

A Lesson in WorldServer

By CKS

Lucy Jobe-Hyerim Ko-Shuyu Luo-Danica Shi-Janic Tang

Goals



Localize “The Gathering Place” website

1. Target languages: Chinese, Korean and Spanish
2. Contents: Four web pages and a logo;
3. Deliverables:
 - a. Target files: xml files and png files;
 - b. Translation memory and term database;
 - c. Question of Pseduotranslation

Let's start at the very beginning

- SOW Client Meeting
- Consistency
 - Format
 - Writing style
 - Terms
- What to include
 - Cost breakdown

Cost Model Setup

General		Access Control					
Name:	CKS_CostModel						
Description:	Cost Model for TMS Project						
Source Locale:	English (United States) <input type="button" value="▼"/>						
Vendor:	<input type="button" value="▼"/>						
Currency:	USD (\$) <input type="button" value="▼"/>						
Scoping Configuration:	CKS_ScopingConfig <input type="button" value="▼"/>						

Target Locale	Cost per Word for Segment Match Range						
	Repetition	No Match (New Word) 0-50%	50-80%	80-100%	100%	ICE Match	Control
Afar	0.00	0.00	0.00	0.00	0.00	0.00	<input type="button" value="Copy"/> <input type="button" value="Paste"/>
Arabic (Egypt)	0.00	0.00	0.00	0.00	0.00	0.00	<input type="button" value="Copy"/> <input type="button" value="Paste"/>
chinese	0.00	0.00	0.00	0.00	0.00	0.00	<input type="button" value="Copy"/> <input type="button" value="Paste"/>
Chinese (PRC)	0.00	0.00	0.00	0.00	0.00	0.00	<input type="button" value="Copy"/> <input type="button" value="Paste"/>
Chinese (Simplified)	0.08	0.11	0.10	0.09	0.08	0.00	<input type="button" value="Copy"/> <input type="button" value="Paste"/>

Cost Model Setup

- Scalability
- Flexibility
- Automatic Adjustments

Target Locale	Asset	Total	ICE Match	100%	100-80%	80-50%	50-0%	Repetition	Cost Estimate (USD)
Chinese (Simplified)	/Client Files/shuyul_PseudoClient_new/Temp/159762/Source/thegatherplace.wordpress.2016-10-14.xml	265	0	0	0	0	254	11	28.82
Rates(USD)	N/A			0	0.08	0.09	0.1	0.11	0.08
Chinese (Simplified)	Chinese (Simplified) Total	265	0	0	0	0	254	11	28.82
Spanish (Latin America)	/Client Files/shuyul_PseudoClient_new/Temp/159762/Source/thegatherplace.wordpress.2016-10-14.xml	265	0	0	0	0	254	11	28.82
Rates(USD)	N/A			0	0.08	0.09	0.1	0.11	0.08
Spanish (Latin America)	Spanish (Latin America) Total	265	0	0	0	0	254	11	28.82
Korean	/Client Files/shuyul_PseudoClient_new/Temp/159762/Source/thegatherplace.wordpress.2016-10-14.xml	265	0	0	0	0	254	11	31.47
Rates(USD)	N/A			0.09	0.1	0.11	0.12	0.09	
Korean	Korean Total	265	0	0	0	0	254	11	31.47
	CKS_DTP	0	0	0	0	0	0	0	45
	Grand Total	795	0	0	0	0	762	33	134.11

Filter Configuration

A filter group for docx and xml

Name	Description	MIME type	Filter
CKS Filter Group	docx, html, xml	application/msword text/xml	DOC Filter XML (NT) Filter

MIME type	Filter	Filter Configuration
application/msword	DOC Filter	Default
text/xml	XML (NT) Filter	CKS_xml_wp

