GALA 2013 MIAMI
THE LANGUAGE OF BUSINESS
THE BUSINESS OF LANGUAGE
Miami Beach
17-20 March 2013

Applying ITS 2.0 to Online MT Systems in HTML5 (MLW-LT W3C)

Pedro L. Díez Orzas (Linguaserve)
Arle Lommel (DFKI)
Daniel Grasmick (Lucy Software)
Pablo Badía Mas (Linguaserve)

@gala_global
#galaconf

Contents

1. Presentation
2. The state of ITS 2.0
3. Online MT System: the MLW-LT use case
4. Machine Side: Sample Applications
5. Human Side: Tasks and Processes
6. Discussion

GALA 2013 MIAMI

THE LANGUAGE OF BUSINESS
THE BUSINESS OF LANGUAGE

Miami Beach
17-20 March 2013

Presentation: MultilingualWeb-LT, ITS 2.0 and HTML5

Brief presentation of subject, goals and speakers

Pedro L. Díez Orzas (Linguaserve)

@gala_global
#galaconf

Outline

- The goal of this session is to arouse interest and to encourage participation about the upcoming:

Internationalization Tag Set 2.0

- ITS 2.0 web metadata standards will directly affect the web globalization and localization sector.

- In the presentation, each speaker will present a different aspect to go from normative to practice.

MultilingualWeb-LT and ITS 2.0

- MultilingualWeb-LT is funded by the European Commission and Administrated by the W3C.
- Part of the project is the Real-Time Multilingual Publication System (RTMPS) use case:

  Content I18N and Advanced Machine Translation
- Project Coordinated by DFKI
- The MLW-LT consortium members of the use case are:

More info about MLW-LT

www.w3.org/International/multilingualweb/lt/wiki/Main_Page

www.multilingualweb.eu

GALA 2013 MIAMI

THE LANGUAGE OF BUSINESS
THE BUSINESS OF LANGUAGE

Miami Beach
17-20 March 2013

The state of ITS 2.0

Arle Lommel (DFKI)

@gala_global
#galaconf

Standardization in W3C

- Internationalization Tag Set 2.0 is developed by the MultilingualWeb-LT Working Group.
- Public requirements gathering in early 2012
- Drafts published starting in July 2012 for public feedback
- Final Call status began December 6, 2012
- Currently editing to reflect feedback

Standardization in W3C

- Implementation-driven approach (more later): all items require *at least* two implementations
- Aggressive schedule to avoid feature creep and slow pace

Status of ITS 2.0 recommendation

- Final call is completed, but feedback is still welcome
- Working draft available: http://www.w3.org/TR/its20/
- Current process requires response to *all* comments (hundreds)
- Final version scheduled for Q4 2013

ITS 2.0 Data Categories (I)

1. Translate*
2. Localization Note*
3. Terminology
4. Elements Within Text
5. Domain*
6. Disambiguation

ITS 2.0 Data Categories (II)

7. Language Information*
8. Directionality
9. Ruby
10. Locale Filter
11. Translation Agent Provenance*
12. External Resource
13. Target Pointer

ITS 2.0 Data Categories (III)

14. Id Value
15. Preserve Space
16. Localization Quality Issue*
17. Localization Quality Précis
18. MT Confidence**
19. Allowed Characters
20. Storage Size

Status of HTML5

- HTML5 and XML are explicit targets of ITS 2.0
- HTML5 recognizes the *translate* data category natively.
- All others categories implemented with its- prefix: requires HTML5+ITS parser/validator (to be added to validator.nu)
- Allows use of HTML5 as authoring format

ITS 2.0 HTML5 implementations

- Quality markup (VistaTEC)
- CheckMate for quality markup.
- Online MT system internationalization (topic today)
- CMS content
- Many others use HTML5 as the target format

Use case overview (I)

1. Simple Machine Translation
2. Translation Package Creation
3. **Quality Check**
4. **Processing HTML5 documents with XML tool chain**
5. **Validation: HTML5 with ITS 2.0 metadata**

Use case overview (II)

6. **CMS to TMS System, with Drupal and Linguaserve's GBC Server**
7. **Online MT System Internationalization with ATLAS PW1 (presented here)**
8. Using ITS for PO files
9. **Browser-Based Review**
10. **Simple Segment Machine Translation**

Use case overview (III)

11. **HTML-to-TMS Round-trip Using XLIFF with CMS-LION and SOLAS**
12. **CMS Implementation of ITS**
13. CMS-Level Interoperability Using CMIS
14. **Annotation of Named Entities**

ITS 2.0 for Web L10N and G11N

- Persistent metadata in Web pages: integrates authoring, localization, and publication
- Unification of HTML and XML workflows (currently separate)
- Makes HTML5 into a “smart” format
- HTML5 group currently working with MLW-LT to define default behaviors for ITS 2.0 in HTML5

How can you contribute?

- Review the specification and provide feedback
- Try out ITS 2.0 implementations in open-source Okapi framework
- Implement relevant ITS 2.0 metadata in your workflow or ask your tools developer to implement it

GALA 2013 MIAMI

THE LANGUAGE OF BUSINESS
THE BUSINESS OF LANGUAGE

Miami Beach
17-20 March 2013

Online MT System: the MLW-LT use case

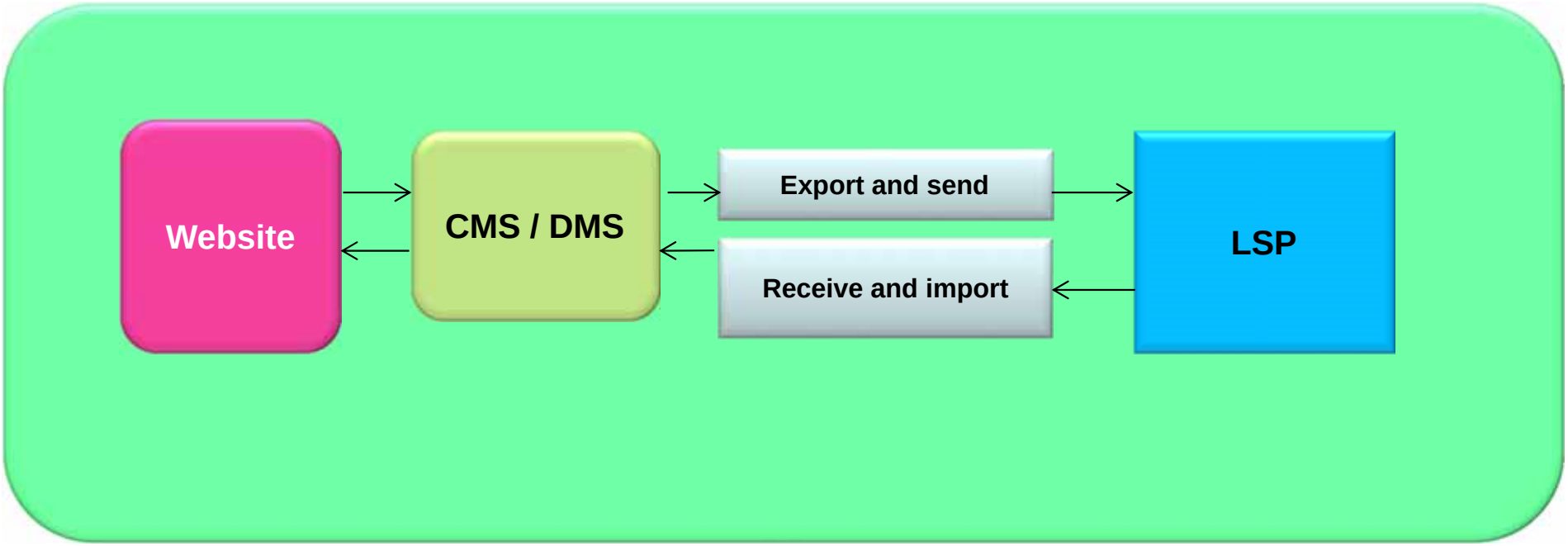
Pedro L. Díez Orzas (Linguaserve)

@gala_global
#galaconf

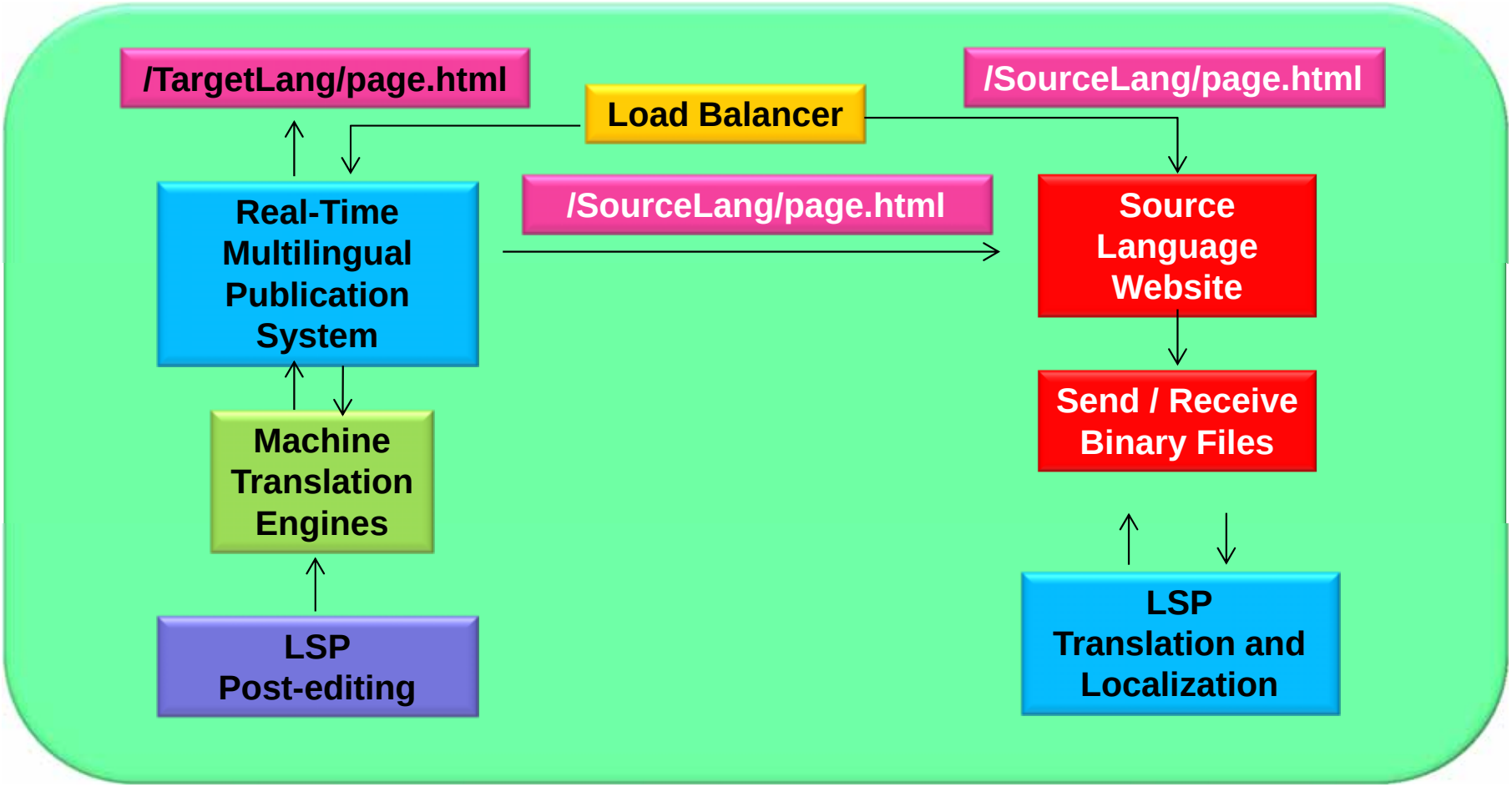
Outline

- Exemplifies how ITS allows an HTML5 Content Author to send instructions about the translation to MT Systems and to a Content Post-Editor through a Real-Time Multilingual Translation and Publication System (RTMPS).
- This RTMPS is connected to different MT Service Providers.

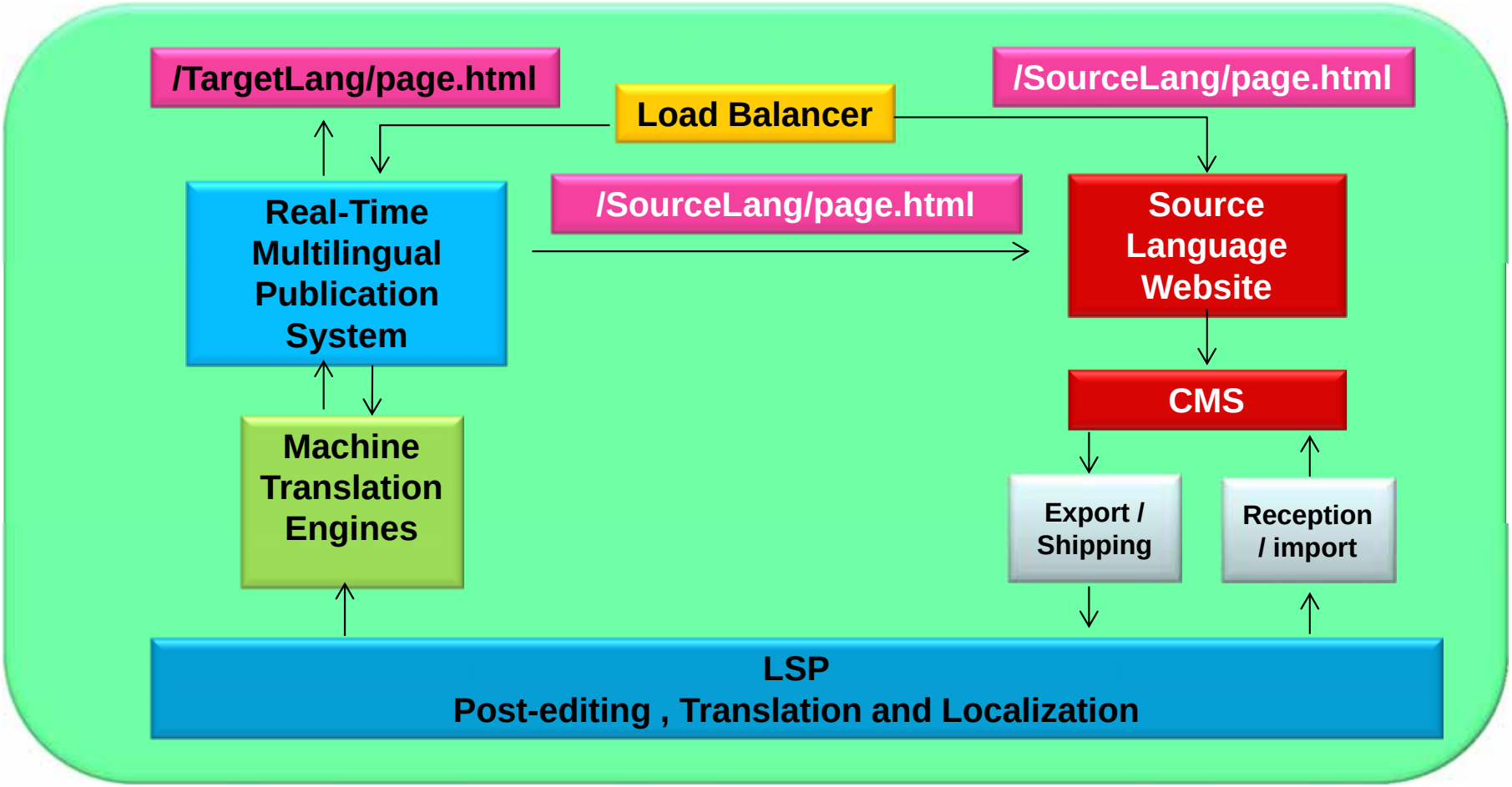
Off-line translation system (OTS)



Real-Time multilingual publication system (RTMPS)



Combined multilingual web translation and publication system

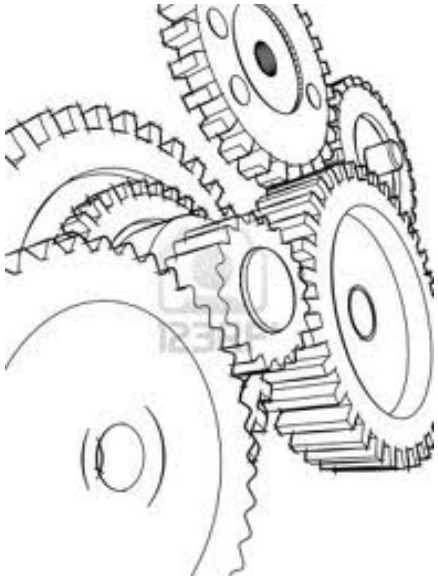


Spanish Tax Agency website

- Spanish Tax Agency
  - Effective application of Spain's tax and custom system
  - Management of tax resources on behalf of other public administrations when required by Law or Agreements
- www.agenciatributaria.es uses a combined multilingual web translation and publication system (OTS and RTMPS)
- More than 5 million of words
- ES, EN, CA, GL, VA
- Machine translation and post-editing.

