From the SelectedWorks of Uwe Muegge

April 4, 2011

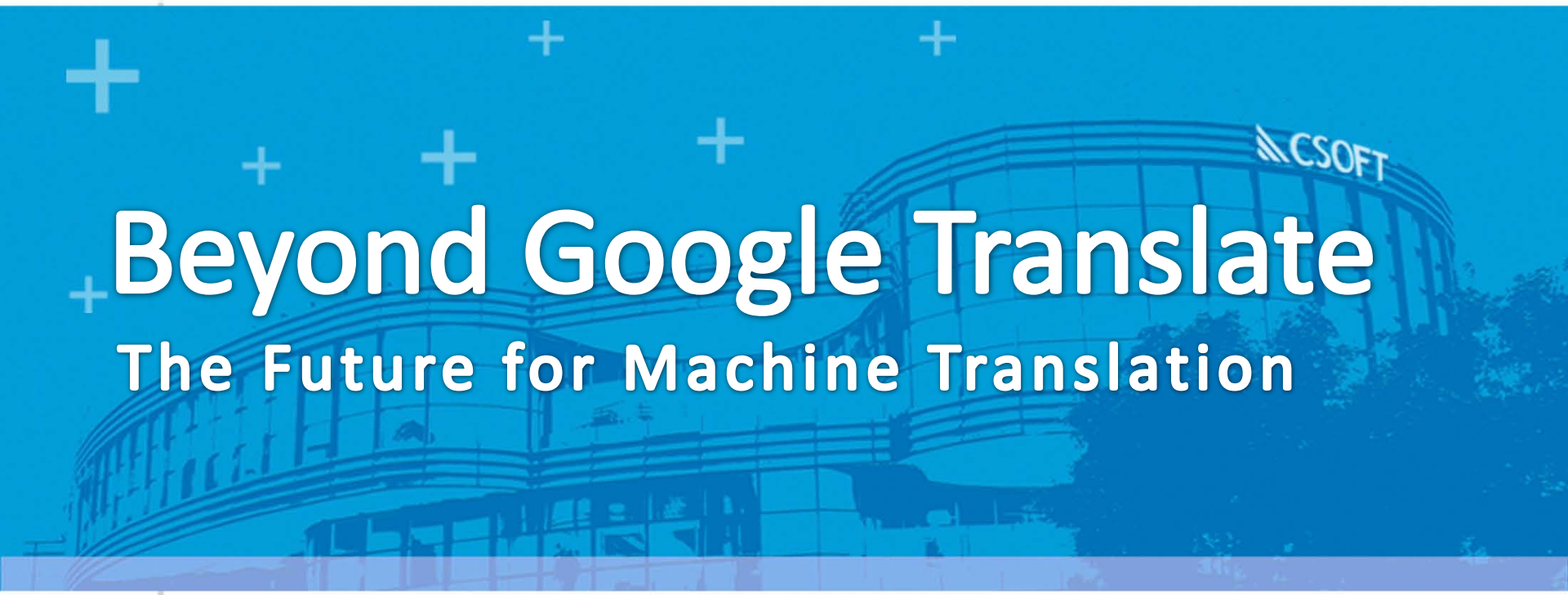
Beyond Google Translate: The future of machine translation

Uwe Muegge

bepress™

Available at: https://works.bepress.com/uwe_muegge/52/

CSOFT™
One World, One Market



Beyond Google Translate

The Future for Machine Translation

Uwe Muegge
CSOFT International

Definition

machine translation

type of software application where the program performs the task of creating a target text in another language without human intervention

How can smaller businesses benefit from machine translation?

Strategy
Machine translation for smaller businesses
Manage Terminology
Deploy a Small-scale MT Solution
Control Authoring

Uwe Muegge

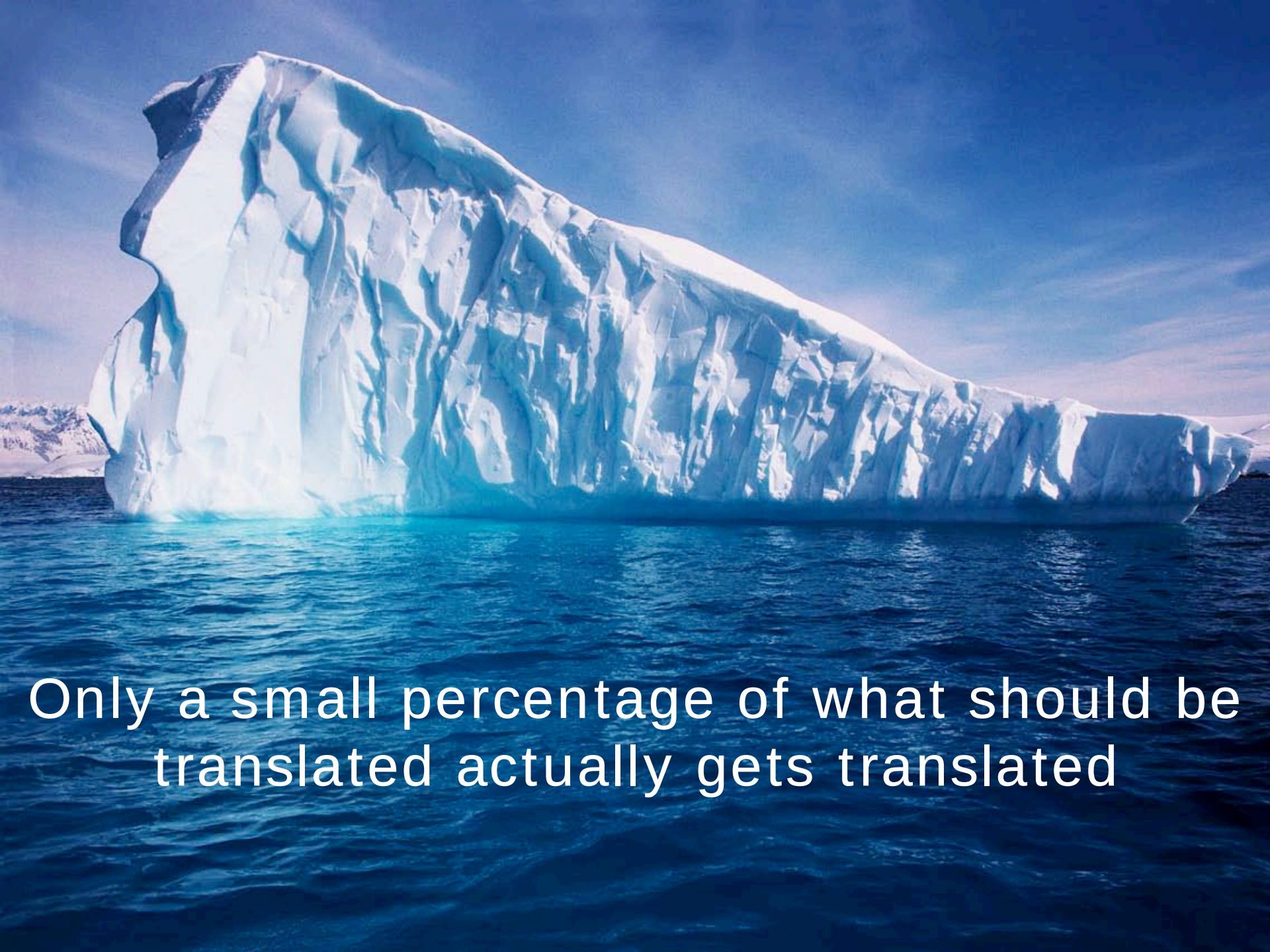
Currently

- CSOFT
  Director, Chief Terminologist
- ISO
  Member, TC 37
- Monterey Institute of Int'l Studies
  Prof., Computer-Assisted Translation

