

Optimizing localization with the right mix of human and AI

Organizations are adopting AI-powered technologies at a rapid pace. Understanding where and how to best apply AI and human expertise to ensure fluent, accurate, and useful translations can provide cost savings and competitive advantage.

Questions posed by: Smartling

Answers by: Hayley Sutherland, Research Manager,
Conversational AI Tools & Technologies



Q1 How are generative AI/large language models impacting the localization space?

There can be various types of generative AI – in this case, we are talking about large language models (LLMs) that have become very good at understanding and generating human language. When used in conjunction with Neural Machine Translation (NMT) and human translators and applied with the proper guardrails and prompt engineering, generative AI has the potential to support advanced capabilities that can provide significant advantages for companies.

These include producing higher-quality, more “human-like” and fluent translations from a combination of NMT, LLMs, and humans at a lower cost; automatically detecting and masking hate speech, harmful or offensive language; and enabling businesses to easily set preferences for tone, style and so forth that are automatically applied.

As in other areas, generative AI has promise for translation use cases, but there are still some caveats and potential pitfalls that organizations need to consider. For instance, there is the now well-known issue of “hallucination” in question-answering, in which a generative AI system can provide false or inaccurate information, particularly when prompted to answer a question for which it has little data.

Similarly, in MT, it can also produce “falsely fluent” translations. Just as its false answers to questions can sound completely plausible but be untrue, a generative AI model used for translation can seem grammatically fluent while actually producing translation errors. However, this weakness of generative AI can become a strength when used in conjunction with NMT; machine translation produces more accurate translations, while generative AI/LLMs can make translations sound more fluent and human-like.

Other concerns such as data privacy and IP challenges should be researched and well understood by any organization considering generative AI. Vendors supporting this technology are working to provide guardrails and other security measures, and buyers should ask questions about how these concerns are being addressed.

Q2 How have advances in AI, particularly neural machine translation (NMT), impacted the quality and cost of translation?

Until fairly recently, machine translation software was largely dependent on handcrafted rules and statistical models. While this helped to improve the speed of some translation workflows, most use cases still required a lot of manual work by humans, from fixing incorrect translations to making stilted machine translations sound more human-like. In the last five years, however, advances in deep learning and machine learning have led to the widespread use of neural machine translation (NMT) to power translation software. NMT uses deep learning (either replacing or in combination with rules-based or statistical models) to train AI models to recognize language patterns and produce increasingly accurate translations.

Advances in AI and the adoption of NMT have vastly improved the quality of machine translation software in recent years. At the same time, as compute power continues to improve, AI-powered capabilities such as NMT have become more accessible and cost-effective for many more organizations around the world.

As a result, NMT can now speed more translation workflows, either by handling them entirely or helping to reduce the time and effort spent by humans on each translation project, without substantially reducing quality. In addition, AI can enhance other aspects of the translation pipeline, such as analyzing MT quality to ensure the best MT model is used for the job, or helping to determine the optimal mix of human- and AI-led translation.

Q3 **How do you measure quality in machine translation, and why is quality measurement important?**

The ability to measure the quality of machine translation is foundational. With quality measurement, organizations can:

- Understand when MT may be a good fit for partial or complete translation
- Understand when to engage humans for translation or post-editing
- Determine which vendor or product offering is best suited for an organization's MT use cases
- Within MT products, determine which model will perform best for each content type, use case, and language
- Optimize quality and cost savings

There are various ways to determine machine translation quality, from BLEU scores (scoring similarity of machine-translated sentences to example sentences provided) to MQM (multidimensional quality measurement) auditing to AI-enabled metrics, including those that leverage LLMs. MQM is a framework that allows users to assess translation quality based on custom metrics selected from eight major dimensions (accuracy, fluency, terminology, locale convention, style, verity, design, and internationalization). While this is a proven method for measuring MT quality, it can be manual and time-consuming, so many organizations today are turning to AI-enabled tools to help automate MT quality measurement.