RTMPS: MLW-LT use case

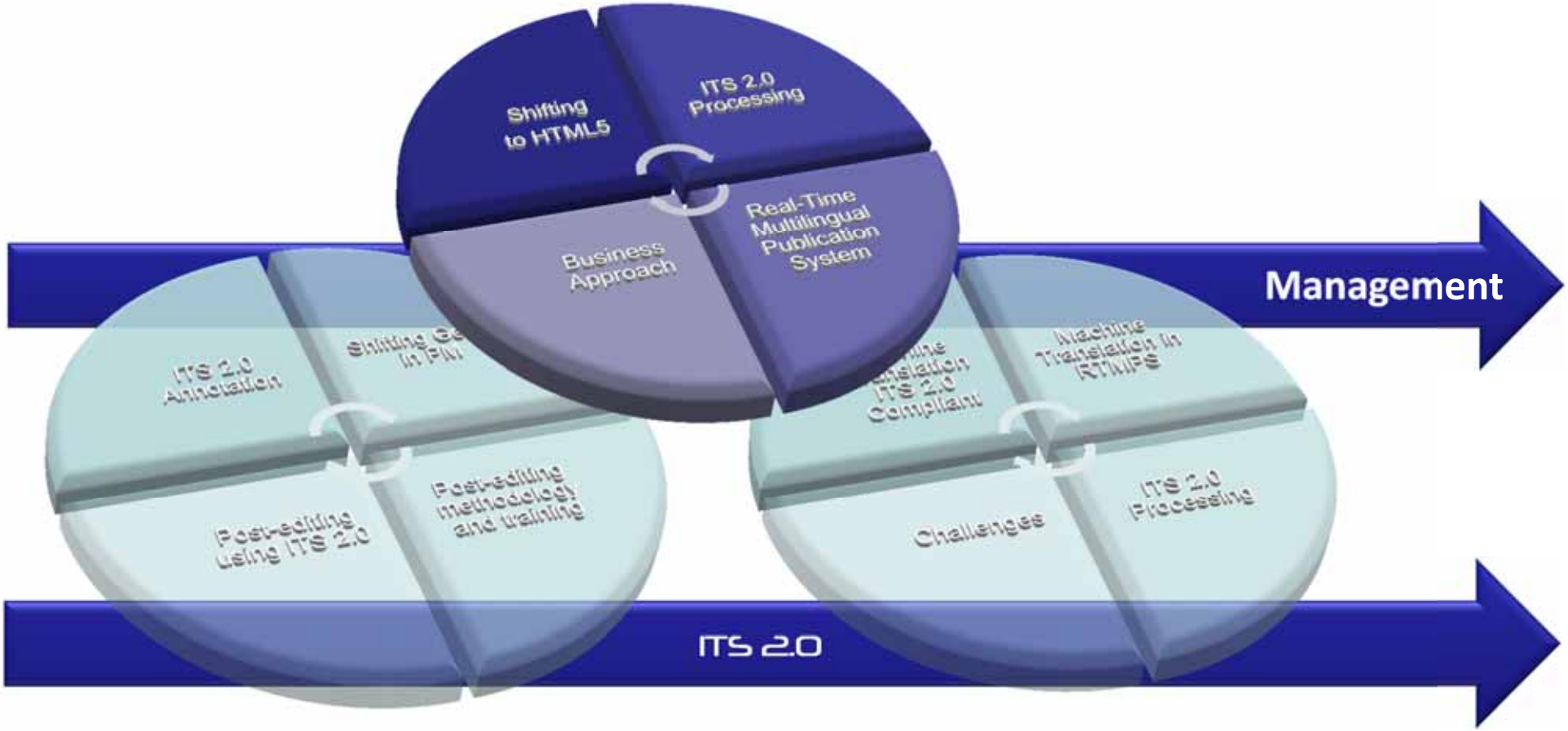
- Spanish Tax Agency, user in the use case RTMP System in MultilingualWeb-LT.
- Use case ingredients:
  - Multilingual www.agenciatributaria.es
  - HTML5
  - ITS 2.0
  - Real-Time Multilingual Publication System
    - ATLAS (Linguaserve's Real-Time Translation System)
    - Lucy Software MT (Rule Based MT)
    - MaTrEx from Dublin City University (Statistical MT)



State of use case (March 2013)

- RTMPS Implementation
  - Previous version (non ITS 2.0) in production since 2010.
  - ITS 2.0 Prototype 100% (ITS 2.0 definition from December 2012)
  - Showcase: pre-production demo
  - ITS 2.0 data categories: 6 (Translate, Localization Note, Language Information, Domain, Provenance, Localization Quality Issue)
- ES-EN total scope: 250 web pages. State:
  - Source language: 30% of target
  - Target language and Post-editing: 30% of target
- ES-FR, ES-DE total scope: 30 web pages. State:
  - Source language: 50% of target
  - Target language and Post-editing: 50% of target
- Testing: Pending

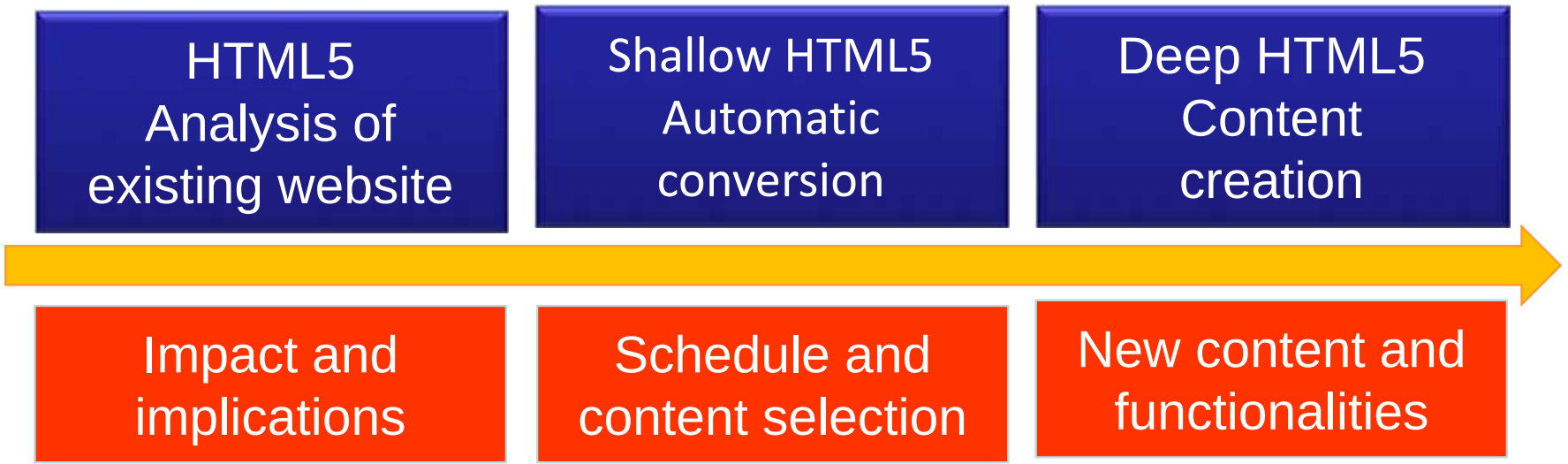
Real-Time Web I18N System use case in MLW-LT



GALA 2013 MIAMI
THE LANGUAGE OF BUSINESS
THE BUSINESS OF LANGUAGE
Miami Beach
17-20 March 2013
MultilingualWeb-LT

(1) Shifting to HTML5: Strategy

- Using ITS 2.0 for HTML requires version 5 according to the current W3C specification.



(1) Shifting to shallow HTML5: Modifications

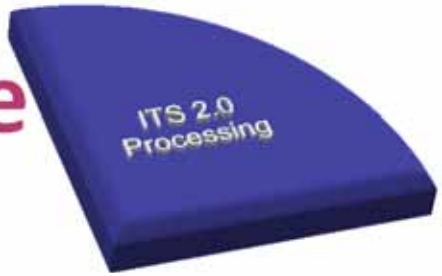
- HTML5 DOCTYPE
- The language page (ISO 639-ISO 3166)
- Self-closed tags not allowed
- Head tags
- Erroneous nesting tags
- Attributes separated by spaces
- Non inclusion of presentation attributes in tags
- Header and body structure needed by tables

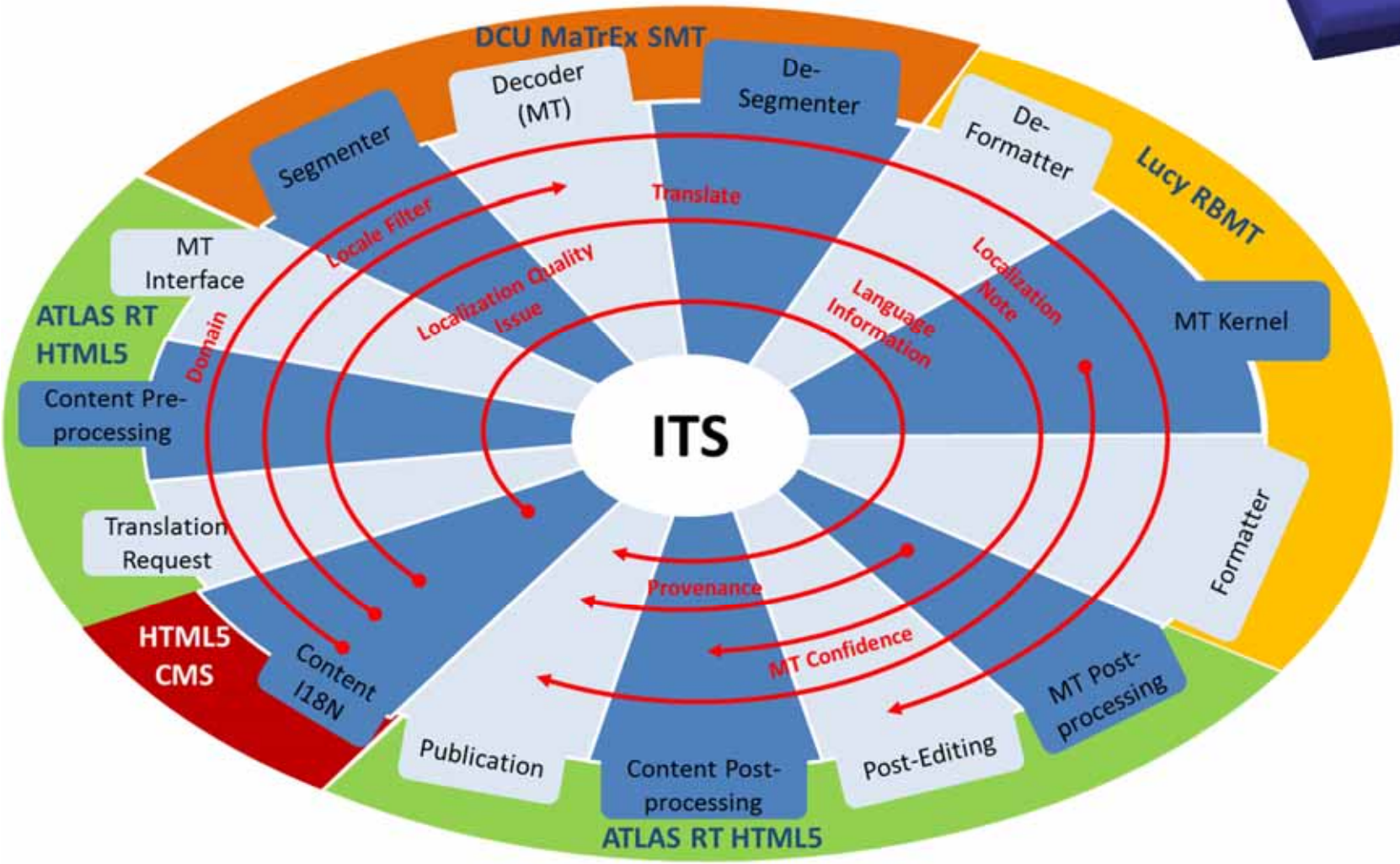
- HTML entities instead of special characters
- URLs cannot contain special characters
- ID attribute cannot contain spaces
- Required attributes (e.g. tag "object" must always have the attributes "data" and "type")
- Assessed attributes (e.g. "rel" attribute of tags "a" and "link" must be one from a closed list)

(1) Shifting to shallow HTML5: Obsolete attributes

Tags	Impact
input	Removed the alt attribute from any input tag that does not contain the attribute "type = 'image'"
div	Cannot define a "name" attribute in a "DIV“ tag
a	Not allowed to define the attributes "name" and "title" in tag "a"
embed and object	Cannot define the attributes: •"Applet" in the "embed" and "object“ tags •"Name" in the "embed“ tag •"Code", "archive", "classid", "codebase", "codetype", "state“ and "standby" in the "object“ tag
table	Not allowed to define the attributes "summary" and "border" in the "table“ tag
img	Not allowed to define the attributes "name" and "border" in the "img“ tag
option	Cannot define the attribute "name" in the "option“ tag.
param	Not allowed to define the attributes "type“ and "valuetype" in the "param“ tag
script	Not allowed to define the attribute “lang” except in “JavaScript", it being case-insensitive in the tag "script"
br	Cannot define the attribute “clear” in the “br” tag
background attribute	No attribute is used to define the "background" in the tags "body", "table", "thead", "tbody", "tfoot", "tr", "td" and "th".

(2) ITS 2.0 usage along the use case





(2) ITS 2.0 data categories: global use (I)

- Translate
  - <its:translateRule selector="//h:*/@title" translate="yes"/>
  - <its:translateRule selector="//h:*/@alt" translate="yes"/>
- Domain
  - <its:domainRule selector="//h:body" domainPointer="/html/head/meta[@name='keywords']/@content" domainMapping="'Economy and Trade' ECON, 'Electronic procedures' ECON, 'Public Administration' ECON, 'Electronic Administration' ECON, Treasury ECON , Taxation ECON, Trade ECON-USER, 'Spanish Tax Scheme' ECON, 'Political Economics' ECON-POL, 'General Vocabulary' GV, 'Non-residents' ECON, Fraud ECON, 'Public institutions' POL"></its:domainRule>

(2) ITS 2.0 data categories: global use (II)

- Language Information
  - <html lang="es">
- Provenance
  - <title its-org="Lucysoftware" its-tool="MTEngine 6.9" its-rev-person="Johanna Blasco" its-rev-org="Linguaserve" its-rev-tool="Transit XV">

Global rules can be internal to the HTML5 file or loaded from an URL

(2) ITS 2.0 data categories: local use

- Translate
  -
- Localization Note
  - STR
- Localization Quality Issue
  - decalrar

(2) ITS 2.0 data categories: processing

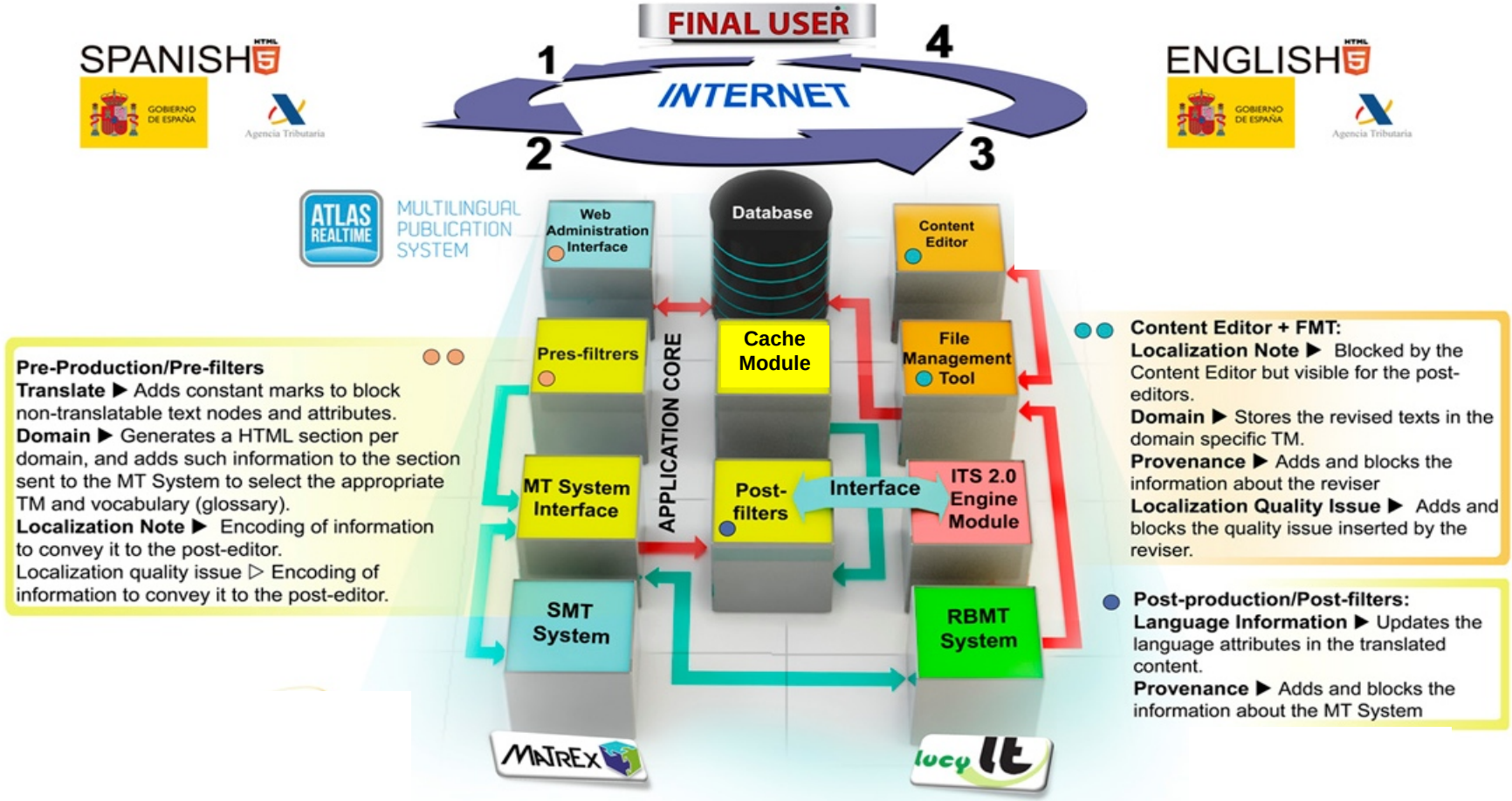
- MaTrEx prototype
- ITS 2.0 LucySoftware prototype
- ATLAS Real-Time: ITS 2.0 prototype
- **ATLAS Real-Time: ITS 2.0 Testing**
- **Spanish Tax Agency Showcase**

