1 Principles and Applications

Machine translation is in our pockets and on our wrists – its convenience needs little explanation. As this technology evolves, translations are also sounding more natural and can be hugely helpful. The problem, as outlined in the Introduction, is that consistency is not one of the technology's fortes. Largely because of the dynamic nature of language itself, machine translation can do well one moment and fail the next. This technology therefore is not perfect, but it is important to address head-on the fact that humans are not perfect either. Even language professionals can make mistakes.

Take the case of a Spanish-speaking woman based in Portland in the United States. On 12 April 2011, the woman started having breathing difficulties. Her husband, also a Spanish speaker, called 911 and spoke to the operator through an interpreter. When asked for their address, the woman's husband replied, '2601 111th Avenue', in Spanish. The interpreter made a mistake and told the operator, in English, '2600 101st Avenue'. The ambulance went to the wrong address. The woman stopped breathing while waiting and died three days later in hospital after her family decided to withdraw life support. A lawsuit was filed in connection with her death seeking a maximum of one million US dollars for her suffering and a maximum of two million US dollars for the suffering and loss caused to the family.[1]

In a separate incident also involving a 911 call, a different interpreter told the call handler that the caller, a Spanish-speaking man based in New Orleans, was being dishonest. The man said he had been shot. While groaning with pain, he appeared to give inconsistent answers and was not revealing his location. The interpreter concluded that the call could be dismissed and told the operator: 'He is telling you a different story every time . . . He's acting ma'am.'[2] Despite the interpreter's characterisation of the call, the man's phone was tracked, and he was eventually found. As it turned out, the man had indeed been shot and needed urgent care. Luckily, he survived. The company that provided interpreting services to the 911 call centre later stated that the interpreter had breached protocol and was no longer a member of their team.

While mistakes like those made in these two 911 calls are difficult to quantify, as these examples show, they do happen. So when discussing the

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:38:37, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009769372.002

risks of machine translation, it is important to recognise that other ways of communicating will also carry some level of risk. Misunderstandings can happen in any type of communication, even the ones that do not involve different languages.

There are nevertheless fundamental differences between hiring a professional language service and using machine translation without proficient supervision. Although professional language services are not risk-free, the risks posed by these services are easier to mitigate. There are official frameworks that outline the type of knowledge and skills that translators and interpreters should have to be considered professionally competent.[3] Translators and interpreters often need to pass exams before they can be certified by professional bodies.[4] There are also standards in place that describe how translation and interpreting tasks should be carried out to ensure a high-quality service is delivered.[5] These measures are not infallible, but they hopefully help to ensure that incidents like those in these two 911 calls happen only rarely.

In the contexts examined in this book, hiring a professional translator or interpreter will almost always be safer than using machine translation without any editing or bilingual supervision. But to frame this subject as simply a choice between language professional and machine would be to oversimplify the problem. Language professionals may be unavailable for several reasons. In some cases, their use may be discouraged by financial or political pressures. Time pressure can also influence decisions, as can users' understanding of the risks, the content of the communication, and the value placed on convenience. This chapter reviews concepts that are central to how these factors can influence the methods used for dealing with a language barrier. In the following sections, I first propose a set of attributes that can be used to distinguish different types of machine translation use. I then look at examples that illustrate the complexity of machine-translated interactions in some public services before looking at key user competencies these interactions require.

Machine Translation-Mediated Communication

The concept of AI-mediated communication is relatively new. Researchers at Stanford and Cornell universities have defined it as 'mediated communication between people in which a computational agent operates on behalf of a communicator by modifying, augmenting, or generating messages' (emphasis omitted).[6] Machine translation-mediated communication can be classed as a subtype of this broader type of AI mediation where translation-generating technologies are the computational agents that modify (i.e., translate) a message. Subsequent work by Emma Goldenthal and a team based at the same universities refers to translation as a category of AI-mediated communication.[7] In their

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:38:37, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009769372.002

categorisation, translation can be divided into 'live translation', 'text-to-text translation', 'online bilingual dictionaries' and 'subtitling'.[8]

As can be seen in the Introduction, my understanding of machine or AI translation in this book differs slightly from the one conveyed by these categories. Specifically, where entries in a dictionary are reproduced directly from a human source – for example, term definitions drafted by human lexicographers or word-in-use examples that are copied verbatim from human-produced texts – I do not consider online bilingual dictionaries to be a type of machine translation. Dictionaries and machine translation systems can nevertheless be integrated, which likely prompted the inclusion of online dictionaries as a category in the typology above. At the time of writing, the online dictionary tool Linguee – the example cited by Goldenthal et al.[9] – has a 'Translate text' link that directs users to DeepL, a machine translation system available on a separate interface.[10] Other translation categories outlined above, especially live and text-to-text translation, are within the scope of my discussion here. To situate these categories within the professional environments examined in this book, a slightly more detailed categorisation is necessary. Below I propose four criteria for describing and analysing machine-translated communication, namely its synchronicity, directionality, remoteness and overtness.

Synchronicity

In Goldenthal et al.'s categories, live translation involves uses of speech while text-to-text translation involves written texts only.[11] As the name suggests, liveness – or synchronicity – is a feature of live translation, but text-to-text machine translation can also be live, for example if individuals show written messages to each other on a screen in real time. Uses of text or speech are therefore not central to my treatment of synchronicity in machine translation use. I instead distinguish between synchronous and asynchronous tasks. I call synchronous tasks those in which a message moves between at least two parties in real time (e.g., talking on the phone). In asynchronous tasks, there is a delay in checking the message or in replying (e.g., sending and receiving an email).

Synchronicity matters for machine translation use for at least two reasons. First, synchronous tasks are more likely to be marked by urgency and improvisation, so machine translation may seem particularly useful for this type of task, especially in situations where seeking human assistance is difficult or impractical. Second, synchronous tasks are the prime locus for multimodality – that is, when multiple modes of communication, such as written, spoken, or non-verbal, are used together. For example, many online meeting tools, such as Zoom or Microsoft Teams, allow participants on a video call to speak in their own languages while captions are displayed on screen with translations of what is said[12] – a mixture not only of languages but also of written and spoken modes

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:38:37, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009769372.002

of communication. Synchronicity in this case will influence participants' approach to the exchange. Team members may seek real-time clarifications or confirmation that they are being understood. The synchronous nature of the interaction is also likely to be conducive to a more spontaneous environment, which possibly modulates participants' tolerance to translation errors.

