(intel)

The Triple Advantage Factor of Machine Translation: Cost, Time-to-Market and FAUT

Will Burgett, PMP®
Localization Program Manager
Intel Corporation
Global Web Operations, GILT

Julie Chang
Localization Project Manager
Intel Corporation
ITFlex Global Language Solutions

Public

Legal Notices

THIS PRESENTATION CONTAINS THE GENERAL INSIGHTS AND OPINIONS OF INTEL CORPORATION (INTEL). THE INFORMATION IN THIS PRESENTATION IS PROVIDED FOR INFORMATION ONLY AND IS NOT TO BE RELIED UPON FOR ANY OTHER PURPOSE THAN EDUCATIONAL. USE AT YOUR OWN RISK! INTEL MAKES NO REPRESENTATIONS OR WARRANTIES REGARDING THE ACCURACY OR COMPLETENESS OF THE INFORMATION IN THIS PRESENTATION. INTEL ACCEPTS NO DUTY TO UPDATE THIS PRESENTATION BASED ON MORE CURRENT INFORMATION. INTEL IS NOT LIABLE FOR ANY DAMAGES, DIRECT OR INDIRECT, CONSEQUENTIAL OR OTHERWISE, THAT MAY ARISE, DIRECTLY OR INDIRECTLY, FROM THE USE OR MISUSE OF THE INFORMATION IN THIS PRESENTATION.

(intel)

The Changing Face of MT

From this….

To this….

Good Morning!

Aloha kakahiaka!

Fully Automatic High Quality Translation (FAHQT)
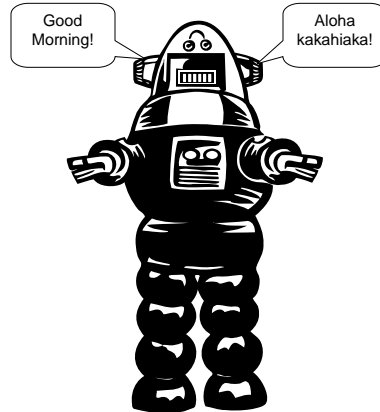
Fully Automatic Useful Translation (FAUT)

The Triple-Advantage Opportunity of Machine Translation for Intel Customer Support

Cost/Resource Schedule/Time

Scope/Quality

- Cost
  - Reached maximum cost efficiencies of human translation
  - Reached maximum budget levels for human translation of customer support content

- Time (Time-to-Market)
  - Takes 10 business days for complete translation cycle even with workflow automation
  - Source and target always 10 business days out of sync

- Quality / Scope (Fully Automatic Useful Translation = FAUT)
  - MT quality has gotten better
  - Users more willing to accept language imperfections in exchange for <u>useful</u> information

Intel's Motivation to Implement MT

Problem

Huge volume of content broadcast to sell products, solve problems, assist customers, avoid legal liabilities, and meet government regulations.

• **In general, 93% <u>will never be translated</u> due to cost**

• **Goal of self-help defeated when customers can't understand English**

• **Many emerging markets English intolerant and can't use English Web sites**

• **Translating customer support knowledge base is an open-ended activity**
  – **No single person has an overview of the entire content and its quality**
  – **Uncertainty of what content needs to be translated**

Opportunity

Use raw and post-edited machine translation to:

• Significantly reduce localization cost
  – From to 30% - 90% reduction (post edited to raw)*

• Dramatically increase translated content customers can access
  – From 10% to 100%*

• Empower customers to find the help they need in their language
  – Enable self help in the locales
  – Increase call deflection

• Deliver fast TTM for localized content
  – From two weeks down to 24 hours for support

• Expand markets:
  – Move funds from MT sites to other languages

* Based on production studies and analysis at Intel Corporation, 2008

(intel)

Business Transformations and Holistic ROI Factors

❑ Maximize integrated content and localization automation

❑ Dramatically reduce localization cost through cost avoidance, reduction in cost per unit, call deflection.

