Deploying novel MT technology to raise the bar for quality: Key advantages and challenges

Johann Roturier

August 2009

Overview

1. Background: MT – Why? How?
2. SPE: A Case Study in Deploying MT technology
3. Challenges and Next Steps

Localization Requirements

Chapter 1

Planning a migration and upgrade strategy

This chapter includes the following topics:

- About migrations and upgrades
- Using Symantec Packager to streamline migrations and upgrades

About migrations and upgrades

Symantec pcAnywhere supports migration from versions 10.5.x to version 12.1 on Windows 2000/2003 Server/XP/Vista. During a migration, pcAnywhere lets you install over the previous version of the product and preserve user-defined settings.

A system restart for migrations and upgrades is required on Vista, but is only required on Windows 2000/2003 Server/XP if system files need to be updated.

Symantec Packager helps you simplify the process of uninstalling previous versions or distributing preconfigured settings to multiple users.

See "Using Symantec Packager to streamline migrations and upgrades" on page 13.

Table 1-1 includes information that you can use as a reference in planning your migration and upgrade strategy.

1

移行とアップグレードの戦略の計画

この章では以下の項目について説明しています。

- 移行とアップグレードについて
- Symantec Packager を使った簡単な移行とアップグレード

移行とアップグレードについて

Symantec pcAnywhere はバージョン 12.0.x から 12.5 への移行を Windows 2000/2003/XP/Vista でサポートします。移行中、pcAnywhere は以前の製品バージョンに上書きインストールしてユーザー定義の設定を保存できます。

Vista では移行とアップグレードでシステムの再起動が要求されますが、Windows 2000/2003/XP ではシステムファイルを更新する必要がある場合のみ要求されます。

Symantec Packager では以前のバージョンをアンインストールする処理または事前定義済みの設定を複数のユーザーに配布する処理が簡単にできます。

p.9 の「Symantec Packager を

い。

移行とアップグレードの戦略を計

Localization Requirements

Translation of ‘user interface’ into Japanese

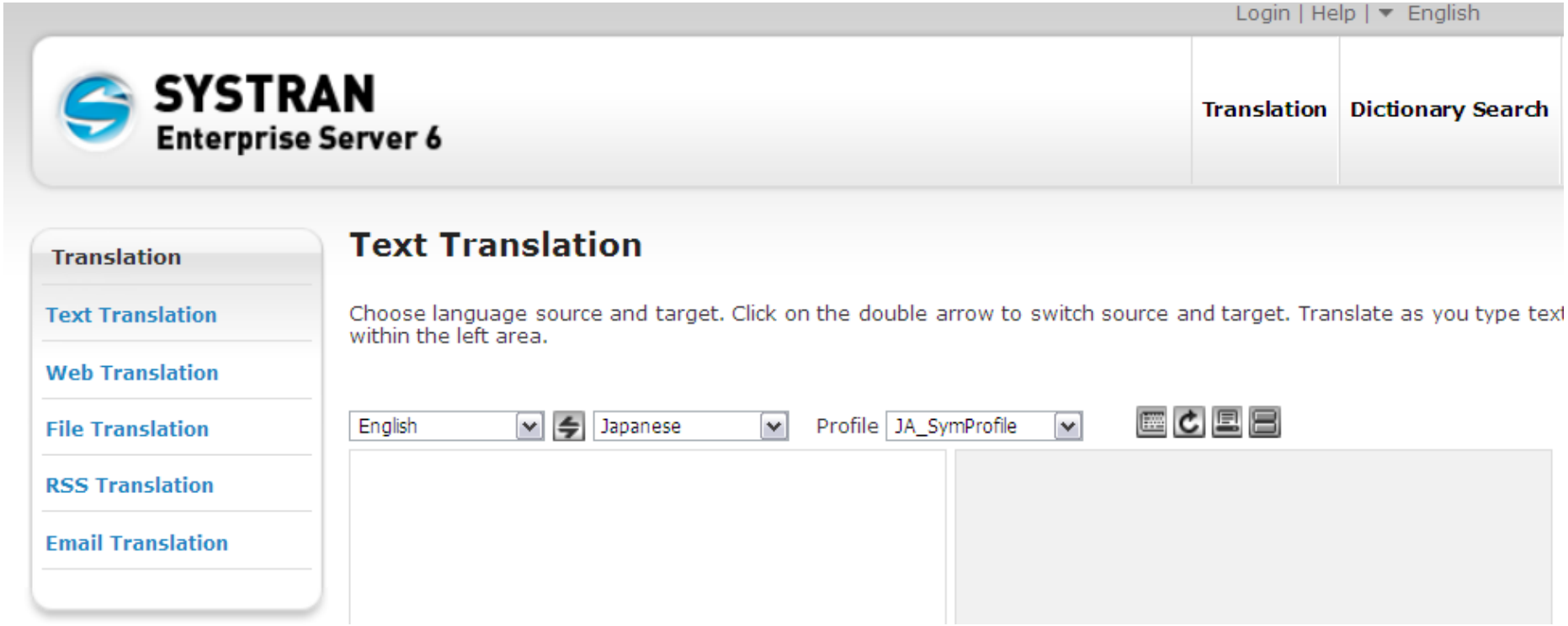
ユーザインタフェイス	ユーザ・インタフェイス	ユーザ インタフェイス
ユーザインタフェース	ユーザ・インタフェース	ユーザ インタフェース
ユーザインターフェイス	ユーザ・インターフェイス	ユーザ インターフェイス
ユーザインターフェース	ユーザ・インターフェース	ユーザ インターフェース
ユーザーインタフェイス	ユーザー・インタフェイス	ユーザー インタフェイス
ユーザーインタフェース	ユーザー・インタフェース	ユーザー インタフェース
ユーザーインターフェイス	ユーザー・インターフェイス	ユーザー インターフェイス
ユーザーインターフェース	ユーザー・インターフェース	ユーザー インターフェース

