

# Requirements of Mobile Web for Indian Languages - Gap Analysis & Way ahead

Swaran Lata<sup>1</sup>, Somnath Chandra<sup>1</sup>, Prashant Verma<sup>2</sup>

<sup>1</sup> Web Standardization Initiative , TDIL Programme, Deit Y, India

<sup>2</sup> Web Standardization Initiative, Deit Y, India

## Abstract

*W3C(World Wide Web Consortium) Mobile Web Best Practices (MWBP) 1.0 is designed to improve user experience of the web in mobile devices. This practice includes DDC (Default Delivery Context), which allows providers to create appropriate experiences in the absence of adaptation and provides a baseline experience where adaptation is used. The paper provides a roadmap for identifying the current usage and potential of mobile technologies in development and presents the gaps between ongoing W3C MWI standards and Indian requirements. Based on the studies specifically in the Indian languages context, the paper proposes some of the changes required in the W3C MWBP 1.0 with respect to Indian languages and proposes the modified W3C DDC defined by W3C Best Practices Working Group(BPWG). Beside these gaps paper also shows some of the additional requirements of India languages. This paper also focuses about the different specifications used by Web Developers to design Websites with content in Hindi and other Indian Languages.*

**Keywords:** LRL, W3C, Mobile Web, DDC, MWBP, Survey, BPWG

## 1. Introduction

India being one of the world's fastest-growing telecommunications market due to a progressive regulatory regime, huge capital outlays for network expansion by operators, reductions in tariffs and cost of handsets. Parallel efforts of enhancing the languages content on the web, coupled by the development of various standards supporting these initiatives is the ultimate goal.

Mobile Web Best Practices W3C (MWBP) [1] 1.0 Basic Guidelines defines the issues that are mostly general in nature and are related to small devices like mobile phones, handheld devices, etc. It specifies Best Practices for delivering web content to mobile devices. The principal objective of W3C MWBP is to improve the user experience of the web when accessed from such devices. It has covered those issues that are general to all web applications and specific to languages of Latin origin. It has covered language issues related to Character Encoding, Content Storage Format, Markup, Page Size, etc.

The Default Delivery Context (DDC) [1] has been determined by the W3C Best Practices Working Group (BPWG), as the minimum delivery context specification necessary for a reasonable experience of the Web. The guidelines that make up the DDC were assessed by applying research, statistical analysis to the data set, and hands-on testing with selected mobile browsers to verify that the guidelines are representative of the population of mobile devices that can access the Internet. In assessing the DDC it was found that some guidelines conflicted with the use of Indian languages.

The proposed modified guidelines of the current W3C MWBP focuses in this paper for specific Indian languages.

This paper presents proposed modified DDC to better represent the population of mobile devices that can access Indian languages websites. The proposed modified DDC is based on the testing of the web pages having different language contents on various networks. This document also aims to deliver comprehensive gaps between MWBP and its requirement in India for enabling mobile web in Indian languages.

## 2. W3C Default Delivery Context

The Default Delivery Context was determined by the W3C Best Practices Working Group, as being the minimum delivery context specification necessary for a reasonable experience of the Web. This specification provides a set of guidelines to assist in mobile web development. It was recognized that devices that do not meet this specification cannot provide a reasonable experience of other one It has also been stressed that many devices exceed the capabilities defined by the DDC and encouraged not to diminish the user experience on those devices by developing only to the DDC specification and are encouraged to adapt their content, where appropriate, to exploit the capabilities of the actual device. The DDC specification consists of nine attributes that are expected to represent the population of mobile devices with Internet capability.

### 2.1 Challenges

The increasing mobile web technology has introduced many usability hurdles for the mobile users:-

- Mobile devices with different screen widths may make webpages look unpresentable
- Web pages in many different languages may not be supported on every mobile devices
- Mobile devices may not be able to render web pages with dynamic contents
- Mobile devices with varying bandwidth support is a concern to define standard page weight
- Lack of standardization of mobile keyboard in Indian languages.
- Lack of standard in transliteration form.

To overcome the difficulties it is important to review the DDC attributes like Resolution width, Character Encoding, CSS, JavaScript, Page Weight w.r.t Indian languages.

