Web Globalization and WSDM Methodology of Web Design

Thomas Appelmans Promotor: Prof. Dr. O. De Troyer

Graduation thesis submitted to obtain a License Degree in Applied Computer Science



WISE - Web & Information System Engineering Department of Applied Computer Science Vrije Universiteit Brussel Belgium Academic Year 2003 - 2004

Web Globalisatie en de WSDM Web Design Methode

Thomas Appelmans Promotor: Prof. Dr. O. De Troyer

Verhandeling voorgedragen tot het behalen van de graad Licentiaat in de Toegepaste Informatica



WISE - Web & Information System Engineering Departement Toegepaste Informatica Vrije Universiteit Brussel België Academiejaar 2003 - 2004

Dedicated to

my sister and my parents.

Web Globalization and WSDM Methodology of Web Design

Thomas Appelmans

Graduation thesis submitted to obtain a License Degree in Applied

Computer Science Academic Year 2003 - 2004

Abstract

The World Wide Web has an international audience and companies realize that if they want to do descent World Wide business they should provide a Web Site that is adapted to the needs of that foreign audience. This goes far beyond translation. Customers seem to prefer Web Sites that are specially designed for their needs and take their cultural properties into consideration.

There are various web design methodologies and strategies, all using different approaches, but none of them incorporates global availability and adaptability. The aim of this study is to integrate this globalization into an existing methodology, WSDM (Web Site Design Method). WSDM is developed by the Web & Information Systems Engineering research group, WISE, at the Vrije Universiteit Brussel.

The incorporation of the globalization issue happens in the different phases of the methodology and takes into account the several issues involved in it. In the first phase the different Localities are identified. In the second phase they are specified and mapped onto the Audiences. The third phase integrates the Localities into the Conceptual Model.

Keywords: Globalization, Localization, WSDM, Conceptual Modeling.

Web Globalisatie en de WSDM Web Design Methode

Thomas Appelmans

Verhandeling voorgedragen tot het behalen van de graad Licentiaat in de Toegepaste Informatica Academiejaar 2003 - 2004

Samenvatting

Het World Wide Web heeft een internationaal publiek en bedrijven realiseren zich dat als ze wereldwijd zaken willen doen ze een Web Site moeten voorzien die aangepast is aan de noden van dat 'vreemde' publiek. Dit gaat veel verder dan enkel vertalen. Het blijkt dat klanten Web Sites willen die de illusie wekken speciaal voor hen gemaakt te zijn en die rekening houden met hun culturele eigenheden. Er zijn verschillende web design methodologiën en strategiëen. Allen hebben ze een verschillen aanpak, maar geen enkel integreert globale toegankelijkheid en aanpasbaarheid. Het doel van deze studie is om globalisatie te integreren in een bestaande methodologie, WSDM (Web Site Design Method). WSDM is ontwikkeld door de Web & Information Systems Engineering onderzoeksgroep, WISE, aan de Vrije Universiteit Brussel.

De integratie van die globalisatie gebeurt in de verschillende fases van de methodologie rekening houdend met de verschillende aspecten die erbij betrokken zijn. In de eerste fase worden de verschillende Lokaliteiten geïdentificeerd. In de tweede fase worden ze gespecificeerd en gelinkt met het Doelpubliek. De derde fase integreert de Lokaliteiten met het Conceptueel Model.

Sleutelwoorden: Globalisatie, Lokalisatie, WSDM, Conceptuele Modellering.

Acknowledgements

I thank my promotor, Prof. Dr. Olga De Troyer, for providing the subject of this thesis and her valuable advice, support, patience and encouragements.

I thank Abdalghani Mushtaha of the WISE research group for providing me with reference work and advice.

My thanks go to my friends and co-students for the fine hours they have given me and their encouragements.

I thank my parents for their moral and financial support. I thank them for giving me the opportunity of studying at the Vrije Universiteit Brussel.

Contents

	Abs	stract		iv
	San	nenvat	ting	v
	Ack	nowlee	dgements	vi
1	Intr	oducti	on	1
	1.1	Globa	lization, Localization, Internationalization and Translation. $\ .$.	1
	1.2	The n	ecessity of globalization in Web Engineering	6
		1.2.1	Internet Usage Statistics	6
		1.2.2	Differences between localities	9
		1.2.3	Reasons for Web Globalization.	9
2	Exis	sting V	Veb Site Design Methods and Globalization.	11
	2.1	Objec	t-Oriented Hypermedia Design Method (OOHDM)	11
		2.1.1	Conceptual Design.	12
		2.1.2	Navigational Design.	13

		2.1.3	Abstract Interface Design	13
		2.1.4	Implementation	13
	2.2	Web N	Modeling Language (WebML)	14
		2.2.1	Data Model	15
		2.2.2	Hypertext Model	15
		2.2.3	Presentation Model	16
	2.3	Objec	t-Oriented Hypermedia (OO-H).	16
		2.3.1	Class Diagram and Use Case Diagram.	17
		2.3.2	Navigation Access Diagram (NAD)	17
		2.3.3	Abstract Presentation Diagram (APD) and refinements	18
	2.4	UML-	based Web Engineering (UWE)	18
		2.4.1	Conceptual Design.	18
		2.4.2	Navigation Design.	19
		2.4.3	Presentation Design.	19
	2.5	Conclu	usion	19
3	Wel	o Site	Design Method (WSDM).	20
	3.1	Chara	acteristic of WSDM	20
		3.1.1	Different Approach	20
		3.1.2	Distinction between Conceptual Design and Presentation De-	
			sign	21

	3.2	WSDM Overview	 21
		3.2.1 The Mission Statement.	 24
		3.2.2 Audience Modeling	 25
		3.2.3 Conceptual Design	 30
		3.2.4 Implementation Design	 32
		3.2.5 Implementation	 33
	3.3	Conclusion.	 34
4	Inte	egrating Globalization into WSDM approach.	35
	4.1	Mission Statement Specification	 35
	4.2	Audience Modeling Extended	 38
		4.2.1 Locality Specification	 38
		4.2.2 Locality Mapping	 39
	4.3	Conceptual Design Revisited.	 40
		4.3.1 Task Modeling	 41
		4.3.2 Navigation Design	 42
	4.4	Implementation Design.	 43
5	Con	nclusions	44
	Bib	liography	46
	Apr	pendix	49

A	\mathbf{Cul}^{r}	tural Differences that harden the Localization Process.	49
	A.1	Text Considerations.	49
	A.2	Text translation.	49
		A.2.1 Left to Right, Right to Left	50
		A.2.2 Date, Time, Size/Weight measurement, Currencies, etc	50
	A.3	Color, Icons, and Images	52
	A.4	Hofstedes Dimensions of Culture and Web Globalization	54
		A.4.1 Power Distance	55
		A.4.2 Collectivism versus Individualism	55
		A.4.3 Masculinity versus Femininity	56
		A.4.4 Uncertainty Avoidance	57
в	Imp	elementation Techniques.	59
	B.1	Abstract Content from Layout Using CSS	59
	B.2	Language Files for Common Strings.	60
	B.3	Localized Content Using DB Locality Flags	62

List of Figures

1.1	Internet Users Worldwide	7
1.2	Internet Usage Growth per Regio	8
2.1	Object-Oriented Hypermedia Design Method	12
2.2	Web Modeling Language	14
2.3	Object-Oriented Hypermedia	16
2.4	UML-based Web Engineering	18
3.1	Overview of the Web Site Design Method	22
3.2	Web Site Design Method Mission Statement Specification	24
3.3	Web Site Design Method Audience Modeling	26
3.4	Web Site Design Method Audience Classes - People who are involved.	27
3.5	Web Site Design Method Audience Class Hierarchy	29
3.6	Web Site Design Method Conceptual Design	30
3.7	Employer Profile ConcurTaskTree	31
3.8	Employer Profile ORM	32

3.9	Conceptual Structural Model	32
3.10	Employer Profile Task Navigation Model	33
3.11	Web Site Design Method Implementation Design	33
3.12	Web Site Design Method Implementation	34
4.1	Global Web Site Design Method Overview	36
4.2	Global Web Site Design Method Audience Modeling Extended	38
4.3	Global Web Site Design Method - Locality Sets	40
4.4	Global Web Site Design Method - Extended Audience Class Hierarchy	41
4.5	Global Web Site Design Method - Customer Address Object Chunk Example	42
4.6	Global Web Site Design Method - Locality Attribute	43
4.7	Global Web Site Design Method - Conceptual Structural Model	43
A.1	The English Al Jazeera Web Site	51
A.2	The original Al Jazeera Web Site	52
A.3	The "Home Page" symbol	53
A.4	Malaysia Tourism Web Site - Power Distance	56
A.5	Guatemalan McDonalds Web Site - Collectivism	57
A.6	Women Excite Web Site - Masculinity	58

List of Tables

A.1 Color Symbolism in Different Cultures							•			54
-------------------------------------------	--	--	--	--	--	--	---	--	--	----

Chapter 1

Introduction

In this introductory chapter we explain in the first section the definition of terms commonly used in the context of globalization. In the second section the reasons for the purpose of this research are explained.

