The innovation and interoperability roadmap for the translation industry

TAUS Research Report

May 2009

|T|A|U|S|

Translation Automation User Society

Readers guide

This report represents the author's interpretation of the industry's collective wisdom based on survey data, group discussions, 30 in-depth interviews, and TAUS articles and reports. For a more objective view, the reader is asked to refer to TAUS World Tour (Q4 2008 - Q1 2009) and online survey (March 2009) results, which are provided in the appendices.

This report outlines the industry's development focus in the coming years, highlighting the types of decisions companies are making, spotlighting a few case histories and predicting how events will unfold.

Underpinning this research is the assumption that there is a fundamental and positive shift from business and innovation models based on control and hierarchy to ones empowered by collaboration and openness.

If you are interested presenting at the TAUS User Conference, Portland, USA, 29-30 October, this report should be used as a guide to help shape your proposal.

If you are interested in finding out more about any of the subjects covered in this report, please visit www.translationautomation.com or write to us at info@translationautomation.com.

References to companies and organizations

Across	Language Grid
Adobe	Language Weaver
Asia Online	Lingotek
Autodesk	Lionbridge
CA	Lucy Software and Services
Centre for Next Generation Localisation	Microsoft
Cisco	Moravia
CIA	MultiCorpora
Clay Tablet	Oracle
Concurrent Technologies Corporation	Pangeanic
Consortium for Service Innovation	PROMT
Dell	PTC
DocZone	SAP
DotSub	SDL
Ebay	Second Life
EMC Corporation	Speaklike
European Commission	Symantec
Facebook	Systran
GlobalSight	Translated.net
Google	VeriSign Inc.
IDEST	Wikipedia
KCSL Linguistic Technologies	XML-INT
Knowledge Accelerators	

Thanks to

Many thanks to everyone who has contributed to this excericse for your openness, time and dedication.

Special thanks to Jaap van der Meer, Andrew Joscelyne and Robert Kuhns for your help and support.

Author- Rahzeb Choudhury

Table of Contents

1. Background

The Information Age has led to insatiable demand for translation services, which cannot be met with existing proprietary business models and the capacity of around 300,000 professional translators worldwide. Translation Automation User Society members are world leaders and catalysts for ensuring better use of technology to increase the capacity and significance of the translation industry.

This report comes at a critical moment as two prerequisites for accelerated innovation and greater interoperability, industry-wide language data sharing and open translation platforms, have become reality. This report outlines the industry's development focus in the coming years, highlighting the types of decisions companies are making, spotlighting a few case histories and predicting how events will unfold.

2. Summary

As you would expect our market survey confirms the global economic slowdown has led to less translation volume (49% of respondents) and less languages being translated (20%). However, whereas some other sectors have been devastated, the picture is for the translation is not dire, with 20% reporting no slowdown and 36% exploring new innovation opportunities.

The headline finding is that over 80% of respondents expect to be using Machine Translation (MT) as part of the mainstream of business within two years, up from 37% today. Counter intuitively, only half of respondents plan to share language data, a prerequisite for improving MT output quality, which is considered the main barrier.

We expect more MT providers to move to service oriented business models, and form partnerships with academic institutions and language service providers. We note that there are clients and providers building on top of open source MT engines, fuelling this movement.

 As technical constraints are overcome, early adopters of language data sharing will rapidly gain cost advantages. When sharing takes off, barriers to entry for MT will come down, more MT languages will become available, and domain specific engines will raise quality. This will improve the competitive environment and super charge the innovation landscape. It will become much easier for entrepreneurs to tap into unmet needs for instant, portable, and personalized translation.

The hype around community translation is settling down and the companies we spoke with have been creative, industrious and level-headed with the crowds that they have been working with. Eighteen percent plan to be using community translation within two years, a reflection of the limits to what volunteers can offer to the business of translating commercial material. Using community translation to improve translation memories and train MT systems is undoubtedly the largest single long-term benefit to the industry.

The take up of open translation platforms is a positive development, whether openness is defined in terms of enabling seamless connectivity between systems, harmonized standards so that there is portability of information and/or open source for fostering greater collaborative innovation. Recent announcements by large, small, old and new companies on plans for more open systems highlights a general trend away from lock-in business models. A quarter of suppliers could be using such systems within two years. This points to a general shift towards Software-as-a-Service based models for language service providers and is the clearest sign yet of progress for an industry reinventing itself.

