Designing Localized Web Sites

Olga De Troyer and Sven Casteleyn

WISE

Department of Computer Science, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussel, Belgium {Olga.DeTroyer, Sven.Casteleyn}@vub.ac.be http://wise.vub.ac.be

Abstract. The term World Wide Web (WWW) emphasizes that the Web is global and many companies realize that this creates new opportunities. A considerable amount of literature on web site development stresses that, in order to attract and retain more customers, it is vital to create different versions of a web site and adapt those versions to the local communities they target. This process is usually called *globalization* and the different web site versions are called *localized* web sites. Although content management systems (CMS) for web sites provide support for multinational web sites, current web site design methodologies do not consider the issue of globalization. In general, the globalization effort is done after the web site is designed and implemented. This makes globalization much harder. In this paper, we show how to extend an existing web site design method, WSDM, to support the design of localized web sites.

1 Introduction

WWW pages are accessible from all over the world. This offers opportunities for companies and organizations to attract visitors from across the country borders and to do business with them. Two different approaches are possible to address this issue: develop one single web site to serve everyone or develop 'localized' web sites for particular localities. The 'one size fits all' approach may be appropriate for particular communities (like researchers) but in general it will be less successful. A considerable amount of literature on web site development stresses that, in order to attract and retain more customers, it is vital to localize a global web site, i.e. to create different web site versions and adapt those versions to the local communities they target. Members of a community do not only share a common language, but also common cultural conventions. Since measurement units, keyboard configurations, default paper sizes, character sets and notational standards for writing time, dates, addresses, numbers, currency, etc differ from one culture to another, it is self-evident that local web sites should address these issues. Some jokes, symbols, icons, graphics or even colors may be completely acceptable in one country, but trigger negative reactions in another country. Sometimes the style or tone of the site's text might even be considered offensive by a particular cultural entity, as a result of which the text needs

to be rewritten rather than merely translated. Next to culturally differences, it may also be necessary to adapt the content to regional differences, like differences in the services and products offered, differences in price, and differences in regulations.

The localization issue is not new. Localization of software has been done for years and research has been performed in this context. Nielsen and Del Galdo [5] stress that localization should encompass more than a 'surface-level' adaptation, by acknowledging underlying cultural differences. The role of culture in user interface has also been addressed by Evers and Day [11]. Barber and Badre [1] detected the existence of cultural markers, i.e. web design elements that are prevalent in web sites of a particular culture (e.g. color, icons, symbols). Sheppard and Scholtz [21] and Sun [23] conducted pilot studies to determine if the absence or presence of cultural markers affects the user's preference or performance. Cultural differences have also been investigated from an anthropological perspective, looking at the intangible nuances of a culture's social hierarchy, individualism, gender roles, attitude towards uncertainty and time-orientation ([18], [10]). This type of research is commonly approached through Hofstede's cross-cultural theory [15]. According to Hofstede, cultural differences are based in deeply rooted values that can be categorized along fundamental dimensions: power distance, collectivism-individualism, five masculinity-femininity, uncertainty avoidance, and long and short-term orientation. His research is based on a large-scale survey carried out between 1967 and 1973 and which covered 53 countries representing the major regions of the world. Marcus and Gould [18] attempt to apply those dimensions to global web interface design, providing suggestions and guidelines to produce successfully localized web sites.

In the context of localization, three different terms are used: globalization, internationalization and the term localization itself. According to LISA (Localization Industry Standards Association) [17] localization of a thing is adapting it to the needs of a given locale. Globalization is about spreading a thing to several different countries, and making it applicable and useable in those countries. Globalization is never all encompassing; you will never cover all the 600 languages on the planet today. In the context of web sites, globalization usually indicates the process of converting a web site to different languages and communities. Internationalization consists of all preparatory tasks that will facilitate subsequent localization. The purpose of internationalization is to make localization easier, faster, of higher quality and more cost-effective. It may include: creating illustrations in which the text can easily be changed; allowing space for translation into languages that require more space; abstracting content from markup; identification and isolation of culturally specific items. Localization adds cultural context to a previously internationalized web site and includes translation. Translation is only one of the tasks of localization but because it is the most cost expensive, time consuming and most vital task it is often used in the same context as globalization, internationalization and localization.

