The Translation Industry in 2022

A report from the TAUS Industry Summit

Amsterdam, March 22-24, 2017

IIIIITAUS

the language data network

Author: Isabella Massardo, Jaap van der Meer
Reviewer: Anne-Maj van der Meer

Design: Anne-Maj van der Meer

Published by TAUS BV, De Rijp, The Netherlands

For further information, please email info@taus.net

Table of Content

1. Executive Summary

TAUS expects that automation in the translation sector will accelerate in the next five years.

We foresee a rapid adoption of automation, both in translation production (machine translation, intelligent TM) and in translation management functionalities (such as project, quality and vendor management) through machine learning.

This automation process will bring along opportunities and challenges. The main opportunity is, of course, the increase in efficiency. The challenges will be the changes in jobs, sharing of data, getting intelligence from the data, working in the cloud. Opportunities also exist in the innovation made possible through translation automation.

Strenghts	Weaknesses
Increases in efficiency	Sharing of data
Opportunities	**Threats**
Getting intelligence from the data Working in the cloud Convergence of technologies	Changes in jobs

All this comes to us through convergence: Convergence of technologies (such as machine translation and speech technologies), but also convergence of markets and business models: freemium and paid, consumer and business, crowd and professionals.

Translation is on its way to become a utility. Buyers and providers of professional translation services need to assess their strategies in light of a changing landscape.

2. How to read this document

The chronology of consultations with TAUS members, analysis of results, reviews with the Advisory Board and brainstorming sessions in World Café style setting at the TAUS Industry Summit on March 22-24 forms the structure of this report. As the figure below illustrates, the 'process' started with a request to all TAUS members to submit ideas for innovation in the translation 'space'.

Sixty different ideas were collected and voted on for importance and relevance. After careful analysis and review the list was 'reduced' to six main topics or drivers of change that together formed the agenda of the TAUS Industry Summit on March 22-24, 2017 in Amsterdam.

At the TAUS Industry Summit a story started to develop that depicted a drastically changing landscape in the coming five years in the translation industry. The story is broken down in ten innovation themes or chapters, each with their own plot but all interconnected.

This report starts with a chapter (chapter 3) on the background – how this report came about: a description of the consultation and how we arrived at the six drivers of change.

In chapter 4 we give a helicopter overview of the evolution of the translation industry from a technology perspective, looking back and looking forward.

In chapter 5 we dive deeper into the six drivers of change: machine learning, machine translation, quality management, data and datafication (including speech), interoperability and academy (human resources and training). Thus far, we are quite analytical.

In chapter 6 we start to extrapolate the facts and trends and project the story of the translation industry in 2022 in ten mini-chapters. Speculative? Perhaps. Plausible? Certainly. Necessary? Definitely. How else can we be prepared for such rapid change in the translation industry.

Industry Consultation

⬇

Sixty ideas

⬇

Six drivers of change

⬇

TAUS Industry Summit
Amsterdam, March 22-24

⬇

The Story of the Translation Industry in 22
Ten Innovation Themes

3. How this report came about

3.1 The Consultation

From November 2016 through January 2017, TAUS conducted consultation surveys among its members.

First, we asked our members to propose ideas for innovation. Sixty different ideas were submitted that were grouped into thirteen categories. We then asked the TAUS community to give a priority rating to these thirteen categories. We received 202 responses , all from decision makers, people on director level and business owners. The chart below indicates the distribution of responses.

Distribution of Responses

Academic   Buyers   Providers   Gov-NGO   Freelancers
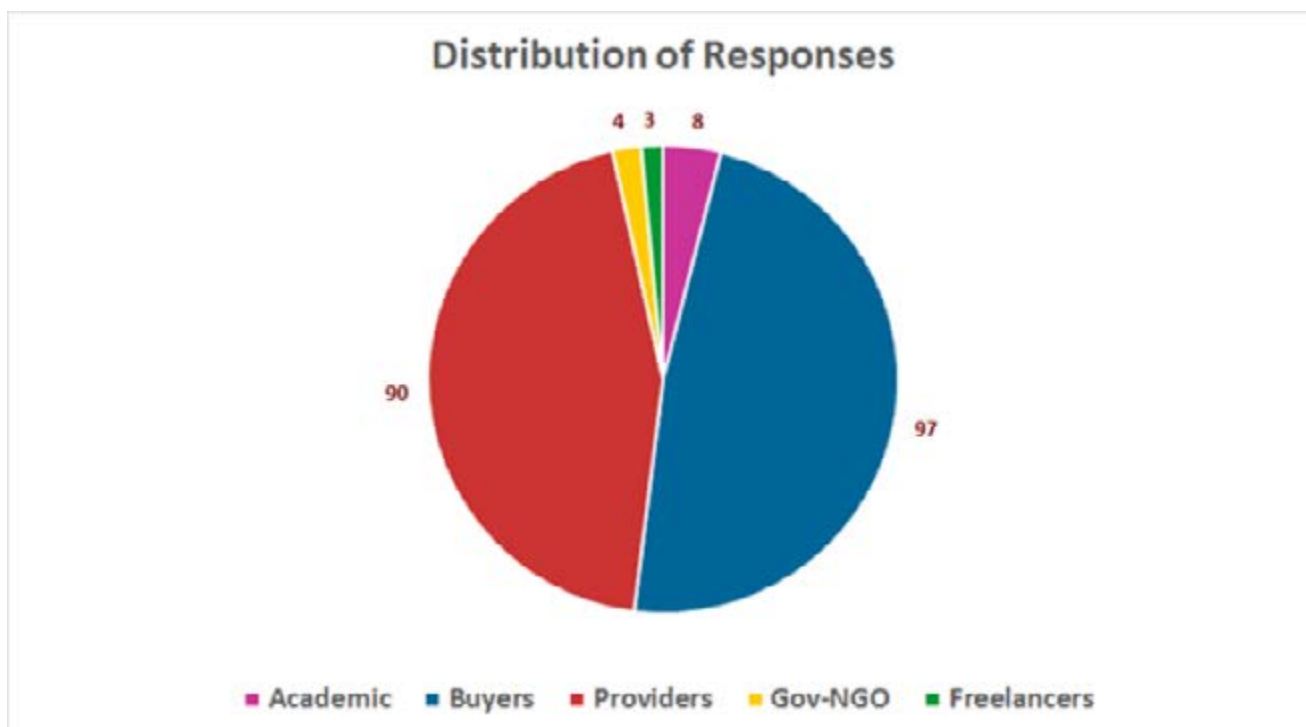
The following chart shows the priorities of the thirteen categories in descending order. The numbers in brackets indicate the numbers of ideas that were submitted in that category.

Machine learning was the clear winner, followed by the Quality Dashboard (DQF) and machine translation (MT) in second and third positions, respectively.

3.2 Selection of topics

Following the consultation and the priority ratings we have selected six topics for the agenda of the Industry Summit, that was held in Amsterdam on March 22-24, 2017. These six topics cover ten of the thirteen categories prioritized by the voters.

Three categories have been excluded due to a low priority rating or lack of relevance: 'Regulations' and 'Open Source Tools' are considered to be out of scope, at least for the TAUS community. On the contrary, 'Driving Adoption' is considered implicit in every theme.

Priorities – All Responses

Machine Learning (9). Quality Dashboard & DQF (11). Machine Translation (4). Intelligent TM (4). Interoperability (4). Data Cloud (7). Translation and IT architecture (4). Open source tools (3). Academy (8). Driving adoption (1). Speech Data (2). Regulations (1). Crowdsourcing (2).
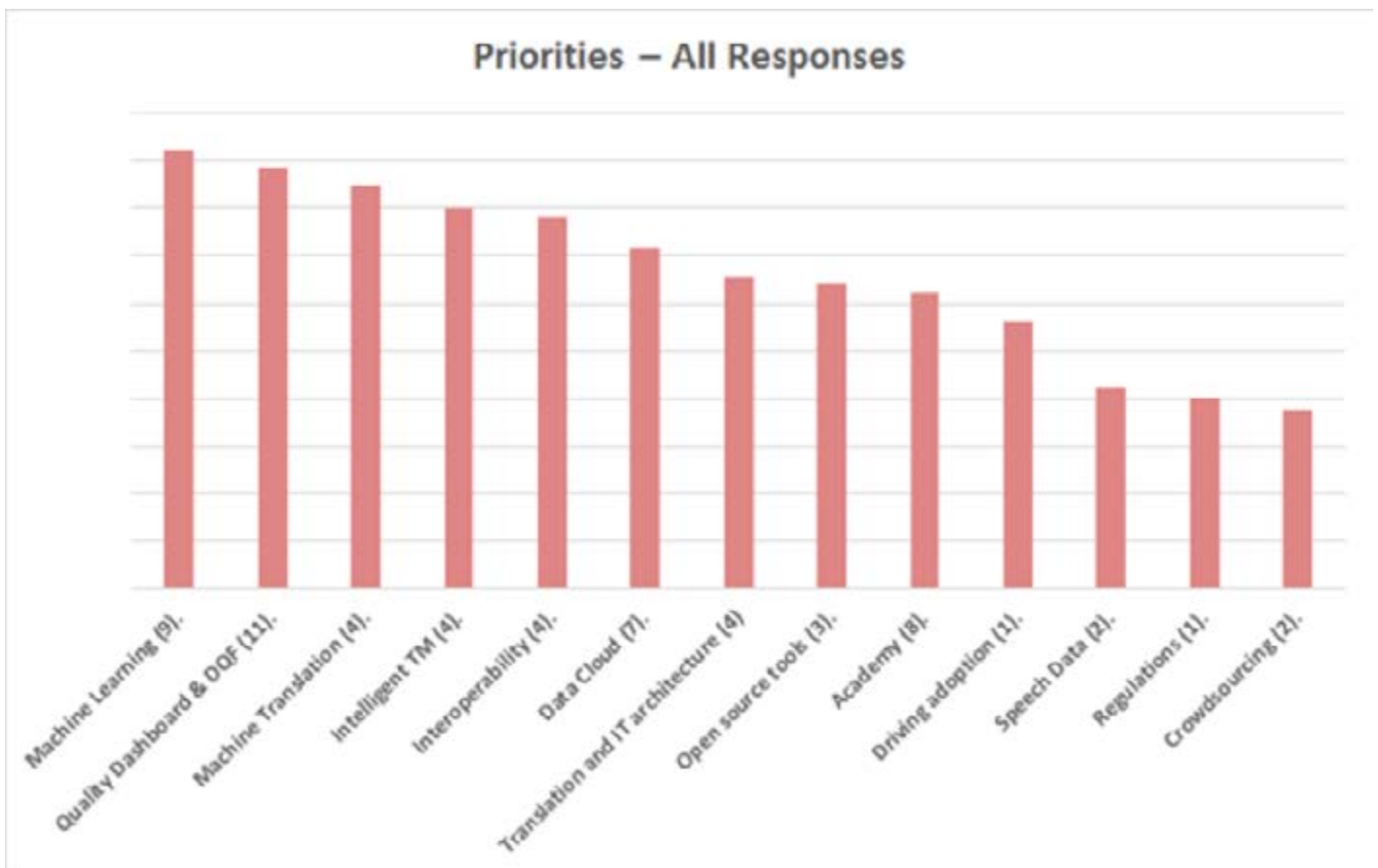
Figure: Thirteen categories with ideas. In brackets is the number of ideas that were submitted in that category.

