P-ISSN : 0853-1692 E-ISSN : 2579-5848
Buletin Al-Turas
website : bit.ly/buletinalturas

Mapping the Landscape of Machine Translation Error Research: A Bibliometric Review (1980-2023)

Sri Dewi Priwarti Siregar*
Universitas Pendidikan Indonesia
Bandung, Indonesia
sridewipriwarti@upi.edu

Mohamad Zaka Al Farisi
Universitas Pendidikan Indonesia
Bandung, Indonesia
zaka@upi.edu

Abdurrahman Do Muhammad Naser
Universitas Pendidikan Indonesia
Bandung, Indonesia
abdurrahmando30@upi.edu

Abstract

Purpose
Despite machine translation enhances communication process, it faces problems to guarantee accurate translation. Efforts to address these problems have been done by many researchers. The study aimed to reveal a comprehensive bibliometric review considering research trends, publishing contributions, and development in 1980-2023.

Method
This qualitative study made uses of a descriptive research design engaging some research on machine translation errors. This study employed bibliometric analysis to comprehensively examine the landscape of previous research on machine translation errors from 1980 to 2023. Based on a Scopus database, 138 publications were initially identified and subsequently refined to 98 articles that were directly relevant to the research topic. The bibliometric analysis highlighted the intriguing patterns and trends in this field.

Results/Findings
The findings revealed that the year 2021 witnessed the highest number of article publications, totaling 13 articles, indicating a growing interest in the topic. Moreover, a notable citation trend emerged in 2011, with 103 citations, signifying the significance and influence of research related to translation errors in machine translation. China notably emerged as the leading country to publish articles on this subject, with 23 publications and 28 collaborative links established with other countries. Among the 98 journals that published research in this domain, 45 of them were classified as Q1 journals, signifying their high impact and scholarly reputation.

Conclusion
The three main aspects comprising errors and human factors, exploration of machine translation, machine learning, and translation languages, and investigations within the field of computational linguistics collectively contributed to a deeper understanding of the complexities and challenges associated with translation errors in machine translation.

Keywords
Bibliometrics, Machine Translation, Machine Translation Errors

*) Corresponding Author

DOI: https://doi.org/10.15408/bat.v30i1.32655 *Print ISSN: 0853-1692 E-ISSN: 2579-5848*

Abstrak

Tujuan
Meskipun dapat meningkatkan proses komunikasi, mesin penerjemah belum dapat menjamin terjemahan yang akurat. Upaya untuk mengatasi masalah ini telah dilakukan banyak peneliti. Namun, tinjauan bibliometric yang komprehensif yang mempertimbangkan tren penelitian, kontribusi penerbitan, dan perkembangannya selama periode waktu tertentu belum dilakukan

Metode
Penelitian ini menggunakan analisis bibliometric untuk memeriksa secara komprehensif lanskap penelitian sebelumnya tentang kesalahan mesin penerjemah dari tahun 1980 hingga 2023. Dengan menggunakan basis data Scopus, sebanyak 138 publikasi diidentifikasi dan kemudian disaring menjadi 98 artikel yang relevan dengan topik penelitian. Analisis bibliometric menyoroti pola dan tren.

Hasil/Temuan
Penelitian ini menemukan tahun 2021 menjadi tahun dengan jumlah publikasi artikel tertinggi, sebanyak 13 artikel, yang mengindikasikan adanya peningkatan minat terhadap topik tersebut. Tren kutipan yang menonjil muncul pada tahun 2011, dengan total 103 kutipan, yang menandakan signifikansi dan pengaruh penelitian terkait kesalahan penerjemahan dengan mesin. China muncul sebagai negara terdepan denfab 23 publikasi dan 28 hubungan kolaboratif terjalin dengan negara lain. Di antara 98 jurnal yang menerbitkan penelitian di bidang ini, 45 diantaranya diklasifikasikan sebagai jurnal Q1.

Kesimpulan
Tiga aspek utama: 1) kontribusi kesalahan dan faktor manusia, 2) eksplorasi terjemahan mesin, pembelajaran mesin , dan bahasa terjemahan, serta 3) investigasi dalam bidang linguistic komputasi yang secara kolektif berkontribusi pada pemahaman yang lebih dalam tentang kompleksitas dan tantangan terkait kesalahan penerjemahan oleh mesin penerjemah.

Kata kunci
bibliometric, penerjemahan mesin, kesalahan penerjemahan mesin

الملخص

الهدف
على الرغم من أنها يمكن أن تحسن عملية الاتصال، إلا أن الترجمة الآلية لا يمكنها ضمان ترجمة دقيقة. وقد بذل العديد من الباحثين جهودًا للتغلب على هذه المشكلة. ومع ذلك، لم يتم إجراء مراجعة ببليومترية شاملة تأخذ في الاعتبار اتجاهات البحث ومساهمات النشر والتطورات خلال فترة زمنية محددة.

الطريقة
و يستخدم هذا البحث التحليل الببليومتري لإجراء فحص شامل لمشهد الأبحاث السابقة حول أخطاء الترجمة الآلية من عام 1980 إلى عام 2023. وباستخدام قاعدة بيانات سكوبس، تم تحديد إجمالي 138 منشورًا ثم تصفيتها إلى 98 مقالة ذات صلة بموضوع البحث. و يسلط التحليل الببليومتري الضوء على الأنماط والاتجاهات.