Access to demos:

http://www.w3.org/International/multilingualweb/lt/wiki/Use_cases_-_high_level_summaryITS 2.0

(3) ITS 2.0 application to a Real-Time Multilingual Publishing System

www.agenciatributaria.es



(3) ITS 2.0 Use case benefits and interactions

- ITS 2.0 Increases user's control and automatic decision processes:
  - Translatability and language pair selection (Translate, Language information)
  - Specific terminology to apply (Domain)
  - Activation rules for post-editing (Localization Note)
  - Quality aspects reported to translation consumer or post-editor (Localization Quality Issue)
  - Post-editors judge quality of translation (MT Confidence)*
  - Identification of agents (Provenance)

(4) RTMPS SWOT

Strengths

RTMPS highly reduces:

- Translation costs (Quality on-demand = MT + post-editing %)
- Management costs
- Delivery time
- Technical: clients do not need to install anything

Weaknesses

Viability dependent on :

- Language combination
- MT system output
- Pre-editing and post-editing methodologies and tools (ITS 2.0 and HTML5 compliance)

Opportunities

Profitability:

- Websites with more than half a million words
- Websites with a very high update frequency

Threats

Control, performance and security:

- The client might lose control of the translation
- Real-Time performance
- Security level

ITS 2.0

(4) Online MT Business Case SWOT

Strengths

Reduction of costs:

- **Translation (MT + post-editing):**
  - **100% post-edited: - 30%**
  - **Depending on % of post-editing cost reduction increases.**
- **Management client side: -80%**
- **Non-invasive technology**
- **Real-Time and fast post-edition**

Weaknesses

Profitability:

- **Sites with more than 4 million words**
- **Several updates per week**
- **Specific Pre-editing and post-editing tools needed**

Opportunities

Viability :

- **ES<>EN, ES<>CA, ES<>GL (and ES<>FR, ES<>PT)**
- **New Project Management approach**
- **"EDI-TA" methodology/training**

Threats

Control, performance and security:

- **ITS 2.0**
- **ATLAS RT Cache**
- **In-house hosting**

(4) Business opportunities rise from needs

- E-commerce
  - Very high volume and rotation
  - Short texts and repetitive descriptions
    - Better for MT
    - Quicker to post-edit
  - Very sensitive to ITS 2.0 benefits
  - Content source independent (HTML from several CMS and other applications)
- Web 2.0 (user content created)
  - GIST translation
  - Immediacy

(4) ITS 2.0 Awareness

- Customers
  - Qualitative and quantitative advantages
- CMS developers and integrators
  - Product adaptations and extensions
- Content creators
  - Training, pre-editing tools and best practices
- LSPs
  - Compliant tools and best practices

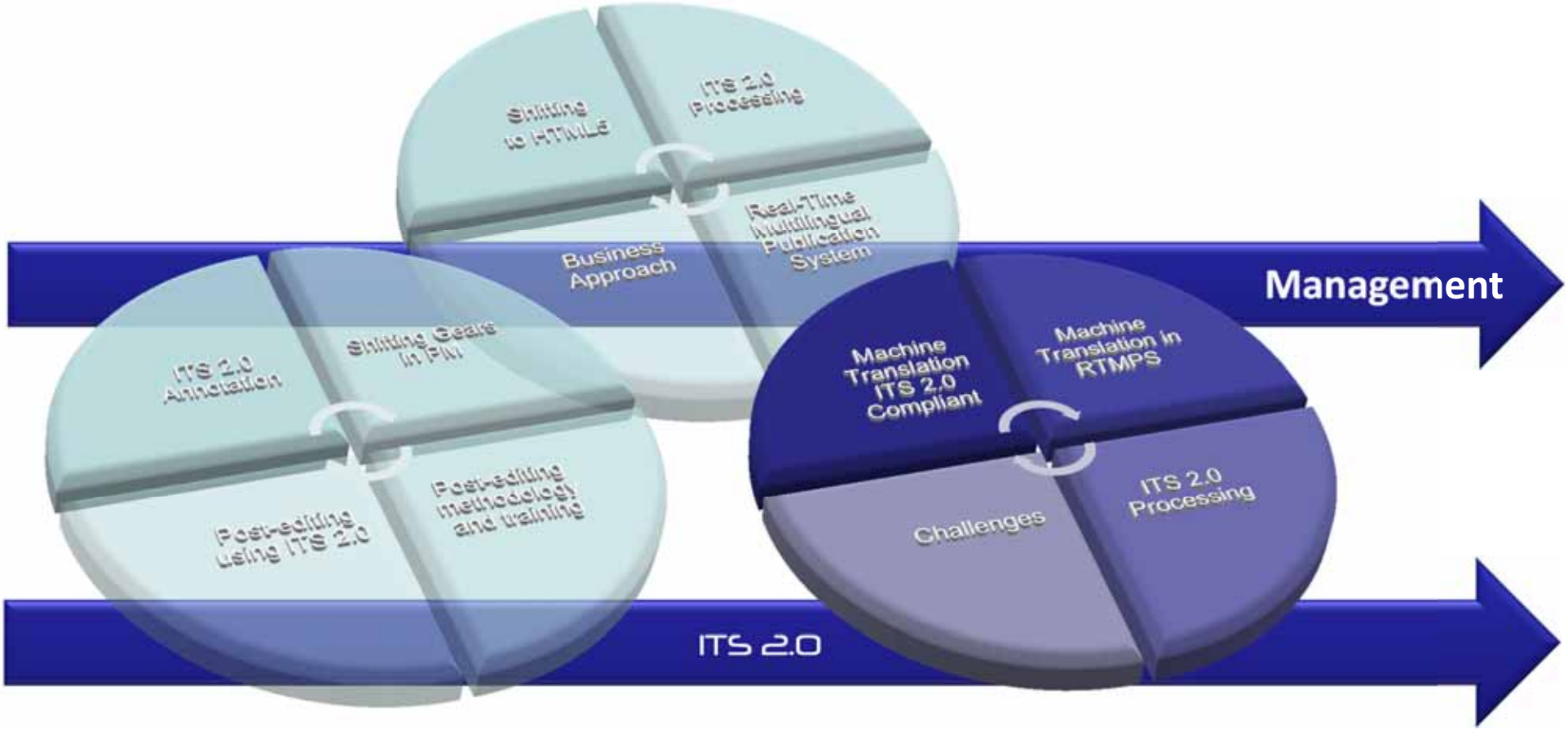
Machine Side: Sample Applications

Daniel Grasmick (Lucy Software)

Outline

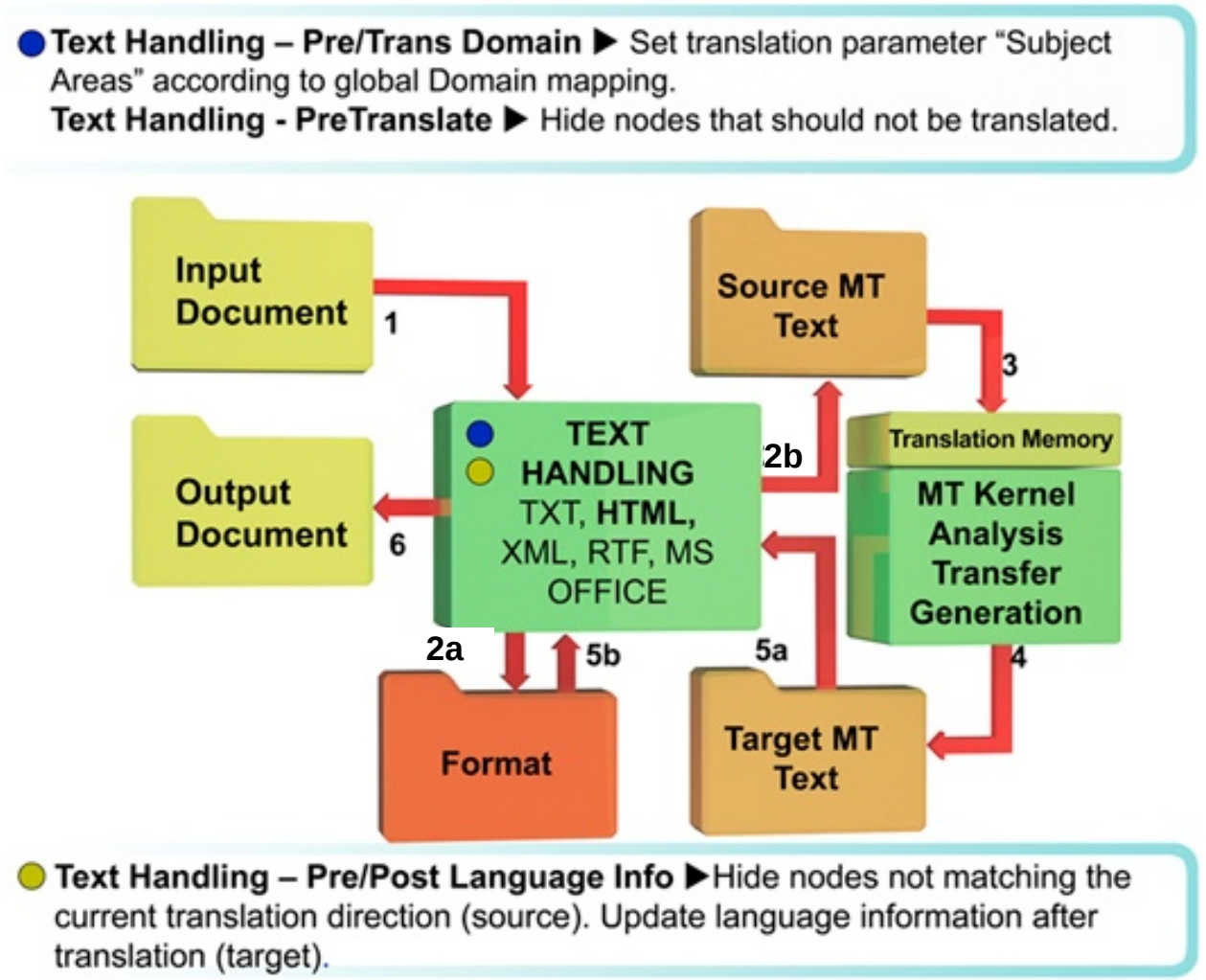
- This part intends to show the integration and application of Machine Translation.
- We will also demonstrate how data categories are used in MT processing as well as the impact and workflow of data categories interaction with other modules.

Real-Time Web I18N System use case in MLW-LT

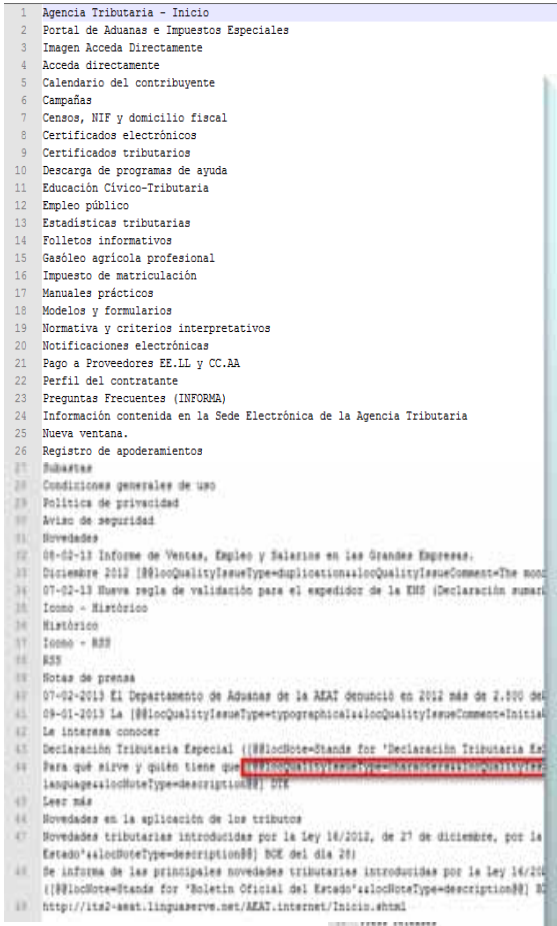


(1) Lucy MT System ITS 2.0 compliant

- Lucy MT on its own can process ITS 2.0 data categories.
- Integrated in the RTMPS, Linguaserve's ATLAS takes care of the ITS 2.0 data category.