Over a third of clients are aiming to be working in a continuous translation environment within two years. This will reinforce advantages for providers with open translation platforms, which offer clients much needed interoperability, and drive down customization costs.

Companies are eager to find cost effective models to localize support content. Human translation, whether by professionals or a community, is only workable in a limited number of cases. Improved MT offers the potential to tap into a huge opportunity to provide localization for Knowledge-Centered Support.

Translation personalization remains largely a research objective in academic circles. Two ingredients for making this happen on a bigger stage, industry-wide language data sharing and open translation platforms, are now reality. Success would mean hugely disruptive innovation and result in world changing impact.

This report shows the rapid progress that has made since October 2008 when TAUS started tracking industry insiders' views on opportunities and challenges to innovation and business models. Survey results and in-depth interviews, combine d with recent developments in the industry help to provide a likely roadmap or at least a compass for the innovation and interoperability agenda that lies ahead. The future of the industry's pricing model and empowering changes to the role of translators remain subjects for review in future reports.

3. Translation automation

TAUS survey results clearly show that machine translation (MT) will soon become a central feature of the localization/translation industry. Thirty-seven percent of respondents are already using MT and all but 18% aim to adopt MT within 2 years.

Benefits

The main benefits are seen as reduced costs and increased efficiency (70%). Almost half of clients informed supporting end users with real time translation is also a key benefit and 40% of suppliers see MT as an important asset for service differentiation.

Applications

The main application is as part of the translation workflow for structured content, such as manuals, usually as a tool to support human translation. Companies already familiar with MT are also using or planning to use MT for real time translation to help end users or support staff when gisting is sufficient.

Where usefulness of translation is the benchmark, integration with blogs, chats, emails, video, voice recognition technology, and handheld devices are options. Wider application for unstructured content is not expected anytime soon.

Barriers

The main barriers for adoption are poor quality MT output (58%) and complexity of customization (45%). To improve quality the following is needed:

- ✓ Statistical MT (SMT) - greater standardization of TMs without losing the benefits of company specific tagging. Much more domain specific language data to train MT engines
- ✓ Rule-based MT (RBMT) - greater standardization of terminology

Companies that already have deployed MT for one or two language pairs inform that complexity of customization becomes less of an issue. However, at this point the cost of multiple MT licenses is often highlighted as a barrier.

Selecting MT engines

Companies' trialing MT engines inform that no one engine meets all requirements and that different engines provide differing results depending on language pair. There is a lively debate on the virtues of RBMT versus SMT, with linguists tending to lean instinctively in favor of the former and engineers to the latter. There is a trend towards providers developing

hybrid solutions that aim to harness the benefits of both methods, such as by PROMT, Lucy Software and Services and Systran.

As Wayne Bourland of Dell aptly puts it, "The big next step is to help customers put it all together. It's too difficult for new entrants right now. The industry has to invest in developing a full suite of services that meets the customers' needs. MT is too piecemeal today. "

During our in-depth discussions a number of people highlighted a need for more objective guidance when selecting MT solutions with the implication that this in itself would reduce the cost and speed up the take up of MT.

MT services models

It is clear that there are potentially lucrative opportunities for providers of MT and supporting services as a new market opens up. However, with clients expressing concerns over the cost of multiple licenses and Google's free translation service disrupting the market, the business model for MT providers is not entirely clear.

At the recent TAUS Executive Forum on open translation platforms participants, including major clients, language service providers and MT companies, tended to agree that a services model is likely to win the day. This might includes fees for computational power, training/ customizing engines, knowledge transfer, sector specialization, pre/post-editing, and data cleaning/management services as well as web integration to an MT engine behind the firewall.

There is a deep and global academic resource base, as seen at the 129 pages long Compendium of Translation Software. We expect to see more partnerships between universities and solution providers in order to bolster research and development efforts. To help MT providers' access markets quickly and ensure comprehensive sets of services are provided, we expect more partnerships with language service providers, and the prime example is the recently announced partnership between Language Weaver and SDL.

Open source is generally regarded as a positive development for MT, instantly slashing start up costs. A number of clients and providers are actively working on solutions using open source as the base, with Moses the stand out choice so far in the SMT space.

