placing the verb at the end. Culturally, Korean writing is more situational, while German writing is more individualistic. German writing often

Input original sentence (English)

Climate change is now affecting every country on every continent. 65/2000

submit

### Forward-Backward Translation

| GPT3.5                                | Back-Translation (google)   | Similar |
|---------------------------------------|---|---------|
| 기후 변화는 이제 모든 나라와 대륙에 영향을 미치고 있습니다.    | Climate change is now affecting every country and continent.      | 0.996   |
| <b>GPT4o</b>                          |   |         |
| 기후 변화는 이제 모든 대륙의 모든 나라에 영향을 미치고 있다.   | Climate change is now affecting every country on every continent. | 0.999   |
| <b>Google - Gemini</b>                |   |         |
| 기후 변화는 지금 전 세계 모든 나라에 영향을 미치고 있어요.    | Climate change is affecting every country in the world right now. | 0.971   |
| <b>Claude3</b>                        |   |         |
| 기후 변화가 이전 모든 대륙의 모든 나라에 영향을 미치고 있어.   | Climate change is now affecting every country on every continent. | 1.0     |
| <b>DeepL</b>                          |   |         |
| 기후 변화는 이제 모든 대륙의 모든 국가에 영향을 미치고 있습니다. | Climate change is now affecting every country on every continent. | 1.0     |

FIGURE 3. User interface of the program.

emphasizes the self, starting with “I,” while Korean writing typically begins by describing the situation.

Although both countries are making significant efforts in implementing climate-friendly policies, their different cultural backgrounds result in varying effects. One such example is the concept of green roofs, which is commonly recognized as an effective urban environmental policy. However, the perception of green roofs differs significantly between Korea and Germany. In Korea, green roofs are often seen as aesthetic enhancements or elements of eco-friendly architecture, whereas in Germany, they are widely regarded as essential for stormwater management and urban biodiversity conservation. These cultural and policy differences highlight the importance of precise translations that preserve both the intent and context of such policies.

This study aims to confirm the consistency of translated statements on CF reduction planning elements in English, Korean, and German while streamlining the traditional back-translation process using modern AI techniques. This approach helps to address the time and resource intensive nature of traditional methods.

## B. INSTRUMENTS AND PROCEDURES

### 1) INSTRUMENTS

The research involved the translation of key statements related to elements of CF reduction planning (e.g., food, energy and waste management policies). These statements

were formulated based on a questionnaire developed to calculate personal CFs based on resident consumption and used in carbon spatial maps, as well as the well-established English statements of the “UN 170 Climate Change Actions” published by the United Nations (UN) in 2020 [17], incorporating the recommendations of local experts in urban planning and environmental sectors. Translation between English, Korean and German was conducted using an application developed with LLM technology, which uses multiple translation models to ensure cross-linguistic consistency.

The study focuses on Korea and Germany due to their distinct policy frameworks and linguistic differences, making them ideal cases for evaluating AI-assisted translation reliability. Multiple AI models were used to evaluate translation consistency, mirroring the process of Brislin’s classic back translation model, which involved employing additional interpreters for the same sentence translation. Furthermore, the researcher reviewed the five translation results to facilitate the utilization of the most appropriate form (Figure 2, Review and revision).

### 2) PARTICIPANTS

The study involved 40 participants, 20 each from Korea and Germany. Participants included government officials, businesspeople, members of civil society organizations, and researchers to ensure diverse input. The participants were categorized as experts and non-experts to assess the effectiveness of the translated statements. The inclusion of non-expert participants (naïve participants) was intended to mitigate the influence of academic goals and elicit translations reflective of lay language, potentially revealing ambiguities in the original questionnaire [3]. Expert and non-expert participants, including native speakers, reviewed the translations to validate linguistic and conceptual accuracy. Their feedback was systematically incorporated into the final versions to enhance precision and usability in diverse contexts.

### 3) PROCEDURE

The app used for translation performed both forward and backward translations, and the results were analyzed using cosine similarity to assess the accuracy and consistency of the translations. The translation results were reviewed by native speakers from both countries (Germany and Korea), including both professionals and laypersons, who assessed the linguistic and conceptual accuracy of the translated statements. The verification of cosine similarity values also helped to decide the closeness of the back-translated texts to the original source statements.

## C. RESULTS

The cosine similarity values for the back-translated texts showed a high degree of linguistic equivalence between the original English statements and the back-translated texts. Statements related to environmental planning policies showed similarity scores between 0.93 and 1.00, showing

a strong preservation of meaning across languages. The analysis showed that the app's translation results were generally reliable, especially for resource-intensive languages such as Korean and German. Most statements, such as those on individual efforts to reduce consumption (0.99 for Korean and 0.98 for German) and support for carbon reduction research (0.997 for Korean and 0.97 for German), showed strong consistency in both forward and back-translations. However, the cosine similarity results varied across the LLM models (GPT-3.5, GPT-4o, Gemini, and Claude 3). For example, the average cosine similarities of GPT-4o, Gemini and Claude 3 on individual efforts were 0.975, 0.978 and 0.974 respectively for Korean.

