



Nunes Vieira, L., O'Hagan, M., & O'Sullivan, C. (2020). Understanding the societal impacts of machine translation: a critical review of the literature on medical and legal use cases. *Information Communication and Society*. <https://doi.org/10.1080/1369118X.2020.1776370>

Publisher's PDF, also known as Version of record

License (if available):
CC BY

Link to published version (if available):
[10.1080/1369118X.2020.1776370](https://doi.org/10.1080/1369118X.2020.1776370)

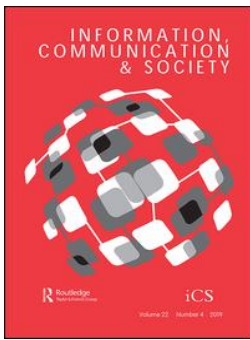
[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the final published version of the article (version of record). It first appeared online via Taylor & Francis at <https://www.tandfonline.com/doi/full/10.1080/1369118X.2020.1776370> . Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: <http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>



Understanding the societal impacts of machine translation: a critical review of the literature on medical and legal use cases

Lucas Nunes Vieira , Minako O'Hagan & Carol O'Sullivan

To cite this article: Lucas Nunes Vieira , Minako O'Hagan & Carol O'Sullivan (2020): Understanding the societal impacts of machine translation: a critical review of the literature on medical and legal use cases, Information, Communication & Society, DOI: [10.1080/1369118X.2020.1776370](https://doi.org/10.1080/1369118X.2020.1776370)

To link to this article: <https://doi.org/10.1080/1369118X.2020.1776370>



© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 16 Jun 2020.



Submit your article to this journal [↗](#)




View related articles [↗](#)



View Crossmark data [↗](#)

Understanding the societal impacts of machine translation: a critical review of the literature on medical and legal use cases

Lucas Nunes Vieira ^a, Minako O'Hagan^b and Carol O'Sullivan^a

^aSchool of Modern Languages, University of Bristol, Bristol, UK; ^bSchool of Cultures, Languages and Linguistics, The University of Auckland, Auckland, New Zealand

ABSTRACT

The ready availability of machine translation (MT) systems such as Google Translate has profoundly changed how society engages with multilingual communication practices. In addition to private use situations, this technology is now used to overcome language barriers in high-risk settings such as hospitals and courts. MT errors pose serious risks in environments like these, but there is little understanding of the nature of these risks and of the wider implications of using this technology. This article is the first structured study of the consequences of uninformed MT use in healthcare and law. Based on a critical literature review, the article presents a qualitative meta-analysis of official documents and published research on the use of MT in these two fields. Its findings prompt calls for action in three areas. First, the review shows that research on MT use in healthcare and law can often disregard the complexities of language and language translation. The article calls for cross-disciplinary research to address this gap by ensuring that a growing body of relevant knowledge in translation studies informs research conducted within the medical and legal sectors. Second, the review highlights a broad societal need for higher levels of awareness of the specific strengths and, crucially, of the limitations of MT. Finally, the article concludes that MT technology can in its current state exacerbate social inequalities and put certain communities of users at greater risk. We highlight this as a persistent issue that merits further attention from researchers and policymakers.

ARTICLE HISTORY

Received 6 June 2019
Accepted 15 May 2020

KEYWORDS

Machine translation; translation in legal and healthcare settings; multilingual communication; computer-mediated communication; translation-mediated communication

Introduction

In the wake of globalisation and a diversifying online population (Wu & Taneja, 2016), communicating across languages, whether in professional or personal contexts, is now a common experience of everyday life. Multilingual communication needs are increasingly met by automatic translation systems, also known as machine translation (MT). These systems have been in development since the 1950s (Somers, 2007). They allow users to obtain virtually instant translations of information they wish to consume or convey in a different

CONTACT Lucas Nunes Vieira  l.nunesvieira@bristol.ac.uk  School of Modern Languages, University of Bristol, Room G97, 19 Woodland Road, Bristol BS8 1TE, UK

© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

language. This can typically be done at no up-front cost to the user.¹ Google Translate, which is currently one of the most popular MT systems freely available online, is used to translate thirty trillion sentences annually across over 100 languages (Kuczmariski, 2018). Over past decades, MT design evolved from rule-based to data-driven systems. The current state of the art in MT technology is neural MT. This is a machine learning methodology that produces highly fluent and idiomatic translations, which may come at the price of lower levels of accuracy compared to previous systems (Castilho et al., 2017). The fact that even advanced MT systems have significant weaknesses highlights the importance of understanding the potential and the limitations of this rapidly evolving technology.

Particularly in high-stakes settings, misuse of MT can have serious consequences. In one recent case, evidence was dismissed in court because consent to perform a police search had been obtained with Google Translate, which raised concerns about the consent's validity (Grosdidier, 2019). In a medical setting, an evaluation of errors that could be caused by MT revealed that the sentence 'your child is fitting' would in one case have been translated to Swahili as 'your child is dead' (Patil & Davies, 2014). Despite the risks of using MT in contexts like these, research on the implications of the widespread and potentially uninformed use of this technology remains sparse. MT use in 'everyday' communication (see Nurminen, 2018) is an emerging research area and there is also a growing body of research on public service interpreting (also known as community interpreting) aimed at making health and legal services accessible across languages (e.g., Angelelli, 2008). Research in the medical (e.g., Das et al., 2019) and legal fields (e.g., Yates, 2006), however, is often undertaken in parallel with, and without the full benefit of, related research in translation studies (e.g., Braun, 2019; Kenny, 2019). This prompted us to investigate the literature on MT use cases from these areas.

This article presents, to the best of our knowledge, the first structured literature review of the implications of misusing MT as a communication tool in medical and legal settings. Our aim is to provide a qualitative meta-analysis of MT's risks and potential in relation to medical and legal communication. Ultimately, we hope to improve the understanding of MT's risks in these fields and stimulate cross-disciplinary research on the societal impacts of MT. The review examines (1) how MT is currently perceived and used in medical and legal settings and (2) how it affects communication in these two areas. We focus on the use of unedited or 'raw' MT output, which has also been called 'Fully Automatic Useful Translation' (see Nurminen, 2018). We examine how MT is used in this way for assimilating and disseminating information as well as for synchronous or asynchronous bidirectional exchanges. Since machine translation (text-to-text) and machine interpreting (MI) (speech-to-speech) both rely on MT systems as the core technology to convert text from one language into another, MT and MI can often intertwine in MT-mediated communication.² Although human translating and interpreting do involve different factors and skillsets, limiting the analysis to just one of these tasks seemed unnecessarily restrictive given the article's broader focus on the technology's use consequences and the fact that outside of translation studies these tasks may not be distinguished.

We structure the remainder of the article as follows. In the next section, we describe the review methodology. We then present the results of our analysis of the two fields selected for investigation. We subsequently provide a discussion of the results and present conclusions and recommendations for future research on MT-mediated communication.

