5

Risk management for content delivery via raw machine translation

Mary Nurminen and Maarit Koponen

Introduction

Technological developments in machine translation (MT) have led to an ever-increasing use of MT as an aid in communicating across languages. In content delivery, MT is most commonly used in a post-editing (PE) process, where a professional translator checks and edits the MT (see, e.g. Guerberof-Arenas 2020). Parallel to this, MT can also be used in its raw, unedited form to help people interact with either content or other people across language barriers. In addition to people using free online MT tools, companies and public organizations may also provide machine-translated content. An organization might choose to deliver information as raw MT so that they can offer more content in more languages and reach new audiences while saving on time and costs. For example, some technology companies have aimed to broaden the range of languages in which they offer product support information by delivering the content in some languages as raw MT (Thicke 2013). In another example, the field of intellectual property rights (IPR) has relied on raw MT for more than a decade (Nurminen 2019) to enable patent professionals to identify relevant patent documents in languages they do not speak themselves.

When MT is used for delivering content, it is important to manage the risks involved. Recent literature (Vieira, O'Hagan and O'Sullivan 2020; Scott and O'Shea 2021) has highlighted the consequences associated with the use of raw MT. Other studies have discussed risk management, although mainly with a focus on scenarios where MT is post-edited (Nitzke, Hansen-Schirra and Canfora 2019; Canfora and Ottmann 2020). Scenarios where raw MT is used, however, entail specific risk considerations. Despite the recently reported improvements of neural MT systems, the possibility of inaccurate translations still exists, and translation errors may entail high risks. At the same time, the use of MT should be balanced against the potential consequences of not translating the information at all, as argued by O'Brien (2020: 313). In this way, delivering information via raw MT can be considered not only a risk but also a mitigation of the higher-level risk of people receiving no information. To better understand the contexts where delivering information via raw MT would be effective, and to manage the potential risks and mitigating factors, a more nuanced understanding would be needed about how people react to and consume raw MT. However, research and discussions of best practices in this area are currently lacking.

This chapter aims to investigate how risk management principles can be applied to scenarios where information is delivered to end-users of products and services as raw MT. We examine risk factors for assessing the use of raw MT, as well as options for managing and mitigating the risks. We contend that a simple checklist of rules that would lead to minimal or no-risk delivery of raw MT content is not possible, and therefore do not aim to propose one. Rather, our goals are to provide general considerations for assessing and managing the risks involved in content delivery via raw MT and to contribute a new viewpoint to existing risk management models in translation. Besides contributing to academic discussions, this can offer a concrete tool for organizations considering using raw MT to communicate across languages. The chapter first provides an overview of risk management principles based on International Standard ISO 31000:2018 and related work on risk management in the context of translation. We then apply the risk management framework to the delivery of content as raw MT, before illustrating these considerations through a concrete use case for raw MT in the patent world. The final section presents the conclusions of this analysis.

Related work: Risk management

Key concepts and framework for risk management

A commonly used risk management framework is the International Standard ISO 31000 (ISO 2018), which defines principles, key concepts and processes. The central concept of risk is defined as the 'effect of uncertainty on objectives' (ISO 2018, term 3.1). The uncertainty arises from a given risk source (term 3.4) and can lead to positive or negative consequences (term 3.6). Risk management according to the ISO standard involves a systematic, iterative process integral to an organization's operation. This process should be customized to each organization's internal context, which relates to the values, strategies and culture of the organization itself, and the external context, which involves social, regulatory, technological and economic factors among others (ISO 2018: 10). The standard also emphasizes the need to account for stakeholder needs and perspectives through active communication and consultation. The overall risk management process of ISO 31000:2018 is illustrated in Figure 5.1.

Within the risk management process, the ISO standard defines risk assessment to encompass risk identification, where the organization examines both tangible and intangible risk sources and consequences, followed by risk analysis to understand the nature of a risk and the likelihood, nature and severity of consequences. Risk evaluation then compares the analysis results to risk criteria, and, based on this evaluation, the organization can decide to take the risk, to avoid the risk by not starting or continuing an activity that causes risk, or to employ some risk treatment measures (see ISO 2018). The standard emphasizes that decisions should not be made on an economic basis alone but should also consider organizational obligations and the diverse stakeholder perspectives.

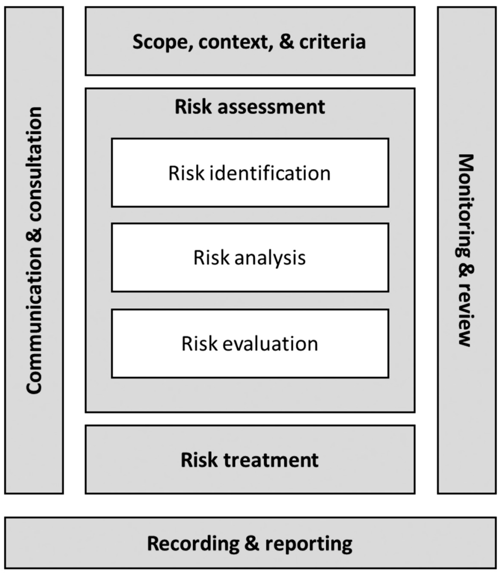

Figure 5.1 Risk management process according to ISO 31000:2018, figure 4.

Reporting and monitoring risk treatment is also important to dynamically address new and emerging risks, and to continually improve the process.

Translation, MT and risk

Relatively little research has been conducted on risk management in the context of (human) translation, but in recent years this topic has received increasing interest. Risk management has been applied at different levels and stages, from an individual translator's management of uncertainty and risk mitigation (see Pym 2021 for an overview) to risk management in terms of business considerations and project management (e.g. Dunne 2013; Zaveckaite and Ulbinaite 2018).

A common view of risk is as something negative, and although risk-taking can also lead to positive outcomes, risk management tends to focus on potential negative consequences and mitigation (see Zaveckaite and Ulbinaite 2018: 1325–6). Translation errors are a central source of risk that can affect 'the author, translator, translation agency, client or end user of the translation' (Canfora and Ottmann 2018: 168), and previous work has examined risks and potential consequences of translation errors in safety-critical situations (Canfora and Ottmann 2018; 2019) or in conjunction with commercial, legislative, normative and political texts (Byrne 2007; Cismas 2010). Potential consequences are discussed further in 'Risk analysis' section.