4.   Language data sharing

Intuitively industry-wide language data sharing (translation memories (TM) and glossaries) makes sense. The benefits of centralizing at enterprise level are well documented and continuing to accrue these same benefits using advanced leveraging technology and better terminology management seems desirable. It is also clear that for machine translation to improve more data is needed to train engines. The vast majority of companies expect to be using MT within two years and a majority also highlighted quality as the main barrier. And yet, only half of all survey respondents intend to share language data, a fundamental building block for improving MT quality, in the coming two years. Our research highlighted a number of reasons for these inconsistent findings.

Business case

Clients (43%) cited the challenge of finding tangible business value as a major barrier. This compares to 30% of suppliers. During interviews, we heard deeply engrained concerns about losing competitive advantage; as well managed TMs are tools to manage costs down by increasing efficiency. There is a lack of awareness of how sharing TMs would be more advantageous than going it alone and why 45 companies, who also compete under the same market conditions, would join forces to form TAUS Data Association (TDA), a global not-for-profit language data sharing organization.

With the current economic climate playing heavily on many minds, there is pressure to put off expenditure (joining fees and administrative overhead) and to adopt a wait and see approach to new initiatives. Only 9% informed they do not see benefits coming from language data sharing. Respondents ranked the main benefits as increasing translation automation and streamlining terminology both at equal first place (56%), followed by reducing translation cost and friction (41%), creating a foundation for innovation (28%), and increasing access to global markets (27%).

Companies that are now focusing on using TMs more efficiently than previously also argue that for them it is too early to start sharing. Some companies also point out that their content is so unique that there is no real value for them from sharing. The immediate action for TDA and its members is further testing and communicating clearly the advantages of language data sharing.

Ownership and intellectual property

Suppliers (64%) and clients (43%) cited uncertainty over the ownership of translation memories as a major barrier. The solution seems like a largely operational one, with logical

question, "Is this worth the effort?" Should documentation be unpublished it would be unlikely to be shared. Once published, there are unlikely to be any confidentiality related reasons not to share relevant TMs. Ownership of TMs usually rests with clients, with agencies and freelance translators signing non-distribution agreements. For clients, translation is a means to an end and there is often little interest in investigating the processes involved. This potentially leaves service providers with a communications task and associated overheads to convince clients that TMs are something to share.

TDA and its members will need to ensure the market is properly informed of the legal framework for the association, data sharing and usage. Those who buy-in to the business case must then be armed with the necessary communications tools to help them convince internally and/or clients/content owners that the profit of sharing is greater than current approaches.

Technical constraints

A number of survey respondents and interviewees highlighted inconsistency in quality levels between TMs and the need for greater standardization of TMs as barriers to leveraging multiple TMs in order to improve translation management systems and MT engines. Others pointed to quality issues with source text being the main barrier to improving TMs and MT.

Inconsistency in quality of TMs appears to be addressed by peer review, filtering, and data cleaning utilities, which are provided by TDA and commercial companies offering similar services. The solutions for standardizing TM come in two forms. Firstly, at some stage in the future the TDA database will host so many TMs that there will be sufficient knowledge to execute a global standard. Secondly, advanced data cleaning services or a technological solution enable the leveraging multiple TMs without losing the benefits to individual companies from tagging.

Beyond the barriers

Translators gain immediately by using the TDA Language Search engine to solve terminological hold ups by comparing usage across TMs from many companies, adding substantially to the current toolkit of dictionaries, websites, in company glossaries and in country experts.

As Manuel Herranz of Pangeanic looks beyond the barriers for language data sharing, "We're expecting substantial quality improvements in our and competitors' machine translation engines, as vast amounts of domain-specific data become available through TDA. It's likely there'll be issues with leveraging multiple TMs if terminology and style are too different. And so certain pre-selection, normalization and standardization processes may be needed."

Within the next few months, it is likely that TDA members will start to make cost savings on translation and terminology management. Access to a giant source of language data will help to lower research and development costs for MT and lower barriers to entry, improving the competitive landscape. Before the end of 2009, we expect more MT solutions with domain specific focuses will enter the market. The bringing together of so much language data into one place will enable harmonization of terminology as organizations can efficiently review and agree usage.