Methodology

Literature search

We performed searches for English-language³ records on Google Scholar and on databases with a more specific disciplinary focus, namely PubMed for healthcare, and HeinOnline and Westlaw for law. We drew on a review of MT development for healthcare (Dew et al., 2018) to fine-tune the keywords used for the searches. All searches were based on the following baseline expression: ‘machine translation’ OR ‘automatic translation’ OR ‘automated translation’ OR ‘online translation’ OR ‘google translate’. On Google Scholar, we combined this expression with other relevant terms in our search for academic publications.⁴ In the Google Scholar search for research on MT use in legal settings, we used: (‘machine translation’ OR ‘automatic translation’ OR ‘automated translation’ OR ‘online translation’ OR ‘google translate’) AND (legal OR law OR lawyer OR judge OR court). For the healthcare Google Scholar search, we used (‘machine translation’ OR ‘automatic translation’ OR ‘automated translation’ OR ‘online translation’ OR ‘google translate’) AND (health OR clinical OR nursing OR medicine OR doctor OR patient). On the discipline-specific platforms and in our case law search on Google Scholar, we used the baseline expression alone. We drew on both HeinOnline and Westlaw for the law search to offset a US focus noted in results returned by HeinOnline. This also addressed the fact that only US case law is currently available on Google Scholar.

The review is limited to records published from the year 2000 onwards, a date filter that was applied to all searches. This largely coincides with the period when MT crossed the one-million-user threshold and gained traction as a widespread freely available online tool (Yang & Lange, 2003, p. 194). March 2019, when we started the review, was the cut-off point for inclusion of any records.

Except for the case law search, results returned by Google Scholar ran into the tens of thousands, so we limited the screening of these results to the first 200 records for each search, which were ranked according to Google Scholar’s ‘relevance’ criteria. By the 100th record the entries largely failed the criteria for inclusion in the study (see below), so 200 records seemed like a conservative threshold. The search results were initially screened for relevance to the topic and to the aims of the review based on the abstracts or text passages containing any of the MT-related terms.⁵ Items that were pre-selected at this initial screening step were considered at the analysis stage when a further subset of records was ultimately retained. The criteria for including a record in the analysis were as follows:

- (1) The records had to contain substantiated evidence either of how MT was being used in a certain context or of how it was perceived or assessed.
- (2) When multiple sources provided similar information, only the most recent or detailed record was retained.
- (3) If a literature review was available, we considered the review itself as an entry without providing a repeated analysis of its internal sources.⁶
- (4) Where reference lists or the authors’ prior knowledge led to relevant records that had not been returned by the searches, these records were manually included provided they met the other criteria.

We note that we did not establish peer-reviewed status as an inclusion criterion. This is because, as the analysis below will show, evidence of MT use implications is often found in professional association publications, official letters and other documents, such as case law, that would not be expected to undergo an academic peer review. Given the second criterion above, it should also be noted that the purpose of the criteria was to ground the analysis in a representative set of evidence of MT use implications for the selected fields. We therefore do not claim to provide an exhaustive bibliography for this subject.

Taking the criteria above into account, the meta-analysis was based on 45 sources, of which 11 were manually included. A flow chart of the review process is presented below in Figure 1.

Analytical approach

Our method for analysing the content is informed by MT research and by healthcare and legal (public service) interpreting research in translation studies, which is currently largely

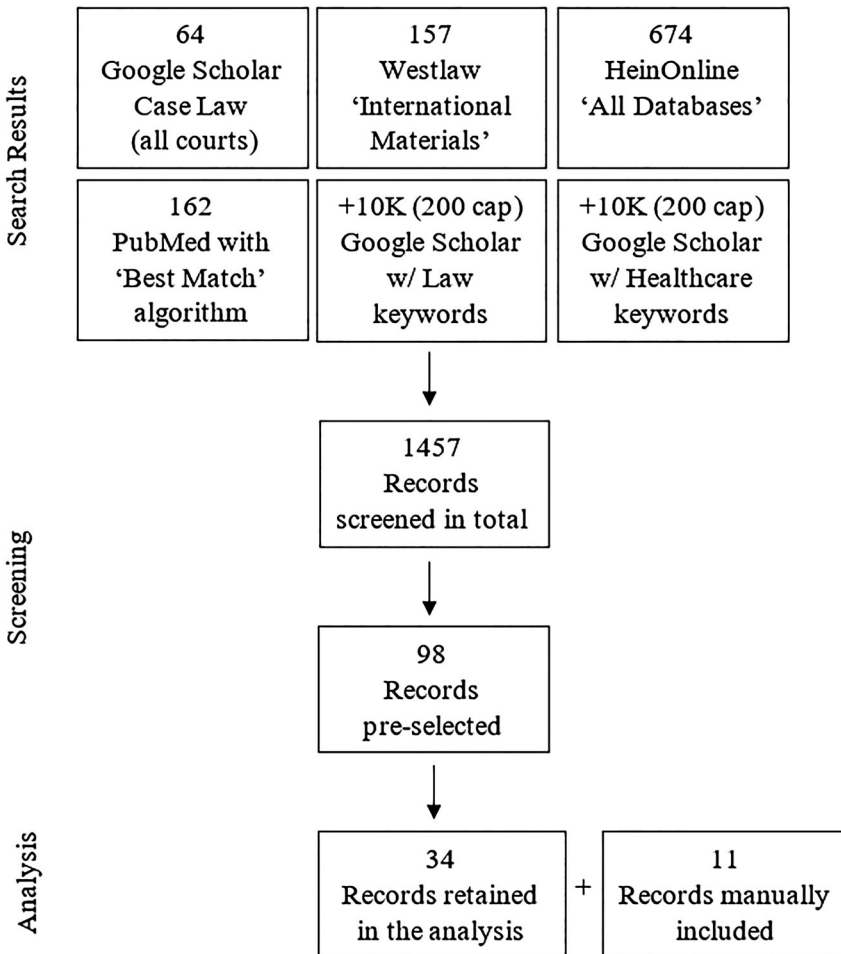


Figure 1. Summary of review process.

concerned with human-based services (e.g., Hsieh, 2016). MT research in translation studies is shedding light on multiple aspects of the technology, including its impact on human translation practices (e.g., Vieira et al., 2019), evaluation methods (e.g., Doherty, 2019) or the notion of translation quality as a matter of fitness-for-purpose (e.g., Bowker, 2019). Meanwhile, research on community interpreting has highlighted translation issues in healthcare and legal settings from cognitive and sociological perspectives. Research in this area often points to the challenges posed by medical and legal contexts, such as retaining consistent quality of service under budgetary constraints, the impact of power relationships between interlocutors, the question of trust, as well as the question of the interpreter's role as conduit versus advocate in relation to professional neutrality (Ozolins, 2015). Our analysis draws on these prior findings about the nature of language-mediated communication in healthcare and legal settings, on the one hand, and the complexity of evaluating MT technology in these settings, on the other. This allowed us to synthesise the literature according to the perception, use and impact of MT in relation to these specialised use cases. Based on this procedure, we present domain-specific findings and three overarching conclusions, which we discuss in detail below.

