

Translation Needs Security or “Why Can’t I Just Google Translate It?”

Linguistic Systems, Inc. and Michael Simon, Seventh Samurai



About 89% of content submitted for translation contains confidential information. Foreign-language translations, if not handled securely, expose law firms and their clients to tremendous financial and legal risk.

Here are some of the risks and how to avoid them.

Free is easy. Free is simple. Free is fun. Free is, most of all, free. So, when you need to translate something — whether it’s client communications or due diligence documents from a cross-border M&A transaction — is it ok to use Google Translate or other public translation tools? After all, they’re free; and it’s not like it costs you anything.

Or maybe it does?

Free does have a cost, at least when it comes to machine translation. Documents that contain confidential information, as in the examples above, need to remain confidential or the cost of free translation can be very high.

If it's your job to work with confidential information, you need to know that free translation services, no matter how easy, simple, or fun, can't always keep that information safe.

And that's not a price you want to pay.

1. Why is security important for translation?

You have many options when it comes to translation. You can use Google's or Microsoft's online translation websites. You can use other public sites, such as Translate.com. You can even use Microsoft's Word Translator plug-in without leaving your Word documents. All of these are free, with no up-front costs.

If you need to quickly translate some personal work or non-confidential business documents, then free services are a great solution to what was once an expensive, difficult process.

However, if you're dealing with confidential information, whether your company's or client's, consider what happens to the information you send to these free services.

Here, for example, is what Microsoft tells you about the Word Translator plug-in and about data sharing to improve the quality of translations.

Data Sharing

You can help Microsoft® Translator suggest better translations. If you turn on Data Sharing, Microsoft® Translator will record which translations you insert into the document when several translations are available. It will also learn from changes you make to suggested translations.

What data is collected?

In order to provide translations for words or sentences you have chosen in the document, this app sends the chosen text to the Microsoft® Translator online service. This is necessary for the app to receive translations from Microsoft® Translator. When you turn on Data Sharing, this app will also collect the following information to improve the quality of translations suggested by Microsoft® Translator:

- *Information on which of the suggested translations you insert into your document;*
- *Changes you make to a suggested translation before inserting it into your document.*

Microsoft and other free services also fail to provide assurances that your data will not be stored and put at further risk. If the information you are translating via free services is confidential, that lack of assurance that your confidential information will be protected should be troubling.

So, who is it that needs to translate confidential information? Here's a list: Law firms; accounting firms; consulting firms; medical professionals; and any company dealing with patents, trade secrets, or potential acquisitions.

There are also those whose data could be used for insider trading, or whose financial info, employee info, GDPR-covered Personal Data, PII, PHI, PCI, or any other private information must be protected.

For all these situations, protecting confidential and personal information – or your clients' – is most likely part of your job, and failing to do that properly can have devastating consequences.

Let's examine the legal profession more closely, as attorneys and litigation managers handle confidential client information constantly, and the penalties for mismanaging it are clear.

How clear? American Bar Association Rule 1.6(c): Confidentiality of Information requires that:

“(c) A lawyer shall make reasonable efforts to prevent the inadvertent or unauthorized disclosure of, or unauthorized access to, information relating to the representation of a client.”

Lawyers who fail to make such reasonable efforts to prevent the disclosure or access to client information can be – and some have been – fined, suspended or even disbarred.

For years, there was a debate within the legal industry as to whether email could be used to send confidential client information. That debate was resolved in favor of using email, but [a 2017 ABA opinion¹](#) updated the requirements recently to include the need for lawyers to use “particularly strong protective measures” in some circumstances.

[Others have questioned²](#) whether Google's free Gmail service can put lawyers into ethical danger because Google harvests data from emails sent through Gmail. Such concerns over email, which has been around for decades, should give lawyers even more cause for concern regarding newer, less-settled systems such as online machine translation.

2. Translation systems routinely handle confidential information that needs to be secure

Concerns over the protection of confidential information aren't just held by lawyers, they are shared by translators as well. The American Translators Association Code of Ethics and Professional Practice requires translators, whether using manual or machine processes to:

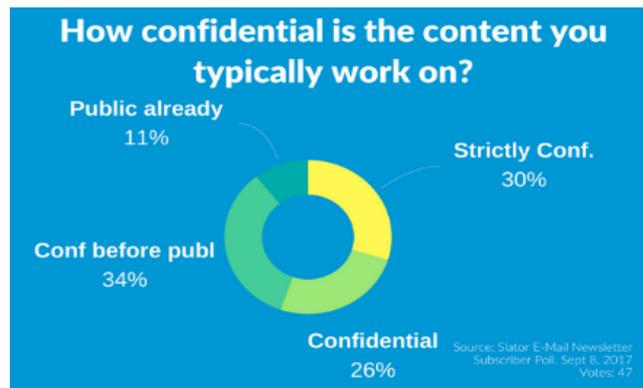
"hold in confidence any privileged and/or confidential information entrusted to us in the course of our work."