Why use MT?

- Increase product terminology consistency
  - Focus on Features and Product Names (which can be enforced through terminology Preparation and machine-translation)
  - Correct Software References (which can be enforced by using a specific MT user dictionary)
- Reduce time to market (TTM) through increased productivity
  - Pre-Translate all content through TM and MT
- Lower localization cost

Production Requirements

- Must be used in conjunction with TM system
- Must support tagged input
- Must be easy to deploy globally (7 languages)
- Input should be controlled (acrocheck project score)



Defining a Production MT Workflow

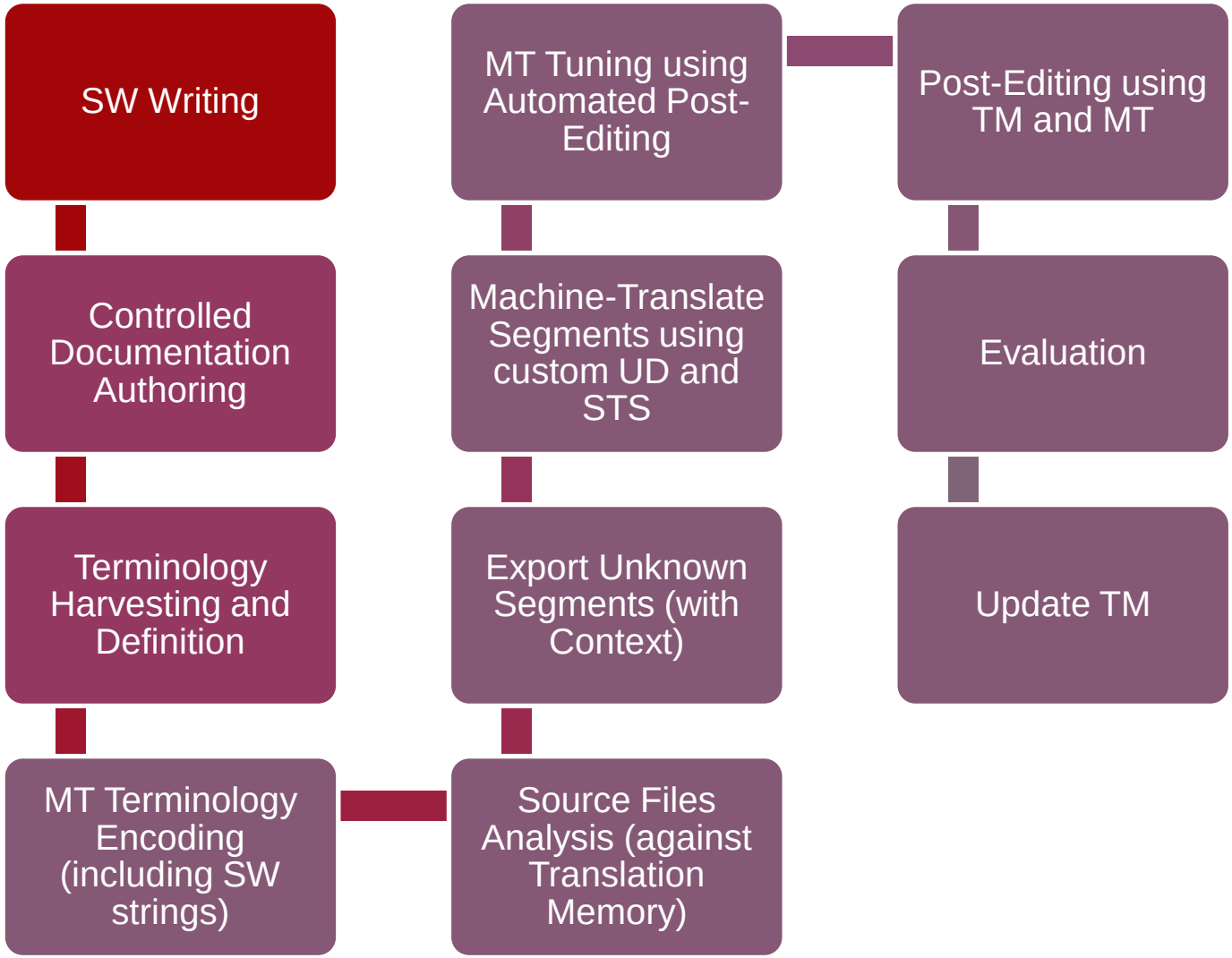
SW Strings

Documentation

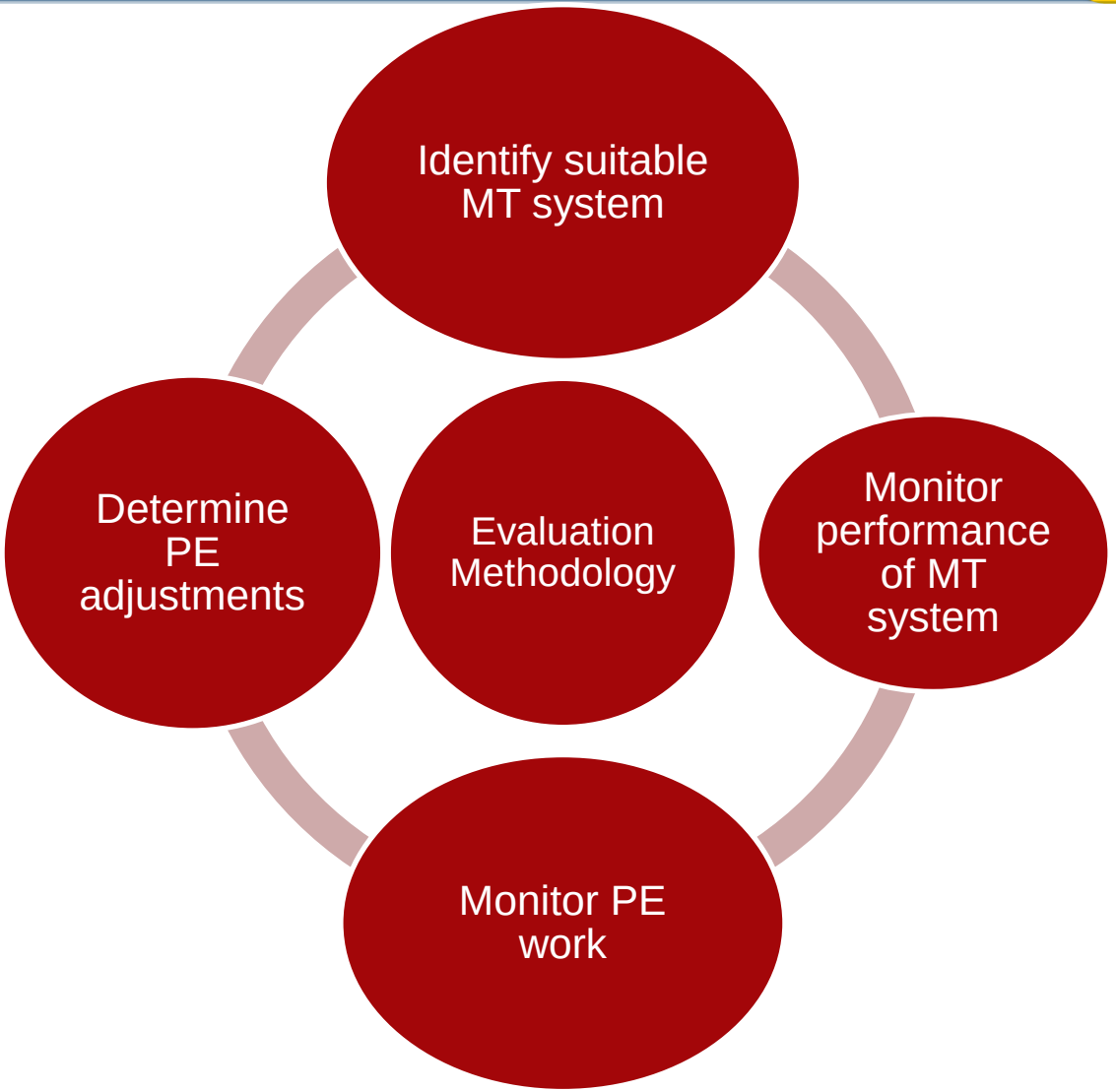
Localisation with MT

TERMINOLOGY

Production MT Workflow



Evaluation


Identify suitable MT system
Monitor performance of MT system
Monitor PE work
Determine PE adjustments
Evaluation Methodology

Timeline

YEAR	STEP	ACHIEVEMENT
2004	Research investigations (Desktop products)	
2005	Initial MT usage for Technical Support translations	Translation Productivity Proven (in-house)
2006	Enterprise system (SYSTRAN v5) and Controlled Authoring Environment (acrocheck 3) deployed	Global access by linguists and writers
2006	User documentation in production (EMEA) with reduced TTM and costs	First large enterprise product (500K words) we localised using MT in 2007 shipped in 7 days (15 days with previous version). Fewer bug fixes to implement in the Help system of this product.
2008	User documentation in production (APJ) with reduced TTM and costs	Reduced costs because PE task is faster than translation task.
2009	Controlled Authoring Environment (acrolinx IQ) deployed Enterprise MT system (SYSTRAN v6) deployed	Opening of ad-hoc translation opportunities Savings have outperformed the investment Quality of output has reached a high level for some languages