Directionality

Bidirectional tasks will be those where participants can all reply to each other – exchanging emails is a typical example, as is submitting a formal request to an organisation or government body (e.g., using an application form) and expecting a response. Unidirectional tasks will be those where information is provided to a receiver or general group of individuals who are not expected to reply. A typical example of this type of task is publishing information on a website.

Synchronous tasks are usually bidirectional by default. They will usually involve at least two participants who communicate with each other by reading/listening and responding to each other's messages – that is, messages can move both ways between sender and receiver. One exception would be a live speech where audience members do not have an opportunity to reply or ask questions. The message in this case flows in one direction from the speaker to the live audience. This type of live communication is unlikely in the contexts examined in this book, so my discussion of directionality is more relevant to asynchronous tasks.

The distinction I make between uni- and bidirectional tasks is akin to what early models of communication refer to as mass and dyadic communication – when information reaches large swathes of people and when it is exchanged between two individuals, respectively. The Westley–MacLean model of communication,[13] for instance, makes this distinction. This model also encompasses communication at an institutional or collective level where the entities sending or receiving messages do not necessarily do so as individuals.[14] Institutional issuers of information are particularly common in uses of machine translation in public services – for example, when government departments machine-translate their website or when official messages are disseminated to the community without a named author.

In many cases, unidirectional communication will be the same as mass communication, but there are nuances of machine translation use that do not quite fit the 'mass' characterisation as usually conceived. For instance, the sociolinguist Philipp Angermeyer has shown that public signage directed at Roma asylum applicants in Toronto has involved the use of Google Translate.[15] The message on these signs has been found to involve prohibitions and aspects of public order. The language of the signs is Hungarian, or ungrammatical 'Google-Hungarian'.[16] The signs clearly single out a specific community even

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:38:37, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009769372.002



Figure 1.1 Sign in Hungarian displayed at the entrance to a recreational centre in Toronto.[17]

though they are displayed openly in a public space. Figure 1.1 shows an example of such a sign displayed at the entrance to a recreational centre. The message, in Hungarian, is ungrammatical and difficult to follow even for Hungarian native speakers, but the gist of it is to stop unauthorised Roma individuals from using the facilities.[18]

This type of signage is unlike typical examples of mass communication such as books or televised content. The message is nevertheless unidirectional since the targeted communities cannot engage with it in a bidirectional exchange (even though they can initiate separate interactions by approaching the relevant institutions to complain about the signs or refer to them in any other way).

Remoteness

Synchronous tasks may happen in a physical space shared by all participants (i.e., face-to-face), or they may happen remotely through communication technologies including the telephone, videoconferencing or online chat. Asynchronous tasks, by contrast, will often take place remotely – with participants in different physical locations.

A shared physical space matters in machine translation use because it may be associated with specific ways of communicating and with perceptions of the technology's convenience. Face-to-face machine translation mediation is more likely to rely on a combination of communicative strategies, which in any event probably involves body language and non-verbal cues. Video technologies also allow interactants to use body language, but the role played by non-verbal communication is more prominent in a face-to-face context. This is because of

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:38:37, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009769372.002

the proximity of participants and the reliability of the contact, which in face-to-face interactions is not subject to internet speed or technical glitches. As for perceptions of convenience, like in other types of synchronous tasks, a shared physical presence will often be associated with spontaneous uses of machine translation in situations where seeking human assistance may be difficult. In some cases, the need for professional assistance can and should be anticipated – for example, in booked medical appointments or other formal interactions where language needs can be identified beforehand. Other face-to-face contexts may be less predictable, and machine translation here may seem like the easiest way of communicating even when its use might not be recommended.

Overtness

An important defining factor of machine translation use is what I call overtness. In face-to-face interactions, uses of machine translation are in most cases overt – that is, it is apparent to all involved that a translation tool is in use. Even in such a context, some participants may not be aware of the automated nature of the translations. But in most cases just by using a smartphone, showing a translation on a screen or playing an audio version of the content out loud, it will be clear that some type of language technology is in action. In asynchronous contexts, by contrast, it is easier for machine translation use to be covert – that is, for those who consume machine translations to do so unknowingly.

Human translations too can be classed as covert or overt.[19] When they are overt, it will be clear to the audience that the material consists of a translation and that an original version of the content exists in a different language. When translations are covert, they are presented as an original. An example of overt translation would be a book that is published in translation where a connection to the original is clear. For example, the original author and the translator may both be mentioned on the cover. The translator may also write a preface that talks about the process of translating the book. An example of a covert human translation would be a medication package insert provided in just one language which is not the language in which the content was drafted. The translation of the insert in this case acts as an independent product.

In relation to machine translation, overtness is not just about the status of the content as a translation, but also about its status as a translation that has been automatically generated. A machine translation will be covert when the sender or message issuer does not acknowledge that a translation tool is in use and/or when the use of machine translation cannot be reasonably presumed from the context. For example, the contents of a website may be machine-translated beforehand and the output of the machine translation tool may be published without any type of acknowledgement that the technology was used. Alternatively, the website publishers may acknowledge that versions of the

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:38:37, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009769372.002

website in other languages are powered by a machine translation service – for example, where a page or section of the website bears the label 'Translated by Google'. The website may also alert its visitors to the fact that the machine translations may contain errors and should not be trusted for making important decisions. It is not uncommon for warnings of this nature to be buried in the website's small print, which from the perspective of many website visitors is not enough for machine translation to be classed as overt since the content may be consumed without a reasonable expectation that it could be defective.

I discuss the issue of overtness in a paper published in 2023 where I highlight the difference between directly operating a machine translation tool and simply consuming the machine translation output.[20] This difference may seem straightforward, but when consumers of the content do not operate a machine translation tool themselves, they may not, as described above, be aware that the technology is in use. A reasonable question to ask at this point is whether it is a problem if machine translation use is not clearly declared. I argue that it is for two reasons. First, assuming that content consumers are minimally familiar with the risks of using unedited machine translations (see Literacies and Meta-Literacies), they are likely to approach the content differently if they are aware that this type of translation is being provided. Second, sometimes the content consumers are expected to be responsible for the use of the tool. When the publishers of a website mention the possibility of mistranslations in the website's terms of use, this type of warning shifts responsibility to consumers of the content while exempting those who deploy the technology. This shift is problematic, but it is even worse if consumers of the content are not clearly notified of the risks because the notification is unclear or difficult to find. In any event, while the machine translation deployers may be seen to be 'doing their part' by warning users, this type of warning is of limited value since visitors to the website are left with no option other than to run the risk of consuming misinformation.

