

## **From crawling to sprinting: Community translation goes mainstream**

**Nataly Kelly, Rebecca Ray & Donald A. DePalma**

Common Sense Advisory

*The notion that “two heads are better than one” is hardly new when applied to translation. The entire corpus of Buddhist sutras was translated into Chinese collaboratively by foreign and Chinese monks over a thousand-year period which began in the 1st century A.D. (Chueung, 2006). However, the dominant model used today for translation in the commercial sector depends on a process that largely inhibits collaboration. This article presents some of the latest findings from research on the state of community translation, based on multiple market research studies carried out over a five-year period, including a comparative analysis of 100 community translation environments and interviews with stakeholders. The research reveals that, over the course of the last several years, translation industry participants have been moving away from the traditional process toward a more dynamic and collaborative model. As community-based models have grown in popularity, distinct types of environments have emerged as well.*

### **1. The TEP model as de facto industry standard**

Over the last few decades, the majority of private-sector translation work has been carried out using the translate-edit-proofread (TEP) model (Kockaert & Makoushina, 2008). The translation step is typically performed by a single translator, followed by a review of the translation by a senior translator who is known as an editor. Finally, a third individual reviews the text to verify that numbers were rendered accurately and that no information appears to have been omitted. The TEP model emerged from the publishing field, and is based on Gutenberg’s printing requirements, in which the author submitted the manuscript, another individual typeset the material, and then a third party reviewed the galley proofs as many times as necessary to prevent any errors from appearing in the final print run (DePalma & Beninato, 2007).

The TEP model has various drawbacks related to translation quality and efficiency. Firstly, the individuals who are located downstream in the production chain usually have less information than those upstream. If the editor knows less about the topic or source text than the translator, he or she is likely to introduce errors instead of correcting them. Another common problem is that the reviewer may not have received the same set of instructions as the editor. The model also can be less than optimal from a

timing perspective. Each individual works on a task before handing it off to the next person in the process, as in an assembly line. As a result, translators, editors, reviewers, and production staff often spend a great deal of time waiting for the person ahead of them to finish instead of advancing the translation project. The TEP model is the most widely used model in the commercial sector today because of the way most translation services are purchased or sold. The vast majority of organizations on the demand side of the market, that is to say, consumers of translation services, do not contract directly with freelance translators—rather, they hire translation companies, also known as language service providers (LSPs) to produce the translations (Kelly & Stewart, 2011). The LSPs manage the translation projects, including the translators, editors, and proofreaders who carry out the TEP steps.

However, some quality steps are not carried out by the LSP, but by the client. For example, the client often assigns an individual to review the translation for accuracy and provide feedback, a process known as “end client review” (Bass, 2006). Typically, this review takes place after the TEP steps have been carried out, and the review is usually conducted by the client’s internal staff, usually located in the country where the target text will be used, or by a partner organization such as an in-country public relations firm or advertising agency. Perhaps the greatest drawback of the TEP model is that it only enables individuals to detect errors at certain checkpoints or at the end of a project. As Malcolm Williams writes in his book on translation quality assessment (TQA), “TQA has traditionally been based on intensive error detection and analysis and has therefore required a considerable investment in both human resources” (Williams, 2004, p. xv). While researchers find that TEP presents various challenges when it comes to quality control, they also highlight the fact that professional-level quality can be obtained through community translation models (Zaidan & Callison-Burch, 2011).

When errors are spotted at the end of a project, it can be extremely costly and time-consuming to fix them. With the TEP model, it is quite common for a client reviewer to spot a problem with a translation after it has already been delivered and the TEP process has been fully completed. Therefore, while the TEP model has been the prevailing model used in the industry, its disadvantages are sufficient in number and severity to cause the stakeholders on both the supply and the demand side—but especially the clients themselves—to consider other solutions.

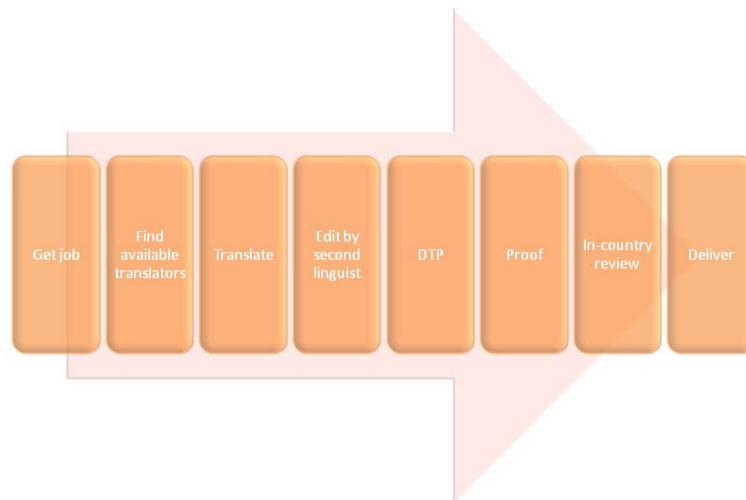


Figure 1: Timeline for Traditional Translation: Each Step Waits for a Hand-Off (DePalma et al., 2007).

