

XML and Localization

Strategies and Best Practices

Scott Bass

President

Advanced Language Translation Inc.

Rochester, NY

www.advancedlanguage.com

- **Happy Birthday to XML! ...Version 1.0 turned five years old on Monday (2/10/03)**
 - See <http://www.w3.org/2003/02/xml-at-5.html>
- **XML is the new wave (for the last five years!)**
- **XML is a standard, not an application**
 - Converting to XML isn't like switching from MS Word to Framemaker
- **Commitment in time, learning and cost to adopt XML for publishing can be high...returns can be great**

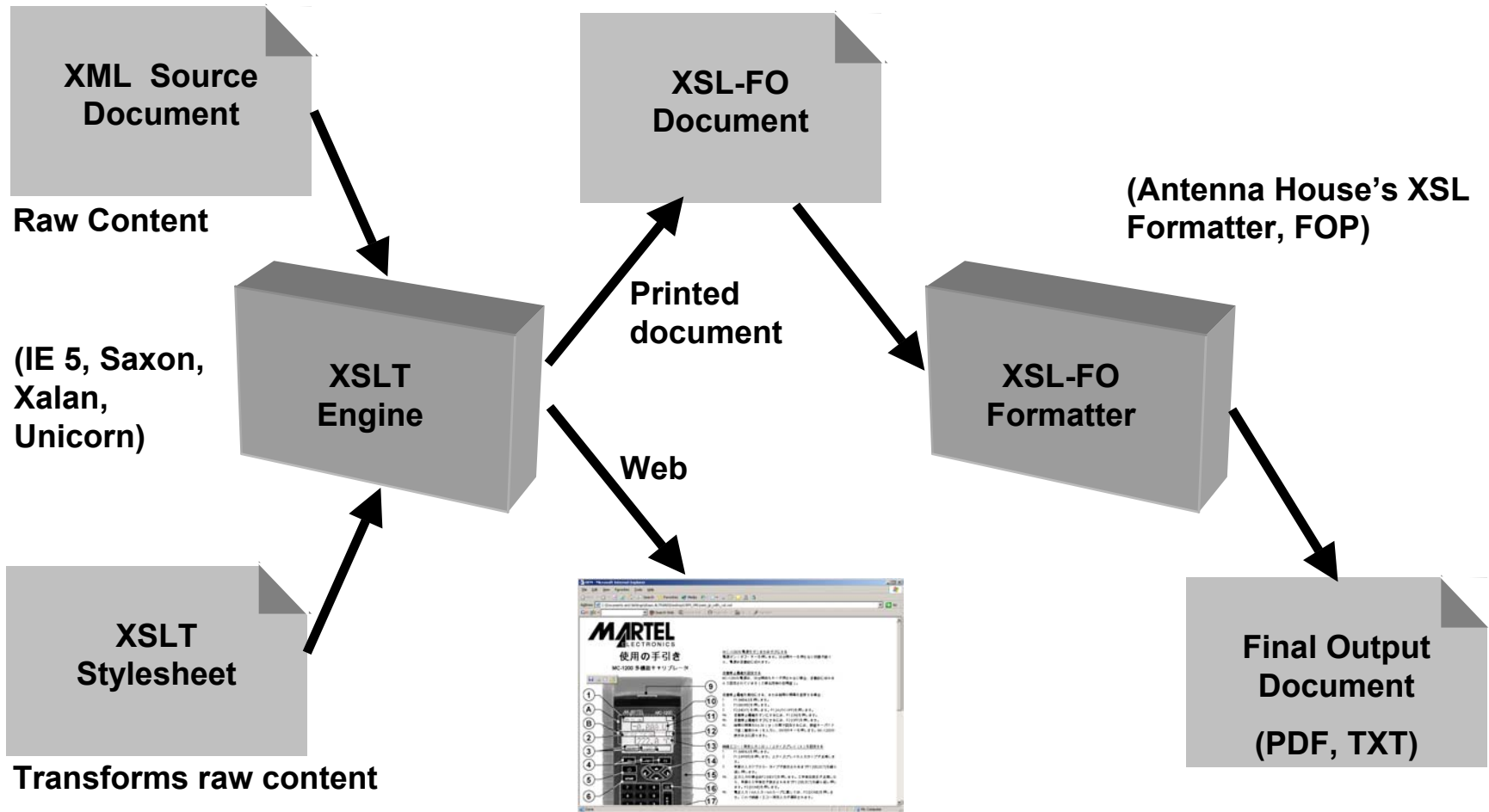
- **Every authoring tool uses markup language...it's usually proprietary**
 - (If you're ambitious try re-writing a Quark document using XTG—Quark's ASCII tagged file format)
- **HTML is a non-standard “standard” for documents on the web**
- **XML is a true standard and can be used for alpha-numeric content, images and voice (VoiceXML)**