1.1 Globalization, Localization, Internationalization and Translation.

Here we will discuss the terms globalization, localization, internationalization and translation. Several definitions can be found in several sources Therefore, it is important to discus the meaning of the different terms for that no misconceptions can occur when reading this thesis.

We start with the commonly accepted meaning of the terms as we can find it in a dictionary. According to the Merriam-Webster online dictionary [1] the meaning of the different terms is the following:

• globalization: to make global

- localization: to make local; orient locally
- internationalization: to make international
- translation: rendering from one language into another

The term translation is rather clear and doesn't need any more explanation, but the other three do. To clarify these definitions more in terms of their meaning in the dictionary we looked up some other definitions:

- global: of, relating to, or involving the entire world, worldwide
- local: of, relating to, or characteristic of a particular place
- international: of, relating to, or affecting two or more nations; active, known, or reaching beyond national boundaries

With these additional definitions the general meaning of the terms globalization, localization, internationalization and translation are intuitively clear. The next question is of course: What do these terms mean in the context of software development and more specific in terms of Website Engineering?

With an exception of the term translation (which is pretty obvious) there is not a single answer to the previous question. In the following paragraphs a brief overview of the various definitions you can find is given.

LISA.org, an organization on Localization Industry Standards Association [2] starts their explanation of the terms with the fact that all of the terms are pretty old. Translation found its origin already in the 14th century, whereas localization, internationalization and globalization find their origin respectively in 1792, 1864 and 1944. Globalization was thus founded towards the end of World War II. They add another word to list of terms in order to explain localization in the context of the industry we are describing here. The dictionary meaning of locale is a place or locality especially when viewed in relation to a particular event or characteristic, but to make it more compatible with our context it is slightly shifted to locale identifies a group of people by their common language and cultural conventions. For example in Belgium you have three locales, you have a French speaking part, Wallonia. You have a Dutch speaking part, Flanders and you have a German speaking part. With the definition of locale in the back of our head their definition of localization is the following:

Localization of a thing is adapting a thing to the needs of a given locale. [2]

Another source, on Langues Connaissances Pouvoir [3] poses that translation is the central activity of localization, but goes beyond the literal translation. Numerous locale details such as time zones, local colors, gender roles, currency, national regulations and holidays, cultural sensitivities, product or service names, and geographic examples must also be made local. Their definition is slightly more industry specific than the more general definition of LISA.org.

Localization primarily includes:

- Translating text content, software source code, web sites, or database content; machine translation may be used in early stages.
- Adjusting graphic and visual elements and examples to make them culturally appropriate.
- Post-production quality control of contents, systems and the integrated product.

WorldLingo.com, a company which main business is localization of web sites, states the following definition in their FAQ, specific for web site localization [4]:

This refers to the process whereby your web site is not only translated into another language but optimized for that particular market making it linguistically and culturally appropriate. As we looked at more context-specific definitions of localization we observed that most of them are based on the definition of the Localization Industry Standards Association (LISA) or are derived from their definition of localization.

Globalization has become a hot topic today in the context of economic globalization, which is countered by the anti-globalization groups. Although this has nothing to do with the definition we are trying to find for the term both parties can agree on the general meaning, namely to make global. According to LISA.org globalization is simply about spreading a thing to several different countries, and making it applicable and useable in those countries. They also suggest that globalization is never all-encompassing; you will never cover all the 6000 languages on the planet today. A globalization effort targets mostly six languages at the time.

Globalization of a thing consists in adapting a thing to the needs of N locales. [2]

Formula: globalization = N * localization

Another more marketing related definition of globalization can be found on the website Langues Connaissances Pouvoir [3].

Globalization is an approach to business strategy that aims to address all of the logistical and organizational challenges an enterprise faces as it expands its supporting content, asses and message across cultures and markets to new clients.

If you are going to localize something a couple of times it is obvious that, to prepare yourself, you are going to take some steps beforehand to make the job easier. Even if you will do it only once; when you want to maintain your product on regular basis, internationalization is simply inevitable. Internationalization is there to make the process of translation and localization easier.

Internationalization of a thing consists in any and all preparatory tasks that will facilitate subsequent localization of said thing. The purpose of internationalization is to make localization easier, faster, higher quality and more cost-effective. [2]

A remark on this definition is that it differs from the more usual definitions you find, namely, internationalization consists of making something language independent and internationalization consists in externalizing localizable items. The flaws in these definitions can be summarized as follows:

- they are tasks, not fundamental definitions;
- they are an incomplete list and too specific e.g. the definition isnt suited for the internationalization of documentation nor for tasks specifically involved in software internationalization;
- they still fail to describe the true nature of the activity;

Though they agree on the general thought that internationalization is planning and implementing products and services so that they can easily be localized for specific languages and cultures, a more practical approach of the definition is found on the website Langues Connaissances Pouvoir [3].

Internationalization may include:

- Creating illustrations for documents in which the text can easily be changed to another language and allowing expansion room for this purpose.
- Allowing space in user interfaces for translation into languages that require more space.
- Creating print or website graphic images so that their text labels can be translated inexpensively.
- Abstracting content from markup in a web application and software.
- Support for international character sets.

• Use of examples with global meaning.

The different sources where globalization, localization, and internationalization are discussed relating to each other agree upon the fact that globalization can only be achieved by the incorporation of internationalization and localization [3].

The efficient process can thus be summarized in the following formula [2]:

Globalization = Internationalization + N * Localization

The three definitions of these terms all involves of a variety of specific tasks. Translation actually refers to specifically linguistic operations, replacing expressions in one natural language into those of another [2]. This means that the definition of translations consists of only one specific task, but because it is the most cost expensive, time consuming and most vital task it is often used in the same sentence as the previous three terms.

1.2 The necessity of globalization in Web Engineering.

In this section we discuss the reasons for globalizing web sites. We will do this by finding an answer on the following questions: How many people use the Internet and where do they come from? Are the differences between Internet users this significant that we need localization? What are the main reasons we should use globalization in Web Engineering?

1.2.1 Internet Usage Statistics.

In 1957 the USSR launches Sputnik was the first artificial earth satellite. In response to that event, the US forms the Advanced Research Projects Agency, better known as ARPA. Thirteen years later four universities form the first 4 nodes of ARPANET with the purpose of fast exchange of information. The roots of the Internet were laid [5].



Figure 1.1: Internet Users Worldwide [7]

Today the Internet is the youngest and fast growing media in the world. In Figure 1.1 you can see the evolution during the past years. There are several reasons for this rise of Internet usage and the demand growing demand for bandwidth [6].

- The network infrastructure becomes larger and better.
- New technologies provide high capacity with spectacular increases of bandwidth.
- Applications using the Internet become more effective. They get more userfriendly and thus more people can use them and know how to use them.
- The content of the web has changed and is changing with users surfing longer on the Net and they bring in more content. The immense amount of information that is available on this worldwide network attracts more people.

As stated in [8] the United States (and subsequently the United Kingdom) were well ahead of the rest of the world when it came to being online. Thus, historically, English has been the lingua franca of the Internet. This however is changing very rapidly. The rest of the world is starting to catch up, or has already catched up. The global network, as we usually refer the Internet to, was longtime a North-American phenomena. In November 1998 still 60% of all Internet users lived in the USA and Canada. In 2002 this figure was already reduced to 30% of the worldwide Internet population.



