



# SYSTRAN

## Automatic Translation Technologies for Enterprise Market and Application Overview

Pierre BERNASSAU  
Marketing Director  
SYSTRAN

*Enhance your communication in multiple languages*



**infor**nortics

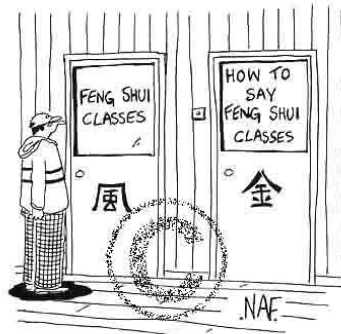
# Automatic Translation is popular



## Trois fois rien *Pessin*



74 LE MONDE 2 > 8 OCTOBRE 2005



The Indians had asked for fire water ... The translation error would prove costly to the trader.



WITH THE AID OF A TECHNICAL DICTIONARY COLIN WAS FINALLY ABLE TO MAKE SOME SENSE OF WHAT THE SALESMAN WAS SAYING



- ➔ Free online translations make popular automatic translation usages (Home, Education, Small Office, Enterprise)
- ➔ Average volume: 30 million translations per day (values based on SYSTRAN technologies only)
- ➔ Google, Yahoo!, Orange, Free, Voila, ...

# Automatic Translation needs are growing



Top Ten Languages Used in the Web ( Number of Internet Users by Language )					
TOP TEN LANGUAGES IN THE INTERNET	% of all Internet Users	Internet Users by Language	Internet Penetration by Language	Internet Growth for Language ( 2000 - 2006 )	World Population 2006 Estimate for the Language
<a href="#">English</a>	29.7 %	322,600,837	28.7 %	135.2 %	1,125,664,397
<a href="#">Chinese</a>	13.3 %	144,301,513	10.8 %	346.7 %	1,340,767,863
<a href="#">Japanese</a>	7.9 %	86,300,000	67.2 %	83.3 %	128,389,000
<a href="#">Spanish</a>	7.5 %	81,729,671	18.7 %	231.1 %	437,502,257
<a href="#">German</a>	5.4 %	58,854,682	61.3 %	113.2 %	95,982,043
<a href="#">French</a>	4.6 %	49,660,498	13.0 %	307.1 %	381,193,149
<a href="#">Portuguese</a>	3.1 %	34,064,760	14.8 %	349.6 %	230,846,275
<a href="#">Korean</a>	3.1 %	32,372,000	45.8 %	78.0 %	73,945,860
<a href="#">Italian</a>	2.7 %	28,870,000	48.8 %	118.7 %	59,115,261
<a href="#">Russian</a>	2.2 %	23,700,000	16.5 %	664.5 %	143,682,757
<b>TOP TEN LANGUAGES</b>	<b>79.5 %</b>	<b>863,981,961</b>	<b>21.5 %</b>	<b>166.7 %</b>	<b>4,017,088,863</b>
Rest of World Languages	20.5 %	222,268,942	9.0 %	500.0 %	2,482,608,197
<b>WORLD TOTAL</b>	<b>100.0 %</b>	<b>1,086,250,903</b>	<b>16.7 %</b>	<b>200.9 %</b>	<b>6,499,697,060</b>



- ⇒ Dozen languages used on Internet, **English is regressing relatively** - strong local content production via blogs, newsgroups and collaborative sites
- ⇒ e-Commerce market has strong increase – direct contact, no boundary – generates translation needs
- ⇒ Internet penetration per language shows diversities and future trends
- ⇒ Content volume grows rapidly: a small portion is localized
- ⇒ Localization process (write up, translation, revision, approbation) is not adapted to Web world which requires a short time to market





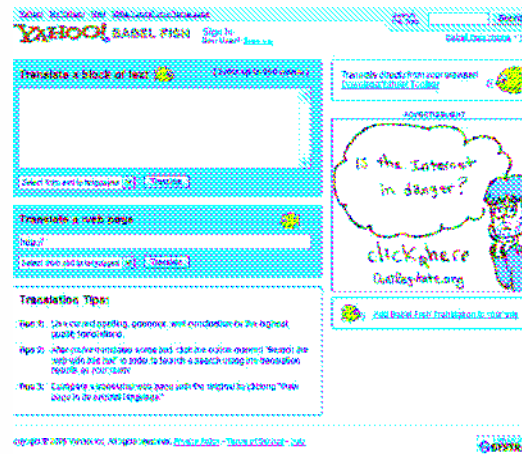
## ⇒ 2 misunderstandings to avoid

### ⇒ Human Translation ≠ Automatic Translation

- The technologies have made great strides but are still far from achieving the same level as professional translation agencies