## 2. Toward a Collaborative Translation Model

In response to the growing presence of online communities and the challenges associated with the TEP model, a collaborative translation model began to emerge in which translation could be performed entirely in a virtual, web-based or cloud-based environment (DePalma et al., 2007). Crowdsourcing is one example of such collaboration methods. In this model, collaborators can be either volunteers, employees, paid professional translators, or a mixture of all of these groups. Under this model, a project manager would first set up a project in the community, running the original document against translation memory files and optionally using machine translation to pre-translate the text. The project manager would then upload material to be translated, check a vendor database for the best resources and invite them to join the “project community,” which is likely to include translators, consultants, client reviewers, and desktop publishing staff. A project community differs from the traditional notion of a “project team” primarily in that project teams for TEP processes are usually organized by language, whereas a project community would include many teams working on different languages but would also allow participants to interact across language groups. In other words, the teams would not be segregated by language in a community-based model. Finally, the project manager would monitor the performance of the community to ensure that questions were

answered, files were available, and deadlines were met. While translation management system (TMS) tools facilitate much of the project management tasks with a traditional TEP process, in community translation environments, it is less typical for the project manager to rely on external software. Instead, the metrics and tracking tools are usually built directly into the environment itself. And, in some cases, specific tools are used to facilitate the community translation work.<sup>1</sup>

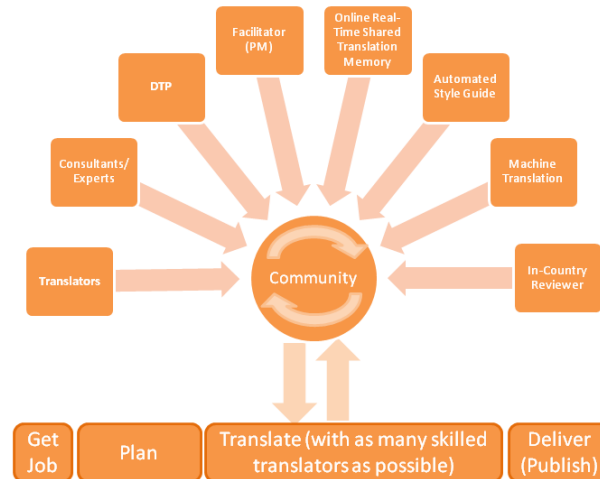


Figure 2: Timeline for Collaborative Translation: No Wait Time Between Activities (DePalma et al, 2007).

Under this model, translators would work asynchronously, for example, by putting 20 translators on a project for five days instead of having five translators work for 20 days. The project managers would rely on the other community members to proofread and improve each other's contributions to the translation. Also, individuals with subject-matter expertise would be invited to participate in the community to vet the translations, ask questions, and suggest improvements. The real-time interaction would enable errors to be corrected earlier in the process. In other words, reviewers would not be left downstream in the process, but could provide input during the process, which would also serve to help with translator training. Through this collaborative process, the editing stage would be eliminated, since quality improvement would take place at essentially the same time as translation. This enables organizations to produce translations much faster than they would by using traditional TEP methods.

## 2.1. The People: Optimizing Client-Vendor Relations and Staff Roles

Within the TEP model, companies adopt a sequential process where—in its simplest form—the work passes from the client to the LSP, who in turn hands it to the translator, who returns it to the LSP, who sends it to a separate editor who is also usually a freelancer, who returns it to the LSP, who conducts a final quality assurance check or “proof,” most typically using internal staff, and then delivers it to the client. In a collaborative environment, these hand-offs become redundant, since tasks do not need to be completed before they are transferred to the next agent. Under the community translation model, the client becomes part of the project community.

The other participant roles also change significantly under this model. Individual translators face seismic changes. They cease to be individuals working alone in their pods, and become active members of a community environment. Suddenly, in order to do their jobs well, instead of resolving their queries by looking terms and phrases up in dictionaries, they can enter questions within an online group environment, where they can ask follow-up questions and submit queries on specific use cases, complete with contextual information, examples, and exceptions. They can also review other people’s translation memory files, something that does not typically happen with the TEP model. Granted, even using the TEP model, translators can access terminology databases and their own translation memory files, along with any that are provided by the LSP for a given project. The major difference with the community-based model is that the translator’s ability to access shared knowledge repositories increases, taking the translators out of the silo of working alone at home and bringing them into a virtual workspace where they can benefit from real-time consultation with peers.

The project manager role also shifts significantly with the collaborative model. Instead of just shepherding files between the various parties, the project manager becomes a facilitator, the person who builds the team, keeps it on task, and brings into the community the resources required for each phase of the project. Most project managers working with the TEP model already use TMS tools in order to carry out the standard project management tasks related to the TEP process. However, because those tasks change in a community-based model, so does the role of the project manager. Some project management tasks become automated through the environment or portal itself. For example, a translation job in an online collaborative setting might be automatically marked as “complete” once a certain number of segments are translated by the group. In a TEP process, a project manager would carry out a manual check before marking a project complete using a TMS tool, and at that point, automatic messages might be sent to different project participants. In a community environment, there

might not be any need for messages to be sent to participants since they can access the same information in other ways—for example, via a dashboard or by hovering over a given project with their mouse.