النتائج
و وجد هذا البحث أن عام 2021 كان العام الذي شهد أكبر عدد من المقالات المنشورة، بواقع 13 مقالا، مما يشير إلى زيادة الاهتمام بهذا الموضوع. ظهر اتجاه ملحوظ في الاستشهادات في عام 2011، حيث بلغ إجمالي الاستشهادات 103 استشهادات، مما يشير إلى أهمية البحث وتأثيره على أخطاء الترجمة الآلية. وبرزت الصين كدولة رائدة مع 23 منشورا و28 علاقة تعاونية مع دول أخرى. ومن بين 98 مجلة تنشر أبحاثًا في هذا المجال، تم تصنيف 45 منها ضمن مجلاتQ1 .

الخلاصة
ثلاثة جوانب رئيسية: 1) مساهمة الأخطاء والعوامل البشرية، 2) استكشاف الترجمة الآلية، والتعلم الآلي، ولغات الترجمة، و 3) التحقيقات في مجال اللغويات الحاسوبية التي تساهم بشكل جماعي في فهم أعمق للتعقيد والتحديات المرتبطة بأخطاء الترجمة بواسطة آلة الترجمة.

الكلمات الرئيسية
القياسات الببليومترية، الترجمة الآلية، أخطاء الترجمة الآلية

Print ISSN: 0853-1692 E-ISSN: 2579-5848 DOI: https://doi.org/10.15408/bat.v30i1.32655

INTRODUCTION

A translation machine is a computer system designed to automatically translate text or speech from one language to another. The significance of translation machines lies in their ability to facilitate cross-language communication, promote cultural exchange, and support international collaboration. In the contemporary era of globalization, machine translation has emerged as a crucial facilitator for cross-language communication (Araújo et al., 2020). Prominent machine translation programs, such as Google Translate and Bing Microsoft Translator, have witnessed a notable surge in popularity in recent times (Jufriadi et al., 2022). This translation technology empowers users to promptly and precisely convert text from one language to another (ElShiekh, 2012). Based on the given elucidation, one can infer that the utilization of machine translation assumes a crucial role, the use of machine translation plays a pivotal role as a crucial facilitator of global connectivity, providing easier access to information from various sources, and expediting communication processes in an era where cross-cultural and cross-linguistic interactions are increasingly prevalent. Additionally, the exploration of challenges and issues related to machine translation, such as translation accuracy, contextual understanding, and ongoing development efforts aimed at enhancing the quality of machine translations, can be considered.

Another perspective, as articulated by Wilks (2009), machine translation is characterized as a computer system designed to autonomously translate text from one language into another. Although translation engines serve numerous purposes in the current era of globalization, translation engines also possess certain drawbacks. Despite its considerable potential to enhance communication, machine translation encounters constraints in ensuring accurate and flawless translations (Almahasees, 2018). Errors in translation may appear in the form of flaws in sentence structure, improper selection of words, or the loss of meaning throughout the translation process.

The results of Muzaffar's (2019) research, which conducted a qualitative evaluation of the Machine Translation System from English to Urdu, namely PBSMT and NMT hosted by Google Translate, revealed that the scores on the assessment aspect in his research did not meet the expected standards. This study demonstrates the necessity for a comprehensive assessment of the studies that have been conducted in this sector to understand the degree and types of errors that exist in machine translation. Vidhayasai et al. (2015) used the object of research in the form of official websites of hotels and airlines to examine potential problems that can develop when using machine translation, such as Google Translate, to translate content in multiple languages. Allue In line with Allue's (2017) research that analyzed the translation of tourism texts and football match reports published online using machine translation, the most representative errors in terms of frequency have been classified at the lexicogrammatical, syntactic, pragmatic, and punctuation levels.

Study on translation andn errors in machine translation offers major advantages. This research can help to design more accurate and effective translation systems by studying the most prevalent translation errors in machine translation. A thorough understanding of translation errors has the potential to influence and expand the disciplines of translation linguistics (Battenburg & Malone, 1989), machine learning (Jervis et al., 2019), and computer science (Zhu et al., 2017). Furthermore, this research will benefit the translation industry, technology businesses (Chaudhari et al., 2020), and ordinary consumers of machine translation to increase translation quality and cross-language communication (Frederking & Taylor, 2004).

The study trend on machine translation errors has expanded dramatically during the past three decades. From 1980 to 2023, researchers and practitioners have been active in conducting studies related to the problem of machine translation errors developed in recent years by technological improvements and the growing importance of machine translation quality.

DOI: https://doi.org/10.15408/bat.v30i1.32655 Print ISSN: 0853-1692 E-ISSN: 2579-5848

Previously, several studies have been conducted to examine translation errors in machine translation. Among them have analyzed common types of translation errors (Abu-Ayyash, 2017; Almahasees & Mahmoud, 2022; Jufriadi et al., 2022), In the study, a qualitative approach was employed to identify common errors in machine translation. The qualitative approach utilized in the research aimed to delve into and comprehend the detailed aspects of these errors, thereby providing a more profound insight into the performance of machine translation. Some research also were conducted to comparing the performance of various machine translation systems (Araújo et al., 2020; Behera, 2019; Halimah, 2018; Muzaffar & Behera, 2019; Qian et al., 2019; Taai, 2011), or identifying factors that contribute to translation errors. The research were conducted through comparative analysis to disseminate the impact of machine translation usage in various usage contexts. The results of the analysis provide a deeper understanding regarding the effectiveness and efficiency of machine translation in meeting the needs of various communication situations. Based on several previous studies that employed qualitative approaches and comparative analyses to examine the utility of translation machines, there has not been research specifically investigating the mapping of the landscape of translation error research in machine translation using bibliometric analysis. This represents a novelty endeavor for conducting further research in this domain.

