Language of the Crowd:
Crowdsourcing in the Translation Industry
White Paper

By Matt Bramowicz, Translation Cloud, LLC

December 2013

Table of Contents

Figures

Introduction

There was a time not too long ago when translation was a solo task, completed by one translator and perhaps a proofreader to catch any errant grammatical mistakes. With the onset of technology, however, there has been great strides in developing new ways to utilize the web's vast connectedness in order to provide translations with more speed and efficiency.

From this technological expansion, crowdsourcing has emerged as one of the more promising, yet divisive methods within the translation industry. Like any emerging trend, there are always going to be proponents and critics who either champion its potential, or criticize its flaws into extinction. However, as the paradigm of the industry itself changes to reflect the global connectedness of the web, it seems more and more likely that the crowdsource solution is becoming less of an innocuous trend designed to gain a competitive advantage, and more of a cost-leveraging necessity to compete on the global scale.

Translation Cloud explores the evolution of crowdsourcing in the translation industry, the current positive and negative attributes, and the ways in which crowdsourcing has successfully shifted the paradigm of the industry.

Changing the System: Outsourcing

Technology is always changing, and with change there will inevitably be resistance to it from those who are accustomed to the old. While this is the case, it does not mean change is always for the better. It is the job of any industry, including translation, to discriminately determine for the sake of its business which changes are useful and which are not. However, those who do not embrace the concept of change will almost certainly be left out in the cold. It's estimated that companies that are in favor of technology grow nearly three times as fast as companies that have mixed feelings about it or are against it.[1]

One of the major industrial changes to have taken place in the United States within the past few decades is that of outsourcing, which is the precursor to crowdsourcing. Outsourcing has consistently gained popularity for many businesses, with most spending 60 to 100 percent of their total localization budget on third-party services.[2]

As demand increased over the past few years for translation services, due in large part to the rising global trend, translation companies themselves began outsourcing many of their projects to the fast growing agencies overseas which offered significantly lower rates. In fact, according to a recent report by Common Sense Advisory, even the United States federal government spent US$4.5 billion on outsourced language services from 1990 through 2010 alone. [3]

IBISWorld estimates that the number of translation service providers have grown at an annualized rate of 4% to 49,405 enterprises within the last 5 years.[4] The increase can be attributed to a number of factors, including the global recession causing many individuals to seek additional wages to supplement their income, and the flexibility the internet provides which allows for translators to easily work from home. By utilizing this outsourcing model, US companies leveraged their newly lowered rates against the growing demand and gained the competitive advantage while still increasing revenue.

However, as the growing demand continued to rise, price quickly became the deciding factor for clients, and the lowered outsourcing rates became the standard rather than the exception.

But while the lower outsourced rates have led to an almost standard business model, outsourcing brings with it greater risks to accuracy and accountability. Since you are dealing with outsourced companies, you never truly get to "meet" the company you are

---

[1] Language Service Provider Growth Factors, Common Sense Advisory
(http://www.commonsenseadvisory.com/AbstractView.aspx?ArticleID=736)
[2] How to Benchmark Your Localization Budget, Common Sense Advisory
(http://www.commonsenseadvisory.com/AbstractView.aspx?ArticleID=754)
[3] U.S. Federal Government Spends US$4.5 Billion on Outsourced Translation and Interpreting Services, Common Sense Advisory
(http://www.commonsenseadvisory.com/Default.aspx?Contenttype=ArticleDet&tabID=64&moduleId=392&Aid=1076&PR=PR)
[4] Translation Services in the US: Market Research Report, IBIS World

procuring services from, rather dealing with them solely through the web or by telephone. This leaves the door open for many "scam" companies to dubiously position themselves as legitimate translation vendors.