The ATA Code of Ethics Commentary goes further to clarify that:

"Clients expect their information to be held in strict confidence. This includes information conveyed in a translation or during interpreting, as well as the very fact that the translation or interpreting has taken place.

It goes without saying that translators and interpreters adhere to all existing international, federal, or state laws or acts concerning confidentiality."

It's easy to understand why translators have this concern for protecting confidential information. As [a recent survey by Slator.com³](#) shows, nearly 90% of content that translators work on is considered "confidential" to some degree. Nearly one third of that content is considered "strictly" confidential – the highest level.



In fact, [a recent story at Slator.com](#) shows how free, unsecured machine translation systems can create dangerous data leaks of confidential information.

In late 2017, employees of Norwegian state-run oil giant Statoil "discovered text that had been typed in on Translate.com could be found by anyone conducting a Google search." Their reaction:

"Wow, what is this?"

Slator went on to report that:

"An expert contacted by the Norwegian state-owned TV station further uncovered 'plans of workforce reductions and outsourcing, passwords, code information, and contracts.'"

Yet, as Slator reported, it's not like Translate.com did anything illegal, or even unexpected, because the "fine print" (as Slator calls it) in Translate.com's online terms and conditions states:

"... while they 'will use reasonable measures to protect any content you provide to us for the purpose of completing the Services,' they 'cannot and do not guarantee that any information provided to us by you will not become public under any circumstances. You should appreciate that all information submitted on the website might potentially be publicly accessible.'"

After the breach became publicized, Slator posted an update of Translate.com's reaction:

"Translate.com attempts to justify the breach by pointing out that 'there was a clear note on our homepage stating: All translations will be sent to our community to improve accuracy' and that 'some of these requests were indexed by search engines such as Google and Microsoft at that time.'"

Finally, they say,

"users can request translations to be removed by emailing support@translate.com."

3. Sometimes "information just wants to be free" is a really bad idea

In 1984, even before the Internet was opened to commercial use, Stewart Brand, the well-known futurist and creator of the Whole Earth Catalog, uttered his famous quote:

"Information wants to be free."

Less famous is the very next thing that Brand supposedly said,

"It also wants to be expensive."

As we've seen, there are plenty of free options available for translation. However, free services have to make money somehow, and that's how your free information can become surprisingly expensive. In addition to their well-known free online translation, Google has translation applications that you can purchase. Otherwise, your data is their product.

Not convinced? Try running a Google search for a product. Here's an example: "Murano glass sink." Just by clicking on one search result, you will probably see it repeatedly in other Google ads and on other sites for days, weeks, or perhaps

months. That's Google (and its advertisers) using your data. Until recently, [Google even allowed this with no restrictions within Gmail⁵](#).

This is Google's business model and, considering that the company (as of the date of this whitepaper) has a market value of $\frac{3}{4}$ of a **trillion** dollars, it works incredibly well. Google also [reportedly paid \\$9 billion to be the default search system on the iPhone⁶](#). Google uses your data to make money, and they tell you upfront (even if in fine print) that they will do so.

Is Google using your data from its Google Translate app to make money as well? Google says that they don't do that, so we accept that. Yet at the same time, we know that Google's CEO was [reported to have stated earlier this year](#) that "their app translates a staggering 143 billion words every day." Considering that Google has already massively monetized search and Gmail, experts are concerned that:

"Given that a lot of people most likely use the translation app while traveling, it's not a stretch to imagine ads . . .

. . . the massive popularity of Google's translation app is unlikely to be ignored forever. Whether Google flips the switch next quarter or two years from now, there's money in Google's language machine."

Even if third parties agree *not* to monetize the data you give them by using their free services, there still may be serious issues. Consider, for example, Article 17 of the EU General Data Protection Regulation (GDPR), known as the "Right to Erasure ('Right to Be Forgotten')":



"The data subject shall have the right to obtain from the controller the erasure of personal data concerning him or her without undue delay and the controller shall have the obligation to erase personal data without undue delay."

Unfortunately, the terms of use of many free translation services, including Google, allow them to keep data indefinitely. If you use such systems to translate "personal data" under the EU definition – one that is very broad – you may find yourself in the difficult position of trying to get it erased by a company that has no obligation or reason to comply.