Despite efforts to address the problem of translation errors in machine translation, no thorough bibliometric review that looks holistically at research trends, publishing contributions, and developments over a longer period has been published. The bibliometric analysis provides a macroscopic image of a vast body of academic literature (van Nunen et al., 2018). Bibliometric tools can be used to evaluate the performance and research patterns of authors, journals, countries, and institutions and to detect and quantify patterns of collaboration between them (Li & Zhao, 2015).

This bibliometric analysis aims to comprehensively examine the research trends related to translation errors in machine translation from 1980 to 2023. The researcher analyzed articles spanning the years 1980 to 2023, as the keywords inputted into the Scopus database identified publications within the time frame of 1980 to 2023. Its primary objective is to map and review the literature on translation errors in machine translation, highlighting research trends, dominant foci, contributions from researchers, institutions, and countries, and summarizing advancements over time.

In this study, several research questions will be addressed, including: 1) What citation trends are associated with translation errors in machine translation? 2) How does the distribution of publications according to journals and quartile values relate to translation errors in machine translation? and 3) What is the research focus related to translation errors in machine translation?. This research will provide a detailed insight into the research landscape of translation errors in machine translation using a bibliometric approach, encompassing the development of research methods, the contributions of scholars, and the opportunities and difficulties that exist in this domain. As a result, this study is likely to lay a solid foundation for future research to improve machine translation quality and better understand the obstacles involved in developing more effective and accurate translation systems.

METHOD

Researchers used the Scopus database to look at data linked to "translation error in machine translation" because of its large number of indispliners. In this methodology section, there are several stages involved in collecting articles from the Scopus database to refine the data that will be subjected to analysis. The initial phase is identification, which is followed by the data screening process, feasibility, and ultimately the inclusion step (Moher, 2009).

The identifying method begins with the researcher entering keywords ("translation error" or "machine translation") into the Scopus database search. Researchers found and received publication information for 318 articles. The second step is to begin the screening procedure through screening. At this level, scholars choose to publish studies in Eng-

Print ISSN: 0853-1692 E-ISSN: 2579-5848 DOI: https://doi.org/10.15408/bat.v30i1.32655

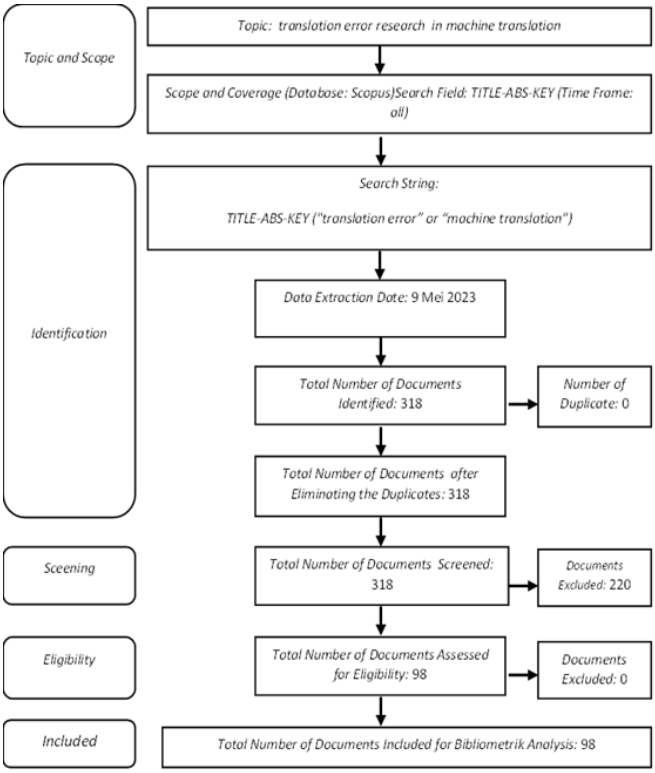

Figure 1. The procedure for data collection

lish in journals. According to the findings of the screening, 98 articles were obtained in line with the criteria.

This study's data was collected on May 9, 2023, during the inclusion stage. In this study, descriptive analysis was employed to examine the articles obtained concerning translation errors. The number of publications relating to translation errors in machine translation from 1980 to 2023 will be graphically displayed using Microsoft Excel Software.

The method section describes in detail how the study was conducted, including conceptual and operational definitions of the variables used in the study. It may consist of research design, description of participants, corpus or object being studied, instrumentation, data collecting procedures, and data analysis. The procedure for data collection will be elucidated in Figure 1.

Publication and citation patterns for translation errors in machine translation are examined by year from 1980 to 2023. Using Microsoft Excel software, the number of publications and average citations per year will be calculated and displayed. PoP software will be used to calculate the h-index and g-index of each article.

Researchers utilize Microsoft Excel software to depict the geographical dispersion of the data that has been collected. Furthermore, researchers employed the VOSviwer tool to examine international cooperation. The researchers displayed journal rankings based on quartiles using Microsoft Excel software. The data acquired from the Scopus database will be classified into (Q1), (Q2), (Q3), and (Q4) categories. According to the data collected, as many as 98 papers have been published in the ranking of the journals stated above.

FINDING AND DISCUSSION

Data collecting for publications on translation errors in machine translation has gone through numerous stages. During the data gathering procedure, 98 articles from 1980 to 2023 that fit the criteria were retrieved; the next stage will be descriptive bibliometric analysis. Publication patterns, citation trends, national and journal distribution, and re-

DOI: https://doi.org/10.15408/bat.v30i1.32655 Print ISSN: 0853-1692 E-ISSN: 2579-5848

search focus will all be covered in greater detail.

