

Internationalization & Localization Testing



HCL

Copyright © 2022 HCL Technologies Limited

Internationalization & localization are subsets of globalization

Internationalization & localization

Internationalization

W3C

MOBILE
WEB

Cascading
Style Sheet

E-
Government

Taking a product and making it linguistically and culturally appropriate to the target locale (country/ region and language) where it will be used and sold"

Localization

Process of generalizing a product so that it can handle multiple languages and cultural conventions without the need for re-design.

Internationalization

Copyright © 2022 HCL Technologies Limited

HCL

Localization Vs Internationalized

Internationalization & localization

➤ Designing and developing in a way that removes barriers to localization or international deployment.

Internationalization

➤ Providing support for features that may not be used until localization occurs.

W3C

➤ Enabling code to support local, regional, language, or culturally related preferences.

MOBILE WEB

➤ Separating localizable elements from source code or content, such that localized alternatives can be loaded or selected based on the user's international preferences as needed.

Cascading Style Sheet

E-Government

➤ It can be localized quickly.

Copyright © 2022 HCL Technologies Limited

HCL

The Internationalization Approach to Globalization

Internationalization & localization

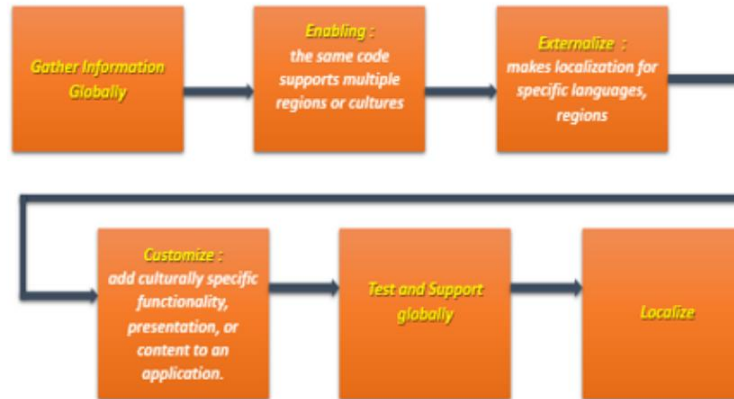
Internationalization

W3C

MOBILE WEB

Cascading Style Sheet

E-Government



Copyright © 2022 HCL Technologies Limited

HCL

Internationalization & localization

Internationalization

W3C

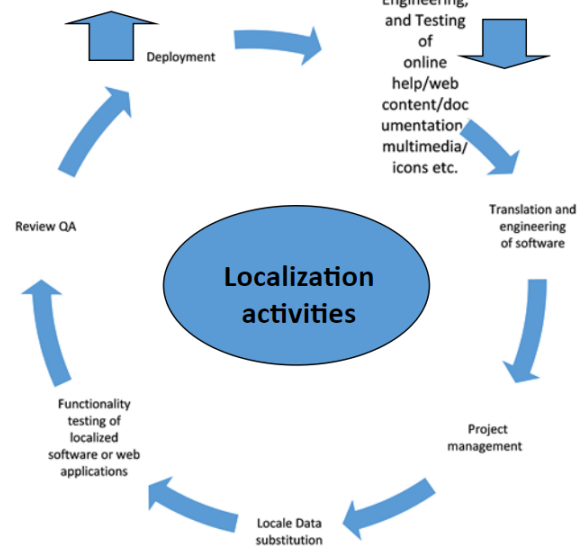
MOBILE WEB

Cascading Style Sheet

E-Government

Localized Application

Internationalized application to be Localized



Copyright © 2022 HCL Technologies Limited

HCL

Complexity of Indian Scenario

Internationalization & localization

Internationalization

W3C

MOBILE WEB

Cascading Style Sheet

E-Government

- India is Multilingual Multi script Country with 22 languages and 11 scripts; population over 1 Billion
- Less than 5 percent of people can read & write English. Over 95 percent population is deprived of the benefits of English-based Information Technology

Issues regarding Indian Languages

- Orthography – Spelling issues
- Pronunciation – may be directly mapped but not always
- One script-many languages
- Many languages – one Script

Copyright © 2022 HCL Technologies Limited

HCL

Internationalization & localization

Internationalization

W3C

MOBILE
WEB

Cascading
Style Sheet

E-
Government



Copyright © 2022 HCL Technologies Limited

HCL

Internationalization Activity in W3C

Internationalization & localization

Internationalization

W3C

MOBILE WEB

Cascading Style Sheet

E-Government

- The W3C Internationalization (I18n) Activity works with W3C working groups to make it possible to use Web technologies with different languages, scripts, and cultures.
- It is to ensure that W3C's formats and protocols are usable worldwide in all languages and in all writing systems.
- The Internationalization (I18n) Activity statement explains concepts relating to internationalization, as well as the current situation and the role within the W3C of the Internationalization Activity.