### ⇒ Professional Automatic Translation Software ≠ Free Online Translators

- Translations done over the Internet are stripped of their context, whereas translations done within a company are based on large sets of texts (ie, translation memories) and dictionaries





## ⇒ Automatic Translation - 2 primary categories of needs

- ⇒ To understand written text in foreign languages
- ⇒ To publish new information in multiple languages

## ⇒ An added value for multiple enterprise applications

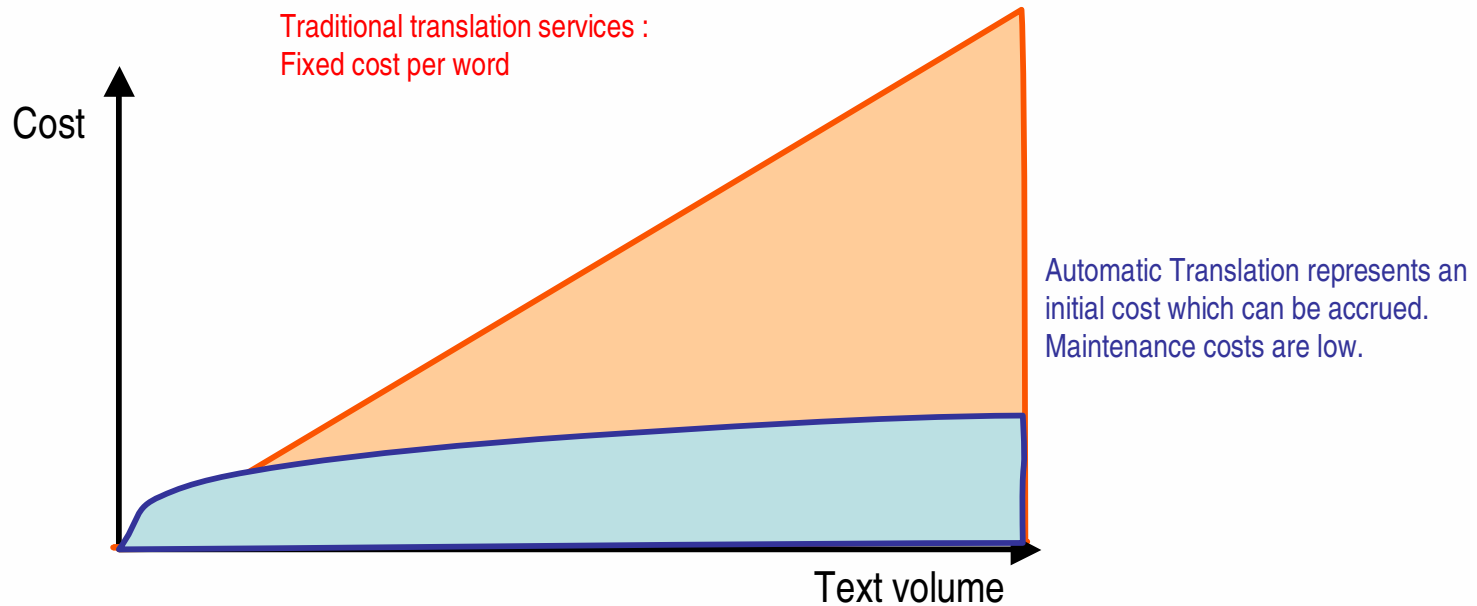
### ⇒ Inbound/inward type communication

- Self-service Translator over the Intranet (Office, Email, Chat, Collaboration)
- Market Intelligence based on foreign language content sources
- Cross-lingual research and translation on knowledge bases