Findings

Publication Trends

Figure 2 depicts the publication patterns for translation errors in machine translation from 1980 to 2023. As illustrated in Figure 2, a total of 98 publications will be published in table form, organized by year of publication:

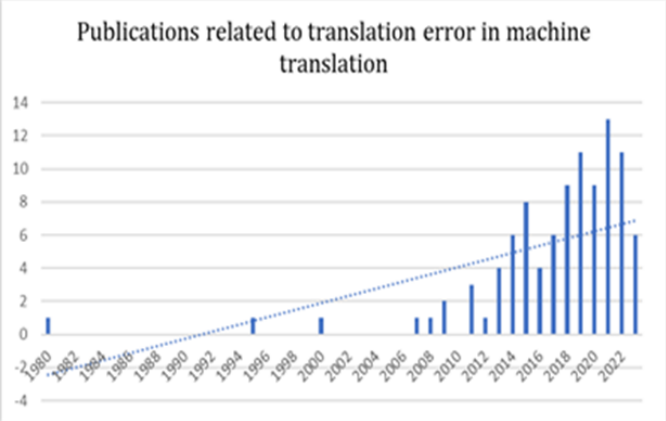

Figure 2. Number of Publications from 1980 to 2023

Based on the aforementioned graph, it is evident that the year 2021 stands out as a pivotal moment in the realm of research about translation errors in machine translation, with a remarkable surge in publications. A total of 13 documents were published during this period, marking a substantial increase in scholarly contributions. It is worth noting that the inception of this research can be traced back to 1980 when the first publication on this topic was documented in the Scopus database.

However, upon closer examination, an intriguing observation arises. The period from 1981 to 1994 appears to be a vulnerable phase, characterized by a lack of publications related to this particular research area. During these years, it seems that the attention and focus of scholars shifted away from studying translation errors in machine translation, resulting in a dearth of research contributions during that time.

Citations Trend

Table 1 will present trends in citations trend to translation errors in machine translation from 1980 to 2023. A total of 98 publications will be observed in total publications per year, grouped depending on the value of NCP, TC, C / P, C / CP, h-index, and g-index, as shown in Table 1.

Table 1 provides valuable insights into the landscape of publications and citations related to translation errors in machine translation. Notably, the years 2021 and 2019 emerge as significant milestones, as they recorded the highest number of referenced publications (NCP), with both years attaining an NCP value of 11. However, when considering the impact of the research, it becomes apparent that 2011 holds greater scientific influence, with a remarkable 103 citations, despite having fewer publications compared to 2021.

An intriguing pattern that Table 1 reveals is the presence of certain years with either no citations or even no publications, particularly within the period from 1981 to 1994. This gap in research activity during those years raises questions about the factors that may have contributed to this lack of scholarly contributions, such as shifting research priorities or perhaps limited resources and technological constraints.

Furthermore, delving deeper into the analysis, it is worth highlighting that from 1996 to 2006, there was a notable decline in publications, with only one recorded during that

Print ISSN: 0853-1692 E-ISSN: 2579-5848 DOI: https://doi.org/10.15408/bat.v30i1.32655

Table 1. Citation Analysis of Publications

Year	T	NCP	TC	C/P	C/CP	h	g
2023	6	-	-	-	-	-	-
2022	11	4	9	0,82	2,25	2	2
2021	13	11	31	2,38	2,81	3	4
2020	9	7	29	3,22	4,14	3	4
2019	11	11	77	7,00	7	5	8
2018	9	9	79	8,78	8,78	6	8
2017	6	5	88	14,67	17,6	3	6
2016	4	3	18	4,50	6	3	4
2015	8	8	86	10,75	10,75	5	8
2014	6	6	68	11,33	11,33	5	6
2013	4	4	64	16,00	16	3	4
2012	1	1	22	22,00	22	1	1
2011	3	3	103	34,33	34,33	2	3
2010	-	-	-	-	-	-	-
2009	2	2	32	16,00	16	2	2
2008	1	1	12	12,00	12	1	1
2007	1	1	81	81,00	81	1	1
2006	-	-	-	-	-	-	-
2005	-	-	-	-	-	-	-
2004	-	-	-	-	-	-	-
2003	-	-	-	-	-	-	-
2002	-	-	-	-	-	-	-
2001	-	-	-	-	-	-	-
2000	1	1	10	10,00	10	1	1
1999	-	-	-	-	-	-	-
1998	-	-	-	-	-	-	-
1997	-	-	-	-	-	-	-
1996	-	-	-	-	-	-	-
1995	1	1	6	6,00	6	1	1
1994	-	-	-	-	-	-	-
1993	-	-	-	-	-	-	-
1992	-	-	-	-	-	-	-
1991	-	-	-	-	-	-	-
1990	-	-	-	-	-	-	-
1989	-	-	-	-	-	-	-
1988	-	-	-	-	-	-	-
1987	-	-	-	-	-	-	-
1986	-	-	-	-	-	-	-
1985	-	-	-	-	-	-	-
1984	-	-	-	-	-	-	-
1983	-	-	-	-	-	-	-
1982	-	-	-	-	-	-	-
1981	-	-	-	-	-	-	-
1980	1	1	20	20,00	20	1	1

Notes. TP=total of publication, NCP=number of cited publication, TC=total citations, C/P=average citations per publication, C/CP=average citations per cited publications, h=h-index, g=g-index

period. This decrease in research output during those years may indicate a temporary dip in interest or engagement within the academic community regarding translation errors in machine translation.

Moving beyond publications and citations, the h-index and g-index values offer additional insights into the impact of specific research contributions. Notably, 2018 stands out as the year with the highest h-index and g-index values, registering 6 and 8, respectively. This suggests that the publication(s) in 2018 had a substantial influence on the field, contributing significantly to the body of knowledge on translation errors in machine translation.

