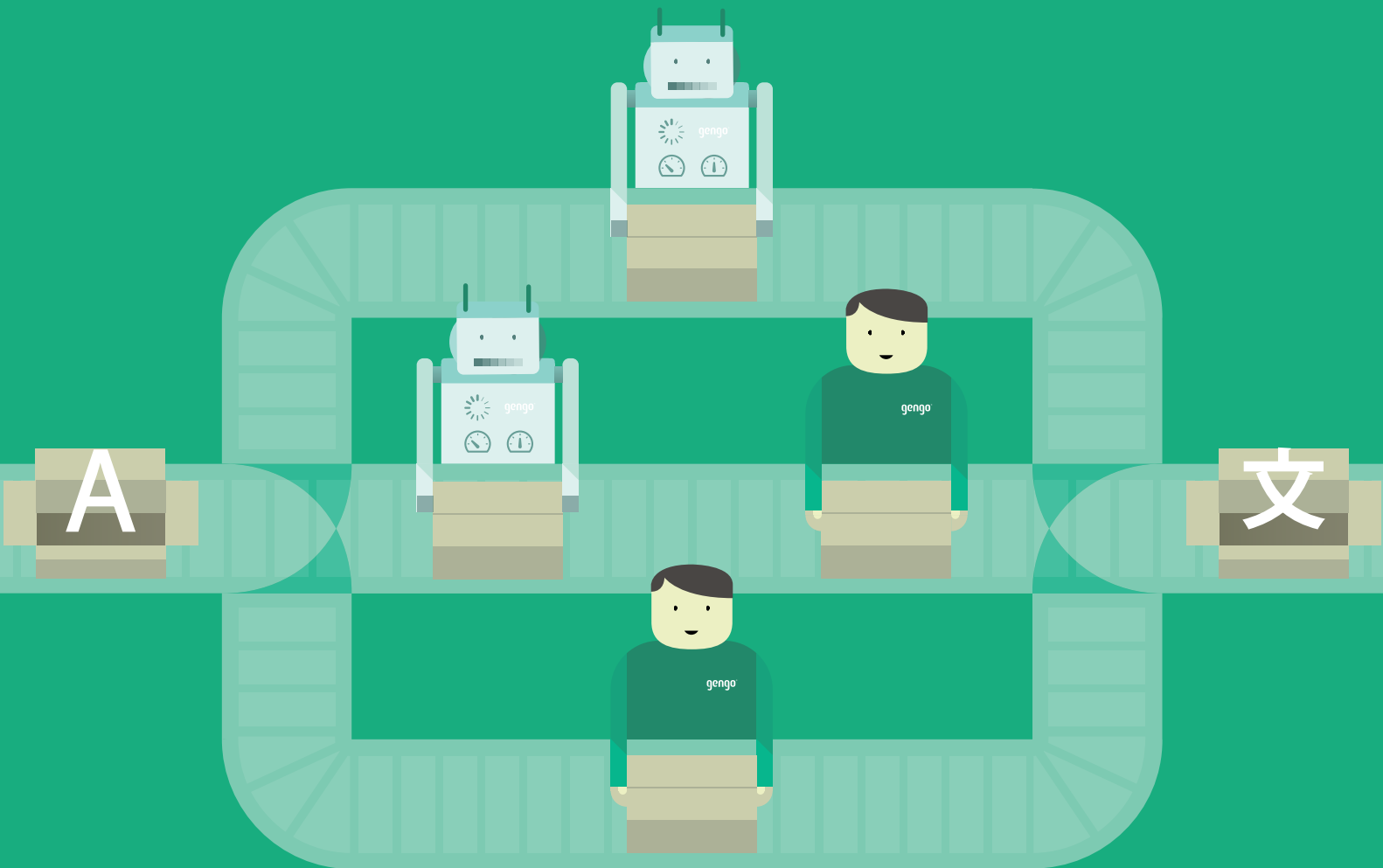


Making sense of translation services

Human translation vs.
Machine translation vs.
Post-edited machine translation



gengo






Introduction

The translation landscape used to be “one size fits all,” with human translation the only available technology. Now, translation can be done traditionally, with a crowd, a machine or a combination of machine and human intelligence via post-editing.