The six topics selected are essential to the business planning strategies of every translation operator, whether on the buy or the supply side. Together they are in fact the drivers of change for the translation industry.

In the next five years the world of translation will be changing because of these drivers. Managers will be challenged to lead, follow or keep up with the innovations that have already been set in motion.

The six drivers of change are:

1. Machine Learning
2. Machine Translation
3. Quality Management *(including DQF and Quality Dashboard)*
4. Data *(including Speech Data)*
5. Interoperability
6. Academy *(human resources and training)*

4. A Helicopter View of the Evolution in the Translation Industry

Globalization exposes us all to content available in an increasing number of languages. We expect translation technology to help us lower the language barriers.

In the past few years, translation technology - covering a broad range of categories - has become a very interesting and dynamic sector for innovators from both inside and outside the industry: new tools, platforms and solutions promising to improve our multilingual capacity have been developed non-stop.

Originally, translation technology was mainly focused on productivity gains. When, in 2007, Google Translate arrived on the scene, Google disrupted the translation sector, even though it had no intention to do so. In addition to surviving over a decade of criticism for alleged bad quality, Google Translate has managed to develop an audience of hundreds of millions of users and to inspire many startups.

Before Google Translate, translation had always been a professional service, affordable almost exclusively to business clients. With the advent of Google Translate, translation has become directly accessible to end-users. Google Translate has raised the importance and significance of translation, and has led to unprecedented innovations in and around the translation industry.

The advances of the last decade have been the prelude to the convergence era, the moment when separate technologies will start sharing resources and interact with each other synergistically, with one common interface on different devices. When convergence comes to full maturity, translation will be universally available on every screen, in every app and on every signboard.

We could say that the path to convergence began in January 2017, when Google announced that Google Translate was available in 103 languages and serving over 500 million users monthly which accounts to 140 billion words per day. In other words, today Google Translate alone translates more words in a day than all human translators in the world translate in a year.

4.1 The Current Landscape
 The history of translation technology is short and fast. It starts with the halt to machine translation (MT) funding in late 1966 following the conclusions contained in the ALPAC report.

The ALPAC report defined MT as being "too expensive, too time consuming, too inaccurate" and sent the clear message both to the public and the rest of the scientific community that the technology in question was hopeless. Since then, MT has always met the opposition of translators, fearing it could take work away from them.

The conclusions of the ALPAC report were mostly due to the large number of available translators at the time, the relatively small amount of texts to be translated, and the inexpensiveness of translators compared to scientists, which made MT uneconomical. The ALPAC report recommended, however, the development of tools to help translators become more productive.

At the end of the 1980s, with the global spread of affordable computer equipment and the first boost to content production, the ALPAC recommendations turned into viable, practical and affordable technology. In fact, that boost spurred the rise of global software and global content that had to be made available in dozens of languages, giving therefore birth to the localization industry. Since then, translation technology has been developing in parallel with the translation industry.

The advances in technology during the 1990s added new capabilities to existing tools, while new tools were released with different packages, functionalities, operating systems and prices, widening the range of choices for buyers.

The 2000s were a decade of rapid growth, and the last few years in particular have been a period of global development. The most prominent position in this period is that of MT. Despite the conclusions of the ALPAC report, research in Natural Language Processing (NLP) continued and MT was part of it. To be fair, the progress of MT has been such that most of today's users cannot even remember the cumbersome efforts that preceded the appearance of Google Translate. And yet, even today, despite its advances and successes, many translation industry insiders still take MT errors as inescapable 'failures'.

True innovations most often come from outsiders and most of today's technologies are heirs of earlier publicly funded projects. As Mariana Mazzucato rightly points out in her book The
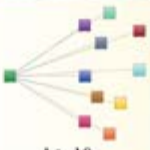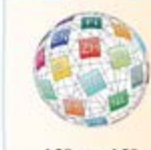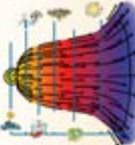
	1980 TRANSLATION	1990 LOCALIZATION	2000 GLOBALIZATION	2010 INTEGRATION	2020 CONVERGENCE	2030 SINGULARITY
INDUSTRY FOCUS	Documents	Software	Simship	Integration in enterprise systems	Embedded in every app, on every screen	Ubiquitous
TYPE OF CONTENT	Paper	Digital	Static Web	Dynamic Web	Personalized	Unlimited
LANGUAGES	1 > 10	1 > 25	1 > 40	6 <> 60	60 <> 60	150 <> 150
TECHNOLOGY	None	TM and terminology software	Workflow (GMS)	MT and Advanced Leveraging	Real-time customized MT	FAUT
DATA	Glossaries	Project TMs	Centralized TMs	Limited shared data	Web crawled	Plug and Play
COMMUNICATIONS	B2B G2C B2C	C2B	C2C	Social	Internet of Things	M2M

Image: Evolution of the Translation Industry Landscape.

Entrepreneurial State, it was the U.S. armed forces that pioneered the internet, the GPS positioning systems and the voice-activated virtual assistants; and it was thanks to the research of academic scientists in publicly funded universities and labs that the touchscreen and the HTML language were developed, just to mention a few examples.

The translation industry and the translation technology industry are not so different, Systran being the first example that comes to mind, because the company survived depressions in the MT market thanks to the financial support from the EU. The translation community has been thriving for decades now on technology and it basks in the belief of operating in a high-tech business, although most of the business models followed are still quite old-fashioned.

As part of the translation technology developments, translation management systems (TMSs) have been designed to meet the basic demands of a typical translation business: cutting operating expenses and, at the same time, replicating the traditional modus operandi. This has led to very interesting abridged applications of project management and enterprise resource planning (ERP) platforms, far from the typical systems of that kind that can be easily found in businesses belonging to any other economic sector. Think of Basecamp, Compuware, Infor, JD Edwards (now Oracle), Jeeves, PeopleSoft, Plex, SAP, Tally: These are all software manufacturers releasing products that could have been adapted for the translation business. The typical translation-oriented management software systems have been developed because the above-mentioned systems are much too complex and expensive for a translation "microbusiness".

4.2 A View on the Future of Translation

Can we ever know what's going to happen next? History is rich of anecdotal wrong predictions, to the point that even Wikiquote has a whole page dedicated to incorrect predictions.

Very few people in the academic and industry field seem to pin down on futurologist Ray Kurzweil's prediction that MT will reach human translation quality levels in 2029. Many people in the translation community were convinced that MT was a hoax only ten years ago, just to acknowledge it is getting better and better today.

In fact, today only a tiny amount of the content that should be translated is indeed translated, while the volume of content requiring translation keeps growing steadily, bringing the demand for translations of publishable quality beyond human scale. Many people in the human translation (HT) space say translators are not good enough to tackle today's challenges, and more and more errors are detected than in the past. MT has advanced enormously since its first development.

The truth is, when we see that some predictions that previously appeared unrealistic come true, we often forget that they were made after a thorough analysis of the moment in which they were formulated and of the recent past. Today, few trends will most probably keep growing in the next few years and will likely become the bridge to the future: the cloud, the datafication of everything, and machine learning.

Machine learning (ML) will be more and more applied to translation, especially when it comes to MT and speech-to-speech (S2S) translation: it will open a new scenario for translation business intelligence with a considerable impact on quality assessment.

Today's most popular "hype" within the translation industry is neural machine translation (NMT). NMT is a new MT technology that has been under research and development since at least 2012, and which matured to the point of open availability during 2016. Organizations that have worked with the technology are convinced that it offers significant advantages over phrased-based SMT in terms of output fluency and accuracy.