Figure 1.2: Internet Usage Growth per Regio [7]

The fastest growing Internet communities are Asia and China. In 1997 China counted 620,000 Internet users, while today (in 2004) this number has increased to 170 million [9]. The use of Internet in Africa is still in the take off phase. In 2002 only 1 out of 250 Africans used the Internet, this is because of the more developed parts of the African Continent, South Africa and North Africa. In North America and Europe one in two uses the Internet [7].

The reasons why in some areas the Internet usage is much higher than in others are the following:

- Africa has fewer telephone lines than cities like Tokyo or Manhattan.
- Some governments ban the use of Internet (as in Myanmar) or restrict and control the access to it (as in Vietnam).

• Even if people have access, most of them living in poor countries would be excluded because of there illiteracy and lack of basic computer skills [7].

So what we can conclude from these paragraphs is that the Internet is growing very fast and that the demographics of Web users has changed dramatically over the past years [10].

1.2.2 Differences between localities.

Now that we know that there is a wide diversity of people using the Internet, we can ask the following question: Are the differences between the Internet users and their localities that significant that we need to create totally different websites for them? Perhaps we can rephrase this question:

Consider your favorite website. How might this website be understood and used in New York, Paris, London, Beijing, New Delhi, or Tokyo, assuming that adequate verbal translation was accomplished? [11]

Is it possible that the user is offended, confused or even alienated with the metaphors, the mental model, the navigation, the interaction, or the appearance that is used in the website?

Localization is much more than mere translation. Appendix A makes this clear with guidelines and examples.

1.2.3 Reasons for Web Globalization.

The first reason for a Web Site to be adapted to another locale is that you want to make your information available to a wider variety of people. Though the main reasons why people/companies want to globalize their Web Site are purely commercial. Underneath a list with reasons collected from several sources:

- Over 100 million people access the Internet in a language other than English. [9]
- 50.4% of Web Users speak a native language other than English. [9]
- Web users are four times more likely to purchase from a site that communicates in the customers language. [12]
- Visitors stay for twice as long. (Site stickiness is doubled). [13]
- Today 50% of all online sales occur outside the US. [14]
- Almost one-third of web sites are presented in a language other than English. [14]
- 37 million Americans do not speak English at home. [15]

Chapter 2

Existing Web Site Design Methods and Globalization.

There are several known existing methods to design websites. There are some older methods like, HDM, and RMM that we will not mentioned here. We will cover the more recent methods like OOHDM, WebML, OO-H and UWE. In the following chapter we will shortly describe each of these methods to present an overview of the methodologies.

2.1 Object-Oriented Hypermedia Design Method (OOHDM).

Object-Oriented Hypermedia Design Method or OOHDM [16,17] is a model-based approach for building hypermedia applications and thus also for engineering websites. Daniel Schwabe, Gustavo Rossi, and Simone D.J. Barbosa first introduced it in 1995 to meet the growing interest in hypermedia design approaches. This design method comprises a four-step process, namely requirements gathering, conceptual design, navigational design, abstract interface design, and implementation. Each step describes a particular design concern and in each step an object-oriented model is built using or enriched using previous iterations.



Figure 2.1: Object-Oriented Hypermedia Design Method [17]

In the following paragraphs we shortly describe each of the steps.

2.1.1 Conceptual Design.

At this stage it is the intention to model the semantics of the application domain. This is done using any well known object-oriented modeling construct and design patterns. No particular method has been specified in OOHDM to produce this Conceptual Design; UML or OMT are suggested. The product of this phase (a class and instance schema built out of classes, subsystems, relationships, and attribute perspectives) can be built using classification, aggregation, generalization, and specialization. The semantics of the application domain need to be approached as neutral as possible, with very little consideration of users and tasks.

2.1.2 Navigational Design.

From the Conceptual Design we can derive the Navigational Design. The main purpose of this phase is building the navigational structure of the application taking the user profiles and tasks into account. The navigational design consists of two schemas, the navigational class schema with nodes, links, indices, and guided tours and the navigation context schema, that is derived from the class schema. The nodes and the links are respectively derived from the conceptual classes and conceptual relationships defined in the Conceptual Design phase. Because of the use of nodes and links the Navigational Design schema represents the path the user can walk during the use of the application. Depending on the view on the application domain different Navigational Models can be representing the same Conceptual Model.

2.1.3 Abstract Interface Design.

The Abstract Interface Model is built using perceptible objects by means of interface classes. These, for their part, are recursively defined or are defined as aggregations of primitive classes (e.g. a text field, a button, etc.). The interface objects depend on the navigational objects defined in the previous step and provide a perceptible appearance. Like with every step in this methodology this one takes us in a higher degree of independence from user-interface technology. Again different Abstract Interface Models can be built from the same Navigational Model.

2.1.4 Implementation.

In the Implementation stage an implementation environment has to be chosen and the different Navigational and Abstract Interface Models have to be mapped on concrete Implementation Models.

2.2 Web Modeling Language (WebML).

The Web Modeling Language is a visual notation for specifying the composition and navigation features of hypertext application.

That is the definition of the Web Modeling Language as stated in [18, 19]. Aldo Bongio, Stefano Cery, and Piero Fraternali presented this methodology in 2000. In the meantime the methodology has evolved and in the following paragraphs a brief description of WebML will be given.

The Web Modeling Language is, like OOHDM [16], a model-driven web design method. In WebML, different models allow the description of a website: the Data Model, the Hypertext Model, and the Presentation Model.



Figure 2.2: Web Modeling Language [19]

A brief description of the different models follows:

2.2.1 Data Model

The WebML Data Model schematizes the data used in the web application so that it is suitable for data design (like in e.g. database design, software engineering, etc.). This is done using a language suited for data modeling. WebML does not propose a language, but is compatible with several of them. The WebML method describes entities and relationships as respectively containers of data and semantic connections between entities.

An entity represents a description of the common features of set of objects of the real world.

E.g. Person, Car, Artist, etc.

A relationship represents the semantic connections between entities.

E.g. the association between an artist and his/her album. XML is used to make a textual representation of the Data Model.

2.2.2 Hypertext Model

The Hypertext Model describes two aspects of the website, namely Composition and Navigation. WebML uses units to represent atomic content elements. There are seven types of them: data, multi-data, index (and its variants multi-choice and hierarchical), entry, and scroller. In the composition aspect of the Hypertext Model we describe which pages compose the hypertext and which content units make up a page. Links make up the navigation of the website. There are two kinds of links, contextual links and non-contextual links. The first kind carries context information, while the second kind has no associated context information.

2.2.3 Presentation Model

Defining the look and the feel of the web pages happens in the last model, the Presentation Model. The methodology does not include a specific model for expressing this phase. Since WebML can be represented using XML, the Hypertext Model from the previous phase can be transformed into a page written in a concrete implementation language such as JSP or ASP.NET.

2.3 Object-Oriented Hypermedia (OO-H).

Object-Oriented Hypermedia Method [20,21] is meant to do interface modeling. Its purpose is to model and integrate high quality interfaces with preexisting business logic modules.



Figure 2.3: Object-Oriented Hypermedia [21]

The OO-H Design Process consists out of different steps:

- 1. Start from Class Diagram and Use Case Diagram.
- 2. Create a NAD (Navigation Access Diagram) instance.
- 3. Generate a Default APD (Abstract Presentation Diagram).

- 4. Apply refinements to APD.
- 5. Generate Deliverables.

We will discuss brief each of the Design Process steps.

2.3.1 Class Diagram and Use Case Diagram.

This is the first step in the OO-H Design Process. It consists of drawing relevant Class Diagrams and Use Case Diagrams using a non-specified data modeling language (e.g. UML). The Class Diagram shows the static structure of the model, it gives an overview of the different object classes and their relationships [22]. The Use Case Diagram depicts the users involved in the system and their actions. It represents a specific way of using the systems [22].

2.3.2 Navigation Access Diagram (NAD).

The Navigation Access Diagram adds navigation and interaction features to the Use Case and Class Diagrams. In order to define the constraints a subset of the standard UML is used, the object constraint language. At least one NAD should be provided for each identified user profile. There should be as many NADs as different views of the system. An NAD consists of four types of constructs: navigation classes, navigation targets, navigation links, and collections. The navigation classes are domain classes with extra user information (permissions and navigational requirements). The navigation targets group the elements of the model that are needed to cover a user navigation requirement. The navigation links depicts the paths a user can follow throughout the system. The collections represent hierarchical connections between navigational patterns. The Navigation Access Diagram contains much more detailed information. For more information we refer to [21].