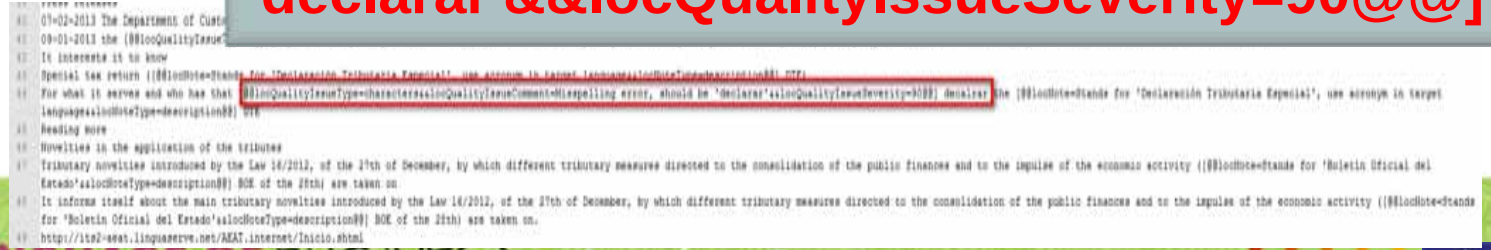


(1) Example Input files I – MT System Files

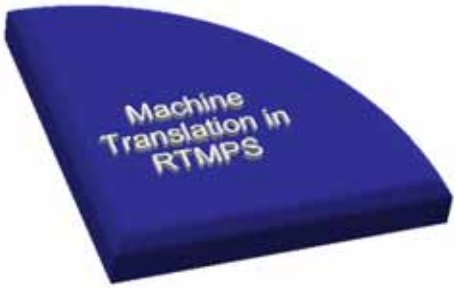


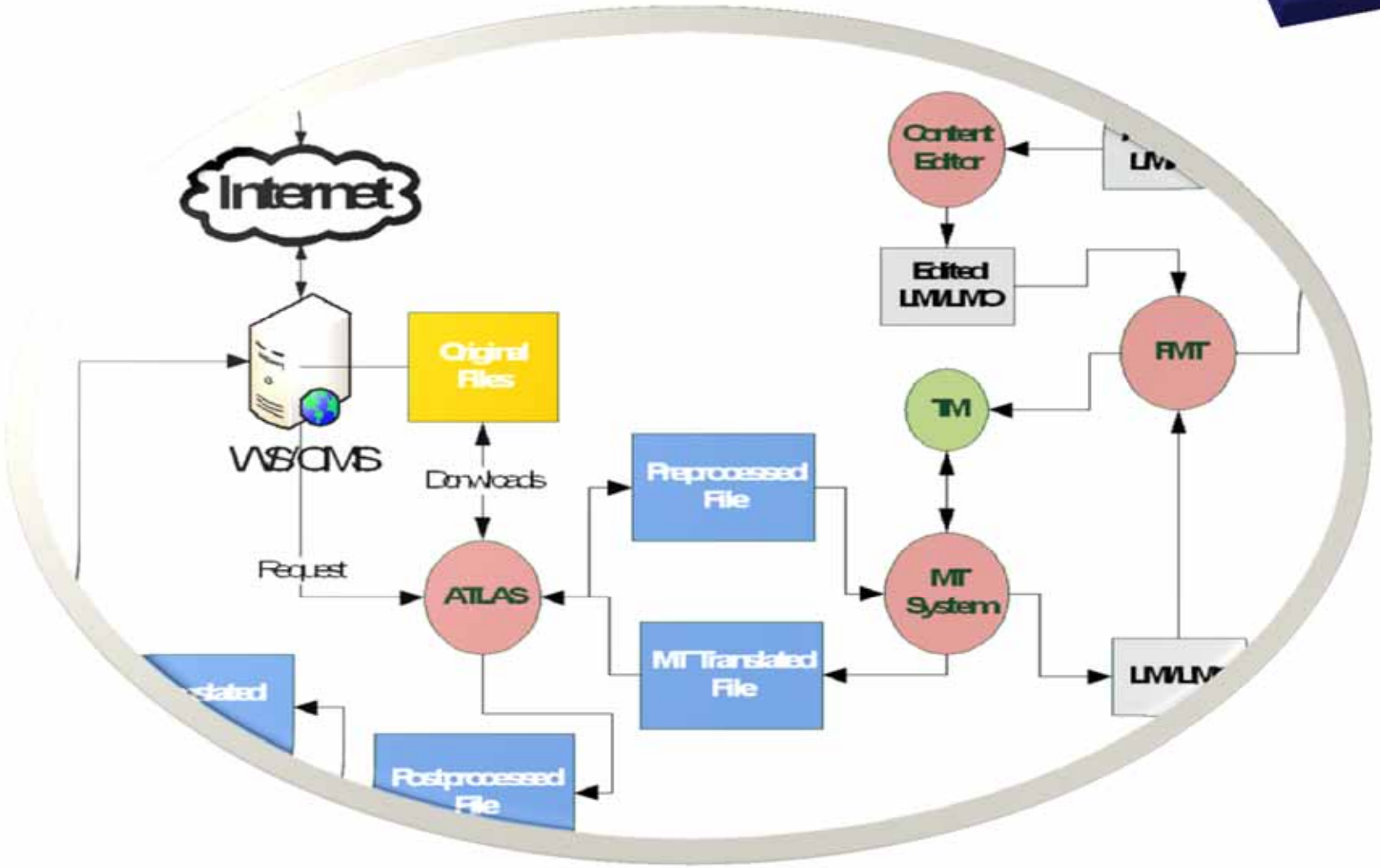
- **Data categories are converted by the RTMPS into a suitable processing format for the MT engine.**
- **Some data categories are used by the MT itself (Translate, Domain).**
- **Some other are passed through by the MT system for subsequent processes:**

[@@locQualityIssueType=characters&&locQualityIssueComment=Misspelling error, should be 'declarar'&&locQualityIssueSeverity=90@@]

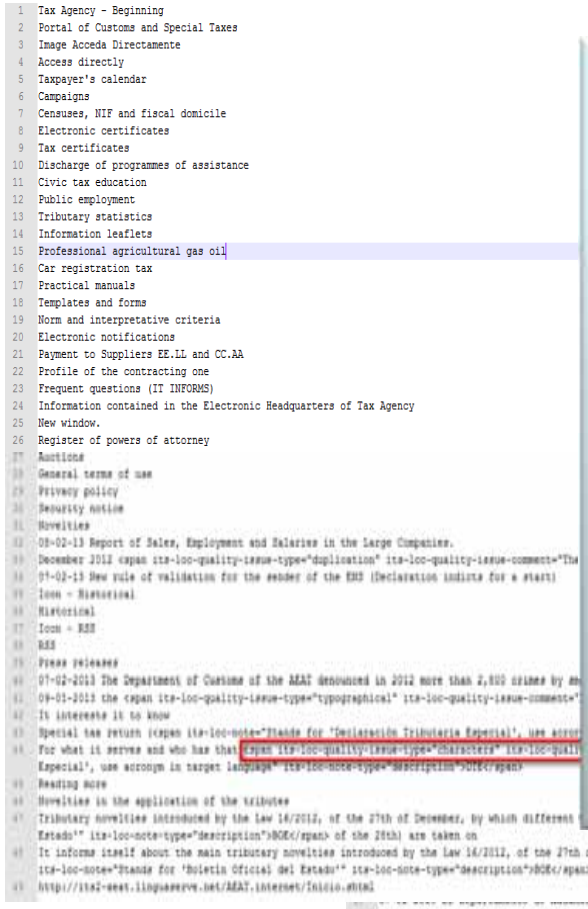


(2) MT system in the RTMPS





(2) Example Input files II – FMT Files



The RTMPS converts the data categories back from the MT to the HTML5 compatible format:

decalrar

GALA 2013 MIAMI
THE LANGUAGE OF BUSINESS
THE BUSINESS OF LANGUAGE
Miami Beach
17-20 March 2013
ITS LT
MultilingualWeb-LT
W3C

(3) ITS 2.0 used in the MT processing in HMTL5 (I)

- Automatic processing of Translate
  - *Default:*
    All elements are translated, attributes not.
  - *Global rules and local "translate" attribute may override the default.*
  - *No-translate nodes are masked as data and will not be processed by MT Kernel.*

(3) ITS 2.0 used in the MT processing using HMTL5 (II)

- Automatic processing of Domain
  - Global rules define a domain property
  - Domain property can be mapped to tool-specific namespaces
  - Lucy LT engine accepts one domain per document ("Subject Area")
  - Subject Area used for word sense disambiguation

(3) Demonstration

- MaTrEx prototype
- **ITS 2.0 LucySoftware prototype**
- ATLAS Real-Time: ITS 2.0 prototype
- ATLAS Real-Time: ITS 2.0 Testing
- Spanish Tax Agency Showcase

Access to demos:

http://www.w3.org/International/multilingualweb/lt/wiki/Use_cases_-_high_level_summaryITS 2.0

(4) Challenges for using ITS 2.0

- Terminology and Disambiguation
  - PE tools and MT systems should be ready to handle the marked terms and their associated information.
- MT Confidence
  - Need reliable MT metrics (e.g. unknown terms) and quality measuring systems
- External Resource
  - Need CAT and PE tools to handle references to external resources to be translated.
  - E.g.: audio & video files would need specific pre-and post-processing

GALA 2013 MIAMI

THE LANGUAGE OF BUSINESS
THE BUSINESS OF LANGUAGE

Miami Beach
17-20 March 2013

Human Side: tasks and processes

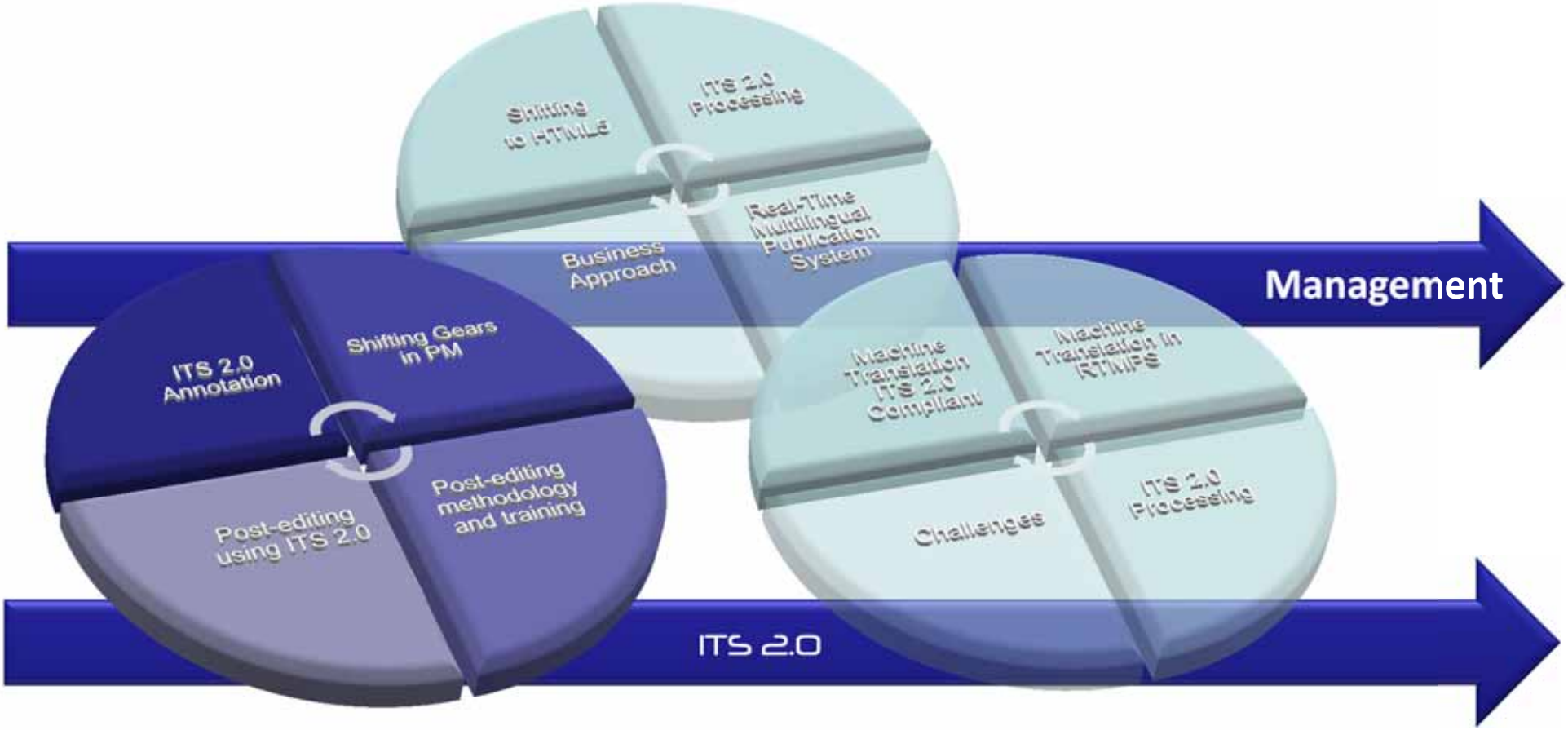
Pablo Badía Mas (Linguaserve)

@gala_global
#galaconf

Outline

- This part describes how a project of this nature changes, both in its management and in the tasks to be carried out, providing examples of post-editing using ITS 2.0 data categories, and showing new possibilities and perspectives.

Real-Time Web I18N System use case in MLW-LT



(1) ITS 2.0 annotation experience

- Strategy adopted in order to annotate the content with ITS 2.0 in an efficient and pragmatic way, considering the pressure and requirements of a real environment.

ITS 2.0

Automatic custom tags conversion

Automatic annotation

Manual annotation

(1) Automatic ITS 2.0 re-use of custom tags

- Custom “no translate” tag already exists in the content and is automatically annotated as ITS 2.0 *Translate* data category:

*
<!--ATLASP1NOTRAD--><a target="_blank" href="http://www.boe.es/diario_boe/txt.php?id=BOE-A-2011-20472">Orden EHA/3552/2011, de 19 de diciembre […] **<!--/ATLASP1NOTRAD-->**
*

*
<a **translate=”no”** target="_blank" href="http://www.boe.es/diario_boe/txt.php?id=BOE-A-2011-20472">Orden EHA/3552/2011, de 19 de diciembre […] *

ITS 2.0

* Also, it is needed the addition of ITS global default rules for known translatable attributes:
  - <its:translateRule selector="//h:*/@title" translate="yes"/>
  - <its:translateRule selector="//h:*/@alt" translate="yes"/>

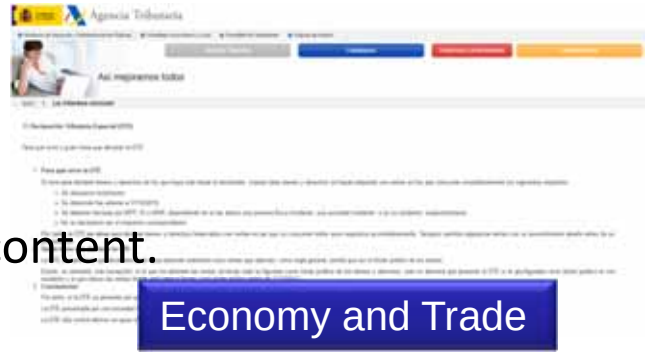
(1) Automatic ITS 2.0 annotation: Domain

1.Extracting relevant domains based on the content.

2. Alignment of the domains with each web page.

3. Use of scripts and regular expressions to annotate the content.

4. Document processing:



i. The selector points to the html root element, indicating that the domain applies to the whole **HTML** document (inheritance).

ii. The **domainPointer** attribute indicates where the domain that applies to the selected content is ("**Economy and Trade**").

iii. The **domainMapping** maps the domain "Economy and Trade" to "**ECON**", which will be sent as an understandable parameter to the MT System.

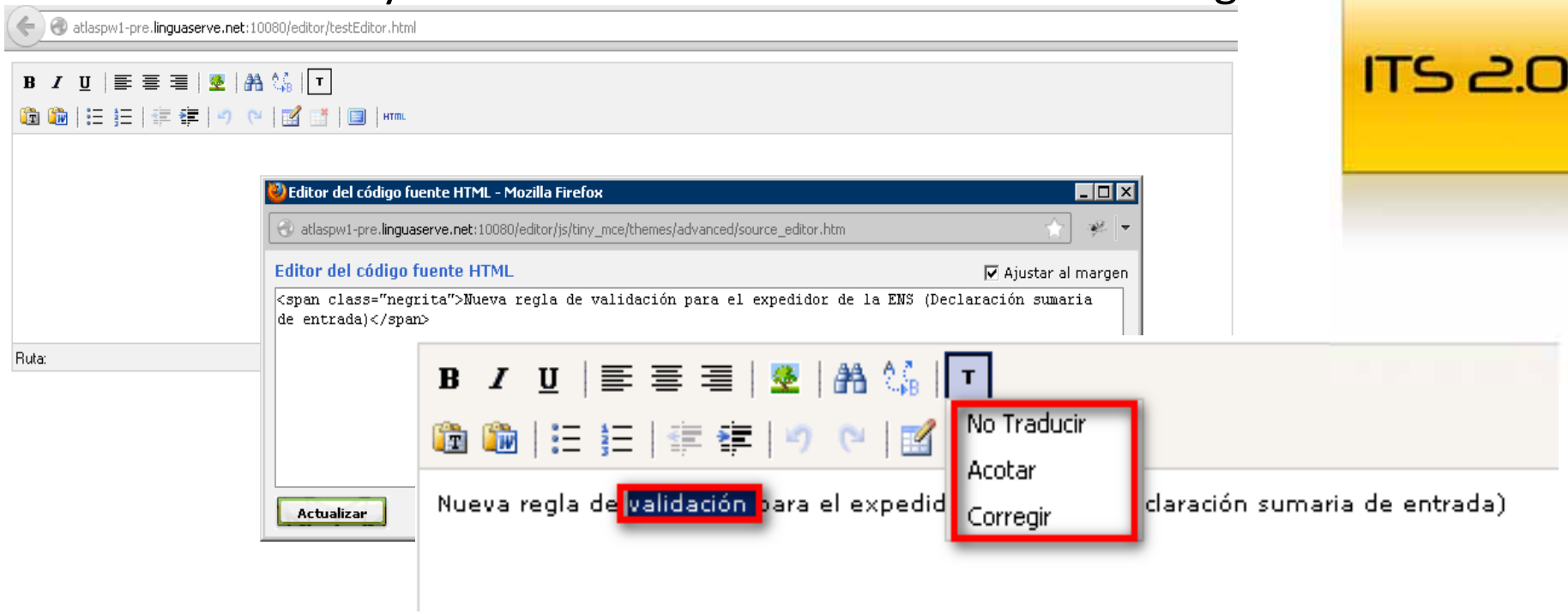
```
<!DOCTYPE html>
<html lang="es">
<head>
<meta charset="utf-8">
<meta name="keywords" content="Economy and Trade"/>
[DOMAIN RULES]
</head>
<body>
[…]
</body>
</html>
```

```
<its:rules xmlns:its="http://www.w3.org/2005/11/its"
xmlns:h="http://www.w3.org/1999/xhtml"
version="2.0">
<its:domainRule
selector="//h:html"
domainPointer="/html/head/meta[@name='keywords']/@content“
domainMapping="'Economy and Trade' ECON,
'Law and Legal Science' LAW, ‘General Vocabulary' GV"/>
</its:rules>
```