The editor/reviewer function also morphs in a collaborative model. In this model, expertise is available not just toward the end of the process, but integrated throughout in the form of real-time shared translation memories and automated style guides. In the TEP process, a style guide is essentially a document with rules regarding grammar, usage, and style. In a collaborative model, many of these rules are built into the environment directly in order to prevent mistakes. For example, a TEP process for a company like Twitter might have a style guide that instructs a translator not to capitalize the word “tweet”. If a translator overlooks this, it might be caught by an editor or proofreader, or perhaps by an end client reviewer, but there is always a chance that it might not be caught. In a collaborative portal, an automatic message would appear if the translator tries to capitalize the word “tweet” to instruct the translator not to do this the moment they type it into the environment. Other rules might prevent the translator from submitting an item with an error until it is corrected, thereby focusing on error prevention instead of error correction. In some cases, authoring or controlled language tools might be used to enforce such rules. In other cases, the rules are built directly into the online environment.

In a collaborative model, human subject matter experts coexist and contribute to high-quality translations in the first pass. In fact, in a collaborative model, subject matter experts do not just do a final review at the end, but rather, are relied upon throughout the process to answer questions in a timely manner so that translators are not left waiting. Translators, thus, no longer rely on the editor’s eyes to catch their mistakes at the end, and must be trained to ask questions to help them translate correctly from the start. Typically, these subject matter experts are the same individuals who would be performing end client review in a TEP process. The difference is that their input is provided at the start of the project and throughout instead of afterward. Thus, the focus shifts from fixing translations that are potentially full of errors to creating a correct translation from the start, with the feedback and guidance of the subject matter experts at an earlier point in the process. It is important to note that subject matter experts and client reviewers are not present for every single project, regardless of which model is used. In other words, there are many translation projects following a TEP process that do not involve any client reviewers. Likewise, there are plenty of community translation projects that do not benefit from client reviewers either. However, for those projects that do require subject matter expertise, a community-based model ensures that said expertise is shared at an earlier point, at the time the translation is actually performed as opposed to after the fact.

Perhaps surprisingly, this model is not necessarily more expensive. Giving translators the ability to collaborate enables them to share their

expertise easily and quickly. The heart of the difference is in the collaboration itself. In online environments, translators can talk with each other and prevent errors from being made in the first place. In a traditional process, a translator working in isolation might make his or her best guess at a given translation. The only possibility for a colleague to catch the error comes further downstream in the process, at the editing stage or perhaps even at the client review stage. In a collaborative model, the mere fact that the translator has access to other human beings with knowledge means that the translator does not have to wait until the editing process to obtain feedback. The translator can simply post a question in an online forum and check the responses. The translator may gain additional insight from people who actually use the product or service for which the material is being translated. This enables the translator to produce the correct translation from the start, as opposed to waiting for someone to spot the error—which might never happen.

The role of the vendor manager becomes extremely important in a collaborative model. Vendor managers that work at most LSPs today tend to manage a large database of translators, editors, and other linguistic workers. Today, most vendor management that takes place is rather simple – resources are organized by language pair and sometimes by location and tool expertise. However, to segment the database and manage it in real time, organizing vendors by more specific criteria, such as gender, subject matter expertise, or national origin. In the case of gender, this might be important for certain types of projects—for example, where the content requires an audio component and a female voice has been requested by the client. Vendors must also be sorted using other criteria, such as past quality scores and client feedback regarding project performance. The real-time nature of the collaborative model makes the role of the vendor manager more important than in the TEP model. The key difference is that in a collaborative model, the work happens in real time, and thus the vendor manager's role becomes more prominent.

The nature of teams is also different in a collaborative model from the way teams work in the TEP model. With the TEP model, translation teams usually consist of translators working within the same language combinations. With a collaborative model, the ability to collaborate and learn from each other can cross language pairs. For example, Spanish translators can leverage solutions found for Portuguese or Italian translations. Japanese translators can be warned up-front about bugs and poorly written or ambiguous source language that the German team finds while localizing and translating. Chinese client-side reviewers can communicate with the project consultants to clarify terminology issues in real-time. For example, a Chinese end client reviewer might warn project consultants that an image or term in Chinese is acceptable in mainland China but has negative connotations or would not be appropriate for Chinese communities in other countries. This finding could affect not only

the Chinese language content but also content written in other languages or problems that could exist in both Simplified and Traditional Chinese.

## 2.2. Process: Improving project management through communities

Speed is a critical issue for most clients of translation services (Kelly, 2011). Timelines are often tight, which means that project managers often resort to individual heroics in order to get work done. In the TEP model, cramming a lot of work into a short timeframe is typical for the delivery of most projects. In contrast, collaborative translation requires more advance preparation and organization before a project starts: recruiting and setting up communities, sourcing and training the best talent, configuring automated style and terminology tools, pre-translating material, and populating translation memory files that can be used for the project. This is most typically done not using the traditional commercial off-the-shelf tools that are common for TEP processes, but rather, by creating environments in which these rules and resources are available directly within the system itself. Nearly all collaborative translation environments that exist today have been developed in-house.