Previously

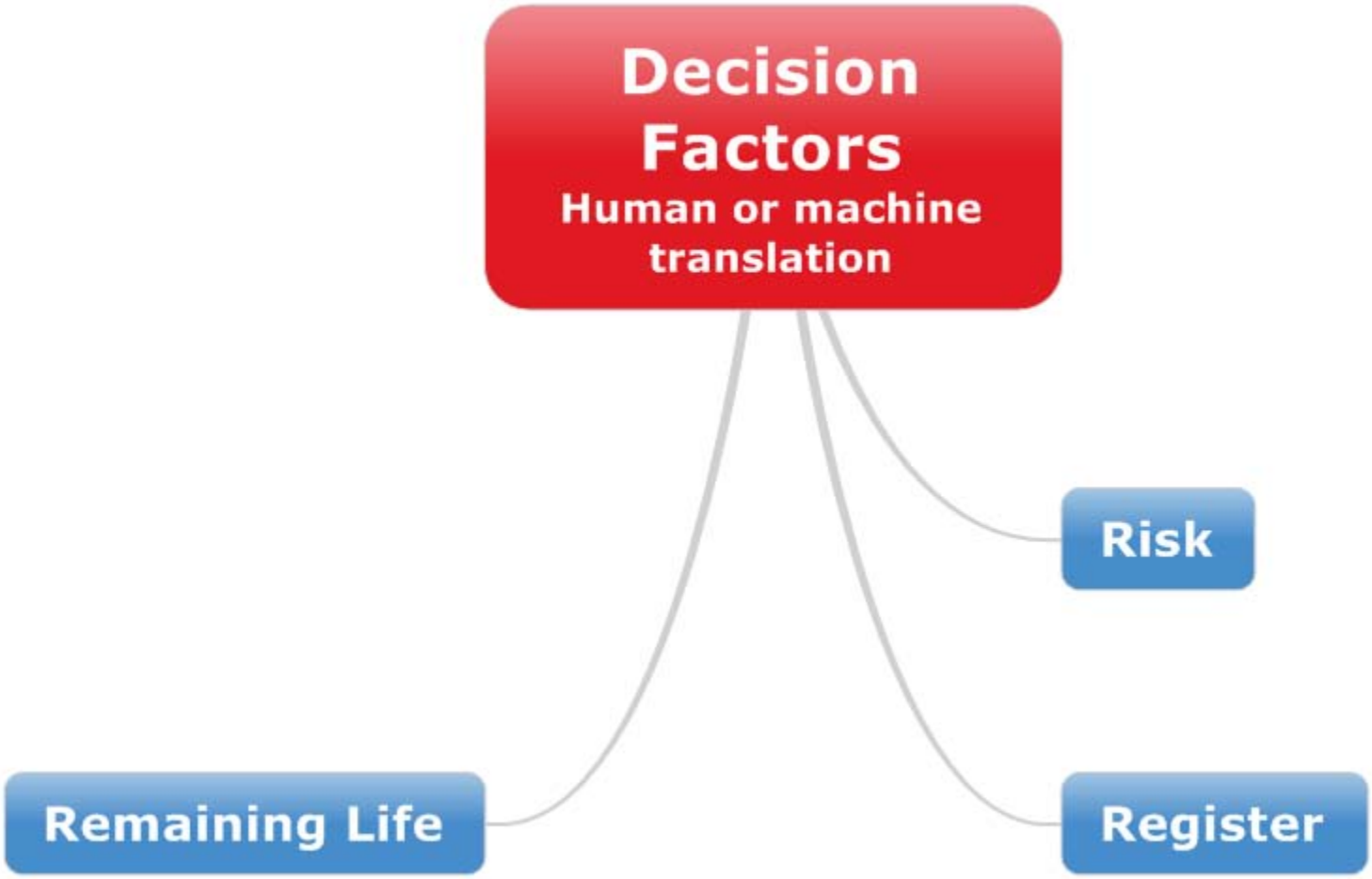
- Medtronic
  Corporate Terminologist
- J.D. Edwards
  Localization Specialist

Why should you be using machine translation?


Only a small percentage of what should be translated actually gets translated

Do you communicate with your clients in the local language - *on all channels*?
E-Mail
Newsletters
Knowledge Bases
Online Help
Training Materials
Website
User Interface
Marketing Collateral
Manuals
Support Cases

How do you determine if a project is suitable for machine translation?


Decision Factors
Human or machine translation
Risk
Register
Remaining Life

Leave to humans: Marketing collateral, corporate websites, regulatory filings, etc.

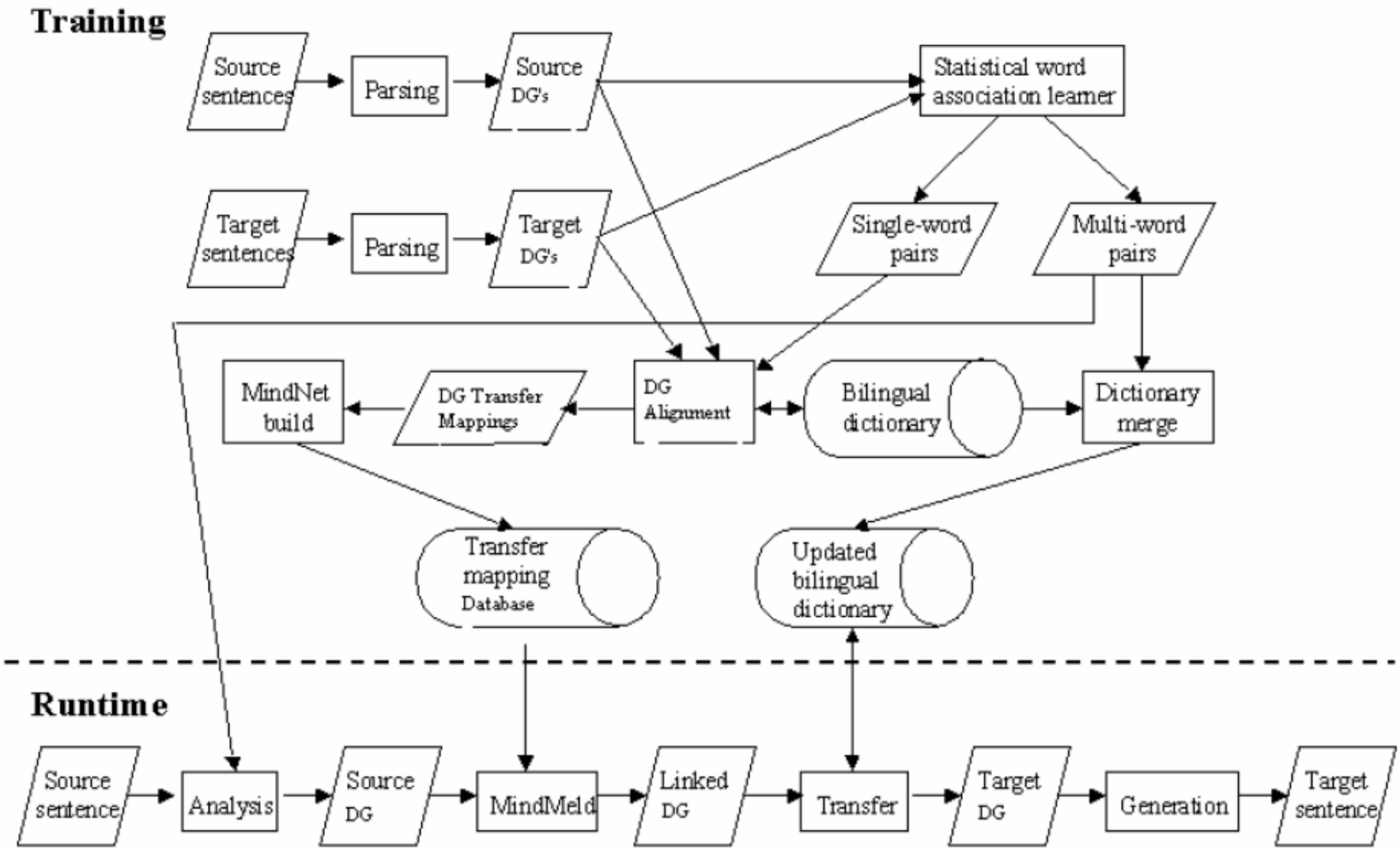
Consider machine translation: Newsletters,
knowledge bases, training materials, etc.