In the longer term many more languages will be added to the MT roster as companies will gain access to the data needed to train engines. Improved quality, greater choice of providers, and the availability of more languages will provide the right ingredients for much easier commercialization of instant, portable and more personalized translation solutions.

Case studies

Advanced Leveraging (AL) technology finds matches at the sub segment level (i.e. phrases) enabling more granular leveraging than TMs. Some AL tools search multiple TMs, pulling together sub-segments from different TMs, and assembling full or partial translation in this way.

IDEST has used the Similis AL tool on European Union content for subjects such as agricultural policies, economic reforms, health and safety issues, and environmental issues when TMs do not provide leverage at the segment level due to differing writing styles between Member States and new text tending not to be incremental extensions of previous publications. Translators have also gained improved terminology support through the automatic identification of source and target terms to increase the consistency of translations.

Concurrent Technologies Corporation has found that using Lingotek's TM suite, LingoSuite, where a number of TMs can be advanced leveraged, led to a two-thirds decrease in translation time and well over 90% accuracy. Lingotek's Suite has been successfully used by governmental organizations, such as the CIA.

Oracle found that in the first six months after deployment, AL led to an additional 20% savings over its use of classic TMs. Oracle is a prime example of successful language sharing, as a succession of acquisitions have been absorbed and TMs incorporated and leveraged.

AL tools are developing rapidly. The European Commission, which is currently looking into applying AL, calls it in-segment leveraging; Lingotek calls AL collaborative leveraging; MultiCorpora, Corpora-based leveraging. KCSL has built a wide range of services around AL.

Sources: TAUS articles and reports; company websites

5.   Community translation

Adobe, Facebook, Google, Microsoft, Symantec, Wikipedia and others have all successfully involved users and customers in community translation. In addition to providing a boost to translation resources they have improved brand perception, user experiences and engaged people to get feedback as well. For Google, improvement suggestions made by the crowd are an invaluable source for training its own SMT engine.

A fifth of clients and 15 % of suppliers expect to use community translation in the next two years, a reflection of the limit to the number o f cases where volunteers will translate commercial products. The scope and significance of helping to increase access to high quality content on the worldwide web is clearly not reflected in these results.

Eri Imai Hagberg of EMC points to some of the more difficult challenges facing companies who are looking to fit community translation in to their overall globalization strategy.

"We are looking at how social media can be used more generally to our advantage, including community translation. We are also working to find ways to integrate TMs and MT into our social media platform."

Approaches to community translation
 During our research the following approaches to community translation came to light.

- ✓ Social network members becoming the community of translators
- ✓ Virally recruited and self organizing crowds
- ✓ Communities of product users who benefit directly from the translation. Often recognized as authors. Most often used for second tier languages such as Russian and Brazilian Portuguese, where there is a shortage of translated material. The process is managed by and post-editing done by vendors
- ✓ Open source product communities with ideological motivations
- ✓ Internal employee communities, such as for in-house content that is not post-edited
- ✓ In business-to -business sectors, resellers/partners who act as the community in smaller markets to benefit from increases in sales revenue/barter arrangements and in some cases share support revenue
- ✓ Communities who post-edit MT output for non commercial content

Quality assurance
There are two approaches to managing quality. The first leaves it to the crowd. Here there may be awards, group voting, peer review, rankings and sometimes a general acceptance

that usefulness rather than publishable quality is the required benchmark. The second approach uses a validation stage involving professional linguists who post-edit.

Issues

Interviewees highlighted difficulty in pressuring a crowd, say to meet a deadline, difficulty in controlling file standards and benefits to the TM database, and firewall issues which are solve d by translation taking place in discussion forums.

The toolkit

Stephan Cocron of VeriSign Inc. crisply suggests exactly what an ideal community translation tool should provide, "Unlimited user licenses, terminology management, clean and intuitive translation interface, translation suggestions, automatic TM updater, MT suggestions, aliases and IDs, admin rights for language managers, in-context translation workbench, simple debugger and resource editor. "

Case studies

By crowdsourcing translation Adobe is able to uncover the optimal use experience for its products, allowing Adobe to build products and markets at the same time. Early adopters in new locales help to rapidly build an ecosystem around products with more content, training resources, and books on APIs etc in more languages. Adobe gains momentum that is not possible using the traditional localization model.