Upon closer inspection of table 1, it is evident that in the year 2018, a cumulative total of nine publications were cited, accumulating a noteworthy 79 citations. This signifies a commendable average citation per publication ratio of nine, thereby emphasizing the importance and influence of the research undertaken in that particular year.

DOI: https://doi.org/10.15408/bat.v30i1.32655 Print ISSN: 0853-1692 E-ISSN: 2579-5848

In summary, Table 1 sheds light on the dynamics of publications and citations in the field of translation errors in machine translation. It highlights noteworthy trends, such as the varying scientific impact across different years, the presence of gaps in research activity during certain periods, and the influence of specific publications on the overall body of knowledge. These insights contribute to a deeper understanding of the evolution and significance of research in this domain.

Geographical Distribution and Relations Between Countries

Figure 3 depicts the geographical spread of the publication's nation of origin. 36 countries publish machine translation errors. The distribution of the nation of origin of the most publications (more than three) may be seen in the figure 3.

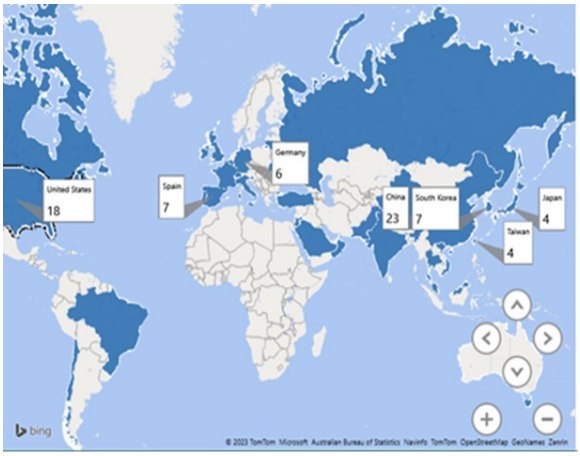

Figure 3. Geographical Distribution of Publications

Based on Figure 3, it becomes evident that China holds the highest influence in the field of machine translation research, specifically concerning translation mistakes. The Chinese government has made a significant contribution to this sector, releasing a substantial number of 23 documents about translation errors in machine translation. This highlights the country's strong presence and active engagement in advancing knowledge and understanding in this area.

Following China, the United States emerges as the second most influential country in this study, with a notable publication count of 18 documents related to translation errors in machine translation. The United States contribution underscores its significant role in shaping the discourse surrounding translation mistakes in the context of machine translation.

In addition to China and the United States, South Korea, and Spain rank as the third most influential countries in this study. Both countries have demonstrated their commitment to the field, with each publishing a notable count of 7 documents on translation errors in machine translation. This showcases their contributions and emphasizes the global nature of research endeavors in this domain.

The findings depicted in Figure 3 offer valuable insights into the leading countries that have significantly contributed to the literature concerning translation errors in machine translation. Notably, China, the United States, South Korea, and Spain have emerged as pivotal players, demonstrating their active engagement and substantial contributions to advancing knowledge and understanding in this field. Through their research endeavors, these countries have provided invaluable insights, propelled advancements, and played a crucial role in shaping the discourse surrounding translation errors in machine translation.

Country collaboration relations

Figure 4 illustrates the trend of country collaboration in the context of machine transla-

Print ISSN: 0853-1692 E-ISSN: 2579-5848 DOI: https://doi.org/10.15408/bat.v30i1.32655

tion research. It's important to note that researchers did not employ a specific criterion to evaluate these interstate linkages. Consequently, even if a country lacks cooperation connections with others, it may still be displayed if it has at least one document associated with it. As a result, Figure 4 includes a total of 36 countries, each contributing to the broader landscape of collaborative efforts in addressing translation mistakes in machine translation.

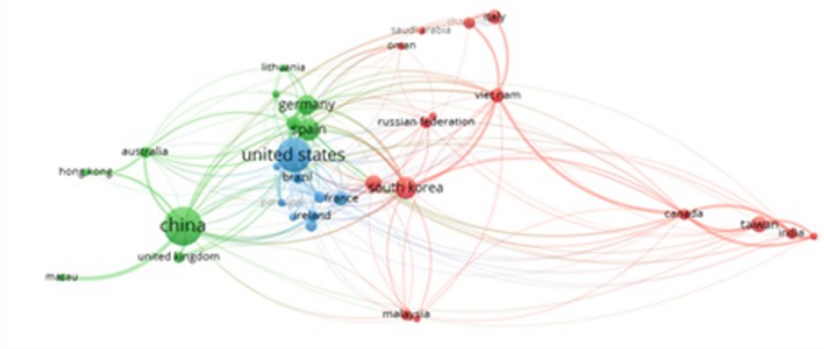

Figure 4: International Relations Publications

Figure 4 offers a compelling visual overview of the collaborative landscape within the realm of translation errors in machine translation. Particularly striking is the prominence of China, the United States, Spain, Germany, and France, which are depicted with larger circle diameters compared to other countries. This suggests that these nations have cultivated extensive and robust collaborative networks with numerous other countries, indicating a strong global presence and active engagement in collaborative research endeavors. This visual representation underscores the interconnectedness and global nature of research efforts aimed at understanding and addressing translation errors in machine translation.

Further insights from the VOSviewer presentation reveal the extent of collaboration between these countries. China, for instance, has established collaboration relations with a total of 28 nations, underscoring its strong network and global engagement in addressing translation mistakes in machine translation. Similarly, Spain has developed cooperation links with 27 countries, further demonstrating its commitment to international collaboration in this domain. The United States, known for its influential role in research, has established cooperative relations with 24 countries, emphasizing its dedication to fostering collaborative efforts in tackling translation errors in machine translation.