### 3. Complexities of Indian languages

There are 22 official recognized Indian languages. The Indian language scripts are mainly stems from ancient Brahmi script and perso-arabic family. The syllable formation possibilities are infinite due to large set of consonant, vowel, vowel signs. Many Conjunct formed are drastically of different look from their consonant counterpart. The conjuncts formation can be either liner or vertical.

- Linguistic Diversity is very rich and wide in India
- One Language –many scripts
- Many Language –one script
- Culturally different depending on region though using same script for different languages.
- Even wide difference for same language across different parts of the country

Some of the areas that need to be focused while building mobile applications in Indian languages are shown below:

1. Presentation: Due to limited screen size and the limited amount of material that is visible to the user, context and overview are lost
2. Input: The versatility of the solution lies in the inputting of Indian languages through the limited keys available on the mobile handset. Mobile device input is often difficult when compared with use of a desktop device equipped with a keyboard.
3. Lack of standardization: Problem of inconsistency and lack of standardization is a very big problem at least for Indic locale.

## 4. Gap analysis of Indian market and W3C MWBP

The current mobile standards do not cover Indian language requirements. The main gaps between Indian market and W3C mobile standards are :

- Indian language requirements are different from International requirements as Indian languages are complex in nature.
- Differences of standardization activities
- Indian language requirements are not incorporated in the same W3C documents.

As the 12 Indian scripts is a big challenge so there is a need to capture the Indian language requirements and the same requirements should be incorporated in the global mobile standards.

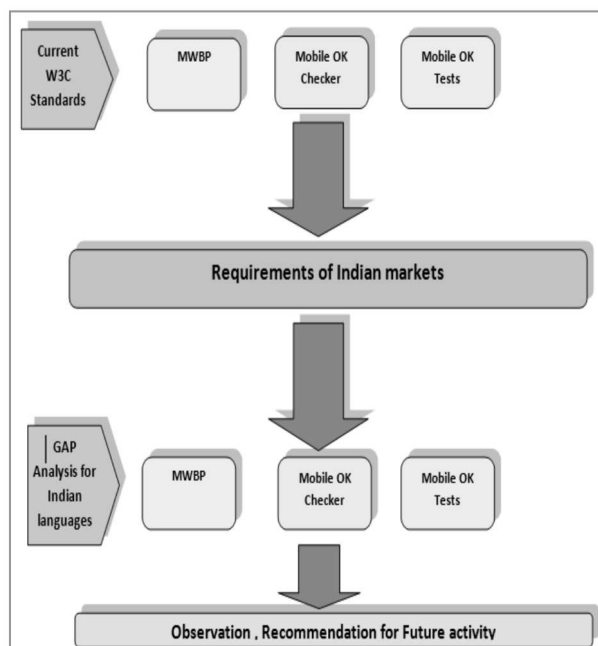


Figure 2: Gap analysis

## 5. Modification required in Indian context

### 5.1 Usable Screen Width

In this era of mobile web, mobile devices have developed considerably and in this development the screen width has also varied a lot. The market has constantly supported 240px screen width, as per the statistics of mobile devices in the market. W3C recommended 120px support in MWBP 1.0. Then with consistent production of 240px screen width mobile devices, covering around 45% of market production every time, 240px screen width covered majority of the market production. In year 2012 the lower screen widths covers less than 5% of the market production whereas 240px still covers majority of the market production.

Hence the new proposed standard is 240px screen width. The yearly market statistics is shown in Fig.3.

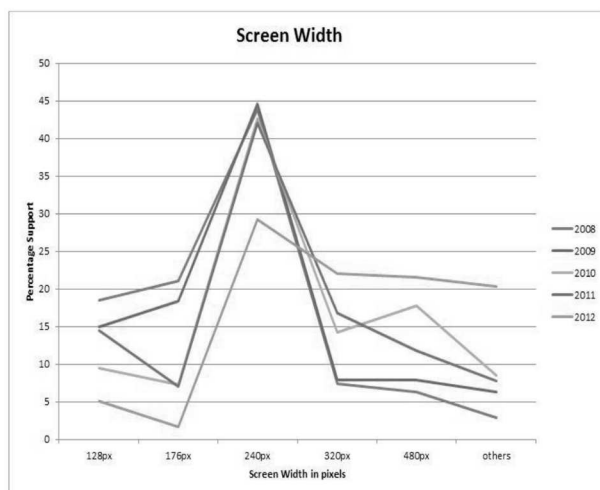


Figure 3: statistics of Mobile Screen Width

Indian users cannot read text properly in small screen width as Indian scripts is complex in nature. Hence the new proposed screen width is 240px.