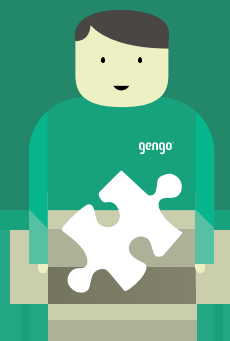
Choosing the right solution and vendor can be complex and stressful. This simple guide will give you the tools you need to confidently reach a decision.



Index

	Solutions explained simply	4
	Things to consider	6
	Human translation (HT)	7
	Machine translation (MT)	9
	Post-edited machine translation (PEMT): No free lunch	10
	Compare the solutions	11
	Questions to ask	16
	Ask your company	17
	Ask a vendor	18
	Questions for all vendors	18
	Questions for HT vendors	19
	Questions for MT vendors	20
	Questions for PEMT vendors	21
	Conclusion	22
	About Gengo	24

Solutions explained simply





A guide to the three main types of translation



Traditional human translation (HT), or newer crowd platform

Human translation (HT) is when a person—rather than a machine—translates text. However, HT can be handled in many different ways. For example, within a traditional agency process whereby files are managed manually, full-time or freelance professional translators work alone or in small groups, and technology has little part in the process; or within a technology-leveraged “crowd” system like Gengo whereby thousands of translators work on digital texts simultaneously using a crowd platform to automatically manage projects. Crowd techniques significantly increase the scalability of HT and reduce the cost associated with traditional models.



Post-edited machine translation (PEMT)

Machine translation reviewed, edited and improved by humans. Post-edited machine translation (PEMT) services normally cost more than MT but ideally less than traditional HT services. The post-editors might be semi-skilled proofreaders or translators with some experience, or even language and/or domain experts.



Machine translation (MT)

A software-based process that translates content from one language to another. Machine translation (MT) services accelerate the process, and reduces the cost, of translating content. It can also increase the availability of translated content. Linguistic quality and accuracy levels vary depending on how well the software is tuned. MT is normally the lowest-cost translation method.

Things to consider





Things to consider



HT Wide variation in cost and approach

Human translation is still by far the most widely-used translation method. But in the past few years, “crowd” platforms like Gengo have radically changed the industry, using technology to scale the process of human translation and making it affordable for a much wider market to go global. This means it’s more important than ever to understand the variation in human translation providers.

There is a huge difference between “boutique” human translation agencies and modern “crowd” platforms like Gengo.

Both types have their place, but ensure you understand the difference between them before beginning. A parallel could be in fashion—when purchasing a suit or a dress, you can choose from a tailor who will consult with you and create a garment by hand over several weeks that will match your needs, or you can buy something off the peg that was produced far more affordably, uniformly and quickly in a factory.

Human translation comes in as many flavors. Due to limited technology and content distribution, the traditional industry was more optimized to the “tailored” end of the spectrum, so most translation providers are more suitable for smaller projects at a higher price. Per-word prices at boutique agencies can be 10 times or more compared to crowd platforms.

Agencies often offer different levels of service.

For example, translation only, proofreading, or TEP, which normally refers to translation, editing and proofreading. You might have complex files and projects to manage, too. Ensure you understand what you are paying for.

The more manual processes and project management used by the agency, the more you will pay over and above the base translation cost. Ensure that you need these extras, or choose a provider who is more suitable. Crowd providers eliminate most of the overheads of a traditional agency through technology, saving cost and allowing scale.

Human translation costs money.

While companies with huge communities like Facebook and LinkedIn have leveraged their users to perform some translation, the vast majority of translation is paid, whether by traditional agency or crowd. Likewise, there is no “free lunch”, and it’s unwise to expect to get like-for-like service at a cheaper price.

Most translators leverage some kind of technology to assist them.

This can range from a simple electronic dictionary, automated glossary or term glossary tools, to sophisticated translation memory technology that stores whole phrases for reuse among translators, thus greatly improving efficiency. This may have a price impact on your project.



Things to consider

There is wide variation in vendors' abilities to integrate with your technology.

If you are managing content and publishing to the web, your vendor should have an API to allow content to automatically be sent to and retrieved from your platform.