Unlike statistical machine translation (SMT) which uses look-up tables to "learn" comparable phrases, NMT learns to translate sentences by using technology called neural networks - several layers of linked "neurons" that operate in symbiosis to roughly imitate the cognitive processes used in the human brain to recognize and learn patterns of information. Like SMT, NMT is trained using parallel data, but due to the "deep learning" carried out by the neural network, the engines are capable of far richer sentence modeling than SMT engines.

The technology is currently being evaluated by MT production specialists. Several MT suppliers have announced the availability of NMT solutions for their clients and LSPs are holding webinars and other events to publicize their findings. Cost of entry is relatively low, and the entire MT industry has the advantage of starting from a level-playing field. The focus will now shift to developing customization wrappers around the core technology, adapting these to specific job requirements and evaluating NMT systems for the full range of translation use cases.

Despite the general enthusiasm, NMT is not the silver bullet that will deliver the ultimate translation solution. However, many of those now riding this new technology wave are talking about a major 'paradigm shift' in MT performance.

Research in neural networks, machine learning, deep learning and the like won't stop mainly because of the irresistible allure they exert. Research will most probably be still both public and privately funded. From the end of WWII to this day, technological research has prospered thanks to sizable public investments that have led to extraordinary achievements with equally extraordinary and extensive bearings.

Defense-related research initiatives may lead to a convergence of technologies such as (neural) MT and speech technologies. And as always, fiction narrative (like the movie Arrival) will set a whole new series of expectations. Evolution in translation technology, however, is going to follow different paths, mostly running without intersecting.

So far, translation technology innovations have produced no real process innovation. In this respect, the improvements in MT quality and the numerous widely available platforms have led to their integration in every professional translation tool, paving the way to a practice as old and as deprecated as MT itself: post-editing.

Estimates of the percentage of post-editing on the total translation production vary wildly, simply because, in the so-called interactive mode, post-editing is now the standard modus operandi for most professional translators, and it is going to stay that way, at least in the short-term.

Traditional, static post-editing may still take some time to become widespread as buyers learn to leverage their data using enriched term bases for sub-segment matching and training SMT engines in order to benefit from fuzzy match patching. As long as NMT will not fully replace SMT, post-editing will remain essential to track and understand issues in the performance of SMT engines as well as to make the necessary adjustments to improve them.

Current translation memory management systems use edit distance to score matching, which hardly compares to human judgment of semantic similarity of segments. There are developments underway towards more 'intelligent' metrics in this respect, derived from NLP experiences.

The next few years may also be open to more innovations in translation memory management, in an 'intelligent' perspective, to embrace multimedia, especially phonetic language data, possibly through enriched metadata capacity. When asked about possible forthcoming innovations, translation industry players usually answer by trusting machine learning, even for quality, which seems to remain the main concern.

Machine learning is a possible basis for predicting the vanishing of TMSs, at least as we know them. TMSs are conceptually outdated, beyond any planned obsolescence. They have failed in helping disintermediation, because their developers have been looking at meeting middlemen's requirements rather than the end customers' needs. However, the drive to disintermediation is going to become sharper and sharper: we expect automation to accelerate, even in translation management functions such as project, quality and vendor management.

Machine learning will play an important role here, too. A machine learning algorithm may be developed using content and human profiling, productivity, terminology identification and harvesting, quality prediction and estimation, financial and project data, to automatically allocate the right resources for optimized leveraging and translator evaluation, as well as to assign tasks to the best resources available. The customization of MT engines too will be done automatically by machine learning algorithms to identify the best matching data.

In this scenario, middleware is going to play a significant role through APIs to manage content flow from one system to another. Also, we expect to see new language-neutral message formats to be developed.

5. Six Drivers of Change

5.1 Introduction

In *The Signals Are Talking*, American futurologist Amy Webb describes a process for identifying new trends while they are still developing and for tracking them as they hit the mainstream. She calls this process "CIPHER", which stands for "Contradictions, Inflections, Practices, Hacks, Extremes, Rarities." Webb looks to the past to inform her conclusions and encourages people to remain skeptical and test their inferences. In the end, being good at making predictions is all in telling trends from trendy stuff.

Today, of the six drivers of change identified in our survey, only one trend will most probably keep growing in the next few years: datafication. Data is the drive to all the technologies that are reshaping our lives and businesses. Data will provide the fuel for the next stage of development in artificial intelligence through machine learning.

The cloud is ancillary to datafication. Given the huge amount of resources needed, it is the necessary infrastructure for data distribution and sharing, neural networks, and deep learning.

No prediction is possible without a generous error tolerance: if estimates of general rates of technological progress are always imprecise, it is fair to say that progress has been happening much more rapidly than ever, and that the curve slope can only become steeper. To put this in perspective: if car makers innovated the way the IT industry has done, today's cars would run 100,000 miles per gallon (over 42.000 km per liter) and cost 50 cents.

As stated in the previous pages, through a survey conducted within the TAUS community, we've identified six drivers of change in the translation industry that must be included in the business planning strategies of any stakeholder, whether on the buy or the supply side of the industry. They are:

1. Machine learning
2. Machine translation
3. Quality management
4. Data
5. Interoperability
6. Academy (HR training)

5.2 Machine Learning

Machine learning (ML) was introduced in the 1950s as a subset of artificial intelligence (AI), to have programs feed on data, recognize patterns in it, and draw inferences from them.

The hyper-abundance of data available today - thanks to the cheaper and more powerful computer hardware - has been infusing new life to old algorithms in an escalation that has led to a wave of workable applications. However, most of us cannot name three applications for machine learning, other than self-driving cars or the voice-activated assistant built in our smartphones. Deep learning is an area of machine learning that refers to many layers of neural networks work-

ing together and benefiting from the abundance of computing resources in the cloud. Most ML algorithms have nothing to do with neural networks, and AI is simply not enough to make computers act like they do in the movies. Not yet, in any case.

2016 was the year when ML went mainstream, with a lot of applications that were almost unimaginable a few years earlier - image recognition and self-driving cars are just two examples.

McKinsey places ML already on the descending slope of the Gartner hype curve, on the cusp between the peak of inflated expectations and the trough of disillusionment. This means that ML will still have a great impact on every organization's strategic planning as it keeps showing promise in delivering a high degree of competitive advantage over the next five to ten years.

Computational power and unprecedented advances in deep neural networks will make data-driven technologies astonishingly disruptive. This might be also the case of MT. In fact, SMT is a machine learning application, and it has been for decades on the forefront of language technologies.

Machine intelligence will play an even increasingly role in everyday life through automation: its effects will vary dramatically across sectors and activities. As a rule, the growth of machine intelligence represents a threat to many human jobs as people will be replaced by intelligent systems. While it will eliminate very few occupations entirely in the next decade, machine intelligence will affect portions of almost any job to a greater or lesser degree, depending on the type of tasks.

In 2013, a study by Oxford economists Carl Benedikt Frey and Michael A. Osborne estimated the probability of 702 specific job categories being replaced by computerization: proofreaders and technical writers were given an 84% and 89% computerization probability, respectively. In 2014, the economist W. Brian Arthur estimated that machines could replace 100 million U.S. jobs by 2025. In turn, McKinsey estimates that automation is going to affect 45% of paid activities, with 30% of activities automated in 60% of all occupations.
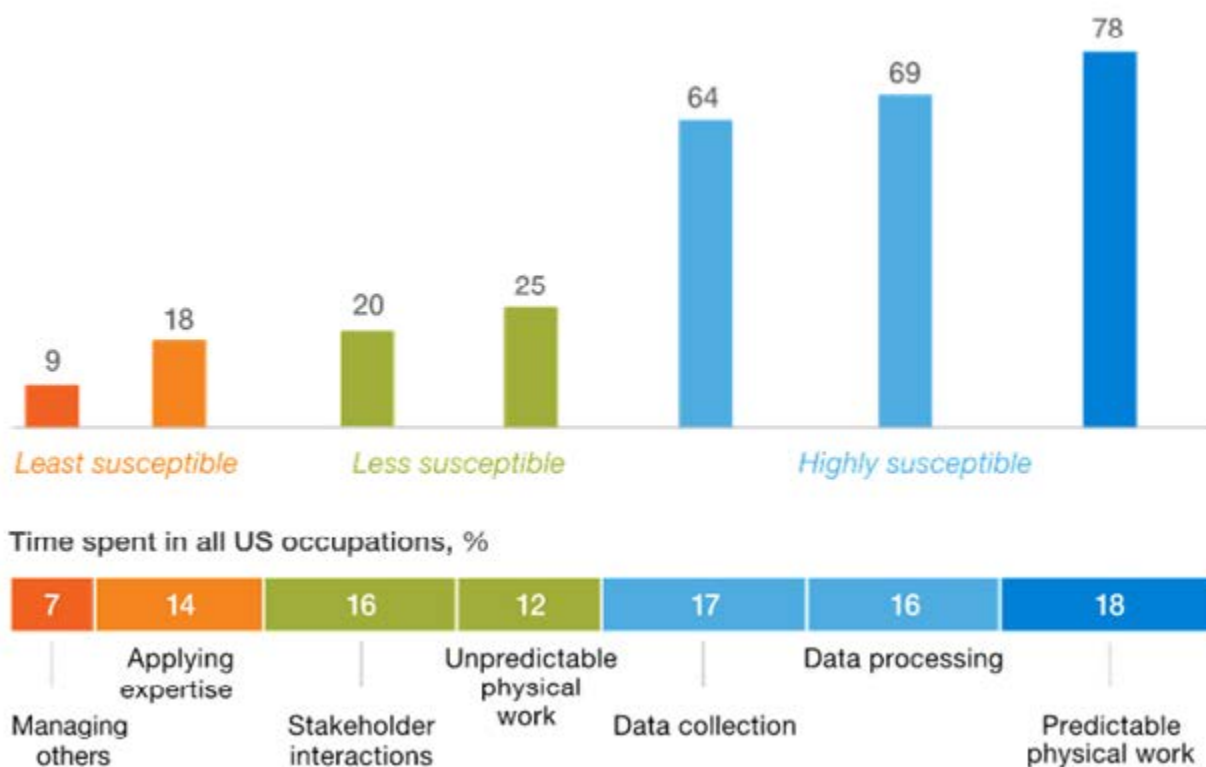
Source: McKinsey & Company

14

Managing positions will remain intact, but most assisting jobs may disappear. The majority of creative jobs is relatively safe while sales jobs could be at risk. The forecast is dubious for technology jobs, again with the more senior jobs being relatively secure, while computer programmers and support workers may likely be replaced.

