

## Evaluation of multilingual websites using localization matrix

Sarika Katkade

M.E Scholar,  
Department of Information  
Technology,  
Pimpri Chinchwad College of  
Engineering, Pune, India  
Email: katkadesarika@gmail.com

Jayashree Katti

Assistant Professor,  
Department of Information  
Technology,  
Pimpri Chinchwad College of  
Engineering, Pune, India  
Email: jayashree.katti@gmail.com

Chandrakant Dhutadmal

Senior Technical officer,  
Centre for Development of  
Advanced Computing  
(C-DAC) Pune, India  
Email: cpdthutadmal@gmail.com

**Abstract**—Creating, implementing and maintaining a huge website can be a terrifying task when the website is within multiple natural languages. The probable communication issues related to website content and structure become even more complex for multi-linguistic sites. The fact is that a majority of Web consumers today don't speak English, and they opt to access websites about companies' services in their own language. Organizations always want to keep a global image for that it required to use the website localization for local users. People can communicate via multilingual websites which need to browse independent screens for each one. It brings one of the most crucial areas exposed up by the period of electronic procedure. Web page localization is a solution to many organizations are leverage to appropriately translate their websites so their international customers can read site content and access information in their native languages. This address to improve culturally adaption of any multi-lingual website that focuses individual user-centric.

**Index Terms**—Localization; Information Retrieval; Translation; Multilingual website

### I. INTRODUCTION

Web page localization is not only doing a simple translation of text [1]. Translating content only resolves partial language problems. Texture and images of website are customized to claim to the target culture. Web page localization means the translating an existing website to local language and culture in the target market [3]. It is the process of approving an online site into a kind of linguistic and cultural environment [4] involving considerably more than performing a translation of text. Website localization should construct a site so that it seems local, to its audience except cultural variations between website inventors and the viewers [4]. While performing localization, main translation process

done by programmer and knowledge of cultural. This localize involves alteration target language and ethnical selection in the textual structure, images ,map and requirement of the site with having consistency of the web site. A target wise accepted web site doesn't need extra worry from visitors of the site to process information. It makes individual theme easier and perceptions toward the site more favourable. The website will be qualifying when the goal of the new website with providing to the locale market in the new locale [4].

The web site localization is the consequence of due to popular use of computer and Internet surfers. Now-a-days, internet as main resource for providing information and facilities. These consumers around the world can communicate in only language. If any business needs to expand globally then web page localization provides platform to it [5]. Localization determines the converting existing product, application or file to specific target market with consideration of requirements, language, cultural preferences. Web page localization is on the top of market to make website progressively important to supplier of a product in that area and replace in foreign markets. Since web design becomes more deadening task if it is developing in many multiple languages with national culture, it will attract online consumer purchasing to access as per locality [7]. Web page creators consider the theme, education level, belief and language, value systems, traditions and habits of the locale culture in order to gain more profit [8]. The website localization is a complicated process and involves three different levels of adaptation.

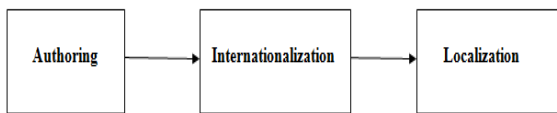


Fig 1: Levels of Adaption

1) Translation: Web page localization processing converting any sort of text used in the language of the state. While doing this process, it is necessary information of the web should be clear and meaningful to avoid ethnical ambiguity. To convert a text, the target preferences must be known [9].

2) Localization: In this, translation process comes first. Even translation of text should be meaningful but also have other activities to show it to local people. It contains to ensure that accepting of text structure, visual display, navigations and images and culturally appropriate for the target area. Many specialized elements which is need to be localized like: current date, current time types, numerical forms, postal address and telephone number formats, foreign currency/money formats, units of measurement and timing speed of website [10]. In localization process, communication does effectively in collaboration with different culture to consider the following items: information construction, theme, routing graphics, photographs, audio, and visual appearance [11].

3) Internationalization: It involves making application program standard that will be fully standardize as per technology, research of the country.

#### A. Localization

A process of properly converting cultural subjects in the internationalization process which proposes for specific local language. It is also included many features considered specifically to the target culture. In this case, localization means adapting existing features to accommodate a particular local things, which turns to be recognized industry standards are currency, language, educational level or income standards depending on kind of the communication [5]. The technological elements localized are the pursuing:

- Current Day and time configuration, calendar settings.
- Currency Standards.
- Monetary-associated information.
- Number operator.
- Telephone number scheme.
- Address scheme including postal codes, lane, states.
- Name formats.
- Different Paper size dimensions.
- Units of measurement.
- Color conventions: Specific Color indicates different things as per country
- Navigation conventions: A 'exit' button might mean differently as per country.