Recently, risk related to the use of MT has received increased attention. Canfora and Ottmann (2020: 59) note that although many of the risks involved are similar to risks in human translation, some are more specific to the use of MT and entail different considerations. These include potential liability and cyber risks, such as the potential exposure of confidential information. As with Canfora and Ottmann's focus on translation in safety-critical contexts (2018, 2019), several studies on MT suggest that a greater consideration of risk is necessary when the texts involved in MT are from high-risk areas. Way (2013) describes the use of raw MT as risky for translating stock reports because of the high financial risk. Both Nurminen and Koponen (2020) and Vieira, O'Hagan and O'Sullivan (2020) cite public health and medicine as high-risk contexts, and Haddow, Birch and Heafield (2021) argue that post-editing of MT is always needed in these settings. Risk in legal contexts has been discussed by Vieira, O'Hagan and O'Sullivan (2020), Scott and O'Shea (2021) and Guerberof-Arenas and Moorkens (2023).

On the other hand, researchers have also pointed out that some environments in which MT could be used might be somewhat resistant to risk (e.g. Way

2013). Nurminen (2021) lists risk-tolerant environments as a contextual factor that positively influences if and how MT gisting is used, specifically citing the patenting process as one such an environment. She describes the paradox of risk tolerance in the environment: since patenting has a wide variety of inherent risks, affordances for mitigating those risks have been built into processes. Those same affordances help to provide an environment which is tolerant to the risks involved with relying on raw MT (2021: 94–5). This is discussed further in 'Risk management in action: The case of patent professionals' section.

A few models have been proposed for classifying different use cases as more or less appropriate for raw MT. Guerberof-Arenas and Moorkens (2023) propose a model for classification that relies on two factors: the shelf life of the content to be translated and its riskiness. They offer a number of suggested use cases that fit into one of four categories: short term/low risk, short term/high risk, long term/low risk, and long term/high risk. However, shelf life and risk seem to be separate factors that are simply put into the same model rather than factors that influence each other. A more comprehensive model for risk assessment is offered by Nitzke, Hansen-Schirra and Canfora (2019: 240), who argue that such a model can be employed for both strategic decisions (whether to use MT at all) and operative ones (whether to use MT for a specific purpose or with a specific text). Their model includes factors related to technology, the sensitivity of the text to be translated, the number of people who might read the translation, the level of quality required of the translation and the number of translators and time available for human translation.

A few studies discuss risk treatment strategies, such as introducing guidelines and legislation (Canfora and Ottmann 2020; Martindale 2020), developing technology that enables risk mitigation (Canfora and Ottmann 2020; Martindale 2020) and promoting MT literacy among users (e.g. Bowker and Buitrago Ciro 2019; Martindale 2020). Besides these articles, however, past research on risk treatment tends to emphasize mitigating risk through post-editing and quality control by human translators.

A different approach to the question of risk and MT gisting is taken by Nurminen (2021: 140–1) in her proposal that the practice of MT gisting could be conceptualized as an exercise in risk management. She maps the risk assessment and treatment processes discerned in her study of patent professionals to ISO 31000 (see previous section) to demonstrate how the processes conform to the standard's description of risk management. She then proposes that conceptualizing MT gisting as risk management offers us a framework for analysing people's use of raw MT, an idea that was first proposed by Pym (2021: 453): 'The categories of risk management invite studies that assess the strategies of all participants in a translation event and use that matrix to try to explain translator decisions.' Nurminen concludes that MT literacy programmes could benefit from including risk management in their syllabi. It would encourage end-users of raw MT not only to acknowledge the risk but also to realize that the risk can be managed through concrete actions.

Risk management considerations for raw machine translation

In this section, we apply the ISO 31000 risk management framework to content delivery via raw MT. Risk is always present in this context: there is no guarantee that a text will convey the meaning intended. The same naturally applies to human translation, but the revision and review controls commonly used to minimize risk in human translation workflows are not available with raw MT. Thorough analysis is therefore required if the use of raw MT is considered, and this section aims to support that analysis. Our discussion focuses mostly on the 'heart' of the process: scope, context and criteria; risk assessment; and risk treatment. It is here that issues that are specific to the delivery of information using raw MT are likely to arise. The other processes – communication, monitoring and reporting – are equally important. In particular, it would be important to implement processes to collect feedback from end-users or internals who notice issues themselves or encounter them in customer comments. For example, Canfora and Ottmann (2018) propose a process for monitoring and analysing incidents related to human translation, and a similar process could be applied to the raw MT situation. Other aspects of risk communication, monitoring and reporting are likely already covered by risk management processes in the organization.

We make certain background assumptions in our discussion. First, the risk management considerations are appropriate for any type of organization, including private businesses, public administration or non-profits, so we use the generic organization to refer to them collectively. Second, the considerations presented are not intended to account for decisions regarding which MT tool to select. Instead, we assume that the organization has already evaluated and identified appropriate MT tools or will do so separately. Information on this process is available both in academic literature (e.g. Canfora and Ottmann 2020), and from commercial technology providers and consultants. We also assume that the MT tool selected would be trained on the organization's own material and would be in the organization's private use. This would offer better quality than generic MT solutions and ensure data protection and privacy.

Communication and consultation

A key part of the risk management framework is communication and consultation with both internal and external stakeholders. This entails involving people with different backgrounds and expertise in order to build a comprehensive and nuanced understanding of different types of risk, the situations where they can occur, and the potential consequences. Needs and expectations with regard to information may vary between stakeholders, as do understanding of MT technology and attitudes towards it. Furthermore, different stakeholder groups may have differing perspectives on risk levels and the acceptability of specific risk treatment options.

Internally, people or groups responsible for content creation and translation would logically be involved in risk management processes focused on MT and information delivery. Other relevant internal stakeholders include managers, people responsible for language technology, and groups who interact with end-users. Consultation can also involve external technology or language service providers or independent consultants, as suggested by Martindale (2020).