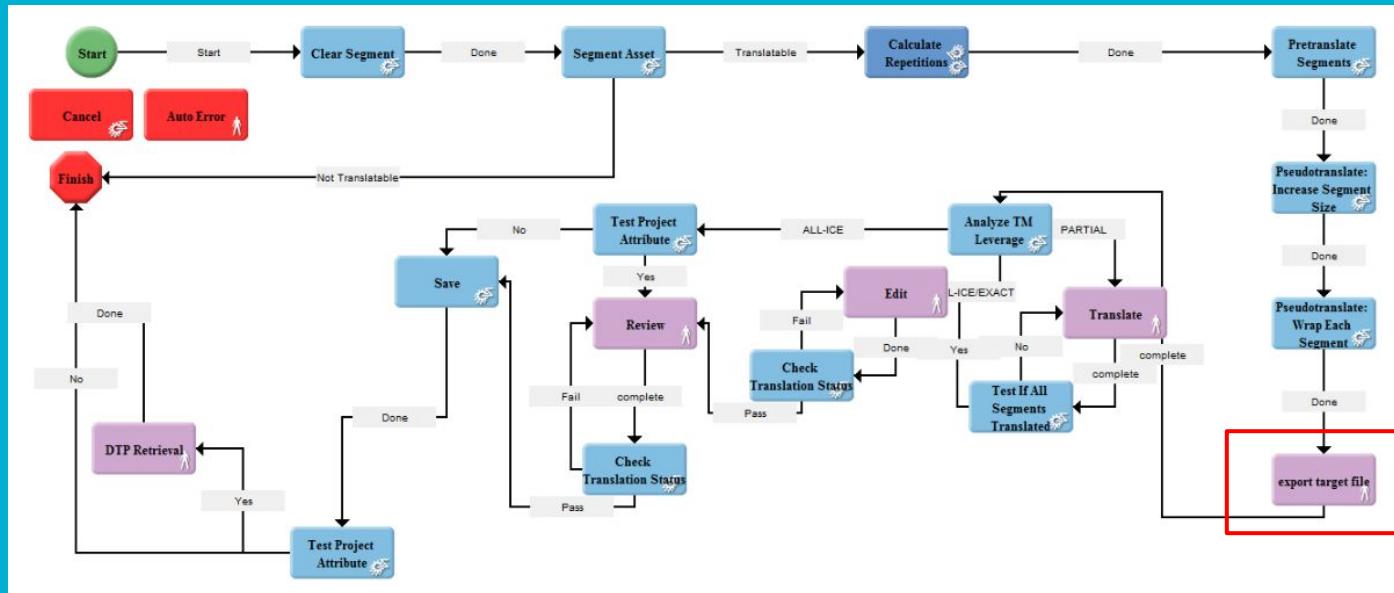
Tag Name	Include/Exclude
TITLE	Include
RSS	Exclude
DESCRIPTION	Include
CONTENT:ENCODED	Include

- ★ **DOC Filter: system default**
- ★ **XML Filter: customized filter for WP-exported xml**

Workflow setup

Tasks: Pseudo, Translation, Editing, Proofreading, DTP

Plan A: Combine all tasks into one workflow



Added a human step to retrieve the pseudo files before it gets pushed to the production steps

Workflow setup

Plan A: Issues

- Pseudo'd segments updated to the Translation Memory (Attr: “Update TM: No”)

Solution: Deleted the “Save” step that follows the pseudo steps.