The collaboration networks depicted in Figure 4, supported by the VOSviewer analysis, underscore the interconnected and cooperative nature of research endeavors aimed at addressing translation mistakes in machine translation. These findings reveal that countries such as China, the United States, Spain, Germany, and France are actively participating in collaborative efforts to bolster understanding, facilitate knowledge exchange, and collectively propel advancements in the field. By fostering such collaborative networks, researchers from around the world can leverage diverse perspectives, resources, and expertise to tackle complex challenges and drive innovation in machine translation research.

Discussion

What citation trends are related to translation errors in machine translation?

The study of translation errors in machine translation has become a crucial linguistic subject to address. This phenomenon not only reflects the complexity of machine translation technology but also carries profound implications for the quality of translation outcomes. Table 1 illustrates the citation trends of translator errors in machine translation from 1980 to 2023. A total of 98 articles were analyzed concerning various aspects of

DOI: https://doi.org/10.15408/bat.v30i1.32655 Print ISSN: 0853-1692 E-ISSN: 2579-5848

translator errors in machine translation. Among these, 36 articles addressed both topics which are translation errors and machine translation. Additionally, 20 articles delved into issues related to translation errors separately, while 42 specifically focused on machine translation. In this context, articles related to translation machines still constitute a substantial body of research, spanning from 1980 to 2023. In Figure 2, it is evident that research pertaining to this topic experienced fluctuating growth from 1980 to 2023. The sharpest increase occurred from 2020 to 2021, followed by a subsequent decline in the succeeding year. On the other hand, the total citation (TC) in table 1 shows the citation trend for translation errors in machine translation from 1980 to 2023. Table 1 shows that articles in 2011 were cited 103 times by having an h-index and a g-index in that year with values of h-index = 2 and g-index = 3. This demonstrates that research conducted in 2011 had a significant impact on studies about translation errors in machine translation. Table 3 shows the three articles published in 2011.

Table 3. Articles published in 2011

No	Author (year)	Title	Sources	Citation
1	(Condon et al., 2011)	Machine Translation Errors: English and Iraqi Arabic	ACM Transactions on Asian Language Information Processing	1
2	(Popovi´ & Ney, 2011)	Towards Automatic Error Analysis of Machine Translation Output	Computational Linguistics	59
3	(Kirchhoff et al., 2011)	Application of statistical machine translation to public health information: a feasibility study	Journal of the American Medical Informatics Association	43

According to table 3, studies by Condon et al. (2011) and Popovi´ & Ney (2011) have been mentioned 1 and 59 times, respectively. This demonstrates that the study was the most mentioned in 2011.

The study conducted by Condon analyzed errors in machine translations of English-Iraqi Arabic dialogues using methods developed for the Human Translation Error Rate measure (HTER). Human annotations were employed to refine the Translation Error Rate (TER) annotations, and approximately 100 translations into each language from four systems were examined. The findings revealed high frequencies of pronoun errors and copula-related errors in translations to English, while translations to Iraqi Arabic exhibited frequent errors in subject/person inflection and closed-word classes. Word order errors were similarly common in both translation directions, with low frequencies of polarity errors. Structural differences between the languages were identified as predictive factors for many errors, along with challenges related to the insertion of lexemes. Resolving problems associated with deictic elements, such as pronouns, may require knowledge of the discourse context.

The study conducted by Popovi & Ney (2011) regarding towards automatic error analysis of machine translation output posits that the assessment and scrutiny of errors in machine translation are crucial yet challenging areas of study. This paper presents a framework for automatic error analysis based on inaccurate word identification utilizing the Word Error Rate (WER) and Position Independent Word Error Rate (PER) approaches. The outcomes of automated error analysis substantially resemble those of human error analysis. The proposed metrics could differentiate between different translation systems and translation systems.

How does the distribution of publications by journal and quartile value relate to translation errors in machine translation?

The distribution of publication rates based on the most viewed quartile value (Q) reveals that out of a total of 98 publications, 45 were published in Q1. Notably, Table 4 be-

Print ISSN: 0853-1692 E-ISSN: 2579-5848 DOI: https://doi.org/10.15408/bat.v30i1.32655

low showcases the researchers associated with the publications that have contributed the most articles relevant to the study of translation errors in machine translation from 1980 to 2023. For further details and insights, please refer to Table 4.

Table 4. Journals with the most articles

Journal Name	Number of articles	Quartile Value
Machine Translation	10	Q1
Applied Sciences (Switzerland)	3	Q2
New Generation Computing	2	Q3

Table 4 offers valuable insights into the publication landscape concerning translation errors in machine translation. Particularly noteworthy is the prominence of the journal "Machine Translation," which has published a significant number of ten articles dedicated to this topic, indicating its leading role in advancing research in this area. Additionally, journals such as Applied Sciences (Switzerland) and New Generation Computing have also made notable contributions, each publishing three articles on this subject. These findings underscore the significance of these journals in furthering our understanding of translation errors in machine translation and highlight potential avenues for researchers to explore and disseminate their work effectively.

The journals listed in Table 4 present promising opportunities as preferred publication outlets for researchers investigating translation errors in machine translation. Leveraging the esteemed reputation and wide readership of these journals, researchers can efficiently communicate their findings, stimulate dialogue, and actively participate in the scholarly discourse surrounding the complexities of translation errors in machine translation. By choosing these reputable journals, researchers can maximize the impact and visibility of their work, contributing significantly to the advancement of knowledge in this field and fostering collaboration within the academic community.

What is the focus of research related to translation errors in machine translation?