Adobe is taking crowdsourcing a step further by building community translation into its applications, and allowing loyal customers to go into the actual software and translate the user interface.

dotSub has successfully enabled people to upload videos and encouraged individual volunteers to write subtitles. One feature-length documentary has been translated into 33 languages by a self recruiting crowd which has also embedded the film into over 1300 URLs.

At Second Life, translation was initially done completely organically by the social network's members. The enthusiasm for and quality of translation led to Second Life tooling up volunteers and organizing programs for localization. Users localize the client, website, wiki help pages, and the knowledge base. Volunteers are acknowledged with announcements, translator of the month leader boards, and spotlighting highest contributors in the official blog and forum. The main motivation for volunteers is the pleasure of making Second Life a better place and sharing the experience with other people who speak the same language.

Source: TAUS articles; Company websites

---

6.   Open translation platforms

Twenty-one percent of respondents are considering using Open Translation Platforms within the next two years. This figure rises to a third for companies already using MT.

Having worked on the client and now the solution side of the industry, Olga Beregovaya of PROMT provides a well-grounded view on why open translation platforms are so important.

"The need to separate infra from lingua is evident. There's often loss of TM leveraging across different systems. The industry must be educated on open translation platforms. When a vendor introduces their own (closed) translation platform, it's misleading and confusing for clients and subcontractors. Platforms need to be transparent."

Defining open

Open was defined at a number of levels by the people with spoke with. These included each or a combination of:

- ✓ Open Application Programming Interfaces enabling connectivity between systems, making it easier for translation resources (human and language) to flow within and across organizations and making it easier to work with multiple vendors
- ✓ Open standards to ensure that leverage is not lost when TMs are combined, ensuring that MT output can improve without costly manual intervention. It is felt that the current standard, TMX 1.4, has been applied according to individual organizational needs and that greater harmonization is needed for future versions
- ✓ Open source to allow free access to technology and achieve success through a collaborative innovation culture. Moses (SMT) and Opentrad (RBMT) were highlighted as two key initiatives open source MT initiatives
- ✓ Open to community translation

Differing priorities

Interviewees who had participated at the recent TAUS Executive Forum on this subject were generally the most optimistic about the industry's prospects for progressing in these 3 areas and an individual company's position tends to influence which type of openness is considered a greater priority. For example those finding it difficult to build business cases for MT due to high upfront costs were more enthused about open source MT initiatives than companies focusing on improving the flow of translation resources between organizations.

A line is drawn

Recent announcements by large, small, old and new companies on plans for more open systems highlights a general trend away from lock-in business models. This points to a general shift towards software-as-a-service based models for language service providers and is the clearest sign yet of progress for an industry reinventing itself.

Case studies

The TAUS Localization Business Innovation White Paper (June 2008) noted companies, such as across, Clay Tablet, DocZone and XML-INT, from outside traditional industry boundaries providing Software-as-a-Service translation/localization workflow solutions. Highlighted here are vendors' Open Translation Platforms which were presented at the TAUS Forum in Edinburgh, March 2009.

- ✓ Asia Online provides SMT services in a pure Software-as-a-Service model, offering features such as data cleaning and preparation tools, on demand SMT engine development, support for both user created and online dictionaries and glossaries, ability to pool data for greater leverage, multiple level domain support and collaborative post-editing environment

- ✓ Eleka has developed an open source RBMT system from the Basque country, called Opentrad, now covering the languages Catalan, Galician, Spanish, French, Basque, based on two different but coordinated designs

- ✓ GlobalSight open source project was launched in January. Integration with Content Management Systems such as Teamsite and Documentum has been completed. Next up are integration of business intelligence features and machine translation

- ✓ Lingotek has developed a platform for collaborative translation, aiming to introduce social networking dynamics, such as group voting and review and genuine sharing of translation resources, into the business of translation

- ✓ The R&D team at Moravia is designing the open translation architecture based on Tiny TM (the open source TM system originated by Frank Bergmann) and a new proposed standard for Computer Aided L10N Project Management System (CALPMS). This standard will allow Moravia to integrate with localization workflow tools, such as]project open[, LTC Worx, and Plunet

- ✓ Lionbridge will open Logoport and Freeway to other vendors, moving to a Software-as-a-Service model and ensuring greater connectivity of its technology with the outside world

- ✓ SDL will connect all its translation tools through open APIs and enabling Translators using Trados to easily connect to TDA and benefit from industry-wide sharing of TMs

- ✓ Translated.net offers industry-shared web-based TM, integration of MT (open source Moses SMT among others), a community-translation model, and the APIs are published

7.   Localization and support convergence

For a growing number of companies localization now encompasses FAQs, support articles, knowledge bases and user-generated content. The content to be translated grows organically as users share experiences, exchanging views on likes and dislikes. Solutions found by support staff in one language are often useful to colleagues who speak other languages. The creation process and positioning of these types of content are at odds with the traditional project-based localization model.