A vital external stakeholder group is of course the people to whom the organization intends to offer raw machine-translated information. The importance of this end-user perspective should be recognized, and the organization should seek to proactively involve them in the early stages of planning to start delivering information as raw MT, as well as in later processes like monitoring. Forms of involvement could include user studies and focus groups, such as those described in Bowker and Buitrago Ciro (2015) or Haddow, Birch and Heafield (2021).

Scope, context and criteria

One of the first steps in implementing risk management step is to establish the scope of the exercise. First the organization needs to determine what unit of content is assessed when considering delivery via raw MT. Assessing the suitability of raw MT for each document produced by the organization would probably be a cumbersome and fractured process. In contrast, different information types involve varying types and levels of risk, so assessing risk for all content collectively is not feasible. Therefore, we assume that the basic unit for risk assessment is a specific content set or information type, such as product support materials or internal announcements. Similarly, the decision model for MT use by Nitzke, Hansen-Schirra and Canfora (2019) uses text type as the assumed unit of analysis. The scope definition should also involve the number of intended target languages and the estimated number of people who would receive the information via raw MT. As noted by Canfora and Ottmann (2018: 172), the wider a translation is circulated and the more target languages are involved, the greater the likelihood of a serious incident (see also Nitzke, Hansen-Schirra and Canfora 2019).

To establish the context, the organization should evaluate external and internal, organization-specific factors that affect the identified risks. External factors include the organization's operating environment or industry; social, regulatory and technological factors; and general perceptions of information and risk. Awareness of and attitudes towards automation, MT and the implications of using raw MT are also important considerations. In some environments, information delivery via raw MT could be viewed negatively, while other target audiences might consider it a sign of innovation. Internal factors include the organization's operational culture and 'risk appetite' (see Nitzke, Hansen-Schirra and Canfora 2019: 241), specifically, its willingness to take risks in information delivery. Such factors can change over time, so analyses should also be updated periodically.

These various factors will inform the criteria for evaluating risks, determining risk likelihood and severity and deciding which risks can be accepted and under what conditions. An important consideration for risk criteria is identifying potential benefits. Being able to deliver more information via MT to more people, especially to those who currently are not offered information in their own languages, can mitigate risks posed by not providing information at all (see O'Brien 2020). The increased reach could also provide a better understanding of what content and languages end-users access the most (Dillinger and Gerber 2009: 11) or increase customer satisfaction and reduce support calls (Thicke 2013: 50). When the output is provided by the organization's own domain-specific MT solution, it can also offer higher-quality information and mitigate the risks posed by stakeholders' use of free online tools (see Canfora and Ottmann 2020). However, potential benefits must be carefully weighed against the risks (Nitzke, Hansen-Schirra and Canfora 2019: 246; Vieira, O'Hagan and O'Sullivan 2020: 12).

Risk identification

One of the most important phases of risk management is risk identification. However, this phase can be challenging because the practice of delivering content as raw MT is not yet widespread and best practices are lacking. In our viewpoint, the most important principle is to carry out the most comprehensive analysis possible. This can be achieved by systematically analysing potential risks through distinct categories. Besides providing a more comprehensive analysis, the approach also makes it easier to identify specific risks and allows people with diverse expertise to participate in analysing areas that are familiar to them without needing to participate in the entire exercise. The need for a holistic approach to risk analysis has been noted in past research as well. For example, Canfora and Ottmann (2018: 171) point out that translation is a complex activity and that serious issues are not usually caused by an individual translation error but result from a series of failures. We propose that analysis focus on the categories of technical, content and context of use.

As discussed briefly earlier, we also propose that a diverse group of people should be involved in analysing these four phases, and that carrying out four distinct analyses allows the appropriate parties to participate in the areas they are most familiar with. In an exercise focused on information delivery in different languages, it is logical that people or groups responsible for content creation and translation would be involved in all phases. In addition, the analysis of organizational or business risks could include managers or other people responsible for the organization's strategy, people responsible for risk management, and representatives from marketing or sales. The analysis of risks related to technology might include people who are responsible for language technology, the organization's other technology and external participants such as the provider of the organization's MT solution, language service providers or, as suggested by Martindale (2020), independent consultants.

The analysis of the content could include those responsible for language technology and representatives from internal groups that rely on the organization's content, for example, those who train customers and other stakeholders. It might also include external participants such as language service providers or external consultants. Finally, the phase devoted to identifying risks in the context of use should include people who study or interact with end-users, for example, people who train customers and other stakeholders, people involved in marketing and sales and members of the user interface team.

A key point in contexts where unedited MT is used is that MT output is not error-free, and although truly misleading translations are rare (Martindale and Carpuat 2018: 16; Martindale 2020: 31), MT will produce errors and predicting where they occur and how they impact the output is difficult. Therefore, errors are not risks but rather risk sources, and actual risk arises when someone relies on faulty information when taking actions and those actions lead to consequences.

The following sections examine possible risk factors related to technology, content and context of use in more detail. Based on prior research, we provide examples that are intended as guidance for identifying risks, not as definitive lists of all risk factors. The risks relevant to each organization are specific to its context, information and target audience and may involve factors differing from the examples given here.

Technological risk factors

Examples of technological risk factors include the quality of the MT output, the language pairs involved and cyber risks. The quality of MT output depends on several factors such as the text types and languages involved (see, e.g. Barrault et al. 2020 for an evaluation of multiple language pairs). Even after MT technology has been chosen and implemented, it would be important to analyse the quality of the MT for the specific information type and language pair an organization is planning to use for raw MT delivery.

A second issue with the quality of MT output is that it is unpredictable, even within one text: it can be good, it can contain clear and frequent errors that lead to confusing or incomprehensible translations, or it can be fluent and easily understandable but nevertheless inaccurate, particularly in the case of the currently predominant neural MT technology. In particular, 'believable output' that is both fluent and plausible can lead the reader to judge the information as credible (Martindale 2020: 25). Problems can then arise when inaccurate translations are trusted by readers. MT output quality also varies depending on the language pair. In addition to a potentially higher likelihood of risk due to more frequent errors, Vieira, O'Hagan and O'Sullivan (2020: 7) suggest that this variability could lead to the unequal treatment of the target audience, or at least the perception of inequality. Byrne (2007: 7) discusses problems caused by faulty English-to-Chinese human translation in a US election process and the ensuing perception that it was 'proof of the indifference of the federal government to the [Chinese American] community'.