2.3.3 Abstract Presentation Diagram (APD) and refinements.

The Abstract Presentation Diagram reflects the abstract page structure of the interface and can be derived from the NAD. It is based on the concept templates, which are expressed as XML documents. The refinement process modifies the previous APD structure. This is a process that has been greatly simplified by the development of a series of patterns and applications that can be used for this purpose.

The Deliverable step implies the development of the concrete application in a chosen implementation language using the previously generates diagrams.

2.4 UML-based Web Engineering (UWE)

UWE is an UML based Web Engineering Approach [23]. It is a methodology for creating the design of Web Applications and consists of three phases, the conceptual design, navigation design, and presentation design.



Figure 2.4: UML-based Web Engineering [23]

2.4.1 Conceptual Design.

In the conceptual design stage a conceptual model of the applications must be build using the functional requirements captured with use cases. Such a conceptual design is build with traditional object-oriented techniques. Using classes, associations, inheritance, etc. The practical representation of the conceptual design is done in the Unified Modeling Language (UML).

2.4.2 Navigation Design.

The navigation design is based on the conceptual design. This design represents the navigational space and the elements that can be used to access navigation. There are several stereotype elements that can be used for building such a navigation design: indexes, guides tours, queries, and menus. The properties of these elements are explained in [24,25].

2.4.3 Presentation Design.

This third phase in the design process consists of two steps. The first step is sketching the content and the look and feel of the nodes. The second step is to combine these views into storyboarding scenarios that can be used to actually build a prototype. There is no precise notation for the sketching and storyboarding techniques, but the UWE approach [23] provides some methodological guidelines and modeling elements.

2.5 Conclusion.

The previously described Web Engineering methodologies are all recent research projects. However, none of them considers globalization issues or one of its aspects: localization, internationalization or translation. Since we live in a world where digital communications has no cultural or national boundaries there is need for decent Web Engineering methodologies to include the globalization aspect.

Chapter 3

Web Site Design Method (WSDM).

In this chapter an overview will be given about the methodology framework of this research, the Web Site Design Method, in short WSDM [31] or for easier pronunciation WiSDoM. In the first section of this chapter we will enumerate the characteristics of this approach and in the second section we will explain the different phases with the help of an example.

3.1 Characteristic of WSDM.

3.1.1 Different Approach.

The Web Site Design Method approach differs significantly from other Web Site design methods. While other approaches take the organization's data or database as a starting point, the so-called data-driven approach, WSDM takes the requirements of the users of the Web Site as a starting point and uses this as a basis for structuring the data and the Web Site. This was called the *user-centered* approach, but since this term is used in the field of HCI in a somewhat different meaning, we will use

the term *audience-driven*. Thus, the Web Site Design Method is audience-driven because it gives consideration to the fact that the target audiences of a Web Site may be composed of different "kinds" of visitors/users [32]. It is based on the fact that different types of users have different kind of needs. And fulfilling these needs leads to better tailored Web Sites, higher usability and greater satisfaction.

3.1.2 Distinction between Conceptual Design and Presentation Design.

In this methodology there is a clear distinction between the conceptual design and the design of the actual presentation (which consists of the implementation language, the grouping in pages, the use of menus, static and dynamic links, etc.). This implies several advantages. The fact that the Conceptual Design is free from any implementation detail and that this is an explicit phase gives us the possibility to allow different presentations/devices. This way the design is not biased by current implementation limitations, but can evolve with the rapid growing web technology. The Conceptual Design can be made available to the audiences of the Web Site or to search engines to increase the usability.

3.2 WSDM Overview.

With the use of an example we will now give an overview of the Web Site Design Method as it is specified in [32]. In Figure 3.1 you can see the different phases of the methodology.

The Mission Statement Specification expresses the purpose and subject of the Web Site, and declares the target audience(s).

The Audience Modeling phase consists of the Audience Classification where we identify and classify the different types of users and the Audience Class



Figure 3.1: Overview of the Web Site Design Method [32]

Characterization where we describe the characteristics of the members of the different Audience Classes.

The Conceptual Design phase consists of two sub-steps: Task Modeling, where we model the information and functional requirements of the different Audience Classes and Navigational Design where we model the (conceptual) structure of the Web Site.

In the Implementation Design step the Page Design, the Presentation Design, and the Logical Data Design is done. These consist respectively of grouping information in pages (following the Navigation Design), specifying the look and feel, and designing the database, an XML DTD, RDF definitions, ...

The last step, the actual implementation of the Web Site, in e.g. HTML,

XML, WML, ...

Consider as an example an online job vacancy site. Typical users of such a Web Site are employers, visitors, and the owners. These different users have different needs and requirements.

- Employers should be able to offer new job opportunities. To do so, they will need a login and password that is provided by the owners of the online job vacancy site. They can request the owners for authorization by providing their coordinates and credit card number to pay for their job advertisements. They can also provide a company logo and a company presentation which will be placed with each job vacancy. Once a company is registered they can start offering jobs. Therefore, they have to provide a job description, a description of the profile they expect of the candidates, a description of the job vacancy, the type of contract, a contact person, etc. Note that not all of these are mandatory. There should also be the possibility of deleting a job offering. Employers should also have the opportunity to search in a CV database for suitable candidates they want for a particular job.
- Visitors should be able to browse through the online job vacancies. They should have the opportunity to search for jobs by function, area (country, state, province, city, ...), statute (laborer, clerk, interim, ...) and language. They should also be able to search for jobs from a specific company. Next thing they may want to do is to apply for a job vacancy. When they do so, they have two possibilities. On one hand they can choose to stay anonymous for the online job vacancy company and do their application by standard email. On the other hand there should be a possibility to register and to create and modify an online CV which is automatically available in the database where employers can search for suitable candidates. The CV is also sent by email to the company of the job vacancy. Job-seekers may also want to create and modify a job profile (what they are searching for), so that when a vacancy is entered by a company which is suitable for that job-seeker, he immediately
receives an email.

• Owners of the company are responsible for the management of the job vacancy Web Site. When a new job has been added, the managers are notified and can inspect the information bout the job. A manager can change the category or subcategory of the job function if necessary. Next to that the manager can decide to notify all people who have the category or subcategory in their job profile.

This example was taken from [33] and will be further used in the following paragraphs to explain the different phases of the Design Method.

3.2.1 The Mission Statement.

The first step in WSDM is to define the *Mission Statement* of the Web Site (Figure 3.2).

Mission Statement Specification

Figure 3.2: Web Site Design Method Mission Statement Specification

The Mission Statement must answer the following questions:

- What is the **purpose** of the Web Site?
- What is the **subject**?
- Who are the **target audiences**?

Without this Mission Statement, which expresses the purpose, subject, and the target audiences of the Web Site, there is no proper basis for making decisions, or for evaluating the effectiveness of the web site and the users will have little idea what the Web Site is for [32].

We can express the Mission Statement for our example website, which we will call JobSite, as follows.

The JobSite mission is to assist employers in meeting their workforce and assist local job seekers to enter, remain, and progress in the workforce using job offers and job profiles.

So, does the mission statement provide an answer to the three question?

Purpose

- Assist employers in meeting their workforce needs.
- Assist job seekers to enter, remain, and progress in the workforce.

Subject

- Job offers.
- Job profiles.

Target Audiences

- Employers.
- Job seekers.

3.2.2 Audience Modeling.

The Mission Statement gives us already a general indication of the audiences involved in the site. For further specification of the target audiences and decision whether or not the audience should be modeled we need this Audience Modeling phase (Figure 3.3). It consists out of two sub-phases, Audience Classification and Audience Class Characterization.

Audience Modeling
Audience Classification
Audience Class Characterization

Figure 3.3: Web Site Design Method Audience Modeling

Identifying Audience Classes.

The Target Audiences abstracted from the Mission statement are:

- Employers.
- Job Seekers.

In the example requirements we speak of yet another Target Audience, the *Site Owners*, which are not included in the Mission Statement because of their supporting nature in the system. We will refer to them as the *Managers*.