### ⇒ Outbound/outward type communication

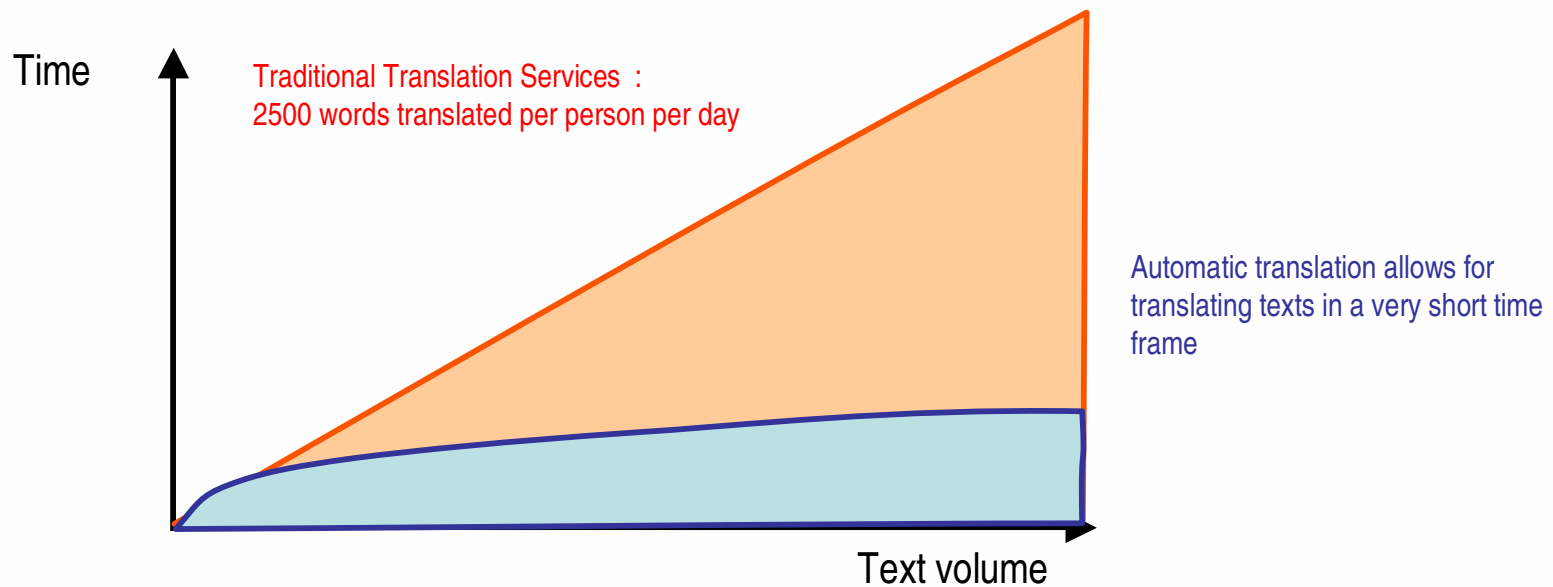
- Web Content Management, Web site localization
- Multilingual production and publication for catalogs and technical documentation
- Multilingual access to product information and technical support knowledge

- ⇒ **Translation costs are independent with volume of text to translate**
  - ⇒ For large knowledge databases and archives
  - ⇒ To handle multiple target languages
  - ⇒ To transform operational costs into investments