HT providers sometimes consider translation as more art than a science.

While translating literary texts or poetry is art, translating for business should be judged objectively, with quality control, speed, adherence to style guides and brands carefully controlled. Make sure your provider understands how to manage these aspects well.

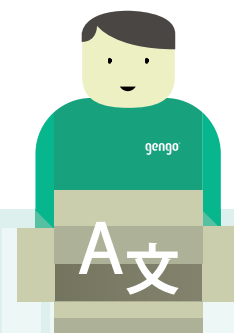
Many HT providers cannot scale.

It's always tempting for a provider to give you their "best" translators at the start, and then struggle to achieve the same quality at scale if you want to translate at high volumes.

Make sure that a trial process is explained with objective quality assessment criteria and that the same methods will be used in the trial as when your project is at full volume. In addition, choose a provider who is set up to handle large volumes. Completely different quality control methods are required compared with at smaller scales.

Some content may be suitable for particular vendors, alternative content for others.

For example, you might wish to have your brand messaging and advertising translated by a high-end transcreation agency whose translators will spend a long time honing specific phrases for high-profile use, at a high cost per word. You might get your other content, such as product descriptions and user reviews translated at scale by a crowd platform set up for affordable, high volume translation.





Things to consider



MT Useful, but limited in scope

MT has long held promise as the holy grail of translation, but still does not normally meet the quality needs for most organizations. However, it's worth staying up-to-date on advances.

The quality of MT is improving, but slowly.

Since the mid-20th century, researchers have continually improved machine translation using statistical, rule-based and other techniques. That said, we are still far away from human-quality translation with MT and its use is still limited.

It's not just Google Translate.

Many first-time customers assume that Google's translation tool is the only or foremost machine translation engine out there. But Google Translate is currently much more focused on consumer usage, has no method for large-volume customers to train their MT, and requires customers to allow Google to use their data, making it an unlikely choice for many enterprises. Instead, providers like Microsoft, Asia Online, Systran and others are more popular in the high-volume space.

Quality can be substantially improved via MT training.

Here, the machine "learns" to translate by studying existing documents, knowledge and historical translations on a domain and then mimics the output style so that it's as close to human translation as possible. This capability is normally only offered to enterprise customers.

MT is normally most suitable for tightly defined subject areas and technical texts.

This is because it can be trained for a much more limited set of vocabulary and usages. Customers with very narrow subject areas are therefore much more likely to get good results versus those with more varied or unpredictable source texts.

The bigger your MT needs, the more complex pricing and service-level agreements can become.

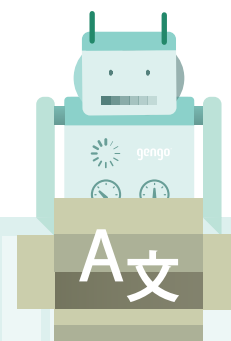
Whereas Google Translate is effectively pay-per-use, higher-end providers will often charge setup fees, hosting fees, and on-site vs. cloud usage fees. And while Google Translate is effectively instant for consumers, turnaround time for high-volume MT can depend on how much you invest.

Quality can vary hugely between language pairs.

Similar languages (like French and Italian) currently produce much better machine translation than distant languages like English and Japanese.

It may be hard to compare per-word costs for MT.

This could be due to various fees like server setup costs and licenses.



Things to consider

PEMT No free lunch

On the face of it, PEMT may seem to be the perfect solution—just get the machines to do the grunt work, then add a dash of human voice on top. However, several key limitations remain.

It's hard for post-editors to improve a poor machine translation.

To turn it into a great translation, it often needs to be completely rewritten. Often, post-editors are not incentivized enough to do this adequately, meaning PEMT translations are not much improved, or not improved at all compared to the original MT. As such, PEMT should not be relied on for stylistic or “human” tone of voice.

The cost of PEMT can be similar to, or higher than, crowd translation.

Because post-editors need to understand and work on the same texts as crowd translators, they have to charge similar rates. This can often mean that there is little or no cost saving with PEMT, combined with an effect whereby the post-editor is reluctant to change even a poorly-styled MT output because it requires significant effort.

The viability of PEMT varies hugely depending on the language pair.

