

White Paper

Machine Translation

CONTENTS

| | |
|--|-----|
| Introduction..... | 1 |
| Brief History of MT at TransPerfect..... | 1 |
| TransPerfect MT Facts..... | 2 |
| Successful MT Implementation..... | 3 |
| Neural MT Technology..... | 3 |
| Implementation Process..... | 4 |
| Human-Centered Approach..... | 5 |
| TransPerfect MT Services..... | 5 |
| Leveraging MT for Your Organization..... | 6 |
| MT Success Stories..... | 7-8 |

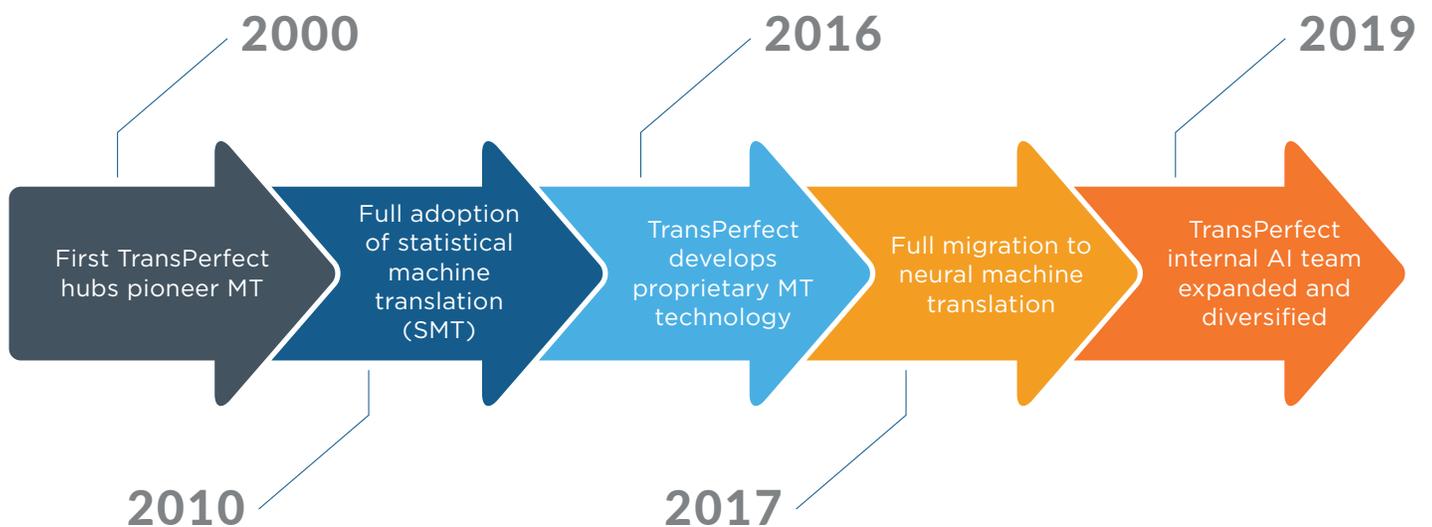
INTRODUCTION

The use of artificial intelligence (AI) in a diverse array of applications is becoming more ubiquitous. In our everyday life, AI applications are used in mobile phones, intelligent speakers, recommender systems, and many other applications. Many of them achieve impressive results, and for a variety of specific tasks they outperform humans. Similarly, the benefits afforded by AI have been employed to tackle a wide range of societal challenges in different domains, ranging from health and hunger to security and justice and from economic empowerment to environmental protection. The increasingly varied applications of AI have earned it a place at the forefront of advances in human development.