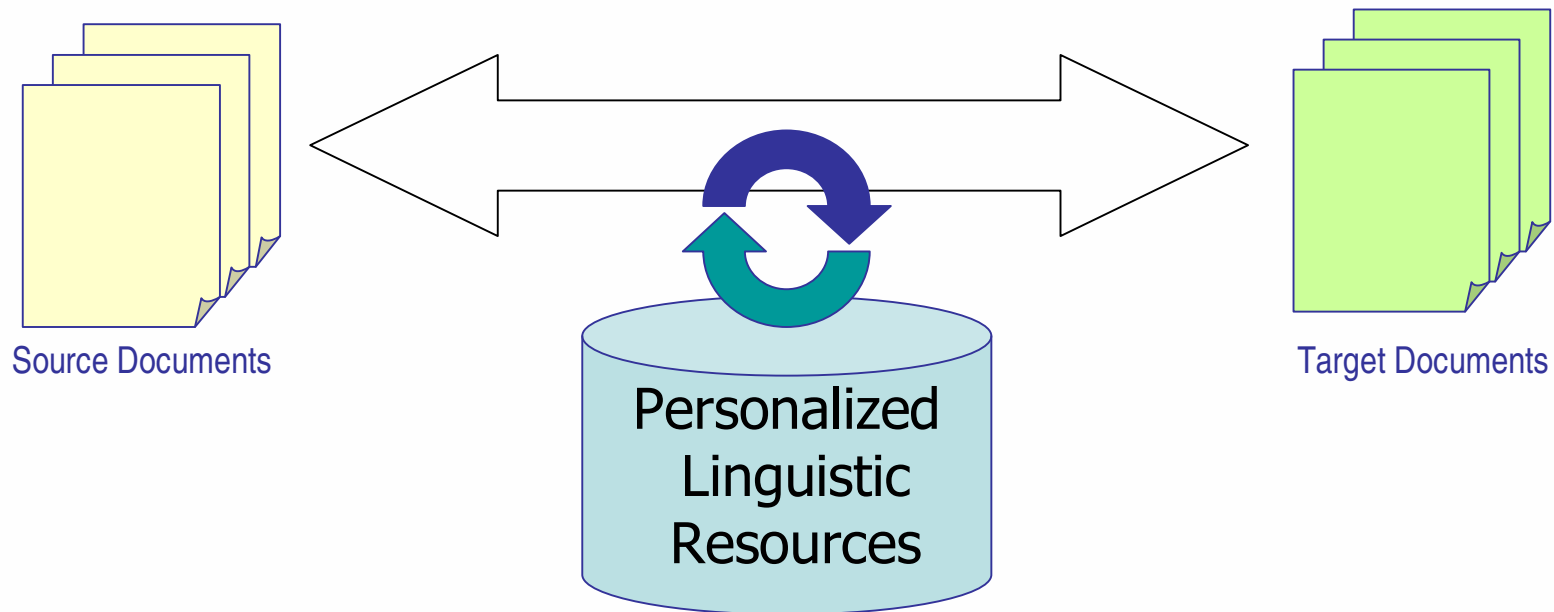


## → Translation speed is non proportional to the volume of text

- For Internet search portals
- For translating dynamic Web pages
- For quickly producing translations of large documents



- ⇒ **Linguistic personalized resources are independent of text**
  - ⇒ Terminology database to be reused for new documents / pages
  - ⇒ Building an enterprise terminology repository
  - ⇒ Multiples usages (inbound/outbound communication)







⇒ We do not always need a perfect translation !

⇒ Utility criterion : is the translated text understandable to the readers ?

⇒ Accuracy criterion : Does the translated text contain the appropriate words?

## 3 main levels of quality

### Publication Quality

#### Professional Translators:

An high level quality for published documents

### Revision Quality

#### Automatic translation with personalized linguistic resources:

Sufficient quality for internal documents or for revisable documents which can be post-edited by experts to produce higher-quality publications

### Browsing Quality

#### Automatic translation with off-the-self industry-specific dictionaries:

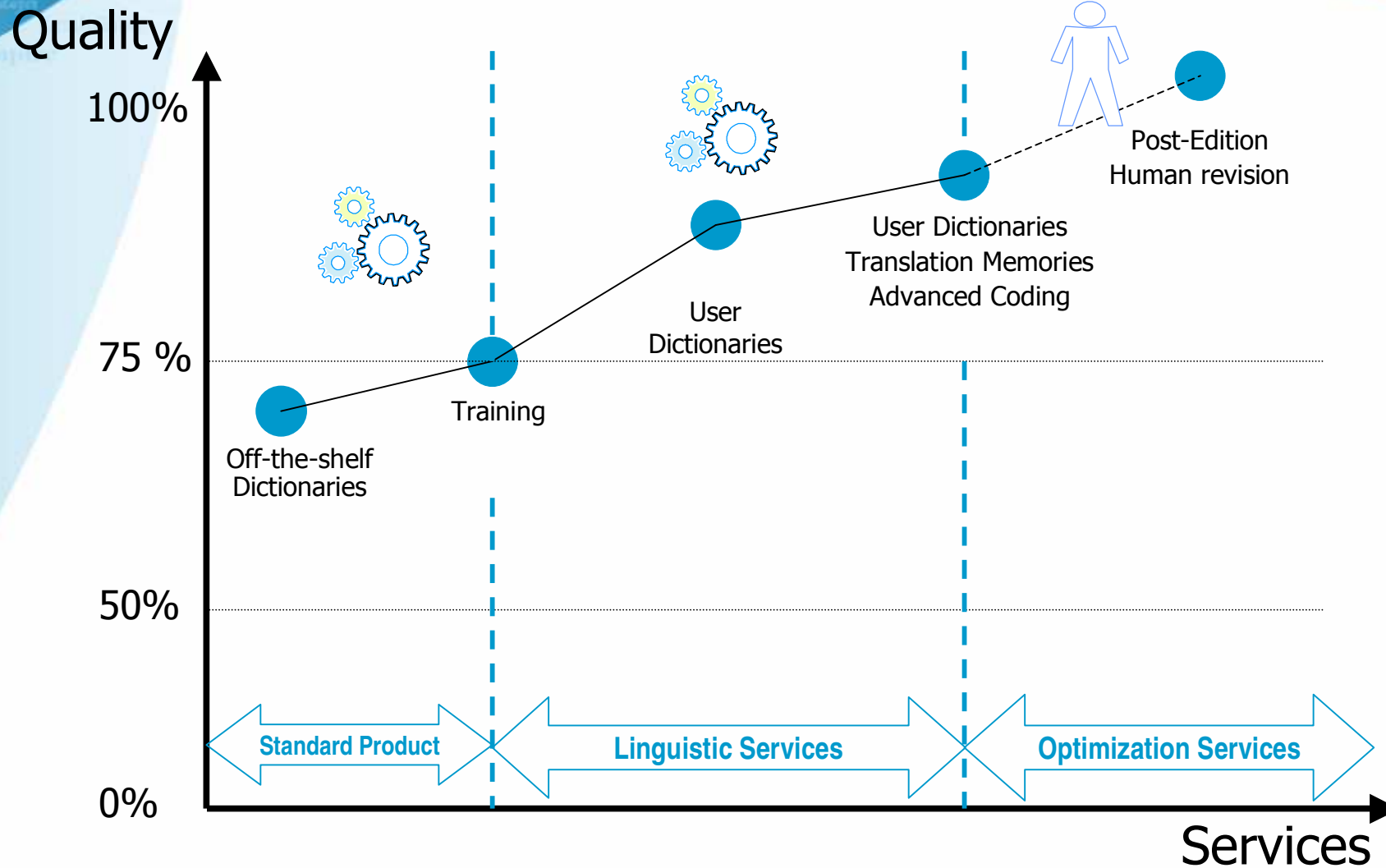
A quality to understand content of foreign documents