The assumption that jobs requiring manual dexterity, creativity, and social skills are the hardest to computerize is already obsolete: new developments in deep learning are making machines more powerful than anticipated, especially in areas relating to creativity and social interaction. However, although technical feasibility is a necessary precondition for automation, it is not a complete predictor that an activity will be automated; the cost of labor and IT infrastructure are two major factors for automation, which might not take place until workforce remains abundant and relatively inexpensive.

For example, at present, almost one fifth of the time spent in US workplaces involves performing specific actions in well-known settings, where changes are relatively easy to anticipate, i.e. predictable physical work. The hardest activities to automate with the technologies currently available are those that involve managing and developing human resources (9% of automation potential) or that apply expertise to decision making, planning, or creative work processes (18%).

When, at the end of 19th century and the beginning of 20th century, electrical engines started to replace steam engines in textile factories, there was only a slight increase in productivity. In our century, one of the biggest—and most expected—technological breakthroughs will come if (or when) machines will develop a human-like understanding of natural language. Computers capable of recognizing concepts in everyday communication could increase the technical potential for automation from 53% of all labor time to 60% in retailing, from 43% to 66% in finance, insurance and healthcare.

In 2013, the Food and Drug Administration approved a machine to administer sedation to patients without the need for an anesthesiologist. An emerging field in radiology is computer-aided diagnosis: a recent study published by the Royal Society showed that computers performed more consistently (almost by a factor of ten) than radiologists in identifying radiolucency (i.e., the appearance of dark images).

W. Brian Arthur called "Second Economy" the portion of the economy where computers do business transactions only with other computers. This Second Economy is upon us, brimming with optimistic entrepreneurs, already spawning a new generation of billionaires, and probably driving much of the economic growth in the coming decades. Arthur also speculates that in 2025, this Second Economy may be as large as the original first economy was in 1995. If it achieves that rate of growth, it will be replacing the work of approximately 100 million workers (the current total employed civilian labor force in the US consists of 146 million workers.)

In the translation industry - as in other industries - many functions will be affected - whether enhanced, expanded or replaced - by ML.

5.3 Machine Translation

How do humans add value? The long succession of wrong predictions about computers means that they have been looked at mostly in the wrong way. The conventional approach of asking what kind of work a computer will never be able to do is a risky route to determining how humans can remain valuable.

SMT was first implemented in the late 1980s and early 1990s. In recent years, it has contributed to the significant resurgence in interest in MT and to a considerable change in perception. Before the release of Google's online SMT engine in 2007, naysayers claimed that excellent MT

quality wasn't going to be possible for centuries; five years later, the claim was 200 years; early this year, 20 years.

The current stage of a few success stories has been preceded and will be followed by an overflow of failures, which will urge many early adopters to give up and others to take action. The problems lie in the fact that too many users are still looking for immediate results, especially when expectations are artfully inflated, and they invest only when providers prove credible and products reliable. Unfortunately, most MT users don't have the correct data at their disposal. This is due to the still prevailing unwillingness of many data producers to hold and protect their data, which is seen as lock-in asset rather than a selling product.

In the past years NMT has been said to be achieving impressive results, and it is more and more often presented as a replacement for SMT. Advances in artificial neural networks are bringing extremely high expectations, suggesting that NMT could rapidly achieve higher accuracy than SMT. Independent evaluators find that NMT translations are more fluent and more accurate in terms of word order compared to those produced by phrase-based systems. They are also more accurate at producing inflected forms of words. However, NMT may perform less well when translating very long sentences. Solutions known as "attention" networks are being used to correct this.

Better quality MT will mean that a broader range of document types and audiences can be addressed. For example, NMT is particularly successful at producing better understandability when handling more vernacular content such as user-generated text, which is less "structured" than formal prose. This suggests that NMT can be usefully applied to speech translation environments, as spoken input tends to be more like user-generated content and less like professionally written texts. It also indicates that NMT might help improve the quality of translating lectures and other types of long-form spoken discourse.

In turn, NMT will help the further expansion of speech-to-speech (S2S) technologies, now available mostly as English-based monolingual systems. Transforming these into multilingual systems implies many deep and expensive changes. Most S2S technologies are still at an infancy stage and confined to university labs. NMT will help bring speech-enabled devices to the streets.

Since translation businesses have a long history of reluctance in investing in technology and innovation, they will most probably wait for NMT to be commercially viable to speed processes and reduce costs. SMT is going to be largely and increasingly pursued along with post-editing. In fact, despite changes in perception, most translation industry players still do not expect the quality of MT output to be hitting the levels of human translation anytime soon.

The operating model of almost all translation industry players is as old as the translation industry itself, with profit margins constantly decreasing in real terms despite the growth of demand. The investments to ripe the many promises coming from the technology side are burdensome, and many businesses lack the critical mass needed. Therefore, MT will lead to the ultimate disruption in the translation industry when, only the premium segment of artsy—and possibly life sciences—translation will remain tradable.

5.4 Quality Management

Space Race has produced an impressive amount of innovations and changes in the last 50 years, paving the road to exponential growth in technologies. The growth is exponential because the results of one stage are used to create the next one. The moment in time when this growth will impact human civilization is rapidly getting closer. This moment is called 'singularity.'

Singularity in translation might be even closer than Ray Kurzweil's prediction of 2029. Thereafter, machines will no longer need humans to learn, having the means to become self-improving. Although far from perfect, the quality of MT output is already acceptable in many cases, and so convenient anyway that users are learning to live with it.

Assessing the level of performance within a supply chain and the level of skills in an organization is vital to every innovation and automation strategy. And yet, translation quality is a vague and deceptive notion, and its assessment is mostly based on perception. The issue of information asymmetry further impedes a comprehensive, all-inclusive rating of the translation effort, and, with the exponential escalation of language pairs and volumes, quality has become a major concern for customers.

Due to the uncertainties intrinsically involved in translation quality assessment, and the fixity of the relevant concepts in the translation community, users seem now willing to accept good-enough MT output, especially for large volumes, delivered virtually in real time. For serviceable MT output with no human intervention downstream, TAUS coined the acronym FAUT (Fully Automated Useful Translation) already in 2007.

Investing in quality-related decision support tools has become essential to gain translation project insights and benefit from MT. With the increasing availability of ever better and cheaper ML platforms, data's potential is extending to translation quality assessment as well. Analytics methods are being applied to translation data to analyze past business performance and combined with common metrics to measure it. A serious issue is represented by the amount and depth of translation data—i.e. project data. Translation data is much less than language data: it has a limited lifespan and, at some point in time, becomes outdated and irrelevant.

To put it all into perspective: the US retailing giant Walmart handles over one million customer transactions per hour and imports them into databases estimated to contain more than 2.5 petabytes of data. On the other hand, the major ten translation companies process approximately 3.8 billion words per year (in bytes, less than half of Walmart's.) On a yearly base, this number corresponds roughly to 0.005% of the overall amount of words processed by Google Translate or 1.6% of the amount of words SDL claims its machine translation engines now translate.

Applying machine learning to data-driven translation quality assessment will be a disruptive innovation that will call for a major shift in conception and attitude, from prescriptive to descriptive, from normative to empirical, through templates of attributes for different combinations of content types and industry sectors.

Data-driven applications in translation quality assessment will go from document classifiers to style scorers, from comparison tools to automatic and predictive quality assessment, from content sampling to automatic error detection and identification. The data-driven approach to quality will require another major attitude shift. Traditionally, translation companies show little understanding of business data. Translation companies are generally reluctant to share their data, even when presented with the real chance of a profitable use, and are generally inclined to hire people whose profiles make them immediately employable in typical positions. For a data-driven approach, a different mindset is required that is usually developed outside the typical recruitment areas of the translation industry.

5.5 Data

The potential of datafication has already become visible in recent years, and it is now in full swing.

Fifty year after Philip K. Dick's *The Minority Report*, 'precogs' are for real. Machine learning pods have taken the place of mutants nonsensically babbling into computers to have their gibberish converted into predictions. Big data instead of punch cards, supporting mathematical, predictive and analytical algorithms for weather forecasts and predictive policing. In the meantime, we are harvesting the yields of massive introduction of ML in many spheres of activity, from self-driving cars to speech and facial recognition.