## 5.2 Character Encoding

In the Mobile technology, the multilingual data handling becomes vital across different layers. Any chosen encoding scheme should consider the following:-

- The encoding scheme should support all possible characters.
- Combinations as per Unicode standard.
- There should be provision to change languages within single message.
- The encoding should be flexible for future Unicode standard.

The UTF-8 encoding is more widely used by websites, emails and many open source applications. It is able to represent any character in the Unicode standard.

## 5.3 Image Format Support

JPEG (Joint Photographic Experts Group) and GIF89a (Graphic Interchange Format) support were recommended by W3C in the MWBP 1.0 (July, 2008), but apart from these PNG (Portable Network Graphics) which support binary transparency and alpha channel transparency is being used widely in Mobile Phones. Hence PNG image format is proposed to be added DDC image support attribute. The yearly market statistics of the PNG support in mobile devices. Scalable Vector Graphics (SVG), provides an XML-based markup language to describe two-dimensions vector graphics. Since these graphics are described as a set of geometric shapes, they can be zoomed at the user request, which makes them well-suited to create graphics on mobile devices where screen space is limited. They can also be

easily animated, enabling the creation of very advanced and slick user interfaces. PNG image format is also adopted as a standard for E-governance applications.

## 5.4 Page Weight

Mobile devices are now able to handle larger web pages. The improvements in the underlying network, browser rendering and available memory in devices have led to support of handling larger web pages. So the 20Kb standard needs to be revised. The fact that mobile devices have now have better supports some problems still exists:-

- It still has only limited types of content support by means of processing and display.
- Mobile data transfer also costs money depending on bandwidth usage.

Hence to deal with above shortcomings, an optimum page weight should be defined. It should be defined such that the web page is downloaded on majority of devices in psychologically acceptable time. Web page navigation is a human-computer interaction, where heuristic evaluation is a testing technique for usability. This optimum page weight value has been established through heuristics that if the page is downloaded and displayed within 4 seconds on the device, then page weight value is accepted. If the download exceeded 4 seconds, a smaller page size is tested for standard. So by following this methodology, the optimum page weight was set to 20Kb in MWBP 1.0.

For developing new standard, we have revised the acceptable time for page download and display. Heuristic cases have been developed by taking into consideration different age groups and classes of users or people. The content of sample pages in Indian language has been chosen carefully, such that it strikes optimum balance and facilitates the survey. Through the survey, we record the satisfactory download time on the dataset for every user.

During the survey the most noticeable thing was different people had different meaning of good, satisfactory, neutral and poor. It was difficult to understand exact meaning, especially from the people of lower class. We considered the most occurring values in the range and ignored very strange values in a particular range due to different mentality. Based on this survey result we accepted the heuristic of 4 seconds and analyzed average page weight that is loaded in a time period of 4 seconds. Average page weight with respect to an average bandwidth of a mobile device was found to be 50Kb with images.

It is proposed to modify the page size limit to 50 KB w.r.t Indian languages for total page weight along with images and stylesheets.

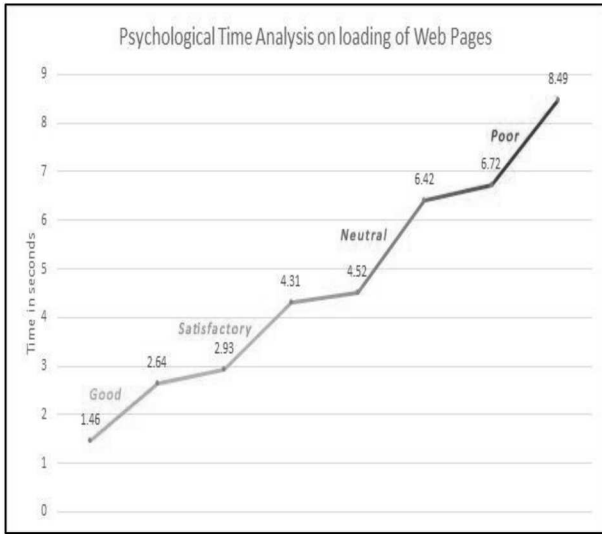
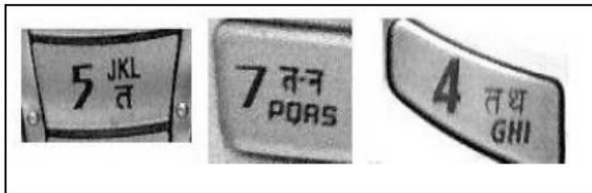


