Translation Technology: a short history and a bit of future

Massimo Ghislandi
September 2014
The concept of machine translation pre-dates translation memory.

Progress made in the 50s and 60s with Academia and Government set it up research and project to produce machine translation.
Early concepts of Translation Memory are started in the 70s. Martin Kay wrote a paper on the concept.

Brigham Young University also contributed to the concept of Translation Memory and ability to check previous translations or repetitions.
Globalization

**Thinking inside the box**

**World merchandise trade**
2012 prices*, $trn

---

**International adoption of containers**

---

<table>
<thead>
<tr>
<th>Ports worldwide</th>
<th>1965</th>
<th>1970</th>
</tr>
</thead>
<tbody>
<tr>
<td>Port labour productivity, tonnes per hour</td>
<td>1.7</td>
<td>30.0</td>
</tr>
<tr>
<td>Average ship size, tonnes</td>
<td>8.4</td>
<td>19.7</td>
</tr>
<tr>
<td>Number of loading ports in Europe</td>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td>Insurance costs†, £ per tonne</td>
<td>0.24</td>
<td>0.04</td>
</tr>
<tr>
<td>Value of goods in transit‡, £ per tonne</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

---

*Deflated by US consumer prices
†Australia to Europe  ‡Hamburg to Sydney

Sources: World Trade Organisation; US Bureau of Labour Statistics; Daniel Bernhofen et al; The Economist
Translation memory technology research grows.

IBM develops a translation memory solution

Trados is formed and develops the first commercially available translation software
The launch of Trados

1984 saw the launch of TRADOS GmbH (TRAnslatIOn & DOccumentatIon Software) as a Language Service Provider by Jochen Hummel and Iko Knyphausen in Stuttgart, Germany to provide translation services for IBM.
while globalization boomed
First version of MultiTerm

1990 brought the first version of one of TRADOS's main applications MultiTerm. MultiTerm was initially created as a terminology management tool for DOS.
The value of terminology

- A properly prepared glossary can
  - save translators as much as 20% of the translation time
  - reduce queries for project managers by 60%
- Or,
  - 5% for translation
  - 9% reduction in labour
  - 50% reduction in translation efforts

Can you build a ROI measurement?
Silvia Cerella Bauer, Terminology Expert and owner of CB Multilingual, offers the following statistics for measuring the potential benefits of using a Terminology solution:

- 3,000 members of staff, only 5% make regular queries (150 people)
- Each one searches the termbase 20 times per month (3k p.m. or 36k p.a.)
- Assume only 66% are successful (23,760 successful queries)
- Assume one search saves 9 minutes of lost productivity. (3,564 hours saved)
- If employees cost an average of €50 per hour:
  - €178,200 saved
    - Increased productivity and more time to do other tasks
The first Translator’s Workbench

1994 saw the first Translator’s Workbench developed to run on Microsoft Windows. TRADOS also sold 200 licenses to the European Commission and a further 200 licenses to Berlitz.

June of 1994 brought the release of MultiTerm Professional 1.5 which included fuzzy matches even when words were incorrectly spelt, a dictionary style interface, faster searches through use of new highly compressed data algorithms, drag and drop content into word processor and integrated programming language to create powerful layouts.

During the same year TRADOS welcomed a computational linguist from the University of Stuttgart, Matthias Heyn, to the organization. Matthias, the architect of the first alignment tool on the market T Align (later known as WinAlign), would eventually become the third partner in TRADOS.
The value of TM

10% Repeated words
Dell Latitude E6330
Owner’s Manual

The value of TM

68% Repeated words
1997 became a milestone year when Microsoft decided to base its internal localization memory store on TRADOS. Microsoft also consequently acquired a share of 20% in TRADOS.

WinAlign was also released as a visual text alignment tool as the first fully-fledged 32-bit application in TRADOS.
Meanwhile... shifts in regional importance

Changes in institutional frameworks allowed trade to become the engine of globalisation, with world trade in manufactured goods increasing more than 100 times in the 50 years since 1955, much faster than the overall growth of the world economy.
2001

**Trados 5**

March **2001** saw TRADOS announcing the launch of TRADOS 5 which provided the next generation in multilingual technology solutions.