This can be because of language similarity and the progress of development of MT. For example, Japanese<=>English MT is far behind French<=>English MT. This makes PEMT much more viable in French<=>English, and often impossible to undertake in Japanese<=>English.

Post-editors can often miss mistakes created by machines.

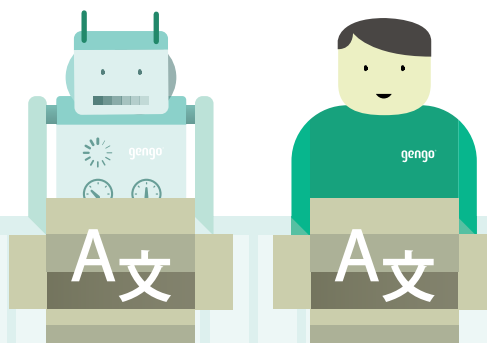
Sometimes these are not obvious grammar errors. For example, a source text might contain a number or qualifier that a machine translation might omit erroneously, changing the meaning of an entire sentence. Unless an editor is diligent, they may miss this oversight.

PEMT is unpopular with professional translators.

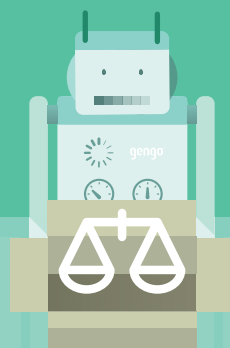
It can turn them into “cleaners” who are simply identifying errors and working with unwieldy and unsatisfying copy. Even newbie translators can quickly become fatigued working with MT. So it can be hard to find the right pool of editors at the right price point, and it can be tricky to retain editors long-term on large projects. Editors may try to game the system if the work is too repetitive or mundane.

When a subject area is specialist, post-editors need to be specialist, too.

For example, a machine translation of a specialist automotive text may be reasonably good if the MT is well trained and the subject area is narrow. But a non-specialist post-editor may miss key errors or even introduce new errors into the text because of their lack of subject-area knowledge. This means specialists may be required, increasing the price of PEMT beyond being just a cheap and simple solution.







Compare the solutions









Compare the solutions





Each method has strengths and weaknesses, depending on your project characteristics. Consider the following factors when deciding on a solution.

	 HT	 MT	  PEMT
The simple answer	<p>Complex projects requiring high touch, particularly creative or literary content.</p> <p>Crowd HT is perfect for high volume projects of everyday content like product descriptions, user reviews and short articles.</p>	<p>High-volume, low-value content in a tightly-defined subject area, normally for an internal audience.</p> <p>Also great for personal use, or to get a gist of an original text.</p>	<p>When machine translation produces an output that is slightly lacking. For instance, in medium-value content within a tight subject area that contains a few errors here and there.</p>
Example of suitable projects	<p>Projects where the human touch is essential.</p> <p>Traditional agency:</p> <ul style="list-style-type: none"> - Complex localization - Marketing materials - Sales brochures - Keynote presentations <p>Crowd platform:</p> <ul style="list-style-type: none"> - Product descriptions for ecommerce - User reviews at scale - Social media content - Support ticket management 	<p>Projects where immediacy and price are the key factors.</p> <ul style="list-style-type: none"> - Internal automotive technical documents - Database feedback - Internal defect reports - Gist for high volumes of legal documents - Inbound support tickets for gist 	<p>Projects where machine translation needs small improvement.</p> <ul style="list-style-type: none"> - Technical documents for low-profile public consumption - Large catalogs of technical material - Internal databases