Check Point

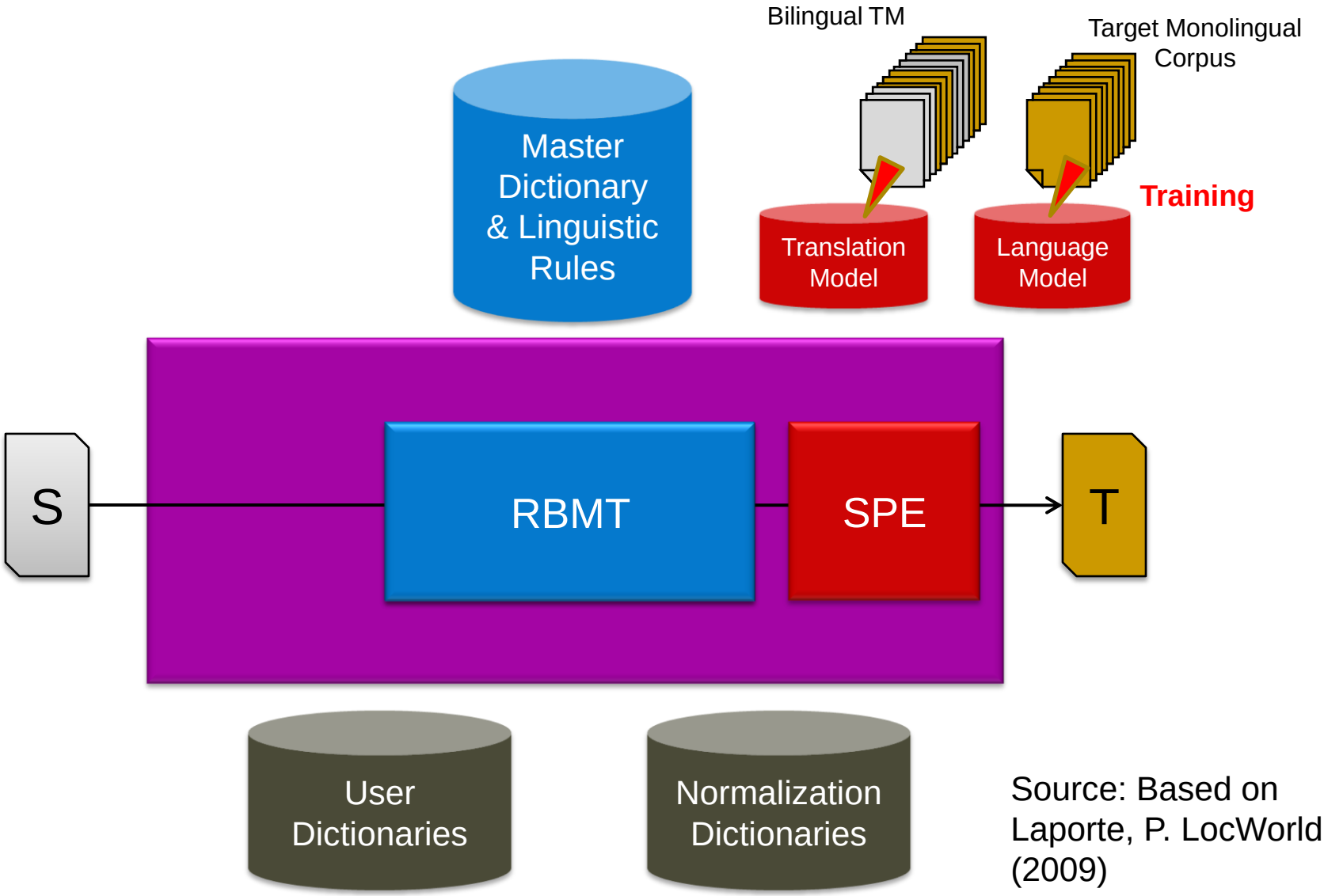
Background: MT – What? Why? How?

2 SPE: A case study in deploying MT technology

Background

- Current MT output
  - Requires repetitive post-editing
  - Lacks fluency
- Need to show continuous improvement to post-editors
  - Current customization strategies:
  - User Dictionaries
    - Difficult and time-consuming to add exceptions (e.g. context clue)
  - Automated Post-Editing
    - Hand-coded (prone to errors)
    - High precision if well-crafted but low coverage
- Idea behind Statistical Post-Editing
  - Learn from Post-Editing activity to build statistical models

New Architecture



Source: Based on Laporte, P. LocWorld (2009)

Case Study

- Collaboration Project with Systran
  - PE Objective
    - Improve overall PE experience in all languages
    - Allow use of SPE for production project
  - Technical Objective
    - Fit into existing workflow
    - Do not deteriorate performance too much
  - Linguistic Objectives
    - Re-case output
    - Preserve key terminology
    - Support for tagged input
    - Show positive Improvement/Degradation