4. Securing your people, processes and technology

There are several points of risk in the translation process, starting with technology, but not just technology alone. These include:

- **Sending information:** Unencrypted information can be intercepted in transmission to the vendor;
- **Revealing information to a vendor:** Vendor's personnel have access to information that is unencrypted and improperly secured against unauthorized access;
- **Trusting vendor assignment to subcontractor translator:** If machine translation is not being used, the translation must be done by a person. Because translation companies cannot hire translators for every language needed, they use subcontractors – creating additional points of risk during transmission;
- **Trusting subcontractor translator:** Vendors that use subcontractors must have processes in place to assure that subcontractor translators can be trusted; and
- **Delivery of information:** When vendors send back the deliverable to the customer, this creates a final point of potential risk in transmission.

For these reasons, it is critical to identify all potential points of risk whether they involve technology, people, or processes so they can be eliminated or mitigated to the fullest extent possible.

In fact, [security experts advise](#)⁷ that people will always be your biggest points of risk. Machine translation helps eliminate some of those risks, in that it takes people out of the equation.

The higher the number of technology systems and data transmissions required, the greater the risk. Therefore, a single platform, that involves a comprehensive translation solution with no external data transfers, is more secure. This ideal secure translation solution should offer:

- Platform security;
- Vendor security.

Platform security would include the following:

- Access protected – only authorized users can access the app;
- Secure encrypted data transfer (256-bit SSL);
- Proprietary machine translation engines – no other vendor;
- Secure encrypted integrations (as with Relativity) via API;
- Encrypted data on servers – even if hacked, the data is unreadable;
- Immediate disposal of MT content upon completion.

Vendor security and confidentiality must be standard procedure. It should include:

- Secure Processes – as evidenced by these certifications:
 - **ISO 9001 Quality Management Systems** – ensures that quality processes and corrective action procedures are always followed;
 - **ISO 17100 Translation Quality Management Systems** – ensures best practices specific to translation management;
 - **ISO 27001 Information Security Systems** – ensures that your projects and clients' personal data are kept permanently secure.
- Translators with signed NDAs;
- Translators and employees have access to data on a need-to-know basis;
- Confidential data cannot be downloaded by translators.

A secure single-platform translation solution is not free, but it can save you considerable exposure and risk in the long run.

5. Our secure solution: Ai Translate by LSI™⁸

Ai Translate by LSI™ is a unique, secure, best-in-class language translation solution. You can choose from up to 6 quality levels and 64 embedded language pairs from one simple interface. (As many as 120 total languages can be ordered.)

Linguistic Systems, the owner and developer of Ai Translate, has earned 3 coveted ISO certifications for quality and data security:

- **ISO 27001** – Information Security;
- **ISO 9001** – Quality Management Systems;
- **ISO 17100** – Translation Quality Management;

These ISO certifications are difficult to achieve, and they require third party testing and re-certification at regular intervals. They represent an ongoing commitment by Linguistic Systems to the highest standards of performance and information security.

Ai Translate is also GDPR-compliant; and our machine translation technology does not retain your data and put you at risk – unlike the most commonly used public translation sites.

Linguistic Systems has also been a long-term partner with Relativity, the leading provider of eDiscovery hosting and review software. The Ai Translate Plugin™ for Relativity seamlessly integrates with Relativity, for a single platform experience within the most popular eDiscovery platform.



Conclusion

You have many options for your translation needs. For those items where information security is not required, you can use a free online app. But when confidential information is involved, and your company's or clients' reputation, money – and maybe even your job – are on the line, consider the *real costs* of “free” translation.

In the end, investing in a more secure translation solution is probably worth it.

Quoted Sources

¹ **LawSitesBlog.com** —

<https://www.lawsitesblog.com/2017/05/aba-issues-major-ruling-ethics-email-electronic-communications.html>

² **ABAJournal.com** —

http://www.abajournal.com/news/article/does_using_gmail_put_attorney_client_privilege_at_risk

³ **Slator.com** —

<https://slator.com/features/reader-polls-mt-quality-lsps-risk-data-breach-universities/>

⁴ **Slator.com** —

<https://slator.com/technology/translate-com-exposes-highly-sensitive-information-massive-privacy-breach/>

⁵ **PCmag.com** —

<https://www.pcmag.com/news/364276/google-third-party-apps-must-stop-scanning-gmail-for-ad-tar>

⁶ **FastCompany.com** —

<https://www.fastcompany.com/90243975/at-12-85-per-iphone-googles-default-search-payment-is-probably-a-steal>

⁷ **BusinessInsider.com** —

<https://www.businessinsider.com/sundar-pichai-google-translate-143-billion-words-daily-2018-7>

⁸ **AiTranslate.com** —

<https://aitranslate.com>