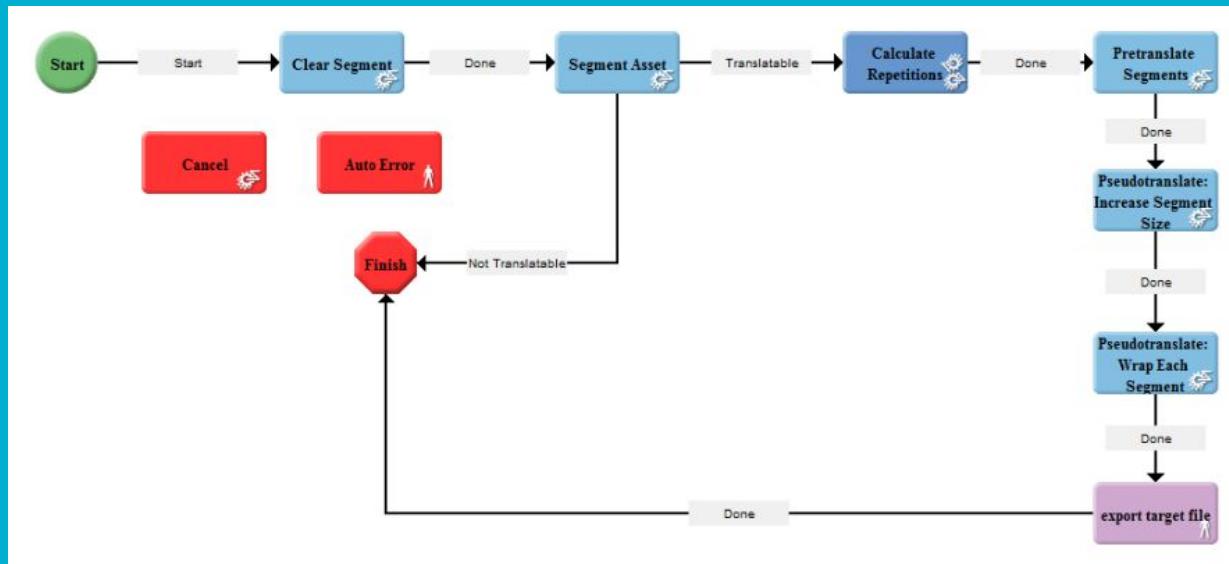
- Pseudo translated segments auto populated in the translation window.
 - > Consider cases where the asset is partially translated

Solution: Purge TMs, setup new workflows.