BRIEF HISTORY OF MT AT TRANSPERFECT

Machine translation (MT) is one of the main AI applications in the field of human language. MT at TransPerfect traces its origins to the beginning of

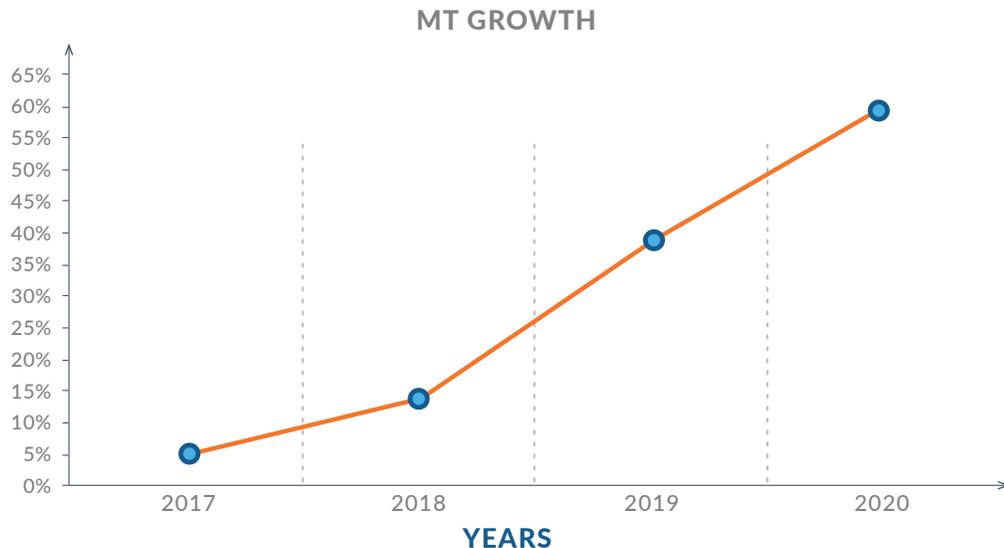
the 21st century with major production hubs in Austin, TX, and Las Vegas, NV, that were specialized in this field and that originally used MT primarily for legal content for e-discovery. At the time, the technology behind our systems was rule-based, but the foundations for statistical MT (SMT) were already laid. Over time, the process was improved with several MT providers, which eventually led to the full adoption of SMT around 2010. A major milestone in TransPerfect MT was the acquisition of Tauyou, a Spanish MT service provider. At the time, Tauyou primarily employed SMT technology and was taking the first steps in neural MT development and paving the way for the development of our proprietary MT technology. In 2017, TransPerfect was fully migrated to neural MT, more or less at the same time as major tech giants such as Google, Facebook, and Microsoft. In conjunction with the overall growth of MT, the internal AI team has diversified to include a research and development department, software engineers, and an implementation team located in several countries around the world. The team is expected to double in size within the coming year to support the growing needs of clients and the industry.



TRANSPERFECT MT FACTS

MT Growth

Currently, over 50% of all of our translation work goes through MT-based workflows, and we have seen constant growth since 2017: 4% in 2017, 15% in 2018, and 42% in 2019. When reflected as revenue, this means that over \$170 million of TransPerfect's revenue is generated by using MT as a first step in the process.

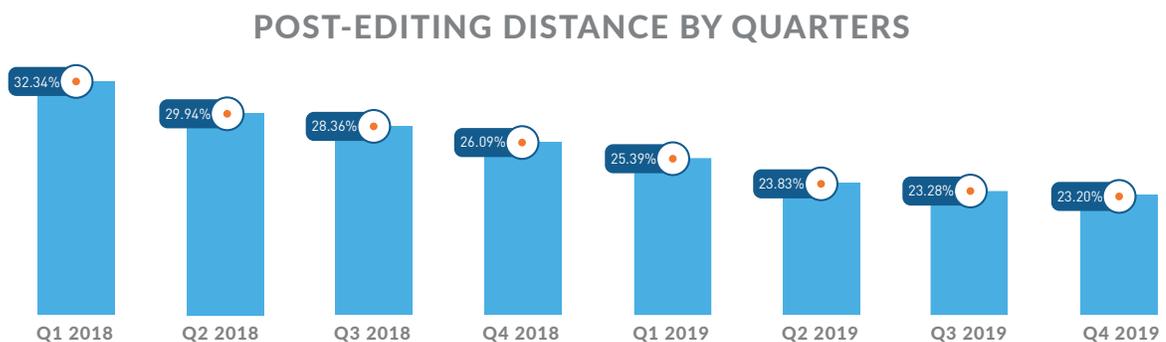


Linguist Certification

More than 75% of our active freelance linguists are certified for MT post-editing services. Every month, around 200 new linguists are onboarded.