As Table 1 shows, several core elements related to project management are different with a collaborative model. The traditional TEP model focuses on *error correction*, whereas the collaborative model focuses on *error prevention*. A community that is optimized for collaborative work has automated systems in place, in addition to a larger number of eyes that can catch errors at the time they are introduced, instead of at a later phase, which is the case with the TEP model. The community serves to “self-correct” as the translation process unfolds. This is a very different process from the one used under the TEP model, which involves catching errors after the translations have been produced. For example, in a typical TEP process, let us assume that a translator uses an incorrect term for a translation, and let us also assume that this term is not available in any of the terminology databases, style guides, or translation memory files. The translator might do this accidentally, simply because the translator cannot locate the term in any existing resources and does not have the ability to submit a query about the term, in the rush to get the translation off to the next phase, in which the editor will review it. The translator will submit the translation with the error, and it is possible that the editor and proofreader will not detect the error either. When the translation reaches the end client reviewer, the end client reviewer spots the error and requests a correction. At this stage, the error may already have been recorded in the translator’s translation memory file and terminology database. So, to purge the error, the translator will make the correction, re-submit the file to the editor, and the process essentially starts over once again. In many cases, more errors and problems are introduced at the time the corrections are made. For



example, an editor might replace a single word but inadvertently include a typo. With a collaborative model, the error would much more likely be caught at the time the translation was actually created, because the translator would have had the ability to ask for help from the other members of the community, including other translators who might have already found acceptable ways of translating a tricky term in a similar language, as well as the same end client reviewers and subject matter experts who would sit at the end of a TEP process instead of being available at the start.

**Table 1:** Process-related issues for collaborative translation

Process Issue	TEP Model	Collaborative Model
Reliance on and advanced preparation	Low	High
How translation activities are performed	Sequential process	Parallel / Simultaneous activities
Quality control focus	Error correction	Error prevention
Information availability	Limited by language pair or project team	Shared across the entire community
Decisionmaking	Select participants make subjective choices based on their individual knowledge and expertise	Large community of translators and experts arrive at group consensus
Feedback provision	Comments flow from all parties back to the project manager (single point of failure)	Real-time tagging allows information to be shared openly by all project participants
Compensation	Price per word model; more words equates to more money regardless of actual effort	Compensation reflects contributions to the process that are not necessarily tied to translated words (contextual insight, clarifications, edits, etc.)

Instead of reporting issues to the project manager as happens in the TEP model, translators in the collaborative model tag and insert comments in real time in the original document, thus helping translators in other languages who might otherwise face the same difficulties. Collaboration

facilitates the elimination of silos of knowledge and makes the whole system or community smarter and more efficient.

The per word/per line/per page pricing model that characterizes the TEP model is not well-suited to the collaborative model, which rewards more than just translating individual words. The collaborative model prizes contributions to the process that include contextual support, answers to questions from the community, clarifications, and edits. This model of rewarding supra-translation activities provides an incentive for community members to contribute beyond just the language transfer. The vast majority of work that is performed through collaborative translation is carried out by volunteers. More often than not, these individuals are not professional translators, but rather, fans or members of existing online communities that exist to support the discussion of a given product, service, or cause. It is important to understand that community translation is not emerging as a “replacement” for the TEP model, at least, not right now, but rather, as a natural consequence of the increased presence of global online environments in which people gather online. Typically, these individuals participate in community translation projects not because they are paid to do so, but because of some other reason. For example, perhaps they are motivated, to see a software product released in their native language, and they realize that by contributing to the translation, they will receive it faster. In other examples, such as the 2010 earthquake in Haiti, they volunteer to help fellow human beings and/or loved ones in emergency situations through community translation. In many cases, the volunteer translators themselves are subject matter experts, with knowledge of the product or topic that is highly specialized. Whereas a professional translator would typically study product documentation and terminology in order to familiarize himself or herself with a product, the individuals who already use these products on a daily basis already understand what different terms mean, and are therefore in some ways more qualified to translate the terminology than professional translators might be.

### **2.3. Technology: Leveraging automation**

In the collaborative translation world, translations are done more efficiently because of technologies that enable these communities to work together seamlessly. In this model, translation happens on the web, all the time, by linguists, subject matter experts, and individuals who are passionate about a particular product, issue, or cause, who can be available regardless of location or ownership of desktop translation tools. One could argue that collaboration is possible even in a TEP model. However, the technology infrastructure available through Web 2.0 allows more extensive and continual collaboration through asynchronous workflow and remote

operations that are much more flexible and reliable than the typical technologies used with a TEP model (Hartley, 2009).

In a collaborative model, translation communities rely on commonly used technology like instant messaging (including via social media platforms), SMS, discussion groups, RSS feeds, social tagging, content categorization, and voting. Since 2005, language software developers have been working on a wide range of translation memory and automated translation innovations to enable translation assets and workflow to reside comfortably and natively in the cloud (Sargent, 2010). In other words, technology developers are working to make it easier for translators and other collaborators to interact in a web-based environment and work on projects without the need to install any desktop software (see Table 2).