The focus of translation error study in machine translation is divided into three main parts: 1) error and human; 2) machine translation, machine learning, and translation (languages); and 3) computational linguistics. The focus of research arising from data related to "translation error in machine translation" can be explained as the understanding and identification of translation errors by humans, the development of technology and machine learning in machine translation systems, and the application of computational computing and computational linguistics in machine translation analysis and improvement.

Focus on understanding and identifying translation errors by humans, such as Herrmann-Werner et al.'s (2021) research, discusses communication between patients and doctors via LTA (language translation apps) as mechanical alternatives, the results of his research reveal some concerns about translation errors that can jeopardize diagnostic decisions and reduce empathy. This study, on the other hand, indicates the viability of employing LTA in undergraduate medical courses. Furthermore, examples of research with a focus on technology development and machine learning in machine translation systems conducted by Loock & Léchauguette (2021) show that applied language students who use machine translation fail to thoroughly correct translation errors. Accuracy errors are easier to spot than fluency errors. Students also tend to ignore fluency mistakes, maybe due to an overreliance on machine translation or a lack of expertise in the target language. To provide students with the abilities required for the critical and professional use of machine translation tools, machine translation teaching methodologies must be enhanced.

DOI: https://doi.org/10.15408/bat.v30i1.32655 Print ISSN: 0853-1692 E-ISSN: 2579-5848

CONCLUSION

In conclusion, the analysis of the findings and discussions leads to several key observations. Firstly, it is evident that 2021 witnessed the publication of 13 documents, indicating a notable interest and research output in the field of machine translation errors during that year. Furthermore, the year 2011 stands out as a significant period in terms of citation impact, with a total of 103 citations attributed to research in this area. On the other hand, ninety-eight articles were scrutinized to explore diverse facets of translator errors in machine translation. Of these, 36 articles comprehensively examined both translation errors and machine translation, while 20 articles specifically delved into the nuances of translation errors. Furthermore, 42 articles concentrated specifically on the intricacies of machine translation.

In examining the influential countries, China emerges as the most prominent player, with 23 publications and 28 collaborative links established with other nations. This highlights China's strong presence and leadership in the subject of machine translation errors. The publication analysis reveals a total of 98 documents related to machine translation errors, with 45 of them categorized as Q1 journals. This indicates a significant proportion of high-impact research within this domain. The research focuses on machine translation error and encompasses three main areas: 1) the study of errors and human factors, 2) investigations into machine translation, machine learning, and translation languages, and 3) advancements in computational linguistics. Additionally, research on "translation errors in machine translation" encompasses the understanding and identification of errors committed by humans, the advancements in technology and machine learning within translation systems, and the utilization of computational and linguistic approaches for analysis and improvement. This underscore concerns regarding errors in language translation applications that impact diagnostic decisions and empathy. It also highlights a tendency among applied language learners to overlook proficiency errors, possibly due to an excessive reliance on machine translation or a lack of proficiency in the target language.

These findings collectively contribute to a comprehensive understanding of the current landscape of research on machine translation errors, emphasizing the importance of addressing errors and improving translation accuracy. By shedding light on prevalent trends and influential factors, this research not only enriches academic discourse but also offers practical implications for stakeholders across various domains. Researchers can leverage these insights to refine existing methodologies and develop innovative approaches in machine translation research. Practitioners, including developers of machine translation systems and language professionals, can utilize these findings to enhance the quality and reliability of translation tools and services. Furthermore, policymakers and educators can benefit from these insights to inform decisions regarding language policy, curriculum development, and the integration of technology in language learning. In sum, the identified trends and influential factors provide valuable insights for researchers, practitioners, and stakeholders in the field of machine translation and related disciplines, ultimately contributing to advancements in language technology and cross-cultural communication.

LIMITATION

The study has several limitations that warrant acknowledgment. Firstly, the data analyzed originates solely from the Scopus database, thus potentially overlooking relevant studies indexed in other databases such as Google Scholar and others. Secondly, the scope of this study is limited to examining translation errors within machine translation, thereby excluding exploration into additional translation fields which may offer valuable insights. Lastly, it's important to note that the data collection concluded on May 9, 2023, which may not encompass subsequent research developments, potentially leading to discrepancies between the findings and the most current research landscape. These limitations underscore the need for future studies to consider broader data sources, explore diverse translation domains, and incorporate more recent research findings for a more comprehensive understanding of translation errors in machine translation.

Print ISSN: 0853-1692 E-ISSN: 2579-5848 DOI: https://doi.org/10.15408/bat.v30i1.32655

REFERENCES

Abdel Azim ElShiekh, A. (2012). Google Translate Service: Transfer of Meaning, Distortion or Simply a New Creation? An Investigation into the Translation Process & Problems at Google. *English Language and Literature Studies*, 2(1), 56–68. https://doi.org/10.5539/ells.v2n1p56

Abu-Ayyash, E. A. S. (2017). Errors and non-errors in English-Arabic machine translation of gender-bound constructs in technical texts. *Procedia Computer Science*, 117, 73–80. https://doi.org/10.1016/j.procs.2017.10.095

Allue, B. R. (2017). The Reliability and Limitations of Google Translate: A Bilingual, Bidirectional and Genre-Based Evaluation. *Entreculturas*, 9(febrero), 67–80.