MT Quality

Average post-editing distance (PED) has improved by around 30% in 24 months: from 32.34% to 23.20%. The graph below displays the average PED, a metric commonly used to evaluate MT quality. PED shows the amount of human effort needed to post-edit the MT output, with lower PED indicating better MT quality.



Our off-the-shelf engines for languages like Arabic and Japanese for life sciences content outperform competitors, as detailed in this blog post.

GlobalLink AI Portal

We released the **GlobalLink AI Portal** for self-service MT to replace publicly available alternatives with confidentiality and security, which has processed tens of millions of words since launch and which generates millions in revenue.

SUCCESSFUL MT IMPLEMENTATION

MT technology is as important for a successful MT implementation as the processes and people involved.

Neural MT Technology

Neural MT (commonly abbreviated as “NMT”) technology is based on deep neural networks, a powerful algorithm that is able to learn from existing translations created by humans. The first neural systems became available in 2016 and have since become the prevailing MT paradigm. The main advantages afforded by neural MT are improved fluency and morphological agreement, which were not well handled by previous models. In addition, neural MT is able to better recognize the context of each word, as well as its meaning.

TransPerfect uses state-of-the-art proprietary neural MT technology, as well as technology for data cleaning, engine preparation and deployment, and ongoing improvement through machine learning and linguistic analysis. Currently, we support over 40 language pairs and are continuously working to expand the list.

We constantly benchmark the quality of our systems against other MT systems available in the market to ensure that our technology is a top contender. However, the best quality can be achieved with MT customization—the process in which we use our clients’ linguistic assets, such as translation memories and glossaries, to “teach” our systems to translate the content accordingly. Our tests demonstrate that MT customization leads to an average quality improvement of 30% over the baseline.

In addition to the benefits offered by customization, ongoing improvement to MT engines can yield up to an additional 30% increase in quality, which can be accomplished through incorporation of newly translated data from linguistic assets (e.g., translation memory), correction of the most frequent changes made by linguists, and implementation of subjective linguistic feedback provided by our translators and subject-matter experts.

Another area of research and development that we are actively working on is MT quality estimation (QE), a technique which allows for a quick estimate of the quality of MT output. QE is useful in a variety of situations, from automatically selecting the best MT engine to estimating the budget for post-editing based on the MT quality.

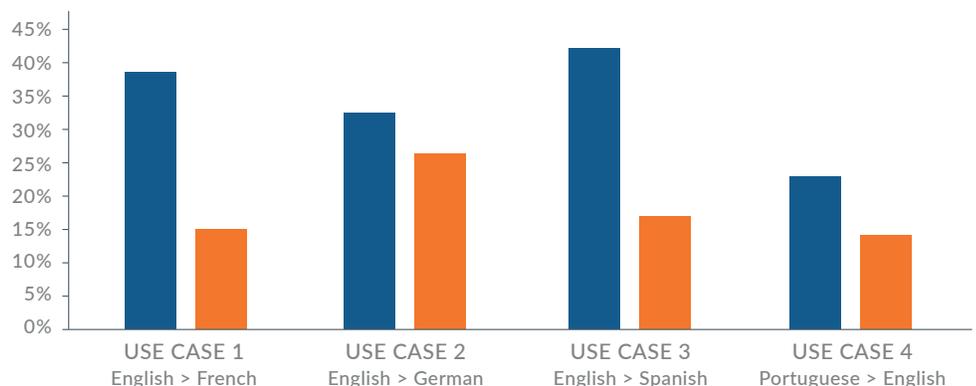
QUALITY IS ALWAYS OUR #1 PRIORITY

As part of our commitment to quality, TransPerfect has ISO 18587:2017 certification for post-editing services.