What are the strengths and weaknesses of Google Translate?

Definition

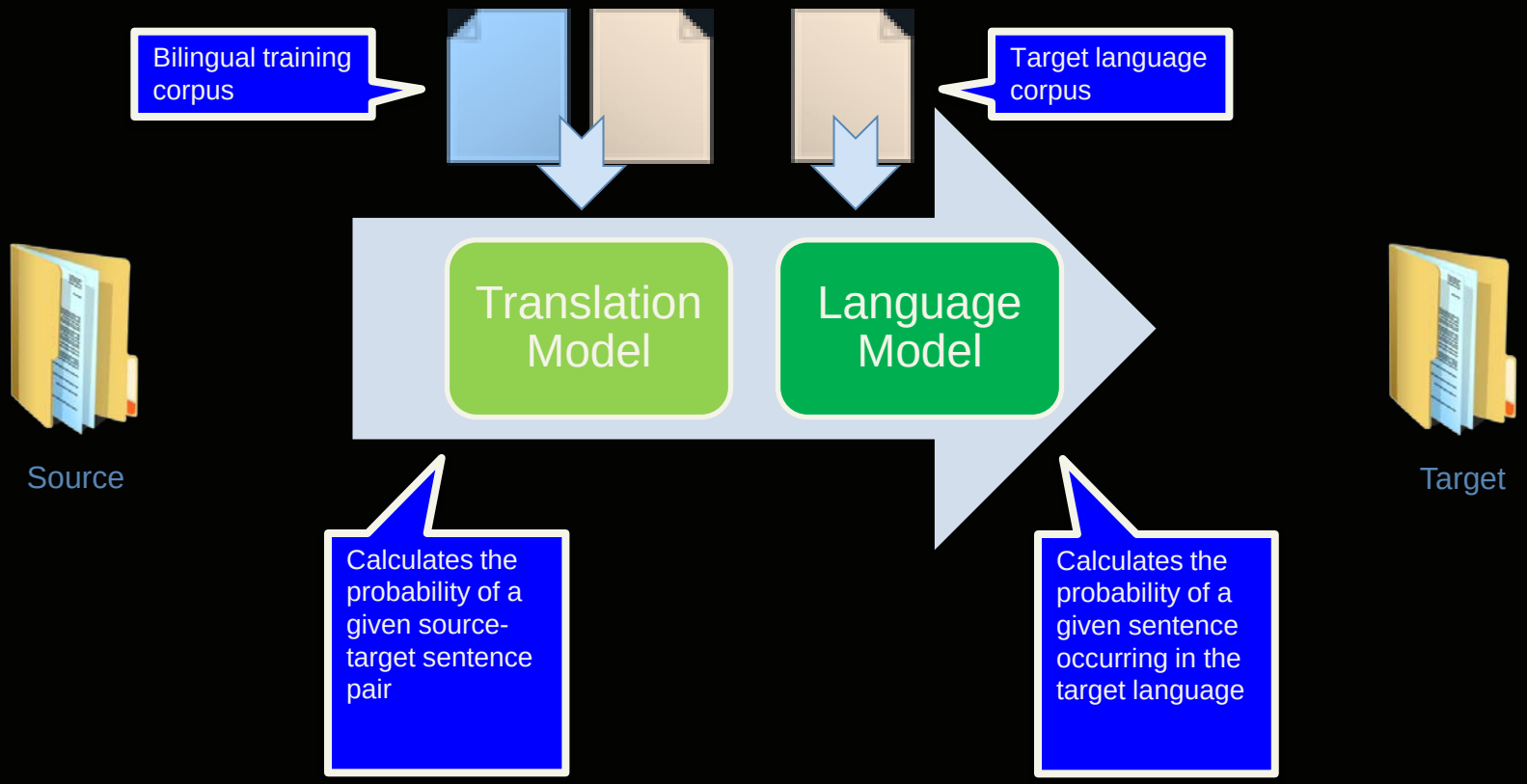
statistical machine translation

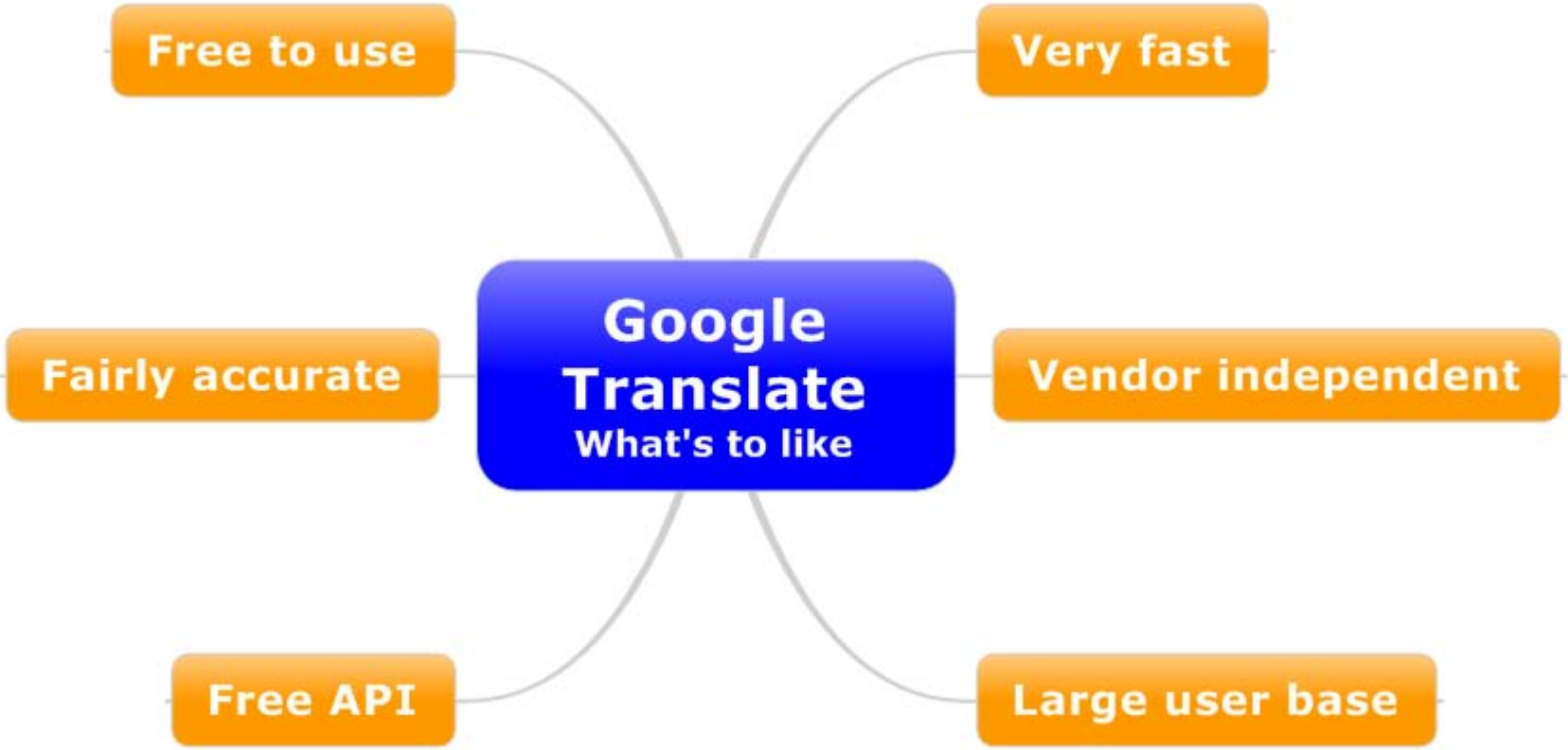
type of machine translation that uses statistical models, which are based on the analysis of large bilingual corpora, to generate sentences in the target language