#### Notes

⇒ Inbound Translation: To understand foreign language texts : Browsing Quality is sufficient

⇒ Outbound Translation: To publish text in other languages : Generate first revision quality level with MT. Then, fully post-edit text to produce high quality documents

# How to optimize translation quality ?





## ⇒ 4 packaged solutions to integrate automatic translation technologies within corporations

- ⇒ Intranet Self-Service Translator
- ⇒ Market Intelligence
- ⇒ Multilingual Web Site
- ⇒ Documentation and localization

## ⇒ Packaged Product and Service

- ⇒ SYSTRAN methodology
- ⇒ Measurable results in few days (Proof of Concept)
- ⇒ Prototyping model



# SYSTRAN Enterprise Solutions



	<b>Intranet Self-Service Translator</b>	<b>Market Intelligence</b>	<b>Multilingual Web Site</b>	<b>Documentation and localization</b>
<b>Dictionary Size</b>	1500-2000	700-1200	4000-6000	5000-8000
<b>Dictionary building method</b>	<b>Extraction</b>	<b>Glossary</b>	<b>Extraction</b>	<b>Extraction Glossary</b>
<b>Products</b>	Enterprise Workgroup Server	Enterprise Global Server	Enterprise Global Server	Enterprise Global Server
<b>Services</b>	<b>5 – 10 days</b>	<b>10 – 15 days</b>	<b>30 – 40 days</b>	<b>40 – 50 days</b>



# SYSTRAN Methodology



Installation

## SYSTRAN Enterprise, off-the-shelf product

### Monolingual corpus analysis, customer sample

#### Monolingual terminology extraction

- Do Not Translate term list (DNT)
- Not Found Word term list (NFW)

User Dictionary Creation

User Dictionary term entry translation

### Bilingual corpus analysis, customer sample

*Translation Memories – Translated documents*

#### Bilingual terminology extraction

- Do Not Translate term list (DNT)
- Not Found Word term list (NFW)

User Dictionary and Translation Memory creation

### User Dictionary to be coded

Customer Approval for most frequent terms within User Dictionary  
Term coding (IntuitiveCoding and ExpertCoding)

### User Dictionary to be tested

Quality Assurance Test

### User Dictionary to be published

Customer environment integration

### Translation Memories

Customer environment integration

## SYSTRAN Enterprise with personalized linguistic resources

# Intranet Self-Service Translator



## ⇒ Benefits

- ⇒ User Productivity
- ⇒ Better international communication, collaboration
- ⇒ Quick and easy tool

## ⇒ Characteristics

- ⇒ Service : 5 to 10 days
- ⇒ User Dictionary creation
  - Terminology Extraction
- ⇒ Training : 2 days
- ⇒ SYSTRAN Enterprise Workgroup Server

## ⇒ Selected Customers

- ⇒ EADS
- ⇒ Saint-Gobain
- ⇒ Philips Morris
- ⇒ Daimler Chrysler
- ⇒ ...