For validation of the machine translation results, the outputs were reviewed by native speakers from both countries. The expert review also corroborated that the translations effectively conveyed the intended meaning without significant loss of nuance or context. However, some statements showed minor discrepancies between the original and translated versions. The expert review recommended revised statements. The revised Korean and German versions were maintained scores above 0.95 and 0.93 respectively, showing that the revisions did not significantly alter the meaning.

In addition, both expert and non-expert native language reviewers found the translations to be understandable and in line with the cultural and technical terminology specific to their fields. There were minimal discrepancies between the expert and lay reviewers, suggesting the robustness of the translations for wider use.

Overall, the high similarity scores confirm the effectiveness of the AI-assisted translation process, supporting its use for accurate cross-cultural research.

## V. CONCLUSION

The results of this pilot study demonstrate the feasibility of using LLM-assisted translation applications for cross-cultural research, especially in complex areas such as CF reduction planning. The high cosine similarity scores indicate that the AI-assisted back-translation process effectively preserves the original intent and meaning of the statements, reducing the need for multiple rounds of human translation. This represents a significant advance over traditional back-translation methods, offering significant time and cost savings.

International comparative survey research is a prevalent methodology across disciplines. However, maintaining research rigor necessitates substantial time and resources for developing universally applicable surveys across languages and cultures. This study aims to enhance the efficiency of cross-cultural research by leveraging recent advances in LLMs to significantly improve Brislin's classing back-translation method.

To date, research utilizing LLMs for back-translation has primarily focused on improving translation performance for low-resource languages and data augmentation techniques for natural language processing [18], [19], [20]. However,

this study is a novel application research that significantly improves the efficiency of existing research methodologies by utilizing AI, differentiating it from previous studies. In the future, the methodology proposed in this study can be effectively utilized in international comparative research involving surveys.

However, some limitations remain. For example, the system's performance may vary depending on the complexity of the subject matter and the specific linguistic challenges posed by lower-resource languages. While the AI-assisted translation approach demonstrated high efficiency and reliability, the accuracy of AI-generated translations can be affected by complex subject-specific terminology, particularly in fields such as environmental policy and urban planning. Additionally, low-resource languages, which have fewer digital training datasets, may exhibit lower translation precision.

Further research is needed to explore how well this method generalizes to other languages or more specialized technical domains. Additionally, while AI-assisted translation provides valuable efficiency, human oversight is still necessary to catch subtle cultural or contextual differences that may not be captured by cosine similarity metrics alone. Future research should explore hybrid methodologies that integrate AI with human expertise to refine translation accuracy. Reinforcement learning-based AI enhancements could further improve linguistic adaptability and context recognition. Moreover, expanding the study to include additional languages and policy domains will be crucial for validating the generalizability of AI-assisted back-translation techniques.

## ACKNOWLEDGMENT

This paper was written following the research work "Comparative study on urban planning elements and policy acceptability to reduce residents' carbon footprint – targeting Korea and Germany" (RR2024-15) and "Development of Planning Support Technology based on Carbon Spatial Map (2025-014(R))", conducted by the Korea Environment Institute (KEI). This study was supported by the Korean National Research Foundation (RS-2025-00517091) and the Korea Agency for Infrastructure Technology Advancement (KAIA) of the Ministry of Land, Infrastructure and Transport (Grant RS-2023-00242291).

## REFERENCES

- [1] E. Cha, K. H. Kim, and J. A. Erlen, "Translation of scales in cross-cultural research: Issues and techniques," *J. Adv. Nursing*, vol. 58, no. 4, pp. 386–395, May 2007, doi: [10.1111/j.1365-2648.2007.04242.x](https://doi.org/10.1111/j.1365-2648.2007.04242.x).
- [2] T.-I. Tsai, L. Luck, D. Jefferies, and L. Wilkes, "Challenges in adapting a survey: Ensuring cross-cultural equivalence," *Nurse Researcher*, vol. 26, no. 1, pp. 28–32, Jun. 2018.
- [3] D. E. Beaton, C. Bombardier, F. Guillemin, and M. B. Ferraz, "Guidelines for the process of cross-cultural adaptation of self-report measures," *Spine*, vol. 25, no. 24, pp. 3186–3191, Dec. 2000. [Online]. Available: [https://journals.lww.com/spinejournal/fulltext/2000/12150/guidelines\\_for\\_the\\_process\\_of\\_cross\\_cultural.14.aspx](https://journals.lww.com/spinejournal/fulltext/2000/12150/guidelines_for_the_process_of_cross_cultural.14.aspx)
- [4] R. W. Brislin, "Back-translation for cross-cultural research," *J. Cross-Cultural Psychol.*, vol. 1, no. 3, pp. 185–216, Sep. 1970, doi: [10.1177/135910457000100301](https://doi.org/10.1177/135910457000100301).