❑ Empower the customer, enable language self help, increase satisfaction

❑ Speed information to customers in their language
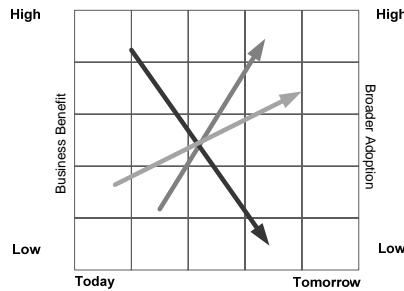
* Based on production studies and analysis at Intel Corporation, 2008

Benefit	Objective	ROI Factor
Cost reduction	Decrease localization cost up to 30% (post edited), 90% (raw, real time) and increase call deflection 10% for customer support. (Net Present Value)*	**HARD**
Time to market	Deliver content to local customers faster: 30% post edit to 100% (raw)*	**SOFT**
Increase productivity	Increase translation volume per head	**HARD**
Empower customers	Enable self help model fully in other languages.	**SOFT**
Dramatically increase localized content	Select 5-20% high value content to localize; MT does 80-95%	**HARD**
Customer satisfaction	Increase geo customer satisfaction on localized sites	**SOFT**
Grow the market	Increase revenue, especially in emerging markets	**HARD**

(intel)

Pyramid of Dependence

Quality Drivers

Cost Savings

- Machine Translation
- Translation Memory
- Human Translation
- Content
- Terminology
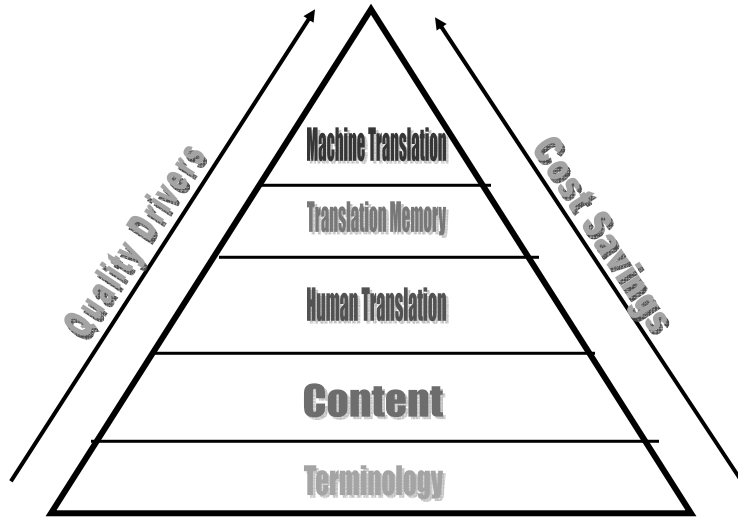
(intel)

MT as Key Part of Global Information Management and Publication

Machine Translation (Raw)

Customer Support

MT + Post Edit

Source

Translation Memory

Automated Localization Workflow

Content Optimization Training → Content Development → Content QA Tools → Content Management System

Target

Terminolgy

Terminology Management
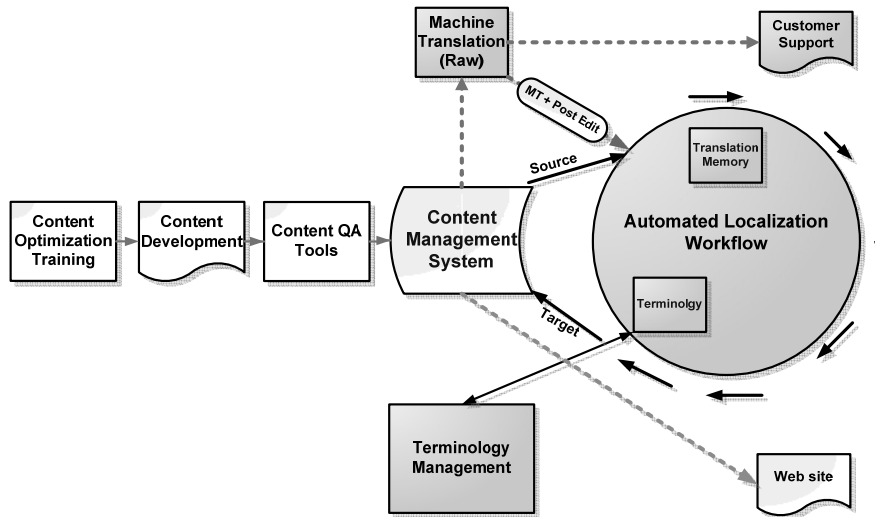
Web site

(intel)

Impact of Collaboration
TAUS Data Association (TDA): Language Data Sharing

Data providers (members)

Users (members & non-members)

Corporate users

Language Service Providers

Governments NGO's

Language data submission

Repository of language data

Language data pooling (members only)

Language search

Corporate users

Governments NGO's

Translation service providers

Writers
Support staff

Developers

Academics

Hosted data
Platform

9 October 23, 2008   Public

intel

Intel's Customer Support Web Site and MT

Content management System
(Maintain English content)

Local customer

Monitor and Compare

Baseline Update
- Detect if English source are added, updated, or removed
- Reset file status

English staging server

Baseline Database

- Pull files that are added/ updated

MT process in MT server
- Get the files that need to be processed
- Pre-processing before translation
- Translation
- Post-processing after translation
- Publish

Publish latest translated content

MT Server

Localized staging server

10 October 23, 2008   Public

intel

Spanish MT Page on Intel's Customer Support Site

Traducción automática | La presente es una combinación de traducciones efectuadas por personas y por sistemas automatizados. Se proporciona solamente como información general, y no deberá considerarse como completa o exacta.