* * *

Distinguishing between the types of machine translation use outlined above is important because these distinctions help to reveal the multifaceted affordances of machine translation tools. Different types of communication can involve different challenges. Synchronous interactions are particularly susceptible to issues of interpreter availability, for instance, especially if the interaction is unexpected. Depending on the language for which assistance is needed, even when the interaction can be planned, it may be difficult to find an interpreter with the right expertise. A search on the UK National Register of Public Service Interpreters reveals that at the time of writing there is a single interpreter on the register in the UK who provides services for Igbo,[21] a language spoken by over thirty million people in Nigeria.[22] Igbo is supported by both Google Translate

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:38:37, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009769372.002

and Microsoft Translator. Although sources other than the National Register can be used to contact an interpreter, if a UK public service professional is taken by surprise by a need for assistance for Igbo, machine translation will seem like an obvious port of call.

The course of action to take in this type of situation will vary, but even in contexts where machine translation use is not being considered there will be ways of mitigating the risks posed by the language barrier. Individuals can point at their language on a list, for instance, a widely used language identification technique that can quickly allow a front-line professional to seek further assistance. Cards containing symbols may also be able to support the interaction. The Department of Health in the Australian state of Queensland has a card of this nature called the Ward Communication Tool.[23] The tool, shown in Figure 1.2 in its Spanish version, has a numerical scale of pain, drawings of the human body as well as words and phrases such as 'blood pressure', 'bleeding' and 'stand up' in over thirty languages. Igbo is not one of these languages, but pictorial information provided by this type of analogue tool may in any case complement machine-translated communication.

In asynchronous communication, the lack of live contact poses different problems. During the Covid-19 pandemic, the website of the Virginia Health Department in the US said in its Spanish version that Covid-19 vaccinations were not required. The Spanish version of the website was a machine translation provided by Google. What the original English text said was that no one would be forced to get vaccinated.[24] As mentioned, a warning would be unlikely to be of much use in this case, so the only effective method of mitigating the error would have been to avoid it in the first place by using a reliable professional service to translate any critical information available on the website. Even written services are often time-sensitive, especially during a pandemic, but those in charge of a website will have more time to plan and seek assistance than, for instance, an emergency nurse.

Each of the communication categories discussed above therefore involves its own pitfalls. And irrespective of what the options may be, machine translation is increasingly likely to jump out as a tool to try, sometimes prematurely or in lieu of safer methods.

The Concept of Usability

The convenience of machine translation tools is closely associated with the concept of usability. The International Organization for Standardization defines usability as the 'extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use'.[25] There are thus three components to usability: effectiveness, efficiency and satisfaction.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:38:37, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009769372.002

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:38:37, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009769372.002

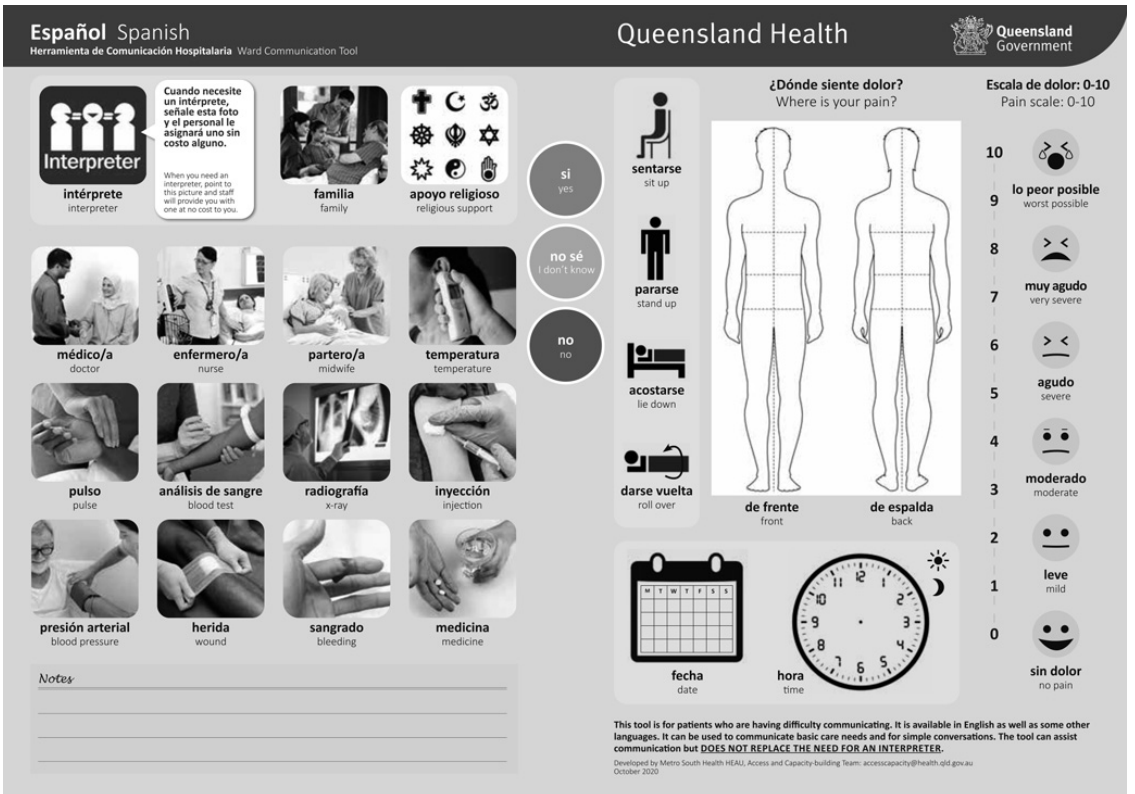


Figure 1.2 Second page of the Ward Communication Tool (Spanish) developed by the Access and Capacity-Building Team at Metro South Health, Queensland, Australia.