Typical scams and scam processes[5]

- Scammers **search** for translators that publish their details on translation portals.
- They **edit** the available resumes/CVs, often (though not always) removing the original translator details such as phone numbers, home address, photos, etc.
- They create a **fake email address** that is very similar to the original translator's name and/or email address – quite frequently a Hotmail address rerouted from a Gmail address.
- Scammers **send** the application with the edited resume to potential client or agency offering services with low rates. There are examples where the cover letter from the original translator was directly reused by the scammer.
- Once the scammer is sent a job, they typically **deliver** machine-translated text, or in some cases actually "hire" the real (scammed) translator to do the job. In the latter case, the real translator receives only a fraction of the actual amount due or doesn't receive any payment.
- When there is a quality complaint, the responsibility is on the scammed translator who in most cases is completely unaware of this practice. This can negatively affect his or her **reputation**.
- There are also companies contacting both freelancers and clients. To the freelancers they offer "marketing" of their resumes and promise regular work; to the clients they offer cheap, fast and high quality translations and other services. While many are legitimate, given the proliferation of websites promising low-cost translation, this is an area of recent concern.

Source: http://elanex.com/EN/wp-elanex-translator-scam.aspx

These sorts of scams occur frequently enough that more innovative solutions have been sought. As a result, businesses have begun to look at other viable alternatives that provide low translation rates, while simultaneously providing a more accurate and trustworthy translation solution.
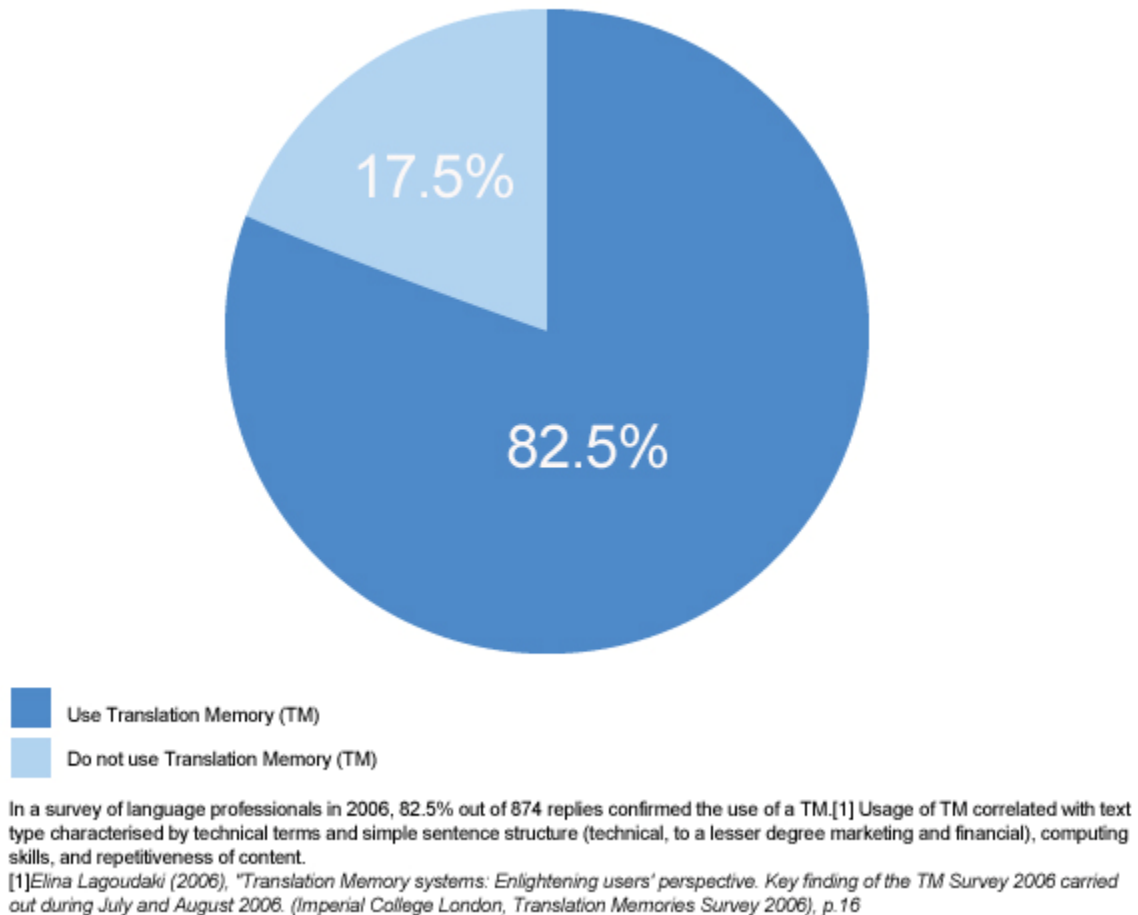
Translation Memory

Translation memory software is one recent advancement that has been increasingly used to help automate industry services, especially for large-volume projects. Translation memory stores translations of words and phrases in its database, and when that word or phrase is repeated, the translation memory system merely pulls the translation from its database rather than requiring the translator to re-translate that portion.