Regardless of the method, the key metric typically boils down to how much effort it would take a human to turn the MT into something acceptable for use. This is also why MT quality measurement is nuanced; "acceptable for use" varies by content type, audience, goal, and so forth. How this will be achieved is something that organizations should strongly consider when selecting a translation vendor, particularly if they are new to measuring MT quality or to leveraging MT to speed localization.

Q4 Why are automation, translation memory and dynamic workflow management such important considerations for localization?

In addition to automating language translation itself, including AI and automation at other stages of the translation workflow can help to speed processes, reduce manual labor, and improve consistency. Among the capabilities are the following:

- Automatically ingesting content for translation via connectors
- Pre-processing translation data to ensure higher quality MT
- Automating translation workflow management, including automating MT engine selection based on content, language, and so forth
- Performing machine translation quality assessment
- Identifying and anonymizing PII

Translation memory can further enhance AI and automation by allowing organizations to re-use previously translated content to quickly train NMT systems, improving quality while reducing cost and time spent training.

Dynamic workflow management is another area where AI can be strategically leveraged to determine the best machine translation engine to use, as well as when and where to incorporate human translation, depending on factors such as content type, goals, target audience, and preferred accuracy thresholds. This helps to optimize the translation process and outcomes, improving quality, costs, and time to value.

Q5 How do you measure ROI from localization?

Ultimately, each organization will need to determine the best ROI metrics to use depending on its localization content type, use case, audience, and goals. If this determination is being done across multiple use cases, or even multiple business areas, business leaders and other key users should be consulted in the development of initial requirements and ROI metrics for any translation service, including MT.

Some examples of potential metrics to determine the ROI of localization, including the use of machine translation tools, include:

- Time to market for translated content
- Improved customer and employee experience ratings
- Increased content usage/re-usage
- Reduction in local support requests
- Improved customer conversion rates in particular regions
- Expanded market reach

Some vendors will provide tools and services to help determine and measure the best ROI metrics for a specific business and its use case. Organizations should decide if they would prefer assistance with this from a third-party technology/services provider or if it will be done completely in-house.

About the Analyst



Hayley Sutherland, Research Manager,
Conversational AI Tools & Technologies

Hayley Sutherland is a Research Manager for Conversational AI and Knowledge Discovery within IDC's Software market research and advisory group.

Her core research coverage includes conversational AI and search, with a focus in AI software development tools and techniques for chatbots and digital assistants, speech AI and text AI, machine translation, embedded knowledge graph creation, search and knowledge discovery, and affective computing (also known as emotion AI).

IDC Custom Solutions

This publication was produced by IDC Custom Solutions. The opinion, analysis, and research results presented herein are drawn from more detailed research and analysis independently conducted and published by IDC, unless specific vendor sponsorship is noted. IDC Custom Solutions makes IDC content available in a wide range of formats for distribution by various companies. A license to distribute IDC content does not imply endorsement of or opinion about the licensee.

External Publication of IDC Information and Data — Any IDC information that is to be used in advertising, press releases, or promotional materials requires prior written approval from the appropriate IDC Vice President or Country Manager. A draft of the proposed document should accompany any such request. IDC reserves the right to deny approval of external usage for any reason.

Copyright 2022 IDC. Reproduction without written permission is completely forbidden.

IDC Research, Inc.

140 Kendrick Street
Building B
Needham, MA 02494
T 508.872.8200
F 508.935.4015
Twitter @IDC
idc-insights-community.com
idc.com

SMARTLING

Smartling's LanguageAI™ platform is revolutionizing the translation and localization of the world's digital content. Named as the top translation management system by CSA Research and by users on G2, Smartling harnesses AI and machine learning to enable the automation of workflows, integrates seamlessly with existing tech stacks and measures and improves translation quality at enterprise scale for a fraction of the cost of traditional translation.

Smartling is the platform of choice for hundreds of B2B and B2C brands, including IHG Hotels & Resorts, State Farm, British Airways, and Lyft. Smartling is headquartered in New York with an office in Dublin.

Smartling, Inc.

1375 Broadway 14th Floor
New York, NY 10018
T 866.707.6278
hi@smartling.com
smartling.com