Results

Healthcare

Perception of MT

In healthcare, although the risks of using MT are acknowledged, the technology is often perceived as the only alternative. In the UK, medical defence organisations (i.e., bodies that specialise in medicine or dentistry-related legal complaints) have warned against the use of MT systems, not least because MT fails to meet standards imposed by the National Health Service (Moberly, 2018a). It is often recognised, however, that there may be situations where other options are not available or are difficult to access (Moberly, 2018b; Narayan, 2013). In a letter that discusses the situation in the UK, a doctor says, 'doctors should try it [MT] when other methods of translation are unavailable' (Wade, 2011). Similarly, a research letter reporting on a study carried out in India claims that MT 'has considerable potential to improve doctor-patient communication when language poses a significant barrier' (Kaliyadan & Gopinathan Pillai, 2010, p. 4). A report from Portugal mentions a successful case where a Ukrainian-speaking patient admitted to hospital with psychotic episodes was successfully treated thanks to MT, though the report points to funding issues and a lack of more appropriate options as motivations for MT use (Leite et al., 2016, p. 966). These reports serve to illustrate how language barriers coupled with funding pressures and other practical difficulties expose doctors and patients to a dilemma where MT, albeit risky, is perceived as the easiest route to cross-linguistic communication.

Use of MT

A review of MT development for healthcare settings was carried out by Dew et al. (2018). They concluded that most initiatives in MT for healthcare were still at pilot level and that concerns about accuracy and a lack of standard evaluation methods still prevent health professionals from using MT as a matter of course. As reports mentioned above show, however, doctors do consider MT use where the circumstances leave them with no

other option. Indeed, a survey of health departments in Northwestern US shows that 70.6% of responding departments did not have a budget for translation and that 30.6% of their staff members had used MT before (Turner et al., 2013, pp. 1379–1380). In a subsequent investigation, it was shown that over half of 34 health professionals who participated in an interview study had used MT before, though only about half of these had used it in a professional capacity while the other half used it for personal reasons, mostly to assimilate information (Turner et al., 2015, p. 142). Participants in this same study expressed concern about MT use, but conceded that there may be contexts where MT can be useful, including in emergency situations that require a fast response (Turner et al., 2015, p. 142). Considering MT use from the patient/client perspective rather than from the perspective of the healthcare provider, Ahmed (2018) observes that the fact that young refugees ‘are often found to rely on Google Translate to facilitate everyday communication’ suggests that this technology could also be used to support refugee healthcare. She goes on to identify the digital divide as a significant problematising factor.

Evaluation of MT

A recent study examines the use of Google Translate for translating emergency department discharge instructions from English into Chinese and Spanish (Khoong et al., 2019). The results showed that 2% of Spanish translations and 8% of translations into Chinese were inaccurate and could potentially cause significant harm (p. 580). In a similar investigation, Das et al. (2019) evaluate the use of MT for translating anticipatory guidance material given to parents (i.e., proactive advice on a child’s health and development). They tested translations from English into twenty other languages and concluded that Spanish and Portuguese were the only languages where Google Translate produced ‘mostly accurate’ results (p. 248). Chen et al. (2017) contrasted human and machine audio-recorded translations of public health information on diabetes from English into Spanish and Chinese, languages spoken by communities where diabetes is particularly prevalent in the US. They asked professional translators to rate audio recordings with three spoken questions from an informational pamphlet. The questions had been interpreted by the machine translation/interpreting system iTranslate⁷ and by professional interpreters. Their main finding is that machine and human versions were comparable for simpler questions but inferior for more difficult ones (Chen et al., 2017, p. 7). In yet another evaluation, Google Translate is assessed as a tool for spoken doctor-patient communications involving English and Mandarin in pre-anaesthetic consultations (Beh & Canty, 2015). An anaesthetist who was fluent in the two languages assessed the speech recognition and the translations for accuracy. While this study concluded that Google Translate is not accurate enough for widespread use, it mentioned the technology as potentially useful in situations where human interpreters are not available (Beh & Canty, 2015, p. 793).

While the studies mentioned above point to some level of MT usefulness for healthcare, it is worth noting that studies like these often run into considerable methodological challenges. Chen et al. (2017) assess just three sentences, for instance, and the simpler sentences among them do not necessarily concern diabetes or indeed a healthcare setting (p. 4).⁸ In Khoong et al. (2019) and Das et al. (2019), the assessment involved back-translating the Google Translate output into English. This is a problematic method. Research on the use of back-translations as a diagnostic tool that can be used to estimate MT quality has shown some correlation in the accuracy of the initial and the back translations, but the

level of accuracy of the back translations was unsurprisingly lower (Shigenobu, 2007, p. 262). Back translations can distort the nature of errors and make it difficult to identify their root causes (see Somers, 2007) so, if the circumstances permit, this is a method to avoid.

Smaller-scale studies involving fewer languages are often able to include more in-depth evaluations. Bedrick and Mauro (2009), for instance, appointed sixteen bilingual clinicians to assess the potential use of Google Translate for translating information about drug side effects into Spanish. While the study was carried out over ten years ago, its findings are not dissimilar to those of more recent research: they concluded that Google Translate had some benefit but that it was not appropriate for unsupervised use given the potential health risk posed by mistranslations (p. 37).

Types of technology tested

The review has also shown that situation-specific interactive devices are taking preference over MT systems for healthcare purposes. A study carried out at Geneva University Hospitals showed how a ‘phraselator’ – a system that uses a decision-tree method to simplify source-language questions and their translations⁹ – can outperform Google Translate in doctor-patient interactions in terms of translation quality, user satisfaction and usefulness in making a diagnosis (Bouillon et al., 2017). Similarly, Parra et al. (2018) present a hand-held system that attempts to improve the experience of those on controlled diets when they travel abroad. The system shows images and ingredient lists that help the user to identify dishes and, for instance, any allergens. A user study showed that users were better equipped to identify ingredients when they used this system instead of Google Translate (Parra et al., 2018, pp. 21–22). Another research group tested a domain-specific tool tailored to medical emergencies (Turner et al., 2019). Like the system proposed by Bouillon et al. (2017), the software presented by Turner et al. (2019) included questions and translations aimed at facilitating emergency care in intercultural scenarios. User studies showed that Spanish and Chinese-speaking participants with low English proficiency preferred the fixed-questions system over Google Translate (p. 6). However, Turner et al. (2019) stress that neither tool was fit for purpose and that improvements in accuracy and usability are still required for safe deployment of the technology (p. 11).

Implications

While the potential implications of MT errors and misuse of the technology in healthcare are significant, we did not come across cases where MT was the documented cause of ill-suited medical advice or other serious healthcare issues. However, it is worth noting that MT use recommendations for this field sometimes fail to provide objective advice. While as previously mentioned MT does not meet National Health Service standards in the UK (Moberly, 2018a), Interpreting and Translating guidance from NHS England simply mentions MT ‘should be avoided’ due to quality concerns (NHS England, 2018, p. 11). Furthermore, results of quality assessments of MT for healthcare tend to be more favourable for language pairs involving English and other Western European languages (Chen et al., 2017; Dew et al., 2018). If, as suggested by the professional letters discussed here, doctors use MT for lack of better alternatives, unequal MT quality across languages could put certain groups of patients at higher risk of misunderstanding and being misunderstood, in a

context where such patients may also be already disadvantaged (e.g., Narayan, 2013). This is a pervasive problem, which we also noted in legal use cases, as discussed in more detail below. Researchers have also observed that MT use puts a considerable burden on both patients and healthcare staff: according to Randhawa et al. (2013), MT should ‘be used very cautiously, and only in clinical encounters with literate patients’; they stress ‘the importance of physician cross-cultural communication skills to recognize and manage dissonance’.