As for classical software, web site globalization is often done once the web site is completely developed and available for a particular community. Content Management Systems (CMS) for web sites usually provide support for multinational web production [4]. However such systems don't provide a methodology for designing the web site. Nowadays, it is recognized that a proper method should be used to design professional web sites. Although, several methods to design web sites (e.g. OOHDM [22], WebML[3], OO-H [13], UWE [16], WSDM[6]) exist, as far as we are aware of,

none of these methods takes globalization issues or one of its aspects (localization, internationalization or translation) into account during the design process. We believe that the globalization process could benefit from taking localization requirements into consideration while designing the web site. If web sites are designed with the need for localization in mind, it may be easier to actually realize globalization because the internationalization activities may already be considered and prepared for during the design process. For this reason, we have extended our own web site design method WSDM in order to support web localization. In this paper, we explain how this has been done. The paper is structured as follows. Section 2 provides a brief overview of WSDM. In section 3, we explain the extensions to WSDM and illustrate them with an example. Section 4 discusses the proposed solution and presents conclusions.

2 WSDM

The design process of WSDM follows an audience driven design philosophy i.e. the design is based on and driven by the requirements of the different types of users. Figure 1 gives an overview of the different phases of the method.



The method starts with the 'Mission Statement Specification'. The mission statement should identify the purpose, the subject and the target audience(s) of the web site. Next, the 'Audience Modeling' phase is performed. In the sub phase 'Audience Classification', the different types of users are identified and classified into so called *audience classes*. During the sub phase 'Audience Class Characterization' the different audience classes are characterized regarding age, experience level, language, etc. The next phase, the 'Conceptual Design', is composed of 'Task

Modeling' and 'Navigational Design'. During 'Task Modeling', the requirements identified during Audience Classification are elaborated and *task models* and *object chunks* are created to model the necessary information and functionality needed to fulfill the requirements. The 'Navigational Design' is used to design the overall conceptual structure of the web site and the navigational possibilities for each audience class. The fourth phase, the 'Implementation Design' contains three sub phases. The 'Page Design' translates the conceptual and navigational design to an actual page structure. The look & feel of the website is defined in the 'Presentation Design'. The 'Data Design' is only needed for data-intensive web sites. In case the data will be maintained in a database, the database schema is constructed. It is also possible that the data will not originate from a database but provided by means of another source (e.g. XML). For pure static web pages, the data design step can be omitted; the actual data will be supplied by the designer during the last step of the method, the actual implementation of the website.

3 Designing Localized Web Sites with WSDM

We will explain how WSDM is extended to allow modeling localized web sites by indicating how each of the different (sub) phases is adapted and illustrate this with an example. For the example, suppose a company wants a web site to offer their products for sale in the US and Belgium. In addition, the product managers of the company should be able to maintain the product information through the web site.

First, we introduce a new concept: *locality*. A locality describes a particular place, situation, or location. Localities are identified by means of a name and a label. Examples of localities are: the US, Japan, and the Flemish community in Belgium.

3.1 Mission Statement Specification

The mission statement is the starting point of the design. The mission statement should identify the purpose, the subject and the target audience(s) of the web site. If we want to be able to take localization into account during the design process, the mission statement should also mention the different localities for which the web site needs to be developed.

For the example web site, we can formulate the following mission statement:

"The web site should allow to increase the sales of the company in the company's country (Belgium) as well as in the US, by allowing people to search for information about the products and to buy the products online. In addition, the web site should act as an easy user interface for the product managers (located in Flanders) to maintain the product information and to keep track of the supply"

From this statement, we can identify:

- The purpose: increase the sales and provide support for maintaining the product information and the supply
- The subject: products of the company
- · The target audiences: potential buyers, and the company's product managers

• The target localities: US, Flanders and Wallonia (Flanders and Wallonia are the two main regions in Belgium each with a different language).

3.2 Audience Modeling

The target audiences identified in the mission statement should be refined into audience classes. Therefore, for each target type of user, the different functional- and informational requirements are identified. Users with the same information and functional requirements become members of the same audience class. Users with additional requirements form audience subclasses. In this way a hierarchy of audience classes can be constructed. The class *Visitor* is always the top of the audience class hierarchy, grouping the requirements all visitors have in common. During 'Audience Class Characterization' the different audience classes are characterized regarding age, experience level, language, etc.

