



ALPHA



White Paper

The big translation technology debate

A round-table discussion on the outlook for translation technology

Technology. Translation. The future. To be blunt: if you have any interest in how these concepts interact, this is an in-depth discussion you probably won't want to miss.

That's because, in a lively debate featuring four of our technology experts, we've asked them to respond to all the key questions in translation technology in 2020 and beyond.

Such as, will machine translation with artificial intelligence replace the human translator? What will translation technology look like in five years' time? What's the best way to translate creatively while working within a translation tool? And will the global tech giants finally take over the translation industry?

ALPHA LTD

St Andrew's House
St Andrew's Road
Cambridge CB4 1DL
United Kingdom

 @thisisalphalive

alphacrc.com

Because the pace of change in the translation industry grows ever greater, we believe an open and unfiltered discussion on what the issues are right now and how the future might look is useful for everyone.

It's a fascinating, frank and unfiltered debate. As you might imagine, there are some common themes but also some divergent views.

You'll hear from **Steen Kesmodel** (Operations Research Manager), **Jon Chan** (Chief Technology Officer), **Guylaine Tritton** (Solutions Architect) and **Jorge Oliva** (Senior Software Designer/Engineer) on everything from the impact of artificial intelligence on translation technology to the development of integrated "super tools".

Of course, at Alpha CRC, we always take the most pragmatic project-by-project approach to how to make the most value of technology to benefit our clients.

But when it comes to lively debate on the future of our industry, we believe there's plenty of room for different points of view. See if you agree...

Q&A: The big translation technology debate

Q: Let's start by clarifying the terminology we're going to be using in this discussion: in terms of translation technology, what do we mean by CAT, TMSs and TEnT tools?

Steen: Although a lot of people still refer to CAT (computer-assisted translation) tools when discussing translation technology systems, it's perhaps not the most relevant terminology these days.

In fact, translation management systems (TMSs) and translation environment (TEnT) tools are more accurate terms when discussing today's technology. A TMS is essentially a server-based tool, while TEnT mostly refers to translators' (desktop) applications.

Q: So, do all TMSs and TEnT tools take the same approach, or are there some that are very different?

Steen: It depends on how you define "approach". The fundamental principles of all TMSs and TEnT tools are segmentation and a database for storing translations (a translation memory, or "TM"). You could say that the advances made in machine translation (MT) are reducing the importance of TMs, at least for some content types. But segmentation, i.e. dividing the text into smaller parts, whether sentences or paragraphs, is certainly still a basic principle.

Jon: You're right, in the last 10 years or so segmentation has definitely appeared as an approach that is common to all TMSs and TEnT tools. This has led to the translation editor being presented as a table of sentences. Segments are certainly the easiest way of storing translations in a memory.

However, the changing design of that memory and the introduction of artificial intelligence (AI) may mean that the "whole-sentence" approach to defining segmentation rules needs to be reconsidered.

Guyllaine: Segmentation is a mindset, perhaps even a habit. Sentence-level segmentation has always been favoured because it yields the most leverage, i.e. it is the most cost-effective for the client and the most productive for the translation provider.

However, the option of segmenting text at paragraph level has been available for a long time. Putting savings aside, segmentation at the paragraph level would give the translator a lot more freedom when high levels of creativity are required. TM concordance searches would still be possible, and quality assurance (QA) settings would function in the same way.



MT workflows would not be affected because MT engines would still process text at sentence level, even when working within the same segment. But then, if we are being creative, do we really need assistance from MT?

I realize that I'm describing transcreation projects here, and that most projects can be seen as standard, but I do think it is too easy to consider ourselves bound by segmentation. It all depends on what the client is prepared to spend.

Steen: Absolutely. I think the conversation around segmentation is still very much ongoing within different sections of the industry. In terms of differences, they can really be summed up by the focus of the various systems, and which role their design aims to support. Very broadly, desktop TEnTs were developed either with the translator in mind (like Trados, DejaVu, memoQ), or as a more or less basic tool to translate packages downloaded from a TMS server (SDLX, Idiom).

Server-based TMS tools, conversely, have shifted the focus to improving the turnaround time and automating the process by getting content directly to the TMS without having to handle files directly (export from the content system, import into the TMS, and back again), although they also frequently try to improve "context" for translators by having more previews of the text, either real (as in TEnTs for some file types) or via screenshots.

But there is also the odd outlier: Kantan Skynet, for example, has taken a radical approach. The entire tool is online: files are automatically run through an MT engine, and segments can then be assigned on a first-come-first-served basis to optimize translation speed. This means a small document can be translated in five minutes by 10 translators doing a few randomly allocated segments each.