Law

Perception of MT

Our legal search revealed mixed levels of awareness of the potential risks posed by MT in this field. Certain US courts have considered the matter and provided guidelines on MT use, but there are also examples of courts and law professionals who put themselves in risky situations in relation to MT. The state court of New Mexico is an example of an institution that has considered MT in more detail. It has a track record of appointing non-English-speaking jurors and has provided MT use guidelines in relation to these appointments. The guidelines state that unedited MT should not be used for materials expected to fulfil a formal role, for example in court proceedings or as exhibits (Chávez, 2008, p. 323). Similarly, the Immigration and Refugee Board of Canada (2014) mentions MT as a non-compliant translation type, and case law shows how machine-translated documents have been dismissed because of this (*X v Re*, 2013, § 26). A district court in California has also previously flagged MT’s unreliability in a case where machine-translated material was presented as evidence (see *NOVELTY TEXTILE, INC. v. WINDSOR FASHIONS, INC.*, 2013).

Alarming, however, there is evidence to suggest that lack of awareness of MT’s risks may be common among law professionals themselves. In *Vasquez v. United States* (2019), a Spanish-speaking federal prisoner whose counsel spoke Spanish tried to withdraw a guilty plea by claiming that he did not understand his situation due to poor translations provided by the counsel. Ironically, the new counsel defending the plea withdrawal criticises the previous counsel’s translations by referring to Google Translate as a viable source: ‘if I don’t know the word, Your Honor, I look up a translation. You find it in Google out there for free. There’s an app there. It’s called Google Translate’ (*Vasquez v. United States*, 2019). In another example, the court itself resorts to MT use: ‘Because Plaintiffs provided no translation of any Polish documents submitted in support of their motion, the Court used the free “Google Translate” service, available at translate.google.com, in order to confirm certain statements’ (*SUPER EXPRESS USA PUBLISHING CORP. v. SPRING PUBLISHING CORP.*, 2017).

Use of MT

Not all legal MT use is condemned, however. In discovery, for instance, the risk of using MT is often considered low. Discovery is a pre-trial phase in legal proceedings that involves the discovery and exchange of evidence and legal information between the parties. MT is often mentioned as a first-pass tool that can be used for triage purposes when the information is available electronically (Foster & Northrop, 2011, p. 45; Nelson & Simek, 2018, p. 19). It is usually emphasised, however, that if the information is to be

put forward as evidence, MT should give way to a professional translation (Giordano, 2013, p. 467). Patent applications constitute another legal context where MT use may be more widely accepted. The Manual of Patent Examining Procedure of the United States Patent and Trademark Office allows examiners to use MT in support of a rejection, for example (USPTO, 2018, § 1207.02). Indeed, there is a long tradition of research and use of MT for patents (see Ceausu et al., 2011; Goto et al., 2013). Nevertheless, even in this context, MT use is not found to be risk-free. Under the European Patent Convention there is no regulation that requires applicants to provide translations of ‘prior art’ (i.e., evidence that an invention is already known), though when requested to do so they may well turn to MT, and there is a precedent which illustrates how this can delay applications and lead to difficulties (Smyth et al., 2015, p. 154).

MT may also have consequences for immigration applicants and in some cases exacerbate issues faced by minorities and vulnerable individuals. In one case, the credibility of an asylum seeker’s application was questioned in the US after machine translations were provided on the application form (Schroeder, 2017, p. 320). Attention has also been brought to confidentiality issues linked to the use of MT by immigration officials. Based on a report conducted by the Office of the United Nations High Commissioner for Refugees (UNHCR, 2014), Oakes (2016, p. 893) cites the use of MT to communicate with child migrants along the US-Mexico border. The report provides details of how an officer entered all questions of an immigration form into Google Translate and read the questions out loud to a girl who had difficulty understanding the questions given the officer’s poor knowledge of Spanish (UNHCR, 2014, p. 35).

In Canada, MT use by couples is often interpreted as a sign that relationships between applicants in immigration cases may not be genuine. In one case, MT use was deemed to imply a prohibitive language barrier and therefore a possible indication of a sham marriage (*Hani v Canada*, 2017, § 2). In another example, MT use was deemed to represent a lack of effort in learning English on the part of the applicant, which in turn was deemed to call into question the genuineness of the relationship (*McDonald v Canada*, 2018, § 7).

MT use can also affect legal confidentiality privileges associated with certain types of communication. One case suggests that the use of Google Translate by married couples could thwart attempts to invoke spousal privilege as a basis for keeping marital communications undisclosed (*US v. Pugh*, 2016). While this avenue was not pursued in this case, this relates to the fact that the very use of online MT can break confidentiality, for instance because the content may be shared with the MT provider (Kenny, 2019). This also raises questions about the use of MT in doctor-patient communications, which are often protected by strict confidentiality rules.

Evaluation of MT

Earlier research on the use of MT for legal purposes focuses mostly on assessments of MT for legal texts (Farzindar & Lapalme, 2009; Kit & Wong, 2008; Yates, 2006). Some of this work is known to overlook the practical difficulties of translation quality assessment. Somers (2007, p. 618) comments on how the context-dependent and subjective nature of translation quality is overlooked in the MT assessment tasks carried out by Yates (2006), for instance. Somers also mentions more practical methodological issues, such as the fact that sentence length was not controlled for in counting the number of errors made by the MT system (p. 617). This is a problem because, assuming equal levels of the intrinsic

severity of the errors, one error in a sentence of ten words is in relative terms more problematic than two errors in a sentence of a hundred.

Wahler (2018, p. 138) mentions minimum requirements on 'how accurate a translation system must be' as a potential solution to the dangers posed by MT use in legal contexts. In principle, establishing minimum accuracy requirements is a desirable move. In practice, however, this is hard to achieve because accuracy in this context does not depend just on the MT system, but also on the intrinsic complexity of the source text and how likely it is to be translated well by MT systems (see Specia et al., 2018). In addition, any appraisals of a system's accuracy will be modulated by characteristics of the assessment method and the context in which the system will be used. Automatic MT evaluations based on the degree of textual overlap between the MT output and corresponding human reference translations (see Papineni et al., 2002), for example, may be a relatively objective way of estimating a system's quality. In legal research, this argument is put forth by Kit and Wong (2008, pp. 317–318). However, Kit and Wong overstate the issue by claiming that automatic evaluation is 'the most authoritative' method of MT assessment (Kit & Wong, 2008, p. 317). This claim overlooks previous research showing how automatic assessments have known problems that can be avoided by a human evaluation (e.g., Callison-Burch et al., 2006). More broadly, statements like these disregard the superficial nature of automatic metrics and the fact that they are, in effect, measures of similarity that do not account for the actual effect or severity of any errors.