Compare the solutions

	 HT	 MT	  PEMT
Complex context	★★★	★	★★
<p>The more complex and important the context of a translation (e.g. the knowledge of surrounding text, the brand, other non-obvious factors), the more you need humans. Crowds can leverage context by following style guides, glossaries and metadata like photography. Extremely high context-sensitive work like transcreation should be left to traditional HT agencies.</p>			
Low volume	★★★	★	★
<p>Overall, if your project is small and expected long-term demand is low, MT and PEMT are unlikely to be good solutions due to the longer ramp-up time and low cost savings.</p>			
High volume (>10M words)	★★★	★★★	★★★
<p>Traditional HT agencies that use a limited number of translators, manual processes, and a 9-to-5 style conventional workflow will struggle to complete high-volume projects within a reasonable cost and timeframe.</p> <p>Good-quality crowd human translation providers like Gengo can match MT and PEMT solutions, using a technology platform to manage large numbers of translators in the cloud. Bear in mind that MT can become expensive at scale, and PEMT providers need to have crowd capabilities like Gengo in order to handle high volumes efficiently (traditional HT agencies just tacking on a PEMT solution will not be adequate).</p>			
Wide-ranging subject area	★★★	★	★
<p>If your project contains an unpredictable variety of topics, it can be impractical to try to train an MT or leverage PEMT, or to achieve good quality. Human translation is essential to handle this kind of issue.</p>			
Repetitive content	★★★	★★★	★★★
<p>At this stage, most HT and crowd HT providers are using translation memory, a technology that allows translators to leverage phrases that were previously translated, saving the customer money and increasing speed and consistency. So repetitive content itself is not a reason to switch to MT.</p>			





Content suitability

	 HT	 MT	  PEMT
Social and casual content	★★★	★	★
Internal technical documents	★★	★★★	★★★
Marketing collateral	★★★	★	★
Legal	★★★	Gist only	★
Product descriptions	★★★	★	★
Support tickets	★★★	★★	★★
Manuals and guides	★★★	★	★★

Output expectations

	 HT	 MT	  PEMT
Gist only	★	★★★	★
Human style/tone	★★★	★	★★
Adherence to brand	★★★	★	★★
Life- or safety-critical	★★★	★	★

Compare the solutions

	 HT	 MT	  PEMT
Risks to consider	The main risk is human error. Traditional agencies have workflows that can be inefficient, especially at scale. Crowd providers try to eliminate these risks and costs through automation.	Machine translation risks are mainly around errors being hard to catch at scale, and unexpected project delays or costs due to mismanaged expectations.	Post-editing combines the risk of machine translation with that of human translation. Additional risk is that the post-editors have insufficient domain experience, potentially introducing further errors or not catching serious machine errors.
Costs to consider	<ul style="list-style-type: none"> - Project management - File handling - Minimum order fees - Per-word translation cost - Project setup fees 	<ul style="list-style-type: none"> - Set-up costs - Server costs - License costs - Per-word translation costs 	<ul style="list-style-type: none"> - Per-word translation cost - Project setup fees - Quality control costs

Questions to ask



? Questions to ask

Ask yourself

Deciding between HT, MT and PEMT depends on your company, project and needs. There's no single approach. The best method is to ask the right questions and dig deep on the answers.

Before contacting a vendor, ensure you have your house in order.

Project size: Is our project or expected yearly volume over one million words?

Most HT agencies can't handle this kind of volume, and many PEMT providers can't either. So a crowd or MT solution may be preferable.

Languages: Are we translating "close" languages that are suitable for MT?

Far-apart languages do not achieve a good MT result, making MT and PEMT bad choices for your project.

Intended audience: Who will read the output, and what are their expectations?

External audiences with high expectations have a low tolerance for mistakes. So MT and PEMT may not be suitable. Internal audiences can be more forgiving of stylistic mistakes and minor errors, making MT more suitable for these purposes.

Turnaround time: How quickly do you need work completed? At what volume?

Most HT agencies aren't built for speed. In addition, setting up a large MT project takes time.

Quality: How are we going to evaluate quality? Who will be responsible?

Being clear on quality up front will ensure you get better performance from your vendor, reduce confusion and increase speed.

Context and quality: Is it important that context is taken into consideration?

If translating work like product descriptions, additional content like photos and videos may need to be taken into consideration by the translator. MT and PEMT may not be suitable.

Price: What price per word are you willing to pay? What overall price?

If you're on a budget, you may be restricted to MT, PEMT or Crowd HT options. Be careful to take into account additional costs that many traditional agencies may include.



? Questions to ask

Ask a vendor

The best vendors have friendly, approachable websites and salespeople prepared to consult freely on the merits of different approaches. Most offer free or low-cost trials, too.

Questions for all vendors

How do you define and measure the quality of a translation?