Figure 3: Psychological Time analysis on loading of Web Pages

### 5.5 Text entry issue

Mobile text entry is often difficult as compared to the desktop as there is very limited space in the mobile keypads. It is difficult to display more than one language on the mobile keypad due to small space available on the keypads. The following figure shows the variation of Hindi characters in different mobile keys :



### 5.6 Scripting Support

A scripting language normally controls the operations of a web page making it more interactive. There are some scripting languages that are used on web for adding dynamic capabilities to World Wide Web pages. Web pages marked up with HTML or XML are largely static documents so web scripting is used to add and modify information on a page dynamically. These languages are generally called Web Scripting languages.

Since there are many different scripting languages of different origin, a standard had to be defined covering majority of the scripting languages over same origin. ECMA Script [5] is the scripting language standardized by ECMA International in the ECMA-262[6] specification and ISO/IEC 16262. The language is widely used for client-side scripting over the web, in the form of several well-known implementations such as JavaScript (developed by Netscape Navigator), JScript (implemented in Internet Explorer) and ActionScript.

## 5.6 Additional requirements for Indian languages

The following future technologies that increase the capabilities to the mobile context are need to be focused for Indian languages :

### 5.6.1 Graphics and layout :

Fonts play an important role for building graphics interface. Indian languages fonts are limited in the mobile devices and have large bandwidth. The W3C WOFF 1.0(Web Open font format ) enable easy to download fonts automatically through style sheets. The future standard WOFF 2.0 compress the size of the font up to 30-40 % of the original size. So this font format plays an important role for Indian languages fonts. The following figures shows the WOFF content results on iphone and android :

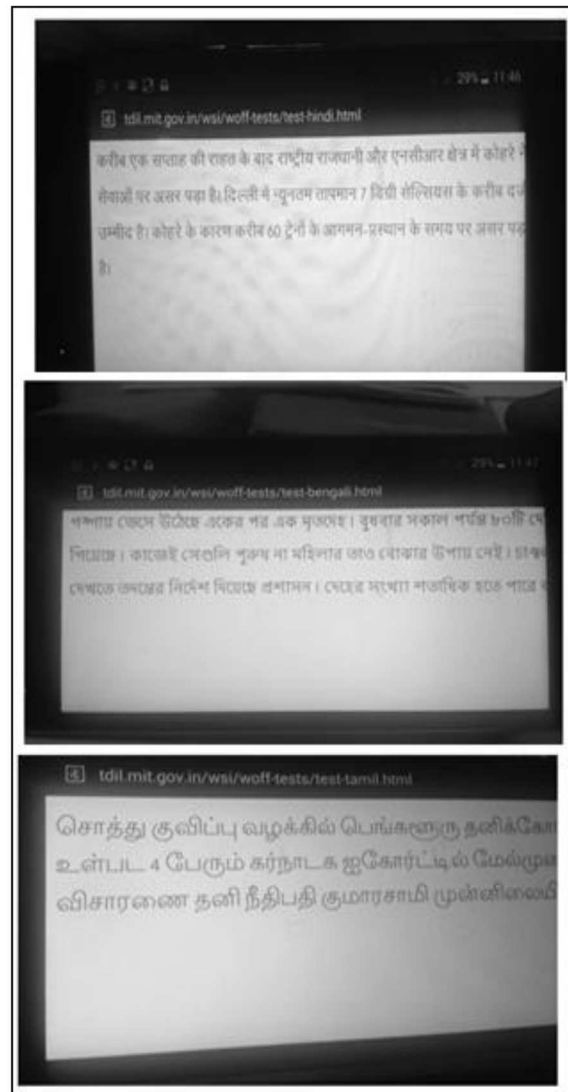


Figure 4: test results on Android 4.4.2



Figure 5: test results on iphone4 Simulator

### 5.6.2 Forms :

The rich forms with HTML is depends on the user input in web applications. It is very difficult to input text in Indian languages due to variation of keyboards and difficult mechanism used in keypads for Indian languages. So it is important to create user friendly virtual keypads for Indian languages and form controls that optimize the user input mechanism specially for Indian languages.

