

Comments and Suggestions on paper of “An Analysis of Asian Language Web Pages”

I think that this paper analyzes and shows statistical value of Asian web pages in virtual world choosing the best possible area of study, I mean here, studying based on country level domain. Showing the statistical value of web is like counting stars in the sky. I would like to comment on this paper discussing achievements from the study and what trends we should go in future basing the following three main aspects of study in it.

1. The Diversity of Asian Languages on the web.
2. Multilingualism and dominant of cross-border languages.
3. Script and Encoding of Asian Languages on the web.

1. The diversity of Asian Languages on the web

Achievements: Using the powerful web crawling tool, Ubi Crawler, we can independently collect awesome amount of Asian web pages and due to Language Identification Module, we can aware living and dying Asian Languages on the web accurately reflecting the current status of Web sites.

According to statistical values in this paper, it shows Digital Language Divide exists at a serious level in Asia Region.

Trend in future: Language specific search engine should be implemented in accordance with the data acquired from this study.

Digital Divide should be bridged to form the inclusive information society striving to develop ICT infrastructure in national level and on the other hand, we should have encouragement in digital literacy ensuring greater mastery of Information Technology.

2. Multilingualism and dominant of cross-border languages

Achievements: Multilingualism and dominant of cross border languages in the web pages of Asian countries are encountered due to the rich diversity of multiple languages even in country level domains.

A rich diversity of written pages in variety of languages shows the lack of localization and standardization of nationally recognized language of a country.

Trend in future: In addition to country level domain, linguistic situations of other domains such as personal, governmental, educational and commercial domains of countries in Asia should be studied basing the IP information of each country.

Standardization and Localization in the above mentioned domains should be implemented in every nation to remain rooted in our mother tongue and to open to learn others.

3.Script and Encoding of Asian Languages

Achievement: This paper highlights on the chaotic situation of Language, Script and Encoding schemes of Asian Languages.

This paper also shows that less penetrating level of Unicode in Asian Countries and multiple scripts in Asian Countries lead to chaotic situations of script and encoding.

Trend in future: We should put into effort to solve such chaotic situations by adapting international standard for Asian Language technologies. The usages of Unicode should be encouraged not only in globally submitted documents such as web pages but also in locally implemented applications so as to bridge digital divide in Asia Region.

We should also have language resources website and language professional association that will survey up to date activity level of languages on Web handling chaotic situation of script and encoding.

Considering the ideas and achievements of this study, I believe that implementation of observation-collection instrument for Asian languages resources on the web will play a pivotal role in the development of linguistic database and cooperation among languages in Asia. Moreover, in the near future, I do hope to emerge a system which allows people from different countries to communicate with one another using their own mother tongues.