Jon: I think another difference is the way in which tools approach the integration of MT. Several are actually introducing AI combined with neural machine translation (NMT). NMT makes much more use of sub-segment analysis than previous MT strategies.

This analysis approach uses “mind maps”, a stalwart from 1980s school lessons, where relationships between nouns, verbs or other sub-segments are mapped in a neural network. The MT engine then applies logic to analyzing the sub-segments to achieve improved sentence translation.

But this strategy of analysis was being applied to CAT tools before the introduction of AI. Advances in AI and MT are now merging into TMS and TEnT tools. MT is very much a separately marketed product requiring much more processing power, but the idea of using TMS and TEnT tools



without integrating MT has become a thing of the past, even though sub-segment matching and swapping can achieve a lot on its own.

This approach, combined with the desire to view and translate text in what appears to be full paragraphs laid out in-context and accessed in the original source document (in-context review – which has always been the panacea), may eventually lead to the demise of the table editor as a tool.

Guylaine: At the other, lower, end of the scale, there are content management systems or digital asset management systems that claim to offer a built-in translation platform. In reality, this means clicking on a button next to a string or sentence to open it for translation, typing in the translation with no assistance in terms of leverage or quality assurance, and clicking on the next sentence.

Q: Do you think the way TMSs and TEnT tools are used has changed in recent years? And how will this change in the future?

Steen: TEnTs have gained more functionality, with built-in on-the-fly MT, but fundamentally they are the same. The downside for translators is that the new generation of translator “workbenches” in cloud-based TMSs (Smartling, Memsource, etc) are a huge step back in terms of functionality.

These will undoubtedly continue to improve but, since the main target for these cloud TMSs are not translators (Worldserver, for example, was aimed at enterprises), developing advanced functionality is not going to be a priority, unless pushed for by a corporate buyer. The functionality of a TEnT might change, even significantly, in the future, but the fundamental principle of tying all essential functions together in a single application will stay.

As long as translation in the broadest sense continues to exist as a craft or an art, there will be a demand for the tools to do it, and the integration of different tools (currently a TM, TB [termbase], MT, etc) in a collaborative environment seems the best way.

Additional functions are likely to include better communication tools (such as integration with messenger apps) and collaboration at the authoring and review stages.

Jorge: I have noticed that lately TMS and TEnT solutions based in a web portal (MemSource, SmartLing, and so on) are gaining ground over native tools like SDL. There seems to be a move towards online working.

Jon: There are TEnT tools where the “assistance” is truly for the purpose of translation and linguistic quality assurance (LQA), and tools where the assistance derives mainly from the benefit of content connectors, sim-ship development and workflow management.

Some very successful software in workflow management includes Smartling and Transifex. Here connectivity to Web CMS and development tools such as BitBucket outshine the ease-of-use to the translator, or integration with AI and MT.

Only memoQ and Trados Studio are true translator workbenches, which focus exclusively on features to aid translation. In between, there is a gamut of online TMSs such as MemSource and XTM Cloud, combining workflow with their own translation tools.



There are also a number of localization-industry workflow management tools such as Plunet, XTRF, LTC Worx with features specific to managing translation workflows, but which rely on memoQ and Trados API (application programming interface) integration for a translator environment. In recent years, there has been a trend towards interoperability via API to the extent that companies now exist whose focus is not only on the localization industry, but whose product consists only of a universal API, for example Xillio.

It's notable that the true behemoths, while deadly serious about controlling the localization and translation industry, do not offer TMS and TEnT tools. Amazon, Apple and Google have all focused their attention on MT and AI.

Guylaine: Perhaps when all the connectors have been built and workflows have been fully optimized, translation technology providers will shift their focus back towards the translation environment. Each TMS and TEnT has its strong points.

Perhaps we will come to a point where competing translation technology providers will have studied each other's systems inside out and will all offer a “super tool”, complete with all the features loved by localization providers, whether they be project managers or translators.

Q: Will TMSs and TEnT tools continue to play a key role in the translation/localization workflow?

Steen: Yes, one way or another. The main thing will probably be an even clearer differentiation of the market, with the “bulk” market dominated even more by MT with AI-driven post-editing at one end of the scale, and “high-end” translations at the other, dominated by subject matter experts (SMEs) in a more evolved sense.

So rather than translators who are only experts in one domain, we will see translators, or transcreators, working in closer collaboration with writers and businesses to emulate the original message in slogans and campaigns as closely as possible.

Of course, translators will still need intelligent tools to store their ideas, research and previous work, as well as guidelines and other items (such as terminology and preferred phrases), so they can be easily recalled when working on a new job.

Jon: I’m not convinced. If we think of the broader idea of computer-assisted translation, the key concept is “assist”. It follows that there needs to be a human being in the process for the computer to assist, and the market and use of MT are reducing the need for human involvement in translation to a minimum.