PED IMPROVEMENT WITH ENGINE CUSTOMIZATION

*Lower PED indicates better MT quality

■ Initial PED ■ PED after customization

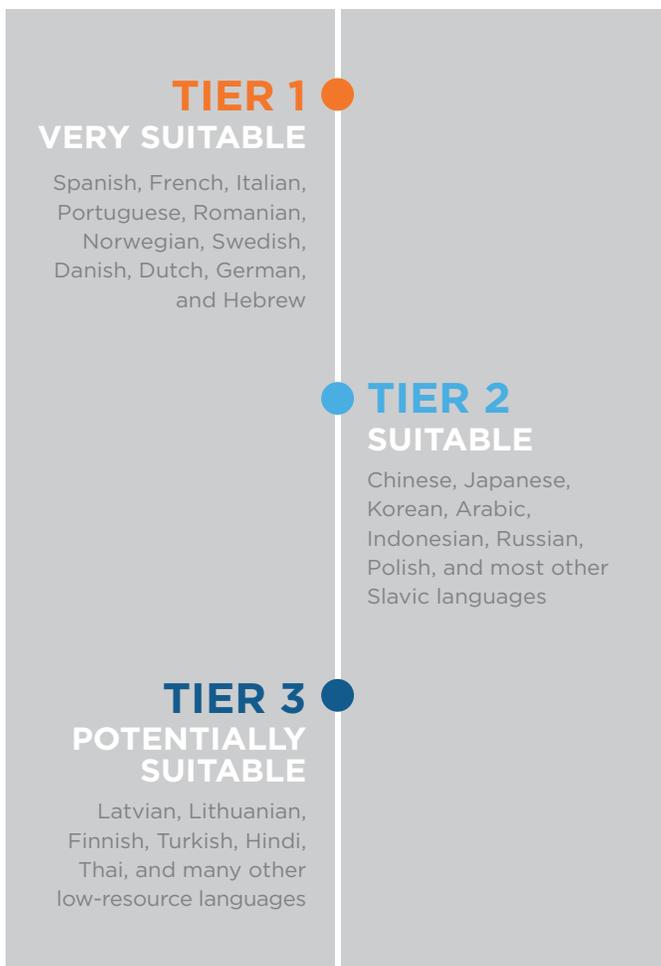


Implementation Process

Tiered-Language Approach

The first question we are asked by clients is what languages work best for MT. As a general rule, it can be stated that the more complex a language, the more challenges it poses to MT. Similarly, the greater the linguistic difference between the source and target languages, the more difficult the task for MT. Although this has improved slightly with the advent of neural MT, differences are still observable depending on the complexity of the language.

For the purposes of MT, we typically group languages into tiers. Using English as a source or target for simplicity, these are the main tiers, in descending order of complexity, for the most common world languages requested for translation (including their variations):

