Linport: toward an open standard for language projects and translation tasks

Alan Melby and Tyler Snow
What is Linport? (this slide and 2 more)

- An acronym for Language Interoperability Portfolio
- Complete and 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
Having worked on various “payload” standards (such as 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:
- 21 parameters for quality translations:
  - Source
  - Target
  - Production
  - Environment
  - Relationships
- Found in: ISO/TS 11669
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 did this source text come 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
    • self-checking/post-editing; revision; review; final formatting; proofreading

[15] additional tasks
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: 21 parameters

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 MultilingualWeb-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 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 service 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
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, service requesters require Linport support in the tools they use or consider using
Why it is important for translators

- Ensures that you will have access to the project specifications (also STS)
- Increases interoperability between tools
- Easier to use your favorite tool because of data-tool separation
Why it is important for translator trainers

- Students learn translation project elements and workflow scenarios
- Structured Translation Specifications are basic to any discussion of translation quality
- All of this in a tool independent learning environment
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 system someday be able to read the structured specifications and tailor the translation accordingly*
What’s upcoming

- Refine & formalize portfolio data model
  - Later: authoring and publication portfolios
- More testing of STS builder
- Develop Portfolio Builder software
  - Roundtrip: portfolio, task packages, portfolio
- Submit to standards bodies:
  - OASIS or ETSI (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)
  › Hopefully a true profile of XLIFF version 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 U.
  - Develop your own reference software
  - Implement portfolios and/or packages in your tool as an early adopter
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!
  - A 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** → **QTLlaunchPad** → 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)
Questions?

Willing to submit project data?
info@linport.org