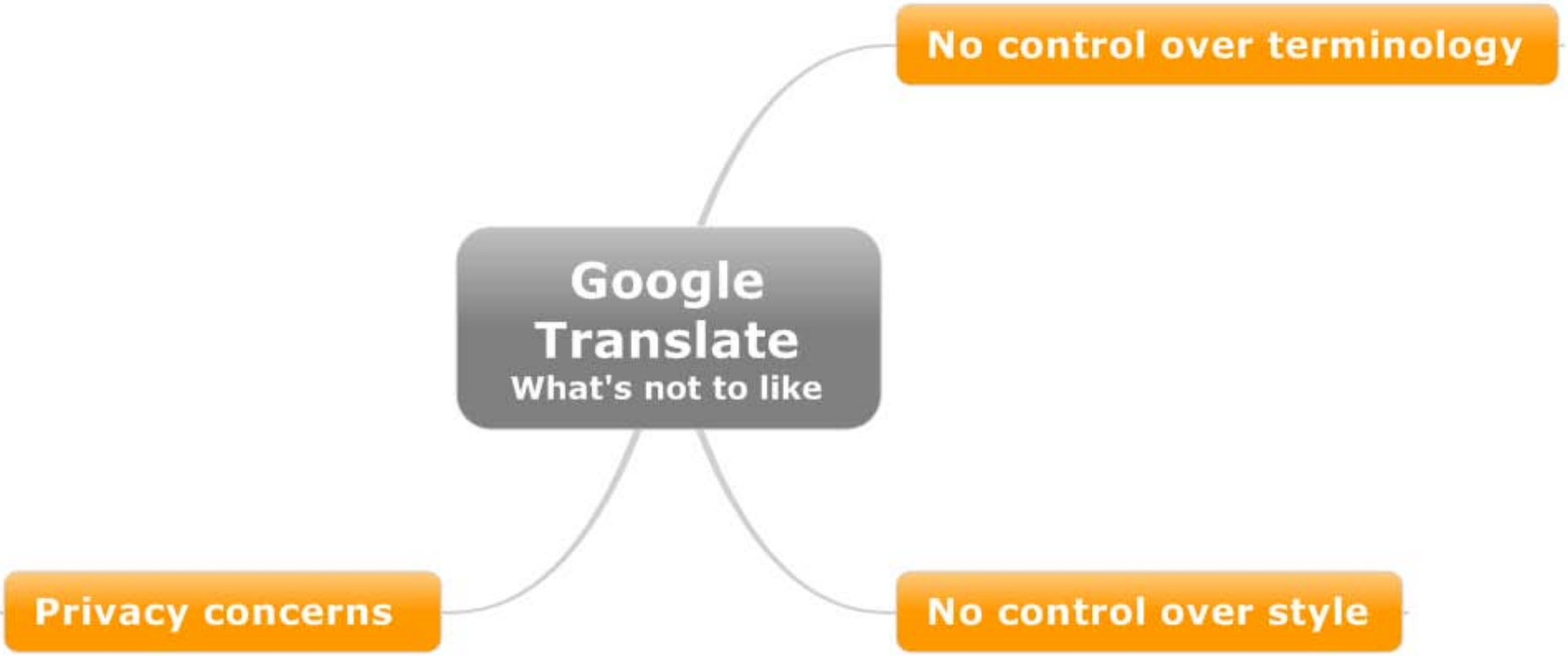


Source: Microsoft

Statistical Machine Translation




Free to use
Very fast
Google Translate
What's to like
Fairly accurate
Vendor independent
Free API
Large user base


Google Translate
What's not to like
No control over terminology
Privacy concerns
No control over style

Google **Language Tools**

Translated search

Type a search phrase in your language. Google will find results in other languages and translate them for you to read.

Search for: [ ] Translate and Search

Search pages written in:
- (•) **Automatically selected languages**
- ( ) Specific languages

My language: English ▾

Example: 1. Search for Bern tourist information.
2. We translate your query into French and German, and find French and German results.
3. Finally, we translate the French and German results back into your language.

Translate text

[ ]

Spanish » English Translate

Google Translate allows you to specify only two things:
1) text
2) language direction

Translate a web page

http://

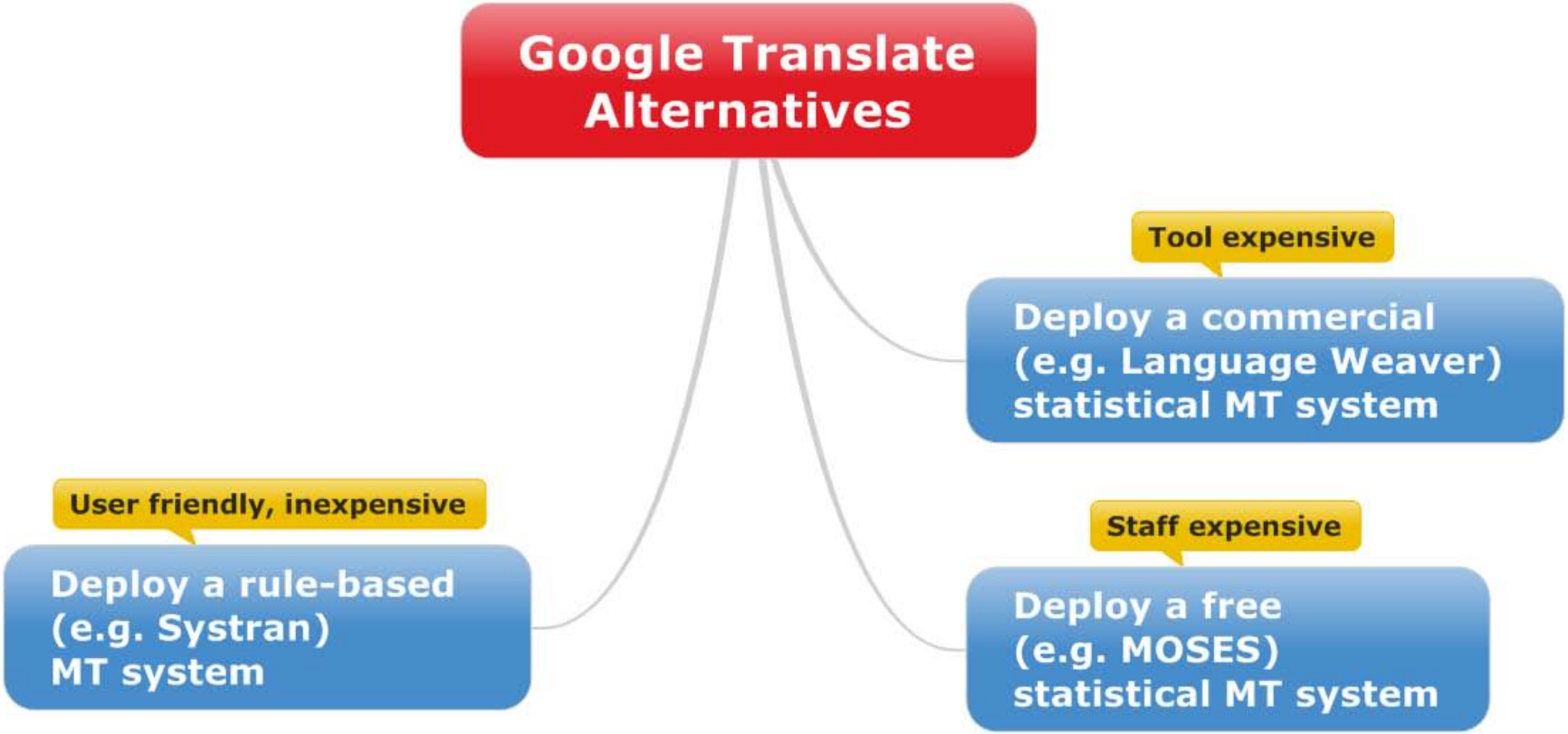
Spanish » English Translate

Use the Google Interface in Your Language

Set the Google homepage, messages, and buttons to display in your selected language via our Preferences page.
Google currently offers the following interface languages:

- Afrikaans
- Akan
- Albanian
- Amharic
- Arabic
- Esperanto
- Estonian
- Faroese
- Filipino
- Finnish
- Kannada
- Kazakh
- Kinyarwanda
- Kirundi
- Klingon
- Norwegian
- Norwegian (Nynorsk)
- Occitan
- Oriya
- Oromo
- Spanish
- Spanish (Latin American)
- Sundanese
- Swahili
- Swedish

If you want more control over your MT system, what are the options?

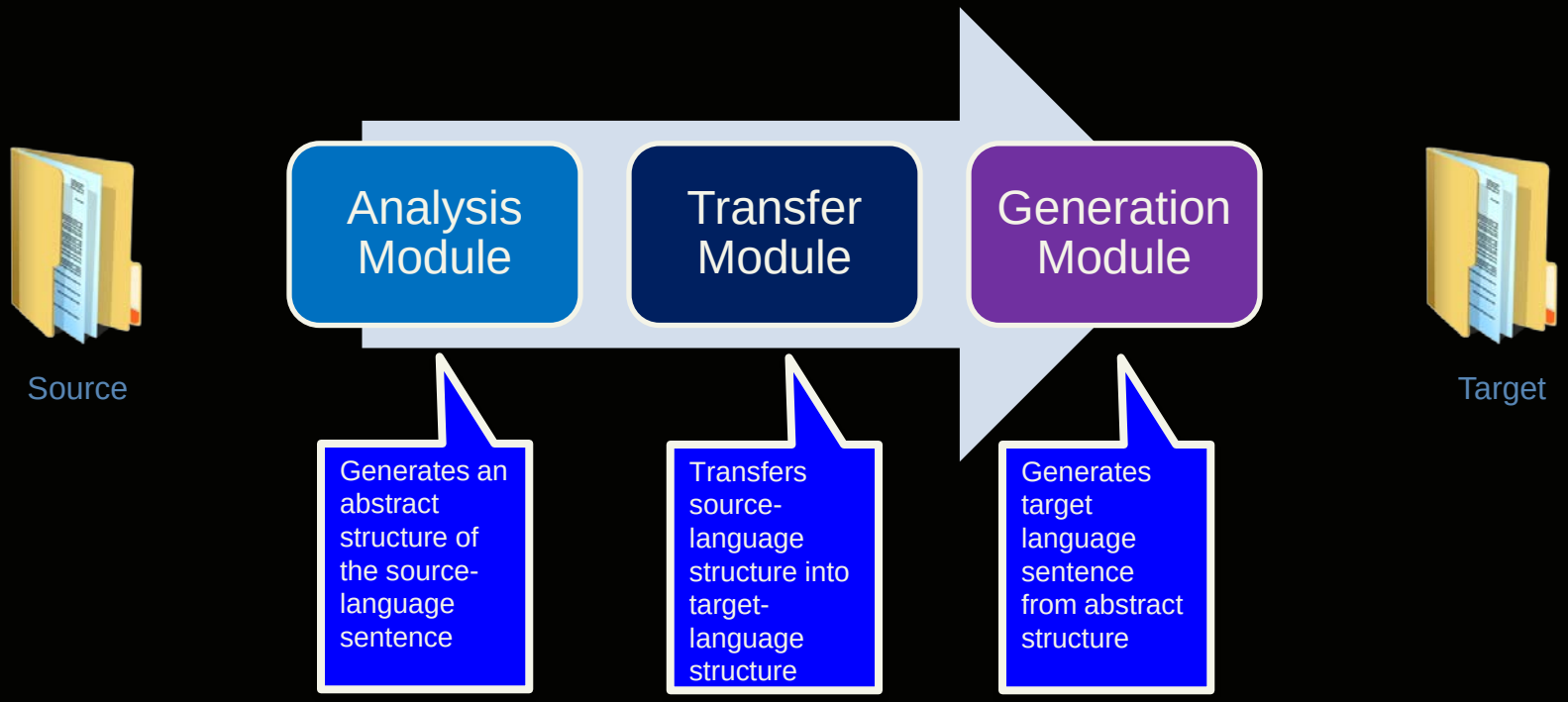

Google Translate Alternatives
Tool expensive
Deploy a commercial (e.g. Language Weaver) statistical MT system
User friendly, inexpensive
Deploy a rule-based (e.g. Systran) MT system
Staff expensive
Deploy a free (e.g. MOSES) statistical MT system

Definition

rule-based machine translation

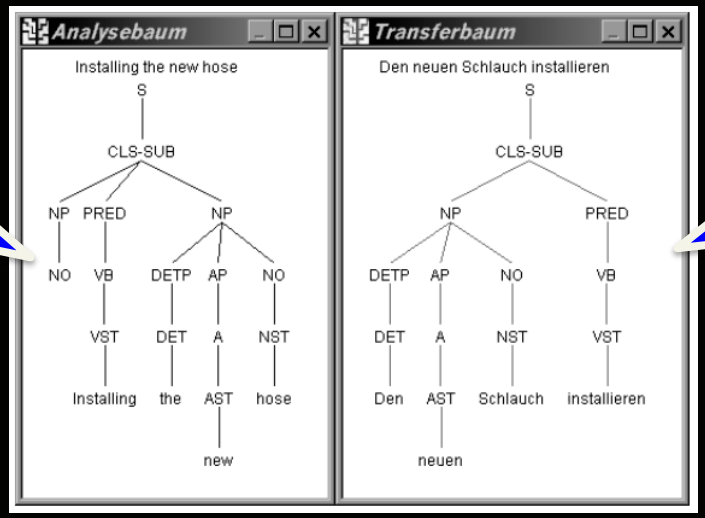
type of machine translation that uses syntactical rules for analyzing sentences in the source language and for generating sentences in the target language