Data has been the fuel of automation, and after entering the automation era at full speed, we are being challenged with many issues.

From a strictly business standpoint, any task that can be systematized and automated should be. The automation of repetitive, blue-collar-type jobs has been underway for a few decades, and no industry is going to be left untouched as ML platforms become more and more affordable. We are going to face the paradox of an exponentially growing number of products and services, manufactured and rendered more and more efficiently, with rising unemployment and underemployment, falling real wages and stagnant living standards.

The major challenge will then be to rethink economic models to accommodate this future. And since this future is fast approaching, we must act as quickly as possible.

Translation data is typically metadata: data about translation that can be harvested downstream the closure of a translation project/job/task, such as content type, language pair(s), domain, subject, number of characters/words/lines, quote/price, scheduled time, time spent, technologies used, translation stats (e.g. source - translation memory match, automatically propagated, machine translated - edited, approved) date and time of last saving, etc.

The analysis of translation data can provide a very valuable insight into the translation processes to find the best resource for a job, to decide what to translate and which technology to use for which content.

Translation data will be more and more frequently generated by algorithms. More data will come from rating staff and KPIs. All these kinds of data will come from ML applied to translation management platforms, which will get rid of human involvement.

Erroneously, also data covering multilingual text resources is labeled as translation data. In fact, language data specifically consists of translation memories, corpora, and lexicographical and terminological collections. Of course, all these resources have metadata too, which could be exploited.

Both kinds of data drive translation automation, and they are clearly becoming strategic. Yet, finding good data is not easy. Firstly, because data sources are heterogeneous; secondly, because this data is still mostly crafted manually; and finally, because it is generally poorly assembled and maintained.

A major problem with translation data—even more than with language data—is in the amount and depth of it: it has a limited lifespan and, at some point in time, it becomes outdated, inaccurate, and irrelevant.

Also, sharing productivity metrics across the translation industry is still problematic, because

translation companies traditionally show little understanding of business data. Stakeholders should become more open and massively start sharing their translation metadata to make it the real big data of the translation industry. This is definitely the next challenge. Also, responsiveness becomes imperative for producers of public data that is generally unfit for reuse.

The most vital issues in this respect will remain education and training, which will be more and more important in order to manage the automatically-generated data and to use internal and external benchmarking for business intelligence.

There is a strong need for data scientists/specialists/analysts, but this profile is still absent from the translation industry. Translation companies should be looking out for these specialists who can mine and use data for automation. This will most probably lead to a further reduction of the number of translation companies that are able to float and thrive in a more and more competitive market.

Data is now a new commodity, and yet, businesses will still hardly be attracted to and willing to pay for larger and larger collections of cold data because they require costly specialized knowledge and tools to process and analyze.

With the emergence of NMT, the translation industry may eventually be no field for big (language) data. On the other hand, project data will become essential, even in a small-data perspective.

Small data is what was once called data and it is not made useless by the advent of big data. Most data in typical daily lives is small data: 'small' simply means that its volume and format make it suitable for human comprehension, i.e. accessible, informative and actionable. In other words, big data is to machines what small data is to people. Small data is what we can see as a result of big data analytics. It is the data humans can grasp and convert into information to take decisions. Rufus Pollock, of the Open Knowledge Foundation, stated that the hype around big data is misplaced: small linked data is where the real value lies. This next decade belongs to distributed models, not centralized ones; to collaboration, not control; and to small data, not big data. If big data is all about finding correlations, small data is all about finding the reason why.

The challenge for the next few years might be the standardization of translation data in order to shape it and make it convenient for users to derive the maximum benefits from it.

5.6 Interoperability

Interoperability is the ability of two different systems to communicate and work together through a common language or interface. It implies open standards and may be developed downstream.

Wide-scale system interoperability across the industry is essential for translation to become a utility. Connecting tools and exchanging files - across different translation tools, translation management systems and content management systems - without loss of leverage and data remains one of the biggest challenges in the translation industry.

While many other industries have flourished thanks to standardization which led to interoperability, automation and innovation, the translation industry has always suffered from a lack of interoperability. This has been costing a fortune for years, both on the client side (in translation budgets) and on the vendor side (in revenues).

The lack of interoperability is not just the price to pay for doing business with multiple vendors and different tools. In fact, the problem has never really been the lack or the maturity of standards, but rather the lack of an organizing body, an umbrella organization capable of leading the effort and monitoring the compliance.

Things have been changing a lot since 2011, [when TAUS published a report on the costs from lack of interoperability in the translation industry](). Many blame the lack of compliance to interchange format standards as the primary barrier to interoperability, and no one believes any longer that true interoperability in the translation industry can be achieved only through awareness programs, education, and certifications. Interoperability should come from the adoption of standards created by consortia and not from the dominance of a market leader. Unfortunately, the efforts of the many groups that have been investigating this challenge do not answer to a clear comprehensiveness strategy.

Translation tool providers have always been opting for a lock-in strategy rather than for interoperability, sometimes providing—or urging their users to provide—APIs to solve integration issues, sometimes offering to support the integration effort. With a fee, of course.

The spreading of MT has forced a breakthrough in the interoperability dilemma, starting a wave of innovation and renewed efforts. Most of these efforts have still been focusing on APIs though, as XML has been established for years as the common language, bringing everyone the industry to find its child formats TMX and XLIFF essentially enough.

So far, most of the many APIs made available are meant to simplify the translation business process and reduce translation management and overhead cost. Only a few have been designed to help disintermediation and facilitate access to services. In the past, there have been several attempts to address this challenge. It is true that the largest area for productivity improvements is workflow automation, but the strongest pressures on efficiency come from customers. Based on experience, it is highly questionable that any initiative taken within the translation industry would be successful since industry players are more interested in fencing their gardens and protecting the status quo.

Once again, the expected new upsurge of MT with NMT might smack the traditional sluggishness of translation industry players or force them to anticipate full disintermediation and reach an agreement over the definition of a standard common interface for translation services.

In this case, we could expect that the most influential buyers of localization and translation services will advance their requests; the technology vendors with the necessary technological and financial resources will fulfill those requests or even introduce their own solutions on the market, just as it happened in the past. On the other hand, customers have invested so much in vendor relationships and technologies that in some ways they have become part of the problem. They also have already heavily invested in workarounds for legacy interoperability problems.

During transition, system integration will still have to be made through a plethora of APIs, defining functionalities that are independent of their respective implementations. Once the major issues of data specifications are solved, we can expect a reduction in the number of APIs.

In fact, losses shall be expected in the exchange of language data through APIs as the definition of standards will be lagging behind the proprietary implementations of translation tool vendors struggling to keep customers locked in. A major breakthrough could be represented by the reduction of the current spectrum of metadata elements and by the standardization of their specifications.

In 2016 TAUS transferred all the work it had done in the previous years on an open translation web services API to GALA. GALA has taken over the lead now and [relaunched the industry initiative under the name TAPICC]().

5.7 Academy

In *You Are Not a Gadget: A Manifesto*, Jaron Lanier wrote, "Any skill, no matter how difficult to acquire, can become obsolete when the machines improve."

Translation education is vocational by definition: it prepares people to work in the trade as translators. None of the skills translation students acquire is anything but sophisticated. We tend to forget that, in 1964, the infamous ALPAC committee emphasized a) that the average salary of translators was markedly lower than that of government scientists; b) that the supply for translators largely exceeded the demand, and c) the need for more basic research in computational linguistics.

Most people's essential skills remained largely the same, from the emergence of agriculture 12,000 years ago to the dawn of the Industrial Revolution in the mid-18th century. The transition to an industrial economy in the Western nations - and the related shift in skill values - took well over 100 years. The subsequent transition to a knowledge-based economy took most of the 20th century.

Today, many players in the translation industry complain about the lack of good translators, but they seem to ignore that, more than in many other academic fields, translation education follows obsolete models that are still shaped for the 20th century.

To make matters worse, the gap between the academic world and the industry is so wide that, when approaching the job market, translation graduates instantly and bitterly realize they don't know much about the actual work they are supposed to do. They also discover that the world is not interested in their basic skills. If you imagine 100 billion words are translated by Google Translate per day and perhaps another 1 million words per other tech giants, not even all human translators combined could compete with this number of words per day.

Education never goes hand in hand with trends because everybody pretends it is a vital component in every innovation, even though it is costly and that it seldom brings immediate returns. Therefore, translation education is often left to public institutions.

On the other hand, translation makes the news only when it is bad. With the ubiquity of translation due to globalization, this means that it is commonly perceived as generally satisfactory. Indeed, as a survey conducted by One Hour Translation in May 2015 showed, most people are incapable of separating the wheat from the chaff.