Step 1 First we are going to consider the activities of the organization related to the purpose of the web site.

Job Mediation is in fact the only activity of the organization. This is done by offering Job Offers to Job Seekers and Job Profiles to Employers.

- Step 2 Then for each activity we have to:
 - 2.1 Identify the people who are involved in the activity.

In Figure 3.4 you can see a scheme of the people that are involved in the organization's activities. We added the Manager Audience Class.

2.2 Restrict them to the target audience. So only consider those people which belong to the target audience formulated in the mission statement.



Figure 3.4: Web Site Design Method Audience Classes - People who are involved.

In our example the people involved in the organization's activities are also part of the Target Audience specified in the Mission Statement.

- 2.3 Divide them into audience classes based on different information or functional requirements.
 - Employer.
 - Information and Functional Requirements
 - * Employer Registering.
 - * Log In.
 - * Create an Employer Profile.
 - * Add/Remove/Update Job Offer.
 - * Browse/Search Job Profiles.
 - * Retrieve a Job Offer application.
 - Job Seeker.
 - Information and Functional Requirements
 - * Job Seeker Registering.

- * Log In.
- * Add/Remove/Update Job Profile.
- * Browse/Search Job Offers.
- * Apply for a Job Offer.
- * Get Notification for a Job Offer.
- Manager.
 - Information and Functional Requirements
 - * Log In.
 - * Notify Job Seekers.
 - * Review Job Offers.
 - * Add/Remove/Update Job Offers.
- 2.4 Decompose the activity if possible.

If it is possible, activities should be further decomposed.

2.5 Repeat Step 2 until no new audience classes are found or no further decomposition is possible.

If decomposition of activities doesn't result into any new audience classes it should be stopped, because it is not useful anymore.

From these Audience Classes we can construct the Audience Class hierarchy (Figure 3.5). This Hierarchy shows all Audience Classes in terms of sub and super classes. The top of the Audience Class hierarchy is always the *Visitor* Class and groups all the requirements common to all Audience Classes. In our example the Visitor just needs to be able to retrieve general information about the Web Site.

Characteristics of Audience Classes.

We talked about the functional and information requirements, but there are still other characteristics that can distinguish the audiences. Members of one audience class may diverge in how the information should be presented to them. In this step



Figure 3.5: Web Site Design Method Audience Class Hierarchy.

of the Audience Modeling a set of characteristics per Audience Class is specified. Some examples of these user characteristics are: level of experience with Web Sites in general, frequency of use, language issues, education/intellectual abilities, age, income, lifestyle,... [32]

If we can distinguish, within one Audience Class, groups of members with different characteristics, we can introduce *audience class variants*. In our example the characteristics can be identified as follows:

• Visitor Characteristics

Visitors of this website are 18 years or older with varying WWW experience.

• Job Seeker Characteristics

Job Seekers are typically 18 or older. Their experience with the WWW may vary.

• Employer Characteristics

Employers are typically of average age (30 or older), they have an average level of experience with the WWW. They are very familiar with the jargon used on a Job Site.

• Manager Characteristics

A managers is someone with advanced knowledge of the WWW and experience in Job Vacancies.

No audience class variants can be identified from these characteristics.

3.2.3 Conceptual Design.

The Conceptual Design phase deals with the conceptual 'what and how'. In the Information Modeling, specially for data intensive Web Sites, we have to ask ourselves: **What** kind of **Information** is needed? With Functional Modeling, for application type of Web Sites, we emphasize on **what functionality** is needed. The Information Modeling and Functional Modeling are part of the Task Modeling (Figure 3.6). In the second step of the Conceptual Design we describe the (conceptual) structure of the web site and model **how** the members from different audience classes will be able **to navigate** through the site [32]. The goal of the Conceptual Design phase is to turn the requirements we identified in the Audience Modeling phase into high level, formal description which can be used later on to generate effective Web Sites.



Figure 3.6: Web Site Design Method Conceptual Design

Task Modeling.

Information and Functional Modeling is done by means of *Object Chunks*. Any modeling technique can be used to do this, but we are going to use ORM or Object Role Modeling with some extra notation to represent the Functional information. In Appendix The Task Modeling step can be summarized as follows:

For each information or functional requirement do

Step 1

- *Define* a task for this requirement.
- *Elaborate* the task into more detail (if necessary).
- Decompose the task into elementary tasks (using ConcurTaskTree).

Step 2

• For each elementary task make an **Object Chunk** that models the information and/or functionality required by this task.

An example will make things clear. Consider the following task:

Create an Employer profile.

This task can be decomposed, in Figure 3.7 you can see the resulting ConcurTask-Tree.



Figure 3.7: Employer Profile ConcurTaskTree

The Object Chunk is thus given in Figure 3.8.

Navigation Design.

This is the second step of the Conceptual Design. During this step we describe the structure of the website and the navigational possibilities for each Audience



Figure 3.8: Employer Profile ORM

Class. This is done by creating a *Navigation Track* for each individual Audience Class. Such a Navigation Track reflects the information, functional and navigational requirements we described in previous steps. All the Navigation Tracks together form the Navigation Model which is used to implement the Web Site.

Tracks are composed of components that represent units of information or functionality, and links that connect those components. In addition, links can be conditional depending on the logical true of a condition [34]. E.g. Figure 3.9 shows a Concep-



Figure 3.9: Conceptual Structural Model - For each Audience Class there is a Navigation Track and these are linked as in the Audience Class Hierachry (Figure 3.5) using structural links.

tual Structural Model of the example Web Site. In Figure 3.10 the Task "Create an Employer Profile" is elaborated in a Task Navigation Model.

3.2.4 Implementation Design

The *Implementation Design* is the next step of the method. As you can see in Figure 3.11 this phase consists of three sub-phases. In the **Page Design** we start from the



Figure 3.10: Employer Profile Task Navigation Model - This Task Navigation Model is derived from the ConcurTaskTree (Figure 3.7).

Navigation Model and we group the information we gathered in pages following the Navigation Design. The 'look and feel' of the Web Site and the layout of the individual pages is specified in the **Presentation Design**. The **Data Design** can be used to specify a logical data schema for a database using the Object Chunks specified in the Conceptual Design phase. If there is no need for a full-fledged database, logical schemas for XML DTD, RDF definitions, etc. can be produced here.



Figure 3.11: Web Site Design Method Implementation Design

3.2.5 Implementation

This last phase of the methodology, *Implementation*, consists of realizing the physical Web Site. First we have to choose an implementation environment (HTML, XML, ...) and then we can convert the result of the Implementation Design to the chosen environment. This can be automated depending on the complexity of the model and the availability of usable tools.

Implementation

Figure 3.12: Web Site Design Method Implementation

3.3 Conclusion.

This rather detailed description of WSDM gives a good overview of the power of the audience-driven approach of this methodology. However, the description given here is not complete and much more information can be found in [30–34].

Chapter 4

Integrating Globalization into WSDM approach.

In the previous chapter we explained the approach of the Web Site Design Method, WiSDoM. In this chapter we will discuss the integration process for integrating Web Site Globalization into the methodology. This will be done analog to the previous chapter by giving an overview of WSDM, but now with emphasis on adaptations made to integrate a Globalization process into it. In Figure 4.1 you can see an updated overview of the different phases of WSDM. In the following sections these updated phases are explained in more detail.

4.1 Mission Statement Specification.

The Mission Statement is defined as an answer to the questions: What is the purpose, what is the subject, and who are the target audiences(s)? This Mission Statement is the start of the methodology and should be a general statement of vision, intent in text form. It is the starting point for designing the web site. If we want to be able to take globalization into account in the rest of the method, the mission statement should mention something about it. Thus it must contain some notion of



Figure 4.1: Global Web Site Design Method Overview.

the culture, the location, the whereabouts, the situation, etc. the target audience(s) are in. It should give an answer to one of these question too:

- What is the location of the target audiences?
- What culture are the target audiences part of?
- In what kind of situation are the different target audiences?
- What are the whereabouts of the audiences?

To make the terms location, culture, situation, whereabouts, etc. clearer we define another, more general, term: locality. The definition as found in a dictionary is the following: Localities describe a particular place, situation, or location [1].

Thus we can add another requirement for the Mission Statement. We will require that it answers the following questions:

- What is the **purpose** of the Web Site?
- What is the **subject** of the Web Site?
- Who are the **target audience**(s)?
- What are the **localities** of the target audience(s)?