---

[5] White Paper: Translator Identity Theft and Scams, Elanex (http://elanex.com/EN/wp-elanex-translator-scam.aspx)

Figure 1.

PERCENT OF LANGUAGE PROFESSIONALS WHO USE TRANSLATION MEMORY

17.5%

82.5%

■ Use Translation Memory (TM)

■ Do not use Translation Memory (TM)

In a survey of language professionals in 2006, 82.5% out of 874 replies confirmed the use of a TM.[1] Usage of TM correlated with text type characterised by technical terms and simple sentence structure (technical, to a lesser degree marketing and financial), computing skills, and repetitiveness of content.
[1]*Elina Lagoudaki (2006), "Translation Memory systems: Enlightening users' perspective. Key finding of the TM Survey 2006 carried out during July and August 2006. (Imperial College London, Translation Memories Survey 2006), p.16*

While some competition exists through various types of translation software, industry operators contend that these systems can still sometimes give poor translations and will not provide serious clients with a proper translation due to variations of context and intention throughout any given document. For instance, "branch" could be used to describe both a tree limb and a division within a company. However, it is not possible for translation memory software to be able to differentiate which translation would be appropriate in a given instance.

Cloud Computing

As technology has advanced over the last decade to the point where network connections have become fast, reliable, and more dynamic, industry specialists began to lean toward more scalable solutions which utilize the full potential of the web. The solution seems to have been found through utilizing cloud-based computing systems, in combination with Software-as-a-Service business models and global micropayment systems such as PayPal.

Since the introduction of cloud computing a few years ago, more and more businesses have begun moving toward cloud-based platforms, and with good reason. Cloud-based systems are much more elastic than their traditional counterparts, whereby a business can pick and choose exactly what platforms they wish to keep internal and which they want to share as part of their outsourcing initiative.

With cloud-based platforms, companies can move away from rigid, self-contained business models and open their work flow solutions to a broader group of workers. Prior to cloud computing, platforms and data were only available internally and systems were confined to working within their own paradigm, which by nature is intrinsically limiting not only in terms of the information itself, but for the company's employees as well. If a company wished to outsource or share anything outside their own internal workforce, all documents, data, programs, etc. must be delivered to an outside party, either individually, or at the most, limited to only a select number of individuals.

Now, all of the information and necessary resources can easily be standardized, which enables companies to move further away from these internal, captive-based labor models and in turn begin to utilize the cloud to its full potential through crowdsourcing with global labor rates, more specialized skill sets and task management processes.

Figure 2.