Do you think you can tell the difference between human translation and machine translation?
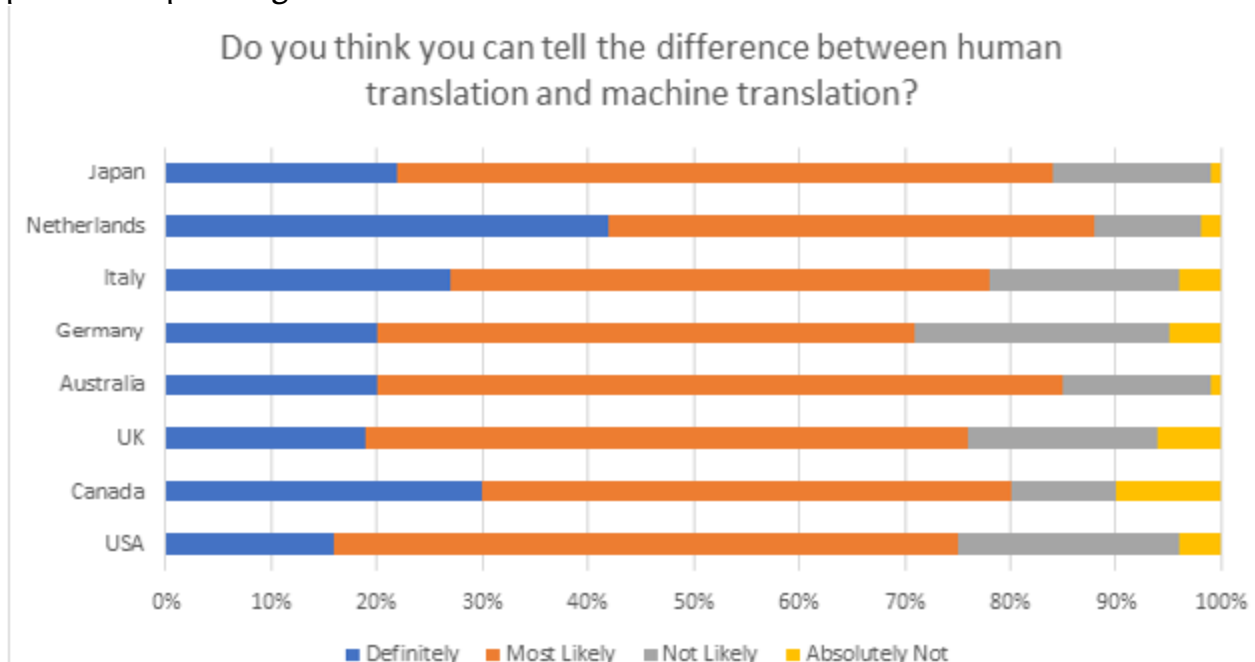
The future may not really need translators, at least not in the old way, as the audience will become even more forgiving for lesser quality of fast-moving content.

A highly-automated localization environment will depend on human skills in quality evaluation, content profiling, cultural advisory, data analysis, computational linguistics, and gradually less and less in post-editing; translating plain text will indeed be a long-tail business.

On the other hand, a recent study of [DeVry University's Career Advisory Board](#) presented at the SXSWedu national education conference in Texas resurfaced a now notorious skills gap in technology. Findings show that only 11% of organizations believe that colleges and universities are providing graduates with skills to meet the tech needs of their businesses and institutions.

According to Zuckerberg Media's CEO Randi Zuckerberg, tech skills are not just programming and coding and engineering; in pretty much any job, whether running a small business like a floral shop or being in marketing, tech skills are necessary. In both the enterprise and the public sector there is a growing need for people with the skills to analyze a business or a procedural problem and solve it by applying technology.

Not surprisingly, Chris Wendt, Group Program Manager for Machine Translation at Microsoft, says that he would not recommend to his children to choose translation as a profession. He'd rather try to redirect them to be a language and culture consultant. In his opinion, transcreation and adaptation between cultures will remain necessary for one more generation, until the differences between our planet's cultures will have been reduced to minor deltas that machines can bridge.

One can no longer afford the luxury of getting too comfortable or complacent: the tech landscape is changing every 2-5 years and the skills one may have now are not necessarily going to be useful in five years.

The success of any innovation depends entirely on the people that are going to nurture, develop, and implement it; in times of exponential growth, education is vital to drive adoption and prepare the next generations of workers. Employers should play a part in closing the skills gap with continued professional training.

It is never too early to prepare for the future; vast workforce and organizational changes are necessary to upend stale business models and related processes.

Any economic advancement involves what the Austrian-American economist Schumpeter called "creative destruction" ("the process of industrial mutation that constantly revolutionizes the economic structure from within, destroying the old one and creating a new one"). Global economic forces are unstoppable just like technology innovation itself. Should we have abandoned ATMs to protect bank tellers?

The simplistic policy answer is better training. But at this pace of change, an improvement of educational system will come perpetually too late. While the academic world continues to follow a creative approach in translation teaching, linguistics has a more scientific address. The adventurous and futuristic side of translation depicted in the movie *Arrival*, for example, has all the traits of fiction and very little in common with the reality.

More than through translation, the language industry should be promoted through better opportunities; and these can only come from real innovations. Translators' pays will be lower and lower as technology improves and the atavistic fragmentation of the translation industry continues leaving customers dominate its dynamics. Linguistics studies may contribute much more

and better to innovation than translation studies. And linguists would probably enjoy many more opportunities than translators.

In the immediate future, translation industry stakeholders may promote new accreditations to improve the industry's credibility and increase 'job transversality', promoting worker mobility across industries: a translation project manager can hardly be placed elsewhere today.

In *Humans Are Underrated*, Geoff Colvin reported that, when conducting their study in 2013, Oxford economists Carl Benedikt Frey and Michael A. Osborne asked employers which skills they would need most in the next five to ten years. The answers did not include business acumen, analysis, or P&L management, but relationship building, teaming, co-creativity, brainstorming, cultural sensitivity, and ability to manage diverse employees. This means that there is no point in trying to beat machines at their own game. What makes people special is their innate propensity for social interaction, this is what makes them most different from the logic-processors. Emotional intelligence will remain one of the most valuable human skills.

Language services today stand on the cusp of a disruptive transformation that will redefine how professional linguists work, forcing them to become far more efficient than ever every day.

Academic institutions still focus on teaching low-value tasks that consume an inordinate amount of their time and energy. Students are still trained to be solution-finders, they are given questions and graded on the quality of their solutions. When developing in their careers, the solutions they propose are still examined, not the questions they have asked.
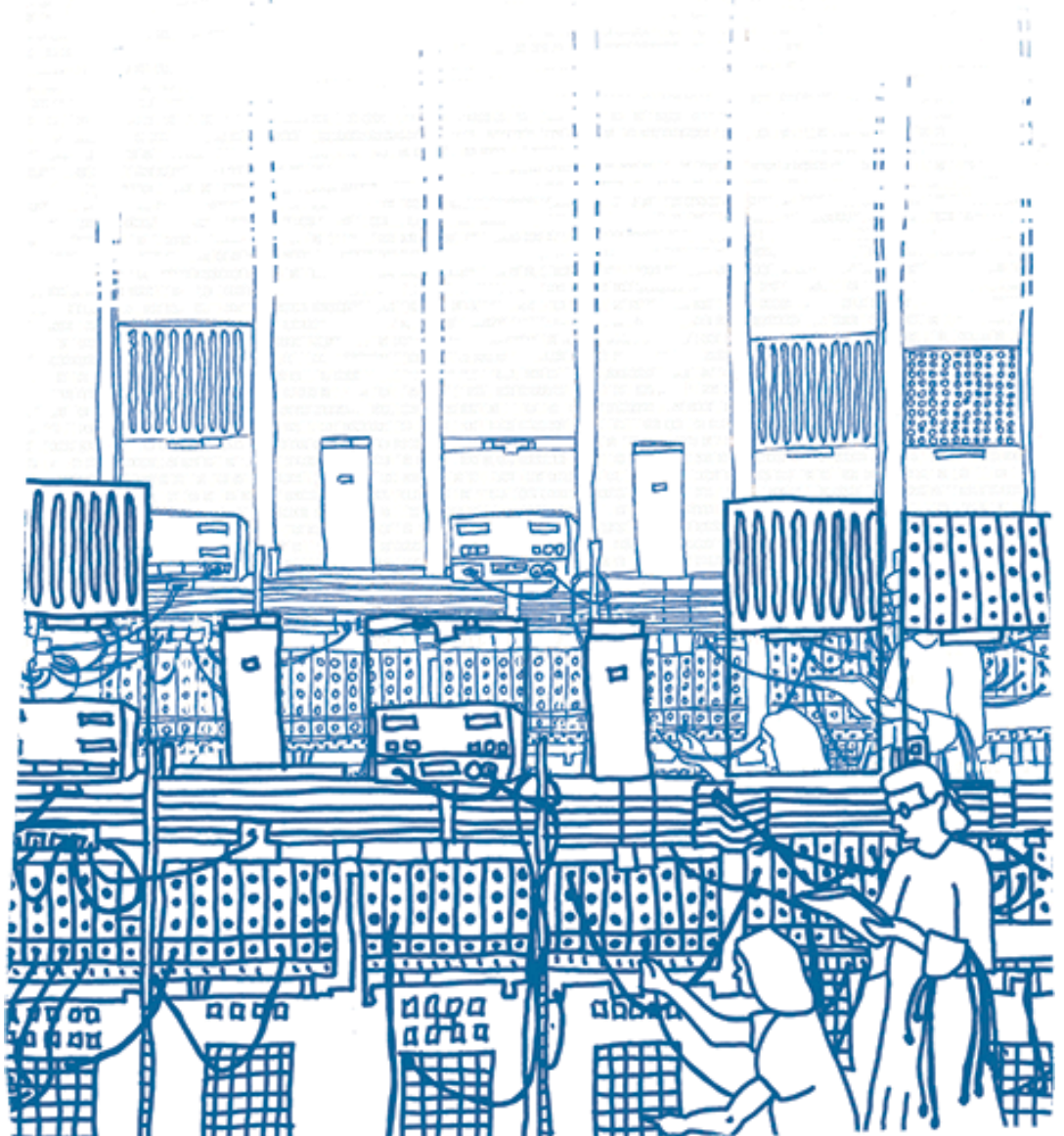
6. Ten Innovation Themes

The six drivers of change discussed in the previous chapter also define the ten innovation themes that we've established for the coming years. With new advances in the fields of machine learning, machine translation, quality management, data, interoperability and academy we can make huge steps in our industry. Startups are popping up everyday and everywhere, introducing the world to new technologies and new ways to use technology.