In our extension of WSDM, we make a distinction between requirements that are typical for the audience class and requirements that are typical for a locality. Requirements that are typical for a locality will be specified separately from those typical for an audience class. Therefore, an additional sub phase called 'Locality Specification' has been introduced. The requirements specified in the Audience Classification should be independent of the localities'. In a similar way a sub phase, called 'Locality Characterization', is added to allow specifying the characteristics of the localities. The characteristics given for the different localities. The order in which the sub phases 'Audience Classification' - 'Audience Class Characterization' and 'Locality Specification' - 'Locality Characterization' is performed is not important because the information they allow to specify is independent.

To express the relationship between the audience classes and the localities another new sub phase is introduced, the '*Locality Mapping*'. This sub phase can only be performed after finishing the four other sub phases. See figure 2 for an overview of the different sub phases of the extended Audience Modeling phase. We now describe the new sub phases into more detail and illustrate them with the example web site.



Figure 2 – Audience Modeling

Locality Specification & Characterization

The requirements and characteristics that are typical for a locality are related to the language, culture, habits or regulations of the locality. Some examples of locality requirements are: an address should always include the state; for each price it should

be indicated if tax is included or not, and if it is not included the percentage of tax that need to be added should be mentioned; all prices should be expressed in EURO. Locality characteristics will typically deal with issues as language use, reading order, use of color, and use of symbols.

We illustrate the Audience Modeling phase with the example web site. There are three audience classes: 'Visitor', 'Buyer' and 'ProductManager', and there are three localities: 'US', 'Flanders' ('FL') and 'Wallonia' ('WA'). Due to space limitations we only give the most important requirements and characteristics.

- Audience Class 'Visitor': Need to be able to browse through the products for sale in the Visitor's country and obtain detail descriptions of those products. *Characteristics*: varying age and varying web expertise
- Audience Class 'Buyer': Need to be able to buy products that are for sale in the Buyer's country. *Characteristics*: varying age but older than 18 and varying web expertise
- Audience Class 'ProductManager': Need to be able to update information and supply information on the products for sale in both countries. *Characteristics*: familiar with the system and the Web
- Locality 'US': Each address should include a state; prices must be in US dollars. *Characteristics*: English speaking
- Locality 'FL & WA': Contact address of the company must be clearly mentioned; prices must be in EURO with tax included; it must be allowed to pay by means of bank transfer
- Locality 'FL': Characteristics: Dutch speaking
- Locality 'WA': Characteristics: French speaking

Locality Mapping

The localities need to be linked to the different audience classes. An audience class may span different localities, e.g. in the example 'Visitor' and 'Buyer' are applicable for all localities. Different audience classes may be needed for a locality, e.g. for the locality 'Flanders' we need the audience classes 'Visitor', 'Buyer' and 'ProductManager'. 'ProductManager' is only needed for the locality 'Flanders'. Therefore, in the 'Locality Mapping', for each locality the audience classes that need to be supported are enumerated. For our example, this results in the following sets. Flanders: {Visitor, Buyer, ProductManager}; Wallonia: {Visitor, Buyer}; and US: {Visitor, Buyer}. Graphically, this can be represented by drawing, for each locality, a box in the audience class hierarchy diagram that includes all the audience classes needed and label this box with the locality' label (see figure 3).



Figure 3 - Audience Modeling

3.3 Conceptual Design

We now describe how the sub phases of the 'Conceptual Design' phase are influenced by the localization requirement.

Task Modeling

During 'Task Modeling', a *task model* is defined for each requirement of each audience class. This is done using an adapted version of CTT diagrams [20] (CTT+). To create such a task model, each task is decomposed into elementary tasks and temporal relationships between tasks indicate the order in which the tasks need to be performed. For each elementary task an object model, called '*object chunk*', is created modeling the necessary information and functionality needed to fulfill the requirement of the elementary task. An extended form of ORM (Object Role Model) [14] (ORM+) is used as language for the object chunks.