Rule-based Machine Translation

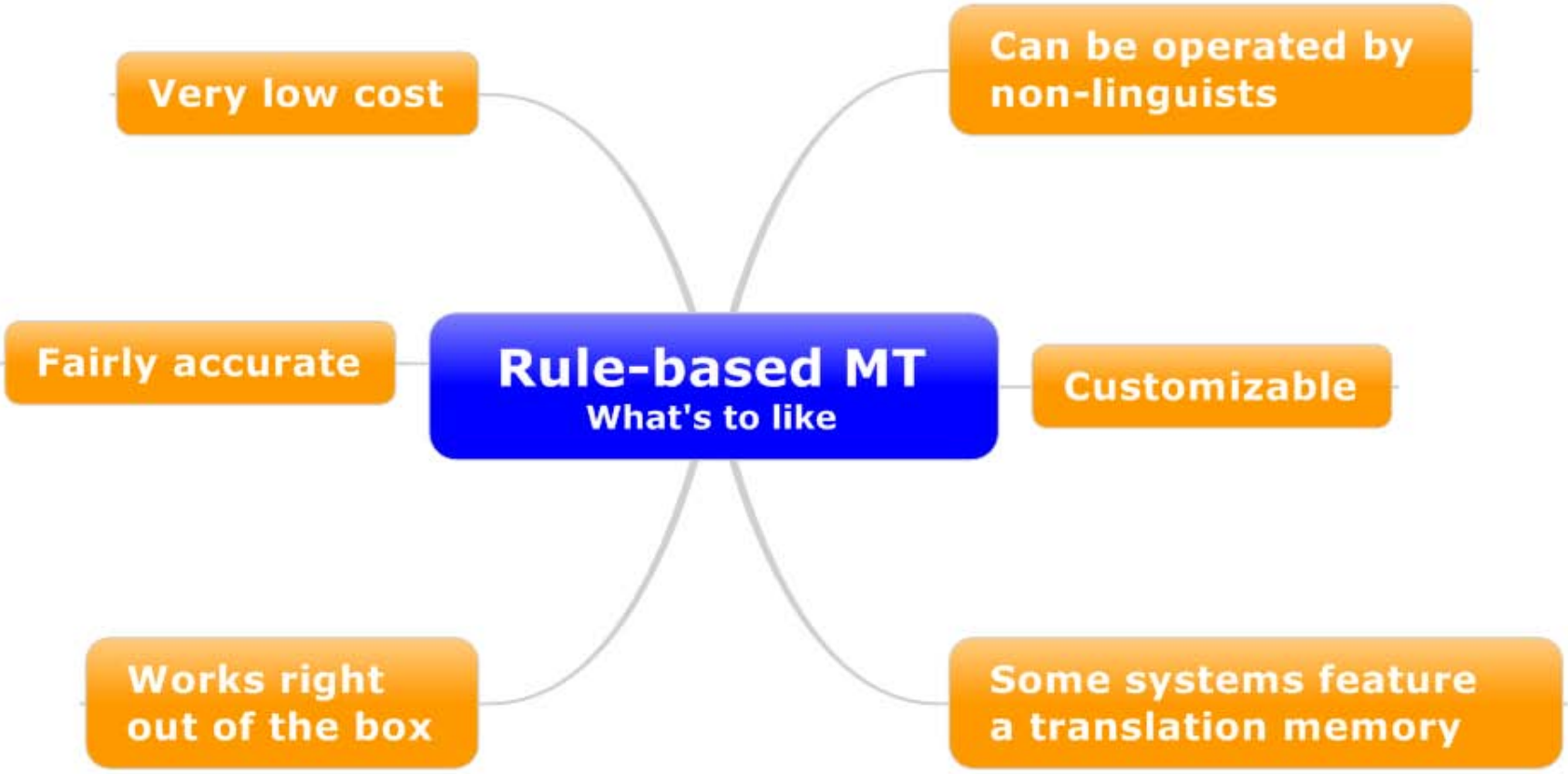


Rule-based Machine Translation

Analysis tree shows syntactical structure of source sentence



Transfer tree shows syntactical structure of target sentence


Rule-based MT
What's to like
Very low cost
Can be operated by non-linguists
Fairly accurate
Customizable
Works right out of the box
Some systems feature a translation memory

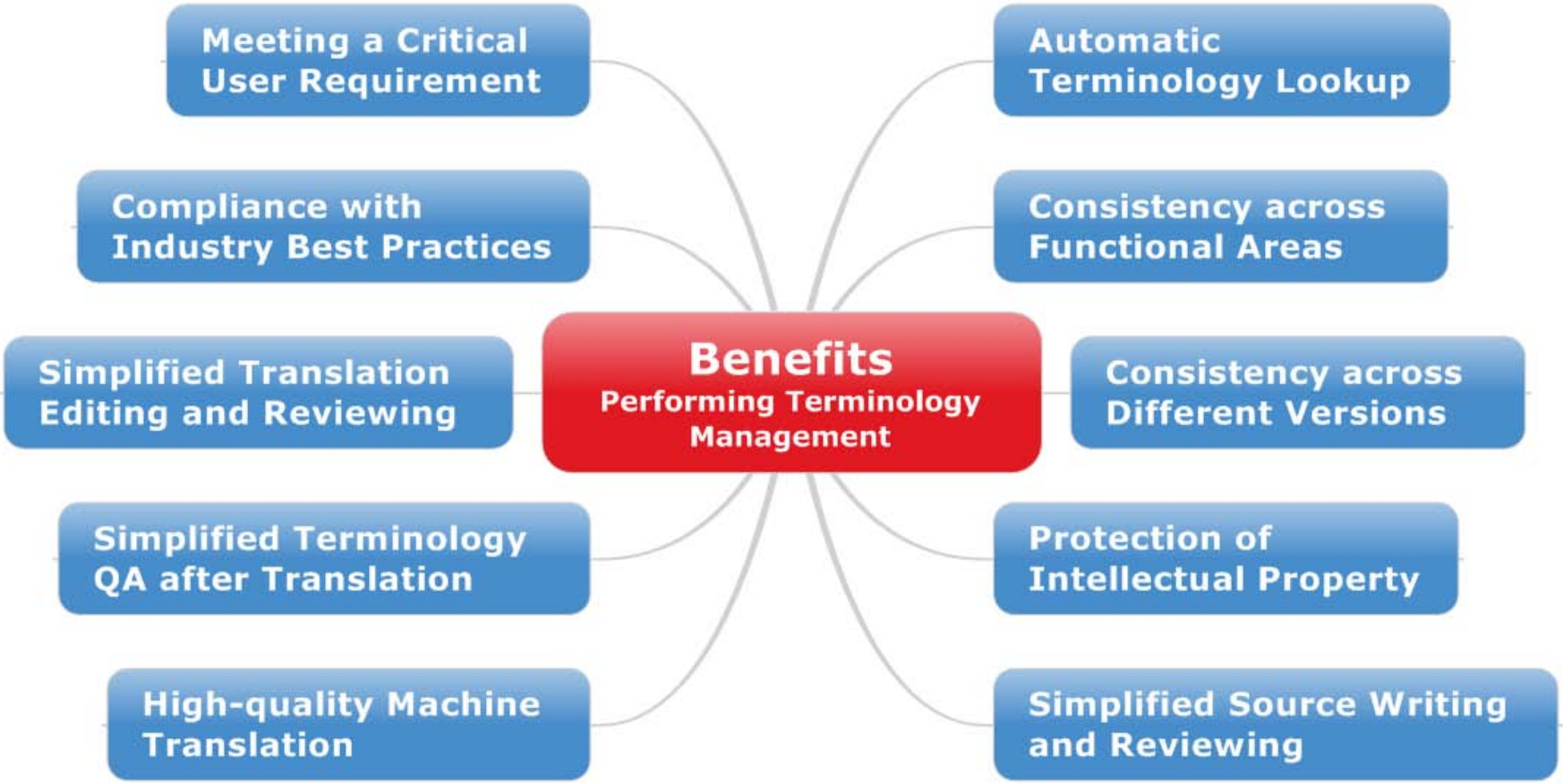
How can you customize a rule-based machine translation system?