We might say that there is an ordinary battle going on between man and machine, with a victory for the latter. And at a deeper layer, there is a fascinating intrigue within our industry at the moment, with many threads about game-changing technologies and trends and an outcome that is perplexing even for all of us who think that they are behind the wheel today.

The translation companies of today will not be the same in 2022. We'll see a split in translation tech and the creative networks, the data factories and the storytelling, the platforms and the boutiques, perhaps sometimes still operating under the same umbrella, but clearly separated in functions. Sounds familiar, this story? Perhaps you are thinking about the paradigm shift in the advertising and marketing industry. Once thought to be so creative, it had its own unique place in an environment of factory and office automation. But now, after a few decades of data storms, the business of the prestigious advertising agencies has changed, fundamentally.

Marketing is automated and driven by data and clicks. The incredibly rapid rise of online ads, razor-sharp marketing, and pay-per-view through companies like Google and Facebook has turned the landscape upside down. Legendary names like Saatchi and Saatchi, McCann Erickson, J. Walter Thompson give us sweet memories of the days of Mad Men, but the creative directors now all report to giant holding companies acting under dull names like Omnicon, WPP, Interpublic and Publicis.

Similar mergers and acquisitions are likely to happen in the relative small translation industry in the coming five years and a convergence with that other creative sector that has fallen victim to data storms - the advertising and marketing industry - would make a lot of sense.

6.1 Algorithmic Management

Machines will be better at almost everything humans do. Like in many other sectors we expect the robots to come into the translation field to enhance the work we do, expand it and ultimately replace us. The "rise of the robots" started very simple with the counting of words of the source document to be translated. Now it has expanded into matching jobs with translators, identifying new terminology, optimizing the leveraging from translation memories, profiling content, sampling for review, deciding on the type of quality evaluation and error types to be checked, status tracking and reporting, invoicing and delivery.

Soon, robots will check the quality, productivity and even predict the quality of jobs yet to be performed. They will track the ROI on the translation of each individual message or segment based on how many users viewed the translation. Not to speak of course of producing the translations itself, speech translation, pay-as-you-go and all other innovations driven by algorithms. These are covered in the other mini-chapters of this story. We have heard many start-ups in our sector already refer to themselves as the Ubers of translation. Well, here you go: self-driving translations will be the norm in '22.

6.2 Datafication of Translation

The datafication of translation started with the [unreasonable effectiveness of data](#) article written by the Google scientists Fernando Pereira, Peter Norvig and Alon Halevy in 2009, or perhaps even earlier when the TAUS Data Cloud was launched in 2008. Translation learns from data. In those early days indeed there was no better data than 'more data'. The English-French Google machine translation engine was trained by a corpus of 100 billion words. Now, with the new generation of Neural MT, very large quantities of data belong to the past. The pursuit of high-quality in-domain translation data will challenge the protectionists and create opportunities for pirates.
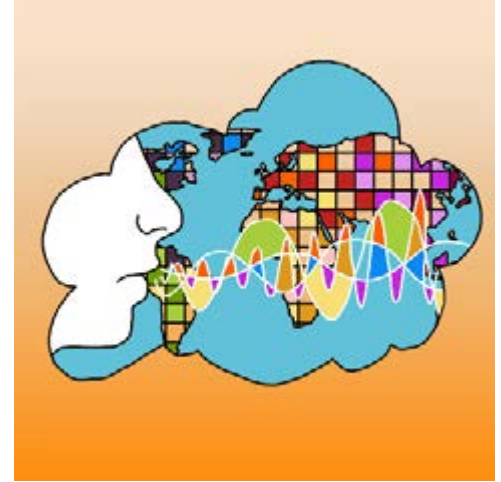
Data has become an obsession, either way, in the translation industry. And it does not stop with translation memory data. We need speech data too. And we want to have the edits and annotations on human as well as machine translations, plus the attributes for content types, industry sectors, translators' locations, the process applied, the technology used. And why not correlate it with the weather reports, the social graphs of the translators and their eye movement tracking? There is always something we can learn from new data.

The internet giants had a competitive edge in translation data, but they spoiled it by polluting their own fishing grounds with machine translations. Now, the hunt is open for new data marketplaces. The European Commission is investing in the Connecting European Facility. But watch out also for the greenfield translation data ventures in China, or perhaps closer to home: the TAUS Data Market.

6.3 Quality Matters

Let's not be blinded by data and technology. The ubiquitous availability of dumb translation will only drive the demand for stories and messages that trigger the user's imagination, that build the customer's brand and engage global and diversified communities. Here is where the synergy between the translation and the advertising and marketing sectors makes the most sense. In an open globalized economy, stories need to be recreated. There is enough bad content out there already that gets generated and translated automatically.

Data is a great help to gain insights into markets and customers. What they like and don't like, where they click and not click, That allows us to take informed decisions where to invest and not to invest when it comes to content creation and transcreation. The way we see it: translators become writers, journalists, and storytellers, cultural consultants, global brand promoters.
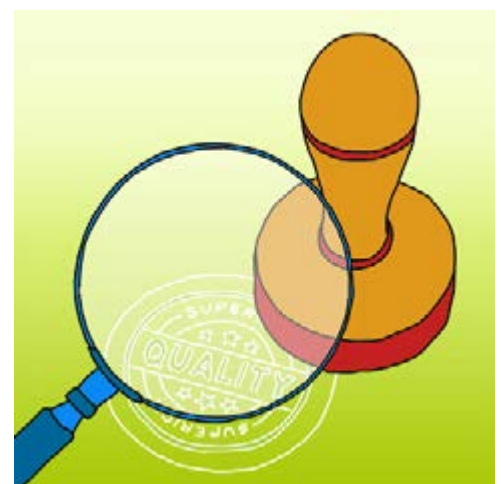
6.4 Convergence

Convergence is the confluence of technologies, business models or markets. It is at these crossroads where the biggest innovations are happening. The best example is still the success of Google Translate: a huge wake-up call for the translation industry. Google had no intention to disrupt the translation industry, was uneasy about it even. But the enormous popularity of the automatic translate button combined with Google Search opened perspectives that some pioneers had dreamt off as early as in the 1970's (see the story about Jean Gachot offering Systran Translate on Minitel in Paris).
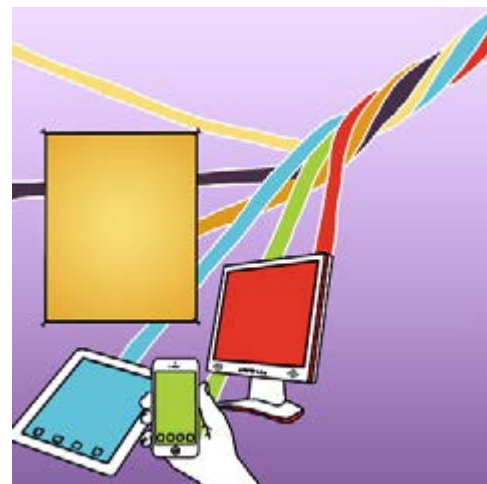
In the next five years, convergence is the thing that will drive the most fascinating and imaginative innovations. The Megaphonyaku and the Wearable Translator in Japan are good examples of how the convergence of machine translation and speech technology can help tourists and travelers manage themselves in a foreign language environment. The Tokyo Olympics of 2020 will likely bring us more surprising translation innovations. It only takes a bit of imagination to think of similar innovations in business and industrial environments. What would you think of machine translation, text-to-speech and video conferencing all coming together in the Microsoft Hololens to support the John Deere field engineer in Vietnam to interact directly with the factory in Germany when he encounters a defect in the tractor? And in a similar way, would the laboratory assistant in China working with a blood testing machine from Roche Diagnostics not be better off when the device speaks to her in Chinese with an instruction for use or a response to a question when asked?

The power of such convergence will particularly be strong in all those cases where today we produce translations that nobody reads because they are not there when and where they are needed.

6.5 Speech

As forecast in the recent TAUS Speech-to-Speech Translation Technology report, we can expect to see good working solutions of speech-to-speech translation coming on the market in the next five years. The Skype Translator and the Japanese wearable devices mentioned above are just a few early examples. People are lazy and generally, prefer listening and speaking over reading and writing. The technology is there and it is working.

What is needed now, are the voices in the required languages and, not to forget, the data to train and automate the speech-to-speech translation systems. This is another new area for language service providers to expand and develop new services, finding and hiring the talents. Spoken translation will pop up in many apps, on wristbands, in glasses and integrated into products from software companies, manufacturers of automobiles, medical devices, and online services.

To scale up for speech-to-speech translation at levels similar to text translation the services and technology companies in our sector will have to develop new processes, workflows, data collections, and tools. Opportunities arise for new specializations.