In our extended version of WSDM, there are also requirements for the different localities. These requirements also need to be considered during task modeling. When constructing the task models, we need to inspect the locality requirements to check if additional or different steps are needed when decomposing a task. If a task needs to be completely different for a specific locality (which is rarely the case), a different CTT must be created and labeled with this locality. If only some additional steps are needed, then these steps are labeled with the localities for which they are needed. In our example, it was specified that in Belgium it must be possible to pay by bank transfer, in the US only payment by credit card is possible. This is indicated in the subtask 'payment method' by means of the labels 'FL' and 'WA' (see figure 4).

When constructing the object chunks, again, we need to inspect the locality requirements to check if additional information is needed. If this is the case, this information is added to the object chunk and labeled with the locality for which it is needed. If the object chunk is created for an elementary task that is labeled, the object chunk itself is labeled in the same way. Figure 5 shows the object chunk 'shipping details'; the State of an Address is only needed for the locality 'US'.

In the object chunks, we should also indicate which information is dependent on the locality. E.g., the value and the currency of the price may depend on the locality. Also the available set of products may be dependent on the locality. This is indicated by labeling the object types that are locality dependent (see figure 6). The set of 'Product' is different for the localities 'US' and 'FL&WA': other products might be available in the US compared to Belgium (FL&WA). 'ProductName' has the three localities US, FL, WA denoting that names of products differ in the three localities.



Figure 4 - CCT 'Payment Method' Figure 5 - Object Chunk 'Shipping Details'



Figure 6 - Object Chunk with Locality Dependent Object Types

In summary, the extended Task Modeling is as follows (the new steps are in italic): For each requirement of each audience class:

- 1. Define a task for the requirement
- 2. Decompose the task into elementary tasks and add temporal relations between the tasks (using CTT+).
 - If the decomposition of the task or a sub-task depends on the locality (expressed by means of a locality requirement) then either make different CTTs and label the CTT with the appropriated locality label or label the nodes that are specific for a locality with the appropriated locality label(s)
- 3. For each elementary task:

Make an object chunk that models the information and/or functionality required by the task (using ORM+)

- If the elementary task is labeled, then also label the object chunk with the same labels
- If a locality requires additional information or functionality (formulated by means of locality requirements) label the relationships that models these requirements with the label of this locality
- If the content of an object type is dependent on the locality, label it with all the localities in which the audience class is involved

Navigational Design

The Navigational Design defines the conceptual structure of the web site and models how the members of the different audience classes will be able to navigate through the site and perform the tasks. Because of the audience driven approach of WSDM, a *navigation track* is created for each audience class. A navigation track can be considered as a sub site containing all and only the information and functionality needed by the members of the associated audience class. If an audience track involves different localities, the track is labeled with all the localities involved (see figure 7).



Figure 7 - Navigation Design

Next, all audience tracks are combined into the *Conceptual Structure* by means of structural links. The structure defined between the audience tracks should correspond to the hierarchical structure defined between the audience classes in the audience class hierarchy. How this is done is described in detail in [2] and will not be discussed here. It is not influenced by localization.

A navigation track is constructed based on the task models developed for the audience class in the Task Modeling. How this is done exactly is outside the scope of this paper and can be found in [8]. Roughly speaking, we can say that for each task model a task navigation model is created and that this task navigational model is based on the structure of the CTT for this task. A task navigational model is composed of *components* and *links*. Because we have labeled the tasks and sub-tasks in the task models where necessary, it is easy to indicate in the task navigational models which components and links are locality dependent.

3.4 Implementation Design

We now describe the impact on the sub phases of the 'Implementation Design'. The 'Page Design' translates the conceptual and navigational design into an actual page structure. The actual page structure is obtained by grouping (or splitting) the components and links from the navigational design into pages. Usually the page structure will be independent of the locality, i.e. for each locality the page structure will be the same. However, if some task models are very different for different localities, a different page structure may be needed. In that case, alternative page structures must be defined and each page structure must be labelled with the locality to which it applies.

The presentation design defines the general look and feel of the web site and for each page in the page design a template is constructed defining the layout of the page. Clearly, in this sub phase we must take the localization characteristics formulated in the 'Localization Characterization' into consideration. Each page belongs to exactly one audience class. Therefore, for each page and for each locality needed for this audience class, a different template should be created. How to take the localization characteristics into consideration in the presentation design is not treated here. This is described extensively in the literature about localization (see e.g. [12], [17], [19]).