Creating/harvesting project-specific terminology

Definition

terminology management

process of controlling key vocabulary


Benefits
Performing Terminology Management
Meeting a Critical User Requirement
Compliance with Industry Best Practices
Simplified Translation Editing and Reviewing
Simplified Terminology QA after Translation
High-quality Machine Translation
Automatic Terminology Lookup
Consistency across Functional Areas
Consistency across Different Versions
Protection of Intellectual Property
Simplified Source Writing and Reviewing

TermWiki

home
- Main Page
- News and Events
- TermWiki Help

popular
- My Glossary New!
- Top Contributors 2011
- TermWiki Toolbar
- Language Jobs
- Trendy Terms

search
Go Search

termwiki
- Dashboard
- Browse Terms
- Add a New Term
- Language Home
- Industry/Domain
- Product Category

EN: trautonium

Definition: early electronic musical instrument developed in 1929 by Friedrich Trautwein

Part of Speech: noun

Usage Status New

Synonym(s):

Sample Image:

Pronunciation:

Industry: Musical equipment

Product Category: Electric musical equipment

Product:

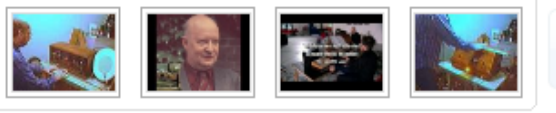
Company:

Related News:

Free, web-based solutions like TermWiki make terminology management easier than ever

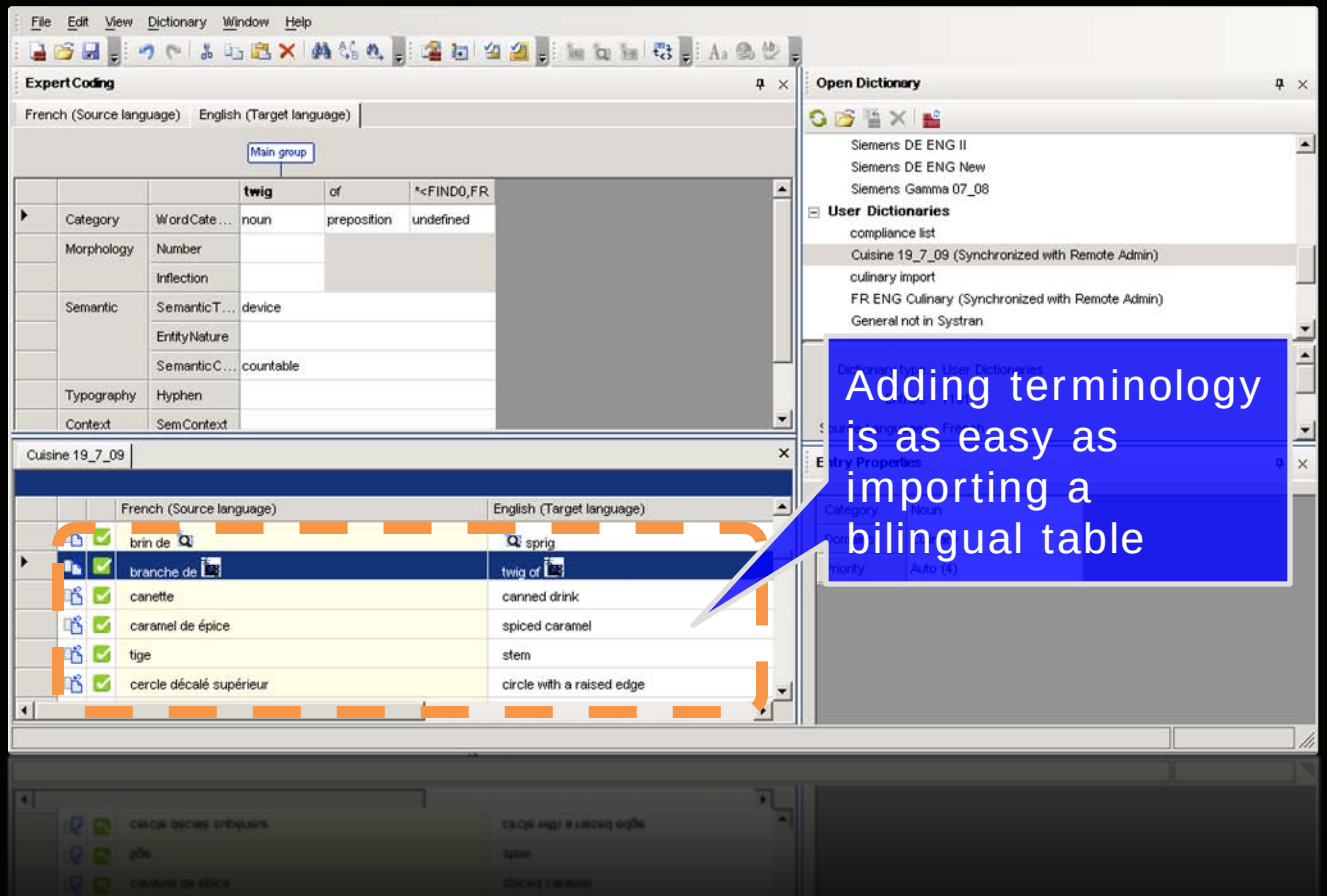
Image search results for **trautonium**

Video search results for **trautonium**



- Created by Uwe on 11/25, 2010, Last edited by Uwe on 11/25, 2010 -

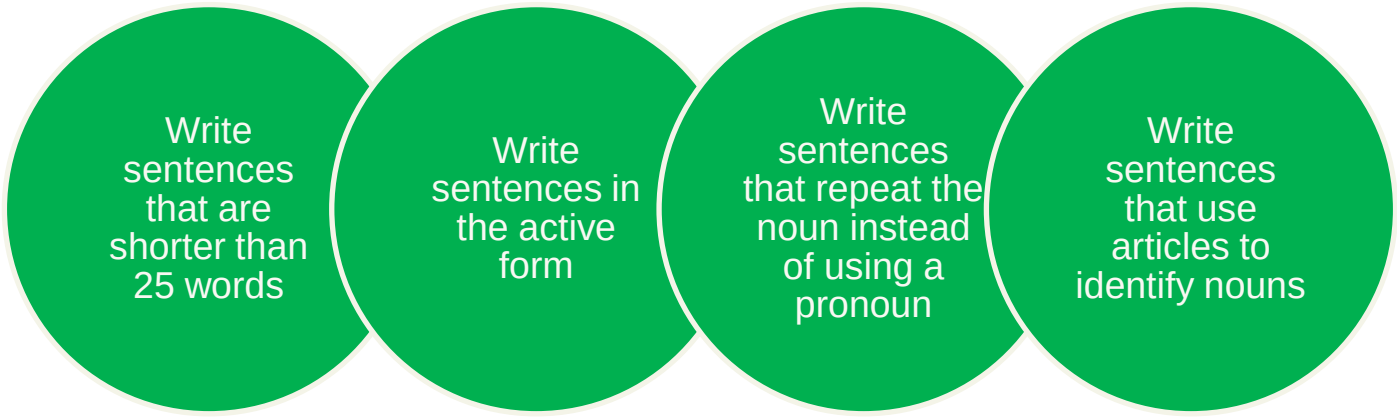
Terminology Component of a Rule-based MT System



What else can you do to optimize translation quality?