- **Easily transform and display content regardless of platform or language!**
- **UTF-8 Encoding is standard...**
- **...But you can use any encoding you want**
- **Facilitates reuse of content (a.k.a. multi-purposing of content)**
- **Easier creation of dynamic content**

- **XML facilitates separating content from form/structure**
- **Great format for storing content**
- **Like having a database without the database and connection problems (an issue for web pages)**
- **It doesn't care what language your content is in—language just becomes another attribute**

- **XML – Extensible Markup Language**
- **XSL – Extensible Stylesheet Language**
- **XSLT – XSL + Transformation**
 - Transform your content to fit its final destination(s)
- **XSL-FO – XSL + Formatting Objects**
 - for creating printable docs
- **XPath – XML document that helps select content**
- **SVG – Scalar Vector Graphics**
 - Embed dynamic text within graphics

XML for Docs and Web Pages



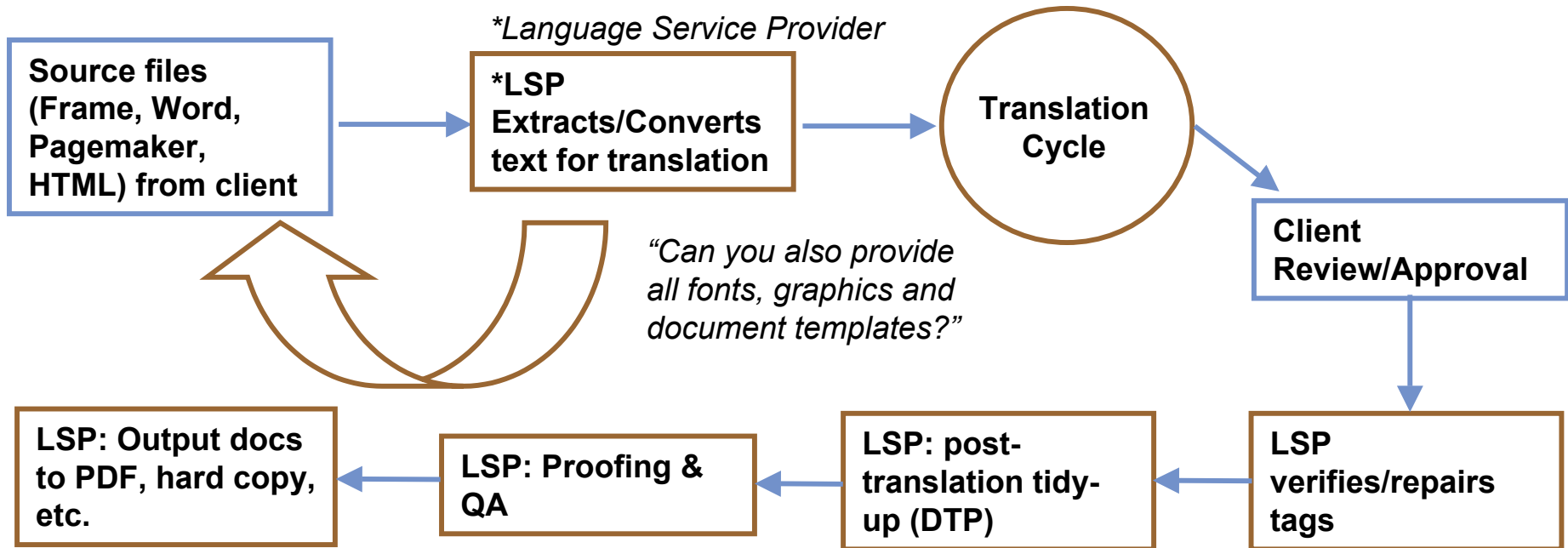
- **Conversion of a simple instruction sheet to be displayed on the web**
 - Step 1: Create content XML file (15 min.)
 - Step 2: Create XSLT file (1 hour)
 - Step 3: Translate XML file into Japanese (text placement only – 15 min.)