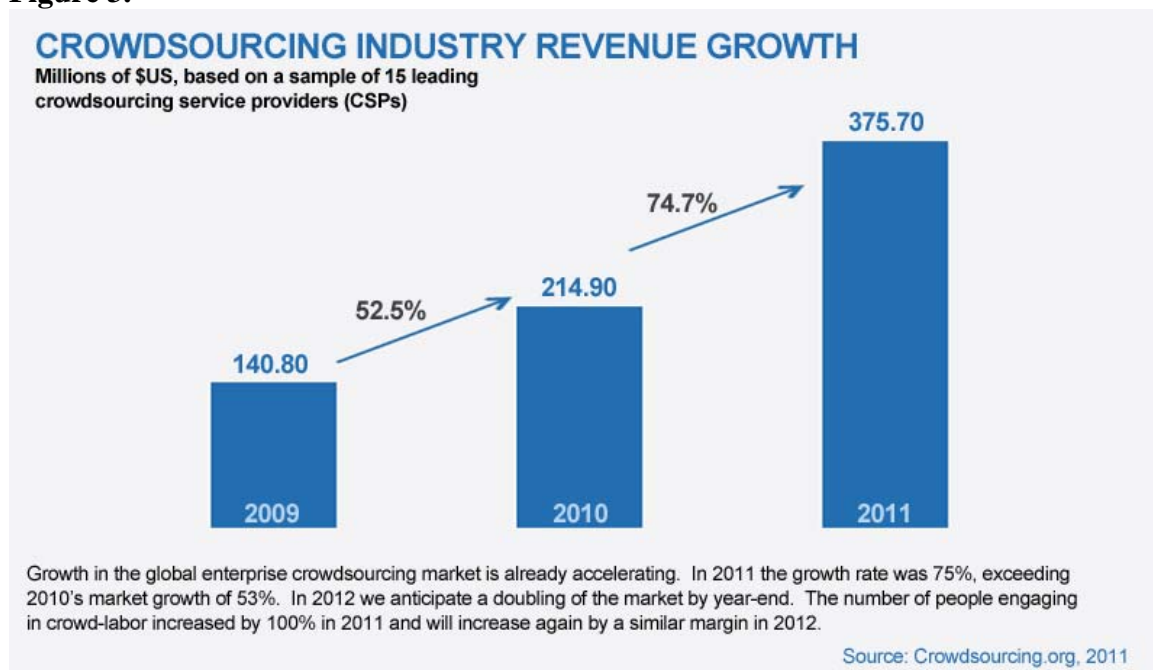**Cloud Computing Industry Statistics**
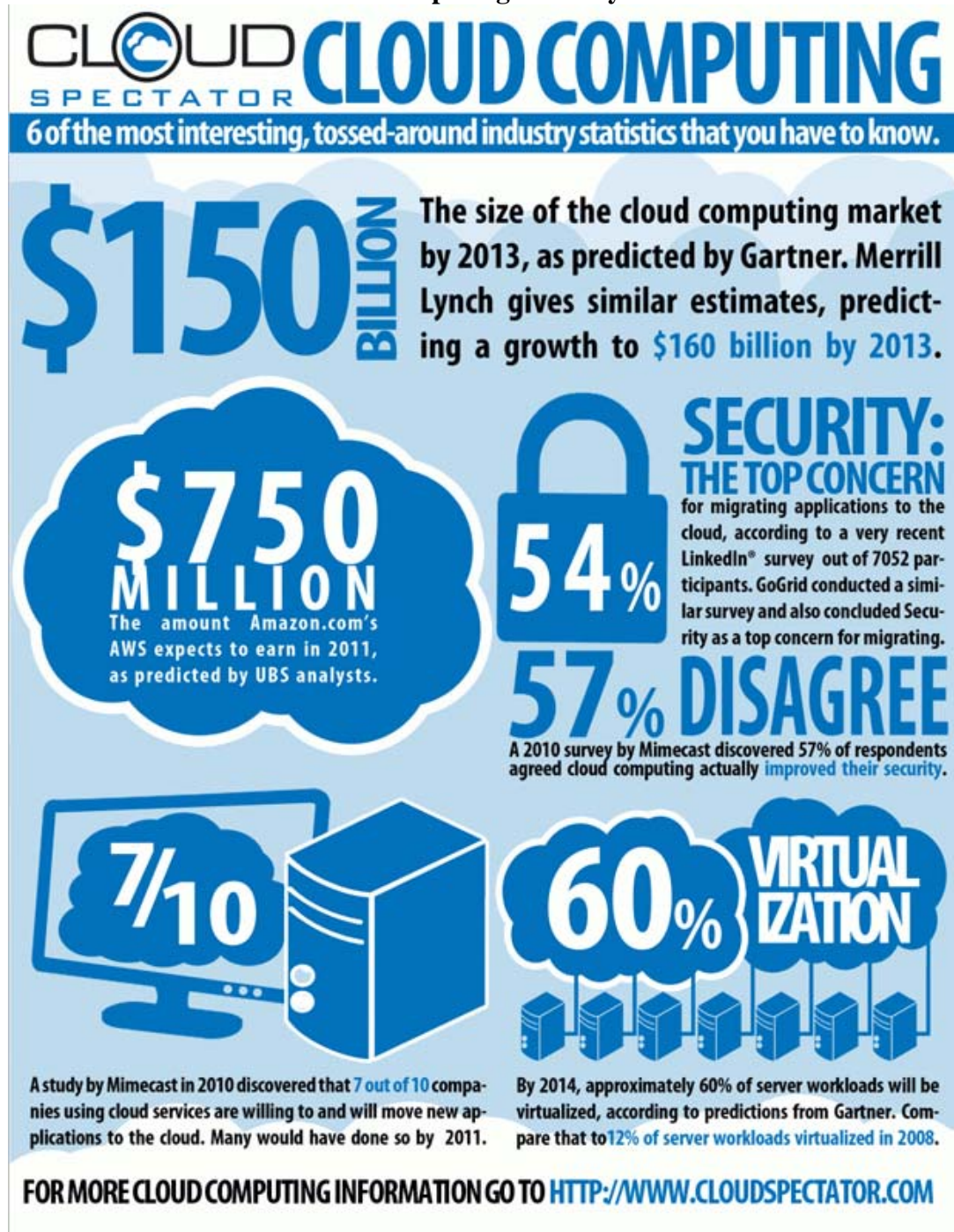
Seven out of ten companies have now switched to cloud computing, with over 80% of new enterprise apps being deployed on cloud platforms. Source: cloudproviderusa.com