In the contexts discussed here, effectiveness consists of successfully achieving a communicative objective. It will often involve the ability to understand or convey information without any adverse effects for the message distributers, receivers or others in society more broadly. Efficiency involves economical or judicious uses of effort, time and money that are commensurate with what is achieved.[26] Satisfaction involves how happy users are with the product – that is, their own assessment of efficiency and effectiveness.

As specified in the definition above, the three components of usability are relative. The same machine translation system can have different levels of effectiveness depending on the 'specified context of use'. As mentioned in the Introduction, this is because the system itself can produce outputs of different levels of quality depending on the content, but also because expectations of how well the system needs to perform will be context-dependent. The variable nature of quality expectations is linked to the notion of fitness for purpose.

In machine translation research, fitness for purpose usually refers to the idea that a translation does not need to be perfect to be useful.[27] No translation, in fact, is ever intrinsically and objectively perfect, so fitness for purpose is not just a cornerstone of machine translation use. It is an important factor in definitions of translation quality more broadly.[28]

Assessments of usability often need to define what fitness for purpose means for a specific scenario. For example, in 2014 researchers at Dublin City University and at the Federal University of Belo Horizonte, in Brazil, wanted to check whether human post-editing improved the usability of instructions on how to configure a piece of software.[29] Each participant in the study was asked to complete a series of software configuration tasks using instructions that were either unedited or post-edited machine translations. The human editing ensured that the machine translations were accurate and understandable, but it left untouched any problems with the translations' style or with how idiomatic they sounded. The researchers measured the effectiveness of the instructions based on successful task completion – that is, the ability to configure the software as instructed. Efficiency was defined as the ratio of successful tasks over total task time. The participants reported their satisfaction using rating scales. The research team also used eye-tracking technology to estimate the level of effort required by the tasks, which can impact task efficiency. The results showed that the participants completed the task more efficiently and with higher satisfaction when the post-edited translations were used. But those who used the unedited machine translations could also complete the tasks successfully. Although the post-edited translations were superior, even unedited machine translations could be considered fit for purpose based on this study.

In another fitness-for-purpose assessment, native speakers of Spanish based in Ottawa were asked to rate translations of the Ottawa Public Library

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:38:37, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009769372.002

website.[30] Texts from the website were presented to participants in four translated versions: (1) unedited machine translation, (2) maximally edited machine translation (where all linguistic issues are rectified, including those that do not affect comprehension), (3) rapidly edited machine translation (where style and idiomaticity issues are ignored), and (4) a translation drafted by a human professional without machine translation assistance. Participants were first asked to consider the translations in a blind evaluation. They were then asked to consider the translations again but this time with details of the time and cost taken to produce them, which was higher the more human input was required. When participants considered the details of cost and production time, the rapidly edited machine translations were their preferred version. Unedited machine translation came second in their order of preference.[31]

Fitness-for-purpose tests of this nature show how open machine translation users are to making compromises. Machine translations are not perfect, and often they do not need to be, so users will often be happy to trade precision for convenience. But there are important caveats. Users may not always be well placed to judge the quality of machine translations. When machine translation is used as a communication tool, users are unlikely to know both the source and target languages well otherwise the tool would not be used in the first place. Inaccuracies can therefore be difficult to identify. In addition, the consequences of using machine translation may be unrelated to translation quality. Especially in the contexts analysed in this book, specific aspects of protocol or legislation may need to be considered, and the consequences of breaching expectations of this nature may only become apparent after machine translation is used.

Furthermore, for most users of machine translation, fitness for purpose is more likely to be judged intuitively based on everyday experience and not based on expertly designed controlled studies like the ones above. In the survey I distributed to UK residents in 2019, satisfaction with machine translation was extremely high.[32] For most survey respondents, machine translation tools were used out of choice rather than for lack of a better alternative. In most cases, the context of use consisted of everyday information consumption tasks, such as browsing social media sites. But those who vouched for the technology's usability sometimes referred to contexts where balancing risks and benefits was not straightforward – for example, in border crossings: 'In my job as a border force officer it [machine translation] is useful as it is quick to access.'[33] Here speed is the key factor in a positive evaluation of machine translation, but speed only serves a purpose if individuals are accurately represented, and misrepresentation, as mentioned, can be difficult to identify. Users' appraisal of fitness for purpose can therefore be skewed. Efficiency can be a stronger driver of satisfaction than effectiveness.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:38:37, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009769372.002

Organisational Factors

The emphasis placed on efficiency is particularly apparent in organisational settings. The first few decades of the 21st century have been marked by crises that put tremendous financial pressure both on businesses and on the public sector. The 2008 financial crisis, the Covid-19 pandemic and more recently wars and geopolitical tensions have all affected institutional budgets. In the UK, local government authorities are a case in point. These local government offices oversee a range of services including social care, housing, schools, libraries and some aspects of public health. These authorities do not have the power to borrow money for routine expenditures. They are funded mainly through local taxation and central-government grants.[34] Research by the Institute for Government, a cross-party think-tank in the UK, shows that the spending power of local authorities in England had a 10.2 per cent drop from 2009/10 to 2021/22.[35] In many parts of the country, the demand for local government services will have increased in the same period. Many services that fall under the remit of local authorities involve communication with the community, whether it be in managing housing, in circulating public health information or in providing care to older adults. The ability of local authorities to provide these services multilingually using language professionals is unsurprisingly under strain.

Guidance provided by the UK's central government has directly targeted language services as an area where local authorities should be making savings. A document issued in 2012 by the Department for Communities and Local Government – currently the Ministry of Housing, Communities and Local Government – gave the following advice to local authority leaders: 'Stop translating documents into foreign languages: only publish documents in English. Translation undermines community cohesion by encourage [*sic*] segregation' (emphasis omitted).[36] Guidance of this nature, albeit historical, can have a lasting effect. This statement frames translation services not only as unnecessary but also as an instrument of social division, a rhetoric that can obstruct access to information and to essential public services.