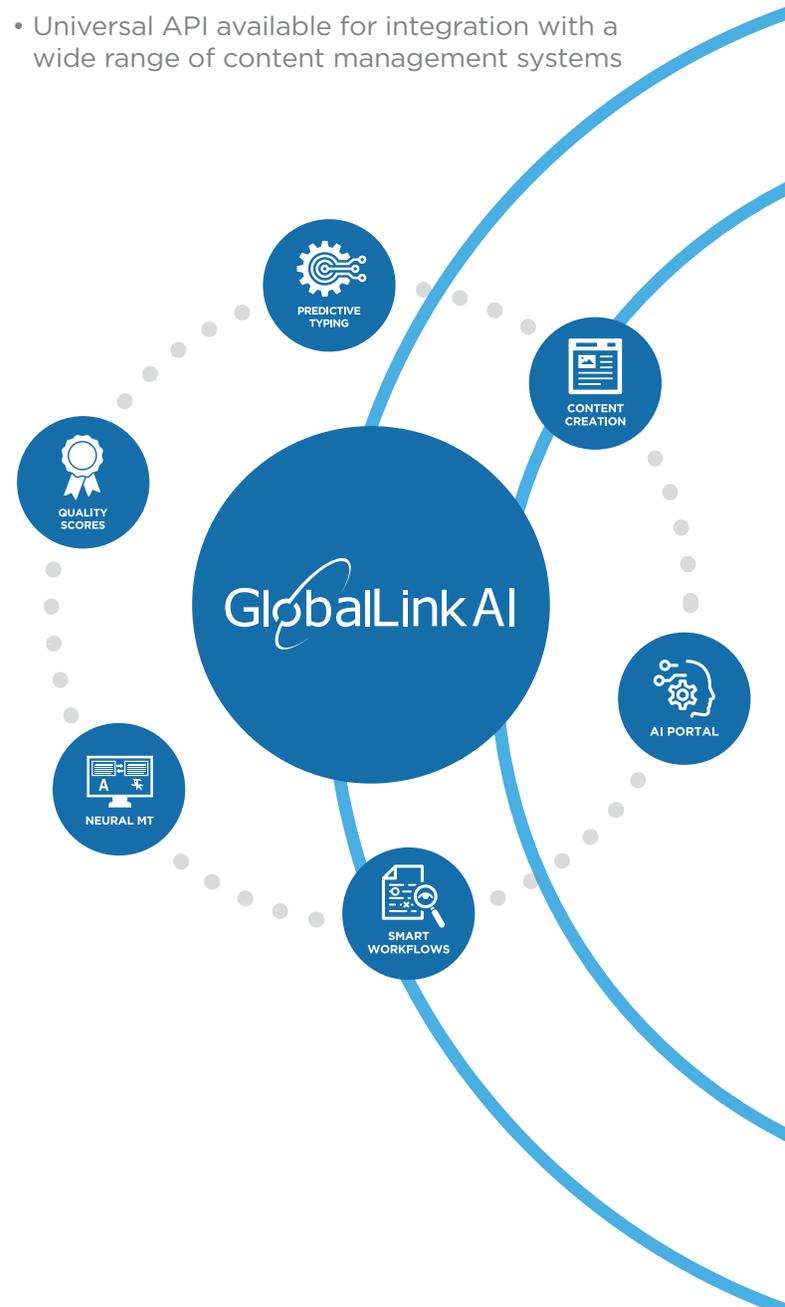


For language combinations that do not include English, our engines typically yield results that are much better than most publicly available options, provided enough data is available for customization.

A Feature in Any Platform

We view MT as a feature available in a variety of platforms for use by clients and internal teams. These include:

- GlobalLink AI Portal – a standalone web-based interface for secure document translation
- Any GlobalLink product such as GL Project Director
- Third-party technology such as Relativity for e-discovery
- Universal API available for integration with a wide range of content management systems



Human-Centered Approach

We advocate for a human-centered approach to using AI in translation, in which our expert linguists are supported by the use of technology. This is why working with the people involved in the process is as important to us as the technology itself. Further, we have developed training programs for our linguist community, internal teams, and clients who use our MT tools.

More than 75% of our active freelance translators are certified for the MT post-editing service, which means that they are not only highly efficient in post-editing, but also MT experts who help us to continually develop and improve our systems.

The TransPerfect AI team combines both technology and linguistic skills to support global MT implementation projects for clients of all sizes and for all business needs.

TRANSPERFECT MT SERVICES

We don't view MT as a replacement for human translation but rather as a way to maximize the business impact of translation and localization. Currently, MT serves as a productivity tool used by translators, and the acceptance of MT by end customers has never been higher.

In addition to the traditional human translation services including one, two, or three steps without machine translation, MT enables us to offer the following service levels and thereby maximize return on investment in translation and localization for our clients:

Full Post-Edited MT

Top-quality translation process with enhanced efficiency provided by MT. In this workflow, machine translation output is brought to human quality by a linguist who is a subject-matter expert, and the final product is equivalent to a human translation-only step. Full post-edited MT might be followed by a proofreading and/or a

quality management step as would be included in a standard workflow. We recommend maintaining the same number of steps as a human translation workflow, with the first step becoming a full post-editing step instead of a translation step. While full post-editing is not recommended for creative texts, any other type of content, including highly visible, customer-facing content, is suitable for this workflow. The goal of this service is to have zero impact on the final quality compared to a standard human translation process.

Light Post-Edited MT

Understandable translation with no major errors but not stylistically perfect, intended for lower-priority content where speed is more important than quality. Linguists will correct the MT with the minimum number of changes only when the target is not understandable, or does not convey the meaning of the source. However, the final translation might sound literal or unnatural.

Customized MT

MT output with no human editing. The output is produced by specialized MT engines trained with client assets, which means improved accuracy and higher quality. In order to achieve meaningful improvement over baseline quality through customization, a minimum of 50,000 sentences are required for training. MT customization leads to significant quality gains and is highly recommended in any scenario in which MT is used without human editing.