When localization is needed, the easiest way to maintain the data is by means of a database (or CMS). In WSDM, the database schema is obtained (during Data Design) by integrating all object chunks into a single schema [9] and map this schema onto a database schema (see e.g. [14]). Here, we also need to take the labeling of the object chunks into consideration when mapping the object chunks into a database schema. We will do this for a mapping to a relational database. Different situations are possible:

- 1. **The complete object chunk is labeled**. This means that the information modeled in the chunk is only needed for the label's locality. This can be reflected by using table (and/or attribute) names that include the locality's label. E.g. TABLE ProductSupply_FL (ProductId INTEGER,)
- 2. A relationship is labeled. This means that the information modeled by the relationship is only needed for the label's locality. This can be reflected by using an attribute name that includes the locality's label. Furthermore, the attribute must allow for null values. E.g. TABLE DeliveryAddress (Street STRING NOT NULL, Nr INTEGER NOT NULL, ..., State_US STRING)
- 3. **An object type is labeled**. This means that the content of the object type is dependent on the locality. We distinguish between entity types and label types:
 - For an entity type, we can add, per locality, a boolean-type attribute to the primary key to indicate if an entity should be considered in the locality or not. E.g. TABLE Product (productId INTEGER, availability_US BOOLEAN, availability_Fl&WA BOOLEAN, ...). The value TRUE for the attribute 'availability_US' indicates that the product is available in the US.
 - For a label type, we have to provide an attribute per locality needed. E.g. TABLE Product (productId,, productName_US, productName_FL, productName_WA,...). An alternative solution is to put the locality dependent attributes in a separated table. E.g. TABLE Product_locality (productId INTEGER, locality STRING, productName STRING, price REAL, ...)

Please note that the mapping described here is only one possible solution. In the proposed mapping, information for different localities is maintained in the same table. If (nearly) all information is locality dependent, it may be better to define different tables for different localities. E.g. Table Product_US (productId, ..., productName, ...) and Product_FL (productId, ..., productName, ...). In this way it is also easier to physically separate the information for the different localities in different databases.

4 Conclusions & Discussion

We have shown how to extend an existing web site method, WSDM, in order to support the localization of web sites. First, a new concept 'locality' has been introduced. A locality describes a particular place, situation, or location. Then, the different phases of the method were adapted to allow for the specification of different localities. The mission statement should also answer the question: What are the different target localities? The Audience Modeling was extended with some extra sub phases to allow to describe the specific requirements and characteristics of the different localities and to link the localities to the audience classes. During Conceptual Design, the locality requirements are taken into consideration and integrated in rest of the design. In the Implementation Design, for each locality a different Presentation Design may be needed and when using an underlying database, the Data Design should take into consideration that for some tables and attributes more than one version is needed. Although, the extension has been done for WSDM, we believe that the principle of using localities as a starting point is generally applicable and may therefore be used by other web site design methods.

We also discuss some of the limitations of the method proposed. First of all, the method is not appropriate for the 'one size fits all' approach to globalization. In that case, it is better to use standard WSDM and to mention in the audience class characterization that the members are from different localities. Then, during Presentation Design, these issues can be taking into consideration.

Sometimes, the approach needed is a mixture of localization and the 'one size fits all' approach. An example of this is our own university web site. 'Potential Students' and 'Researchers' are two of the target audiences. For researchers, we want to use the 'one size fits all' approach, but for potential students we want to localize. Our local students should be addressed in their local language; foreign students should be addressed in English and should only see information that is applicable to them (e.g. only the English programs). We can achieve this dual approach by defining two localities ('Locals' and 'Foreigners') for the audience class 'Potential Students' and no localities for the audience class 'Researchers'. In the audience characterization of the 'Researchers' we can state that the language must be English and that it is an international audience.

Next, the approach that we follow assumes that the differences in type of content and structure between the localized web sites are rather small. If the local web sites need to be substantially different, this method may not work well. In that case too many requirements from the audience classes need to be moved to the locality specification and those requirements may not all express requirements that are "specific" for the locality. Then, it may be better to define different audience classes, e.g. 'US-Tourists' and 'Europe-Tourists' instead of one audience class 'Tourists' and two localities 'US' and 'Europe'.

