



How translation professionals can make the choice on investing in an MT software or system solution

Jeff ALLEN

MultiLingual magazine - Editorial Advisory Board

LINGUIST-List - Industry Advisory panel

SAP BusinessObjects- Engineering Tools/Process Expert

<http://www.linkedin.com/in/jeffallen>

23 September 2010

Association of Translation Companies (ATC) Annual Conference 2010

London, UK



Outline

- Not here to discuss if MT is a possible option (but rather an appropriate option for your context)
- How can you make well-informed business decision on MT ?
- Who should you involve at language service provider company/ agency decisions about MT ?
- How much upfront investment (time, knowledge transfer, money) is needed?
- What can you expect on worthwhile results and return on investment ?
- What MT system types and solutions to consider? ("Which MT software/system is best for me ?")
- Links and case study references for MT
- Question/Answers



Factors to consider for using MT system/software

Choosing an MT system or software follows similar guidelines to qualifying a translation job mixed with buying CAT TM software

- ▣ Translation approach
- ▣ Documentation (or speech) type (procedures, descriptive, diagnostic troubleshooting, etc)
- ▣ Domain/field (medical, automotive, insurance, financial, etc)
- ▣ General text or very technical text
- ▣ User/audience type
- ▣ Number and types of language directions
- ▣ (Semi-)Automated translation system types
- ▣ Levels of human-machine interaction
- ▣ Means for capturing human knowledge and recycling it
- ▣ Mapping system features/modules to translation approach types to be used



Who should you involve in MT decisions ?

- **Marketing** – do they want to position the approach of the company as using MT
- **Sales and Pre-sales** – how do they want to position service and solution offer, and pricing
- **Account Managers** - how will implementing MT influence already regular customer and their needs and requests
- **Project Managers** – will the project managers be able to find the resource with skills on MT systems when specific skills are necessary for existing resource base
- **Production managers** – how does the implementation of MT combine and affect existing production processes and skilled people?
- **Finance** – how much do you/your company want to invest financially (cash, training, external freelance time, etc)
- **Support** – your customer feedback volume will increase. How do you plan to scale up support for this?



Machine Translation system types

- **EBMT (Example-based MT):** Translation Memory (TM) applications are EBMT-like. Several academic and government-funded EBMT research systems. Also Microsoft's Data-Driven MT. Moses SMT now has EBMT wrapper module.
- **RBMT (Rule-based MT):** majority of commercial MT systems
 - Non-customizable: nearly all vendors have link or box on website
 - Customizable: SYSTRAN, PROMT, Reverso, PowerTranslator, LEC, SDL Transcend, Compendium, LucySoftware, etc
- **S(B)MT (Statistical-based MT):**
 - Non-customizable: Google Translate, Microsoft Translator
 - Customizable: several opensource systems (Moses, Apertium), + some customizable commercial MT offers (Language Weaver, Asia Online, PangeaMT, Applied Language Solutions)
- **Hybrid MT :** different hybrid flavors (SYSTRAN V7, Asia Online, PROMT, LinguaSys, Applied Lang Solutions), depends on primary system. Previous govt funded & other research systems were called MEMT (Multi-engine MT)
- **KBMT (Knowledge-based MT):** semantic category approach in few customized industrial MT systems (CMU KANT) & few commercial systems (CIMOS, Translingo)



Can MT and Translation Memory (TM)/CAT tools be combined?

- **WordFast (TM)** : can link to Google Translate, SYSTRAN, PROMT
- **Alchemy (TM)** : can link to SYSTRAN, PROMT (but specific versions)
- **Trados (TM)** : PROMT has special P4T plug-in. SYSTRAN also claims to have plug-in.
- **STAR Transit (TM)**: SYSTRAN claims to have plug-in

- **SYSTRAN MT**: built-in TM memory
- **PROMT MT** : built-in Associated memory.
- **GT4T**: can mix Google with Word-based TM CAT tools.
- **AsiaOnline**: Analyzing TM data is a fundamental concept



How much investment effort / costs to plan for ?

● Cash/Credit

- ▶ very different levels if you choose an SMT system vs a pure RBMT system vs a Hybrid system.
- ▶ Licensing schemes have been shifting over past 2 years to make it feasible for LSPs to become customers

● Training of existing resources

- ▶ Linguistic-based RBMT systems require training to be effective. DO NOT try to just do it on your own.
- ▶ - No publicly advertised training courses by vendors.
- ▶ - ProZ has a new MT webinar series on MT
- ▶ - Plan for 1-day to 1-week + ongoing seminar training

● SMT approach has expertise included in Project costs

- ▶ Training from SMT perspective is the cycle of making the machine learn from the texts
- ▶ Still need to help posteditors understand their new approach

● Ramp-up productivity time

- ▶ No training: Some have said ROI took 18 months without training.
- ▶ Trained: Some LSPs stated productivity gains in 4 weeks -user training



Web sites with Practical info MT topics for LSPs

● Translatorscafe MT forum

- ▶ <http://www.translatorscafe.com/cafe/MegaBBS/forum-view.asp?forumid=44&catlock=8>

● Proz MT forum

- ▶ http://www.proz.com/forum/machine_translation_mt-844.html

● LinkedIn MT-related forums

- ▶ [Automated Language Translation Group](#) , [Localization Professional](#) , [Post Editing Service Level Group](#) , [Translation Automation](#) , [Bay Area Machine Translation Users Group](#) , [SYSTRAN User Group](#) , [PROMT User group](#)

● MultiLingual Magazine product reviews

- ▶ <http://www.multilingual.com/>

● Facebook MT-related groups

- ▶ [Reverso](#) , [SYSTRAN](#) , [Reverso.net](#) , [Promt Translator](#) , [Skype Translate](#) , [Common Sense Advisory, Inc](#) , [Localization World Conference](#) , [XTM International](#) , [XTRF Translation Management Systems](#) , [ProZ.com](#) ,

● Yahoo Groups

- ▶ http://tech.groups.yahoo.com/group/SYSTRAN_users/
- ▶ http://tech.groups.yahoo.com/group/PROMT_users/
- ▶ http://tech.groups.yahoo.com/group/Reverso_users/

● Allen Keys 2 Languages website (includes lang tech evaluation)

- ▶ <http://www.allenkeys2languages.org>

● Other sites

- ▶ Other threads and posts on [LANTRA-L](#) , [Xing](#) , [Viadeo](#) , etc



MT Case studies

SYSTRAN

- [2007\[1\]\[1\].12.Symantec.pdf](#)
- [2008\[1\]\[1\].01.AvonAndSomersetPolice.pdf](#)
- [2008\[1\]\[1\].02.LincolnshireCountyCouncil.pdf](#)
- [2008\[1\]\[1\].04_SYSTRAN_etude-de-cas-symantec.pdf](#)
- [2008\[1\]\[1\].06_SYSTRAN_case-study_EADS.pdf](#)
- [2008\[1\]\[1\].06_SYSTRAN_etude-de-cas_EADS.pdf](#)

PROMT

- [alain_champagne_case_study_english.pdf](#)
- [Patrick_Roye_freelance_translator_case_study_english.pdf](#)
- [MT-dictionary-building-casestudy-jeff-allen_v100.pdf](#)
- Conference paper by Jeff Allen at AMTA 2004

SDL

- [KbT-System-case-study-cnh-assist.pdf](#)



Questions/Answers