The use of MT also involves cyber risks (e.g. Canfora and Ottmann 2020; Nitzke, Hansen-Schirra and Canfora 2019) which centre around the use of external systems that are not secure and may expose the organization's non-public information to others. This cyber risk is lessened if the organization employs a secure MT solution and the machine-translated information offered to external parties is already publicly available.

Content risk factors

Another set of risk factors involve the content or information type that an organization is considering delivering as raw MT. The genre, subject matter and content creation processes should be examined.

The first concern is whether the subject matter involves safety-critical information. Information meant to be used in situations where the safety of the end-user or others may be endangered is particularly risky due to the potentially severe consequences (see Canfora and Ottmann 2018; Vieira, O'Hagan and O'Sullivan 2020). As an example, Byrne (2007: 17) discusses a case where a translation error in a product manual led to unsafe behaviour when the product malfunctioned. Safety-critical information types would therefore be unsuitable for delivery via raw MT (Nitzke, Hansen-Schirra and Canfora 2019; Canfora and Ottmann 2020). Other content types less suited for MT would include information with potential financial, legal or political ramifications (see Byrne 2007; Cismas 2010; Vieira, O'Hagan and O'Sullivan 2020).

A second factor concerns whether the genre is conducive to producing high-quality MT output. Although evidence on which genres lead to better MT output is somewhat unclear, creative texts such as literature, marketing and advertising are commonly suggested as poor candidates for MT (see Nitzke, Hansen-Schirra and Canfora 2019 or, for an industry perspective, Densmer 2020). Conversely, content types in which translations are required to follow the source text strictly would be more suitable for MT.

A related factor is the content creation process. In general, content that is created in a standard way is considered to be better suited for MT than content that varies greatly (Nitzke, Hansen-Schirra and Canfora 2019). In contrast, using raw MT to deliver content authored by a variety of people using non-standard methods is riskier. Nitzke, Hansen-Schirra and Canfora (2019: 243) outline a general rule that texts (and processes) suited for translation memories are also good for MT if they do not contain other risks.

Contextual risk factors

To date, little attention has been devoted to the contexts in which MT is used, although research suggests that context influences how people use and receive raw MT (see Nurminen, 2021). Relevant contextual factors include the size of the intended group of end-users; their qualities, such as familiarity with the subject matter and MT; their goals and processes when using raw MT; and factors related to the environment in which MT would likely be used, including access to possible auxiliary tools and materials.

The user's knowledge of the subject matter can positively impact the reception of raw MT (Nurminen 2021), whereas a low level of familiarity can make MT gisting more challenging. For example, Nitzke, Hansen-Schirra and Canfora (2019: 245) suggest that customer service agents could have sufficient background knowledge to understand a question posed via raw MT messages, but the customers who lack the same knowledge might have difficulty understanding the agents' machine-translated replies.

Another factor is the users' MT literacy, meaning their knowledge regarding how MT works, how it can be used and what the implications of using it are (Bowker and Buitrago Ciro 2019). A low level of MT literacy could make the users more prone to trusting MT uncritically, while a better understanding of the technology can help them recognize potential errors. On the other hand, a low level of MT literacy could be associated with a negative perception of MT and cause the users to not accept MT use at all.

Contextual factors like the user's goals are also important. If the user of the raw MT will act on the information, for example, install an electronic device, then defective translations in the user manual can lead to unsafe situations (see Cismas 2010: 493; Byrne 2007: 17). Contextual factors may offer additional considerations for safety-critical content. For example, a text of medium safety risk might be considered less risky if the target audience has a high level of expertise in the subject matter. Caution is required in such considerations, however, because the audience and their level of expertise may not be fully known to the organization (cf. Pym 2021: 451–2). The users' goals also impact how acceptable they find raw MT. For example, users may be more accepting if the purpose is to find practical information about services than if the information somehow impacts the rights of a language minority (see Bowker and Buitrago Ciro 2015).

Environmental factors such as access to other resources also affect the users' means and motivation to verify raw MT output when they encounter an unclear passage or suspected error (see Martindale 2020). The easy availability of auxiliary language resources like glossaries or alternate translation versions – or a colleague who speaks the source language – could help verify unclear passages. Higher risk is involved if the end-users of the raw MT are assumed to work in an environment without such verification aids.

Risk analysis

Each identified risk is then analysed in terms of likelihood, potential consequences and their severity. As noted, the main risk that someone may use incorrect information delivered via raw MT and that may lead to consequences. To define likelihood levels and types of consequences, the organization could apply a risk matrix like the one outlined by Canfora and Ottmann (2019).

Likelihood of occurrence

The likelihood of a risk (in terms of an incident leading to some consequences) can be categorized using different levels. For example, Canfora and Ottmann (2019: 83) use the likelihood levels of very likely, probable, occasional, remote and improbable. Estimating the likelihood of translation-related risks is challenging because, beyond anecdotal reports, clear and reliable evidence of faulty translations leading to severe consequences are difficult to find (Canfora and Ottmann 2018; Byrne 2007: 3). Canfora and Ottmann (2018: 169) note that this lack of information can be dangerous due to a human tendency to underestimate the likelihood of (negative) incidents and ignore incidents with low likelihood but potentially serious consequences.

As discussed earlier, likelihood of risk may be higher for some texts and information types. Additionally, even within one text, certain parts might carry higher risk than other parts (see Pym 2015). Another factor is the MT output quality. Frequent MT errors increase the likelihood of faulty translations leading to consequences. Conversely, fluent and believable MT output may increase the likelihood that the user places unwarranted trust in the output. Finally, the exposure of the translation needs to be considered, as the size of the audience and number of languages affect the likelihood of an event.

Possible consequences

As in the case of human translations, consequences arising from faulty MT output can involve physical harm like injury and death, legal consequences, property damage, financial damage, embarrassment, damage to reputation or trust and impaired or prevented communication (see Byrne 2007; Cismas 2010; Canfora and Ottmann 2018: 171; Canfora and Ottmann 2019: 81–2; Vieira, O'Hagan and O'Sullivan 2020: 10; Scott and O'Shea 2021: 26–7). Various consequences are also tied to each other in that the same incident can cause more than one consequence, and one consequence can lead to others.