Some organizations have already amalgamated support and localization departments. Autodesk, Cisco and Ebay began trying to localize support content with the help of MT as far back as 2003-2004, initially with disappointing results; they all started retrying within a year.

Karen Combe of PTC echoes the sentiments of many executives, "If MT could be implemented, either hosted or self service, the technology could be used by our customer support team also, but we need to see clear benefits."

Most consumer market companies now trialing MT confirmed that should the output provide enough of a gist, they would use MT as a tool for localization-on-demand for support content. Companies such as Knowledge Accelerators and Speaklike already offer translation solutions tailored to meet customer support needs. A handful of companies are also using professional translators on support content, whilst building up resources for MT engines in parallel.

A huge opportunity

Human translation, whether by professionals or a community, is only workable in a limited number of cases. Improved MT offers the potential to tap into huge opportunity to tap into localization for Knowledge-Centered Support.

TAUS is working with the Consortium for Service Innovation to create a forum for the sharing of ideas and experiences on the convergence o f localization and support. Future case studies and articles will report on developments in this evolving space.

8.   Continuous translation

Companies using iterative/agile development approaches are the front runners in this change scenario. In the past the main examples included Autodesk, Oracle, SAP and Symantec. They have since been joined by a number of other companies that we spoke to during our research. Quicker turn around with translation already taking place during development is helping to generate revenue, reduce costs and improve quality. The more automated processes the larger the cost savings over time.

Just over a third of clients and a fifth of suppliers from our survey are aiming to use/offer continuous translation as part of their mainstream translation business in the coming two years.

Implementing continuous translation requires a focus on architecture and process re-engineering. Takatoshi Adachi of ca is very clear about the pitfalls;

 "The biggest issue is integration of tools. There isn't a set of tools out there that connect easily. We relied on lots of customization by our engineering team to make it happen."

The jury is still out on whether the traditional pay per word pricing model will be replaced by hourly rates or another approach. The booking of time blocks for expected busy periods is becoming more popular. Counting words will remain a key productivity metric.

Challenges
There needs to be a strong vision and commitment to building a robust process, such as with the unique approach at Oracle. A number of those we spoke to highlighted that the initial cost of customization to connect workflow, authoring, and other systems is too high. Scalability, in terms of user licenses and database, needs to be considered from the outset.

The road ahead
The next steps include more open solutions that enable connectivity and portability without heavy customization. For many companies already working in continuous translation mode, integration of MT as well as further efficiencies from automation are on the agenda.

Case studies

Symantec built an in house development program, cutting inefficiencies by automating whenever possible, and largely meeting targets to cut costs by 30% and half time taken. In the last three years, translation volumes have grown from 10 to 30 million words per annum. This comprised 80% product documentation and 20% marketing material, inter alia. Headcount has remained virtually flat. The introduction and implementation of MT has largely amortized the upfront technology investment. The only outsource cost is post-editing. The actual price of 'translation' halved in two and half years.

The Oracle Translation Factory is the centralized platform enabling localization across Oracle's complete product range. The platform is tightly integrated with development providing a fully automated continuous workflow. A centralized TM containing all of Oracle's translation assets across 35+ languages helps to ensure a very high reuse rate of translations and optimized turnaround times, ensuring thousands of simship releases a year. The TM is effectively a multilingual content management system; a live leveraging tool. In-line with Oracle's growth strategy, this scalable translation platform has allowed Oracle to integrate their new products and many acquisitions and reap the same benefits of cost reduction, improved time to market and high level of quality.