MT System

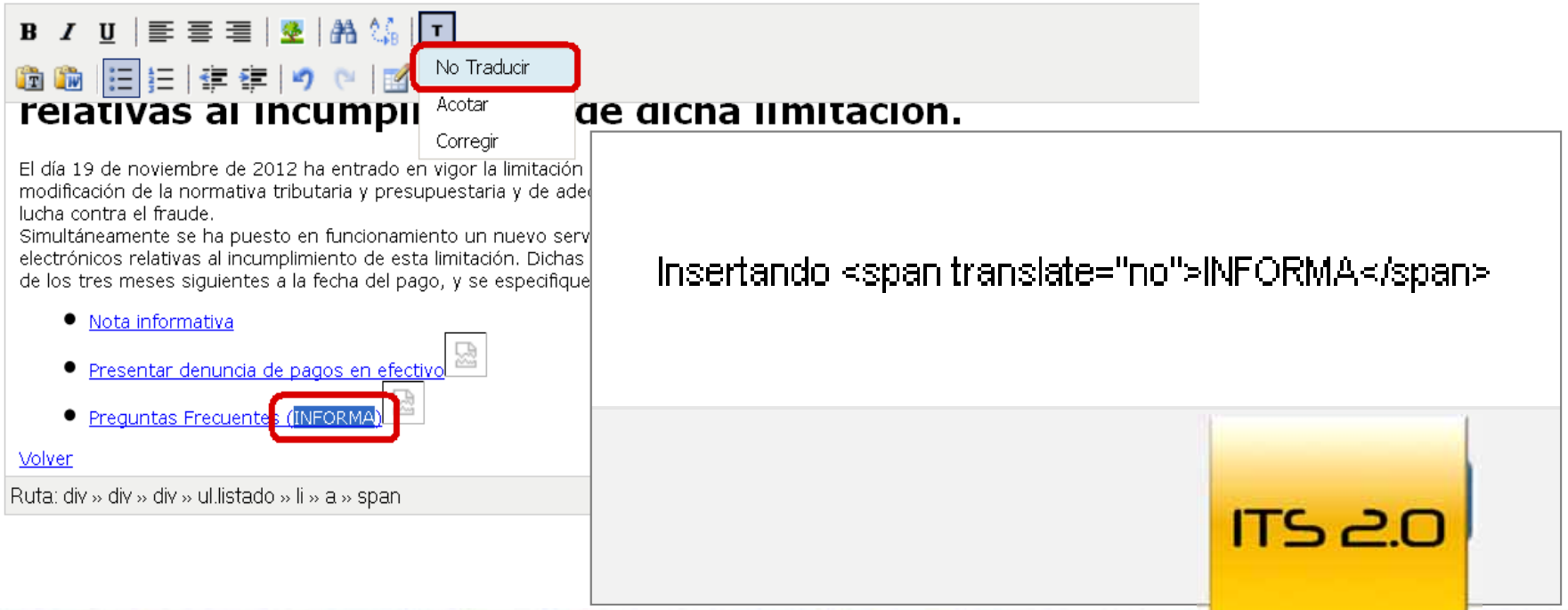
(1) Manual ITS 2.0 annotation

- Quick and pragmatic approach:
  - New HTML Editor plugin created for the ITS 2.0 manual annotation for open source HTML Editor
  - User-friendly interface for the manual insertion of tags



(1) ITS 2.0 Manual annotation: Translate

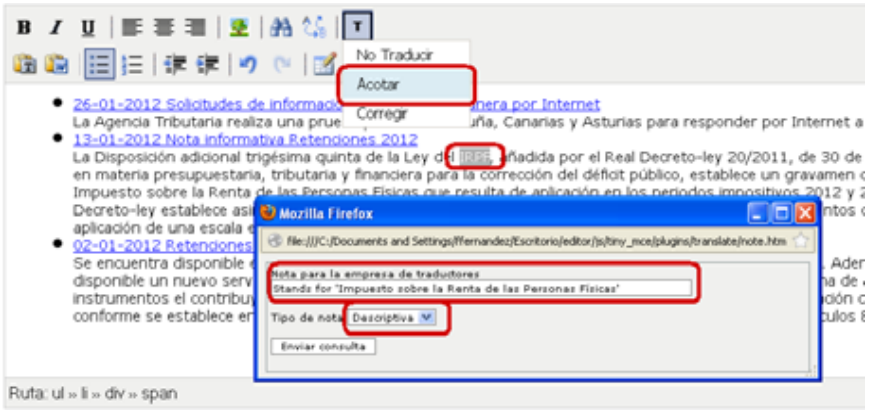
- The author must only select the non-translatable element, click on the insertion icon (T) and click on the annotation type: ***No Traducir***.



GALA 2013 MIAMI
THE LANGUAGE OF BUSINESS
THE BUSINESS OF LANGUAGE
Miami Beach
17-20 March 2013
MultilingualWeb-LT
W3C

(1) ITS 2.0 Manual annotation: Localization Notes

- Use of the annotation type ***Acotar***: The author inserts the annotation text into the box and the software will automatically create the tag.
- The pull-down menu is used to choose the type of localization note. It can either be description (*descriptiva*) or alert (*alerta*).




```
<p>La disposición trigésima quinta de la Ley del <SPAN ITS-LOC-NOTE="STANDS FOR 'IMPUESTO SOBRE LA RENTA DE LAS PERSONAS FÍSICAS ', USE ACRONYM IN TARGET LANGUAGE" ITS-LOC-NOTE-TYPE="DESCRIPTION">IRPF</SPAN>
</p>
```

(1) ITS 2.0 Manual annotation: Localization Quality Issue

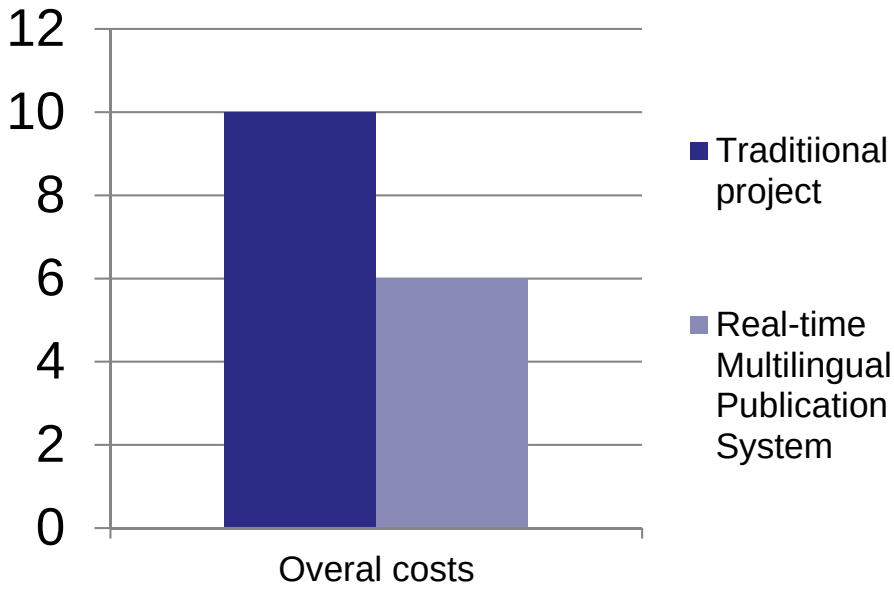
- Use of the annotation type ***Corregir***: The author chooses a type of issue from a pull-down menu, inserts a comment into the box (*Comentario*), chooses a severity level between 0 and 100 (*Severidad*) and an optional link to a reference document (*documento de referencia*), and the software will automatically create the tag.

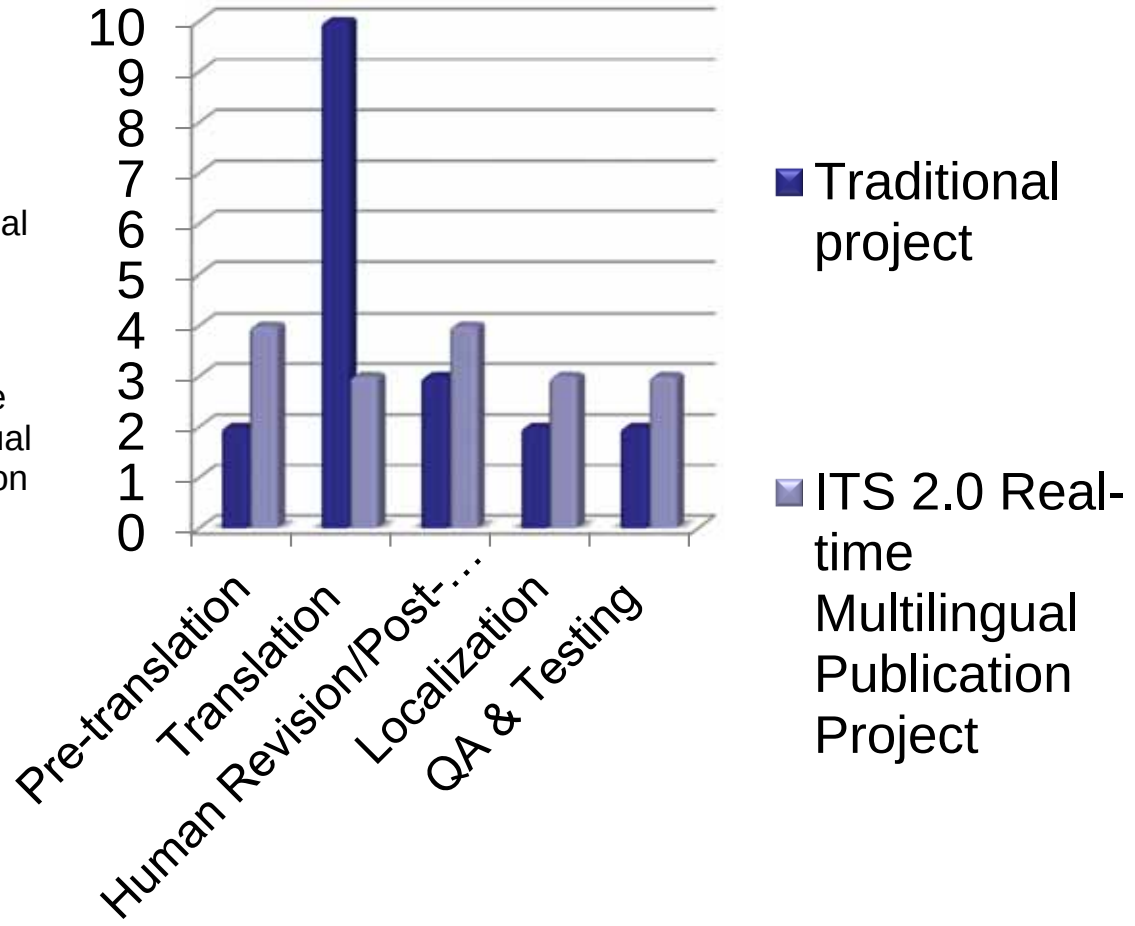
Online filing can be done by the interested party or by someone representing them. In both cases, an electronic certificate X.509.V3 issued by the **<SPAN ITS-LOC-QUALITY-ISSUE-COMMENT="HAS PREVIOUSLY BEEN TRANSLATED AS 'ROYAL MINT'. PLEASE BE CONSISTENT." ITS-LOC-QUALITY-ISSUE-TYPE="INCONSISTENCY" ITS-LOC-QUALITY-ISSUE-SEVERITY="70">NATIONAL COIN AND STAMP FACTORY</SPAN>**.

(2) Shifting Gears: Impact in project management

- Basic differences in project management
  - Offline Web Translation System and Translation services
    - Proactive planning and reactive execution
    - Terminology for the specific domain
  - Online Web Translation System
    - Proactive planning and execution
    - Application of the concept “Textual Universe”
    - Terminology from text and MT system

(2) Shifting gears: New cost structure

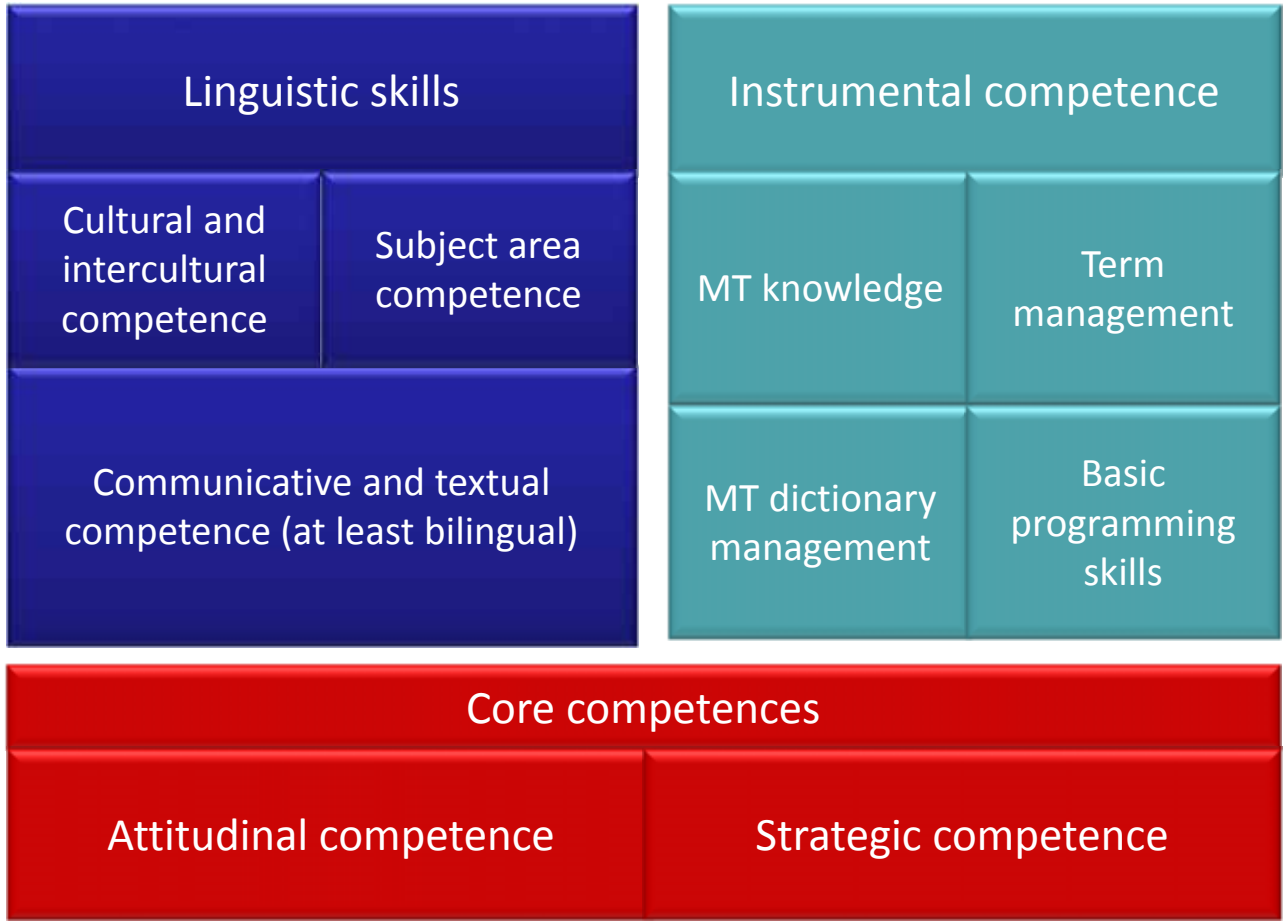




(2) Shifting Gears: new tasks

- New translation tasks
  - Textual analysis and definition of post-editing rules
  - Terminological analysis, management and codification
  - Post-editing of Machine Translation
  - MT error reporting
- New localization tasks
  - Prior programming HTML localization
  - Management of binary files for localization
  - Proactive recurrent testing and fixing