**Table 2:** Technology-Related Issues for Collaborative Translation

<b>Technology Area</b>	<b>TEP Model</b>	<b>Collaborative Model</b>
Communication with other group members	E-mail / Telephone	Discussion group, chat, social and professional network, Skype, Wiki, Groove
Translation memory	Desktop translation memory applications like Trados and Déjà Vu	Next-generation translation memory tools from companies like Alchemy, Elanex, Kilgray, Lingotek, Lionbridge, and XML-Intl
Style compliance	Style guides	Authoring tools such as acrocheck and AuthorAssistant
Content creation	Standalone desktop content creation tools such as Word and FrameMaker	Open documents; XML everywhere; DITA

### 3. Facebook's Early Adoption of the Collaborative Translation Model

Several companies—including Facebook, Sun, Microsoft, and Plaxo—were early adopters of collaborative translation models and participated in detailed interviews with the authors about their models, including their reasons for adopting community translation processes, their challenges, lessons learned, and their recommendations for others. The findings from all four of these early adopters were documented extensively in a more comprehensive report, which includes the case study presented below

(DePalma & Kelly, 2008). Of these community translation pioneers, perhaps no company's model has received as much attention as that of Facebook. The firm even went on to apply for a patent for the specific community translation platform that it developed (Wong et al., 2008). For this reason, we have elected to describe Facebook's experience in more detail in the section that follows, focusing on the elements of its adoption of a collaborative model that serve to highlight the differences between this model and the traditional TEP model.

Facebook blended its initial reliance on translation agencies with the linguistic skills of an enthusiastic community whose members wanted to use Facebook in their languages. The company's translation process evolved into a four-step process:

- (1) **Translation.** Facebook users translate strings and sentences in the interface and help files. The site has many members, so there can be multiple translations for the same English text;
- (2) **Voting.** Members of the community vote on the translation alternatives. Popular translations rise to the top of the selection pool. Facebook maintains a leader board that shows who's who among the volunteer translators. Facebook members can override the most popular translation, but only with a compelling justification, given that these end users are considered to be the de facto experts;
- (3) **Collaboration.** Facebook's user-translators review and solve trickier or more difficult translations on discussion boards;
- (4) **Review.** Professional translators review all translations. They review which community members generate the most popular translations and the scores they receive.

Facebook provides a platform for developing applications that site members can install and run. The company used this same capability to deliver the Translation App, the application that underlies its collaborative translation process. The application works as a home page and staging point for the translation community. It highlights Facebook's three major localization steps—glossary creation, content translation, and testing and verification—and clearly indicates the progress the translation community is making toward the completion of each step, using a dashboard-type display.

Facebook built various quality control mechanisms directly into the Translations App. It automatically presents glossary entries with definitions and approved translations for the technical terms found in each translated element. And, it puts automatic checks in place to verify that capitalization and punctuation conform to style specifications for different element types. For example, if a common term such as "Like" must be capitalized in order to be considered correct, the tool would not allow a translator to submit a translation of this term without the term being capitalized. It would flag it as unacceptable and prompt the translator to fix it before it would be

considered an acceptable submission. Professional reviewers use the same application as the community translators.

Determining what content to translate and into which languages can be one of the most daunting tasks for any company. To focus its planning efforts, Facebook added two triage steps to its decision-making process. Facebook began by determining which content assets had to be offered in other languages. Its analysis found 300,000 words destined for other languages, comprised of 125,000 words in user interface elements, another 125,000 in legal and help content, and 50,000 words of miscellaneous content (DePalma & Kelly, 2008). These categories reflected top translation priorities, with user interface text as the first—and in some cases, the only—content to be translated. When it comes to language selection at Facebook, not all languages are created equal. Each receives a level of support based on its strategic importance, determined largely by the number of local internet users. Supported languages get full internal support. Unsupported languages get lower levels of support or no support at all.

Facebook wished to accomplish three specific goals with its community translation initiative. First, the company wished to deliver localized sites in a fraction of the time required with conventional methods such as TEP. Notably, cost reduction was not among the factors. Second, the company wanted to ensure high quality, and to use the most appropriate terminology for the audience in question. Third, the company was motivated to bring the social networking tool (i.e., Facebook website) to more communities around the world (DePalma & Kelly, 2008).

Facebook found that collaborative translation delivered dramatic improvements in speed. Initial discussions with language service providers revealed that using traditional processes would enable the first translation to be delivered in a matter of months. Facebook believed that the size of the community would allow it to deliver the translations in days or weeks instead. Using collaborative translation, volunteers completed the Spanish and German localizations in just one week, while the French team took just 24 hours to produce its language variant (DePalma & Kelly, 2008).

The Facebook translation system brought together a unique combination of people, processes, and technology. The company needed translators with linguistic skills and subject-matter expertise, both of which could be found in its user community. Facebook confirmed that collaborative translation offered significant quality benefits because its community provided it with expert reviewers. The resulting translations met the company's requirements—they were clear and unambiguous and conveyed the meaning of the original, but they did not sound unnatural or forced.