Localities are specified as a (part of a) geographic location, a culture, a nationality, etc. Keep in mind that one target audience can have different localities and vice versa as we will see later on in the example.

In this chapter we will use the example from previous chapter, but we will add a globalization factor. Suppose the JobSite is meant for the US market as well as for the Japanese market. Companies and Job Seekers from US and from Japan must be able to use the JobSite.

Considering the Mission Statement from the previous chapter:

The JobSite mission is to assist employers in meeting their workforce and assist job seekers to enter, remain, and progress in the workforce using job offers and job profiles.

We can adapt it as follows:

The JobSite mission is to assist employers in the US and in Japan in meeting their workforce and assist **local** job seekers to enter, remain, and progress in the workforce using job offers and job profiles.

This way the localities can be identified from the statement: the US and Japan.

4.2 Audience Modeling Extended

We already know what the definition of a locality is, we also know that a locality is actually a part of the audience. A target audience takes part in one or more localities. This is an important observation in modeling the different localities and mapping them onto the target audiences. Therefore we've extended the original Audience Modeling with two extra steps (Figure 4.2). The first one is **Locality Specification**. Different Localities may have different requirements and characteristics and in this step we will describe these. The second one is **Locality Mapping** where according to the Mission Statement the Localities are mapped onto the Audience Classes.



Figure 4.2: Global Web Site Design Method Audience Modeling Extended.

4.2.1 Locality Specification

Each Locality specified in the previous step (the Mission Statement Specification) has its own requirements and characteristics. These requirements and characteristics depend on the language used in the Locality and/or are culturally or ethnic related. These requirements and specifications can be described, just like in the Audience Classes, on different levels: the Information Level, the Functional Level, and the Usability Level.

Information Level describes what information should be available specifically for a Locality. Extra information that isn't needed in every other Locality. E.g. extra menu-item with Social Responsibilities on a company Web Site in countries with high Collectivism (see Appendix A).

- **Functional Level** describes what extra functional requirements are needed in a Locality. E.g. extra/less fields when an address has to be entered.
- **Usability Level** describe the specifications on how the information should be represented. According to the diversity of the Localities this level will be the most elaborated. E.g. the language used in the Locality, the reading order, the colors, symbolism, etc. For a comprehensive list of example we refer to Appendix A.

Some of the Usability Level specifications may be translated into usability requirements, while others may be used later on in the implementation phase to guide the design of the "look and feel" of the Web Site.

4.2.2 Locality Mapping

The Localities need to be linked to the Target Audiences. This is done using the Audience Class Hierarchy and *Locality Sets*.

A *Locality Set* is a (mathematical) set which holds one or more Audience Classes, these are the elements of the set. In Figure 4.3 you can see an example. This Audience Class Hierarchy represents 6 Audience Classes: Visitor, Audience-A, Audience-B, Audience-C, Audience-D, and Audience-E. There are also 4 Localities and the same amount of Locality Sets: Locality-A, Locality-B, Locality-C, and Locality-D. In this example Visitor is the only element of Locality-A and must be only available in Locality-A. Audience-E is element of Locality-A, Locality-C and Locality-D and must be available in those three Localities.

In our example of the JobSite we can state the following: Occasional Visitors, the Job Seekers, and Employers should be able to see the Web Site in their own Locality, thus these Target Audiences should be able to see the Web Site in the US Locality



Figure 4.3: Global Web Site Design Method - Locality Sets

or in the Japan Locality. The site Manager lives in the United States and he must only be able to see the site in the US Locality. The result can be viewed in Figure 4.4.

4.3 Conceptual Design Revisited.

In the previous chapter we explained the different steps of the original Conceptual Design phase. In this section you find an overview of the adaptations made to the original phase. First we consider the Information and Functional Modeling in the Task Modeling step and second the adaptations to the Conceptual Structural Model are explained in the Navigation Design.



Figure 4.4: Global Web Site Design Method - Extended Audience Class Hierarchy

4.3.1 Task Modeling.

In the Locality Specification phase we described in the Information Level and in the Functional Level the extra requirements about each Locality. In the task modeling phase it are those requirements, Information and Functional, which we are going to model, thus we need to take the specifications we described in the Locality Specification phase into account.

In the Functional Modeling step the specifications mentioned in the Functional Level has to be defined as a Task. This Task has than to be elaborated, if necessary, into more detail. Next, we decompose the Task into elementary tasks using ConcurTask-Tree. These steps are not different from those described in the previous chapter with the exception that the source of the task is not the Functional Requirements described in the Audience Modeling, but in the Functional Level in the Locality Specification.

The Information that is needed on the Web Site is described using Object Role Model (ORM). Since we are adapting the Web Site to support multiple Localities, information is perhaps going to be different. The information that is going to be dependent on the different Localities has to be noticeable in the Object Chunks created in this phase. These Object Chunks act as a basis for the Conceptual Schema from which eventually a database can be created.

Consider the following example: We take two Localities, the US and Belgium. In Belgium a postal address consists of a street, a street number, a city and a zip code. In the US you also need to specify the state in which you live. In Figure 4.5 you can see how we represent the fact that only in the US Locality the state object is needed: A label indicating the Locality is put on the role connection.



Figure 4.5: Global Web Site Design Method - Customer Address Object Chunk Example

4.3.2 Navigation Design.

The Navigational Design describes the structure of the Web Site and a model of the way users will be able to navigate through the site. The result of the Navigation Design is a Conceptual Structural Model containing the different Navigation Tracks based on the different Audience Classes and the Audience Class Hierarchy. Since we mapped our Localities on those Audience Classes and all the activities these Audiences are involved in, there has to be notion of those Localities in the Conceptual Structural Model of the Web Site.

The notion of Localities can be added to the Navigation Design by adding Locality Labels (Figure 4.6) to the Audience Tracks. Each Locality has its own Locality Label. The Locality Label has to be applied only to the tracks, the components, of which tracks are composed, inherit the property of the Locality.



Figure 4.6: Global Web Site Design Method - Locality Attribute

In our example this will result into the following Conceptual Structural Model for the JobSite (Figure 4.7)



Figure 4.7: Global Web Site Design Method - Conceptual Structural Model

4.4 Implementation Design.

The different phases of the Implementation design are analogue to the original ones we described in the previous chapter. The only difference is that in the Page, Presentation, and Data Design the different Localities have to be taken into account. In the Appendix B you can find some common Implementation Techniques that can help with the Implementation Design.

Chapter 5

Conclusions

The subject of this thesis is Web Globalization. We presented answers on questions like: What is Globalization in the context of the World Wide Web? Why is it necessary? What are the recent existing Web Design Methodologies? Do they deal with Web Globalization?

It turned out that none of the existing web site development methods available in the literature consider the issue of Web Globalization. Nevertheless, it is an important issue that goes beyond content translation. Therefore, we have extended an existing web site design method, WSDM developed by the Web & Information Systems Engineering research group at the Vrije Universiteit Brussel, in such a way that, when needed, a web site designer can take into account, right from the beginning of the design that globalization is required.

First, a detailed overview of WSDM is given. Next, we presented our approach by integrating the Web Site Globalization issues in the WSDM approach. To do so, first we introduced the concept of Locality. A Locality is a particular place, situation or location. This concept allows us to identify, next to the target audience(s), the different Localities for which the web site need to be developed. This is done during the Mission Statement Specification phase. In the second phase of the design (Audience Modeling), the requirements and characteristics of the Localities can be specified. Then the Localities are mapped on the target Audience(s) (the Audience Classes). In the third phase, Conceptual Design, the requirements of the Localities are integrated into the Conceptual Model by labeling roles in the Object Chunks and navigational tracks with the Localities for which they are relevant.

The impact of the extensions to support Globalization, on the method is minimal. The specification of the Globalization can be compared to an aspect (as in Aspect Oriented Programming). It is possible to add or remove it from a design without touching the main part of the design, e.g. the requirements of the different audience classes are first modeled whatever their Locality may be. The extensions proposed are also inline with the philosophy used by WSDM, which is the Audience-driven approach. This approach gives due consideration to the needs of the different customers (called audiences).