Copyright © 2022 HCL Technologies Limited

HCL

W3C Internationalization Resources

Internationalization & localization

Internationalization

W3C

MOBILE WEB

Cascading Style Sheet

E-Government

➤ Internationalization of Web design & Applications

- Character Model for World Wide Web
- Authoring Techniques for XHTML & HTML
- Authoring CSS
- Unicode in XML

➤ Internationalization of Web Architecture

- Language tags and Local Identifiers
- Internationalization Tag Set

➤ Internationalization of XML

- Best practices for XML Internationalization

➤ Internationalization of Web Services

- Language tags and Local Identifiers for World Wide Web

Copyright © 2022 HCL Technologies Limited

HCL

Activities to be undertaken by W3C India Office

Internationalization & localization

Internationalization

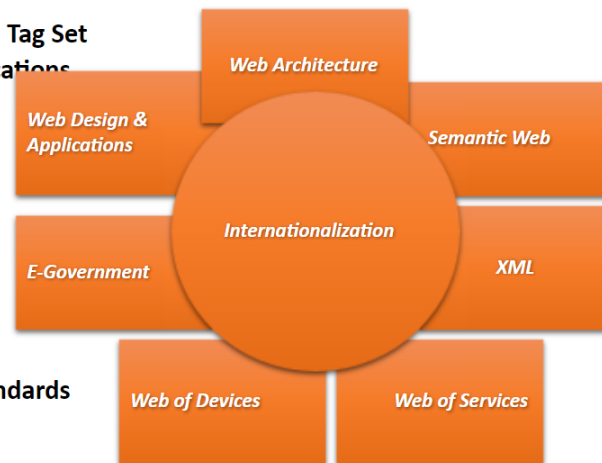
W3C

MOBILE WEB

Cascading Style Sheet

E-Government

- ❖ Internationalization
 - ❖ Internationalization Tag Set
- ❖ Web Design and Applications
 - ❖ Styling
 - ❖ Html
 - ❖ Xhtml
 - ❖ Wai
- ❖ Web Architecture
 - ❖ XML
- ❖ Semantic Web
 - ❖ OWL and RDF
- ❖ XML Technology
 - ❖ XML associated standards
- ❖ Web of Services
 - ❖ SOA
- ❖ Web of Devices
 - ❖ Mobile Web Initiative
 - ❖ Voice
- ❖ E-Government
 - ❖ Use cases



Copyright © 2022 HCL Technologies Private Limited

HCL

MOBILE WEB

Internationalization & localization

Internationalization

W3C

MOBILE WEB

Cascading Style Sheet

E-Government

Challenges :

- Adopting right encoding scheme
- Availability on handsets
- Usability of Mobile Web Browser
- Web support of all Indian languages
- Study on specific requirements for Indian languages for W3C Mobile Web Best Practices
- Must support standards and specifications
- Access to all handset features

Copyright © 2022 HCL Technologies Limited

HCL

MOBILE WEB

Internationalization & localization

Internationalization

W3C

MOBILE WEB

Cascading Style Sheet

E-Government

Issues on Mobile Web

- *Character Encoding*
- *Bandwidth and Cost*
- *Presentation Issues*
- *Input*
- *Device Limitations*
- *Lack of standardization*
- *Fonts*
- *Backward Compatibility with Legacy Devices*
- *Lack of standardization*
- *Rendering Issues*

Messaging Issues

- *Lack of availability for all characters.*
- *There is no guarantee that a message encoded will be displayed properly at the receiving terminal.*
- *Issue of Multiple Script -one language not addressed.*
- *Standardization of glyph support, syllable composition logic is also an important aspect and is dependent on the implementation level of handset manufacturer.*
- *Legacy Systems*

HCL

Copyright © 2022 HCL Te

MOBILE WEB

Internationalization & localization

Internationalization

W3C

MOBILE WEB

Cascading Style Sheet

E-Government

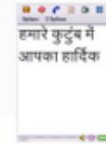
Issues in Mobile Keypads

- Multi-tap issues

Too
char

Hindi Alphabets mapping in Mobile

2	अ आ इ ई उ ऊ ऋ
3	ए ऐ ओ औ अः
4	क ख ग घ ङ
5	च छ ज झ ञ
6	ट ठ ड ढ ण
7	त थ द ध न
8	प फ ब भ म
9	य र ल ण व श ष स ह



char No way to know which

- Dictionary Based issues

- Difficult to learn and operate for the target segments.
- Different spelling for , , even , many permutations. Which is the one to be mapped