(intel) Trabajar Divertirte Asistencia Acerca de Intel Cambiar de ubicación [Buscar]

Productos Tecnología Comunidades Descargas Integradores

Asistencia y descargas
Explorar por producto

Procesadores

Búsqueda
Asistencia y descargas

○ Toda la asistencia
◉ Esta categoría

¿Ha respondido a sus preguntas esta información?
○ Si ○ No

Procesadores
Los procesadores multi-core preguntas y respuestas

Dónde puedo obtener más información sobre dual-core y quad-core?
Encontrará más información sobre estas tecnologías aquí:

- Tecnología de doble núcleo
- Tecnología quad-core
- Tecnología multi-core

Por qué los gráficos de la CPU Windows® Administrador de tareas no idénticos?
No se trata inusuales si tiene un procesador con varios núcleos. Cada cuadro representa uno de los núcleos en el procesador. No habrá 2 gráficos para el uso de dos núcleos y 4 procesadores gráficos históricos para procesadores quad-core etc. Los gráficos se no idénticos debido a que muy pocos aplicaciones se divide su carga de procesamiento completamente forma uniforme a través de la núcleos. Programas más avanzados (por ejemplo: Juegos) son diseñados para aprovechar la tecnología de múltiples núcleos y hará un mejor trabajo de propagación la carga de procesamiento a través de la núcleos.

(intel)

Machine Translation Performance Indicators: Cost, Time, Scope

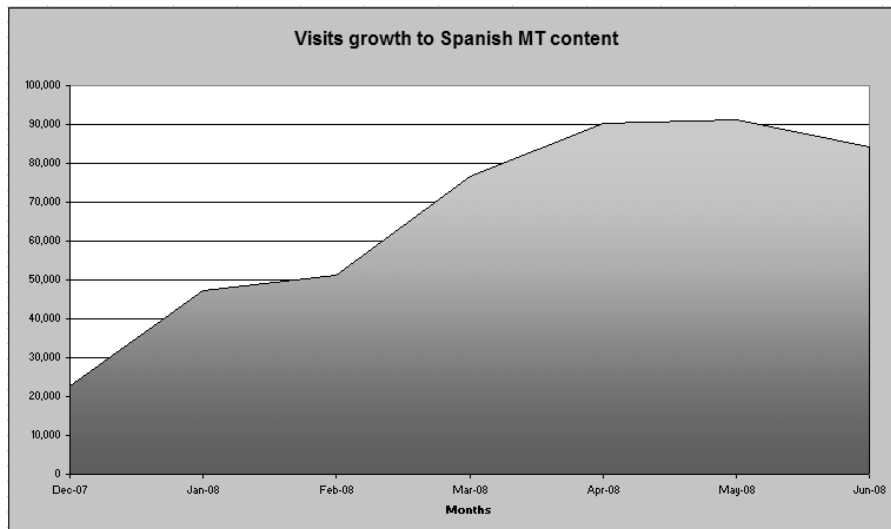| Indicator	Baseline	Objective	Actual
Cost / word	X	10% of X	5% of X
Number of Files	2,000	16,000	16,585 (Q1) 12,530 (Q2)
Number of Words	1,100,000	4,000,000	4,564,000 Q1 3,745,000 Q2
Translation cycle	10 business days	24 hours	24 hours
Call Deflection	XXXX cases	10% case reduction	In progress

(intel)

Visits to Spanish Customer Support Site Since MT

Visits growth to Spanish MT content

Machine Translation Quality

- **BLEU Score**
  - ENU >Spanish: 69.93
- **Customer satisfaction survey**
  - Did this information help answer your question?
- **Human Evaluations**
  - 5 independent QA evaluators are asked to rank MT quality from 10 knowledge base contents, approx 300 sentences.
  - Using scale 1-4 to evaluate the areas of MT function, performance, and translation accuracy.
  - Evaluating the same content after each MT re-training
  - Translation accuracy evaluation scale:

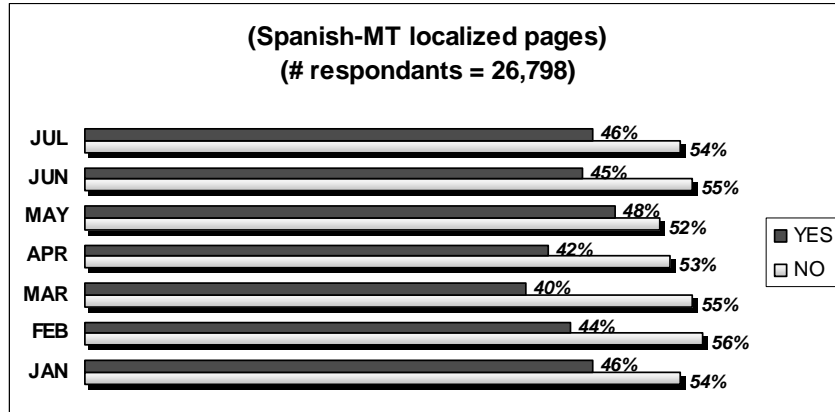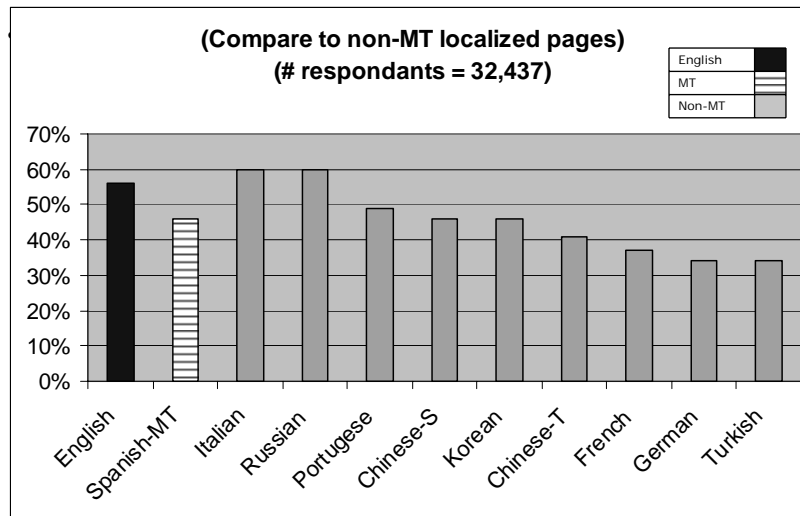|	
Excellent	Translation accuracy is 90% or above, all information is included
Good	Translation accuracy is 80%, content information is comprehensible
Fair	Translation accuracy is 60%, some information transferred accurately
Poor	Translation accuracy is below 60%, content is not comprehensible

Customer Survey – Raw MT Publishing

Did this information help answer your question?

(Spanish-MT localized pages)
(# respondants = 26,798)

Month	YES	NO
JUL	46%	54%
JUN	45%	55%
MAY	48%	52%
APR	42%	53%
MAR	40%	55%
FEB	44%	56%
JAN	46%	54%

Legend: ■ YES □ NO

(intel)

Customer Survey on Intel Customer Support Site

(Compare to non-MT localized pages)
(# respondants = 32,437)

Legend:
- English (black)
- MT (striped)
- Non-MT (gray)

Category	Value
English	56%
Spanish-MT	46%
Italian	60%
Russian	60%
Portugese	49%
Chinese-S	46%
Korean	46%
Chinese-T	41%
French	37%
German	34%
Turkish	34%

(intel)

Results of LAR Spanish MT system integrated into Intel Corporation's Customer Support Site

- Launched late Q4 '07

- Increased Spanish content by factor of 8X (~2000 files before, vs. ~16000 total files after): 100% Spanish site versus 12%

- Nightly static publishing schedule updates every 24 hours instead of 2-3 weeks
  - 800 to 1500 files change every two weeks
  - Updates processed through MT every 24 hours
  - MT dictionary refreshed every 6 months
  - Minor corrections made through MT backend

- Based on results and geo approval, stopped 95% of Spanish human translation starting in Q2 '08

- Remaining core content of human translation will keep both TM and MT finely tuned

- More languages coming in 2008

17 October 23, 2008   Public

Achieving the Triple-Advantage of Machine Translation

- Cost
  - Reduced cost per word by 95% using raw MT output

- Time-to-Market
  - Reduced customer-support translation cycle from 10 business days to 24 hours

- Fully Automatic Useful Translation (FAUT)
  - Answered customer's question: 44% MT versus average 47% human translation versus 56% English

18 October 23, 2008   Public

Public