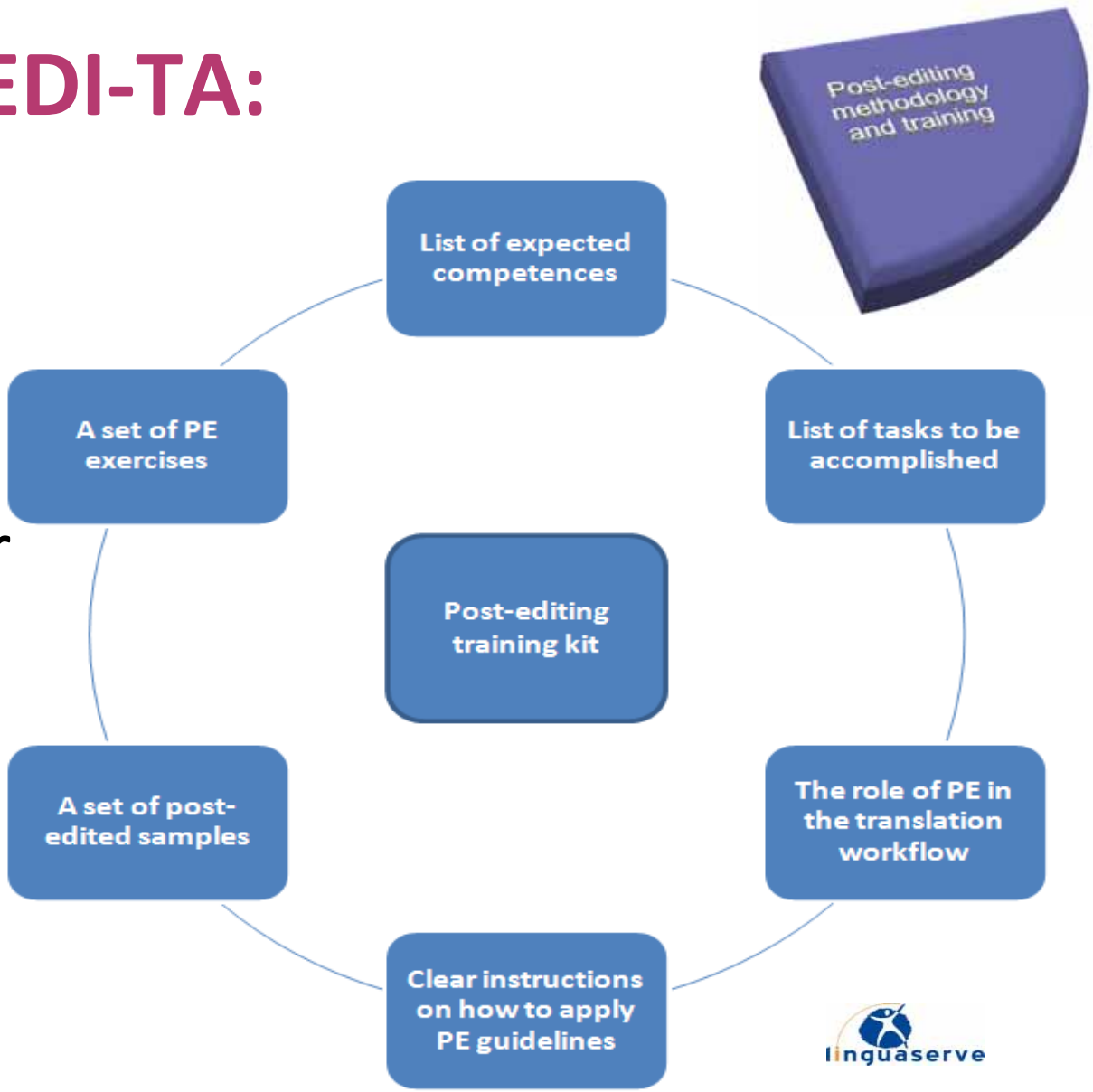
(2) Team capabilities and profiles: post-editor



PE skills and competences
(Source: Rico and Torrejón, forthcoming)

(3) Post-editing EDI-TA: Training

- The post-editing training kit includes a number of skills that extend beyond of those of a usual translator or reviewer



(3) Post-editing EDI-TA Methodology: Step 1: Preliminary analysis

- **Objectives**
  - Establishing PE patterns for each language combination
  - Report on recurrent MT errors that can be fixed prior to starting the PE process.
  - Report new terms to be included in project glossaries.
- **Outcome:**
  - **Data analysis for post-editing guidelines specification.**
    - Language independent PE rules
    - language dependent PE rules,
    - system specific PE rules
    - client specific rules.
  - **PE guide.**

(3) Post-editing EDI-TA Methodology: Step 2: post-editing MT output

- **Objectives**
  - Post-editing MT output according to the specifications established in phase 1, and in the PE guide.
  - MT errors to report later in the process.
  - Responding to output quality as required by the client.
- **Outcome**
  - Post-edited output.
  - List of terms to be updated in project's glossaries.
  - List of observations toward the PE guidelines, should any of these be updated.

(3) Post-editing EDI-TA Methodology: Step 3: error reporting & quality control

- **Objectives**
  - Reporting feedback to allow for MT improvement and/or source content optimization
  - Conducting a quality control following *on demand* client expectations.
  - Collecting samples of post-editing issues in order to facilitate training of other fellow post-editors in the team, and keeping up-to-date the reference material.
- **Tools**
  - MT error reporting
  - Quality control template
  - PE guidelines comments
  - Term registration

(3) Post-editing EDI-TA Methodology: Types of rules

- Dependent /Independent-Language rules
- Dependent /Independent-System rules
- Dependent /Independent-Client rules
- Other ?

Basic Language independent rules

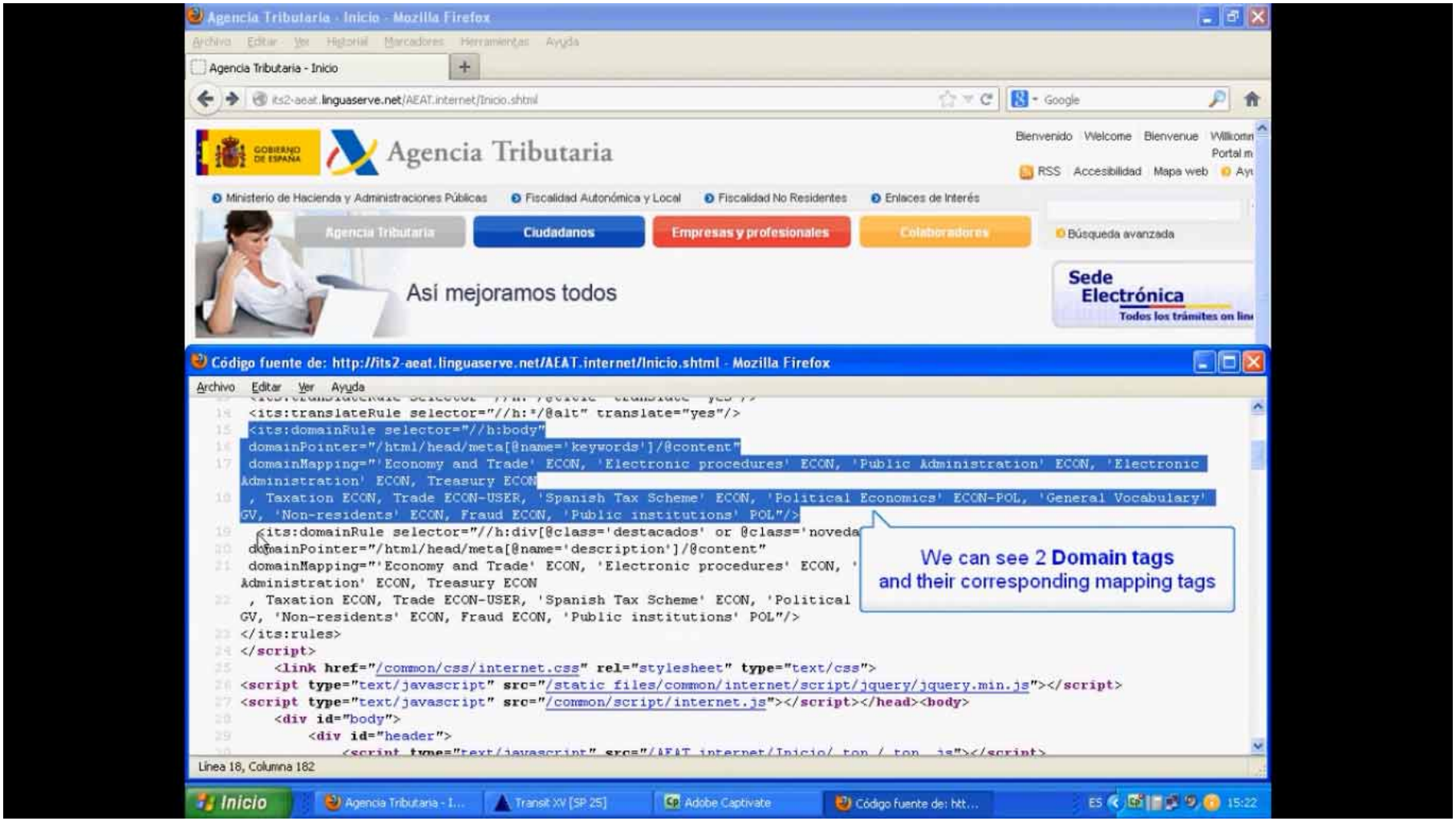
Category	Rule	Description
Addition	Omission	Fix any omissions as long as they interfere with the message.
Deletion	Syntax	Delete any element which interferes with the message.
Movement	Syntax	Fix any linear order, only if it interferes with the message.
Movement	Syntax	Fix any incorrect linear order.
Replacement	Syntax	Fix any incorrect phrase structure.
Replacement	Morphology	Fix any morphological mistakes.
Replacement	Lexicon	Fix wrong lexicon.
Replacement	Terminology	Fix wrong terminology.

(3) Post-editing EDI-TA Methodology: ITS 2.0 metadata that helps people

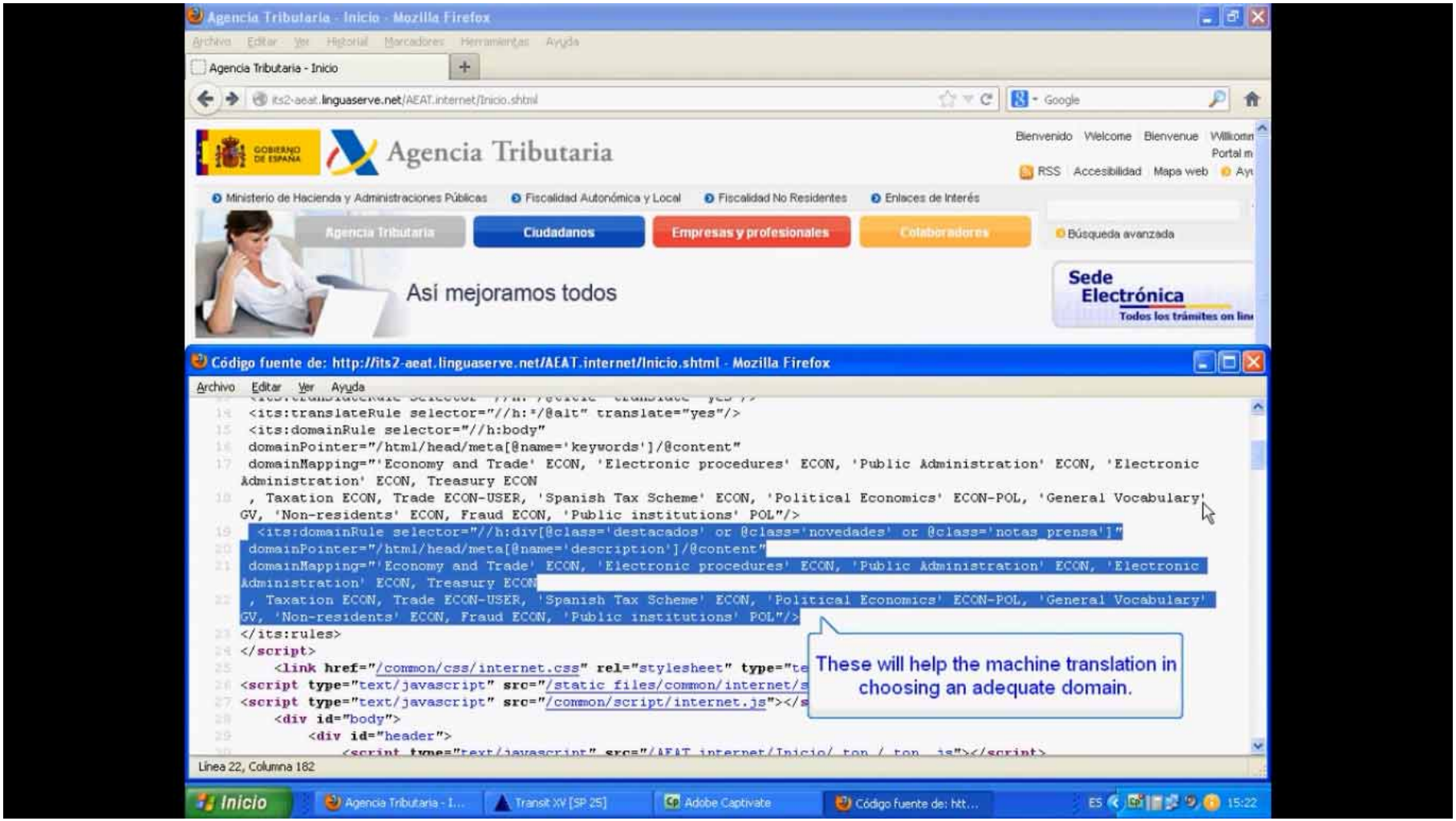
- Providing contextual information
  - Language information
  - Domain
  - Provenance
- Containing activation PE rules
  - Translate
  - Localization note
  - LocalizationQualityIssue*

(4) Demonstration: PE using ITS 2.0

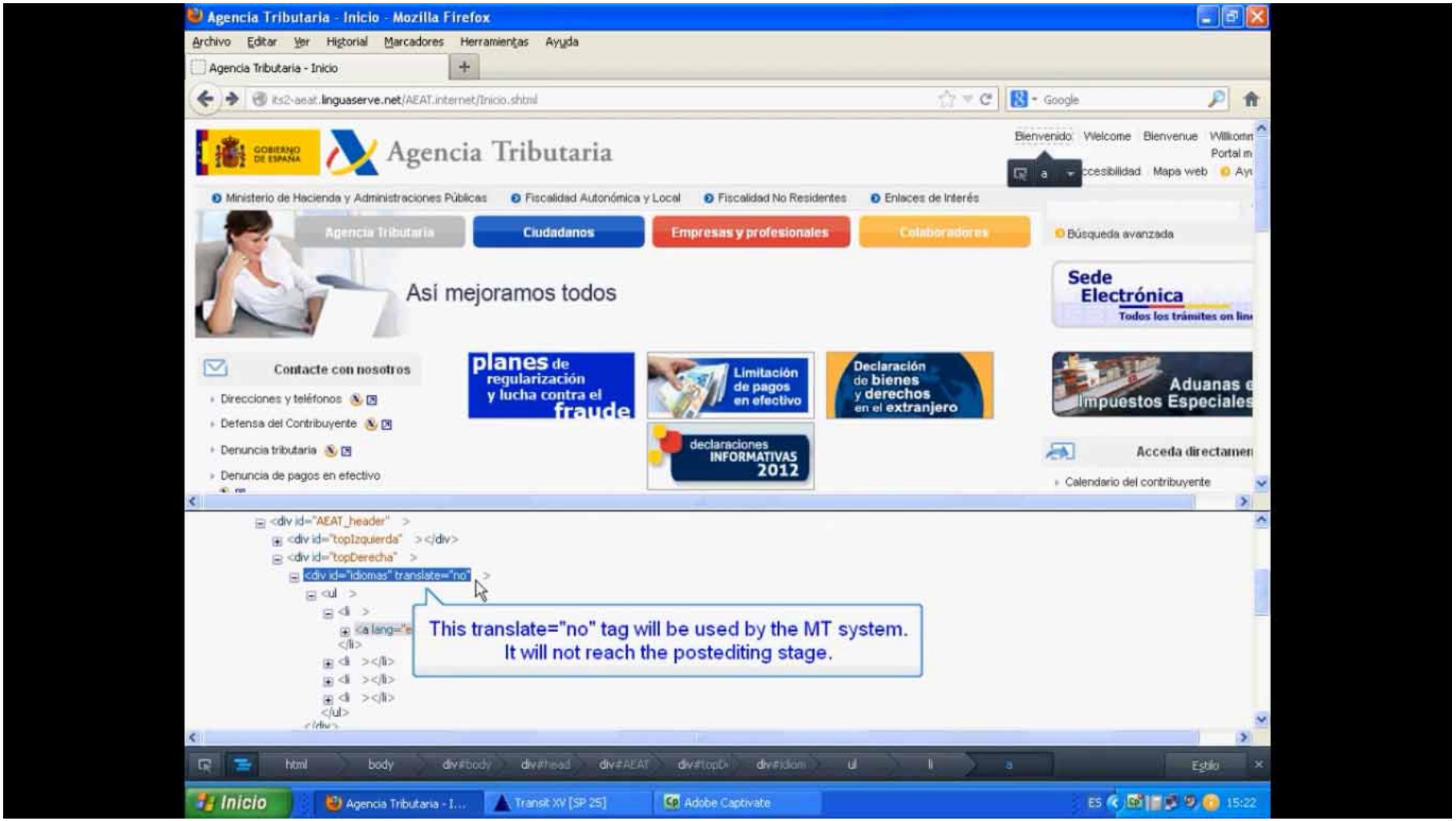
Demo 1/9


Agencia Tributaria - Inicio - Mozilla Firefox
Agencia Tributaria
Así mejoramos todos
Código fuente de: http://its2-aeat.linguaserve.net/AEAT.internet/Inicio.shtml - Mozilla Firefox
<its:translateRule selector="//h:*/@alt" translate="yes"/>
<its:domainRule selector="//h:body"
domainPointer="/html/head/meta[@name='keywords']/@content"
domainMapping="'Economy and Trade' ECON, 'Electronic procedures' ECON, 'Public Administration' ECON, 'Electronic Administration' ECON, Treasury ECON
, Taxation ECON, Trade ECON-USER, 'Spanish Tax Scheme' ECON, 'Political Economics' ECON-POL, 'General Vocabulary' GV, 'Non-residents' ECON, Fraud ECON, 'Public institutions' POL"/>
We can see 2 Domain tags
and their corresponding mapping tags