Bibliography

- [1] Merriam-Webster Online Dictionary, http://www.m-w.com/.
- [2] Localization Industry Standards Association, http://www.lisa.org/.
- [3] Globalization, Localization, Internationalization and Translation, http://philip.pristine.net/glit/.
- [4] WorldLingo: Web Site Localization, http://www.worldlingo.com/.
- [5] Hobbes Internet Timeline the definitive ARPAnet & Internet history, http://www.zakon.org/robert/internet/timeline/.
- [6] Six Reasons for the Bandwidth Boom, George Runkle, http://www.fool.com/dripport/2000/dripport001025.htm, 2000.
- SocioSite: PECULIARITIES OF CYBERSPACE: INTERNET USE(ERS), http://www2.fmg.uva.nl/sociosite/websoc/demography.html.
- [8] How To Internationalize Your Web Site, http://www.sdlintl.com/.
- [9] Global Internet Statistics, http://www.glreach.com/.
- [10] Mcracken, D. ,Wolfe, R.: "User-Centered Website Development A Human-Computer Interaction Approach", Pearson Prentice Hall (2004)
- [11] Marcus, A. ,Gould, E.: "Cultural Dimensions and Global Web User-Interface Design: What? So What? Now What?", In Proceedings of the 6th Conference on Human Factors and the Web, Austin, Texas (2000)

- [12] 8e6 Technologies :: Internet Filtering for the large Enterprise, http://www.ldc.com/.
- [13] Forrester Research: Technology research and advice, http://www.forrester.com/.
- [14] Messaging Online, http://www.messagingonline.com/.
- [15] United States Department of Health and Human Services, http://www.os.dhhs.gov/.
- [16] Object-Oriented Hypermedia Design Method, http://www.oohdm.inf.pucrio.br:8668/.
- [17] Schwabe, D. ,Rossi, G.: "Developing Hypermedia Applications using OOHDM", In Workshop on Hypermedia Development, Pittsburgh, USA (1998)
- [18] Web Modeling Language, http://www.webml.org/.
- [19] Ceri, S. ,Fraternali, P. ,Bongio, A.: "Web Modeling Language (WebML): a modeling language for designing Web sites", Computer Networks, Amsterdam, Netherlands (1999)
- [20] Object-Oriented Hypermedia, http://gplsi.dlsi.ua.es/iwad/ooh_project/.
- [21] Gomez, J., Cachero, C.:: "OO-H Method: Extending UML to Model Web Interfaces", The Web Engineering Research Group, University of Alicante, Spain (2003)
- [22] Glossary, http://myphliputil.pearsoncmg.com/.
- [23] Hennicker, R. ,Koch, N.: "Modeling the User Interface of Web Applications with UML", Institute of Computer Science Ludwig-Maximilians-University of Munich, Germany (2001)
- [24] Baumeister, H., Koch, N., Mandel, L.: "Towards a UML extension for hypermedia design.", In Proceedings of UML'99, France, (1999)

[25] Hennicker, R. ,Koch, N.: "A UML-based Methodology for Hypermedia Design.", In Proceedings of the Third International Conference UML, York, UK (2000).

Springer-Verlag, 2000.

- [26] The Principles of Marketing, Philip Kotler, Pearson Education Limited, 2001.
- [27] History & info Overview of many calendars, http://webexhibits.org/calendars/.
- [28] Optimal Web Design, http://psychology.wichita.edu/optimalweb/international.htm.
- [29] Stengers, H.: "The World Wide Web as a Globalizing Catalyst, Consequences for Cross-Cultural Web Interface Design", Erasmus College, Brussel (2003)
- [30] WiSDoM homepage, http://wsdm.vub.ac.be/.
- [31] Lekova, A. ,De Troyer, O.: "Improved Navigation Through Extended XML Links", In Proceedings of the Third International Conference on Information Integration and Web-based Applications & Services (IIWAS 2001), Linz, Austria (2001)
- [32] De Troyer, O.: "Audience-driven web design", In Information modelling in the new millennium, Eds. Matt Rossi & Keng Siau, IDEA GroupPublishing, ISBN 1-878289-77-2 (2001)
- [33] De Troyer, O. ,Bille, W.: "Design Methods for Internet-based Systems, Assignment 3: design and implement a website", WISE Laboratory, VUB, Brussels (2003)
- [34] Najjar, J.: "Business Rules & WSDM Methodology of Web Design", Wise Laboratory, VUB, Brussels (2002)
- [35] World Wide Web Consortium, http://www.w3c.org/.
- [36] FullXML, http://www.fullxml.com/.

Appendix A

Cultural Differences that harden the Localization Process.

Localization is much more than mere translation. The following paragraphs will make this clear with guidelines and examples.

A.1 Text Considerations.

Adapting the text to fit a certain language and culture is a major consideration, perhaps the most important step in localization.

A.2 Text translation.

Translation of text on websites can happen in two different ways. It can by done by humans exclusively, or you can do a semiautomatic translation: the rough part can be done using machine translation, and humans do the finer work. Machine translation alone does not suffice for a high quality result. Take into consideration that translation goes beyond the vocabulary. In the past embarrassing errors have occurred.

• A slogan of the Coors Brewery, Turn it Loose became Suffer from diarrhea in Spanish [10].

Other examples to show that we must consider the communication of concept in context of a culture:

- In Arabic countries a day begins at sundown, not at midnight [10].
- In Japanese, culture the number 13 is not considered unlucky. There it is the number 4 [10].

Another important issue in text translation is the way of communicating information. In the United States they like straightforward and specific prose, in contrast to Japan, where a phrase like Buy Now! would be considered extremely distasteful [10].

A.2.1 Left to Right, Right to Left.

Also reading orders vary among different locales. In most of Europe and America the reading order is from left to right, but in Middle Eastern languages for example, it is from right to left. This has also a big influence on the layout design.

In Figure A.1 and in Figure A.2 you can see an example different reading orders.

A.2.2 Date, Time, Size/Weight measurement, Currencies, etc.

Today more than six different calendars are used worldwide. Using the right date notation can prevent many misunderstandings.



Figure A.1: The English Al Jazeera Web Site - The text order is from left to right.
- http://english.aljazeera.net/

E.g. We are now in the year 2004 and I was born on February 20th in the year 1982. In the Chinese Calendar it is the year Heisei 16 and I was born in Showa 57. This is the year of the Monkey. I was born in the year of the Dog, 8100 days ago.

Other types of calendars are the Indian Calendar, the Islamic, the Jewish, the Ethiopian, the Persian, etc. They all differ in writing style and are based on different counting systems [27]. Even in the same calendar you can have different writing styles. In the United States, the interpretation of 04/05/2004 is April 5, 2004, while most of the rest of the world will interpret it as May 4, 2004. That is why it is important that you use a writing style that can be interpreted unambiguously.

Time is another text feature that we must take in consideration in the localization step. The most significant difference there is the use of a 12-hour clock (with the A.M. and P.M. abbreviations) and the use of a 24-hour clock. The globe is divided into different time zones; so if you mention time, dont forget to mention the time zone. E.g. UTC, which means Coordinated Universal Time.

Other examples are in the notation of size and weight measurement. E.g. Kg vs.



Figure A.2: The original Al Jazeera Web Site - The text order is from right to left - http://www.aljazeera.net/

Pounds, Km vs. Miles, etc.

Monetary differences need also special attention. Not only the currency symbol, but also conversion needs to be taken into account. E.g. 450.00 EUR equals 544.77 USD, monetary symbols, etc.

Many other culturally or locale specific text features have to be taken into consideration, many of them specific to a certain area. We mentioned the most frequent occurring ones.

A.3 Color, Icons, and Images.

In China, a bride who wore white would shock the assembly and cause pain for her parents, because white is the color worn to funerals and has a strong association with death. In China and parts of India, red is the preferred color for a bridal dress; in the United States, this would have a strong negative connotation. [10] A company tried to lay connections with Taiwan and got them in an embarrassing way. They were looking for more foreign sales and arrived in Taiwan with green baseball hats they were going to distribute freely as gadgets. The moment they chose for their commercial adventure was a month before the elections in Taiwan and green was the color of the opposition. Even worse, after the trip the visitors discovered that an ancient Taiwanese habit says that a man who wares a green hat has cheated on his wife. The leader of the delegation told us later: I had no idea what went wrong with the green hats, but that trip gave us certainly more insight in the gigantic differences between certain cultures. [26]

As you can see identical colors can have different meanings between different localities. The feelings and thoughts associated with the color differ between cultures, nations, or population groups. It is important to use the correct colors if you dont want to create a wrong impression. In Table A.1 you can find a small overview of color symbolism in some different cultures.