I think traditional CAT tools will only survive on the sidelines of the localization industry, where humans will continue to translate in very specialist fields and will always require some assistance to speed up productivity.

Guylaine: I agree with Steen. It is too early to say that the translator is becoming extinct. Their role has already changed and will continue to change, but language specialists will always be required. I am not sure machines are capable of making the final call yet.

Q: Do you think TMSs and TEnT tools have adapted enough to MT developments? If not, what could they do better?

Steen: There are currently two types of MT engines. There are out-of-the-box solutions, like Google Translate, DeepL Translator, and Microsoft Translator. These are available via a browser or an app for anyone on holiday, and as plugins to TMSs and TEnTs

The other type is custom engines that require training and ongoing optimization by specialist engineers. These types of engines are still not widely available for individual translators or small language service providers (LSPs), but even this aspect is becoming easier to manage; some basic customization and feedback features are already available in some TEnTs. This is certainly an area where we are likely to see some growth.

Jon: In recent years, both memoQ and Trados have reinvented themselves as always-connected user platforms which control and channel access to many localization tools in the Cloud. This was their way of offering a connection to MT and AI, but also Wikis, glossaries and translation communities.

It is an attempt to keep the human editing environment relevant and to control revenue by selling added-value access to services which are increasingly done by other companies or indeed without any human input at all.

Guylaine: Google, Microsoft and DeepL offer customized solutions and make it possible to train one's own engines but these are tied into subscriptions. Some free MT services do however provide access to domain-based engines, so even though customization is not possible, it is nevertheless possible to connect to say, an English to German automotive MT engine.

Q: What changes do you see in the future for TMSs and TEnT tools?

Steen: As I said before, major communication and collaboration features will be integrated. These features already exist to an extent, but mostly via API. This means it's something larger organizations can handle, for example by connecting a TMS to [workplace collaboration tool] Slack. The communication features are also likely to branch out to include query-handling (sending queries directly to developers and content creators) and bug-fixing.

Jorge: You are right about the need for API in these tools, Steen. In my opinion as a developer, it's important for all these web solutions to deploy an API layer so that the users can customize and adapt the workflow and how the tool responds on their side.

In fact, recently I have spent a lot of time using this API layer to adapt the client workflow in our specific TMS (memoQ). This enables us to use our preferred tool, and then upload and integrate the translation in our clients' proprietary tools.

Jon: SDL recognized the direction of travel and invested as much as it could in AI technology which it placed in its Language Cloud. However, localization companies are too small to protect themselves against the true owners of AI.

In fact, it seems unlikely that localization companies will survive by owning this more expensive technology and limiting access to their own technology via a TEnT. Presently, Google charges a subscription for a professional Google Translate account. It is uncertain whether this model will persist or whether it will again become free.



In terms of encouraging people to freely contribute to the AI development process and to the data corpus owned by Google, it must compete with different financial models for collecting this data, such as Amazon's Alexa and Apple's Siri.

Bixby was a little late to this party, but now it is established, Google might decide that encouraging people to pay for and use new tech gadgets, which essentially crowdsource data, is a better model than subscription for services.

It only takes one of these behemoths to decide that a high standard of NMT ought to be free, or near free to incorporate into any API, and this will sound the death knoll for translation technology providers and localization companies struggling to afford such technology to make their software appealing.

At which point the focus will be on the user interface (UI) and user experience (UX) for a lightweight editor that sits on top of any product and provides access to AI cloud services. We may find that such an editor is provided free in all web browsers in the future.

Guylaine: There is one feature I still think is missing in TMSs and TEnT tools where MT is concerned. Some MT providers now push confidence scores about the quality of the MT output back into the translation environment.

If the score is 95%, the translator knows that they won't need to spend much time post-editing the segment. If the score is 75%, they know they will need to spend a little longer. But they still do not know which part of the segment they need to focus on.

MT will have taken a huge leap when translators are no longer able to differentiate a fuzzy match from an MT segment. The translator will be given an accurate quality score and will know which parts of the segment they need to concentrate on.

Q: Is there any new technology that will replace TMSs and TEnT tools?

Steen: AI is likely to play a greater role in MT, to assess quality and manage advanced workflows, but as specialized software for translation by specialist translators, a TEnT is unlikely to disappear any time soon.

Jon: NMT and AI are the future of the localization industry. Texts will be processed by NMT and then post-edited by AI. There will be no need for humans to process more than 90% of translations; they will simply be retranslated by machine each time, without recourse to specific product TMS.

