

Creating the world's largest Translation Memory

Marco Trombetti
Translated
MT Summit XII
Ottawa, August 2009

<http://mymemory.translated.net>



About Translated

Multi Language Vendor

Localization in 80 languages, 9,000 customers

Mission

Removing administrative overhead: faster and cost-effective. Make translators more productive.

Focus

Web content and Information Technology

Value Proposition

Cost reduction through workflow automation

My wish-list for the perfect TM Platform

MyMemory

Millions of segments, still FAST



Web 2.0, Wiki-like crowd contributions



One-click upload/download TM



Adaptive Matching, A priori probability



Propagate changes real-time



Fully integrated Statistical MT



Open standard



The Solution



Free,
Collaborative
Web Translation
Memory

Barriers to contribution



Privacy and IP



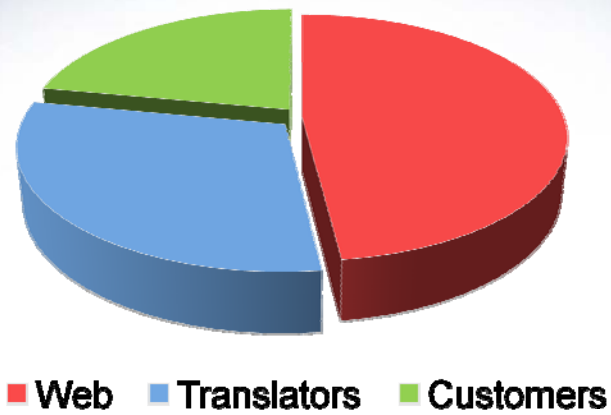
Time vs Value

Will people contribute?

Yes! 30 words per second for 1 year

1 Billion words

Sources



Benefits

LSP & Clients

- ♥ More control
- ♥ Free contribution from pro & crowd
- ♥ Quick reference and QA

Translators

- ♥ Quick relevant reference
- ♥ Improved organization and usability
- ♥ Free memories

Technology Challenges

- ✔ Scalable Real Time Indexing
- ✔ Adaptive Ranking
- ✔ Cost-Effective Scalability

Generating the interest



Pre-Populate



Make contributions easy



Make it free

Protecting IP

- ✔ **Private Memories**
- ✔ **Hiding Proper Nouns**
- ✔ **One-click report abuse**

Quality

- ✔ **Feedback from the crowd**
- ✔ **QA Correlation**
- ✔ **Automatic QA**

Is 1 Billion words is enough?

No. HT industry produces

**Tens of billions of
words per year**

Next Step

More Data & Technology

- ✔ More data extraction from the web
- ✔ Real Time SMT training
- ✔ Open the data for free to MT developers

QA crowd platform

- ✔ Fast and Cost-effective MT Evaluation
- ✔ Parallel Corpus evaluation and clean-up

Thank you!



Questions?

Marco Trombetti | marco@translated.net

More info: <http://mymemory.translated.net/doc/>