**Accomplished by an advanced XML user with intermediate skills in XSL.*

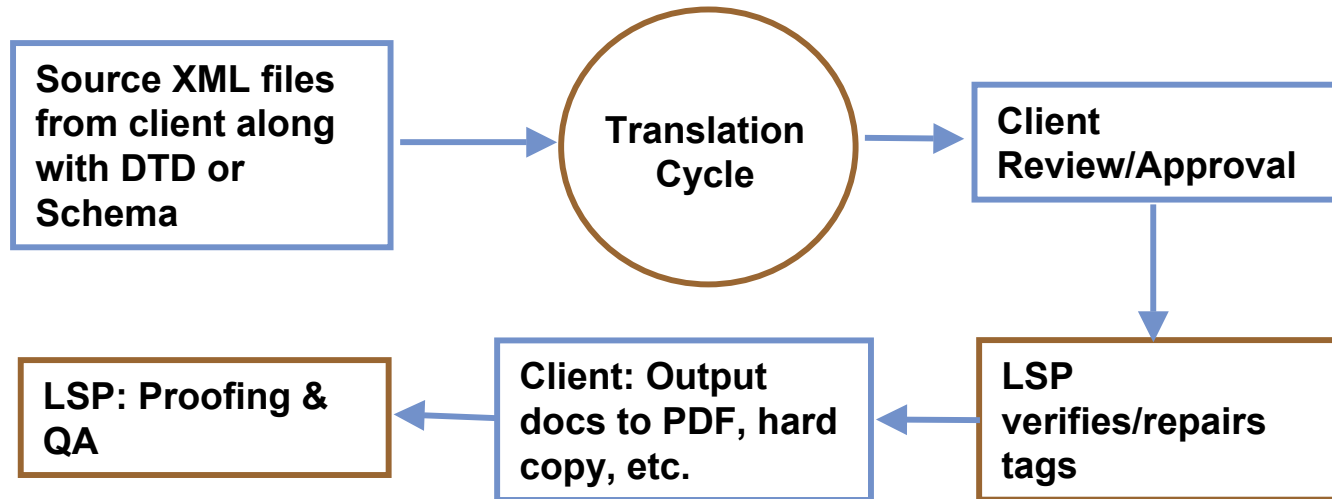
- **Content is separated from “format”**
 - Translation moves faster
 - Lower pre- and post-translation setup/processing costs
- **Standardized translation process regardless of document type**
- **Reduced desktop publishing costs for translated documents**
- **Better reuse strategies during authoring will reduce translation costs**

- **Some “Cons”**
 - Translation lacks context, reference documents are required.
 - XML internationalization must be precise and complete BEFORE translation starts
 - Greater burden on document management processes and infrastructure, since content is now more “dynamic”
 - Graphics with embedded text are still an issue (SVG can be a solution)

Translation Process: Traditional Workflow



- This process typically utilizes email or FTP to transfer files.
- It requires manual tracking of the movement of files.
- There is no automatic logging of activity.
- Difficult to track project metrics for large projects.