Crowdsourcing

The concept for crowdsourcing is to allow the internet public to contribute information to solve the desired task. Due to the influx in online users (estimated to be about 3 billion by 2016), the internet has opened the doors to an entirely new workforce with varying expertise and skill sets.

While crowdsourcing may have already started to gain popularity in small doses amongst industry professionals with the proliferation of cloud-computing platforms, crowdsourcing has recently seen a massive surge in primary use by many businesses. In fact, according to Crowdsourcing.org, the trend seems to be almost doubling each year.

Figure 3.

CROWDSOURCING INDUSTRY REVENUE GROWTH
Millions of $US, based on a sample of 15 leading crowdsourcing service providers (CSPs)

Growth in the global enterprise crowdsourcing market is already accelerating. In 2011 the growth rate was 75%, exceeding 2010's market growth of 53%. In 2012 we anticipate a doubling of the market by year-end. The number of people engaging in crowd-labor increased by 100% in 2011 and will increase again by a similar margin in 2012.

Source: Crowdsourcing.org, 2011

Of course, with the availability of high tech tablet and smartphone computing, it's not hard to see why the crowdsourcing trend is appealing to workers. It allows the opportunity for work to be performed virtually anywhere at any given time, while still within a fully secure platform.

The below graph shows the growth in which workers are effectively embracing the crowdsourcing method, and the growth relative to various industries over the past few years.

Figure 4.
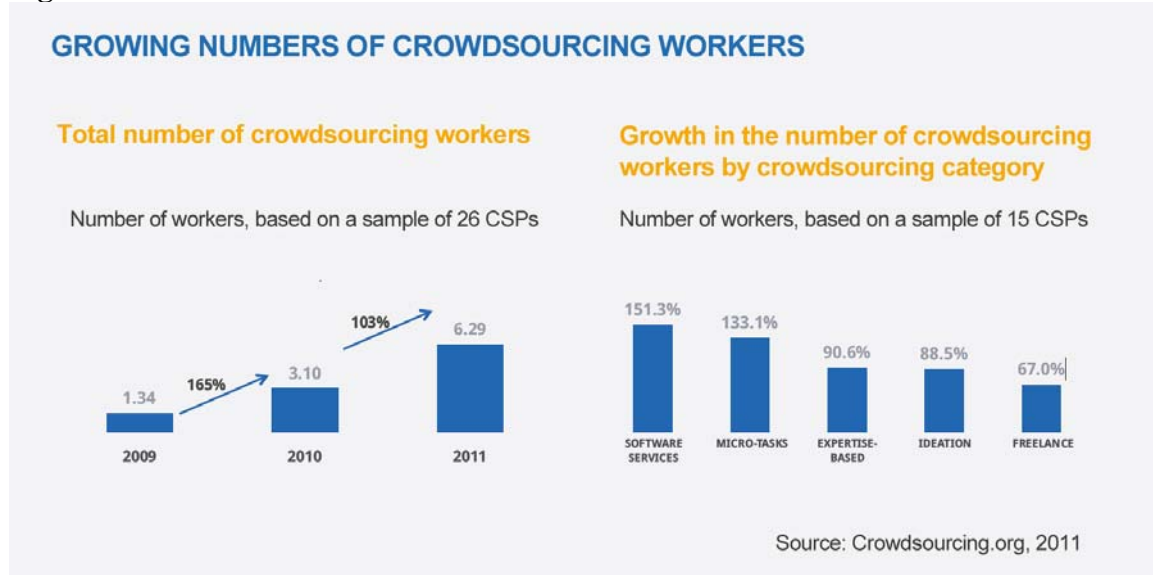
Crowdsourcing can be particularly useful within the translation industry, as it provides the opportunity to easily reach users from varying backgrounds and language groups and therefore harness the crowd to deliver faster and less expensive translations. According to Rob Vandenberg, "[Crowdsourcing] introduces a whole new group of translators and can dramatically lower translation costs. As a result, the ability to translate content will no longer be limited to companies with large budgets but will be open to small businesses and even micro-writers such as hobby bloggers."[6]