Among all diversifications might be covered through print media conveying meaning of translation. Some extra would possibly come below the heading of 'translation' in an exceptionally practical method; however now not all of things in translation could include all the technical and advertising and marketing decisions that are encompassed by way of the concept of localization.

#### B. Motivation

The need for localization nowadays is huge from every issue, because it brings blessings to industries, clients, and government, without a doubt to everybody. Even extra importantly, usually localization saves lives in nations where human beings have trouble on get right of access to existence saving facts in their local language. The need of localization to industries, as their products should be depleted everywhere within the global. The earnings are an awful lot greater while the product is customized to the needs and preferences of a locale. It is a great idea to localize country or vicinity-centric of the user, as it brings additional valuation to your product and suggests which you appreciate their neighborhood history, language particularities, subculture and traditions. One of the

motive purpose behind localization is the concept that it facilitates expands your business.

### C. Literature Survey

The internet studies shows that all websites have reality multilingual internet site proposes different language vocabularies and data can be posted in a language-unbiased style, whilst appropriate culture-based (linguistic) data permitting to transfer languages may be saved one after the other. On this sense, the multilingual website records [3] can be realized as a layer of languages and resources on top of the prevailing connected facts architecture with the aid of adding linguistic data and its vocabularies in multi languages, using with overlapping between records with labels in multilingual languages and facilities to dynamically technique and traverse linked records throughout different languages [5]. A different graph based technique is a combination of the link structure and the translation path between the distinctive language variations of a multilingual website. Common methods for robots varies from performing a fully automated systems that helps to perform all task towards to semi-automatic structures that vantage current content which include input text or maybe prerecorded speech inside the transport of the final online robot[7][11].

In a business website, many viewers able to access products of online website which will describe in proper details and examine it with special language. Most of the companies will analyze purchaser comments in multiple languages for analyzing review comments [8][12]. Simply, comments in a couple of languages raise problems in studying the fabric. To analyze multilingual comments of purchasers categorize into: affirmative and negative sentiments. To evaluate it accordingly comments technique used as “Multilingual Sentiment classification (MSC)”.

## II. PROPOSED METHODOLOGY

### A. Multilingual Linked Data

A cloud data with a layer of multilingual languages which does overlapping between multilingual applications with linguistic statistics providing a set of services for growing websites. It provides access to link a different linguistic multilingual statistics [1]. A set of statistics that could be exported as related records. The data silos can be multilingual or monolingual. For producing multilingual linked facts a few offerings are performed to transform these statistics into multilingual linked facts. Further, few changes are needed to move lingual discovery and representation of mappings between related facts of vocabularies of different languages and datasets expressed in target languages [5]. Such cross-cultural linkage things are accessed by using localization services. These localize services which definitely translate vocabularies or ontology which used within the linkage information of cloud into variety of local languages.

### B. Graph model for Localization Pattern

To purpose of graph model along with the data aggregation to extract local patterns in a multilingual website, a tiny website test some subparts of different multilingual websites [2]. The Japanese website related to computer science research center having web pages in both language English and Western. The most important part of the Japanese city and a community site developed in Linux, launching various articles proposed in eleven languages and relying on its members for the translations of websites[6][10]. Languages of website followed the global or local web content pattern, with all pages having global information were available both in English and Western, the English language sub-contract graph becoming a sub chart of Japan one. Graph model follows the mirror patterns where both vocabulary versions followed the same structure.

### C. Dynamic generated Localization

An international web application consider web text structure and image captions provide randomly just like it treats actual web data elements as financial values. On this values translators can work in concealment of the web application

and HTML CODE page templates [2]. The technique proposes to modify strings tables to every local language. It is to codec locale language properties such as filename, linguistic properties which load the appropriate data file according to Locale's getting language method. The language files are simple key-value pairs. Take an example, Any webpage file, as properties would have title of page as Welcome to Sample page whereas a form setting file would have deal with a test page.

#### D. Reusable, Interactive, Multilingual online avatars

Online robot should try to deal with both the sound track and visual content required to implement the character. An audio track file would be in any target language is allocated specific moves and stored in the database. An online robot consists a skeleton including a library of movements are stored in database to contribute assets [4]. Whenever the user tries to make a movement, it requests in a specific language, the inter-related voice as per movements is gathered from the database and performed by robot. A voice recording of specific cultivate scripts are analyzed and combined to provide gestures and movements related to the events in the conversation [8].