Implications

As the examples above show, MT is now used in serious legal settings and this can have equally serious implications. Surprisingly, official recommendations in this respect are, at best, limited. A recent article points to a lack of US federal guidelines on how to establish that those who produce written translations to be used in court are qualified (Wahler, 2018, pp. 110–111). Low awareness of the risks posed by MT can have striking legal consequences. In two recently reported cases, transport police officers in the US used Google Translate to ask Spanish-speaking motorists for consent to search their vehicles (Grosdidier, 2019, p. 94). In both cases, the police found illegal substances in the vehicle and charged the motorists with a crime. However, in efforts to nullify the search consent, the use of Google Translate was later challenged in court as not enough to overcome the language barrier. As mentioned in the Introduction, in one case the motion was dismissed. In the other, however, the evidence was suppressed (Grosdidier, 2019, p. 94).

MT has also been a factor in libel cases. In a British court, a claimant in a defamation case alleged that defamatory information published in Serbian would be accessible to English language readers through a Google Translate link provided by the publisher, which the judge found to be a serious issue worthy of being considered (*Ahuja v. Politika Novine I Magazini Doo*, 2016, § 65).

It is also worth noting that, as in healthcare settings, the budgetary appeal of free online MT can often present a dilemma in legal contexts. Clients may have low proficiency in the language in which their legal case is being processed. They may not be able to afford professional translations either. If, in such a situation, lawyers cannot themselves read the original version of relevant documents, MT may offer an easy option. On the other hand, if lawyers cannot verify the original content themselves, they will not be able to attest to the accuracy of the MT output. This means that by using MT, lawyers could be found to

fall short of standards required by the duty of care they owe to clients (Wahler, 2018, p. 112). As a response to these risks, it has been proposed that the same regulations currently in place in the US for the use of interpreting in legal settings should be applied to written translations (Wahler, 2018, p. 131). Notably, Wahler also refers to the notion of controlling client intake (p. 130). This would involve avoiding taking on cases that require translations. While Wahler does not promote this practice, it cannot be ignored that the risks of MT use in this context have the potential to exacerbate social and linguistic inequalities by discouraging lawyers from working with clients who do not speak the language(s) of the country where they happen to live or where they find themselves, an issue that may also affect healthcare practices, especially in the private sector. This again shows how MT is a double-edged sword: it can make multilingual content more accessible but at the same time, owing to its limitations, pose a greater risk to certain communities.

Conclusion

The review for healthcare revealed that the use of MT in health communication is marked by high, and often urgent, demand as well as by a circumstantial lack of workable alternatives. Our findings suggest a reasonable level of awareness of the risks posed by indiscriminate use of MT in medical settings, but the technology is nevertheless regarded as a last-resort option. The literature also suggests interactive phrase dictionaries to be potentially more promising in healthcare settings than MT systems, although there is no standardised method for evaluating the technology in these contexts.

In legal settings, there is evidence of how MT use can influence the decision-making in critical legal situations. The use of MT has led to appeals and affected immigration applications and other court judgments. Given the seriousness of these issues, we were surprised by the scarcity of efforts to promote greater awareness of the risks of MT technology for this field. Attitudes to MT in legal circles also struck us as more ingenuous compared to perceptions of MT in medical settings. We noted that, relative to research in healthcare, the use of MT in legal contexts is an even more under-researched topic where the nature of specific risks and consequences is still taking shape and could be far-reaching. The low levels of awareness of MT use implications observed in legal settings are a somewhat counter-intuitive and therefore worthy finding, given the otherwise strictly regulated nature of this field. We highlight the use of MT in legal contexts as an area of priority with ample scope for future research.

Table 1 summarises key field-specific findings from the literature review. Below we present three overarching findings in relation to MT use and awareness of its implications. These findings serve as specific calls to action aimed at mitigating the risks of MT use in high-stakes settings.

MT-mediated communication merits robust cross-disciplinary research

We have found evidence that research on MT use in health and legal settings tends to underestimate the complexities of language use and language translation. Research on public service interpreting and translation in healthcare and legal contexts offers sophisticated analytical frameworks to understand the needs of the stakeholders in these settings, who form a complex web of communicative relationships. The fact that some legal

Table 1. Findings on Perception, Use and Impact of MT in Medical and Legal Settings.

	Medical settings	Legal settings
Perception of MT	Last resort; High level of awareness of risks	Easy alternative; Inconsistent levels of awareness of risks
Use of MT Situation	High demand and in emergencies	High demand; differentiated risks between behind-the-scenes use (in discovery; patents) vs public-facing (court proceedings; interviews)
Types of technology	Custom-made systems such as interactive phrase systems being tested	Off-the-shelf systems predominate
Evaluation of MT	Variety of approaches, including the use of back translations to evaluate the quality of MT	Lack of specified goal; automatic evaluation promoted as ideal and non-problematic
Implications	Duty of care could be breached; Legal implications arising from failed communication affecting medical outcome; potential for private healthcare facilities avoiding patients who need language facilitation	Duty of care could be breached; potential miscarriages of justice; Legal implications of MT use in its own right (used as evidence of language competency or intimacy level in immigration assessment); potential for lawyers to decline clients who need language facilitation

translations can stand a word-for-word approach can be taken to justify the use of MT in this field, as in the case of patent translations. However, the specific role played by non-verbal cues in court interpreting, for instance, where each side often wants to find gaps in the other's argument, highlights how it is paramount to consider all communicative clues. Such capabilities are currently beyond the scope of any MT system, and the concept of interpreters as conduits (Ozolins, 2015) needs to be carefully addressed when implementing MT for these settings. In MT research, evaluation is a growing research area (Doherty, 2019) that can offer a context-sensitive analysis of the technology in medical and legal settings beyond the use of back translations. Whether the back translations are performed by human translators or, worse still, using MT itself, in translation and MT research this method is known to be problematic (see Somers, 2007). We therefore stress the need for MT assessments to be context-dependent and to take account of the text's real-world purpose (Doherty, 2019). The analysis also suggests a gap in the understanding of the role played by human translators or interpreters in interlingual and intercultural communication. This in turn can lead to a misguided understanding of the extent of MT's and MI's capabilities, as observed in some of the content reviewed. There is, therefore, a need for language-related research from specialised domains, like healthcare and law, to draw on evidence concerning the workings of language and translation themselves and not just on information about the specialised context in which the translations are used.

Higher MT literacy is required across society

Especially in high-stakes contexts like the ones discussed above, using MT requires the user to weigh the benefits of the technology against the risks it may pose. MT-mediated communication should therefore presuppose some level of awareness of the technology's limitations and capabilities as well as of the implications of its use in different settings. Having such awareness has been referred to as a matter of being 'literate' in the use of

the technology (Bowker & Buitrago-Ciro, 2019; Williams, 2006), which is also linked to the broader concepts of machine learning and artificial intelligence literacy (Long & Magerko, 2020).

Conceptualisations of literacy in an MT-use context allude to the core assumption that MT's efficacy as a communication tool can vary depending on how and when it is used. This assumption underpins the present review to the extent that we focused on the status quo of the perception and use of MT in the specific settings analysed. This represents a user-centred standpoint, concerned with MT's repercussions for people in society. From this perspective, it is clear from the analysis above that it is not MT itself that is intrinsically risky or problematic. Rather, it is the (lack of) awareness of what this technology can and cannot do that poses a fundamental risk.