In 2021, I sent freedom-of-information requests to UK local authorities with particularly linguistically diverse populations. These requests revealed that the official budget for translation and interpreting services in Slough, a town and local authority west of London, was 600 pounds per year from 2018/19 to 2021/22.[37] Slough has a population of nearly 160,000,[38] of whom over 20 per cent do not have English as their main language.[39] Actual expenditure with translation and interpreting during the period covered by my enquiry far exceeded the planned budget. Information sent to me by Slough Borough Council showed that in 2019/20, the council spent over 6,000 pounds on translation and interpreting.[40] Even accounting for central government requests for savings to be made in this area, the language services budget of the council was clearly insufficient.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:38:37, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009769372.002

The same type of financial pressure facing local authorities affects other parts of the UK public sector. Before issuing the savings guidance mentioned above, the UK government had decided to outsource court interpreting services as a way of cutting costs. This decision has been examined by several parliamentary investigations that criticised the working conditions and levels of pay offered to interpreters.[41] According to lawyers and language professionals, outsourcing has downgraded the quality of interpreting available in UK courts. Significant misrepresentation in a trial has in one case been blamed on the outsourcing model.[42]

The low level of financial support for language services is likely to increase reliance on machine translation in public service settings. At the time of writing, versions of the Slough Borough Council website in languages other than English are unsurprisingly machine-translated,[43] as are the websites of police forces,[44] hospitals[45] and GP surgeries.[46] These uses of machine translation will have arisen from decisions taken by individuals, but it is important to note that in all these cases the technology is organisationally deployed. It is integrated into official information outlets that are directly controlled by government departments and public service providers.

The difference between individual and organisational uses of machine translation matters for several reasons. For one thing, the standards expected of organisational and of personal communication will usually be different. For example, if someone is involved in a traffic accident in a foreign country and a passer-by decides to help the victims using the Google Translate phone app, the help provided will not usually be subject to any type of official standard. There will be no formal expectations of how effective the communication needs to be, and most people would be grateful that a stranger decided to help and were making their best effort to communicate using the app. When the emergency services arrive, however, if they were also to use a machine translation tool, the situation would be different. If machine translation errors in that case were to have a negative effect on those involved in the accident, like the risks faced by the nurse mentioned in the Introduction, there would be broader questions about the standard of care provided, any language accessibility policies that may or may not have been in place, or more generally the extent to which the emergency services should be held accountable for mistranslations and their consequences.

A second reason why the difference between personal and organisational machine translation use is important concerns the way in which the technology is presented to the translation consumers. When AI translation tools are integrated into official websites or used by public service professionals, the technology is presented with the legitimating authority of the institution. Civilians or members of the public may be less likely to question the trustworthiness of the translations in that case.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:38:37, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009769372.002

In a chapter of the *Routledge Handbook of Translation and Migration*, I argue that keeping personal and organisational uses of machine translation separate is an important aspect of ensuring responsible deployment of these tools.[47] The appeal of efficiency – as a financial imperative and a driver of user satisfaction – will be an important marker of communicative contexts where translation tools are a first-line method of overcoming language barriers. Front-line professionals resort to these tools of their own accord when the need arises, an instinct that needs to be anticipated and appropriately discussed. When in these cases organisations fail to have official guidance in place, which is often the case,[48] they risk assenting to any uses of machine translation by omission instead of giving the delicate risk-benefit ratio of these tools due consideration.

Migration

One context where organisations tend to overlook the complexities of machine translation is migration. Migration is closely intertwined with uses of machine translation in public services, where language barriers can arise in interactions between institutions and newcomers to the country. On the one hand, migrants are often portrayed as a group for whom machine translation is particularly beneficial. Technology developers proudly show how their tools can be useful to refugees,[49] and refugees themselves will often refer to Google Translate as an essential tool in the process of integrating into new communities.[50] On the other hand, public institutions, especially immigration departments, can latch onto the availability of machine translation in ways that can negatively affect migrants' well-being and jeopardise their immigration statuses.

A recent example of institutional reliance on machine translation of this nature involves the UK's attempt to address an unprecedented backlog of applications for asylum.[51] Asylum applications usually involve a formal interview, but to speed up processing the UK government decided to request applicants from certain countries to complete a written questionnaire instead. In 2024, the questionnaire route was used for applicants from Afghanistan, Eritrea, Libya, Syria, Yemen and Sudan.[52] The questionnaire was only available in English, and only English could be used to complete it. Those who did not have sufficient knowledge of English to understand or complete the questionnaire were therefore faced with a problem, to which the government proposed a few solutions. Instructions to be provided to applicants included the following passage found in a template dated 23 February 2023:

> If you do not speak, write or understand English, you can use online translation tools. What you provide us with doesn't need to be perfect: we can always ask you for further details either by writing to you or ringing you to collect further information, or by

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:38:37, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009769372.002

inviting you to an interview. A friend who does understand English can also assist you to explain why you are claiming asylum, but they must not provide you with immigration advice.[53]

The proposed methods to overcome the language barrier are therefore online translation tools and seeking help from English-speaking friends. The UK Minister of State for Immigration at the time provided similar instructions in a written response to public concerns about the questionnaires. The response, dated 2 May 2023, included the following statement:

> Where the claimant is unable to read and or write in English then they could choose to seek support with understanding the language through community links such as charity organisations, non-governmental organisations, friends, family, online translation tools and other networks.[54]

The instruction to use machine translation – that is, online translation tools – is not unreasonable. It is important to note that applications received through this streamlined route could not be rejected based on the questionnaire alone. If the government was satisfied by the information provided in the questionnaire, claims could be approved. If potential issues were identified, the next step was to enquire directly with the applicant or ask them to attend an interview.[55] On first impression, therefore, it may look like there were sufficient safeguards in place to mitigate the risks posed by machine translation since translation errors could not directly lead to a refusal.

But the reality faced by asylum applicants was quite different. For those whose lives may literally depend on the outcome of these claims, using machine translation tools may seem like too high a risk to take. The languages spoken by individuals who were requested to complete the questionnaire can be particularly difficult for machine translation to get right. Languages that are not well resourced in terms of textual data are more likely to be ill-served by machine translation technology since machine-learning models will have less material to learn from (see Chapter 6). Machine translation research is attuned to this issue and efforts to optimise the technology for low-resource languages are ongoing.[56] But the fact remains that social and linguistic inequalities are built into these tools as a result of how they are developed. Some languages are not supported by the technology at all. At the time of writing, Bedawiyet, Bilen, Kunama, Nara, Saho and Tigré, some of the languages spoken in Eritrea,[57] are not available in Google Translate.