6.6 Bite-Sized and Long-Read Translations

A demarcation line is becoming clearer in the next five years, between bite-sized translations and long-read translations, between product localization and transcreation. Product localization has its own characteristics. Gone are the days of once-a-year big product releases. We already live in an age of continuous delivery. Tasks and jobs become smaller and smaller, down to a segment and phrase level. The concept of a project is blurring. Processes become more agile and integrated with product development.

The quantum leap in machine translation comes at a rescue. But whether or not machine translation is used, translation is always on and virtually real-time. From a service and technology provider perspective, platform integration seems to be inevitable. Several innovative cloud-based platforms with drag-and-drop translation features at various quality levels have appeared on the market in recent years. They will challenge the status quo and force many vendors and buyers to follow.

But what if more than a human touch is needed, when the story needs to be recreated or transcreated in other languages and for other communities? Translators become writers and cultural consultants, brand promoters. They become crucial to the success of a product in a new market. The cascaded supply chains come under pressure and disintermediation will become a theme, again. New ventures like Translate and Create will challenge established translation service houses. Convergence with publishing and advertising and marketing services seems a natural way to go. It will be interesting to see how the market evolves around this widening gap of automatic and creative translations.

6.7 The Quantum Leap

In 2022 we will look back at a five-year sprint of Neural MT that has brought an unprecedented improvement in machine translation quality. If it could be expressed in a percentage, the experts will say that five years of NMT is equal to the twenty years of SMT that preceded. It is spooky in the way that even the researchers often don't know what sparked these rapid improvements. The machines have become self-learning and take a thousand decisions to arrive at better results.
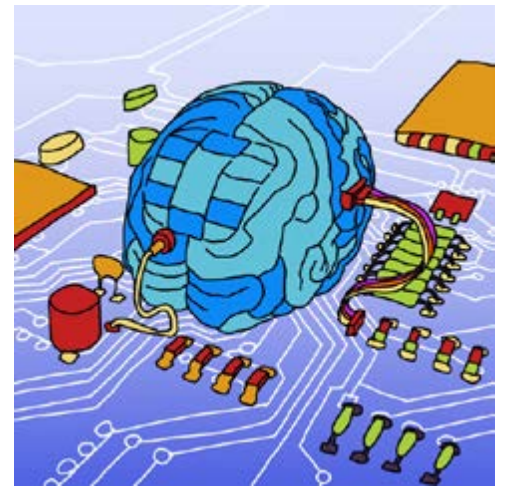
What's more: the Deep Learning technology takes any data - monolingual, bilingual, audio and even videos - to build versatile engines that can do lip reading (and translation) and even translate between languages for which no direct bilingual data were available. The machine translations are more fluent and natural than what we are used to today, hiding potential inaccuracies from our blind eyes. Neural MT has lent itself very well to the emergence of more speech-to-speech translation apps, thanks to the fluency features and the lower burden on storage capacity.

All of this leads to a reality that all published content is translated in at least fifty languages by default or at least virtually, meaning that it is available on-demand, real-time and probably free. This is a giant step from today's reality, even though today (2017) the total volume of machine translation output is already 500 times larger than the total translation production of all human translators together.

Rumour has it that the success of Neural MT is followed by a new breakthrough in MT technology that some already refer to as Quantum MT. The Quantum MT generation may be able to add the precision to machine translation that is needed to bridge the accuracy gap in the current Neural MT systems.

6.8 The Long Tail
In 2022, the translation industry will be much less constrained when it comes to the spread of languages. Ongoing globalization will continue to open markets. Populist trends in politics will hardly stop the pursuit of more customers around the world by businesses both from Western and Asian countries. In fact, what we will see is that the ease of e-commerce brings customers closer, also to small and medium-sized companies, stimulating further growth in global trade.

The China One-Belt One-Road program, for instance, is primarily focusing on the countries along the historic Silk Road, adding quite a few languages to the mix, covered by the translation industry. If today a global enterprise covers on average 25 languages, we speculate that this will double in the next five years. The ubiquitous availability of machine translation can fill the gap to a large extent.

The new generation of Neural MT systems not only improves performance but also adds new languages faster and easier because it is less dependent on very large quantities of translation pairs. The so-called zero-shot approach to building machine translation engines lets developers build engines out of unpaired data of different, but related languages. Ironically enough, rather than putting translators out of a job, we think that the availability of bulk machine translation in new languages will stimulate the demand for creative long-story translations. The interest in new cultures and languages will grow as a result of the technology.

6.9 Pay as you Go
The ubiquitous availability of translation as a utility, often free but not always adequate or good enough, triggers new business and pricing models. Why not charge only when the user clicks and consumes the translation? Why not differentiate the pricing depending on content profiles, the number of clicks? Or perhaps introduce a metric that helps us determine when a high-quality human translation is needed.

The language service companies have become used to reinvent their business and refresh their service portfolios and will continue to do so in the next five years, only more rapidly and radically. The fixed price per word model will go away. Translation tech firms will charge based on the use of their platforms. The creative translation firms will rather price on an hourly basis depending on the talents that are needed.

6.10 Users First
Particularly in the translation industry, we have lost sight too often of why we translate, whom we translate for, and how the translation is used. Too often translations are produced as an obligatory item in old-fashioned push or publishing models without much care for usability and findability. This must change if we are to follow the trends towards more democratic and user-centric business models. Service providers in the translation sector are likely to play a more

important role in consulting on usability and cultural diversity. In fact, there can be tremendous merit in engaging users in defining features, labeling, translation and terminology of services and products in new markets. Users become the new talents.

6.11 Epilogue

In 2022, when we look back on the story of the translation industry in the past five years, we will not see such a smooth journey through the ten chapters highlighted in this story. Of course not. There will be hiccups and fall-outs, trials and errors, and severe competition. The providers in the translation industry continue to spend an inordinate percentage of revenue on old-fashioned sales. The industry will suffer from the Bodo Dilemma*: an abundance of tools, technology, data and innovative solutions combined with a painstaking shortage of talents.

Data becomes increasingly important, while the technology edge is diminishing as a result of more open source solutions and sharing of the latest advancements through academic papers. The overriding trends towards data, data-sharing, machine learning and the cloud will, on the one hand, lead to fascinating innovations, growth, and maturity but fire up concerns over privacy and security on the other hand. We may see translation blocking as a phenomenon, similar to add-blocking, to protect companies and supply chains from unwanted and unsolicited free translations.

Some providers will prosper offering secure and closed translation services and technologies on-premise. But in the long run, in the next lustrum of the industry perhaps, the cloud may prove to be irresistible even for the most paranoid, especially if and when the new quantum computer technology delivers the ultimate safe security. Who knows?

*The Bodo dilemma is named after Bodo Vahldieck, Quality Manager at VMware, who expressed his frustration at the TAUS Industry Summit about not being able to find young talents who are willing and able to come and work with the fantastic localization technology suites at his company.

7. Further Reading

TAUS Reports and Articles

https://www.taus.net/think-tank/reports/translate-reports/taus-speech-to-speech-translation-technology-report

https://www.taus.net/think-tank/reports/translate-reports/taus-machine-translation-market-report-2017

https://www.taus.net/think-tank/reports/translate-reports/taus-translation-technology-landscape-report-2016

https://www.taus.net/think-tank/reports/translate-reports/lack-of-interoperability-costs-the-translation-industry-a-fortune

https://www.taus.net/blog/is-there-a-market-for-mt

External Links

http://www.nytimes.com/2010/01/15/books/15book.html?pagewanted=all&mcubz=0

http://ieeexplore.ieee.org/abstract/document/4804817/

https://www.amazon.com/Humans-Are-Underrated-Achievers-Brilliant-ebook/dp/B00OZ0TLBK

http://www.mckinsey.com/~/media/McKinsey/Industries/Advanced%20Electronics/Our%20Insights/How%20artificial%20intelligence%20can%20deliver%20real%20value%20to%20companies/MGI-Artificial-Intelligence-Discussion-paper.ashx

http://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf

8. About the Authors

Isabella Massardo

Isabella Massardo has worked in the translation industry since 1988, as a translator, trainer and writer. She writes regularly on topics relevant to the business of translation, with a particular focus on technology and innovation. Her copywriting includes marketing materials for a range of global clients, from language service providers to tech companies. Isabella is an ECQA-certified Terminology Manager. She holds a master's degree in Russian Language and Literature from the University of Parma, Italy.

Jaap van der Meer

Jaap van der Meer was the founder and CEO of some of the largest global translation and localization service companies in the 1980s and 1990s. In 2005 he founded the Translation Automation User Society (TAUS). TAUS is an innovation think tank and platform for industry-shared services for the global translation and localization sector. Many of the largest IT companies, government translation bodies and their suppliers of translation and localization services and technologies are members of TAUS. TAUS offers among others a platform for translation quality evaluation and benchmarking and a platform for pooling and sharing of translation memory data. Jaap van der Meer has written many articles over the years about the translation industry.

Reviewer

Anne-Maj van der Meer

Events Director

About TAUS

TAUS, the language data network, is an independent and neutral industry organization. We develop communities through a program of events and online user groups and by sharing knowledge, metrics and data that help all stakeholders in the translation industry develop a better service. We provide data services to buyers and providers of language and translation services.

The shared knowledge and data help TAUS members decide on effective localization strategies. The metrics support more efficient processes and the normalization of quality evaluation. The data lead to improved translation automation.

TAUS develops APIs that give members access to services like DQF, the Quality Dashboard and the TAUS Data Market through their own translation platforms and tools. TAUS metrics and data are already built in to most of the major translation technologies.