Facebook devised a system with many conventional translation steps, such as style specification, glossary development, and linguistic quality assurance. The organizers added two important elements—immediate feedback and correction. Facebook also wanted to simplify the job for its voluntary translators, so its software platform eliminated the need for

manual processing beyond translation. Due to its use of collaborative translation, Facebook was able to extend the benefits of localization to a larger international community. More Facebook users were able to interact with the service in their own languages; the process itself was conducive to creating greater brand awareness through social networking and word-of-mouth marketing. For example, a U.S.-based member working on a translation project for Colombian Spanish could easily invite friends in other countries to join the network and participate in the translation project for other national varieties of Spanish.

Application developers also benefited from Facebook's foray into international markets. The company gave Facebook members who are interested in localizing their own applications access to the same technology, language assets, and processes that it used for its own localization efforts. Facebook applied its vision and expertise in social networking to mobilize communities of user-translators. Several facets of its program stand out. One of the noteworthy features of its model is that the translation takes place fully within the community itself. While a language version is "under construction", only those members working as user-translators on that translation project can actually see what's going on with the target language.

Another unique element is that Facebook offers multiple variants of Latin American Spanish. This approach not only customizes the content for the distinct user groups, but reduces the chances of community infighting due to one nationality dictating the fate of the terminology and idioms that other members would have to live with. Facebook also uses real-time tracking to stay on top of progress. The company monitors the translation process closely, keeping this information in front of translators as a way of marking progress and keeping participants focused on the tasks at hand. The company also gives great visibility to the volunteers who do the most to make translation happen. Each language has a leader board that monitors progress and recognizes the leading contributors. It logs the total number of winning words and phrases they have submitted and the number of votes they have received.

However, in spite of its many benefits, the Facebook project also had its share of problems. First, the company found that it had to actively pursue contributions and engagement in order to keep contributors active. Second, the maintenance of the platform itself was a large chore, involving access control, scalability, and security issues. And, of course, the company struggled with some of the same quality assurance concerns that all companies deal with. Facebook also faced criticism of its user-based translation effort, primarily from observers who viewed the initiatives as a form of doing translation on the cheap. However, it is important to consider the fact that Facebook realized minimal or no cost savings from its community efforts. Its large investment in technology offset most of what it saved from getting translations from its members. It did benefit financially

from translating unsupported languages, automating the localization process, and prioritizing text segments to be translated, so that localized sites can be operational sooner. In fact, commercial language service providers for supported languages translate more than half of Facebook's content, and they are paid for it. The company also engages these providers to evaluate the work of its volunteer translators (DePalma & Kelly, 2008).

#### **4. Community translation enters the mainstream**

Although community translation, also commonly referred to as crowdsourced translation due to its use of a "crowd" or a "community" to perform translation tasks, is still in its infancy from some perspectives, it is definitely spreading and becoming integrated into more traditional translation workflows. A recent study of 104 community translation platforms reveals a typology with three different environments (Ray & Kelly, 2011). The first is driven by a cause (often charitable), the second is geared toward a for-profit product or service, and the third is centered around outsourcing.

In the widely-used *cause-driven* type of collaborative translation environment, people choose content that interests them, and they translate it at their convenience. The material may relate to the latest world disaster, non-profit activities, newspaper content, videos, or anything else that attracts their attention. Volunteers are generally not remunerated in any way, but they are often recognized in some public fashion on the organization's website.

Media content is a popular cause-driven source of crowdsourced translation in many parts of the world. It covers printed content (newspapers, magazines, books), as well as audio and video. Also known as "fansubbing" or "fandubbing" when applied most typically to animation films, these sites are common in markets where there is a great demand for foreign language content that creators either don't have the resources to localize or haven't yet found what they believe to be the appropriate business model to do so.

Implementations in the *product-driven* category of collaborative translation are generally projects in which for-profit companies recruit and manage a crowd. The members of the crowd must often match a specific profile to be chosen to translate websites, software, documentation, or other content. Companies such as Adobe have adopted this model (Ray & Kelly, 2011). The volunteers are often remunerated through free products, services, or promotional merchandise from the company. As in other forms of the collaborative model, they are recognized in some fashion through leader boards, the company website, or credited on the piece of content they have helped to translate.

There are also outsourcing portals that offer crowdsourced translation as either their principal revenue driver or as one of their service offerings.<sup>2</sup> Anyone can purchase crowdsourced translation from such firms. The people who perform the language-related services are generally remunerated with actual money, not just in giveaways, praise or recognition within the community.

The actual working environment for translators, editors, and reviewers depends somewhat on the model under which they are working, as well as the type of content to be delivered. Our review of 104 community translation environments showed that there are currently three basic environments: wiki- or forum-based, database-driven, and full-blown collaborative translation platforms. The wiki-based model tends to be the simplest, but is very difficult to scale beyond just a few languages and one or two products or services. It requires no other investment than some type of wiki software or a forum, plus a part-time community moderator. It also makes it fairly easy to gauge how committed a user community might be if a larger crowdsourced initiative were launched. The database type was by far the most common environment of the sites we reviewed. These environments give users either a simple interface or a full-fledged dashboard to carry out the translation work.

