



CONVERGING METHODS IN MACHINE  
TRANSLATION

---

MADISON, OCTOBER 13, 2008

ROUND TABLE AT LOCALIZATION WORLD

CONVERGING METHODS IN  
MACHINE TRANSLATION

MADISON, OCTOBER 13, 2008

---

### **SUMMARY**

---

Half a century of history in machine translation is known for its debates and disputes about the most effective way of cracking the problem of automatic language translation. It is only since the technology is finding its way into more and more day-to-day applications that the academic divides seem to be fading away. Different from researchers, users are merely interested in what works and what does not work in the most practical sense. Dogmas are replaced by pragmatics. Statistical, example-based or rule-based ... the latest generation of machine translation engines combine different flavors, and they start to perform better.

This TAUS Translation Automation Round Table will feature different approaches to machine translation with guest speakers from some of the leading MT companies: Systran, language Weaver, ProMT and Asia Online. Participants in this round table will learn about the opportunities and barriers to using MT in localization for user manuals, service literature, knowledge bases and other applications. The round table includes hands-on exercises with different MT engines and benchmarking of performance on productivity and quality scores.

Discussions will focus on some of the key questions for everyone who is considering to make machine translation part of its production infrastructure. What does it cost? How can I measure and benchmark the quality of output and the productivity of my post-editors? How can I best train and customize my engines? What does it cost me to maintain the system? Will there be more and better MT engines coming on the market? How difficult is it to develop an MT engine?

---

### **BENCHMARKING**

---

Participants in this TAUS round table are requested to bring along their laptop computer and a short text (100 – 200 words) in HTML or txt format for use in a post-editing and benchmarking exercise. Through this hands-on experience participants will learn how different approaches to MT perform on their own text material and they will learn about quality evaluation methodologies and benchmarking statistics.

---

### **TAKE-AWAYS**

---

Participants in this TAUS workshop will take away new insights and knowledge covering the following areas:

- User cases from TAUS members in machine translation.
- The latest developments in the leading MT companies.
- Customization and training of MT systems.
- Quality evaluation and benchmarking methodologies.

---

AGENDA: 9.00 – 17.00

---

Following is the agenda for this TAUS round table. The meeting starts at 9.00 and ends at 17.00.

- 1) Introductions (9.00 – 9.20). One minute introduction by each participant covering:
  - Company, department, place in organization.
  - Current level of translation automation
  - Plans and ideas for translation automation
  - Objectives for the meeting
- 2) Overview of Translation automation (9.20 – 10.00). Translation technology and the evolution of the industry, User cases from TAUS members, Technology roadmaps, Best practices. An introduction by Jaap van der Meer.
- 3) DIFFERENT APPROACHES TO MACHINE TRANSLATION AND QUALITY EVALUATION (10.00 – 10.30). Overview of machine translation systems and quality evaluation in a historical perspective and present use. By Heidi Depraetere.

**Coffee Break – 10.30 – 10.45**

- 4) Building a Continuous Evolution SMT Model (10.45 – 11.30). This session will describe an approach to SMT that integrates human post-editing and corrective feedback in real-time. It will also describe some tools that help users to understand how SMT engines make decisions and how to influence the engine's learning with small amounts of data. This analyze-correct-relearn cycle can produce rapidly evolving MT engines that produce FAUT faster than any of the alternatives. By Kirti Vashee.
- 5) STANDARDS, TRANSPARENCY AND NEXT GENERATION OF MT (11.30 – 12.15). Alex Yanishevsky, Senior Integration Architect at Prompt Americas, will talk about Prompt philosophy and vision, the importance of supporting standards and overall portability of MT solutions, and finally share Prompt's take on adding SMT modules to Prompt RBMT engine. By Alex Yanishevsky.

**Lunch - 12.15 – 13.30**

- 6) Converging methods at systran. (13.30 – 14.15) SYSTRAN's hybrid machine technology includes integrated advanced post-edition tools which increase productivity. Combined with a training module leveraging monolingual and/or bilingual resources, users are able to incrementally shorten the time to generate high quality translations. We will demonstrate this process and discuss customer successes on productivity gains and quality scores. By Jean Sennelart.

- 7) SMT: Requirements for Success. (14.15 – 15.00) Because of the technology that drives SMT, some companies are better suited for this type of translation software than others. There are barriers to success, and several requirements to ensure a successful deployment. This presentation will uncover the requirements for success and the criteria that make one company better suited for SMT than others. It will also discuss the benefits that sharing data can have on the results from a SMT system and how post-editing adds long-term, significant value to a SMT deployment. By Hannah Grap.

**Coffee break - 15.00 – 15.30)**

- 8) hands-on exercise (15.30 – 16.15). Using a MT web service, participants will test several MT engines on their own texts, do post-editing and measure and benchmark productivity and quality.
- 9) Strategies to take away. (16.15 – 16.45) Group discussion on lessons learned. Moderated by Heidi Depraetere.
- 10) Closure (16.45 – 17.00) Evaluations and next steps.