2002

**Trados launch free online service**

In April **2002** TRADOS launched Translationzone.com. The new website was launched as a free online community designed to serve the needs of translators, their customers and translation agencies.
Worldwide product trade
Digital Explosion begins
Digital Explosion and language

- China: 458,262
- United States: 199,099
- Europe: 429,611
- Latin America (exc. Brazil): 104,823
- India: 81,792
- Japan: 73,656
- Brazil: 70,220
- Russian Federation: 66,955
- Southeast Asia: 66,224
- Australia: 15,913
- Rest of World: 200,500
Trados 6

TRADOS 6 was launched in April 2003, with TRADOS 6.5 following in October of the same year. As well as expanding its product portfolio, TRADOS was growing in size with 186 members of staff in 2003.

Trados TeamWorks

TRADOS launched TeamWorks, a localization process management solution to transform the way localization and translation teams worked. TRADOS TeamWorks was an industry-first solution to consider and optimize the localization process in its entirety, from creation to delivery, unlike point-solutions that only assist with specific tasks.
Translation Leverage

Translation leverage: The percentage of content in new or revised Web pages for which a firm already has a translation in its translation memory.

![Translation Leverage Diagram]

- **Annual translation cost without translation memory**
- **Annual translation cost with translation memory**
- **Annual savings with translation memory**

**Year one:**
- 3 M words to translate
- 40% translation leverage
  - $360K
  - $144K
  - $216K

**Year two:**
- 4.25 M words to translate
- 50% translation leverage
  - $510K
  - $255K

**Year three:**
- 5.50 M words to translate
- 60% translation leverage
  - $660K
  - $396K
  - $264K

**Note:** Translation costs at $0.12 per word

**Source:** Forrester Research, Inc. (“Effective Translation Management Saves Time, Reduces Cost” July 2005)
Common Sense Advisory estimated that the market for outsourced language services was US$ 8.8 billion worldwide in 2005

Predicted growth for US$ 12 billion worldwide in 2010

Trados is acquired by SDL

In 2005 TRADOS was acquired by SDL plc for £35 million. The acquisition provided extensive end-to-end technology and service solutions for global information assets. The successful integration of SDL and TRADOS created a clear market leader in Global Information Management Solutions.
Digital Content

- 13ZB (+40% Y/Y)
- >4ZB (+50% Y/Y)

Zetabytes (ZB)

SDL Trados 2006 launched

2006 proved another milestone year for SDL TRADOS. February saw the launch of SDL Trados 2006, bringing together SDL and TRADOS technology in one unified product for the first time.

SDL Trados 2007 launched

SDL Trados 2007 was launched in Spring 2007, providing many new features to increase efficiency and quality throughout the translation supply chain.

A year to celebrate!

SDL Trados celebrates 25 years of translation memory technology.

2009 saw the release of SDL Trados Studio 2009, not only the best of Trados and SDLX, but the arrival of the next generation in translation memory which was set to revolutionize the way people work.
**SDL OpenExchange launched**

**2010** saw the launch of SDL OpenExchange, a unique open industry platform enabling 3rd party developers to build and market apps and plug-ins for users of products such as SDL Trados Studio.
The journey to evolution continues!

SDL Trados Studio 2011 is released. Based on the revolutionary 2009 version, this new release transforms how people review, as well as bringing enhancements in terms of increased leveraging, faster performance, increased productivity and hundreds of refinements to bring users and ultimate experience.

SDL Passolo 2011

A new version of SDL Passolo, the software for localization experts, is released, bringing team collaboration to the next level.
2014

SDL Trados Studio 2014

SDL Trados Studio 2014 hits the market and changes the way translators see their translation environment. Taking the hugely powerful Studio platform into the simple and intuitive user environment strengthened SDL’s role in providing the language industry with the very best translation technology.