Impaired or prevented communication with the target audience could be caused, on the one hand, by MT output that is difficult to understand or contains clear errors. Conversely, impaired communication can be caused by output that is easy to understand in that it is fluent, but it nevertheless contains errors. Neural machine translation, which is currently the predominant MT technology in use, can produce output that sounds fluent but that contains errors (see Martindale and Carpuat 2018). Both Martindale and Carpuat (2018) and Rossetti, O'Brien and Cadwell (2020) found that fluent MT output engendered trust among MT users. Martindale (2020: 25) identifies 'believable output' as a risk and defines believability as a combination of fluency, plausibility and human judgement of credibility. Problems can arise when texts contain translation errors, yet they are trusted by readers.

When someone uses information that is incorrect due to translation errors, they may incur property damage, financial damage or physical harm, which in turn can result in questions of liability and legal consequences for the organization that provided the information. As noted by several researchers (e.g. Canfora and Ottmann 2020; Vieira, O'Hagan and O'Sullivan 2020) currently legislation is just starting to react to liability questions concerning MT, leading to uncertainty in predicting and planning for potential consequences. Some laws and regulations concerning human translation do exist (for a more detailed discussion, see Byrne 2007), but it is unclear whether or which of these might also apply in cases in which information is translated by a machine.

The final example risk could also be viewed as the consequence of all the risks listed above. Impaired communication, damages to property or body, and questions of liability can cause embarrassment and loss of reputation of the organization. As illustrated by Zaveckaite and Ulbinaite (2018: 1336), although language and translation are often seen as less important support features to the main product or service, they have direct impact on how the user perceives that product or service. If the end-user finds an error in the translation (see Pym 2015: 78), loss of credibility of the translation can potentially also affect the credibility of the organization offering the information in the raw MT scenario (see Vieira, O'Hagan and O'Sullivan 2020). In addition to a loss of reputation due to concrete issues that end-users notice, an organization's reputation might also suffer from a general negative attitude towards MT. For example, Baumgarten and Cornellà-Detrell (2019: 18) discuss a case where translation errors drew negative attention to the use of MT for translating subtitles in Catalonia, leading to backlash from the viewers towards the TV channel. Also, simply the knowledge that MT has been used may lead the users to perceive the content more negatively (Asscher and Glikson 2021).

Severity of consequences

Death and major physical harm are generally viewed as the most severe potential consequences. Martindale (2020: 22) labels the most severe consequences 'dangerous catastrophic failures' while Google designates high-risk content as YMYL or Your Money or Your Life (Google 2022: 26). Canfora and Ottmann (2018: 176) deem fatal or lifelong injuries as most severe, followed by other injuries, and then material or financial damage, legal consequences or damage to reputation. If the same incident leads to multiple consequences, severity should be assessed according to the most severe (see Canfora and Ottmann 2019). However, such severity categories are not straightforward and depend on situational factors. Impaired communication is often considered the lowest in severity, but in the voter information case discussed by Byrne (2007: 6), impaired communication had rather important political ramifications.

How severe the consequences will be is affected by the various factors. As noted previously, consequences can be particularly severe for safety-critical content (see Canfora and Ottmann 2018). Even for a given content type, consequences can vary from negligible to very severe depending on situational factors (see Cismas 2010: 493). According to Martindale (2020), potentially dangerous situations are characterized by 'believable' MT output, lack of means or motivation to verify the output on the part of the users, and use cases where the user will act on the information provided.

Risk evaluation

The risk analysis results are then evaluated against the risk criteria set earlier (see section above on scope, context and criteria). Based on the risk likelihood, potential consequences and their severity, as well as the organization's context and risk appetite, the organization can decide that some risks may be tolerated while others are not. Using a risk matrix (see Canfora and Ottmann 2019: 82–3), the organization might, for example, deem occasional risk with low-impact consequences like impaired communication to be acceptable, but even improbable risks with potentially catastrophic consequences to be unacceptable. Pym (2021: 446) likens such evaluation to distinguishing 'kittens' (risks with low likelihood and lesser impact) from 'tigers' (risks with high likelihood and severe consequences): one can live with kittens, but protective measures are needed with tigers.

Factors for this evaluation again include information type and content as well as contextual factors. As noted, raw MT is better suited to low-risk texts

(Nitzke, Hansen-Schirra and Canfora 2019: 242), and not advisable for safety-critical information, at least until a clearer understanding can be formed of its use and users. Evaluation might additionally include factors like resources and time available for translation and expected life span of the translation (see Nitzke, Hansen-Schirra and Canfora 2019: 244–5). However, decisions should not be based on time or financial aspects alone. Finally, evaluation should also balance the risks of providing raw MT against the risk that no translated information is provided and the risks posed by customers turning to generic online systems.

Risk treatment

Based on the risk evaluation, the organization can decide to avoid the risk, take the risk, remove the risk source, change the likelihood or consequences of risk or share the risk with other stakeholders. When risks are deemed unacceptable, the primary way to avoid risk is to not use raw MT for a specific text type or purpose. Conversely, if the risk level of raw MT is low enough to accept (cf. 'kittens' in Pym 2021: 446; see also Canfora and Ottmann 2019), the organization can decide to take the risk without additional measures. Removing the risk source could mean changing the process so that the machine translation is checked by humans, which differs from the use case in this chapter (see Nitzke, Hansen-Schirra and Canfora 2019 for details).

The organization can also decide that the risk involved in raw MT use is acceptable with some further treatment. The following sub-sections provide examples of measures categorized as changing the likelihood or the consequences of a risk or sharing the risk with other stakeholders. Effective risk management generally requires combining multiple measures at different levels (Canfora and Ottmann 2018: 179–80) and should account for user behaviour, such as the tendency to use default options (see Canfora and Ottmann 2020: 65). Many of the proposed measures could be implemented without a large investment of resources.