Case Study

- Steps
  - Cleaned TMs
    - Removed Segments with comments
    - Checked Key Terminology
  - Sent TMs and resources to SYSTRAN
    - TMs: 40 K translation units
    - User Dictionaries (around 2 x 10K per language pair)
  - Received SPE models
    - Deployed on staging server
    - Used Rules + SPE as Hybrid System
  - Validated with:
    - Automatic Scores on test set (5K translation units)
    - Human Evaluation on subset of test set (100 segments)
    - PE pilot project

Validation Results (Automatic)

Language/Score	Baseline Systran 6.06	Systran 6.06 (with default Symantec UDs)	Systran Hybrid (with default Symantec UDs)
Italian	33.84 (BLEU 0.1553)	46.42 (BLEU 0.3014)	56.24 (BLEU 0.4442)
French	40.92 (BLEU 0.2399)	51.91 (BLEU 0.3890)	56.72 (BLEU 0.4606)
Japanese	39.46 (BLEU 0.1703)	45.94 (BLEU 0.2336)	58.88 (BLEU 0.4110)
Simp. Chinese	44.40 (BLEU 0.2344)	52.67 (BLEU 0.3201)	58.14 (BLEU 0.3721)
German	29.26 (BLEU 0.1338)	37.67 (BLEU 0.2224)	45.94 (BLEU 0.3223)

Validation Results (Human)

Category/Lang.	JA	CS	FR	IT	DE
Not Found Words	0.7	2.8	-	1	-
Simple Terms	**5.3**	3.9	2	5	1
Phrases	2	0	1	7	11
Meaning	0.54	1.23	1	0	0.74
Determiners	-	5	7	0.5	1.12
Prepositions	1.34	1.85	5.75	6	1.19
Pronouns	-	0.5	1	8	1
Tense	1	1	7	10	2
Number	1	-	1	0.5	0
Gender	-	-	1	-	0.5
Other Grammar	2.75	-	3	26	0.25
Punctuation/Case	0.2	-	0	1	0.07
Word Order(Short)	0.17	-	1	0.7	9.5
Word Order (Long)	1	-	-	3	0.86
Tags	1	-	-	-	-
Average	**1.45**	**2.2**	**3**	**3**	**2.24**

Validation Results (PE)

- PE Results
  - 5 K words
- Simplified Chinese
  - Progress in fluency and meaning
- French and Italian
  - Throughput was improved slightly
  - Overall experience was a little better
  - Use SPE models for large production project (200K+)

Checkpoint

- Background: MT – What? Why? How?
- SPE: A case study in deploying novel MT technology
- 3 Challenges and Next Steps