Ask this to ensure that the vendor has an objective, understandable way that they define quality, so you can have a reliable feedback loop where your needs are matched.

Can we see an unfiltered set of translation examples?

It's easy to provide examples of your best work, but it's much more revealing to see an unfiltered set of real translations. If your provider is uncomfortable with this, it may be a red flag.

What is your price per-word?

Do you have minimum ordering fees?

Many HT services will have minimum fees. Find out up front.

What capacity of translation can you handle?

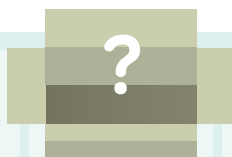
Many HT and PEMT providers can't handle high volumes, or will take an unrealistically long time to handle large projects. Crowd providers and MT providers are normally more capable.

Do you have an API? What kind of integrations are available?

Providers with an API and integrations with other platforms (like CMSs or ecommerce platforms) make it much easier to manage your content, and show their tech-savviness.

Do you have a selection of brand-name customers using your service?

It should be easy for your provider to give you names and references of similar customers.



? Questions to ask

Questions for HT vendors

There are thousands of traditional HT vendors. To sift the good from the bad, you want to ensure that you're talking to a vendor who has objectivity, a good understanding of technology, and experience in your area, combined with a straightforward pricing approach.

What is your average turnaround time?

What is the variation by language?

Vendors should indicate readily and with statistically-significant data to back them up.

What is your normal workflow?

What parts are manual and what are automated?

Try to get an indication of how scalable their process is. Manual may mean higher cost and lower speed. Crowd providers will normally fully-automated most project-management tasks.

How will you manage our company's translation memory?

As you build up a memory with larger projects, your provider should have a clear answer on whether they will re-use this for other clients.

Are there extra costs we need to be aware of?

Find out if you may be charged for hidden extras like project management, file-handling, technology integration or more.

Specifically for Crowd HT vendors

Crowd providers should be absolutely on top of their technology, ready to scale with you, and obsessed with objective, scalable techniques to manage quality and performance. The best providers use leading-edge techniques, are friendly and open, and have experienced sales staff who can consult with you.

Do you have a purpose-built technology platform?

Innovators like Gengo built their platform from the ground up to provide crowd translation, making the process hyper-efficient and completely scalable.

How do you maintain style and brand, like glossaries and style guides?

The best crowd providers have glossary tools built into their technology platform so their translators can follow your style.

Can you leverage translation memory?

Crowd providers may have different levels of support for translation memory. This is more than just a checkbox so ensure you have a full understanding.

Are your operational metrics publicly visible so that we can compare?

The best crowd providers display their metrics publicly, for example Gengo's Open Data site allows the public to see their performance metrics at any time. How does your speed compare to other providers? Many providers claim to be "the fastest". Ensure you have real, hard data about turnaround time, rather than unverifiable claims.

How do you manage quality across language pairs?

Crowd providers need scalable quality control techniques that are much more robust than small agencies'. Ensure your crowd provider has rigorous, objective control.

What kind of account management and post-sales service can I expect?

While automation and technology is essential for crowd providers, you need an account manager you can trust. The best crowd providers have sales teams providing white-glove service.

? Questions to ask

Questions for MT vendors

Your objective is to figure out if MT is right for you, what quality you can expect, and what the investment of time and money will be to get the results you want.

Is MT suitable for my content and needs? Why?

Good MT providers will be honest about whether you're a good match, and refer you to a crowd, traditional HT or PEMT provider otherwise.

What is the ranking of language pair quality for your system? How do my languages fit in?

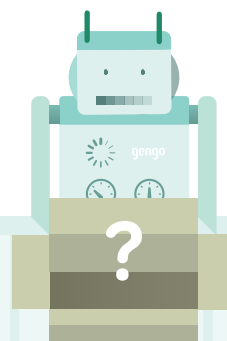
MT providers should be honest and upfront about quality of MT between specific language pairs. The best providers will coach you on the factors affecting quality for your content.

What is the training process, how large of a corpus is required, and how long does it take?

High-quality MT providers will ensure you have a realistic timeframe and expectations on how quality may improve with training over time.

What are the costs I need to consider at each stage of scale?