The full-fledged collaborative translation platforms are the least common, but offer the most benefits, since they are built with crowdsourced translation in mind. This means that there is an underlying translation platform that provides some type of terminology management, the possibility of translation memory re-use, access to machine translation, or application programming interfaces to allow for fairly easy integration. These platforms are also usually more flexible and feature-rich when it comes to integration with existing content management systems. The downside is that this model may require more upfront and ongoing investment in maintenance or fees for cloud services. However, it pays off in the medium to long term by allowing for scalability that the other models cannot allow. Best practices for all of these environments are starting to emerge in the areas of workflow, tool design, personnel requirements, community engagement and support, remuneration, recognition, and governance.

## **5. Conclusions regarding the future of crowdsourced translation**

In spite of growing adoption rates on the demand side of the market, crowdsourced translation is not yet very popular among language service providers and groups that represent professional translators—in other words, the supply side of the market feels that this model is a threat to the status quo (Kelly, 2009; Zetzsche, 2009). However, on the demand side of the market, research carried out on early and recent adopters of



crowdsourced translation shows that these firms are realizing its many benefits (Ray et al., 2011). In particular, crowdsourced translation allows them to engage meaningfully with potential customers and constituencies on a long-term basis.

One important element of crowdsourced translation is that it empowers the actual end users by allowing them not only to determine which terms and phrases are most accurate for their content, but to decide which content to translate in the first place. Adobe, for example, has already gone a step further with its collaborative translation in China by encouraging its users to choose the content that they think should be translated, rather than mandating what they should do. Sometimes, not all levels of content must be translated—for example, certain sub-menus that are buried deep within a software program might rarely be accessed by users, and so perhaps might not merit translation. Norton has integrated a terminology component to encourage its community to suggest, discuss, and validate terms. This ensures that the terms that will appear in its products are not merely terms selected by whichever individual translator happened to work on a project, as would usually be the case in a TEP process, but rather, terms that have been suggested by actual consumers of their products and services and are more likely to be understood by other prospective customers. HootSuite allows its communities to determine which languages should be implemented next. Once it reaches a certain number of community members that use a given language, they use this feedback to initiate a community or collaborative translation process (Ray et al., 2011).

In general, smaller high-tech companies (such as HootSuite) and non-profit organizations (such as Kiva.org) make up the majority of early and current adopters for collaborative translation. However, the pressure to provide more local products, services, and content at a faster rate is forcing many organizations to consider more innovative solutions, such as collaborative translation. Of course, collaborative translation is not the only solution that clients are evaluating, but rather, one of many potential options that they are exploring. Other solutions to address the challenges of speed, scale, and scope include increased usage of automated or machine translation, centralized procurement of translation services, and enhanced investment in technologies, such as global content management systems, terminology databases, translation memory tools, authoring software, and TMS products. Some organizations are also employing more in-house translation staff in order to meet the demand for quick turn-around translation within their own time zones. In other words, collaborative translation should be viewed, not as a dangerous development that threatens the profession of translation, but rather, for what it is—the market's response to specific challenges that are not being satisfactorily addressed through traditional TEP models.

It should also be noted that it is still quite rare, on a project by project basis, for community translation to “displace” work that could

otherwise be handled via a TEP model by professional LSPs. The fact is that community translation processes are used within online communities and lends themselves to work that can readily be handled in such spaces. In other words, these projects are handled within communities that already exist—it would not be possible for organizations to simply round up volunteers and convince them to join communities for their causes or products, nor would it be cost-effective to do so for the majority of projects. Community translation is emerging for very specific purposes and in very narrowly defined contexts. However, more and more communities are being created each day—not only on product- and cause-specific websites that are managed directly by the client organizations, but also on social networks that are external to the organizations themselves. As the world becomes increasingly connected, and as internet users for more languages become accustomed to having the ability to access content in their own languages, the willingness to sit back and wait for organizations to go through a traditional TEP process will continue to fade. If individuals already know a product and speak multiple languages, they will be tempted to go ahead and suggest terms and translations themselves in order to get the product into their language(s) faster.

Take, for example, a bilingual Latino in the United States who uses an ancestry tracking website and wants to capture family history data from a Mexican grandparent who only speaks Spanish. This person is motivated to get the interface translated into the grandparent's language and will likely be happy to spend some of his free time in order to capture his own family history data more easily. Or, consider the Chinese- and English-speaking user of a mobile phone application who would love to have her children play the game, but her children have not yet achieved full proficiency in English. Would she be willing to translate 50 phrases into Chinese, her native language, so that she can share this game with her children? Or, take the case of the mobile health worker in India who treats patients who only speak Tamil, but the simple preventive steps that she needs to share with them are available only in Hindi. Would she be willing to translate this information into her native tongue in order to save the lives of more of her patients, versus repeating this information to them verbally over and over? There are many reasons why people volunteer to translate information in community translation environments. In some cases, it gives them access to content they wish to access, enjoy, or share in other languages. For many languages, the content would not actually be translated without community translation, because the market demand does not exist in some cases.

