



Directorate-General for Translation

Standards for multilingual web sites

MultilingualWeb.eu, 4-5 April 2011, Pisa, Italy

M.T. Carrasco Benitez



Multilingual Web Sites (MWS)

- Web sites with multilingual parallel texts
 - Texts that are translations of each other
- Most EU institutions are MWS
 - <http://europa.eu>
- Necessary for a real multilingual WWW
- Great practical relevance

Principles

- **Break nothing**
 - Forget about changes to the basic standards
 - Enter the dance without causing any missteps
- **It will take a lot of work and time**
 - To create the standards and programs

Two sides of the same coin

- **User**
 - Readers of web sites
 - **Admittedly, users are often content contributors**
- **Webmaster**
 - Builders of web sites
 - **An idealised function that are in fact many functions**

- **This is a simplification**
 - The real world is more fuzzy

Present situation

■ User

- No consistent multilingual user interface
- User have different experience when visiting unrelated MWS; and even in the same web site
 - **The user should always have the same experience in MWS**

■ Webmaster

- Bespoke systems are built each time
- It should be possible to built a new MWS with standard off-the-shelf software
 - **Existing software are not interoperable**
 - **Each CMS keep the data in their way**

Point of views

■ User

- Mostly monolingual or get a second-best language
- Language selection is just hurdle
- Like localization (L10N)

■ Webmaster

- Must manage a multilingual system
- Building and maintaining systems with parallel texts is complex
- Like internationalization (I18N)

Characteristics

- **User**
 - Grand public
 - A small additional complexity could be disastrous
- **Webmaster**
 - Professionals
 - A bit more adventurous, but not much

Interfaces

■ User

- Interface browser ↔ server
- Standards: well established
 - **Additional guidelines are needed**

■ Webmaster

- Interface server ↔ content
 - **Even the directory based content is not fully compatible**
- Standards: practically none

User 1

- User interface - best practice guide
- Interface browser ↔ server
- Main functionalities
 - Get the best page
 - See other variants (linguistic versions)
- Implementations
 - Browser side
 - Server side

User 2

■ Browser side

- Language button in the browser
 - Using HTTP header fields
 - Implementable with current specifications
- Transparent Content Negotiation (TCN), cookies, etc
 - Very good, but little used – it might be too smart
 - Language is just one of the dimensions in TCN

■ Server side

- A reserved URI – `http://example.com/mypage/variant`
 - This will show all the variants for the page “mypage”
 - Another approach is a federation of server a la Wikipedia
 - `http://de.wikipedia.org` - `http://fr.wikipedia.org`

User 3

- Translation request
 - As potentially available language – a variant
 - Immediately or referred
 - This will integrate machine translation
- Variants might have to be extended
 - Previous versions could be a variant dimension
 - Or a new *metaresource* concept would be required
- It should also be valid for wikis

Webmaster 1

- New standards as they do not exist
- Interface server ↔ content
 - General requirement - not limited to MWS
 - Web server ↔ content management systems (CMS)
- Solve MWS and leave hooks for the others
 - Otherwise, it might take forever

Webmaster 2

- **Content ↔ content generation**
 - Basic functionalities might be part of MWS
 - **Page generation**
 - A full translation system should be out of scope
 - **Further interfaces or hooks**
- **The full framework**
 - Authorship, Translation and Publishing Chain
 - **ATP-chain**
- **Translation is expensive**

Next step

- Create a new working group
- It should be in the I18N Activity of the W3C
 - What is the procedure?
- If the W3C is not interested
 - In an exiting one, such as the IETF
 - A dedicated initiative

To do list

- Two "standards"
 - *Multilingual web best practice guide for user interface*
 - *Multilingual web best practice guide for content management*
- Paper + running code
 - Reference development
- Implementing early mechanisms already in the standards and server side
 - Language button
 - Variant URI

End

- Interested parties have a look to
- MultilingualWebSites.org