Sources: TAUS reports

# 9.	Personalization

In this context, personalization refers to automated translation that is focused on individual users or sub-groups requirements. Examples would include translation of web pages, chats, and social networking sites with very targeted use of terminology that is culturally attuned. This would entail going beyond domain trained MT engines to ones trained for very specific products families or niche groups. There would be myriad new opportunities for commerce, ranging from cost reductions to easier access to new markets. The gains for society would potentially be far more profound with people of similar interests being able to form close bonds and transcend language barriers for example.

A new agenda

 None of our interviewees expected this to be possible anytime soon. However, such personalization for better inter-cultural communication is taking place at the Language Grid initiative and on the agenda at the Centre for Next Generation Localisation. Two ingredients for making this happen on a bigger stage, industry-wide language data sharing and open translation platforms, are now reality. This potentially provides the data to fuel and the connectivity needed in order to collectively the leverage language data, creating a foundation for personalization on a wide range of subjects. The next step is finding an effective way for companies/ organizations/specific groups to work together to harmonize and streamline terminology. Success would mean hugely disruptive innovation and have a world changing impact.

10. Appendices

World Tour findings

TAUS conducted a series of round table meetings in Europe and USA from October 2008 to February 2009. In break-out sessions participants debated the opportunities, threats and barriers for innovation in the localization industry. In the tables below we list the arguments that received the highest scores.

Opportunities	%	Votes
Expanding need for translation – new opportunities – offering choice – specialization	27%	61
Business model around communities	18%	41
Offering/leveraging technology/MT to enhance translator's productivity	14%	31
Improve time-to-market	10%	24
Share translation memories – Global Memory – Increase capacity	9%	21

Threats	%	Votes
Loss of LSP role – failure of current business model – customers can bypass LSPs	13%	33
Loss of quality – language degradation	12%	30
Power of translator – resistance to change	9%	22
Fragmented competition – technology divergence – no standard	8%	19
Prices going down to zero	8%	19

Barriers	%	Votes
Changing the business model – LSPs resistance to change	19%	51
Attitude of translators – slow technology adoption	11%	30
High level of investments required – lack of ROI data for LSPs	8%	21
Lack of standards – insufficient standards in User-Generated-Content	6%	15
No control over source content	5%	14

Market survey results

The online market survey was conducted during March 2009.Two hundred and eleven business owners and decisions makers responded, comprising: 65 (31%) buyers, 129 suppliers (61%) and 17 (8%) consultants. Respondents came from automotive, healthcare, pharmaceutical, and technology sectors. Technology sector respondents were a large client majority. Respondents are based in North America, Europe, Asia and South America. Percentages add up to more than one hundred as respondents could select multiple options.

1. What is the effect of the economic slowdown on your translation business?

	User	Supplier	Other	Totals
A. None	15.4%	20.9%	23.5%	19.4%
B. Less volume of content translated	32.3%	**58.1%**	41.2%	48.8%
C. Less languages translated	26.2%	17.1%	11.8%	19.4%
D. More outsourcing	10.8%	7.8%	0.0%	8.1%
E. Increased automation	21.5%	15.5%	29.4%	18.5%
F. New opportunities for exploring innovation	**46.2%**	31.8%	29.4%	36.0%
G. Other	16.9%	9.3%	23.5%	12.8%

Other – most often refers to price pressure.

2. Which of the following technologies and/or innovations will your company apply in mainstream translation business in the coming two years?

	User	Supplier	Other	Totals
A. Machine translation	49.2%	**55.8%**	47.1%	**53.1%**
B. Community translation	20.0%	14.7%	29.4%	17.5%
C. Language data sharing	**50.8%**	50.4%	35.3%	49.3%
D. Continuous translation	32.3%	20.2%	23.5%	24.2%
E. Open translation platforms	14.8%	25.6%	17.6%	21.3%
F. None of these (no change)	6.2%	18.6%	23.5%	15.2%
G. Other	16.9%	7.8%	17.6%	11.4%

Other – most often refers to improving translation memories, management or authoring systems.

3. Most of the respondents to a previous TAUS survey agreed that the word-based pricing model in the translation industry is out of date and does not support the objectives of innovation, efficiency and quality service. Please indicate whether you agree or not, and share any suggestions if you like for new business models to be applied.

	User	Supplier	Other	Totals
A. We agree that the business model of word-based pricing is not very effective	**56.9%**	**45.7%**	41.2%	**48.8%**
B. The current word-based pricing and business model serves us well. no need to change	29.2%	**45.0%**	23.5%	38.4%
C. Suggestion for business model innovation	13.8%	9.3%	35.3%	12.8%

Most suggested a hourly model. A few people suggested moving to a model based on value of content or complexity of work, and others to service-based models.