Workflow setup

Plan B: Separate tasks into two workflows

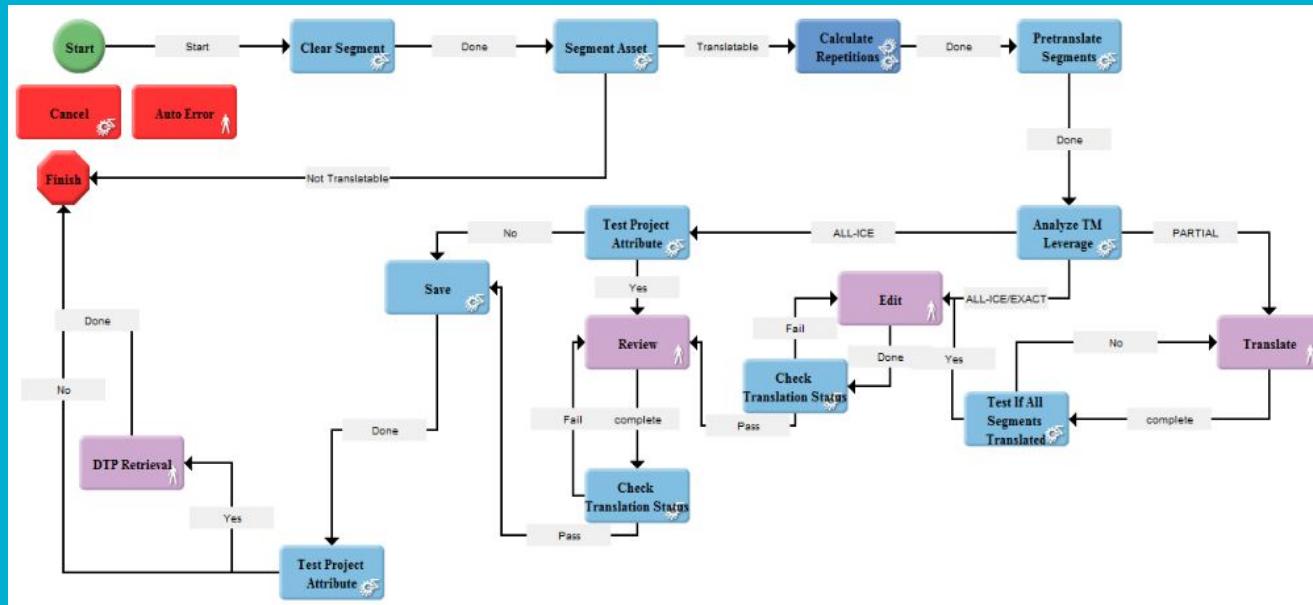
1) Pseudo translation workflow



- ★ 3 Pseudo Workflows
- ★ Increase segments
CN +20%
ES +30%
KO No expansion
- ★ Wrap segments
Prefix with [你üñはコ

Workflow setup

2) Production workflow (TEP + DTP Retrieval)



- ★ Analyze TM leverage for corresponding production phase.
- ★ Check translation status to make sure each step is completed.

Pseudo testing

ZH

[你üñは□Home+] [你üñは□About Us+] News [你üñは□Resources++]

[你üñは□Grant from Community Foundation++++++]

March 5, 2015 by hyerimko

[你üñは□On February 27, 2015 the Community Foundation for Monterey County awarded a grant to The Gathering Place for the purposes of Organizational Development and Transition to an independent Nonprofit status.++++++]

■ News

ES

[你üñは□Home+] [你üñは□About Us+++] News [你üñは□Resources++]

[你üñは□Grant from Community Foundation++++++]

March 5, 2015 by hyerimko

[你üñは□On February 27, 2015 the Community Foundation for Monterey County awarded a grant to The Gathering Place for the purposes of Organizational Development and Transition to an independent Nonprofit status.++++++]

■ News

KO

[你üñは□Home] [你üñは□About Us] News [你üñは□Resources]

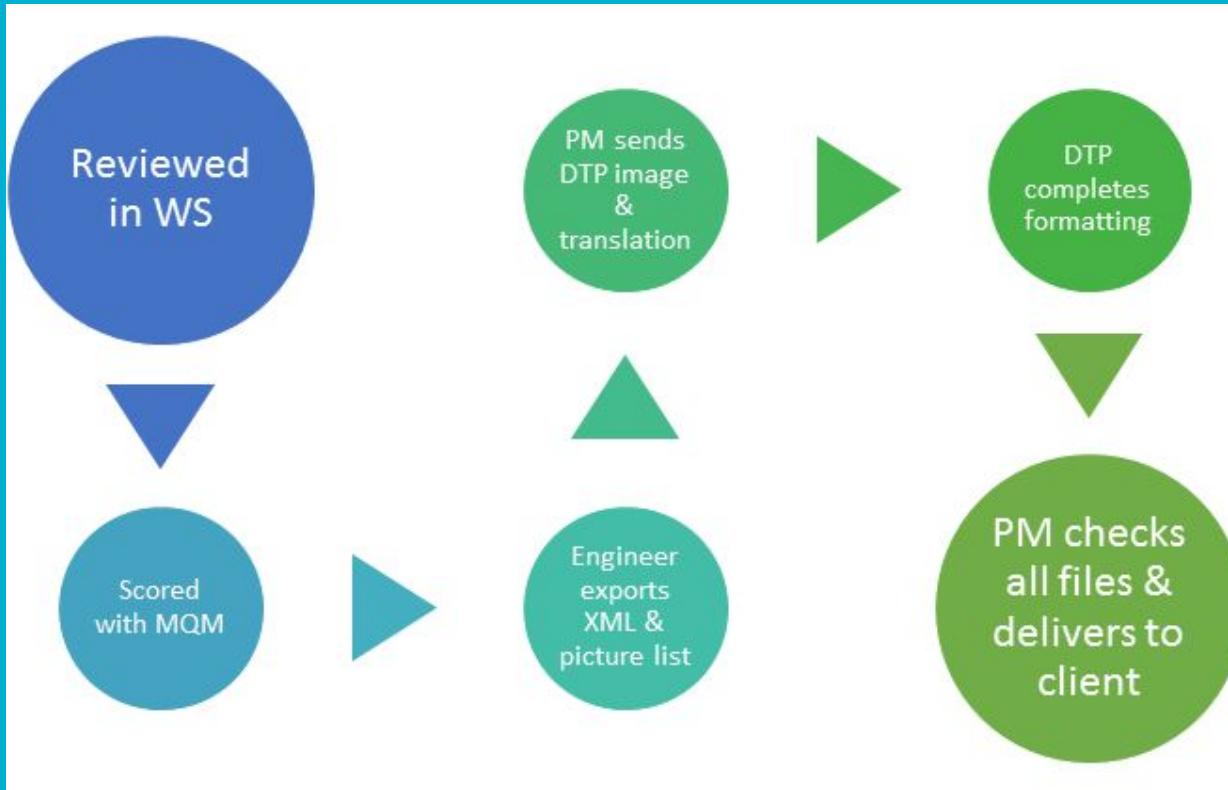
[你üñは□Grant from Community Foundation]