Even if their language was not supported by machine translation and they could not use English or another language that could be machine-translated to English, applicants could ask community members for help, as suggested by the government. But here too there was a problem. In the UK, only individuals with specific professional credentials can provide immigration advice. It is a criminal offence to advise applicants without these credentials,[58] and the

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:38:37, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009769372.002

distinction between helping someone to complete a form in English and advising them on their application is at best a difficult one to define. Research by the Human Rights Law Centre at the University of Nottingham has shown that applicants have been approaching immigration charities in the UK for help with translating their responses, but charities may refuse help to avoid the risk of being accused of providing unregulated immigration advice.[59] Asking asylum applicants to use machine translation therefore involves more complexities in practice than it may seem to in theory, which is all the more problematic given that claims could be withdrawn if the questionnaire was not submitted within thirty days (plus a ten-day grace period).[60]

The assumption that machine translation is an easy solution to the language barriers facing UK asylum seekers is a good example of how personal and organisational machine translation use can get conflated. Here, the use of the technology is outsourced to applicants when in fact it stems from an institutional decision to not consider applications in languages other than English. The onus of overcoming the language barrier is entirely on the applicant. The government is not willing to accept its own guidance and machine-translate questionnaire responses from other languages into English.

Direct uses of machine translation by immigration departments are in any case also problematic. In a case reported by the *Guardian* in September 2023,[61] an individual coming from Brazil spent six months in an Immigration and Customs Enforcement facility. He spoke Portuguese but could not read or write it. Staff at the facility, who did not speak Portuguese, used speech translation to communicate with him, but the tool was unable to recognise his regional accent. He did not understand why he had been taken to the facility. He had also been separated from his family and did not know where they had been sent. Other detainees tried to help him to prepare a written application using machine translation, but the tool did not do a good job of it – for example, Belo Horizonte, a Brazilian city where he had lived, was translated literally as 'beautiful horizon'.

In the United States, machine translation is now part of routine border checks. Customs and Border Protection (CBP) staff – the officers who speak to travellers at the border – can use a tool called CBP Translate. The tool is powered by Google, but it has its own interface and has been developed specifically for CBP use.[62] Officers can access the tool at their workstations or on smartphones issued by the government.[63] The tool transcribes travellers' and officers' speech. Translations of the transcript are then displayed on screen and can also be played out loud if an audio version of the content is available.[64]

On first impression, using a tool of this nature for border checks may again seem like a good idea. Routine questions asked at the border tend to be simple and easy to answer, such as where travellers are coming from, how long they intend to stay and the purpose of their visit. The US government has also put measures in place to work around some of the risks posed by the tool. Officers

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:38:37, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009769372.002

are not allowed to use the tool to collect sworn statements[65] or to decide whether a law enforcement action is required.[66] Human translations need to be used for these purposes. Additionally, travellers need to consent to the use of the tool. Officers display a message in travellers' own language explaining the purpose of the tool and saying that travellers can choose to use a human interpreter instead.[67] Waiting for a phone interpreter can delay entry to the country, so assuming that the tool is fit for purpose – even if not completely accurate – it can be beneficial for travellers as well as for the CBP.

As is usually the case, however, the deployment of the tool involves more nuanced shortcomings. The message presented to travellers fails to give any warning about the possibility of mistranslations. The text of the message reads as follows:

> Thank you for using the Translation Application – this mobile device allows proficient communication with U.S. Customs and Border Protection (CBP) Officers in your native language. Through this application, voice data, transaction information, and transcripts of conversations between you and the CBP Officer will be collected and retained according to the published retention schedules. You may decline to use this application and request a human translation.[68]

As discussed previously, machine translation works well in some cases and not so well in others, and the factors that can affect translation quality are extremely varied. The promise of 'proficient communication' conveyed in this message is therefore problematic.

The Department of Homeland Security states that 'CBPOs [CBP officers] are not required to use CBP Translate and can use other options, such as a phone translation service, if they believe the translations are inaccurate.'[69] While some inaccuracies may be obvious – and while admission to the United States cannot be denied based solely on this tool – if the tool is required to begin with, it follows that the officers' ability to identify mistranslations will at best be limited. So not only is official information provided to service users potentially misleading, but also on-the-fly evaluations of the technology rest on individuals who are not well placed to judge its effectiveness. Here again efficiency takes priority.

Google Translate Does Not Work Alone

On 3 November 2022, members of a police force were patrolling a bus station in Omaha, in the US state of Nebraska. The group, including Task Force Officer (TFO) Jaworski and Special Agent (SA) Iten, were in plain clothes but were wearing their body cameras and carrying their guns and badges. A bus coming from Denver and heading to Chicago stopped at the station for a change of drivers. All passengers had to disembark for the change, but they could leave their luggage on board. A twenty-three-year-old man who had got off the bus

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:38:37, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009769372.002

caught TFO Jaworski's attention. The man had made multiple trips to the toilet and was behaving in a way that was considered suspicious. TFO Jaworski approached the man, presented his badge and proceeded to communicate in English, but the man, a Spanish speaker, could not speak English. TFO Jaworski did not speak Spanish and decided to use the Google Translate app on his phone. He used the app's conversation feature, which transcribes what is said, displays a translation on screen and allows an audio version of the message to be played out loud in the target language – a synchronous, predominantly spoken use of the tool.

The use of Google Translate initially went well. TFO Jaworski told the man he was not in any trouble and then asked a few basic questions, such as where the man was coming from and going to. At one point there was confusion around how long the man would stay at his destination. According to Google Translate, the man first replied, 'few weeks, a few days', but the man subsequently clarified it was days rather than weeks. SA Iten spoke Spanish, so TFO Jaworski decided to ask his colleague for assistance. SA Iten entered the conversation at that point and started to interpret for TFO Jaworski. Even with SA Iten's assistance, there were communication issues at times, but SA Iten rephrased his words in Spanish when the man did not understand what was said. At one point SA Iten did not know the Spanish word for 'room' and asked his colleague for help. TFO Jaworski then turned to Google Translate again to ask the man about searching his bags: 'Would you want to do it in a private room out of the public?' SA Iten continued to communicate with the man in Spanish and asked whether he consented to having a sniffer dog check his luggage. SA Iten touched his nose when asking this question and kept checking if the man understood what was happening.