4. Machine translation is being introduced as a useful technology in more and more translation environments and processes. What do you see as the main barriers to effective use of MT technology?

	User	Supplier	Other	Totals
A. Poor quality of MT output	**56.9%**	**60.6%**	47.1%	**58.3%**
B. Cost of MT licenses and implementation	35.4%	34.9%	23.5%	34.1%
C. Lack of post-editing resources	30.8%	33.3%	29.4%	32.2%
D. Complexity of customization	43.1%	48.1%	41.2%	46.0%
E. Other	24.6%	17.8%	23.5%	20.4%

Other – myriad reasons given, including : MT business models, poor quality source content, inability to support in- line tagging, lack of reliable quality assessments, translators unwillingness to do post-editing, and lack of language coverage .

5. Machine translation is being introduced as a useful technology in more and more translation environments and processes. What do you see as the main benefits of MT technology?

	User	Supplier	Other	Totals
A. Service differentiation, MT as a new service	26.2%	41.1%	29.4%	35.5%
B. Real-time translation to support end-users	47.7%	33.3%	29.4%	37.4%
C. Cost reduction and efficiency improvement in translation	**73.8%**	**68.3%**	**64.7%**	**69.7%**
D. Other	15.4%	10.1%	29.4%	13.3%

Other - translating more content and faster.

6. Language data sharing (i.e. translation memories and terminology) is being introduced as a good practice that stimulates interoperability and innovation in the translation industry. What do you see as the main barriers to sharing your translation memories with all stakeholders in the global translation industry?

	User	Supplier	Other	Totals
A. No clear business value	**44.6%**	29.5%	23.5%	33.6%
B. Objections from main stakeholders	32.3%	38.0%	52.9%	37.4%
C. Ownership of translation memories is unclear	**43.1%**	**63.8%**	35.3%	**55.0%**
D. Other	36.9%	20.2%	35.3%	26.5%

Other – well managed TMs are a competitive advantage, inconsistency in quality of TMs, uniqueness of own TM, differences in term use in same domain, and risk of deteriorating quality.

7. Language data sharing (i.e. translation memories and terminology) is being introduced as a good practice that stimulates interoperability and innovation in the translation industry. What do you see as the main benefits from sharing your translation memories with all stakeholders in the global translation industry?

	User	Supplier	Other	Totals
A. Reduction of costs and friction in translation memory management overhead	40.0%	41.9%	41.2%	41.2%
B. Opportunities for increased translation automation (MT and advanced leveraging)	56.9%	**55.0%**	**64.7%**	56.4%
C. Create a foundation for innovation in translation business and development of new services	26.2%	27.9%	29.4%	27.5%
D. Streamline and unify terminology in vertical industry	**64.6%**	50.4%	58.8%	**55.5%**
E. Increase access to new markets and new customers worldwide	21.5%	26.4%	47.1%	26.5%
F. Other	7.7%	7.8%	17.6%	8.5%

Other – no benefits seen.

8. Times of economic slowdown urges us to rethink the business in general and create new ideas. New services may be developed that will lead to new successful business. Is your company looking at developing new services, such as:

	User	Supplier	Other	Totals
A. Post-editing MT services	20.0%	**67.4%**	29.4%	**49.8%**
B. Language data hosting services	13.8%	27.1%	17.6%	22.3%
C. Translation data cleaning and management services	27.7%	37.2%	23.5%	33.2%
D. Other	**49.2%**	20.9%	70.6%	33.6%

Other – mainly clients for whom this question is not relevant. A few people suggested consultancy and services.

9. Do you already deploy machine translation in your organization?				
	User	**Supplier**	**Other**	**Totals**
A. Yes	27.7%	40.3%	41.2%	36.5%
B. No	**72.3%**	**59.7%**	58.8%	**63.5%**

10. Do you expect to be using MT in:				
	User	**Supplier**	**Other**	**Totals**
A. one year	**42.6%**	**45.0%**	46.7%	**44.4%**
B. 2 years	27.7%	41.3%	40.0%	37.4%
C. Never	29.8%	13.8%	13.3%	18.1%