But the benefit of crowdsourcing doesn't stop at the ease of accessibility of different language service providers. With the need for agility within the translation industry itself, companies can utilize crowdsourcing for such contextual services as HR, accounts receivable, accounts payable, and even project coordination.

Crowd-based Translation Tools

Crowd-based translation tools are designed to make the process of crowdsourcing translations easier by leveraging more control for both the company and for the individual translators. Crowd-based translation tools break up a given project into smaller, more manageable sections and these sections are then accessed by multiple people at once from all over the world. Content owners can easily track the progress made and manage the quality on an ongoing basis, usually within real-time or with a short delay. With the assistance of other features like translation memory, translators

---

[6] Crowd Leader Rob Vandenberg: How Crowdsourcing Advances Industries (http://dailycrowdsource.com/crowdsourcing/crowd-leaders/475-crowd-leader-rob-vandenberg-how-crowdsourcing-advances-industries)

also can be prompted with translations in order to ensure quality and consistency throughout, even if the translator is not too familiar with the subject matter.

Some tools provide the ability to embed translation platforms directly into a client's website, which allows for a completely seamless integration. By doing this, it allows the translator to view the context and structure as a whole, rather than simply seeing the segments in a word document or some other separate file format, which often makes it difficult to understand the context, and especially, how the translation fits aesthetically on the site.

Some companies that embrace the crowdsourcing platform in its entirety currently include Smartling and Gengo. In the last 12 months, Smartling has tripled revenues, added 120 enterprise customers and doubled its employee count. Smartling has plans to hire another 100 employees in 2014, and even recently secured $24 million from investors for their cloud-based translation software.

Crowdsourcing Motivation

While this big push towards crowdsourcing platforms have been the impetus for companies such as Smartling and Gengo, the concept of crowdsourcing can also be utilized for intrinsic methods of translation services. Individuals have proven that they are willing to volunteer their services if they are actively engaged in the subject matter, or are rewarded for their services with non-monetary incentives.

Companies such as Twitter and Facebook have successfully utilized crowdsourcing while translating the content of their sites. They knew their non-English speaking fans wanted to use a version in their native language, so both companies left it up to their users to essentially create the versions themselves. The reward of having one of their favorite sites in their native-language was enough for those language users to volunteer their time to translate the content.

Particularly in Twitter's case, the company utilized a community of hundreds of thousands of users to initially translate the site into French, Italian, German and Spanish in 2009, and then later added Korean, Japanese, Russian, Turkish and Portuguese.

Laura Gomez, Manager of Localization at Twitter stated, "We believe that investing in the community of translators is an integral part of our international growth and success...they know our product best and capture the Twitter spirit through their translations."[7] Matt Sanford, Twitter International went on to state, "In the case of Twitter's Translation Center we rely heavily on our user's prior knowledge of Twitter itself."[8]

---

[7] Laura Gomez (http://www.quora.com/Laura-i-G%C3%B3mez/answers/Internationalization)
[8] Matt Sanford (http://www.quora.com/What-key-benefits-did-Twitter-and-Facebook-uncover-in-crowdsourcing-their-translations)

One of the core tenants that make crowdsourcing translations successful is the audience's engagement in the subject matter itself. By opening up the process to only a handful of potential translators, you're limiting yourself to a group that may or may not be fully aware of your company's core message or brand. However, by engaging the full population, it is much more likely you will be able to find translators who already have an understanding of your company and therefore will provide a much more accurate translation.