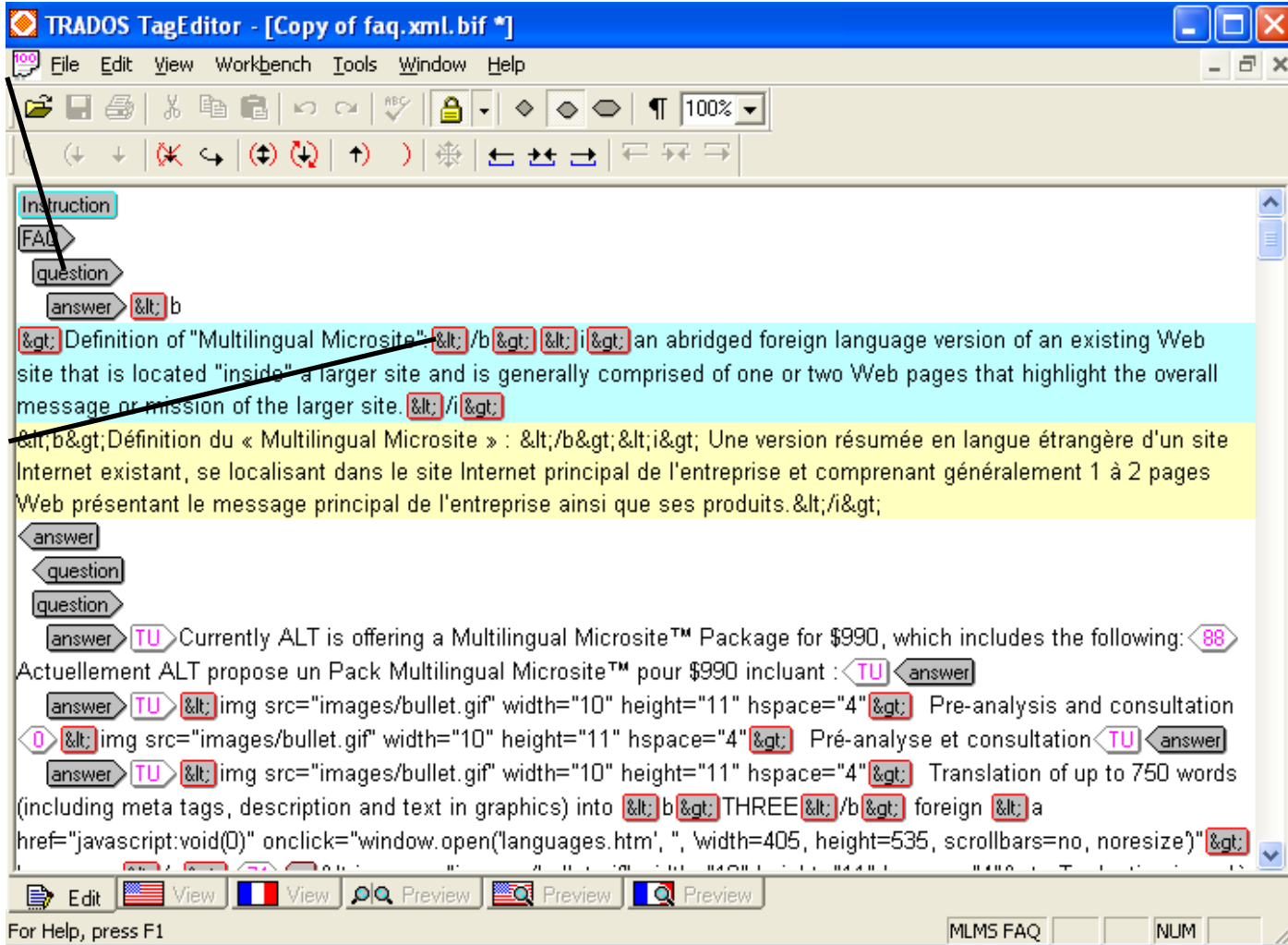


- **This process allows for integration of the translation cycle into existing CMS operations. (e.g. SOAP and SQL triggers can automate the movement of XML files)**
- **Integration into CMS enables automated tracking of files/content.**
- **Project management burden can be diminished for both client and LSP PM's.**

Translation Process: Translating XML

External
XML tags
are
truncated to
reduce
clutter

Internal tags
appear and
can be
placed in
the
translation



TRADOS TagEditor - [Copy of faq.xml.bif *]

File Edit View Workbench Tools Window Help

Instruction
FAT
question
answer Definition of "Multilingual Microsite" <i > an abridged foreign language version of an existing Web site that is located "inside" a larger site and is generally comprised of one or two Web pages that highlight the overall message or mission of the larger site. </i >
 Définition du « Multilingual Microsite » : <i > Une version résumée en langue étrangère d'un site Internet existant, se localisant dans le site Internet principal de l'entreprise et comprenant généralement 1 à 2 pages Web présentant le message principal de l'entreprise ainsi que ses produits. </i >
answer
question
question
answer TU > Currently ALT is offering a Multilingual Microsite™ Package for \$990, which includes the following: <88
Actuellement ALT propose un Pack Multilingual Microsite™ pour \$990 incluant : <TU <answer
answer TU Pre-analysis and consultation
<0 Pré-analyse et consultation <TU <answer
answer TU Translation of up to 750 words
(including meta tags, description and text in graphics) into THREE foreign

Edit View View Preview Preview Preview

For Help, press F1 MLMS FAQ NUM

Is it (about) time for XML?

- **Is XML for your company?**
- **Is your current process broken?**
- **Are your needs enough for a full XML implementation?**
- **Is it the right time for XML?**

When do you make the leap to XML?

- **Reuse of text for use across various media is high**
- **Content is dynamic**
- **Structure lends itself to “granulating” information**
- **High publishing volumes of repetitive information**
- **Got a CMS? Then get XML.**

- **Redesigned web site to be fully XML enabled.**
- **Whole site was driven by a single .asp page**
- **Translated content for a given topic stored in same file with all other languages**
- **No limit to the number or complexity of the languages**

- **Site speed and efficiency increased**
- **Localization costs for the site decreased dramatically....**

- **Site could NOT be read by web spiders nor indexed by search engines**
- **Result: a wonderfully dynamic, state-of-the-art site that no one could find!!!**
- **Lesson learned: You can be too smart ...**



Advanced Language Translation Inc.

Cascade Center

25 N. Washington St.

Rochester, NY 14614, USA

Within North America: 1-800-218-9024

Outside the USA: 585-697-0462

Fax: 585-697-0467

e-mail: info@advancedlanguage.com

www.advancedlanguage.com