Change the likelihood or consequences of the risk

The likelihood of a risk can be reduced by helping users achieve the best possible understanding of raw MT content, and outcomes can be changed by aiming to ensure faulty information will not lead to severe consequences. To accomplish this, users could be encouraged to check an auxiliary source to verify a raw MT text, or in critical situations, even be required to do so (see Martindale 2020: 38). What follows are examples of treatment options organizations could consider.

Increase MT literacy among users. MT literacy can help raw MT users understand risks, recognize common errors and follow best practices (Bowker and Buitrago Ciro 2019; Nurminen 2019; Vieira, O'Hagan and O'Sullivan 2019; Martindale 2020). Organizations could embed information, tips and hints in the interfaces through which users access raw MT, or even offer training in MT use. Guidelines on using raw MT can also raise MT literacy and promote best practices.

Show the source document. People using MT often translate from languages in which they have some proficiency, and they tend to look at the source text when reading raw MT. This can contribute to their ability to understand raw MT (Nurminen 2019, 2020). Especially if the organization publishes content in a language in which the target audience might have some proficiency (e.g. English), users could benefit from having that language version easily viewable, either side by side with the raw MT output, or by hovering the mouse over the corresponding target passage.[1] Pictures and other multimodal elements of the source text also provide important information when using raw MT (Nurminen 2019, 2020). Some MT solutions translate full documents, showing the original multimodal elements in place (e.g. Microsoft; see Doss Mohan 2021). Alternatively, users should be given other ways to quickly refer to the source document.

Provide easy access to multiple MT outputs. MT users sometimes translate the same text with different tools to improve their understanding of it (Nurminen 2019, 2020; Martindale 2020), and Gao et al. (2015) found that providing MT users with two outputs is an effective way to improve understanding. Access to multiple MT outputs could be offered by embedding a second MT tool directly into the interfaces used to access raw MT (see an example in Nurminen 2019: 37). Another option is automatically generating output with multiple engines and showing all outputs to end-users (Martindale 2020).

Provide access to other language resources. Verifying the correctness (or incorrectness) of specific words or passages can be helpful in MT gisting. The organization can encourage such verification by providing easy access to resources. For example, the organization's own glossaries could be translated (by humans) into the languages in which raw MT is delivered and opened to end-users. Technology could also be used to highlight terms or offer links between source and target versions of terms. Integrating access to freely available resources such as online dictionaries can also enable easy verification of raw MT content (Martindale 2020: 41).

Help users recognize errors. Some technological solutions can flag output with potential errors. For example, MT confidence estimation (or quality estimation) produces an automatically estimated score for the quality of a translated piece of text (see Guerberof 2020: 347). Highlighting passages with low scores can prompt MT users to approach them with caution and verify the information (Martindale 2020: 42).

Offer an alternative path to information. Whenever possible, MT users should be offered a second way to get information in case of MT failure. Commonly, people are offered the option of contacting support services, and this should continue to be an option even after information is delivered in users' own languages via MT.

Share the risk

The organization can also decide to share the risk with other stakeholders, most importantly, with the people reading the raw MT content. Sharing the risk means ensuring that the end-user is aware of the risk and obtaining confirmation of that awareness for reasons of liability. The following are examples of treatment options organizations could consider.

Make MT transparent. People should always be aware that they are reading raw MT output. If users request MT themselves, they are naturally aware of it, but if they are offered pre-translated content, it may not be obvious to them. Therefore, clear labels and disclaimers are needed. People might also save machine-translated documents for future use or share them with others (Nurminen 2020). Therefore, raw MT should be properly labelled as such. Disclaimers, explanations and warnings placed in critical points of the end-user's process of accessing raw MT are also important for increasing MT literacy among users (see previous section).

Confirm that end-users are aware of risks. Asking end-users to confirm that they understand the risks involved in using raw MT might be appropriate in addition to warnings and disclaimers. For example, users might be required to check a box confirming their understanding before proceeding to raw machine-translated content.

Change the external environment. Sharing risks could also mean joint risk management initiatives with other organizations that deliver content as raw MT. For example, organizations in an industry might join forces to implement industry-wide guidelines and best practices for raw MT publishing, develop MT literacy programmes, sponsor sessions on MT gisting in conferences and user group meetings, or even introduce legislation if needed (Canfora and Ottmann 2020).

Monitoring and review, recording and reporting

An important part of the risk management framework in ISO 31000 is the ongoing monitoring, review and reporting of risks and risk treatment. Explicit and transparent communication about risks is important (see Zaveckaite and Ulbinaite 2018: 1338) for monitoring. People within the organization should be encouraged to report any MT-related problems or errors they notice themselves or encounter in customer comments. For human translation, translators themselves are in key positions to report translation errors, including near-misses that were caught during revision (see Canfora and Ottmann 2018: 178), but no such revision stage exists for raw MT use. For this reason, it is vital to proactively collect feedback and experiences from the end-users of the organization's machine-translated texts, for example, in the form of surveys. The organization should also implement easy procedures for end-users to report MT errors and encourage them to use this option regardless of whether major consequences occurred.

It is also important to establish clear procedures for identifying incidents involving raw MT use. This will provide the organization valuable information on risks and consequences, and will facilitate the assessment of the effectiveness of risk treatment. Canfora and Ottmann (2018) propose a process for monitoring and analysing incidents related to human translation, and a similar process could be applied to the raw MT situation. In addition to major incidents where more or less serious consequences have occurred, Canfora and Ottmann (2018: 181) emphasize the importance of analysing 'near-miss' incidents where a translation error could have caused a potentially hazardous situation even if no consequences occurred. For example, a customer might report on a translation mistake they encountered in the raw MT that would have led to severe consequences if not noticed. Observing and analysing such cases is important because near-misses and major incidents share the same root causes, and addressing the causes helps reduce the likelihood of more serious events (Canfora and Ottmann 2018: 172–3).

Risk management in action: The case of patent professionals

To illustrate the principles of risk management, this section examines the use of raw MT for information in a real-life use case involving patent professionals in the intellectual property rights (IPR) field. This group has actively relied on raw

MT to understand patent documents needed in their work for approximately a decade (Nurminen 2019: 32). The background for our discussion of risk management is formed by a study conducted in 2018–19 by Nurminen, who interviewed nine Scandinavian patent professionals about how raw MT is used in their working environments (for more detailed results, see Nurminen 2019, 2020; 2021).

