Linport: Interoperable Translation and Localization Project Containers (A Progress Report)

Josep Bonet Heras & Olaf-Michael Stefanov
On behalf of Alan Melby and Tomas Carrasco Benitez
15-17 May 2013, Nairobi, Kenya
What is Linport? (this slide and 2 more)

- An acronym for Language INteroperability Portfolio
- Complete, interoperable container solution for both:
  - Translation projects (portfolio, n-lingual)
  - Tasks within projects (package = 1 project task)
an idea whose time has come...
... and a project

- With one main deliverable: a “blueprint”
  - Blueprint = draft standard to be submitted to an industry standards body
- And secondary deliverables
  - Open-source reference software for proof-of-concept and for implementers
- Hosted by LTAC Global (a non-profit organization)
- Open to all interested parties
- **Not** a competing translation tool
Why are we (Alan and Tomas) involved

• After working on various “payload” standards
  • e.g. TMX and TBX
• it is logical to work on containers to put them in

• To promote Structured Translation Specifications (STS) as part of the payload. STS is a set of
  › 21 parameters for quality translations
  › Found in: ISO/TS 11669
Structured Translation Specifications (STS)

- 21 parameters for quality translations:
  - Source
  - Target
  - Production
  - Environment
  - Relationships

- More info, can be found in: ISO/TS 11669
Why should JIAMCATT orgs be involved

- As major tool users and in cases such as the EC also as developers, we see and promote the value in a structured, interchangeable container for both project payload & specifications
- As multilingual organizations, large and small, we need to appreciate the complexity and difficulty in communicating project requirements
- ... and seek ways to reduce both complexity and difficulties
Source Parameters describe the source text

- [1] textual characteristics
  - (a) source language; (b) text type;
  - (c) audience; (d) purpose
- [2] specialized language
  - (a) subject field; (b) [monolingual] terminology
- [3] volume (words, characters, or other measure)
- [4] complexity (non-native author; graphics; etc.)
- [5] origin (where source text came from)
Target parameters state linguistic requirements

- [6] target language information
  - (a) target language; (b) [bilingual] terminology
- [7] audience (sometimes differs from source)
- [8] purpose (sometimes difference from source)
- [9] content correspondence
  - (summary/full; overt/covert; etc.)
- [10] register (often implied by text type)
- [12] style
  - (a) style guide; (b) style relevance
- [13] layout
Production parameters state tasks to be performed by whom

- [14] typical production tasks
  - (a) preparation
  - (b) initial translation (human or machine)
  - (c) in-process quality assurance

- [15] additional tasks
  - self-checking/post-editing; revision; review; final formatting; proofreading
Environment parameters

- [16] technology
  - (any interoperable tool? Particular tools?)
- [17] reference materials
  - Style guides, translation memories, etc.
- [18] workplace requirements
  - Anywhere? Designated secure facility?
Relationship parameters

• [19] permissions
  • (a) copyright; (b) recognition; (c) restrictions;

• [20] submissions
  • (a) qualifications; (b) deliverables; (c) delivery [method]; (d) deadline

• [21] expectations
  • (a) compensation; (b) communication
Webpage about the 21 parameters which make up **STS** – Structured Translation Specifications

www.ttt.org/specs
Container contents besides the structured specifications

- Source text(s)
- Target texts when done; or bi-text format
- Resources to carry out translation (per specs)
  - Terminology
  - Translation memory
  - Style guide
  - What else?

One group working on **What else**, esp. for MT, is the **Multilingual Web-LT** Working Group drafting the **ITS 2.0 (Internationalization Tag Set)** standard for submission to **W3C**, the World Wide Web Consortium.
Portfolios

- Can be split into bilingual packages
  - Each package is a **TIPP** (from IN!)
- A package is associated with one task
  - Initial translation (human or machine)
  - Revision of Human Translation (HT) or post-editing of raw MT
  - Review
  - etc.
- After tasks are performed
  - Response packages can be merged back into a revised portfolio
Why Linport is important to translation project managers

- No need to “re-package” data as it passes between translation tools
- No lost specifications
- Thus, increased efficiency and quality
Linport from the perspective of content owner/originator

- Can send out the same project portfolio when working with multiple translation providers (internal or external)
- A portfolio with specifications is useful for archiving and assessment even if the translation system is entirely on-line
- A project portfolio is not tied to any particular translation tool
Why it is important for translators

- Ensures that they will have access to the project specifications (also STS)
- Increases interoperability between tools
- Easier to use your favorite, or the most practical tool for the occasion because of data-tool separation
Why Linport is important for machine translation developers

- Standardized way for computers to extract the necessary data, perform machine translation, and return results
- Task object for Translation API
- Research challenge: MT systems may eventually be able to read the structured specifications and tailor the translation accordingly
Linport and Tool developers

- Some tool developers will be early adopters
  - Some will implement only the TIPP package
  - Others will also implement portfolios
- Most tool developers will wait until Linport is an industry standard
- At some point, customers will require Linport support in the tools they use or consider using
What’s upcoming for Linport?

- Refine & formalize portfolio data model
  - Later: authoring and publication portfolios
- More testing of STS builder
- Develop Portfolio Builder software
  - Round-trip: portfolio, task packages, portfolio
- Submit to standards bodies:
  - OASIS or ETSI or elsewhere (to be decided)
  - ISO (fast track after industry standard)
- Implementation by tool developers
What’s related?

- Quality Translation Launch Pad
  - An EC-funded project
- XLIFF:doc from IN! (automated workflow)
  - A profile of XLIFF version 1.2
- TAUS
  - Translation API
    - For transmission of portfolios and packages
How to get involved

- Join the Linport community
  - www.linport.org

- Share real translation project data
  - Urgently needed!
  - Must be non-confidential or sanitized

- Help with proof-of-concept software
  - Test apps developed at Brigham Young University
  - Develop your own reference software
  - Implement portfolios and/or packages in your tool as an early adopter
Questions?

Willing to submit project data?
--> info@linport.org
Glossary of acronyms used

- **DGT** – Directorate General for Translation
  - A part of the EC; JIAMCATT Partner
  - [ec.europa.eu/dgs/translation](http://ec.europa.eu/dgs/translation)
- **EC** – European Commission
- **ETSI** – European Telecommunications Standards Institute
  - [www.etsi.org](http://www.etsi.org)
- **IN!** – Interoperability Now!
  - Group working to improve the interoperability of tools and technology within the localization industry
  - [code.google.com/p/interoperability-now](http://code.google.com/p/interoperability-now)
- **ISO** – International Standards Organization
  - JIAMCATT Partner
  - [www.iso.org](http://www.iso.org)
- **Linport** – The Language Interoperability Portfolio Project
  - [www.linport.org](http://www.linport.org)
- **LISA** - Localization Industry Standards Association
  - Ceased to exist March 2011
- **OASIS** - Organization for the Advancement of Structured Information Standards
  - [www.oasis-open.org](http://www.oasis-open.org)
  - **QTLP → QTLaunchPad →** Quality Translation Launch Pad.
- **TAUS** – Translation Automation
  - [www.translationautomation.com](http://www.translationautomation.com)
- **TIPP** – Translation Interoperability Protocol Package – an **IN!** project
- **XLIFF** – XML Localisation Interchange File Format
  - XLIFF 1.2: [docs.oasis-open.org/xliff/xliff-core/xliff-core.html](http://docs.oasis-open.org/xliff/xliff-core/xliff-core.html)