5 References

- 1. Barber, W. and Badre, A.: Culturability: The Merging of Culture and Usability. In: Proceedings of the 4th Conference on Human Factors and the Web, at http://www.research.att.com/conf/hfweb/proceedings/barber/ (1998)
- Casteleyn, S. and De Troyer, O.: Structuring Web Sites Using Audience Class Hierarchies. In: Proceedings of DASWIS 2001 (ER) workshop. Yokohama, Japan (2001)
- 3. Ceri S., Fraternali P., and Bongio A: Web Modeling Language (WebML): a modeling language for designing Web sites. In: WWW9 Conference, First ICSE Workshop on Web Engineering, International Conference on Software Engineering (2000)
- 4. CMSWatch, at http://www.cmswatch.com (accessed date: May 2004)

- 5. Del Galdo, E. M. and Nielsen, J. (Eds.) International User Interfaces. John Wiley & Sons, New York (1996)
- De Troyer, O. and Leune, C.: WSDM : A User-Centered Design Method for Web Sites. In: Computer Networks and ISDN systems, Proceedings of the 7th International WWW Conference, Elsevier (1998) 85-94
- 7. De Troyer, O.: Audience-driven web design. In: Information modeling in the new millennium, Eds. Matt Rossi & Keng Siau, IDEA GroupPublishing (2001)
- De Troyer, O., Casteleyn, S.: Modeling Complex Processes for Web Applications using WSDM. In: Proceedings of the IWWOST2003 workshop, Oviedo, Spain (2003)
- De Troyer, O., Plessers, P. ,Casteleyn, S.: Conceptual View Integration for Audience Driven Web Design. In: CD-ROM Proceedings of the WWW2003 Conference, Budapest, Hongary (2003)
- Dormann, C. and Chisalita, C.: Cultural Values in Web Site Design. In: Proceedings of the 11th European Conference on Cognitive Ergonomics ECCE11 (2002)
- Evers, V. and Day, D.: The Role of Culture in Interface Acceptance, In: Proceedings of Human Computer Interaction, Interact'97. Chapman & Hall: London, (1997) 260-267
- 12. Globalization, Localization, Internationalization and Translation (GLIT), at http://philip.pristine.net/glit/ (accessed date: January 2004)
- Gómez, J., Cachero, C., and Pastor O.: Modelling Dynamic Personalization in Web Applications, In: Third International Conference on Web Engineering (ICWE'03), LNCS 2722, pages 472-475. Springer-Verlag Berlin Heidelberg (2003)
- 14. Halpin, T.: Information Modeling and Relational Databases, 3rd edition, Morgan Kaufmann (2001)
- Hofstede, G.: Cultures and Organizations: Software of the Mind. McGraw-Hill, New York (1997)
- Koch, N and Kraus, A.:. The expressive Power of UML-based Web Engineering. In Second International Workshop on Web-oriented Software Technology (IWWOST02), Schwabe D., Pastor, O. Rossi, G. and Olsina L. (eds) (2002) 105-119
- Localization Industry Standards Association (LISA), at <u>http://www.lisa.org</u> (accessed date: January 2004)
- Marcus, A. and West Gould, E.: Crosscurrents: Cultural Dimensions and Global Web User-Interface Design. ACM Interactions 2 (4) (2000) 32 – 46
- McCracken D., D. Wolfe R.J.: User-Centered Website Development. A Human-Computer Interaction Approach, Pearson Prentice Hall (2004)
- Paterno, Mancinii and Meniconi,: ConcurTaskTress: a Diagrammatic Notation for Specifying Task Models. In: Proceedings of INTERACT 97, Chapman & Hall (1997)
- 21. Sheppard, C. and Scholtz, J.: The Effects of Cultural Markers on Web Site. In: Proceedings of the 5th Conference on Human Factors and the Web (1999)
- Schwabe, D and Rossi, R.: An Object Oriented Approach to Web-Based Application Design. Theory and Practice of Object Systems 4(4). Wiley and Sons, New York, ISSN 1074-3224) (1998)
- Sun, H.: Building a Culturally Competent Corporate Web Site: An Exploratory Study of Cultural Markers in Multilingual Web Design. In: Proceedings of the 19th Annual International Conference on Computer Documentation. ACM Press: New York (2001).