Internationalization of Softwares

Assignment
on

Web Internationalization

Submitted for partial fulfilment of the Degree of
Masters
of
Information Technology
(Software Architecture)
Batch:2014-2016

double dash: That will result in a single long dash

sapce between : and 2

Submitted By:
Harjot Kaur Mann
n9275398

can't n be capital

Queensland University of Technology
Brisbane, Australia

Contents

Do we need "in" here?

Why colon?

1   Web Internationalization

1.1   First Taste: An Introduction

A terminology which is bringing awareness to cosmopolitan in the field of World Wide Web. Internationalization takes everything under the control of world. Scientifically, it enables the application to run with different vocabularies and according to their rules and customs. To arrange your web content obtainable in all other different languages and civilization for its easy accomodation for the enjoyer is known as Web Internationalization. This method of internationalization is also known as translation and localization enablement. Creating multilingual websites becomes possible with this technique, which is a powerful tool to serve the customers and buyers with better serviceability. The languages impels the text order, fonts and the accent. The crucial parts of building a website is the content and navigation but the main thing under consideration is the language which is written on the web pages and their built-in properties. There are certain specified and authorized standards which are used to represent a web site, these are Unicode. The Unicode history and the implementations for coding are provoded to the developers which are very important to take the results for building bilingual or tri-lingual websites.

It is notable here that a website possesses three folds: Content, Presentation and Behaviour.

1. **Content Layer:** The web pages contains the main content which users read when they visit the website. The text and the images are the main portion which are inhered in the content and the arrows needed by the readers for exploration of the website are also included in it. To build up the content layer, HTML is used in web development and it is also used to design the web document.

2. **Presentation Layer:** This layer presents your document to the primers. CSS handles this sheet and the approach of the document, that how it will be shown and with which media form.

3. **Behaviour Layer:** In this layer, the functioning of the webpage is achieved. For illustration, if AJAX is used, it means Javascript is enabling the functioning of the page. For backend results, Java or PHP is needed to produce the results on the demand of the user's click.

Relation to Localization

Internationalization is very much closely related to the process of Localization which is to adapt a product or website to a particular language, country or customs. This is ometimes called a locale. Internationalization and Localization are also defined by i18n and l10n respectively. This is because the Internationalization contains 18 characters between i and n and similarly in Localization, it is having 10 letters between l and n. To deine these terms more clearly, an example of is relation can be the resourceBundle class of Java. This permits for getting the locale of some information like calenders for currency. In java if resourceBundle is present there, then it means the website is internationalized.

1.2   Upbringing and Implication

According to the demonstration of Sasaki and Phillips in 2005 ans 2006 respectively, internationalization is not considered as a specific feature, it is a requirement to design the software or website in general. Chinese believes that you can use any language to buy but you must sell in customer's language. This leads the actual demanding to plan for the multilingual and intercultural websites. In 2000, the Queens Borough Library in New York stated that the half of the habitants uses different languages at their homes to talk to each other rather than English. Due to the acquisition of this information, forces the developers to think about the internationalization of websites so that they can extend its scope to those folks who

speaks Non-English languages. But this type of translation needs much responsibility, so they need to be handle it with much care to establish their reputation in the eyes of customers that they are providing an excellent assistance and accurate assertion to their users disregarding their roots and culture. One such historical example of Internationalization is of Bosch. It inaugrated its first office of international sales in London in 1898. Its sales generation was 88 percent by 1913 over Germany.

Scott describes that In the next 10 years from 2007, there will be 70 percent of the Shopping malls will be bulit by selling to the Indians, Chinese and Russians.

How Unicode comes into play?

While designing websites, you should ensure that your website will be fitting with other cultures and languages, which leads to the existence of Unicode. This is basically a standard for ciphering industry. Implementing the persistent portrayal of content, disrespect to the script, is the main objective of this title. For examplify, there are some languages like English, arabic or Chinese, whether we write or read from left-to-right or vice versa, are indulged. Unicode uses almost ninety different scripts and billions of characters. It can be processed with differenct encodings and these are UTF-8, UTF-16 and UCS-2. The most basic character encoding style for Unicode is the UTF-8. For emails and websites, this is the default system for encoding and this will assure your website's adaptability with any language. An absolutely new term, to set the binary codes or characters. The written content of several distinct languages can be shown, handled and altered with a system called Unicode Worldwide Character Standard, which is looked after by Unicode Consortium.

Joe Becker from Xerox and Lee Collins and Mark Devis from Apple, these were the people who based unicode in 1987, by inspecting the production of Universal Character set. In the character enoding, the bytes will be generalized to characters. As the Unicode understands the scripts and languages very well and hence it manages the mulitlingual content. When Unicode is used in forms, databases and scripts, it is very effective but if you will not specify the character encodings properly, it will be difficult to read the text completely by the user.

There are many most widely used scripts which are encoded by unicode and it unifies historic and regional variations in the markup languages like XML. These variations are called glyphs. These glphys are unified into the singular characters by Unicode. Han unification is the most important example of glyphs uniifcation, which connects the different glyphs to the variants of Chinese, Korean and Japanese to one character.

1.3   Status in Quo

Upto 2011, the third part of the world was connnected to the Internet. In Asia, the 42 percent cyberspace is residing, which is the largest population using World Wide Web while North America is 13.5 percent of the Earth's community, 24.4 percent from Europe and 20 percent from other countires including Australia. The number of people using Arabic is marked upto 1500 percent in the spent 10 years. To properly localizing the websites upto the international level, big companies have big budgets also. It means they have their own country domains for example, in India .in is used in domain to represent it as an Indian website. However, its diffcult to set their revenues upto the international endeavours for every company because it costs million dollars. So its difficult to built the websites which are internationally friendly.

So to build a world ready design, all the websites should feel locally relevant. Yunker, the co-founder of Annual Web Globalization Report Card noticed that Google is the best in building scalable products. It is obvious that Google is very strong proffessionally to lacalize its full text in all the forty languages. Without getting the full knowledge of a project, to commence a project will be a complete failure. Therefore,

it is important for the team to know the leeway and blow of the attempts. Your team shoudl divide all the areas of the application and proper analysis should be done.

1.4   Technical Come-outs

The internationalization issues of markup language and related technologies are the in the centre of interest. It is very compulsory to look at the technical issues while making a website. It is quite good to spend a quality of time specially on testing it so that you can ensure that its working good with all the devices. There are some other technical issues which you need to take in consideration.

When an application is built, every little part of it is affected rather than limiting to the presentation layer. It is concluded from the experience, that globalization of the web service has mainly two ways, and these are: Internationalization and Localization. Both these parts are foursquare to the web application layers that shown as in the diagram.

The following detail will elaborate these affairs.

Looks space before The

Figure 1: Internationalization and Localization orthogonal to web application layer

1.4.1   Issues related to Database Layer

A database is an important part of every website, so it is important to take care of the following points:

1. **Character Set:** The character set features include:

   - The data or content which is coming in should be sustained by the database. All the language conditions should be concerned by the character set. There are some standards which supports the English as well as some other European languages like ISO 8859, but the problem is it is not promoting Chinese, so Unicode can be used.

   - Refactorization should be there, so that there will be no need to cahnge the code and t recompile it when localized.

   - The data should be handled by the foreign incorporates.

2. **Migration of Data:** It is not diffcult to migrate the data to database because, its already encoded in the ASCII format but still to export and import data to shift the character set, there are number of tools accessible from database merchants.