#### E. Multilingual Sentiment Classification on Textual Data

To investigate an multilingual websites having product review in multilingual dataset with bilingual texts [5]. The classification is a formation associated with an ontology and textual structure category. The technique to classify multilingual text proposed as "Multilingual Sentiment Category (MSC)". This technique works by performing two main processing steps: lingual separation of multilingual data and sentiment category. The overview of the MSC

methodology can be presented as Fig.2

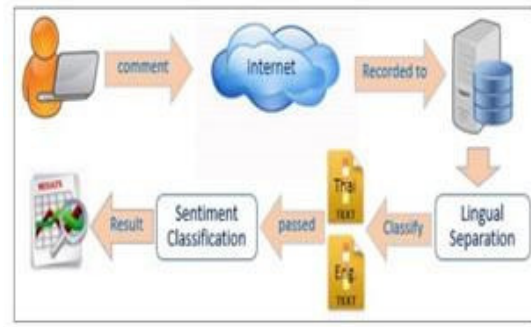


Fig 2: Overview of the SSC methodology [5]

### III. COMPARISON ANALYSIS

#### A. Techniques

Ontology Localization involves ontology mapping [1] helps to mapping between cross linguistic information. Linked structure and translation path [2] between different language helps to understand a multilingual website. To analyze multilingual data set, multilingual segment classification [4] is needed.

#### B. Methods

Multilingual sentiment classification [3] for analyzing bilingual texts. Ontology localization[12] of an adapting culture context of data. Online avatar uses speech of recording and 3DSMax, Maya to improve flexibility and effectiveness.

#### C. Advantages

The multilingual model [4] provides interactive to bilingual speech and to analyze it. Cross linguistic helps to dynamically mapping of data based on sentiment classification. A transition graph helps to identify appropriate cultural adaption.

### IV. CONCLUSION

An individual can't always understand all languages other than local language. To maintain a site which does appropriately localization is more meaningful than just translation. It helps to recognize the scope of improvement in multilingual websites. The different techniques proposed to analyze pattern of localization and conversion with maximum result. It can help to reduce errors and bugs in multilingual websites and

definitely will validate user friendliness of region-centric websites

#### REFERENCES

- [1] Jorge Gracia, Elena Montiel-Ponsoda , Philipp Cimiano , Asunción Gómez-Pérez, Paul Buitelaar , John McCrae, “Challenges for the multilingual Web of Data”, *Science, Services and Agents on the World Wide Web 11 (2012)* 63–71,2012.
- [2] Julien Bourdon and Toru Ishida, “A Graph Based Model for Understanding Localisation Patterns in Multilingual Websites“, 2011 Second International Conference on Culture and Computing, 2011.
- [3] Terence Parr, “Web Application Internationalization and Localization in Action”, ICWE’06, July 11-14, 2006, Palo Alto, California, USA, 2006.
- [4] Charlie Cullen , Carl Goodman , Paula McGloin , Anna Deegan, Evin McCarthy, “Reusable, Interactive, Multilingual Online Avatars”, 2009 Conference for Visual Media Production, 2009.
- [5] Jantima Polpinij, “Multilingual Sentiment Classification on Large Textual Data”, 2014 IEEE Fourth International Conference on Big Data and Cloud Computing, 2014.
- [6] Luo Yuansheng, “Multi-lingual Text Clustering Method Using Bilingual Semantic Correspondence Analysis”, 2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), 2016.
- [7] P. Tonella, F. Ricca, E. Pianta, and C. Girardi, “Restructuring multilingual web sites”, in *Software Maintenance, Proceedings. International Conference on. IEEE*, 2003, pp. 290–299.
- [8] R. George, K. Nesbitt, P. Gillard, and M. Donovan, “Identifying cultural design requirements for an australian indigenous website”, in *Proceedings of the Eleventh Australasian Conference on User Interface*, Volume 106. Australian Computer Society, Inc., 2010, pp. 89–97.
- [9] P. Sandrini, “Website localization and translation”, *LSP Translation Scenarios: Selected Contributions to the EU Marie Curie Conference Vienna*. Citeseer, 2007.
- [10] J. S. Modha, H. G. S. Pandi, S.J. Modha, “Automatic Sentiment Analysis for Unstructured Data”, *International Journal of Advanced Research in Computer Science and Software Engineering*, Vol. 3, Issue 12, 2013.
- [11] F. Bravo-Marquez and M. Mendoza, B. Poblete, “Meta-Level Sentiment Models for Big Social Data Analysis”, *Knowledge-Based Systems*, 2014.
- [12] D. Spohr, L. Hollink, P. Cimiano, “A machine learning approach to multilingual and cross-lingual ontology matching”, *Proc. of ISWC 2011*, to appear, Bonn, Germany, October 2011.