Guylaine: I don't believe TMSs and TEnT tools will disappear any time soon. Perhaps we need to broaden our scope and think in terms of language rather than translation or even transcreation. I have not come across many platforms that make it possible to create content in several languages within the same project.

We are no longer talking about source and target languages here but of several source languages, each following specific brand guidelines. In October this year [2020], [Adobe entered into a strategic collaboration with Skyword](#), and this is definitely something to watch for the future.

Q: Do you think the TMS/TEt tool landscape will still be dominated by the same big players in 10 years' time? Are there any new kids on the block that are doing well?

Steen: The only relatively new software I know of is Kantan's Skynet, though it could be said that there is nothing new in any of the parts; they've brought together familiar elements and combined them to create something different. There is MT and AI to assess the output, combined with a radical "marketplace" concept where registered translators work on individual segments to make incredibly fast turnaround times possible (by translating segments out of context).

The new cloud-based TMSs like Smartling, XTM, Memsource and so on, are all doing business and to some extent replacing previous technology like Worldserver.



Software development has also changed a lot, so classic CAT tools like Passolo and Catalyst are no longer required – most source files can now be handled directly in TMSs/TEtTs and there is less of a need to do QA "engineering" separately (dialog box resizing, menu hotkey clashes), due to

dynamic resizing. However, it is still an issue in some apps and for smaller displays, but increasingly being resolved by building better rendering/preview in the translation editor.

Jon: There are no really big players in translation technology. SDL Trados, the industry leader by far, has just been bought by RWS – one of the largest localization enterprises – but these companies are minnows compared to the big tech giants that wield cash assets far greater than the GDP of most countries.

These corporations are not investing in TEtT tools. Their focus is on data corpuses and AI. One generation ago, it seemed as if human language was the hardest challenge technology faced. Nowadays, speech and language are small change to the latest supercomputers being built.

Guylaine: Translation technology providers are still being founded, for example Lokalise in 2017 and Smartcat in 2018. They offer cloud-based translation platforms and are popular with clients because of the level of connectivity they offer via APIs. Their translation environments do not yet compare with the likes of memoQ or Trados Studio, but they will only get better.

The translation landscape is changing and TEtT tools may have reached a plateau, but the translation industry will continue to evolve.

Q: As we wrap up, is there anything else you think might happen to TMSs/TEt tools and the way they are used in the future?

Steen: Tied to the general trends we've been discussing, we will probably see an even greater differentiation between those translators happy to job on new systems like KantanMT and other Cloud systems where you have no control over the tools (bar a few options for customization), and the technically advanced translators who continue to value having their own customizable tools where they can build and maintain their own assets.

This will be an issue for translators in the future, whether working in-house for a single company, for an LSP or as a freelancer. Few will like their main tools being locked to the jobs they do, and being unable to build and maintain their own assets – or use a tool that suits their work habits as SME, or simply as an experienced translator. Few professionals in any line of work like to use tools of others' choice. Chefs always use their own knives...

Jon: CAT is a very limited concept and will likely diminish in importance until it is vanishingly small. Human translation is likely to go the way of classical music. It will be taught alongside literary theory and translation, as an art for upper class entertainment. Traditional CAT tools will become specialist, expensive software with very low demand.

Guyline: I think we will witness a convergence of sorts. Traditional CAT tool providers will increasingly try to emulate cloud-based platforms. Cloud-based platforms will look at the features available in offline translation environments once they have exhausted the added value they can offer their clients, and will try to implement these features into their systems.

Biographies

Jonathan Chan, Chief Technical Officer, Alpha CRC



Over 25 years in the localization industry, Jonathan has purchased, configured and deployed almost every major localization globalization management system (GMS) on the market. Jonathan joined Alpha CRC in 1994 and, as Chief Technical Officer, he is part of a consultancy team which ensures customers are provided with the correct technology to integrate smoothly with their content management systems.

Steen Kemodel, Operations Research Manager, Alpha CRC



Steen started working at Alpha in 1996 as a translator and Scandinavian language manager. After gaining first-hand knowledge of most translation tools on the market from the translator's point of view, Steen moved to his current position in the Operations Team at Alpha where he spends most of his time helping translators get the most of their tools and improve the workflows.

Guyllaine Tritton, Solutions Architect, Alpha CRC



Guyllaine began her career as a Localization Engineer in 1998 and joined Alpha CRC in 2002 as a Senior Project Manager. Guyllaine has held the positions of Head of Project Management and Operations Manager at Alpha CRC.

As a Solutions Architect, Guyllaine's role is to work as a consultant to strategic accounts. She is also part of Alpha CRC's on-boarding team and advises on best practices, workflows and technology.

Found that interesting?

If you're not already signed up to our newsletter, register on our [News](#) page to receive lots more great articles that will help you on your journey to going global.