The review shows a clear demand for higher levels of this type of awareness. The use cases reviewed also demonstrate, however, that low awareness of the risks posed by MT cannot just be attributed to isolated instances of behaviour. Institutional budgetary pressures as well as rudimentary or non-existent official guidelines all contribute to uninformed MT use. This means that navigating the risks of MT-mediated communication requires a concerted effort to raise MT awareness across society from both individual and institutional perspectives.

The concept of MT literacy provides a framework within which to promote greater awareness of opportunities presented by MT and of its limitations. To date, however, MT literacy has been conceptualised predominantly in relation to academic or scholarly practices (Bowker & Buitrago-Ciro, 2019; Williams, 2006). Raising awareness of the potential risks of MT applications in academic contexts, including at educational stages that can shape individuals' approach to language and language technologies from an early age, is indeed strategic. Nevertheless, MT technology evolves fast, and users who are no longer in education and who do not speak other languages may be precisely those who are at higher risk of being affected by MT's limitations. We therefore make two suggestions aimed at raising awareness of the implications of MT use. First, we call for MT literacy to be promoted across intercultural contexts involving doctor-patient communication, legal cases and other high-stakes settings. Second, we stress a need for robust standards regarding the situations in which MT use is and is not admissible. Our review shows that, at present, guidelines on MT use tend to be applied on an ad-hoc and field-dependent basis. We reviewed documents and studies that reveal a lack of robust directives on the use of MT by lawyers and healthcare professionals. We therefore see room for significant interventions from specialists in MT, translation studies, communication, law, health and other areas to collaboratively shape guidelines that can equip professionals and other members of the public with the tools required to minimise the risks posed by MT.

More efforts are needed to democratise MT development across languages

Finally, a recurrent concern apparent across the studies included in our analysis is the way in which MT can exacerbate inequalities. For one thing, it is worth noting that MT research suffers from a disproportionate availability of data and resources for a relatively small number of the world's languages, which has been an ongoing concern for years (Jones et al., 2000). Our analysis confirms that unequal MT development continues to be a serious challenge.

On the other hand, it should also be noted that as MT technology advances it may present opportunities for democratising multilingual communication in ways that hitherto have not been possible. For instance, there have been efforts to reduce the reliance on the need for parallel bilingual texts in MT development by leveraging more easily available monolingual data (Lample et al., 2018). If efforts to increase data efficiency continue, it will become easier for MT to provide higher-quality results for less-spoken world languages for which data and other resources are scarce. This is a critical issue that merits careful attention in future MT research from a technological as well as a sociological perspective.

Notes

1. This technology can also be referred to as ‘automatic translation’, ‘automated translation’ or simply ‘online translators’. Typical examples of MT systems freely available online include Google Translate (<https://translate.google.co.uk/>) and Bing Microsoft Translator (<https://www.bing.com/translator>).
2. MI requires additional technologies for speech-to-text (speech recognition) and text-to-speech (speech synthesis) conversions. In some cases, uses of these technologies cannot be strictly classed as just translation or interpreting, such as reading written machine-translated text out loud to communicate or using speech recognition to generate written translations.
3. Restricting the review to English-language records risks excluding relevant information available in other languages. That said, an earlier attempt to use the most comprehensive bibliographic database for Japanese, CiNii, for scholarly Japanese and English publications, resulted in 24 articles on MT use in healthcare and no records for MT use in legal settings. After considering the trade-offs of expanding the search to other languages, we decided to focus on English-language databases, with English search terms, as a suitable initial method to investigate the problem at hand.
4. Patents were excluded from the Google Scholar searches.
5. The first author devised the criteria and carried out the screening and initial analysis. The remaining authors fine-tuned the analysis and carried out subsequent searches to check for coverage issues and any relevant records that could have been missed. All authors contributed to the overarching shape of the review.
6. This procedure was applied in just one case: that of the review by Dew et al. (2018), which unlike the present article examines MT’s state of advancement and feasibility for healthcare rather than perceptions of the technology or its societal impacts.
7. See <https://www.itranslate.com/>.
8. The simpler sentences in the assessment were ‘What should they be?’ and ‘What actions should I take to reach these goals?’ (p. 4).
9. Some of these are not dissimilar from earlier MT-based conversational systems (see e.g. Wahlster, 2000). While systems based on set phrases are different from MT proper, we include them in the analysis when they are compared to MT as a communication tool. We also noted studies (e.g. Albrecht et al., 2013) where this comparison is not made; such studies do not always mention machine translation as a viable option, suggesting that either there may be limited awareness in professional settings of the range of translation options available, or that in some environments MT is excluded from the outset due to perceptions of its usability.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Lucas Nunes Vieira is a Lecturer in Translation Studies with Technology at the University of Bristol, UK. He researches cognitive and sociological aspects of using machine translation in communication and human translation processes [email: l.nunesvieira@bristol.ac.uk].

Minako O'Hagan, PhD, is the Discipline Convenor for Translation Studies at the School of Cultures, Languages and Linguistics at the University of Auckland. She specialises in applied translation studies with a technology focus, including game localisation and non-professional translation [email: minako.ohagan@auckland.ac.nz].

Carol O'Sullivan is a Senior Lecturer in Translation Studies at the University of Bristol. Her research interests include translation and technology, particularly, to date, the history of audiovisual translation [email: carol.osullivan@bristol.ac.uk].

ORCID

Lucas Nunes Vieira  <http://orcid.org/0000-0003-3038-4001>

References

Sources in the Meta-Analysis

(* Manually added)

- *Ahmed, R. (2018). Challenges of migration and culture in a public health communication context. *Journal of Public Health Research*, 7(2), Article 1508. <https://doi.org/10.4081/jphr.2018.1508>
- Ahuja v. Politika Novine I Magazini Doo, [2015] EWHC 3380, 1 W.L.R. 1414 (High Court, Queen's Bench Division 2016).
- Bedrick, S. D., & Mauro, A. (2009, November 14–18). *A multi-lingual web service for drug side-effect data*. AMIA Annual Symposium proceedings. AMIA Symposium, 2009, San Francisco, CA, USA (pp. 34–38).
- Beh, T., & Canty, D. (2015). English and Mandarin translation using Google Translate software for pre-anaesthetic consultation. *Anaesthesia and Intensive Care*, 43(6), 792–793.
- *Bouillon, P., Gerlach, J., Spechbach, H., Tsourakis, N., & Halimi, S. (2017, May 28–31). *Babeldr vs Google Translate: A user study at Geneva university hospitals (HUG)*. Proceedings of the 20th annual conference of the European association for machine translation (EAMT), Prague. EAMT.
- Ceausu, A., Tinsley, J., Zhang, J., & Way, A. (2011). Experiments on domain adaptation for patent machine translation in the PLuTO project. *The 15th annual conference of the European association for machine translation (EAMT 2011)*, 30–31 May 2011, Leuven, Belgium: European Association for Machine Translation. <http://doras.dcu.ie/16412/>
- Chávez, E. L. (2008). New Mexico's success with non-English speaking jurors. *Journal of Court Innovation*, 1(2), 303–328.
- Chen, X., Acosta, S., & Barry, A. E. (2017). Machine or human? Evaluating the quality of a language translation mobile app for diabetes education material. *JMIR Diabetes*, 2(1), Article e13. <https://doi.org/10.2196/diabetes.7446>
- Das, P., Kuznetsova, A., Zhu, M., & Milanaik, R. (2019). Dangers of machine translation: The need for professionally translated anticipatory guidance resources for limited English proficiency caregivers. *Clinical Pediatrics*, 58(2), 247–249. <https://doi.org/10.1177/0009922818809494>
- Dew, K. N., Turner, A. M., Choi, Y. K., Bosold, A., & Kirchhoff, K. (2018). Development of machine translation technology for assisting health communication: A systematic review. *Journal of Biomedical Informatics*, 85, 56–67. <https://doi.org/10.1016/j.jbi.2018.07.018>
- *Farzindar, A., & Lapalme, G. (2009). Machine translation of legal information and its evaluation. In Y. Gao & N. Japkowicz (Eds.), *Lecture notes on artificial intelligence* (Vol. 5549, pp. 64–73). Springer-Verlag.