Control authoring

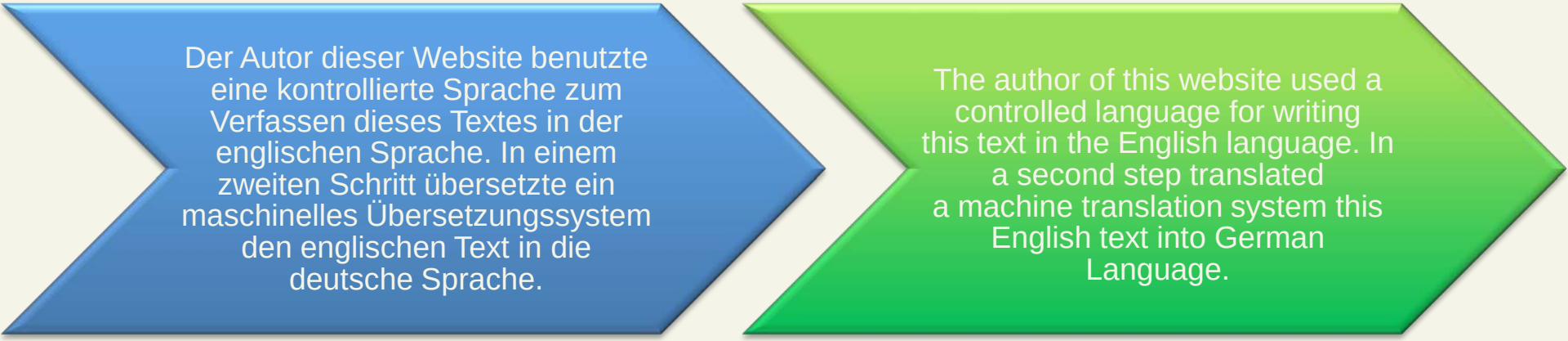
Sample Writing Guidelines



Quick Demo

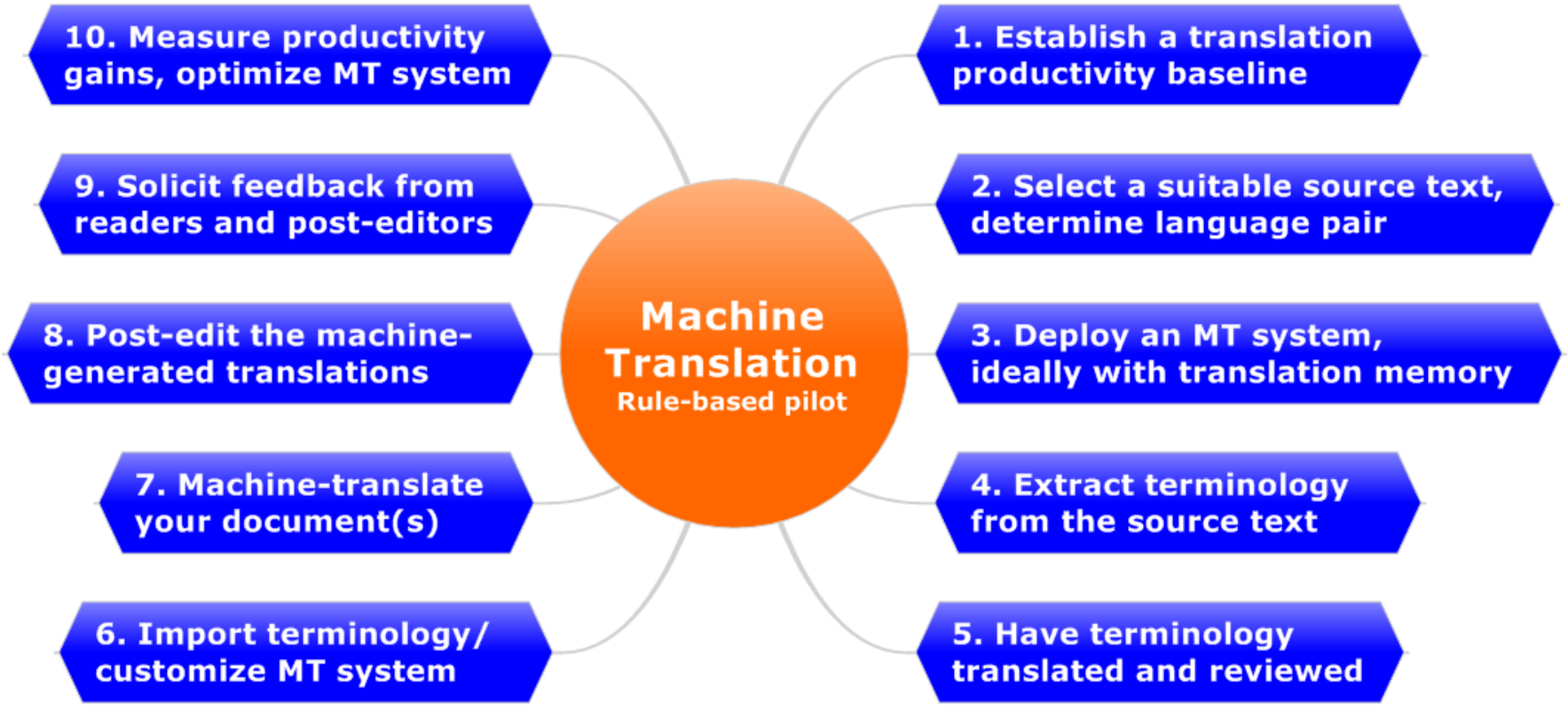
muegge.cc
Maschinenübersetzung
Kontrollierte Sprache
Übersetzungsnormen
Englisch
Deutsch
Französisch
Maschinenübersetzung
Einführung
Kontrollierte Sprache
Regeln für die Maschinenübersetzung
Übersetzungspublikationen
Übersetzungsnormen
Informationen über diese Website
Informationen über den Autor
Google-Anzeigen
Übersetzer
Deutsche Sprache
3 Sprache
Arabisch Sprache
Google-Anzeigen
Deutschland Überweisung
Online senden: Zeit und Geld sparen An jede Bank für nur $4.99 überweis
www.xoom.com/germany
muegge.cc
Anmerkung:
Der Autor dieser Website benutzte eine kontrollierte Sprache, um diesen Text in der englischen Sprache zu schreiben. In einem zweiten Schritt übersetzte ein maschinelles Übersetzungssystem den englischen Text in die deutsche Sprache. Schließlich korrigierte ein menschlicher Übersetzer die Übersetzungsfehler der Übersetzungssoftware. Der Korrekturprozess erforderte beträchtlich weniger Zeit als ein traditioneller Humanübersetzungsprozess.
Produzieren Sie eine virtuelle Version dieser Website in einer anderen Sprache.
Ein maschinelles Übersetzungssystem übersetzt die komplette Website in ein paar Sekunden.
Der Autor dieser Website hat eine kontrollierte Sprache verwendet, um diese Website zu schreiben. Diese Website ist für eine maschinelle Übersetzung geeignet.
Klicken Sie auf die gewünschten Sprachen, um diese Website mit dem maschinellen Übersetzungssystem von Systran zu übersetzen.
Deutsch > Englisch.
Deutsch > Französisch.
Deutsch > Italienisch.
Deutsch > Portugiesisch.
Deutsch > Spanisch.
Englisch > Arabisch.
Englisch > Chinesisch.
Englisch > Deutsch.
Englisch > Französisch.
Englisch > Griechisch.
Englisch > Niederländisch.
Englisch > Italienisch.
Englisch > Japanisch.
Google-Anzeigen
Call Germany from USA 3c
No Fees nationwide access, call log 500 m free Call log, recharge bonus
www.3longdistance.com
Dirk Vincken Text&Bild
Erstklassige Fotos, anspruchsvolle Texte, knackige Reportagen
www.dirk-vincken.de
Beglaubigte Übersetzungen
Russisch Dolmetscherin in München Ausführliche Informationen unter
www.lay-translations.de
J & M Dolmetschen
Dolmetschen Beratung Organisation Professionelle Diplom-Dolmetscher
jandm-konferenzdolmets...
Übersetzungen Dolmetschen
Qualitativ, bewährt, kostengünstig Fachübersetzungen,

Sample Translation (Systran)



Strategy
Machine translation for smaller businesses
Manage Terminology
Deploy a Small-scale MT Solution
Control Authoring

What are the steps to implementing a rule-based MT program?


Machine Translation
Rule-based pilot
1. Establish a translation productivity baseline
2. Select a suitable source text, determine language pair
3. Deploy an MT system, ideally with translation memory
4. Extract terminology from the source text
5. Have terminology translated and reviewed
6. Import terminology/ customize MT system
7. Machine-translate your document(s)
8. Post-edit the machine-generated translations
9. Solicit feedback from readers and post-editors
10. Measure productivity gains, optimize MT system

You can make machine
translation work
for you!

Thank you for your attention!

Uwe Muegge
uwe.muegge@csoftintl.com
Twitter: @UweMuegge

CSOFT™
One World, One Market