Almahasees, Z. M. (2018). Assessment of Google and Microsoft Bing Translation of Journalistic Texts. *International Journal of Languages, Literature and Linguistics*, *4* (3), 231–235. https://doi.org/10.18178/ijlll.2018.4.3.178

Almahasees, Z., & Mahmoud, S. (2022). Evaluation of Google Image Translate in Rendering Arabic Signage into English. *World Journal of English Language*, 12(1), 185–197. https://doi.org/10.5430/wjel.v12n1p185

Araújo, M., Pereira, A., & Benevenuto, F. (2020). A comparative study of machine translation for multilingual sentence-level sentiment analysis. *Information Sciences*, 512, 1078–1102. https://doi.org/10.1016/j.ins.2019.10.031

Battenburg, J., & Malone, J. L. (1989). The Science of Linguistics in the Art of Translation: Some Tools from Linguistics for the Analysis and Practice of Translation. In *Language* (Vol. 65, Issue 2). https://doi.org/10.2307/415364

Behera, P. (2019). *A Qualitative Evaluation of Google ' s Translate : A Comparative Analysis of PBSMT and NMT Systems*. *18*(49042), 180–192.

Chaudhari, B. S., Zennaro, M., & Borkar, S. (2020). LPWAN technologies: Emerging application characteristics, requirements, and design considerations. *Future Internet*, *12*(3). https://doi.org/10.3390/fi12030046

Condon, S., Parvaz, D., Aberdeen, J., Doran, C., Freeman, A., & Awad, M. (2011). Machine translation errors: English and Iraqi Arabic. *ACM Transactions on Asian Language Information Processing*, *10*(1). https://doi.org/10.1145/1929908.1929910

Frederking, R. E., & Taylor, K. B. (2004). Machine Translation: From Real Users to Research. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 1317). Springer.

Halimah. (2018). Comparison Of Human Translation With Google Translation Of Imperative Sentences In Procedures Text. *Bahtera: Jurnal Pendidikan Bahasa Dan Sastra*, *17*(1), 1–23.

Herrmann-Werner, A., Loda, T., Zipfel, S., Holderried, M., Holderried, F., & Erschens, R. (2021). Evaluation of a language translation app in an undergraduate medical communication course: Proof-of-concept and usability study. *JMIR MHealth and UHealth*, *9*(12). https://doi.org/10.2196/31559

Jervis, A. J., Carbonell, P., Vinaixa, M., Dunstan, M. S., Hollywood, K. A., Robinson, C. J., Rattray, N. J. W., Yan, C., Swainston, N., Currin, A., Sung, R., Toogood, H., Taylor, S., Faulon, J. L., Breitling, R., Takano, E., & Scrutton, N. S. (2019). Machine Learning of Designed Translational Control Allows Predictive Pathway Optimization

DOI: https://doi.org/10.15408/bat.v30i1.32655 Print ISSN: 0853-1692 E-ISSN: 2579-5848

in Escherichia coli. *ACS Synthetic Biology*, *8*(1), 127–136. https://doi.org/10.1021/acssynbio.8b00398

Jufriadi, Amalia Asokawati, & Thayyib, M. (2022). The Error Analysis of Google Translate and Bing Translator in Translating Indonesian Folklore. *FOSTER: Journal of English Language Teaching*, 3(2), 69–79. https://doi.org/10.24256/foster-jelt.v3i2.89

Kirchhoff, K., Turner, A. M., Axelrod, A., & Saavedra, F. (2011). Application of statistical machine translation to public health information: A feasibility study. *Journal of the American Medical Informatics Association*, 18(4), 473–478. https://doi.org/10.1136/amiajnl-2011-000176

Li, W., & Zhao, Y. (2015). Bibliometric analysis of global environmental assessment research in a 20-year period. *Environmental Impact Assessment Review*, *50*, 158–166. https://doi.org/10.1016/j.eiar.2014.09.012

Loock, R., & Léchauguette, S. (2021). Machine translation literacy and undergraduate students in applied languages: Report on an exploratory study. *Revista Tradumatica*, *19*, 204–225. https://doi.org/10.5565/rev/tradumatica.281

Moher, D. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement (Chinese edition). *Journal of Chinese Integrative Medicine*, *7* (9), 889–896. https://doi.org/10.3736/jcim20090918

Muzaffar, S., & Behera, P. (2019). *A Qualitative Evaluation of Google ' s Translate : A Comparative Analysis of PBSMT and NMT Systems*. *October*, 180–192.

Popovi´, M., & Ney, H. (2011). Towards Automatic Error Analysis of Machine Translation Output. *Computational Linguistics*, *70*(8), 4943. https://doi.org/10.1162/COLI

Qian, M., Liu, J., Li, C., & Pals, L. (2019). A Comparative Study of English-Chinese Translations of Court Texts by Machine and Human Translators and the Word2Vec Based Similarity Measure's Ability To Gauge Human Evaluation Biases. *Proceedings of Machine Translation Summit XVII Volume 2: Translator, Project and User Tracks*, *2*, 95–100.

Taai, L. Al. (2011). *A Comparison of Arabic Literature Translation into English and Swedish A Comparison of Arabic Literature Translation into English and Swedish*.

van Nunen, K., Li, J., Reniers, G., & Ponnet, K. (2018). Bibliometric analysis of safety culture research. *Safety Science*, 108(November 2016), 248–258. https://doi.org/10.1016/j.ssci.2017.08.011

Vidhayasai, T., Keyuravong, S., & Bunsom, T. (2015). Investigating the Use of Google Translate in "Terms and Conditions" in an Airline's Official Website: Errors and Implications. *Pasaa*, 49(January-June), 137–169.

Wilks, Y. (2009). *Machine Translation*. Springer.

Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. *Proceedings of the IEEE International Conference on Computer Vision*, *2017-Octob*, 2242–2251. https://doi.org/10.1109/ICCV.2017.244

Received (06-06-2023) *Accepted (29-03-2024)* *Published (31-03-2024)*

Print ISSN: 0853-1692 E-ISSN: 2579-5848 DOI: https://doi.org/10.15408/bat.v30i1.32655