Depending on the provider and your needs, you may use cloud-based, self-hosted or other solutions. You might pay consultation fees or additional usage fees.



? Questions to ask

Questions for PEMT vendors

The main objective for you to establish is whether post-editing will provide better quality than MT alone, will scale, and whether the quality for price will outmatch HT, especially Crowd HT.

What qualifications and experience do your post-editors have within our industry?

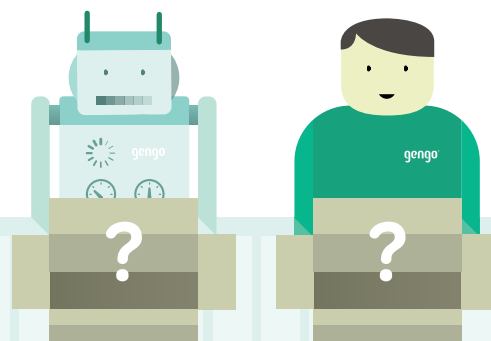
It's important to balance cost against the risk that post-editors are not experienced enough in your content to provide value. More dangerously, with specialist content, inexperienced post-editors may make serious errors. Ensure you have an honest conversation with your provider to establish understanding.

How do you measure your post-editors' improvement of quality over MT?

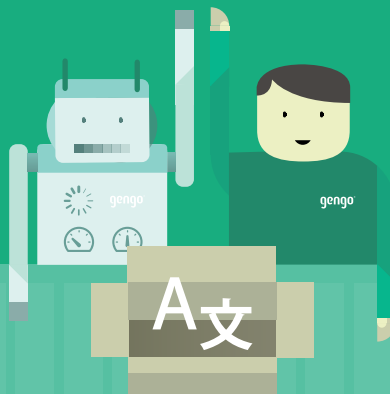
You need to ensure you are getting what you pay for, and reaching an actual improvement over the raw MT when using post-editing. Too often, clients may pay for post-editing and receive very little improvement over the original MT.

What is your error rate?

This is a leading question, but designed to find out if the PEMT provider has a solid understanding of how well they detect errors by their editors and in their MT. How suitable are our proposed languages for PEMT? A good provider will be honest and up-front about whether your project is suitable for PEMT.



Conclusion



You can confidently choose the right translation process for your project by using the information in this white paper, and by following the five simple rules below.

1. Understand the difference between HT, MT and PEMT

You're always in a better position if you are knowledgeable about your options. Vendors may push you to a particular solution that suits them, rather than giving you a choice, so be prepared in advance.

In brief: Use HT when you need a human touch. Use MT when you need scale and immediate translation without a human touch. Use PEMT sparingly to soften MT and reduce errors.

2. Ensure you're ready

Make sure your organization has asked itself the right questions before contacting a vendor. Simple things like "who will judge quality?" can cause huge headaches if not figured out in time, no matter which process you use.

3. Experiment before making a big commitment

It's best to know your options before you need to order a huge project. So make small trials, ensuring any trials are "real" and use production-level resources.

4. Use different solutions for different purposes

There is no "one size fits all" approach, and while sticking with one provider for all your needs may seem convenient, you can often get better results elsewhere, often at a far more advantageous price. So do your homework.

5. Understand the huge HT gains made by crowd providers

The traditional translation industry still offers good quality for complex or "high-touch" projects, but crowd providers like Gengo have opened up a huge new industry for good quality, human-style translation at scale. Make sure to build your knowledge and experience with this new style of translation.

Still not sure?

Contact Gengo for a free consultation 

About Gengo

Gengo is a global, people-powered translation platform enabling everyone to read and publish across languages with one click. Gengo's unique platform consists of a diverse and growing community of more than 15,000 qualified translators from 114 countries working across 37 languages and 63 language pairs. The professional human translation platform represents a huge leap in quality compared to machine translation, while enjoying faster delivery times than traditional service providers. Through its web offering and open API, Gengo enables businesses of any size to quickly and easily connect with a global audience at scale. Clients and partners include TripAdvisor, Rakuten and The Huffington Post. Founded in 2008, Gengo is headquartered in Tokyo, Japan, with an office in Silicon Valley.

Learn more at gengo.com 



gengo