**SYSTRAN - WebServer 5.0 - Traduction - Windows Internet Explorer**

http://eads.systran.fr/eads/main?s=trans&task=ed240c6ffa14b785-1dfafd1

EADS

**EADS Translator** CONTACT

[Cliquez ici pour afficher le résultat de la traduction précédente.](#)

1. Traduire de  
Français en Anglais

2. Sélectionner votre domaine EADS spécifique (Aerospace ou Defence)  
EADS Aerospace Engineering

3. Sélectionner les données à traduire :

Saisir ou coller le texte ci-dessous

Texte traduit ci-dessous

Ceci est un test pour les matinées découvertes

This is a test for the morning discovery

Une page Web

Un fichier

**Conditions d'utilisation :**

- OUI : pour comprendre des documents rédigés dans une langue étrangère.
- NON : pour traduire un texte dans une langue étrangère.
- Veuillez contacter votre service de traduction pour toute question.

Formats acceptés : DOC, PDF, TXT, RTF, HTML, XML. [Pour plus de précisions, cliquez ici](#)

4.

Powered by SYSTRAN

Internet 100%



## → Benefits

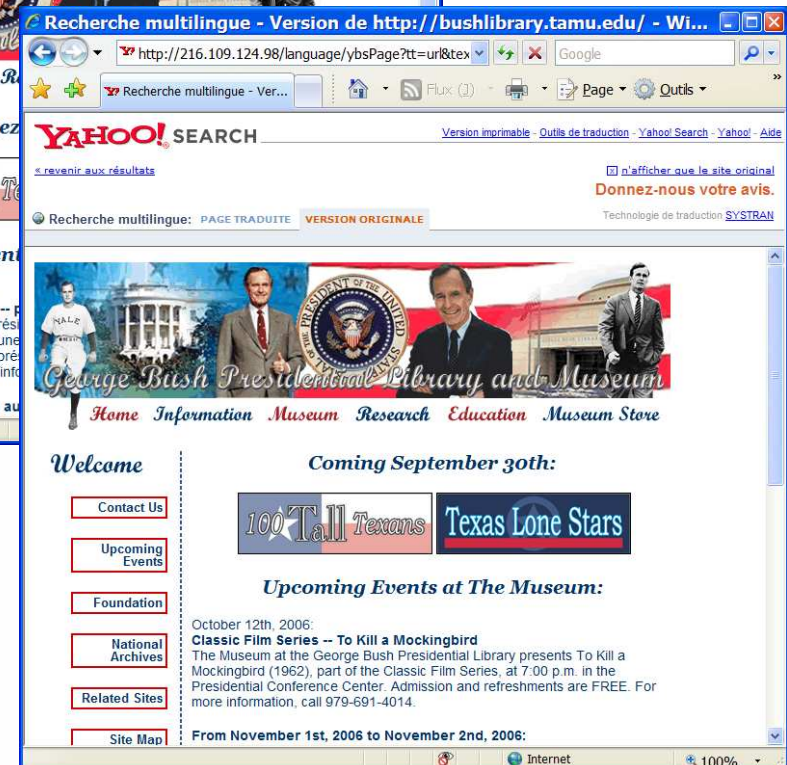
- Broader market intelligence
- Foreign information research
- Open vision on foreign language resources

## → Characteristics

- Service: 10 to 15 days
- User Dictionary creation
  - Glossary and customer terminology integration
- API integration with search engines and market intelligence tools
- Training : 2 days
- SYSTRAN Enterprise Global Server