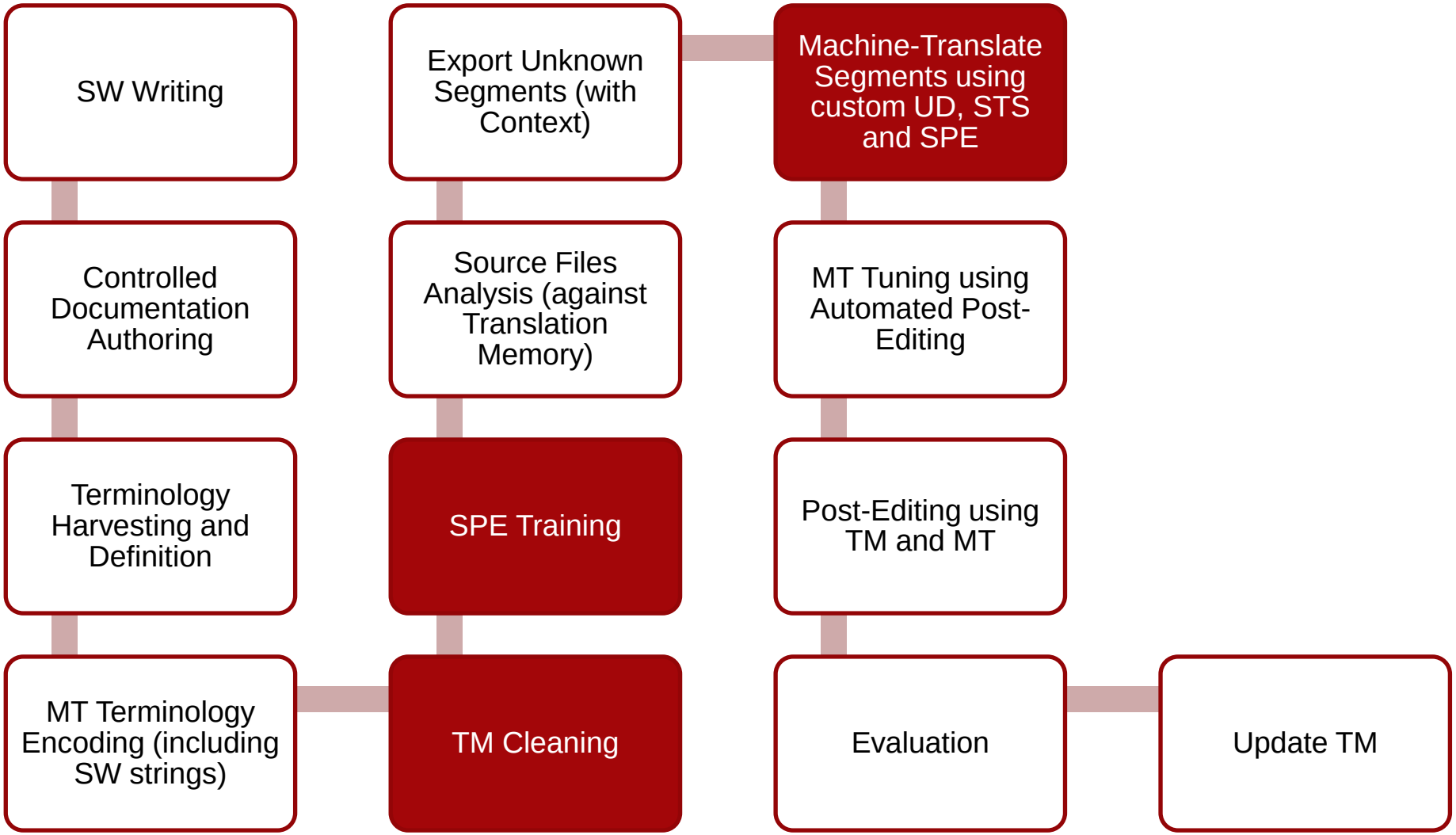
Challenges

- Japanese PE feedback
  - Quality was impacted in places, yet high number of unchanged segments
  - Missing words in output sentence (for example, important negative words such as “not” disappeared) and words that were not in original sentence were used in output
- German PE feedback
  - Errors seem less predictable so less possible to think in terms of making global corrections
  - Still 35% faster than translating from scratch
- Explanations
  - Use of “free” translations in training data
  - Not enough model filtering

Challenges

- Synch MT resources
  - TM
  - SPE Models
  - MT UDS
  - Automated PE module
- Analyze changes brought by SPE
- Maintain performance

Challenge: MT Workflow Update



SPE Developments

- Reduce degradations to a minimum
  - Investigate tuning based on human judgments AND automated scores
  - If not possible, relying on a pure rules-based output may be preferable for certain types of sentences
    - This decision should be made by the MT engine based on its confidence of the MT outputs it can produce.

- Investigate further TM cleanup and management
  - To isolate segments that are not worth re-using at a sub-segment level.
  - Selecting quality data rather than large data sets seems preferable when the objective is to improve PE experience.

Summary

- Good return on investment from MT and SPE
  - Good output quality for most languages
  - Statistical element has already helped raise the bar for quality for some languages (in large production projects)
- Additional opportunities for ad-hoc content localization using MT
  - Technical Support Documentation (including UGC)
  - Training materials
- Close collaboration with post-editors is key
  - Understand better PE task
  - Use PE activity and MT error analysis to optimize MT systems

Confidence in a connected world.

Thank You!

Johann Roturier

johann_roturier@symantec.com