- Foster, T. J., & Northrop, S. A. (2011). A lawyer's guide to source code discovery. *The Federal Lawyer* (February 2011), pp. 42–46. <https://www.fedbar.org/wp-content/uploads/2011/02/feature3-feb2011-pdf-1.pdf>
- Giordano, S. S. (2013). It's all Greek to me: Are attorneys who engage in or procure legal translation for their clients at risk of committing an ethical violation? *Quinnipiac Law Review*, 31(2), 447–487.
- Goto, I., Chow, K. P., Lu, B., Sumita, E., & Tsou, B. K. (2013). Overview of the patent machine translation task at the NTCIR-10 Workshop. In N. Kando, & T. Kato (Eds.), *Proceedings of the 10th NTCIR Conference, June 18–21, 2013, Tokyo, Japan* (pp. 260–286). National Institute of Informatics.
- Grosdidier, P. (2019). Lost in translation: Google Translate might not always be the best way to obtain search consent. *Texas Bar Journal*, 82(2), 94.
- Hani v Canada (Minister of Citizenship and Immigration), CarswellNat 5166, No. MB3-04842 (Immigration and Refugee Board of Canada, Immigration Appeal Division 2017).
- *Immigration and Refugee Board of Canada. (2014, December 8). *Refugee protection division practice notice: Compliance with refugee protection division rules*. <https://irb-cisr.gc.ca/en/legal-policy/procedures/Pages/AviPraNotRpdSprComRulResReg.aspx>
- *Kaliyadan, F., & Gopinathan Pillai, S. (2010). The use of Google language tools as an interpretation aid in cross-cultural doctor-patient interaction: A pilot study. *Informatics in Primary Care*, 18(2), 141–143. <https://doi.org/10.14236/jhi.v18i2.764>
- Khoong, E. C., Steinbrook, E., Brown, C., & Fernandez, A. (2019). Assessing the use of Google Translate for Spanish and Chinese translations of emergency department discharge instructions. *JAMA Internal Medicine*, 179(4), 580–582. <https://doi.org/10.1001/jamainternmed.2018.7653>
- Kit, C., & Wong, T. M. (2008). Comparative evaluation of online machine translation systems with legal texts. *Law Library Journal*, 100(2), 299–321.
- Leite, F. O., Cochat, C., Salgado, H., da Costa, M. P., Queirós, M., Campos, O., & Carvalho, P. (2016). Using Google Translate© in the hospital: A case report. *Technology and Healthcare*, 24(6), 965–968. <https://doi.org/10.3233/THC-161241>
- McDonald v Canada (Minister of Citizenship and Immigration), CarswellNat 6057, No. VB6-00279 (Immigration and Refugee Board of Canada, Immigration Appeal Division 2018).
- Moberly, T. (2018a). Doctors are cautioned against using Google Translate in consultations. *BMJ*, 363, Article k4546. <https://doi.org/10.1136/bmj.k4546>
- Moberly, T. (2018b). Doctors choose Google Translate to communicate with patients because of easy access. *BMJ*, 362, Article k3974. <https://doi.org/10.1136/bmj.k3974>
- Nelson, S. D., & Simek, J. W. (2018). Running with the Machines: Artificial intelligence advances bring benefits, threats to practice of law. *Montana Lawyer*, 43(8), 18–21.
- *NHS England. (2018). Guidance for commissioners: Interpreting and translation services in primary care. NHS England / primary care commissioning. <https://www.england.nhs.uk/wp-content/uploads/2018/09/guidance-for-commissioners-interpreting-and-translation-services-in-primary-care.pdf>
- NOVELTY TEXTILE, INC. v. WINDSOR FASHIONS, INC., No. CV 12-05602 DDP (MANx) (Dist. Court, CD California 2013). https://scholar.google.co.uk/scholar_case?case=378347259667692690
- Oakes, J. (2016). U.S. immigration policy: Enforcement & deportation Trump fair hearings-systematic violations of international non-refoulement obligations regarding refugees. *North Carolina Journal of International Law*, XLI(4), 833–918.
- *Parra, A., Haddad, A. W., Boutin, M., & Delp, E. J. (2018). *A hand-held multimedia translation and interpretation system with application to diet management*. <https://arxiv.org/abs/1807.07149v1>
- *Randhawa, G., Ferreyra, M., Ahmed, R., Ezzat, O., & Pottie, K. (2013). Using machine translation in clinical practice. *Canadian Family Physician*, 59(4), 382–383.
- Schroeder, J. L. (2017). The vulnerability of asylum adjudications to subconscious cultural biases: Demanding American narrative norms. *Boston University Law Review*, 97(1), 315–348.
- *Smyth, D., Barker, R., & Belcher, T. (2015). Rage against the (translation) machine. *Journal of Intellectual Property Law & Practice*, 10(3), 153–154. <https://doi.org/10.1093/jiplp/jpu257>