Raw MT

MT output with no human editing produced by one of our off-the-shelf MT solutions. This is MT that does not take into account terminology or style from our clients. This service is typically employed for 'gisting' (understandability), particularly in scenarios where the content types are very generic and do not have a heavy client focus.

LEVERAGING MT FOR YOUR ORGANIZATION



Strategic Migration

Are you translating via human translation but want to move to an MT post-editing workflow for strategic reasons?

We can perform an **MT pilot**, which will enable us to assess the impact of MT on the turnaround time and cost, depending on the language pair and content type destined for MT.



Budget Optimization

Are you interested in optimizing your budget for globalization and shortening time to market?

MT can assist by offering an array of service levels (raw MT, light post-editing, full post-editing, and customized MT) to reduce time spent on the translation phase and optimize budgets to deliver quality results as efficiently as possible.



Customized, Secure Solution

Do you want to replace the use of Google Translate at your company with a secure option?

TransPerfect's proprietary GlobalLink AI Portal provides a customizable, secure solution to suit your company's security and quality needs.



Expanded Capacity

Does your company need to translate large volumes of content for e-discovery?

Integration of MT solutions into platforms like Relativity support expedited translation of large batches of content.



Integration for Developers

Are you a developer interested in integrating MT in your platform(s) or tool(s)?

TransPerfect's Universal MT API supports integrations with your product.

MT SUCCESS STORIES

Leading European Hotel Chain (Travel/Hospitality)

Goal

A major player in the hospitality and tourism industry, this leading European hotel chain was seeking to optimize translation expenditures at the same time that TransPerfect proactively proposed MT and post-editing workflows to meet similar goals.

Solution

To support translation of predominately Spanish source content into five primary and five secondary European languages, engine training with robust client translation memories and an MT post-editing workflow was proposed for the former, and raw MT was proposed for the latter.

Result

Following successful implementation of a tiered MT approach based on business needs and a technology migration, a discount was offered to support cost reductions sought by the client. Additional solutions were proposed to further the efficiency gains introduced by MT, including leveraging our automatic language converter combined with the trained Spanish > Portuguese engine to extend cost and TAT savings to this language pair.

High-End Fashion House (Retail)

Goal

A luxury fashion house requested support from TransPerfect with MT as a way to save their translation budget for other, more strategic business initiatives like high-impact localized marketing materials.

Solution

A pilot was proposed to evaluate the quality of product descriptions translated into German, which was met with positive results. During the pilot, the engines were trained with client translation memories, and following the success of the pilot additional languages were added.

Result

In addition to nearly doubling productivity and the overall efficiency of the workflow, a discount was offered to the client for MT services resulting in cost savings of 15% on product content. Additionally, linguists reported feeling very pleased with the results of the MT output, and now German, Spanish, Italian, and Dutch are all supported.

Major Clinical Research Organization (Life Sciences)

Goal

During an RFP process, the AI Portal was proposed to provide several tiers of MT quality intended to meet different business needs at the enterprise level as a secure alternative to free, generic MT engines.

Solution

Case studies were presented to illustrate the benefits of a customized MT solution tailored to the specific business needs outlined by key stakeholders, and a pilot was proposed to evaluate the viability of the AI Portal solution.

Result

Following the success of the pilot and deployment of the AI Portal solution, MT services were expanded to other areas of the company, resulting in gains in efficiency and turnaround time and overall reductions in cost.

Global Insurance Provider (Legal)

Goal

A US-based global insurance company was seeking a solution to centralize and support the translation of an immense amount of foreign language policies into English for processing.

Solution

To support this unique business need, an integrated TransPort and AI Portal solution was proposed that would fully centralize the translation and project management process, allowing users to submit, track, and download machine translated policies.

Result

Following a successful implementation to a select group of teams, the solution was expanded to other business units through the creation of an intranet page directly linking to TransPort + AI Portal, supporting over 7,000 users worldwide and processing over 20 million words per year. Further, the solution was broadened to accommodate foreign into foreign language translation requests, illustrating the adaptability of TransPerfect MT solutions to meet the diverse needs of a global company.