In conclusion, the differences between the TEP model that is the status quo in the market today and the emerging community translation model are vast. However, to understand these differences, it is important to also highlight the market drivers that are causing community translation to gain popularity, as well as to acknowledge that the scenarios in which

community translation is used, while somewhat limited at the present time, will become increasingly common as time goes on.

## References

- Bass, S. (2006). Quality in the real world. In K. J., Dunne (Ed.), *Perspectives on localization* (pp. 69-84). Amsterdam: John Benjamins.
- Bauer, W. (1964). *Western literature and translation work in communist China*. Hamburg: Institute für Asien Kunde, Alfred Metzner Verlag.
- Cheung, M. P. Y. (Ed.). (2006). *An anthology of Chinese discourse on translation*. (Volume 1: From Earliest Times to the Buddhist Project). Manchester: St. Jerome.
- DePalma, D. A., & Beninato, R. S. (2007, December). *Collaborative translation*. Retrieved from <http://www.commonseadvisory.com/AbstractView.aspx?ArticleID=934>
- DePalma, D. A., & Kelly, N. (2008, August). *Translation of, by, and for the People*. Retrieved from <http://www.commonseadvisory.com/AbstractView.aspx?ArticleID=889>
- Hartley, T. (2009). Technology in translation. In J. Munday (Ed.), *The Routledge companion to translation studies*. New York: Routledge.
- Kelly, N. (2011, January). *Pick up the pace: Speed as a differentiator for LSPs*. Retrieved from <http://www.commonseadvisory.com/AbstractView.aspx?ArticleID=1310>
- Kelly, N., & Stewart, R. G. (2011, April). *Translation vendor management: What clients really think about language service providers*. Retrieved from <http://www.commonseadvisory.com/AbstractView.aspx?ArticleID=1407>
- Kelly, N. (2009). *Freelance translators clash with LinkedIn over crowdsourced translation*. Retrieved from <http://www.commonseadvisory.com/Default.aspx?Contenttype=ArticleDetAD&tabID=63&Aid=591&moduleId=391>
- Kockaert, H., & Makoushina, J. (2008). Zen and the art of quality assurance: quality assurance automation. In *Proceedings of the thirtieth international conference on translating and the computer Conference Information: Translating and the Computer* (30<sup>th</sup> ed), ASLIB, London.
- Munro, R (2010). *Crowdsourced translation for emergency response in Haiti: the global collaboration of local knowledge*. Paper presented at the Association of Machine Translation for the Americas Conference, AMTA, 2010. Retrieved from <http://amta2010.amtaweb.org/AMTA/papers/7-01-01-Munro.pdf>
- Ray, R., & Kelly, N. (2011, February). *Trends in crowdsourced translation*. Retrieved from <http://www.commonseadvisory.com/AbstractView.aspx?ArticleID=1316>
- Sargent, B. B. (2010, February). *The new vortex around TMS-in-the-Cloud*. Retrieved from <http://www.commonseadvisory.com/AbstractView.aspx?ArticleID=834>
- Williams, M. (2004). *Translation quality assessment: an argumentation-centred approach*. Toronto: University of Ottawa Press.
- Wong, Y., Grimm, S. M., Vera, N., Laverdet, M., Kwan, T. Y., Putnam, C. W., Oliván-Lopez, J., Losse, K. P., Cox, R., & Little, C. (2008). *Community translation on a social network*, U.S. Patent Application Serial No. 329288. Washington, DC: U.S. Patent and Trademark Office.
- Zaidan, O. F and Callison-Burch, C (2011). Crowdsourcing translation: Professional quality from non-professionals. In Association for Computational Linguistics (Ed.), *Proceedings of the 49th Annual Meeting of the Association for*

*Computational Linguistics*, (pp. 1220-1229). Portland, Oregon, June 19-24, 2011.

Zetzsche, J. (2009, November/December). About missed opportunities. *ATA Chronicle*. Retrieved [http://www.internationalwriters.com/toolkit/09\\_Nov\\_Dec\\_MissedOpps.pdf](http://www.internationalwriters.com/toolkit/09_Nov_Dec_MissedOpps.pdf).

---

<sup>1</sup> Examples of community translation tools include but are not limited to: CrowdIn (<http://crowdin.net/>), CrowdSight (<http://www.welocalize.com>), GetLocalization (<http://www.getlocalization.com>), GlotPress (<http://trac.glotpress.org/>), LingoTek (<http://www.lingotek.com>), Transifex (<http://www.transifex.net/>). Inclusion in this list does not constitute an endorsement by the authors.

<sup>2</sup> Examples of outsourced community translation providers include but are not limited to: Crowdfunder (<http://crowdfunder.com/>), Microtask (<http://www.microtask.com>), MyGengo (<http://www.mygengo.com>), OneHourTranslation (<http://www.onehourtranslation.com>), ServioTranslate (<http://www.serv.io/translation>), SpeakLike (<http://www.speaklike.com>), and Zhubajie (<http://www.zhubajie.com>). Inclusion in this list does not constitute an endorsement by the authors.