## → Selected Customers

- Yahoo!
- USA Intelligence Community
- Europol





## ⇒ Benefits

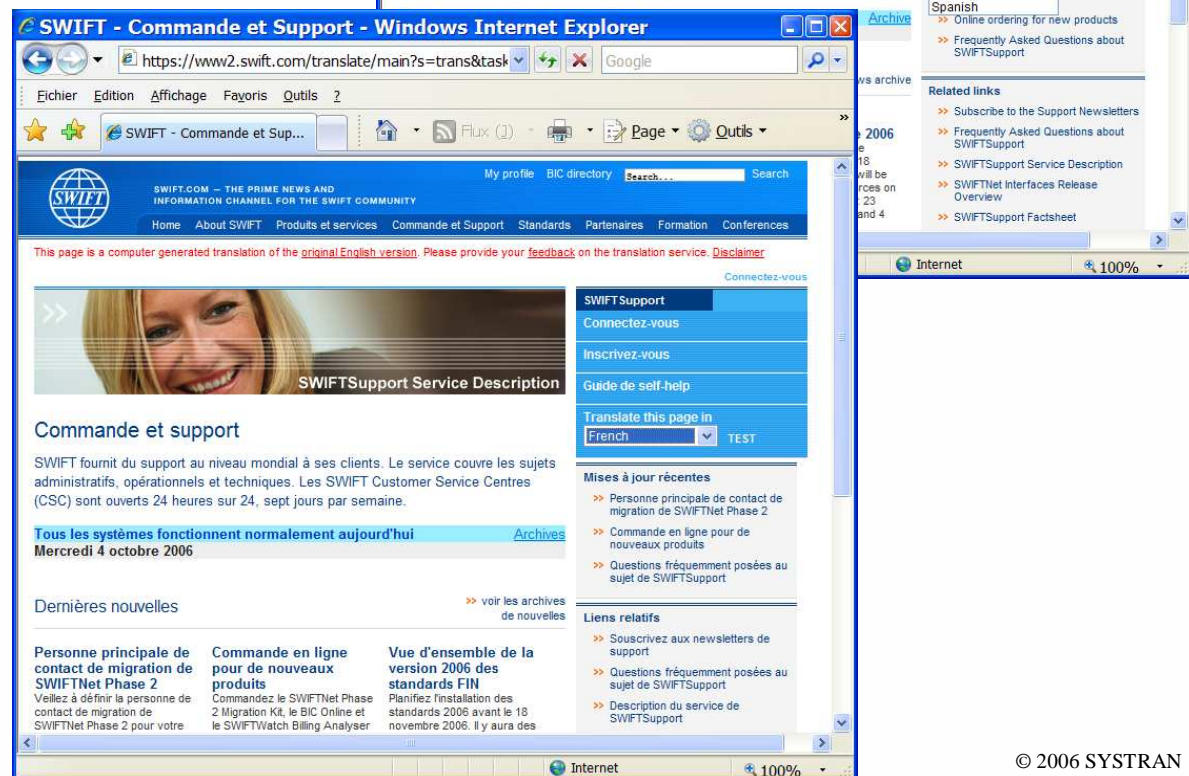
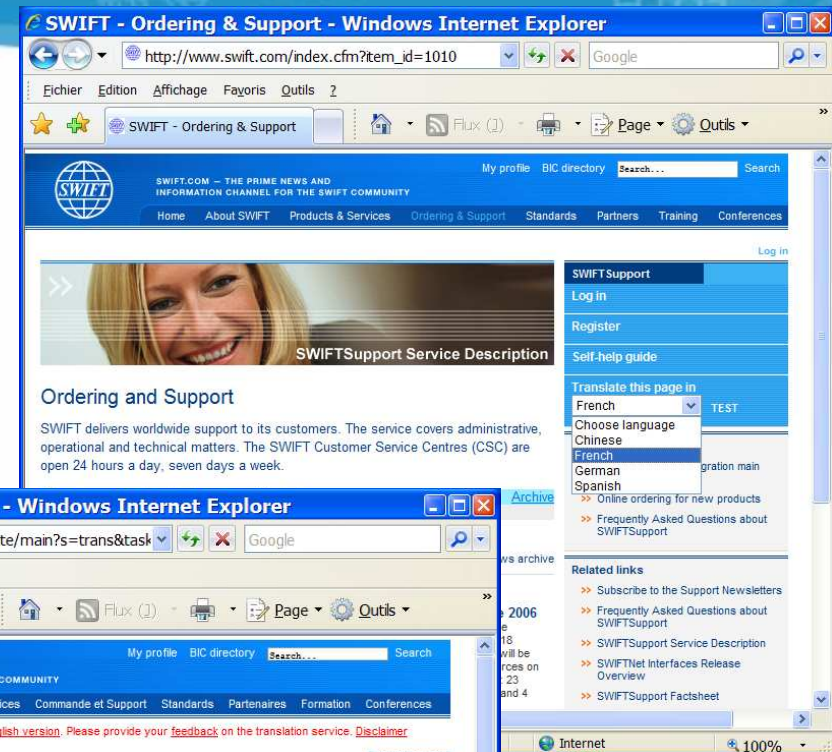
- ⇒ Consistent and aligned content for frequently updated Web sites
- ⇒ Scalable model to add additional languages
- ⇒ Short time to market

