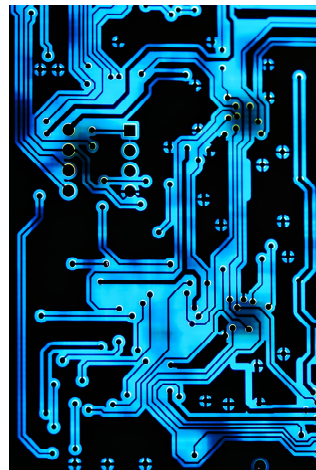# Top 10 Machine Translation FAQs*

In 2009, at the Localization World conference, Microsoft launched the challenge: if you're not using machine translation, why not? Companies such as IBM, Intel, Symantec, Cisco, Autodesk, Adobe, Fortis Bank and Siemens have already launched Machine Translation (MT) pilots, and have begun integrating MT into their workflows. 2010 promises to be the strongest year ever for MT as new programs and new processes prove it capable of delivering quality translations at an accelerated rate, for significant cost savings.

What is the buzz all about? Here are the top 10 questions companies are asking today:

1 What is machine translation?

Machine translation (MT) refers to a translation carried out by a software program. There are literally hundreds of MT programs, but they tend to fall into one of two camps: rules-based systems, which translate according to grammatical rules, and statistical systems, which base their translations on patterns derived from the application of algorithms to the analysis of vast quantities of similar texts. Systran is a Rules-Based system (RBMT) while Google's online translation server is Statistical (SMT).

Raw machine translation output, such as that obtained from Google, Babelfish or other online engines, is generally of low quality, and not to be confused with enterprise systems. MT that is trained on company- and domain-specific terminology results in superior quality. Lexcelera's **Customized MT** can be used to provide understandable quality for content such as online customer support material. Lexcelera's **Optimized MT** goes one step further for documents such as user manuals that require irreproachable quality: in this process, expert post-editors improve the raw output so that it is virtually indistinguishable from a fully human translation.

2 Why is everyone talking about MT all of a sudden?

Undeniably, the economic crisis is helping to fuel the surge of interest in MT as localization managers are being told to do more for less. However, non-budget-related factors are also driving the uptake of MT including the increasing importance of international customers to companies' top – and bottom – lines, the growing volumes of content to translate, the need for speed to reach global markets and the quest for greater process automation.

3 Will Machine Translation replace human translators?

Yes and No. Certain types of content, such as online knowledge bases, require only "understandable" quality to be successful. In these cases properly customized MT can be published as is, without human intervention. However, for the majority of cases where optimal

---

*Frequently Asked Questions

quality is required, MT functions as a productivity tool for human translators; in this case, post-editing the raw MT output is dramatically faster than translating from scratch.

4 What are the advantages of MT?

Enterprises who use machine translation typically obtain:

- Increased productivity,
- Faster time to market,
- Better terminological consistency and even translation quality,
- Significantly reduced costs.

5 What kind of savings should I expect from MT?

Optimized Machine Translation that is trained on client- and domain-specific terminology can result in 10 to 30% cost savings for fully human quality, and up to 95% for gisting quality. A software publisher with a well-established in-house MT system found that by combining process automation with MT:

- Productivity doubled
- Costs were cut by 30%
- Translation volumes (80% documentation) increased from 10 million words per year to 20 million without an increase in headcount
- The actual price of translation was halved in 2.5 years.

6 How do you train MT to give you the quality you need?

Training of Statistical Machine Translation (SMT) is an engineering activity and consists of feeding in millions of segments of bilingual text, glossaries, etc. so that the engine will learn by these examples. With Rules-Based MT (RBMT), where the system already masters the grammar of the source and target language, training is a specialized linguistic task including term extraction to build dictionaries, then coding and testing the result until the system is finely tuned.

7 You say MT can offer even better quality than fully human translation. How is this possible?

MT *improves quality* over existing human translations when there is:

- Expert training of the engine on company- and industry-specific terminology
- Source documents that are grammatically correct and use consistent terminology
- Automatic normalization of source
- Leveraging of existing translation memories (TMs)
- Trained human post-editors
- Virtuous circle of feedback to improve the engine in real-time
- Ongoing dictionary maintenance with new/modified terminology
- Automated post-production
- Updating of translation memories with improved content

8 I'm confused between TM and MT: what's the difference?

Translation Memory (TM) systems retain a copy of each phrase that has been translated, so that when new source material contains text that has been previously translated, the TM system pulls out from its memory material that has been translated. This may be a sentence that is exactly the same (a 100% match) or a sentence that is similar (a fuzzy match). Machine Translation (MT), on the other hand, generates a translation for entirely new source text.

The best process maximizes leveraging from existing translation memories – approved segments – and MT.

9   What MT program is best for me?

The best engine will depend on the language pairs you need, the type of documents and the workflow that MT needs to fit into as well as legacy content (TMs, bilingual and monolingual files). Lexcelera organizes pilots to determine which system will give the best results.

10   What kind of content works best for MT?

Machine Translation is best suited to large volumes of text with relatively unambiguous sentences such as:

- User manuals
- Procedures (SAP, etc.)
- Engineering reports
- Specifications
- Online help and documentation
- Knowledge bases
- Emails
- Streaming updates
- Customer support documents

Lexcelera, founded in Paris in 1986, was the first translation and localization provider in France to receive ISO 9001:2000 quality certification. Today Lexcelera (Eurotexte Group) has offices in London and Paris, and is a pioneer in Machine Translation. Lexcelera's services include consulting, training of MT engines, dictionary creation, updates and maintenance, technical support, user training, terminology management and linguistic post-editing.

For further information, or for a free, non-binding **Proof of Concept Pilot**, please contact:

Lori Thicke
General Manager
lori.thicke@lexcelera.com
+33 (0)1 55 28 88 09