March 5, 2015 by hyerimko

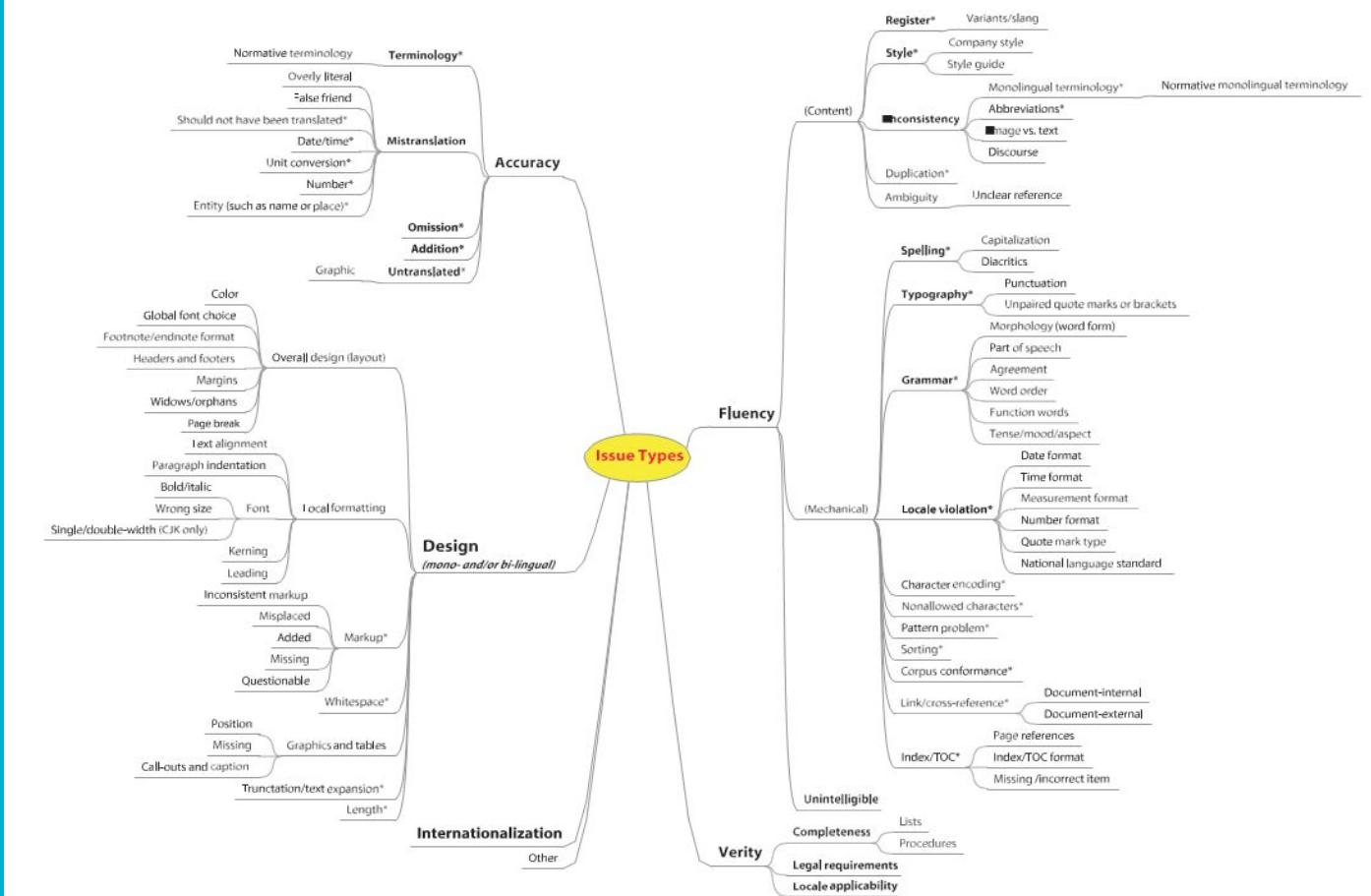
[你üñは□On February 27, 2015 the Community Foundation for Monterey County awarded a grant to The Gathering Place for the purposes of Organizational Development and Transition to an independent Nonprofit status.]