Demo 2/9

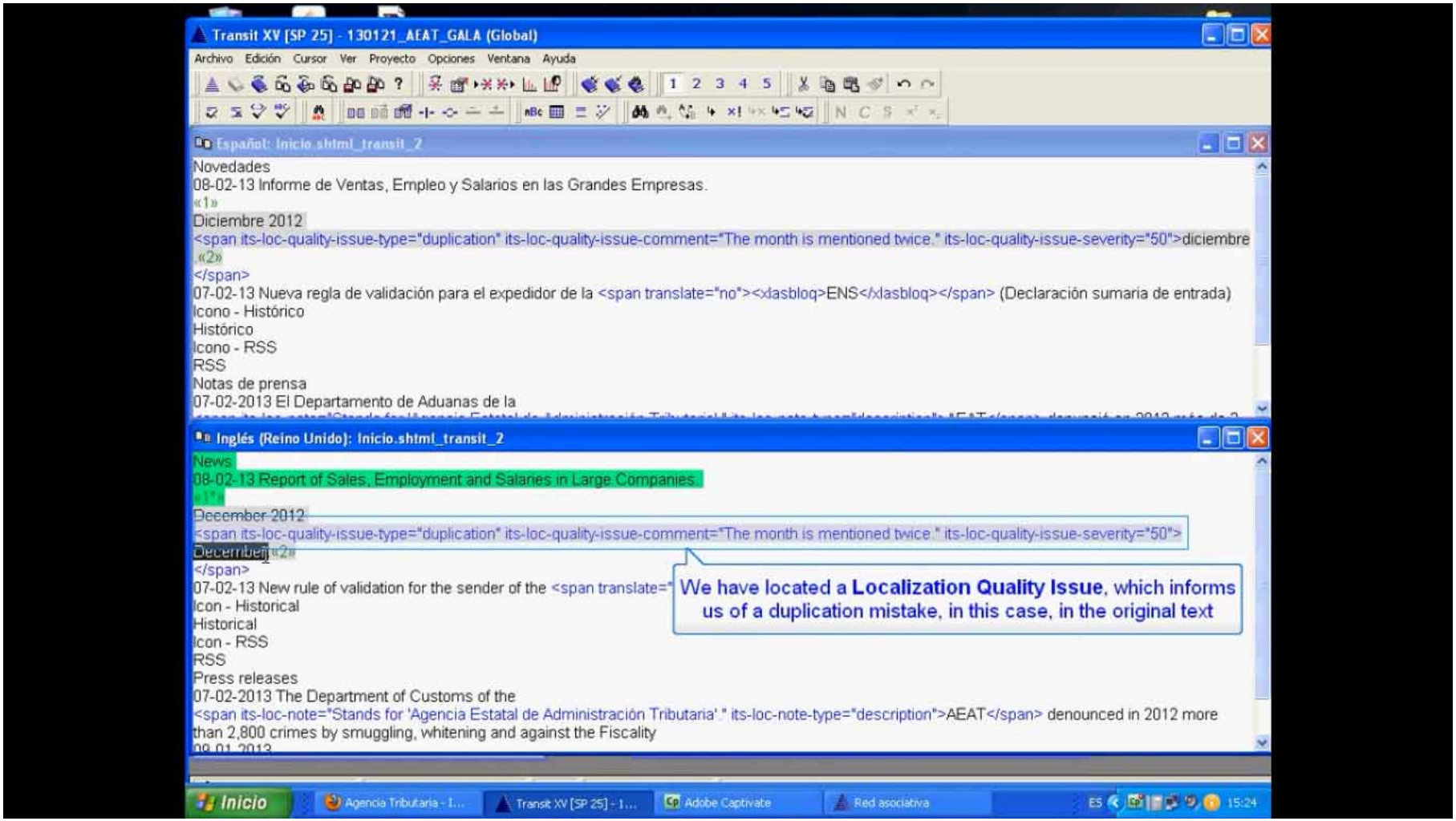

Agencia Tributaria - Inicio - Mozilla Firefox
Agencia Tributaria
Así mejoramos todos
Código fuente de: http://its2-aeat.linguaserve.net/AEAT.internet/Inicio.shtml - Mozilla Firefox
These will help the machine translation in choosing an adequate domain.

Demo 3/9


Agencia Tributaria - Inicio - Mozilla Firefox
Archivo Editar Ver Historial Marcadores Herramientas Ayuda
Agencia Tributaria - Inicio
Agencia Tributaria
Ministerio de Hacienda y Administraciones Públicas
Fiscalidad Autonómica y Local
Fiscalidad No Residentes
Enlaces de Interés
Ciudadanos
Empresas y profesionales
Colaboradores
Así mejoramos todos
Sede Electrónica
Contacte con nosotros
Direcciones y teléfonos
Defensa del Contribuyente
Denuncia tributaria
Denuncia de pagos en efectivo
planes de regularización y lucha contra el fraude
Limitación de pagos en efectivo
Declaración de bienes y derechos en el extranjero
declaraciones INFORMATIVAS 2012
Calendario del contribuyente

This translate="no" tag will be used by the MT system.
It will not reach the postediting stage.
Inicio
Agencia Tributaria - I...
Transit XV [SP 25]
Adobe Captivate

Demo 4/9

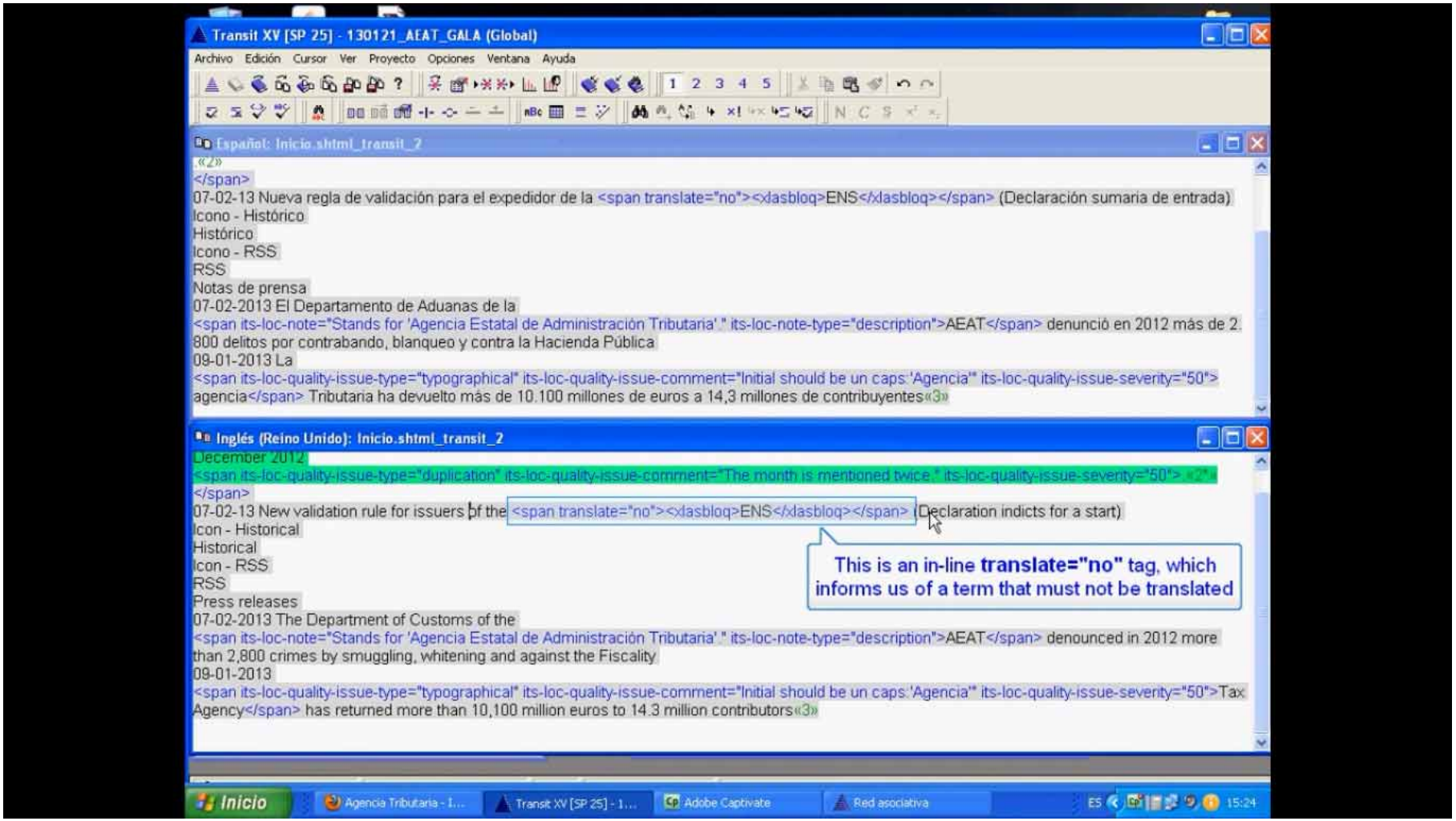

Transit XV [SP 25] - 130121_AEAT_GALA (Global)
Archivo Edición Cursor Ver Proyecto Opciones Ventana Ayuda
Español: Inicio.shtml_transit_2
Novedades
08-02-13 Informe de Ventas, Empleo y Salarios en las Grandes Empresas.
«1»
Diciembre 2012
diciembre
«2»

07-02-13 Nueva regla de validación para el expedidor de la <xlasbloq>ENS</xlasbloq> (Declaración sumaria de entrada)
Icono - Histórico
Histórico
Icono - RSS
RSS
Notas de prensa
07-02-2013 El Departamento de Aduanas de la
Inglés (Reino Unido): Inicio.shtml_transit_2
News
08-02-13 Report of Sales, Employment and Salaries in Large Companies.
«1»
December 2012

December
«2»

07-02-13 New rule of validation for the sender of the <span translate="
Icon - Historical
Historical
Icon - RSS
RSS
Press releases
07-02-2013 The Department of Customs of the
AEAT denounced in 2012 more than 2,800 crimes by smuggling, whitening and against the Fiscality
09-01-2013
We have located a Localization Quality Issue, which informs us of a duplication mistake, in this case, in the original text
Inicio
Agencia Tributaria - I...
Transit XV [SP 25] - I...
Adobe Captivate
Red asociativa
ES
15:24

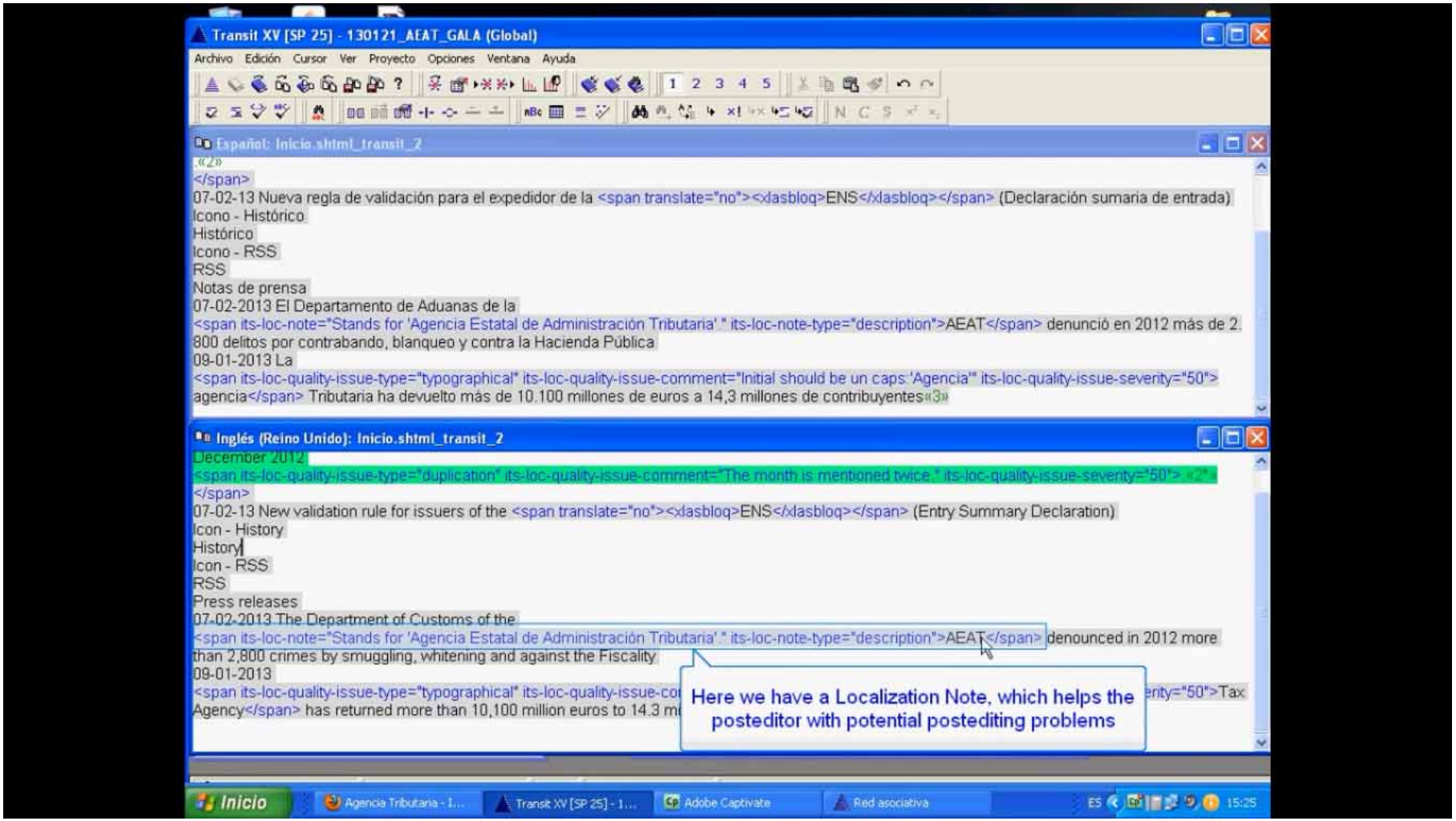
Demo 5/9


Transit XV [SP 25] - 130121_AEAT_GALA (Global)
Archivo Edición Cursor Ver Proyecto Opciones Ventana Ayuda
Español: Inicio.shtml_transit_2
«2»

07-02-13 Nueva regla de validación para el expedidor de la <xlasbloq>ENS</xlasbloq> (Declaración sumaria de entrada)
Icono - Histórico
Histórico
Icono - RSS
RSS
Notas de prensa
07-02-2013 El Departamento de Aduanas de la
AEAT denunció en 2012 más de 2.800 delitos por contrabando, blanqueo y contra la Hacienda Pública
09-01-2013 La
agencia Tributaria ha devuelto más de 10.100 millones de euros a 14,3 millones de contribuyentes«3»
Inglés (Reino Unido): Inicio.shtml_transit_2
December 2012
«2»

07-02-13 New validation rule for issuers of the <xlasbloq>ENS</xlasbloq> (Declaration indicts for a start)
Icon - Historical
Historical
Icon - RSS
RSS
Press releases
07-02-2013 The Department of Customs of the
AEAT denounced in 2012 more than 2,800 crimes by smuggling, whitening and against the Fiscality
09-01-2013
Tax Agency has returned more than 10,100 million euros to 14.3 million contributors«3»
This is an in-line translate="no" tag, which informs us of a term that must not be translated
Inicio
Agencia Tributaria - I...
Transit XV [SP 25] - I...
Adobe Captivate
Red asociativa
ES
15:24