## 6. Modified Default Delivery Context for Indic Languages

Based on the above study , the proposed DDC for Indian languages with W3C defined DDC is shown in the Fig 6.

The major modification is required in screen width, total page weight, image format support, script for Indian languages so that the users can use mobile web pages seamless in Indian languages.

| DDC attributes            | W3C MWBP 1.0 DDC                        | Proposed DDC for Indian languages |
|---------------------------|---|-----------------------------------|
| Usable screen width       | 120 pixels, minimum                     | 240 pixels, minimum               |
| Markup language support   | XHTML Basic 1.1                         | XHTML Basic 1.1                   |
| Character encoding        | UTF-8                                   | UTF-8                             |
| Image format support      | JPEG and GIF 89a                        | JPEG, GIF 89a and PNG             |
| Maximum Total Page Weight | 20 KB                                   | 50 KB with images                 |
| Colors                    | 256 Colors, minimum                     | 256 Colors, minimum               |
| Style Sheet Support       | CSS Level 1 and CSS Level 2 @media rule | CSS Level 1 and CSS Level 2       |

|        |   |             |
|--------|---|-------------|
|        | together with the handheld and a ll media types |             |
| HTTP   | HTTP 1.1  | HTTP 1.1    |
| Script |   | ECMA script |

Figure 6: Default Delivery Context Specifications

## 7. Conclusion

W3C Mobile Web Initiative has provided a platform to improve user's mobile web experience. In this work we have presented results from an extensive characterization study of mobile web content having specific focus in Indian languages. Also users all over the world do not have homogeneous nature and thinking. These factors results in great calculations to conclude which technology is best adapted by users. Only considering technical support and ease of the technology cannot conclude standards. Moreover standards always need frequent reviews, because technology changes, leading to change in standards with time. The paper proposes the change in some of the DDC attributes w.r.t Indian languages. This paper summarizes results the various requirements that needs to be incorporate in the W3C MWBP and these changes may reflect to W3C Mobile OK Checker, so that it checks that localized web content can be delivered in a manner that is conformant with Mobile Web Best Practices.

## References

- [1] <http://www.w3.org/TR/mobile-bp/>
- [2] <http://www.w3devcampus.com/mobile-web-and-application-best-practices-training/>
- [3] <http://wurfl.sourceforge.net/>
- [4] <http://triin.net/temp/html-elements.html>
- [5] <http://www.ecmascript.org/>
- [6] <http://www.ecma-international.org/publications/standards/Ecma-262.htm>
- [7] <http://dev.opera.com/articles/view/javascript-support-in-opera-mini-4/>
- [8] <http://www.mozilla.org/en-US/mobile/5.0/releasenotes/>
- [9] <http://www.quirksmode.org/m/table.html>
- [10] [http://www.statowl.com/web\\_standards\\_javascript\\_version\\_support.php?l=1&&interval=month&chart\\_id=5&fltr\\_br=&fltr\\_os=&fltr\\_se=&fltr\\_cn=&timeframe=ytd&timeframe=custom2008-092011-11](http://www.statowl.com/web_standards_javascript_version_support.php?l=1&&interval=month&chart_id=5&fltr_br=&fltr_os=&fltr_se=&fltr_cn=&timeframe=ytd&timeframe=custom2008-092011-11)
- [11] [http://docs.blackberry.com/en/developers/deliverables/11844/Feature\\_JavaScript\\_512506\\_11.jsp](http://docs.blackberry.com/en/developers/deliverables/11844/Feature_JavaScript_512506_11.jsp)
- [12] <http://www.w3.org/Mobile/mobile-web-app-state/>