This is exemplified in Matt Sanford's statement, "Twitter has also experimented with Language Service Providers and they do a great job at some types of tasks, while at others they don't. One quick example to highlight the difference in community translation quality versus professional translators: Twitter uses the word "unfollow" to label [a button]. If you ask an LSP to translate this they will likely return a phrase such as "Nicht mehr folgen" in German. If you ask a translator directly they will probably point out that "unfollow" isn't a word. Twitter's language is very informal in English and community translation has helped us keep that playful tone in every language (even the ungrammatical "Entfolgen" in German)."

Wikipedia is another website which not only utilizes crowdsourcing for its main content, but also to translate existing content into other languages as well. Wikipedia accomplishes this by compiling its pages for translation into two separate categories: the "translate" section which lists articles identified as having content in another language, and the "cleanup" section, which lists articles that have been translated but require post-editing or proofreading.

Figure 5.

Translated Articles on Wikipedia

Subject	Number of French articles	Percentage of total (French)	Number of Spanish articles	Percentage of total (Spanish)
Biography	20	58.8%	20	33.3%
Arts (TV, film, music, fashion, museums)	3	8.8%	8	13.3%
Geography	2	5.8%	12	20%
Transportation	2	5.8%	4	6.7%
Business/Finance (includes company profiles)	2	5.8%	2	3.3%
Politics	1	2.9%	4	6.7%
Technology (IT)	1	2.9%	1	1.6%
Sports	1	2.9%	1	1.6%
Education	1	2.9%	1	1.6%
Science	1	2.9%	0	0%
Architecture	0	0%	2	3.3%
Unknown	0	0%	3	5%
Other	0	0%	2	3.3%

| Subject	Number of French articles	Percentage of total (French)	Number of Spanish articles	Percentage of total (Spanish)
Total	34	99.5%	60	99.7%

Figure 5: Subjects of translated articles listed on Wikipedia:Pages needing translation into English (May 15, 2011-May 15, 2013)

As the data from the table suggests, the majority of the translations from French and Spanish listed on Wikipedia's 'Pages needing Translation' are biographies. While some of these biographies are of historical figures, most are of living people. The arts were also another popular topic for translation. In the translations from Spanish, articles about cities or towns in Colombia, Ecuador, Spain, Venezuela and Mexico (grouped here under the label "geography") were also frequent. So it seems that the interests of those who have been translating articles from French and Spanish as part of this initiative have focused on arts, culture and politics rather than specialized topics such as science, technology, law and medicine. This goes to show just how important personal engagement in the subject matter is for the user to willingly volunteer their time and translation knowledge.

In fact, in a survey of 76 crowdsource translators, intrinsic motivations are the most often cited reasons for volunteering their services.

Figure 6.