Later in the interaction, Google Translate was used again by TFO Jaworski and then by a third officer. The team had to seek a warrant to search the man's luggage. They eventually discovered that the man was carrying fentanyl in his suitcase. The man was charged with knowingly possessing a controlled substance with the intention to distribute it. Like in the law enforcement example mentioned in the Introduction, the man attempted to have the evidence suppressed in court. Among the allegations made by the defence, it was argued that the affidavit that supported the search warrant implied that the man had been evasive when answering the officers' questions. According to the affidavit, the man said that he would stay at his destination 'a few days to a few weeks'. This phrase appeared in the affidavit without quotation marks, potentially as a description of what had been said rather than as a verbatim quote. According to the man, it was Google Translate that caused the confusion about the duration of his stay.

The events above are described in *United States v. Gonzalez-Moreno*.[70] The motion to suppress the evidence was denied. The analysis by the judge stated that 'Using the Google Translate app or SA Iten speaking Spanish, officers had

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:38:37, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009769372.002

a conversation with Defendant during which they all appeared to understand one another.'[71] In this case, therefore, the use of machine translation was legally deemed successful. But further to this legal outcome, this case illustrates the range of strategies that are often used in this type of interaction. Google Translate was not the only communication method at play in the exchange between the man and the officers. Google Translate and SA Iten complemented each other. Like in many face-to-face conversations, the interaction was further supported by attempts to rephrase what was said, by contextual information, and by non-verbal cues, as researched widely in the fields of conversation analysis[72] and communication accommodation theory.[73]

In conversation analysis, the process of correcting or editing what is said to ensure understanding is called a repair.[74] Seeking clarification about answers and rephrasing questions are all examples of repair-type techniques that the officers employed. The officers' attempts to adjust their way of communicating with the man are also consistent with the concept of convergence in communication accommodation theory. Convergence, according to this theory, involves adjusting one's communication to become more like that of one's interlocutor,[75] which in the example above arguably involved the officers' attempt to use Spanish in the first place.

Interpreting research describes specific strategies that can be employed to check understanding in situations like this one. For instance, one of the interactants may be asked to repeat in their own words what they understood of what was said to them.[76] In policing contexts where it may be hard for civilians to repeat legal statements in their own words, it may be more appropriate to split the message in small chunks and check understanding of each chunk separately.[77] While it goes without saying that the officers above did not perform at the level of professional interpreters, they did use their own intuitive methods of adapting the communication in a way that in the eyes of the court was enough to ensure understanding.

This interaction nevertheless raises important questions. Why did the officers not use a phone interpreting service? They had to wait for the warrant before searching the man's luggage, so it is reasonable to presume that they could also have waited for a professional interpreter. Especially for a widely spoken language such as Spanish, issues with interpreter availability would have been unlikely. Similarly, the officers could have relied on professionally provided pre-translated information. It would have been standard to use written material in Spanish to read the man his rights, for instance, but the officers did not do that, and the reason is quite mundane – they had 'misplaced' the Spanish version of the document:

> SA Iten stated he did not have a rights advisory form in Spanish and asked TFO Jaworski to use the Google translate app . . . SA Iten needed to continue with the rights advisory,

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:38:37, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009769372.002

but since he had misplaced his DEA [Drug Enforcement Administration] Spanish rights advisory form, SA Iten replied he could 'fake it' (i.e., make his 'best effort') to advise Defendant of his rights in Spanish without reference to the form.[78]

The officers' instinct to 'fake it' points to a broader approach to communication where the question of language is often an afterthought. Furthermore, while the officer's knowledge of Spanish weighed on the judge's decision to find in favour of the state, attempts to complement machine translation use in this way are not always successful. A review I co-authored in 2020 mentions the case of a United States immigration officer who attempted to interview a Mexican girl using Google Translate.[79] The United Nations report that describes this interaction explains that the officer used Google Translate to translate questions from a written form. The form was in English and the officer tried to ask the questions out loud in Spanish by reading machine translations provided by Google.[80] The report describes that the translations 'often did not make sense and the officer's ability to read and pronounce in Spanish was very poor'.[81] The girl being interviewed had difficulty understanding the questions, and her replies, including details of health problems, were not fully understood by the officers.[82] While in some instances users may be able to at least in part mitigate machine translation deficiencies based on their own knowledge of the context and the languages involved, it is also possible that the user and the tool, rather than complementing each other, can exacerbate each other's deficiencies.

The consequences of using machine translation therefore depend greatly on the tool–user interaction rather than strictly on the tool. As will be discussed further in Chapter 2, relationships between humans and technologies are often socially constructed in such a way that human actions change the effects of the technology, and the technology changes the effects of human actions. A lot therefore depends on the user and their knowledge of themselves and of their own limitations, but also their knowledge of the tool and of its risks.

Literacies and Meta-Literacies

Being able to identify and mitigate the risks of any type of AI presupposes a level of critical familiarity with AI tools that can be treated as a type of literacy. Different definitions of AI literacy have been proposed in the literature. Most of them allude to competencies such as knowing what AI is and how it works, being familiar with its limitations and ethical implications, as well as knowing when and how to use it.[83] A widely used definition of AI literacy is the one proposed by researchers at the Georgia Institute of Technology, who call it 'a set of competencies that enables individuals to critically evaluate AI technologies; communicate and collaborate effectively with AI; and use AI as a tool online, at home, and in the workplace'.[84] Among sixteen competencies encompassed by this

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:38:37, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009769372.002

definition are 'recognizing AI', 'understanding intelligence' and 'ethics'.[85] The researchers at Georgia Tech also identified a series of design considerations that stem from these competencies, such as making design choices that increase transparency and promote the understanding of AI.[86]

The competencies outlined in the Georgia Tech definition are crucial to responsible deployments of machine translation, even if the proposed design considerations are currently largely absent from applications of the technology. Transparency, for instance, is not a feature of how machine translation is often used on public service websites, where its use is in many cases poorly indicated.[87] Promoting understanding of how machine translation works is not an inherent design feature of machine translation tools either. Nor is it common in institutional uses of the technology. The message displayed to travellers at the US border, for instance, can in fact do the opposite by promising proficient communication rather than explaining what machine translation is and acknowledging that it makes mistakes.