SDL Studio GroupShare 2014

SDL Studio GroupShare 2014 is launched offering localization professionals enhanced project and asset sharing technology. Providing easy integration with SDL Trados Studio and SDL MultiTerm, GroupShare 2014 reduces daily administration by automating many manual tasks, reduces errors and maximizes resource sharing.
Translation is falling behind

One zettabyte ($2^{70}$) of digitized information exists today

2.5 quintillion bytes come into being every day

0.00000067% of the digitized information created every day is Translated
The cost of not translating

It is estimated that 11% of exporting European SMEs (945,000 companies) may be losing business because of identified communication barriers.
### Table 16: Post-Edited Machine Translation Is Part of Most LSP Futures

Source: Common Sense Advisory, Inc.

<table>
<thead>
<tr>
<th>“Do you offer post-editing for machine translation?” (PEMT)</th>
<th>Percent of Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Freelancer</td>
</tr>
<tr>
<td>Yes</td>
<td>21.75%</td>
</tr>
<tr>
<td>No</td>
<td>62.81%</td>
</tr>
<tr>
<td>No, but plan to in the near future</td>
<td>15.44%</td>
</tr>
</tbody>
</table>
Consumer Acceptance of MT

75% of people use free MT tools
Web users take advantage of free translation tools due to the greater accessibility and integration of Machine Translation solutions

93% use Machine Translation to understand English
Less than 10% of non-native EN speakers get by without using MT software to translate websites they visit
Proven Productivity Gains

MT + PE = Faster than Full Human Translation

- **University Research**
  - English to German post-editing of authoring software documentation – **30%** faster than human translation
  - English to Spanish post-editing of software documentation – **25%** faster than human translation
  - English to French, Italian, German, Spanish of design and engineering software documentation – **74%** faster than human translation

  *Source: Post-Editing Machine Translated Text in A Commercial Setting: Observation and Statistical Analysis by Midori Tatsumi, MSc., School of Applied Language and Intercultural Studies - Dublin City University*

- **SDL Labs:**
  - Averages out at **30%**
Post-Editing Process

Ready
- Read source text, then MT output
- Determine usable elements

Set
- Build round MT output
- Do not under- or over-edit
- Focus on accuracy

Go
- Ensure all elements present
- Correct punctuation, grammar, terminology
Tips for Post-Editing Baseline Output

- Consider MT suggestions but check terminology compliance carefully
- Capitalize on the fluency which is an advantage of SMT output
- Identify verbs, expressions and other sentence parts that can be re-used
- Check that style is consistent
- Look out for generic translations of domain-specific terms that will need to be amended
Tips for Post-Editing Vertical Output

- Take advantage of style which is often appropriate for different projects within a domain
- Re-use technical terminology where appropriate (e.g., new project terms not found in the reference material)
- Perform a sanity check of your project specific terminology, as the vertical’s terminology might differ in some cases
Tips for Post-Editing Customized Output

1. Capitalize on terminology as it will be client-specific.
2. Take advantage of style as it will mimic the client-specific preferences.
3. You may need to focus on fluency and style when the syntax is more complex.
The Language Services Market is Growing

In 2014 it is worth $37.19bn
Globalization

Consumer behaviours

Digital explosion

Language

Revenue & Brand

41
What’s happening to the market?

○ Overall volumes continue to grow but **prices are down** (avg $0.13/word)
○ The market for translation technology is worth **$1,287m**
○ **PEMT** market is slowly growing – currently at about 3% of the LS market, worth of **$1,238m** ($841m last year).
○ Two technology trends: Machine translation and cloud-based software
○ Foreign exchange issues affect market sizing.
○ **17.45%** of LSPs sell Post Edited MT
  – (although the number is probably higher but LSPs do not identify this as a different service from Translation)
Globalization continues
Digital content will continue to boom

Opportunity for Big Data

- Digital Universe
- Useful if Tagged and Analyzed

Translation supply chain
Productivity, efficiency