Crowdsource Motivation
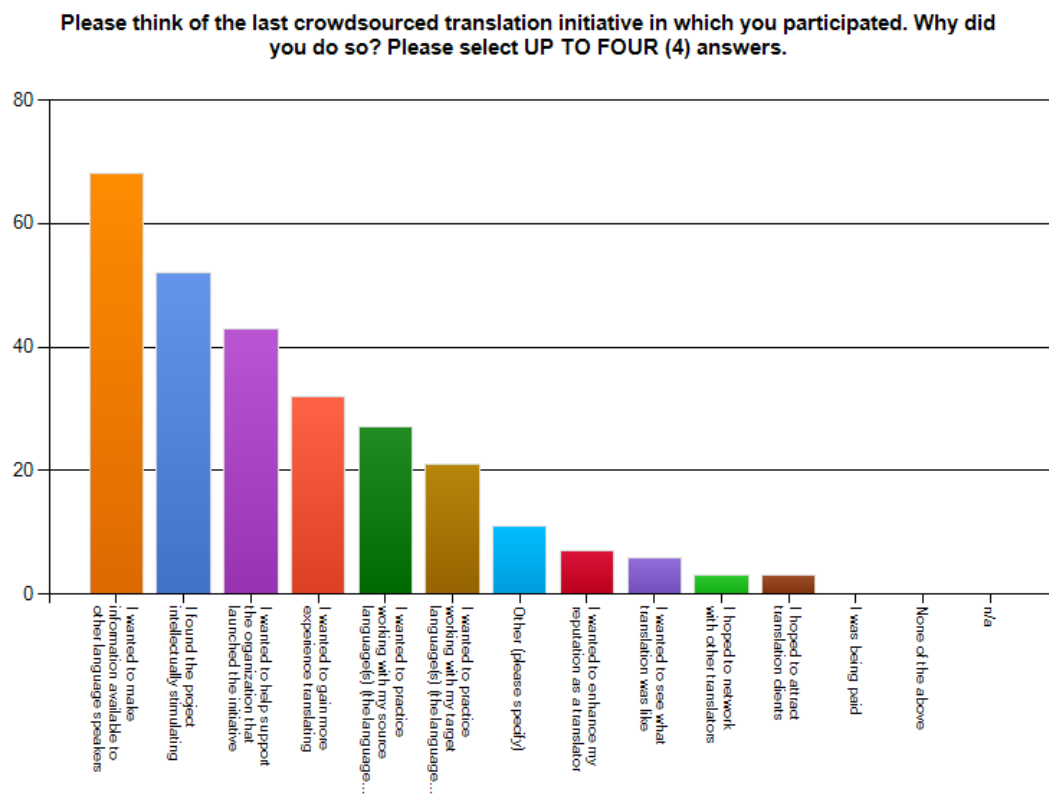
As the chart shows, intrinsic motivations (making information available to others, finding intellectual stimulation in the project, and supporting the organization that launched the initiative) were the motivations most often cited by the translators as the reasons for engagement. However, a significant number also had extrinsic reasons for participating: Wanting to gain more experience translating or practicing their language skills.

Crowdsourcing Mobile App Translation

A relatively new company which focuses on providing both intrinsic and extrinsic value to volunteer crowdsource translators is Ackuna. Ackuna is a crowdsource translation platform for localizing different mobile and tablet apps. As mobile apps continue to gain in popularity on new and emerging devices, there is an increased need for app translation. Ackuna took on this issue by developing a free-to-use system which allows developers to upload their native files and have a community of translators translate the app's text into multiple other languages.

For the translators who volunteer their time, Ackuna provides intrinsic motivation through the apps themselves. If an app looks particularly interesting on the Ackuna website, and the developer has uploaded the appropriate screen shots and descriptions to make figuring out the context as easy as possible, many translators will gravitate toward wanting to volunteer their services so that they can both be a part of bringing that app to life for their native language, and so that they can also even use the app once it officially launches in that language.

As an extrinsic motivation, Ackuna is also a social networking platform for translators and language enthusiasts to simultaneously practice their craft on real-world examples, as well as promote themselves within the site through reputation points and earned badges given out by Ackuna for accurate translations and for reaching certain milestones within the system. The ranking system is in place to provide both quality assurance to the app developers, and also for the translators to "keep score" of their work to showcase to other potential paying clients.

So far, Ackuna's system is working out quite well. Since its inception, Ackuna's userbase has grown by almost 1200%, with over 6000 current users and over 42,000 successfully translated segments.

Figure 7.
Ackuna's User Growth

Site Statistics

In Summary

As the translation industry continues to grow, innovation is key for any company to compete on the global scale. Where outsourcing alone was once the go-to method for companies to leverage their costs to increase profit margins, it has certainly lost the effectiveness it once had. With the lowering of standard rates and the uncertain accuracy levels it brings with it, it seems that necessity really is the mother of invention.

As a result, there has been significant interest within the translation industry to explore more innovative solutions. Many successful companies such as Gengo, Smartling, and Ackuna have developed their own innovative crowdsourcing methodologies and programs which utilize the crowdsourcing concept, but significantly build upon it to meet their own needs for their target audiences.

While there are still some stumbling blocks to overcome, it seems like as a whole the industry is headed in the right direction with its attention on crowdsourcing. All companies have reported crowdsourcing to be a successful endeavor, with few negative attributes. The mitigating factors (meeting clients' budgets and time frames while ensuring accuracy of the translations) are key points to focus on, but without taking into account the needs of the translators, any innovative approach will be doomed to fail. By focusing on the needs of the company as well as both the intrinsic and extrinsic needs of the translators, companies can continue to be successful even within the vastly changing atmosphere of the translation industry.