## ⇒ Characteristics

- ⇒ Service: 30 to 40 days
- ⇒ User Dictionary Creation
  - Terminology Extraction
- ⇒ SYSTRAN Style Sheet
- ⇒ API integration with Web Publishing tools
- ⇒ Training: 4 Days
- ⇒ SYSTRAN Enterprise Global Server

## ⇒ Selected Customers

- ⇒ Swift
- ⇒ Manutan





# Documentation and Localization



## ⇒ Benefits

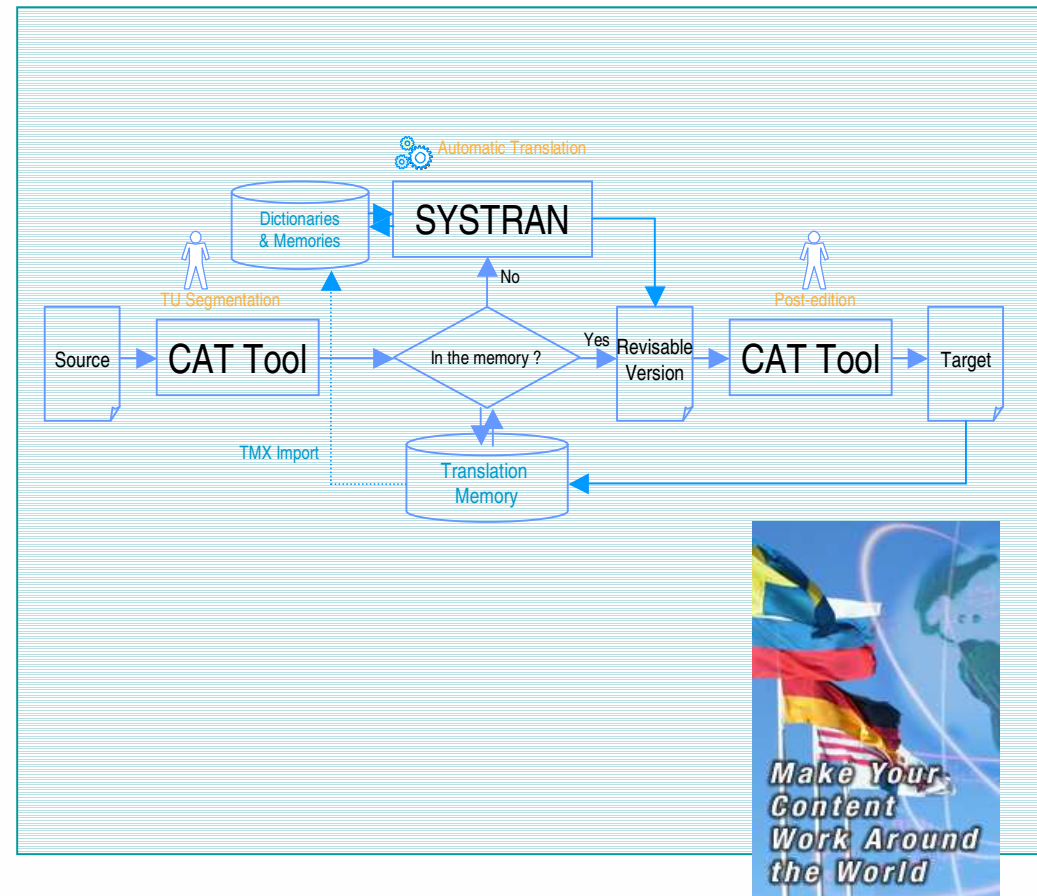
- ⇒ Speed up localization workflow
- ⇒ Complementarities with computer-aided translation solutions
- ⇒ Reuse and consistent translation

## ⇒ Characteristics

- ⇒ Service: 40 to 50 days
- ⇒ User Dictionary Creation
  - Customer glossaries and business terminology integration
- ⇒ Translation Memories integration (TMX)
- ⇒ API integration with computer-aided translation tools (Trados, Star, Idiom...)
- ⇒ Training : 4 days
- ⇒ SYSTRAN Enterprise Global Server

## ⇒ Selected Customers

- ⇒ Symantec
- ⇒ Cisco
- ⇒ Ford



For more information...



⇒ Visit our Web site : [www.systransoft.com](http://www.systransoft.com)