- Somers, H. (2007). The use of machine translation by law librarians – A reply to Yates. *Law Library Journal*, 99(3), 611–620.
- SUPER EXPRESS USA PUBLISHING CORP. v. SPRING PUBLISHING CORP, No. 13-cv-2814 (DLI)(JO) (Dist. Court, ED New York 2017). https://scholar.google.co.uk/scholar_case?case=4176961653124460868
- Turner, A. M., Brownstein, M. K., Cole, K., Karasz, H., & Kirchoff, K. (2015). Modeling workflow to design machine translation applications for public health practice. *Journal of Biomedical Informatics*, 53, 136–146. <https://doi.org/10.1016/j.jbi.2014.10.005>
- Turner, A. M., Choi, Y. K., Dew, K., Tsai, M. T., Bosold, A. L., Wu, S., Smith, S., & Meischke, H. (2019). Evaluating the usefulness of translation technologies for emergency response communication: A scenario-based study. *JMIR Public Health and Surveillance*, 5(1), Article e11171. <https://doi.org/10.2196/11171>
- Turner, A. M., Mandel, H., & Capurro, D. (2013). Local health department translation processes: Potential of machine translation technologies to help meet needs. *Annual Symposium Proceedings. AMIA Symposium*, 2013, pp. 1378–1385. <https://www.ncbi.nlm.nih.gov/pubmed/24551414>
- *UNHCR. (2014). *Findings and recommendations relating to the 2012–2013 missions to monitor the protection screening of Mexican unaccompanied children along the U.S.-Mexico border*. Regional Office Washington, D.C. for the United States and The Caribbean. http://www.immigrantjustice.org/sites/immigrantjustice.org/files/UNHCR_UAC_Monitoring_Report_Final_June_2014.pdf
- *USPTO. (2018). *Manual of patent examining procedure*. USPTO United States Patent and Trademark Office. <https://mpep.uspto.gov/RDMS/MPEP/current#/current/d0e122292.html>
- US v. Pugh, F. Supp. 3d, 15-CR-116 (NGG) (Dist. Court, ED New York 2019) 162: 97. https://scholar.google.co.uk/scholar_case?case=10165936483680238055
- Vasquez v. United States, No. 3: 16-cv-2623-D-BN (Dist. Court, ND Texas 2019). https://scholar.google.co.uk/scholar_case?case=4394896296660864715
- Wade, R. G. (2011). Try Google Translate to overcome language barriers. *BMJ*, 343(nov15 2), Article d7217. <https://doi.org/10.1136/bmj.d7217>
- Wahler, M. E. (2018). A word is worth a thousand words: Legal implications of relying on machine translation technology. *Stetson Law Review*, 48, 109–139.
- X v Re, CarswellNat 10832, No. TB3-01748 (Immigration & Refugee Board of Canada (Refugee Protection Division) 2013).
- Yates, S. (2006). Scaling the tower of Babel Fish: An analysis of the machine translation of legal information. *Law Library Journal*, 98(3), 481–502.
- Other Sources**
- Albrecht, U. V., Behrends, M., Schmeer, R., Matthies, H. K., & von Jan, U. (2013). Usage of multi-lingual mobile translation applications in clinical settings. *JMIR mHealth uHealth*, 1(1), Article e4. <https://doi.org/10.2196/mhealth.2268>
- Angelelli, C. V. (2008). *Medical interpreting and cross-cultural communication*. Cambridge University Press.
- Bowker, L. (2019). Fit-for-purpose translation. In M. O'Hagan (Ed.), *The Routledge handbook of translation and technology* (pp. 453–468). Routledge.
- Bowker, L., & Buitrago Ciro, J. (2019). *Machine translation and global research: Towards improved machine translation literacy in the scholarly community*. Emerald Publishing.
- Braun, S. (2019). Technology and interpreting. In M. O'Hagan (Ed.), *The Routledge Handbook of translation and technology* (pp. 271–288). Routledge.
- Callison-Burch, C., Osborne, M., & Koehn, P. (2006). Re-evaluating the role of BLEU in machine translation research. In *Proceedings of the 11th Conference of the European Chapter of the association for computational linguistics, 3–7 April 2006, Trento, Italy* (pp. 249–256). Association for Computational Linguistics.
- Castilho, S., Moorkens, J., Gaspari, F., Calixto, I., Tinsley, J., & Way, A. (2017). Is neural machine translation the new state of the art? *The Prague Bulletin of Mathematical Linguistics*, 108(1), 109–120. <https://doi.org/10.1515/pralin-2017-0013>
- Doherty, S. (2019). Translation technology evaluation research. In M. O'Hagan (Ed.), *The Routledge Handbook of translation and technology* (pp. 339–353). Routledge.

- Hsieh, E. (2016). *Bilingual health communication: Working with interpreters in cross-cultural care*. Routledge.
- Jones, E., Thomas, N., & King, A. (2000). *Linguistic diversity on the internet: Assessment of the contribution of machine translation* (Report No. PE 289.662 /Fin.St.). Luxembourg: European Parliament - Directorate General for Research. [http://www.europarl.europa.eu/RegData/etudes/etudes/join/2000/289662/DG-4-JOIN_ET\(2000\)289662_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/etudes/join/2000/289662/DG-4-JOIN_ET(2000)289662_EN.pdf)
- Kenny, D. (2019). Machine translation. In M. Baker, & G. Saldanha (Eds.), *The Routledge encyclopedia of translation studies* (pp. 305–310). Routledge.
- Kuczmariski, J. (2018, November 28). A new look for Google translate on the web. *The Keyword*. <https://www.blog.google/products/translate/new-look-google-translate-web/>
- Lample, G., Ott, M., Conneau, A., Denoyer, L., & Ranzato, M. A. (2018). *Phrase-based & neural unsupervised machine translation*. <https://arxiv.org/abs/1804.07755>
- Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. *CHI 2020, April 25–30, 2020, Honolulu, HI, USA* (Paper 598). <https://doi.org/10.1145/3313831.3376727>
- Narayan, L. (2013). Addressing language barriers to healthcare in India. *The National Medical Journal of India*, 26(4), 236–238. <http://archive.nmji.in/archives/Volume-26/Issue-4/SFM-II.pdf>
- Nurminen, M. (2018, September 12). Machine translation in everyday life: What makes FAUT MT workable? *TAUS eLearning Blogs*. <https://blog.taus.net/elearning/machine-translation-in-everyday-life-what-makes-faut-mt-workable>
- Ozolins, U. (2015). Ethics and the role of the interpreter. In H. Mikkelson, & R. Jourdenais (Eds.), *The Routledge Handbook of interpreting* (pp. 319–336). Routledge.
- Papineni, K., Roukos, S., Ward, T., & Zhu, W. J. (2002). BLEU: A method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting on association for computational linguistics, 7–12 July 2002, Philadelphia, PA USA* (pp. 311–318). Association for Computational Linguistics.
- Patil, S., & Davies, P. (2014). Use of Google Translate in medical communication: Evaluation of accuracy. *BMJ*, 349(dec15 2), Article g7392. <https://doi.org/10.1136/bmj.g7392>
- Shigenobu, T. (2007). Evaluation and usability of back translation for intercultural communication. In N. Aykin (Ed.), *Usability and internationalization. Global and local user interfaces. UI-HCII 2007. Lecture notes in computer science* (Vol. 4560, pp. 259–265). Springer.
- Specia, L., Scarton, C., & Paetzold, G. H. (2018). Quality estimation for machine translation. *Synthesis Lectures on Human Language Technologies*, 11(1), 1–162. <https://doi.org/10.2200/S00854ED1V01Y201805HLT039>
- Vieira, L. N., Alonso, E., & Bywood, L. (2019). Introduction: Post-editing in practice – process, product and networks. *The Journal of Specialised Translation*, 31, 2–13. https://www.jostrans.org/issue31/art_introduction.php
- Wahlster, W. (Ed.). (2000). *Verbmobil: Foundations of speech-to-speech translation*. Springer-Verlag.
- Williams, L. (2006). Web-based machine translation as a tool for promoting electronic literacy and language awareness. *Foreign Language Annals*, 39(4), 565–578. <https://doi.org/10.1111/j.1944-9720.2006.tb02276.x>
- Wu, A. X., & Taneja, H. (2016). Reimagining internet geographies: A user-centric ethnological mapping of the world wide web. *Journal of Computer-Mediated Communication*, 21(3), 230–246. <https://doi.org/10.1111/jcc4.12157>
- Yang, J., & Lange, E. (2003). Going live on the internet. In H. Somers (Ed.), *Computers and translation: A translator's guide* (pp. 191–210). John Benjamins.