Another non-textual aspect on a website that is tricky to localize is icons. You all know the icon with the house. If you click on it, you will be deferred to the main page of that website. The symbol of the house refers to the Home of the website, the Home Page. But in France they refer to the main page of a website as Welcome Page (*page dacceuil*). In Germany they call it Starting Page (*StartSeite*), in Spanish it is the Beginning (*inicio*). Thus, in these languages, an icon of a house does not convey the message [10].



Figure A.3: The "Home Page" symbol does not convey its message in all cultures.

The use of symbols, pictures, and colors in a website can cause great difficulty in its localization and has to be done very careful.

	\mathbf{Egypt}	China	Japan	India	France
Red	Death	Happiness	Anger, Dan-	Life, Cre-	Aristocracy
			ger	ativity	
Blue	Virtue,	Heavens,	Villainy		Freedom,
	Faith, Truth	Clouds			Peace
Green	Fertility,	Ming Dy-	Future,	Prosperity,	Criminality
	Strength	nasty,	Youth,	Fertility	
		Heavens,	Energy		
		Clouds			
Yellow	Happiness,	Birth,	Grace,	Success	Temporary
	Prosperity	Wealth,	Nobility		
		Power			
White	Joy	Death, Pu-	Death	Death, Pu-	Neutrality
		rity		rity	

Table A.1: Color Symbolism in Different Cultures [10]

A.4 Hofstedes Dimensions of Culture and Web Globalization.

In connection with the fact that cultures, even within some countries, are very different, Hofstede has done some major research. Geert Hofstede, a Dutch cultural anthropologist, conducted during 1978 and 1983 detailed interviews among hundreds of IBM employees in 53 different countries. From the data he gathered from the analysis of the interviews he formulated a theory that world cultures vary along consistent, fundamental dimensions [11]. Hofstede identified four dimensions. Each of the 53 countries was assigned with a rating from 0 to 100 for each of these four dimensions. In the following paragraphs you will find a short description of the cultural dimensions Hofstede identified.

A.4.1 Power Distance.

According to [11] the Power Distance dimension refers to the extent to which less powerful members expect and accept unequal power distribution within a culture.

Some characteristics of cultures with high Power Distance [11, 28]:

- They tend to have centralized political power and deep hierarchies.
- Hofstede claims that these types of societies emphasize authority, experts, certifications and official logos, security, and an acceptance for restrictions to information access.

Cultures with low Power Distance tend to view subordinates and supervisors as closer together and more interchangeable. There is less difference between higher and lower hierarchies. Equality is expected and generally desired. In the ranking, Hofstede made [28], Malaysia is the number one with the highest Power Distance. Austria has the lowest PD.

In Figure A.4 you can see a nice example.

A.4.2 Collectivism versus Individualism.

The second dimension, collectivism versus individualism, refers to, according to [28], the degree to which a culture emphasis the self and immediate family over the society at large. Sites from countries such as the United states promote a very modernlooking, youthful, and individually successful looking design, while on the other hand collective cultures feel more comfortable with sites that promote the history of the company and how it helps the Guatemalan society in its whole [28]. In the ranking, Guatemala scores the best on collectivism.

In Figure A.5 you can see a nice example.



Figure A.4: On a Malaysian tourism site (http://www.tourism.gov.my/) the first item you find in the navigation menu is corporate information with pictures and names of the Board of Directors and the Organization Structure.

A.4.3 Masculinity versus Femininity.

The MAS Dimension, or Masculinity versus Femininity Dimension, refers to gender roles [11]. It refers to the degree to which traditional masculine roles of assertiveness and competition are emphasized [28]. The highest and lowest ranked countries in the MAS index are respectively Japan and Sweden. So according to the figures in Japan the traditional distinctions between males and females are strongly maintained, while in Sweden they tend to collapse the distinctions and overlap gender roles [11].

In Figure A.6 you can see a nice example.



Figure A.5: The Guatemalan site of McDonalds (http://www.mcd.com.gt/) has an item in its navigational menu that mentions the role McDonalds is playing in the Community. It contains a list of the Social Responsibilities McDonalds is fulfilling in Guatemala. What does McDonalds do for the collective?

A.4.4 Uncertainty Avoidance.

The last cultural dimension Hofstede identified is Uncertainty Avoidance. In [28] they refer to this dimension as the degree to which individuals have anxiety about uncertain events. Low UA (Uncertainty Avoidance) cultures are more relaxed and prefer more informal business arrangements. In Web Sites you can see a low UA by great complexity of the website and less control over navigation. In the UA Index Greece has the highest UA level and Singapore the lowest.

The last word isnt said about these cultural dimensions. It was already a long time ago that the interviews were conducted and since than much has changed, or hasn't it? It is really hard to find decent examples that fit each cultural dimension. More information about that you can find in [29].



Figure A.6: The fact that the Portal Site .excite has a separate Web Site for women in Japan (http://women.excite.co.jp/) which it does not have in other countries with lower MAS index.

Appendix B

Implementation Techniques.

In the following paragraphs we describe some techniques to make the implementation of Web Sites for different Localities a lot easier.

B.1 Abstract Content from Layout Using CSS.

Cascading Style Sheets (CSS) is a simple mechanism for adding style (e.g. fonts, colors, spacing) to Web documents [35]. Using these Style Sheets gives you the possibility to completely divide the Layout of your Web Site from the Content. The possibilities of CSS are almost endless. More information can be found on the Web Site of the World Wide Web Consortium [35].

This simple example shows us how we can change the reading order of the text on a website using CSS.

p {
 text-align: left }

If this style is attached to a web-page all the text within a paragraph html tag will be aligned to the left.
p {
 text-align: right }

In this example it will be aligned to the right.

B.2 Language Files for Common Strings.

On Web Sites most of the textual content is the same in the different Localities, it is just the language in which it is represented that is different. This can be used to abstract these different strings into variables in a common file and replace them on the web-page with a call to the string in that file.

This can be done by naming all the common strings used in the Web Site and gathering them, together with their translation in separate files. E.g. this is an excerpt from a XML file which stores all the common string from a Web Site with its English translation. The strings can be called by using the position number they appear in.

```
<?xml version="1.0" encoding="UTF-8"?>
<language>
<name>English</name>
<translator>
<name>John Doe</name>
<email>john-doe@email.com</email>
</translator>
<encoding>utf-8</encoding>
<strings> <!-- main menu, start @ position 1 -->
<string>Website Administrator</string>
<string>Browse site</string>
<string>Settings</string>
<string>Pages</string>
```

```
<string>Contents</string>
<string>Surveys</string>
<string>Quotations</string>
...
```

The following excerpt is the same as the previous but from the Croatian language file.

```
<?xml version="1.0" encoding="windows-1250"?>
 <language>
   <name>Croatian</name>
   <translator>
     <name>Croatian John Doe</name>
     <email>croatian-john-doe@email.com</email>
   </translator>
   <encoding>windows-1250</encoding>
   <strings> <!-- main menu, start @ position 1 -->
     <string>Administracija Web Stranica</string>
     <string>Pregled weba</string>
     <string>Postavke</string>
     <string>Stranice</string>
     <string>Sadraj</string>
     <string>Ankete</string>
     <string>Tipovi</string>
     . . .
```

As you may have noticed in these language files the encoding (utf-8 and windows-1250) is stored also. These examples were taken from [36].

B.3 Localized Content Using DB Locality Flags.

Now that we have created Localized layouts and that we have localized the common strings in the different localities, we also want to create localized content. if the content of our Web Site is extracted from a Database (Microsoft Access, SQL, etc.) this can be done by using Locality Flags. Depending on the Locality we are in the flag changes and different content is extracted from the Database.

E.g. in our example we have two Localities, US and Japan. The parameter appended to the page's URL tells the code in the web-page to extract the US related content from the Database.

index.asp?loc=US

This is analogue for the Japanese Locality but with a different flag.

index.asp?loc=JAP