In an IPR process such as patenting an invention or defending a patent, one of the most important tasks of patent professionals is to review all potentially relevant patent documentation. This review can involve hundreds of documents in various languages. Documents written in languages the patent professional does not speak are read as raw MT, and the goal is to understand each document well enough to evaluate whether it is relevant to the IPR case or not. If raw MT is not sufficient for making this decision, patent professionals can request a human translation. However, time and costs limit the number of documents that can be translated by humans. On the other hand, overlooking a relevant document can lead to negative consequences and is considered one of the main risks patent professionals face. These conflicting needs and risks are weighed against each other when deciding whether to rely on raw MT or choose human translation.

Nurminen (2019) concluded that this decision-making was a process of risk assessment. Although the process was described as more informal and manifested through patent professionals' own thought processes and discussions with their colleagues, the way risks are evaluated and weighed against benefits nevertheless reflects more formalized risk management principles. In what follows, we examine the use of raw MT by patent professionals through the lens of the risk management principles presented in this chapter.

To begin with, the patent professionals displayed a good knowledge of the scope and context in which they evaluated risk in decisions concerning translation. They discussed their own organizations' views on risk and were aware of the criteria that determined if risks were acceptable. They also displayed a familiarity with the external context of IPR work and its various stakeholders and processes.

When faced with the need to understand a potentially relevant patent document written in a foreign language, patent professionals described identifying the risks of using raw MT. For example, they pinpointed risk factors related to the IPR process in which the document would be used and risks related to their own initial understanding of the machine-translated version. They then analysed the identified risks, considering likelihood and potential consequences, as in this example that was voiced by informant #4 in Nurminen (2019: 36):

> If we make the wrong decision and allow a product to the market which does not have freedom to operate, there is a risk of using time and money and goodwill in a court case and potentially being responsible to cover the damages of a client.

Finally, they evaluated the identified risks against their organizations' risk criteria and decided whether to rely on the machine-translated document or to incur further time and cost by ordering human translations.

If the treatment decision was to take the risk by relying on raw MT, patent professionals might continue to try and achieve a better understanding of the raw MT document in question, for example, through discussions with technical experts or accessing further material such as multimodal components of the original patent document (see Nurminen 2019: 37). In the IPR context, the risk of using raw MT is always shared through full transparency. All stakeholders are kept fully aware of the general use of raw MT and of which texts are authored by humans and which are raw MT.

Conclusion

This chapter set out to examine risks and risk management in situations where an organization intends to deliver information to end-users as raw MT. Such practices are still relatively new but will likely increase as technology advances and people adjust to new ways of communicating and receiving information. Although such practices offer the benefit of making more information available in more languages to more people, those benefits must nevertheless be weighed against risks. As illustrated in this chapter, applying the risk management principles and framework defined in ISO 31000:2018 can help organizations identify and analyse potential risks, evaluate whether the risks are acceptable, and determine how to treat the risks. Examples of risk factors, analysis and evaluation criteria, and proposed treatment options are provided in various sections of the chapter and illustrated with the real-life use case of MT gisting by patent professionals. Nevertheless, we emphasize that risk management requires analysis of the organization's specific operating context and active communication with a diverse group of stakeholders.

Although risk management can offer a lens through which translation processes and actions can be analysed, as suggested by Pym (2021: 447) and Nurminen (2021), risk management in the specific context of content delivery using raw MT has to date gone unexplored. This chapter contributes a new viewpoint to considerations of risk in various types of translation and contributes to practical risk management planning and decision-making in organizations that use raw MT to deliver content or are considering doing so.

The nascent nature of research on MT gisting and its inherent risks means that our approach focuses more on initiating a conversation than offering final solutions, and we look forward to further discussions and contributions on the topic. Future research could concentrate on examining specific areas of risk management in greater detail. Additionally, empirical investigations into how risk is conceptualized in actual environments in which content is delivered as raw MT would be welcome.

Note

1 For example, the European Patent Office's MT tool, Patent Translate, has this function.

References

Asscher, O. and E. Glikson (2021), 'Human Evaluations of Machine Translation in an Ethically Charged Situation', *New Media and Society*, 25 (5): 1087–107, https://doi.org/10.1177/14614448211018833.

Barrault, L., M. Biesialska, O. Bojar, M. R Costa-Jussà, C. Federmann, Y. Graham, R. Grundkiewicz, B. Haddow, M. Huck, E. Joanis, T. Kocmi, P. P. Koehn, C. Lo, N. Ljubešić, C. Monz, M. Morishita, M. Nagata, T. Nakazawa, S. Pal, M. Post, and M. Zampieri (2020), 'Findings of the 2020 Conference on Machine Translation (WMT20)', *Proceedings of the Fifth Conference on Machine Translation*, 1–55. Available online: https://www.aclweb.org/anthology/2020.wmt-1.1.

Baumgarten, S. and J. Cornellà-Detrell (2019), 'Translation and the Economies of Power', in S. Baumgarten and J. Cornellà-Detrell (eds), *Translation and the Global Spaces of Power*, 11–26, Bristol: Multilingual Matters.

Bowker, L. and J. Buitrago Ciro (2015), 'Investigating the Usefulness of Machine Translation for Newcomers at the Public Library', *Translation and Interpreting Studies*, 10 (2): 165–86, https://doi.org/10.1075/tis.10.2.01bow.

Bowker, L. and J. Buitrago Ciro (2019), *Machine Translation and Global Research: Towards Improved Machine Translation Literacy in the Scholarly Community*, Bingley: Emerald Publishing.

Byrne, J. (2007), 'Caveat Translator: Understanding the Legal Consequences of Errors in Professional Translation', *The Journal of Specialised Translation*, 29 (7): 2–24, https://doi.org/10.1037/023079.

Canfora, C. and A. Ottmann (2018), 'Of Ostriches, Pyramids, and Swiss Cheese: Risks in Safety-Critical Translations', *Translation Spaces*, 7 (2): 167–201.

Canfora, C. and A. Ottmann (2019), 'Who's Afraid of Translation Risks?', in H. E. Jüngs, L. Link, K. Schubert, and C. Zehrer (eds), *Challenging Boundaries: New Approaches to Specialized Communication*, 73–92, Berlin: Frank & Timme GmbH.