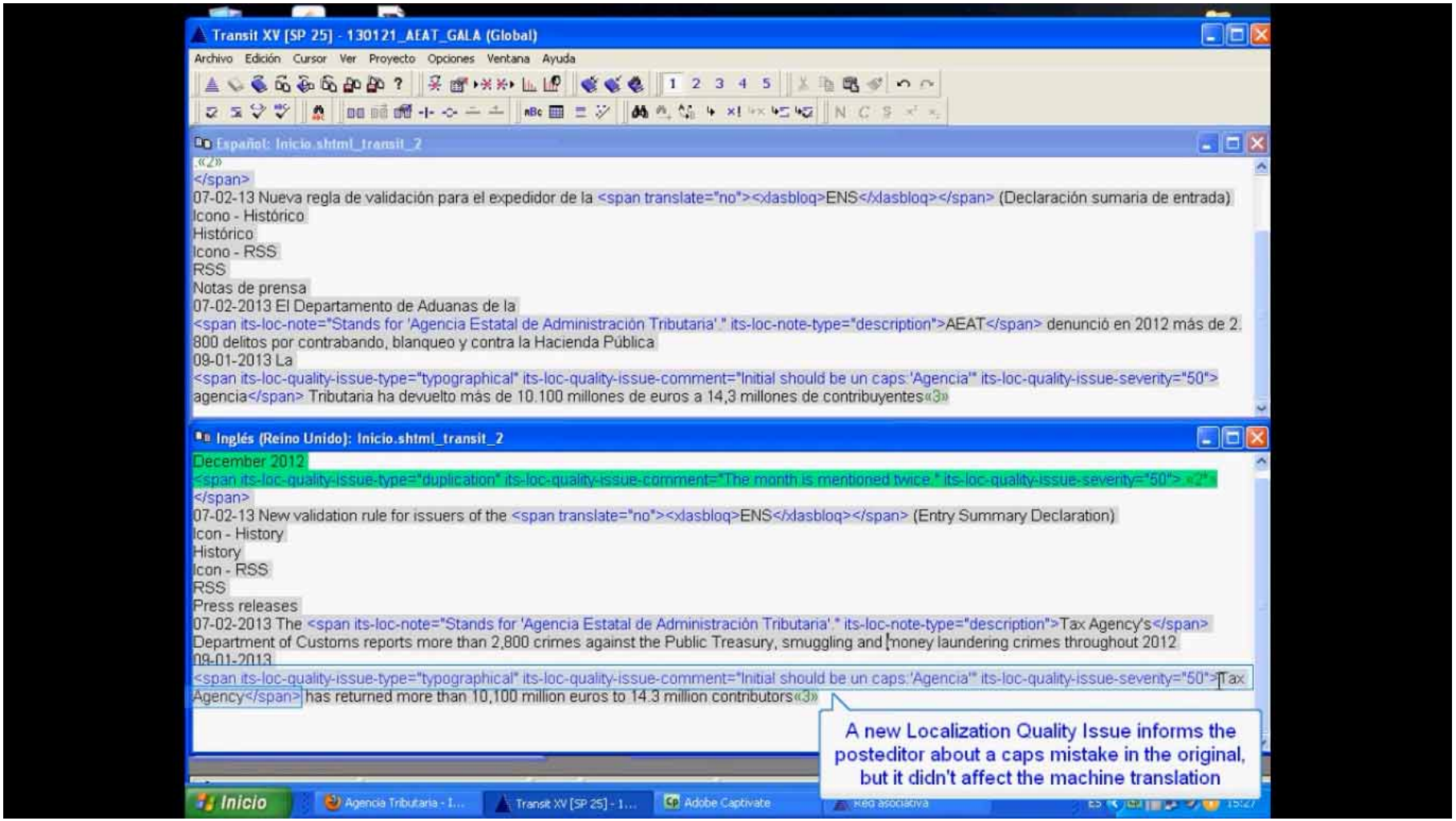
Demo 6/9


Transit XV [SP 25] - 130121_AEAT_GALA (Global)
Archivo Edición Cursor Ver Proyecto Opciones Ventana Ayuda
Español: Inicio.shtml_transit_2
«2»

07-02-13 Nueva regla de validación para el expedidor de la <xlasbloq>ENS</xlasbloq> (Declaración sumaria de entrada)
Icono - Histórico
Histórico
Icono - RSS
RSS
Notas de prensa
07-02-2013 El Departamento de Aduanas de la
AEAT denunció en 2012 más de 2.800 delitos por contrabando, blanqueo y contra la Hacienda Pública
09-01-2013 La
agencia Tributaria ha devuelto más de 10.100 millones de euros a 14,3 millones de contribuyentes«3»
Inglés (Reino Unido): Inicio.shtml_transit_2
December 2012
«2»

07-02-13 New validation rule for issuers of the <xlasbloq>ENS</xlasbloq> (Entry Summary Declaration)
Icon - History
History
Icon - RSS
RSS
Press releases
07-02-2013 The Department of Customs of the
AEAT denounced in 2012 more than 2,800 crimes by smuggling, whitening and against the Fiscality
09-01-2013
<span its-loc-quality-issue-type="typographical" its-loc-quality-issue-co
erity="50">Tax Agency has returned more than 10,100 million euros to 14.3 mi
Here we have a Localization Note, which helps the posteditor with potential postediting problems
Inicio
Agencia Tributaria - I...
Transit XV [SP 25] - I...
Adobe Captivate
Red asociativa
ES
15:25

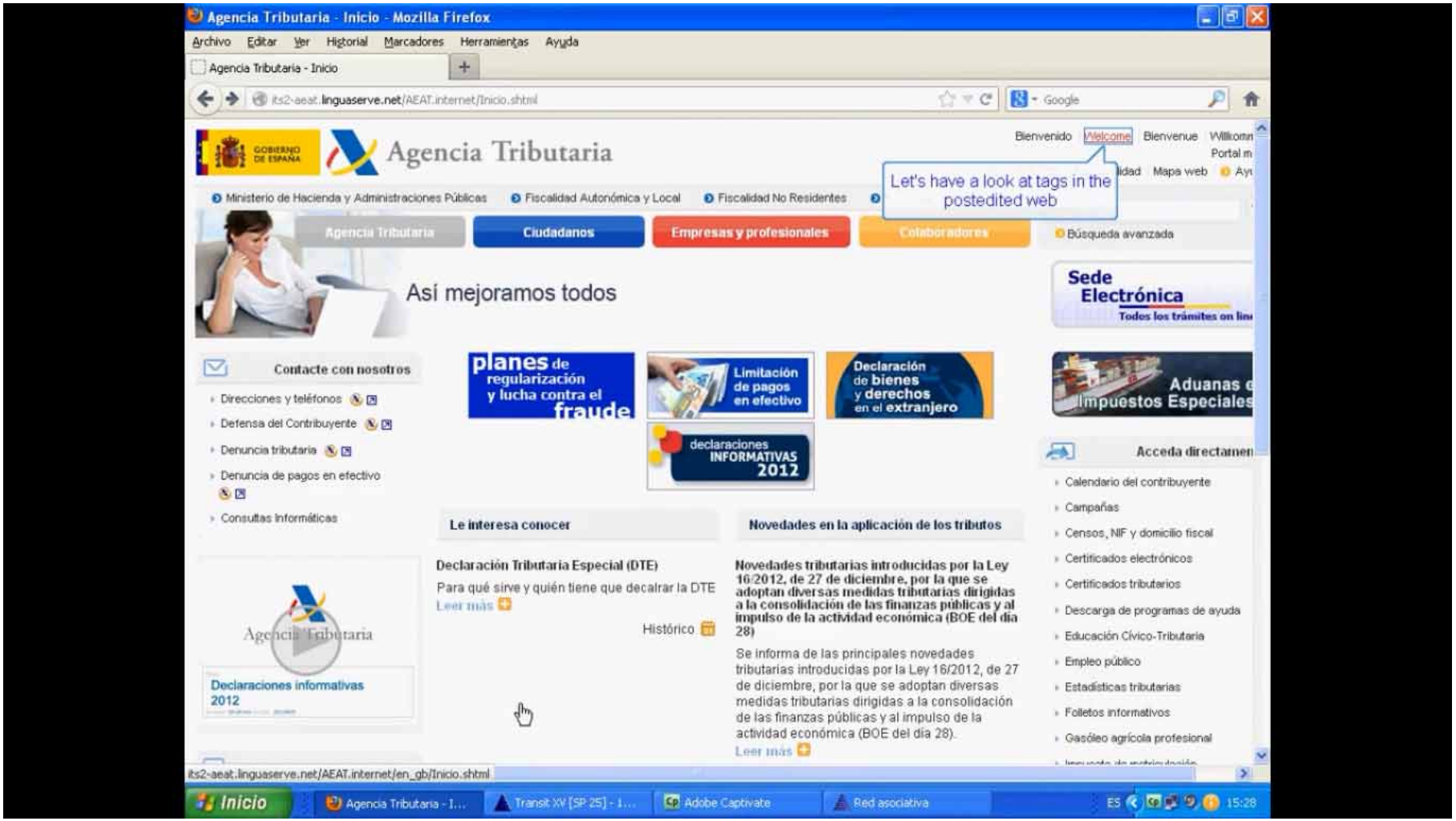
Demo 7/9


Transit XV [SP 25] - 130121_AEAT_GALA (Global)
Archivo Edición Cursor Ver Proyecto Opciones Ventana Ayuda
Español: Inicio.shtml_transit_2
«2»

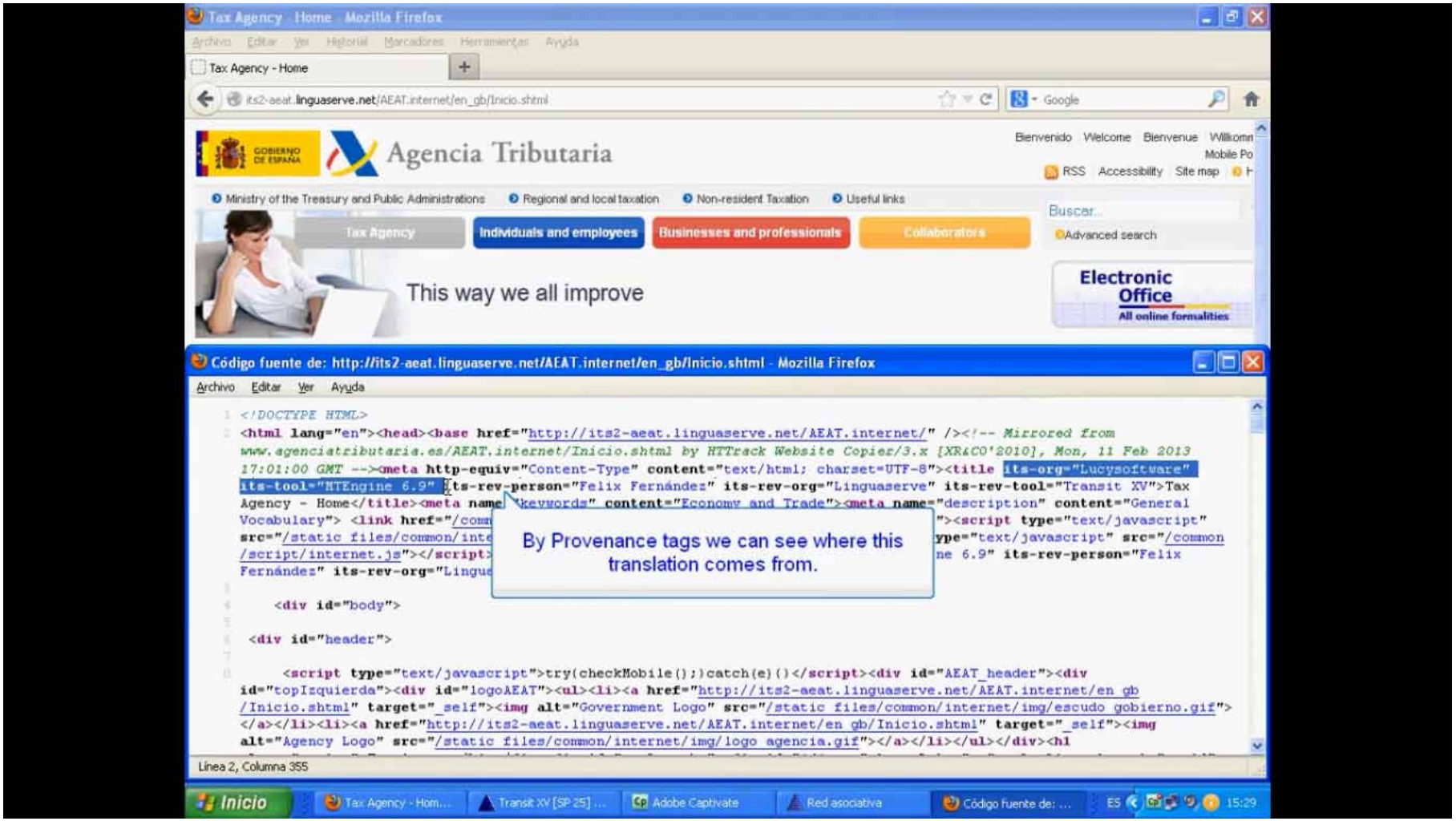
07-02-13 Nueva regla de validación para el expedidor de la <xlasbloq>ENS</xlasbloq> (Declaración sumaria de entrada)
Icono - Histórico
Histórico
Icono - RSS
RSS
Notas de prensa
07-02-2013 El Departamento de Aduanas de la
AEAT denunció en 2012 más de 2.800 delitos por contrabando, blanqueo y contra la Hacienda Pública
09-01-2013 La
agencia Tributaria ha devuelto más de 10.100 millones de euros a 14,3 millones de contribuyentes«3»
Inglés (Reino Unido): Inicio.shtml_transit_2
December 2012
«2»

07-02-13 New validation rule for issuers of the <xlasbloq>ENS</xlasbloq> (Entry Summary Declaration)
Icon - History
History
Icon - RSS
RSS
Press releases
07-02-2013 The Tax Agency's Department of Customs reports more than 2,800 crimes against the Public Treasury, smuggling and money laundering crimes throughout 2012
09-01-2013
Tax Agency has returned more than 10,100 million euros to 14.3 million contributors«3»
A new Localization Quality Issue informs the posteditor about a caps mistake in the original, but it didn't affect the machine translation
Inicio
Agencia Tributaria - I...
Transit XV [SP 25] - I...
Adobe Captivate

Demo 8/9


Agencia Tributaria - Inicio - Mozilla Firefox
Archivo Editar Ver Historial Marcadores Herramientas Ayuda
Agencia Tributaria - Inicio
its2-aeat.linguaserve.net/AEAT.internet/Inicio.shtml
Agencia Tributaria
Let's have a look at tags in the postedited web
Ministerio de Hacienda y Administraciones Públicas
Fiscalidad Autonómica y Local
Fiscalidad No Residentes
Ciudadanos
Empresas y profesionales
Colaboradores
Búsqueda avanzada
Así mejoramos todos
Sede Electrónica
Contacte con nosotros
Direcciones y teléfonos
Defensa del Contribuyente
Denuncia tributaria
Denuncia de pagos en efectivo
Consultas Informáticas
planes de regularización y lucha contra el fraude
Limitación de pagos en efectivo
Declaración de bienes y derechos en el extranjero
declaraciones INFORMATIVAS 2012
Aduanas e Impuestos Especiales
Acceda directamen
Calendario del contribuyente
Campañas
Censos, NIF y domicilio fiscal
Certificados electrónicos
Certificados tributarios
Descarga de programas de ayuda
Educación Cívico-Tributaria
Empleo público
Estadísticas tributarias
Folletos informativos
Gasóleo agrícola profesional
Le interesa conocer
Declaración Tributaria Especial (DTE)
Para qué sirve y quién tiene que declarar la DTE
Leer más
Histórico
Novedades en la aplicación de los tributos
Novedades tributarias introducidas por la Ley 16/2012, de 27 de diciembre, por la que se adoptan diversas medidas tributarias dirigidas a la consolidación de las finanzas públicas y al impulso de la actividad económica (BOE del día 28)
Se informa de las principales novedades tributarias introducidas por la Ley 16/2012, de 27 de diciembre, por la que se adoptan diversas medidas tributarias dirigidas a la consolidación de las finanzas públicas y al impulso de la actividad económica (BOE del día 28).
Leer más
Declaraciones informativas 2012
its2-aeat.linguaserve.net/AEAT.internet/en_gb/Inicio.shtml
Inicio
Agencia Tributaria - I...
Transit XV [SP 25] - 1...
Adobe Captivate
Red asociativa
15:28

Demo 9/9


Tax Agency - Home - Mozilla Firefox
Tax Agency - Home
Agencia Tributaria
Ministry of the Treasury and Public Administrations
Regional and local taxation
Non-resident Taxation
Useful links
Tax Agency
Individuals and employees
Businesses and professionals
Collaborators
This way we all improve
Electronic Office
All online formalities
Código fuente de: http://its2-aeat.linguaserve.net/AEAT.internet/en_gb/Inicio.shtml - Mozilla Firefox
By Provenance tags we can see where this translation comes from.
Línea 2, Columna 355
Inicio

(4) Challenges for RTMPS using ITS 2.0 Pre-editing

- Exploring best practices using ITS 2.0 data categories.
- Exploring training and methodology in content creation.
  - ITS 2.0 capabilities, usage, and training kits.
- Creating tools
  - Full HTML5 compliance and ITS 2.0 annotation facilities.
  - Writing tools for content quality, and controlled language for post-editing output adaptation.

(4) Challenges for post-editing and ITS 2.0 Post-editing

- Exploring best practices using ITS 2.0 data categories.
- Improving EDI-TA PE methodology and training.
  - Contextual, activation, and identification rules using ITS 2.0.
  - Specific language/system/client dependent and independent PE rules and training.
- Creating /extending tools.
  - Requirements from EDI-TA and showcase experience.
  - Specific PE functionalities and ITS 2.0 assistance and viewing functions for post-editors.

Q & A

More information about MLW-LT and use cases:

www.w3.org/International/multilingualweb/lt/wiki/Main_Page

www.multilingualweb.eu

pedro.diez@linguaserve.com

arle.Lommel@dfki.de

daniel.grasmick@lucysoftware.com

pablo.badia@linguaserve.com