- [5] S. Y. McGorry, "Measurement in a cross-cultural environment: Survey translation issues," *Qualitative Market Res., Int. J.*, vol. 3, no. 2, pp. 74–81, Jun. 2000, doi: [10.1108/13522750010322070](https://doi.org/10.1108/13522750010322070).
- [6] R. W. Brislin, "Comparative research methodology: Cross-cultural studies," *Int. J. Psychol.*, vol. 11, no. 3, pp. 215–229, Oct. 1976, doi: [10.1080/00207597608247359](https://doi.org/10.1080/00207597608247359).
- [7] H. Wang, H. Wu, Z. He, L. Huang, and K. W. Church, "Progress in machine translation," *Engineering*, vol. 18, pp. 143–153, Nov. 2022, doi: [10.1016/j.eng.2021.03.023](https://doi.org/10.1016/j.eng.2021.03.023).
- [8] J. Son and B. Kim, "Translation performance from the user's perspective of large language models and neural machine translation systems," *Information*, vol. 14, no. 10, p. 574, Oct. 2023, doi: [10.3390/info14100574](https://doi.org/10.3390/info14100574).
- [9] A. HENDY, M. Abdelrehim, A. Sharaf, V. Raunak, M. Gabr, H. Matsushita, Y. J. Kim, M. Afify, and H. H. Awadalla, "How good are GPT models at machine translation? A comprehensive evaluation," 2023, *arXiv:2302.09210*.
- [10] J. Yan, P. Yan, Y. Chen, J. Li, X. Zhu, and Y. Zhang, "GPT-4 vs. human translators: A comprehensive evaluation of translation quality across languages, domains, and expertise levels," 2024, *arXiv:2407.03658*.
- [11] OpenAI. *GPT-3.5 Turbo API Reference*. Accessed: Mar. 11, 2025. [Online]. Available: <https://platform.openai.com/docs/models/gpt-3.5-turbo>
- [12] OpenAI. *GPT-4o API Reference*. Accessed: Mar. 11, 2025. [Online]. Available: <https://platform.openai.com/docs/models/gpt-4o>
- [13] Google. *Gemini Models*. Accessed: Mar. 11, 2025. [Online]. Available: [https://ai.google.dev/gemini-api/docs/models/gemini?\\_gl=1\\*1qfema8\\*\\_up\\*MQ\\*\\_ga\\*MTIwMTQwNzQ3NS4xNzQxNjI0MTAx\\*\\_ga\\_P1DBVKWT6V\\*MTc0MTYyNDEwMS4xLjAuMTc0MTYyNDEwMS4wLjAuMzEyMjMwNjkx](https://ai.google.dev/gemini-api/docs/models/gemini?_gl=1*1qfema8*_up*MQ*_ga*MTIwMTQwNzQ3NS4xNzQxNjI0MTAx*_ga_P1DBVKWT6V*MTc0MTYyNDEwMS4xLjAuMTc0MTYyNDEwMS4wLjAuMzEyMjMwNjkx)
- [14] Anthropic. *All Models Overview*. Accessed: Mar. 11, 2025. [Online]. Available: <https://docs.anthropic.com/en/docs/about-claude/models/all-models#model-names>
- [15] DeepL. *DeepL API*. Accessed: Mar. 2025. [Online]. Available: <https://www.deepl.com/en/products/api>
- [16] A. Singhal, "Modern information retrieval: A brief overview," *IEEE Data Eng. Bull.*, vol. 24, no. 4, pp. 35–43, Jan. 2001.
- [17] United Nations. *170 Actions To Combat Climate Change*. Accessed: Mar. 11, 2025. [Online]. Available: <https://sites.ungeneva.org/170actions/climate/>
- [18] K. Yin Hong, L. Han, R. Batista-Navarro, and G. Nenadic, "CantonMT: Cantonese to English NMT platform with fine-tuned models using synthetic back-translation data," 2024, *arXiv:2403.11346*.
- [19] B. Liao, S. Khadivi, and S. Hewavitharana, "Back-translation for large-scale multilingual machine translation," 2021, *arXiv:2109.08712*.
- [20] S. Edunov, M. Ott, M. Ranzato, and M. Auli, "On the evaluation of machine translation systems trained with back-translation," 2019, *arXiv:1908.05204*.



**JI-BUM CHUNG** is currently an Associate Professor with the Department of Civil, Urban, Earth, and Environmental Engineering, Ulsan National Institute of Science and Technology (UNIST), Republic of Korea. His research interests include risk communication/management, disaster management, and AI applications.



**TAEHYUN KIM** is currently a Chief Research Fellow with the Division for Environmental Planning, Korea Environment Institute (KEI), Republic of Korea. His research interests include sustainable and resilient urban/environmental planning and design.

...