The description of the CBP tool and uses of other tools discussed so far show how crucial the concept of AI literacy is. And machine translation too – as a specific type or application of AI – has its own formulation of literacy. Literate machine translation users are described as being able to:

- comprehend the basics of how machine translation systems process texts;
- understand how machine translation systems are or can be used (by oneself or by other scholars) to find, read, and/or produce scholarly publications;
- appreciate the wider implications associated with the use of machine translation;
- evaluate how (machine) translation-friendly a scholarly text is;
- create or modify a scholarly text so that it could be translated more easily by a machine translation system; and
- modify the output of a machine translation system to improve its accuracy and readability.[88]

These competencies reflect many of the broad principles laid out by definitions of AI literacy, but here with a focus on academic writing. The academic focus is justified, among other factors, by the dominance of English in academia and the barriers this poses to those who need to read or publish research in English as a non-native language. While greater linguistic diversity may be preferable in academic contexts, machine translation can help scholars with low English proficiency to disseminate their work to a wider audience.

Academic communication is very different from the critical service contexts discussed in this book, however. The efficiency imperative is stronger in the contexts analysed here. Whether it be in border crossings, hospitals or police patrols, speed will often be important, which in turn often calls for spoken interactions and a seamless integration of machine translation into the conversation. Officers at the US border, for instance, can use the CBP tool to capture

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:38:37, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009769372.002

and translate speech as well as to machine-translate the content of physical documents using the cameras on their government phones.[89] This is a more dynamic use of the tool than using it for academic writing. Editing machine translations to improve accuracy – one of the competencies outlined above – may be possible for scholars with some knowledge of both source and target languages, but as mentioned previously would at best be challenging for public service professionals without that kind of multilingual skill.

Broader conceptualisations of machine translation literacy have therefore been proposed.[90] In the context of crisis communication, the concept of literacy has been associated with trust.[91] While this association comes with a call for more empirical research in this respect,[92] it can be presumed that the more users know about the technology and its limitations, the less likely they are to trust it uncritically. In crisis settings, or emergencies more generally, the relationship between trust and machine translation literacy is a complex one. On the one hand, trust in information is widely recognised as crucial to effective communication[93] If trust is compromised because of machine translation use, this could hinder communication even in cases where the machine translations are accurate. On the other hand, machine translation accuracy is never guaranteed, so individuals will do well to be cautious and to not trust the technology completely when the translations cannot be checked for accuracy via other means (see Chapter 5).

Technology itself can be used to check accuracy. Researchers have developed techniques that can be used to automatically estimate the level of quality of machine translations.[94] These techniques are based on methods that resemble those used to develop machine translation, namely by having machine-learning models identify patterns in existing data and then apply what is learnt to new content. These models provide estimated quality scores by analysing the translation itself and its corresponding source language version – for example, by considering the length of the message in each language, the type of words they include as well as other, more complex features of the content. Models trained to identify patterns in these features and how the features correlate with human judgements can be used to flag mistranslations. Large language models too can be asked to provide estimates of how accurate their translations are. Meta (formerly Facebook Inc.) filed a patent application in 2015 for a system that applies a quality estimation methodology to decide whether machine translations should be presented to users automatically. Only translations with estimated quality scores that cross a certain threshold are selected to be automatically displayed.[95] This type of scoring can be a useful tool, but like most technologies based on machine learning, quality estimation is not guaranteed to be accurate. To be used judiciously, it also requires AI literacy competencies.

As mentioned, using language professionals does not necessarily guarantee accuracy either. The skills required to overcome language barriers in the

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:38:37, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009769372.002

contexts discussed so far are therefore quite varied. Familiarity is needed not only with machine translation technology but also with translation itself as a service and process.[96] The skillset of literate machine translation users should also include elements of other types of literacy, such as information and data literacies.

An information-literate individual successfully 'recognizes the need for information', 'evaluates information', 'identifies potential sources of information', 'develops successful search strategies', among other competencies.[97] In relation to the presence of AI tools in digital spaces, this type of literacy increasingly concerns the ability to recognise the need for, and to identify, 'the truth'.[98] As AI technologies are increasingly capable of altering the content of texts and images and of reproducing individuals' voices, being able to identify misinformation becomes harder and more dependent on acute critical skills. In relation to machine translation specifically, content consumers would do well to evaluate the information they are given and consider the range of methods that might have been used to provide it even if these methods are not overtly described. End users or consumers of machine translation would also do well to engage critically with their sources. Even when these sources seem reputable, they may have different priorities or be ill-advised, whether it be government institutions that instruct asylum applicants to use machine translation or official messages that are supposed to be informative but which overstate the technology's capabilities.

Data literacy competencies overlap with many of the AI, machine translation and information literacy skills outlined above. Data literacy has been defined 'as the component of information literacy that enables individuals to access, interpret, critically assess, manage, handle and ethically use data'.[99] This definition underpins the data literacy component of the AI literacy framework proposed by researchers at Georgia Tech.[100] The focus of this definition is on direct experiences of data use, but knowledge of what can be classed as data and of how it is used is also central to machine translation literacy.

Awareness of the privacy implications of using machine translation systems (see Chapter 5) is a particularly important aspect of being familiar with the concept of data. Machine translation tools, especially those that are available free of charge on the internet, may process users' data in ways that they do not realise. Textual data provided to a machine translation tool as the input to be translated has been found openly available on the internet before.[101] Depending on the system, users also compromise their intellectual property rights by sharing content with the tool.[102] Some of these possibilities may be explained in the tools' terms of service and/or privacy policies, but given the usual length and complexity of these documents users are unlikely to study these policies in detail.[103]

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:38:37, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009769372.002

My research on how machine translation users approach the matter of privacy shows that users may in fact favour machine translation use over sharing information with a language professional.[104] Yet if data is overshared with machine translation tools by accident – for example, by uploading the wrong image to a machine translation app – users quickly realise the potential reach of the technology over their private lives.[105] Understanding the value of data is therefore crucial to interactions with machine translation systems, particularly for public service professionals who will be bound by expectations of confidentiality.

Many of the competencies of AI, machine translation, information and data literacies are about knowing what one needs to know about. Information literacy, in particular, can be framed as a type of 'meta-literacy'.[106] In machine translation use contexts, meta-literacy concerns knowledge not only of machine translation itself and of how to use it, but also knowledge of the different types of knowledge required to judge the technology's risk-benefit ratio for different scenarios. Being machine-translation literate – and thereby information and data literate – is thus an ongoing learning process that involves being sensitive to the challenges of different communicative circumstances but also to how the technology itself evolves over time. These literacies and meta-literacies are essential for the discussion provided in this book. I come back to them in passing at different points in the book and in more detail in Chapter 5.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:38:37, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009769372.002