Canfora, C. and A. Ottmann (2020), 'Risks in Neural Machine Translation', *Translation Spaces*, 9 (1): 58–77, https://doi.org/10.1075/ts.00021.can.

Cismas, S. C. (2010), 'Risk Assessment and Mitigation in Drafting Translation Contracts', *Proceedings of the International Conference on Risk Management, Assessment and Mitigation*, 491–5.

Densmer, L. (2020), 'Machine Translation and the Scale of Emotional Weight', *RWS*, 4 February. Available online: https://www.rws.com/blog/machine-translation-and-the-scale-of-emotional-weight/

Dillinger, M. and L. Gerber (2009), 'Success with Machine Translation: Automating Knowledge-Base Translation, Part 2', *Client Side News*, February: 9–11.

Doss Mohan, K. (2021), 'Translator Announces Document Translation (Preview)', *Microsoft*, 17 February. Available online: https://techcommunity.microsoft.com/t5/ai-cognitive-services-blog/translator-announces-document-translation-preview/ba-p/2144185 (accessed 23 January 2023).

Dunne, E. S. (2013), *Project Risk Management: Developing a Risk Framework for Translation Projects*, Doctoral dissertation, Kent State University. Available online: https://etd.ohiolink.edu/apexprod/rws_etd/send_file/send?accession=kent1368700402&disposition=inline (accessed 23 January 2023).

Gao, G., B. Xu, D. Hau, Z. Yao, D. Cosley, and S. R. Fussell (2015), 'Two is Better than One: Improving Multilingual Collaboration by Giving Two Machine Translation Outputs', *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 852–63. Available online: https://dl.acm.org/doi/10.1145/2675133.2675197

Google (2022), 'Search Quality Rater Guidelines: An Overview'. Available online: https://services.google.com/fh/files/misc/hsw-sqrg.pdf

Guerberof-Arenas, A. (2020), 'Pre-Editing and Post-Editing', in E. Angelone, M. Ehrensberger-Dow, and G. Massey (eds), *The Bloomsbury Companion to Translation Industry Studies*, 333–60, London: Bloomsbury.

Guerberof-Arenas, A. and J. Moorkens (2023), 'Ethics and Machine Translation: The End User Perspective', in H. Moniz and C. Parra Escartín (eds), *Towards Responsible Machine Translation: Ethical and Legal Considerations in Machine Translation*, 113–33, Cham: Springer.

Haddow, B., A. Birch, and K. Heafield (2021), 'Machine Translation in Healthcare', in Ş. Susam-Saraeva and E. Spišiaková (eds), *The Routledge Handbook of Translation and Health*, 1st edn, 108–29, Abingdon, Oxon; New York: Routledge.

ISO (2018), *ISO 31000 Risk Management*, Guidelines, Geneva: International Organization for Standardization.

Martindale, M. J. (2020), 'Responsible Gist MT Use in the Age of Neural MT', in *Proceedings of the 14th Conference of the Association for Machine Translation in the Americas, Workshop on the Impact of Machine Translation*, 18–44. Available online: https://www.aclweb.org/anthology/2020.amta-impact.2/

Martindale, M. J. and M. Carpuat (2018), 'Fluency Over Adequacy: A Pilot Study in Measuring User Trust in Imperfect MT', *Proceedings of the 13th Conference of the Association for Machine Translation in the Americas (Volume 1: Research Papers)*, 13–25.

Nitzke, J., S. Hansen-Schirra, and C. Canfora (2019), 'Risk Management and Post-Editing Competence', *Journal of Specialised Translation*, 31: 239–59.

Nurminen, M. (2019), 'Decision-making, Risk, and Gist Machine Translation in the Work of Patent Professionals', *Proceedings of the 8th Workshop on Patent and Scientific Literature Translation*, 32–42. Available online: https://www.aclweb.org/anthology/W19-7204/

Nurminen, M. (2020), 'Raw Machine Translation use by Patent Professionals: A Case of Distributed Cognition', *Translation, Cognition & Behavior*, 3 (1): 100–21, https://doi.org/10.1075/ts.00025.nur

Nurminen, M. (2021), *Raw Machine Translation in Context: How Context Influences the Use and Reception of Raw Machine Translation*, Doctoral dissertation, Tampere University, Finland.

Nurminen, M. and M. Koponen (2020), 'Machine Translation and Fair Access to Information', *Translation Spaces*, 9 (1): 150–69.

O'Brien, S. (2020), 'Translation Technology and Disaster Management', in M. O'Hagan (ed.), *The Routledge Handbook of Translation and Technology*, 304–18, Abingdon: Routledge.

Pym, A. (2015), 'Translating as Risk Management', *Journal of Pragmatics*, 85: 67–80, https://doi.org/10.1016/j.pragma.2015.06.010.

Pym, A. (2021), 'Translation, Risk Management and Cognition', in F. Alves and A. Jakobson, *The Routledge Handbook of Translation and Cognition*, 445–58, Abingdon: Routledge.

Rossetti, A., S. O'Brien and P. Cadwell (2020), 'Comprehension and Trust in Crises: Investigating the Impact of Machine Translation and Post-Editing', *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation*, 2: 9–18.

Scott, J. and J. O'Shea (2021), 'How Legal Documents Translated Outside Institutions Affect Lives, Businesses and the Economy', *International Journal for the Semiotics of Law*, 1–43, https://doi.org/10.1007/s11196-020-09815-5

Thicke, L. (2013), 'Revolutionizing Customer Support through MT', *MultiLingual*, 49–52.

Vieira, L. N., M. O'Hagan, and C. 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*, 24 (11): 1515–32, https://doi.org/10.1080/1369118X.2020.1776370

Way, A. (2013), 'Traditional and Emerging Use-Cases for Machine Translation', *Proceedings of Translating and the Computer*, 35, London: Aslib.

Zaveckaite, A. and A. Ulbinaite (2018), 'Assessment Criteria of Project Risk Management in Language Translation Service Companies', *Technological and Economic Development of Economy*, 24 (4): 1323–43, https://doi.org/10.3846/20294913.2017.1295287