■ News

Finalization Phase



MQM Scoreboard



Source:
Multidimensional Quality Metrics
 (Lommel, Burchardt, Uszkoreit)

Information about the test		Severity multipliers		These values may be overridden by the user										
Word count (source):	231	1	Minor (default = 1)		<input type="checkbox"/> I am performing a source quality check									
Word count (target):	375	3	Major (default = 5)	5		0								
		4	Critical (default = 10)	10		4								
		Target Text Issues						Source Text Issues						
Issue Type		Weight (default = 1)		Issue counts			Penalty	Target Subscore		Issue counts			Penalty	
				Minor	Major	Critical	Raw	Adj.		Minor	Major	Critical	Raw	Adj.
Accuracy														Adjusted (Delta) Scores
<input type="checkbox"/> Bilingual Terminology														
<input checked="" type="checkbox"/> Mistranslation		5					0	0	100.0%					100.0%
<input checked="" type="checkbox"/> Omission		10					0	0	100.0%					100.0%
<input checked="" type="checkbox"/> Untranslated		10				1	1	10	97.3%					97.3%
<input type="checkbox"/> Addition														
Accuracy subtotal							10		97.3%					97.3%
Fluency														
<input type="checkbox"/> Unintelligible														
Content														
<input type="checkbox"/> Monolingual Terminology														
<input checked="" type="checkbox"/> Inconsistency		5		1			1	5	98.7%					98.7%
<input type="checkbox"/> Register														
<input checked="" type="checkbox"/> Style		5					0	0	100.0%					100.0%
Mechanical														
<input checked="" type="checkbox"/> Spelling		1					0	0	100.0%					100.0%
<input checked="" type="checkbox"/> Orthography		1					0	0	100.0%					100.0%
<input checked="" type="checkbox"/> Grammar		1		3			3	3	99.2%					99.2%
<input checked="" type="checkbox"/> Locale convention		5					0	0	100.0%					100.0%
Fluency subtotal							8		97.9%					97.9%
Verity														
<input type="checkbox"/> Fidelity														
<input type="checkbox"/> Completeness														
<input type="checkbox"/> Legal requirements														
<input type="checkbox"/> Locale applicability														
Fluency subtotal							0		100.0%					100.0%
TOTAL							18		95.2%					95.2%

Lessons learned

1. Meeting with the client

- a. Scope of the work
- b. Workflow
- c. Deliverables

2. Using WorldServer?

Thank You