MOBILE WEB

Internationalization & localization

Internationalization

W3C

MOBILE WEB

Cascading Style Sheet

E-Government

Suggestions

- *In terms of internationalization, operators must support appropriate character encoding on the signaling channel which would allow all characters of the world to be represented.*
- *Need of investigation and study of major issues for enabling Mobile web in Indian languages*
- *Standardization of mobile media also required to be addressed taking into consideration of specific requirements of each of Indic languages.*

Copyright © 2022 HCL Technologies Limited

HCL

Internationalization & localization

Internationalization

W3C

MOBILE WEB

Cascading Style Sheet

E-Government

- Road Map :

- Character encoding as per Unicode Standard
- Enable Mobile Web in Indian languages
- Initiation of study for Mobile Web Best Practices 1.0 with respect to
- requirements for 4 Indian languages : Hindi, Bangla, Marathi, Tamil

PRONUNCIATION LEXICON SPECIFICATION

Internationalization & localization

Internationalization

W3C

MOBILE WEB

Cascading Style Sheet

E-Government

- Homograph:
 - Most languages have words with different meanings but the same spelling (and sometimes different pronunciations), called homographs. For example, in Hindi the word कनक (घत्तूरा) and the word कनक (सोना) have identical spellings but different meanings.
- Homophones:
 - Similar pronunciation but different meanings (and possibly different spellings), for instance "बली" (बलवान) and "बलि" (बलिदान).
 - Similar pronunciation, similar meaning with different spelling

"हिंदी" and "हिन्दी"

Roadmap :

Initiation of study for PLS 1.0 with respect to requirements for 3 Indian languages : Hindi and Bangla, Tamil

Copyright © 2022 HCL Technologies Limited

HCL

Cascading Style Sheet

Issues :

Internationalization & localization

Internationalization

W3C

MOBILE WEB

Cascading Style Sheet

E-Government

Drop Letter in Indian languages

- Issues for Indian Languages with respect to first character used in Hindi, Malayalam, Bengali, Telugu and Gujarati etc

The image shows three examples of text in different Indian languages, illustrating the 'Drop Letter' issue. The first example is in Hindi, showing the text 'बिजे पि-र सदापति राजनाथ सिंह' where the first character 'बि' is not properly rendered. The second example is in Bengali, showing the text 'বিজে পি-র সাদাপতি রাজনাথ সিংহ' where the first character 'বি' is not properly rendered. The third example is in Malayalam, showing the text 'വിജെ പി-ര സാദാപതി രാജനാഥ സിംഗ്' where the first character 'വി' is not properly rendered. The text is displayed in a standard font, and the issue is highlighted by the fact that the first character is not properly rendered or is missing.

Cascading Style Sheet

Internationalization & localization

Internationalization

W3C

MOBILE WEB

Cascading Style Sheet

E-Government

- Underlining of characters
 - There are some examples of Indian languages in which Matra's are not readable due to underlining of characters

Hindi - अन्य भाषाओं में भी अनुवाद

Punjabi- ਚੰਨ

Bengali- ভাই মূল্যে আর্কাইভ একই ওলি মালতি।

Gujarati - સરદાર ગુજરી

Marathi- मराठी मुला मुलींची नावे

Tamil- நீரிற்சுமிநி யிளமை நிறைசெல்வம்

Telugu - కెరనా టెల్యూ TV9 వీగెం " డ్యూం డి డింకర్ " వేర్

Copyright © 2022 HCL Technologies Limited

HCL

Cascading Style Sheet

Internationalization & localization

Internationalization

W3C

MOBILE
WEB

Cascading Style Sheet

E-Government

- Vertical arrangements

चौ	य	or	य	or	यक्	श	श	
द	क्ता		क्		ता	क्ति	or	क्
			ता					ति

- **Formatting issues :**
 - Horizontal justification



- *Bullets and Numbering*

W1	U+0000	W9	U+000A, F 3
W2	U+000B	W4	U+000A, F 3
W3	U+0007	W5	U+000A, F 2
W4	U+0009	W6	U+000A, F 3
W5	U+0009	W7	U+000A, F 3
W6	U+0009	W8	U+000A, F 3
W7	U+000A	W9	U+000A, F 3
W8	U+000A	W10	U+000A, F 3
W9	U+000A	W11	U+000A, F 3
W10	U+000A	W12	U+000A, F 3
W11	U+000A	W13	U+000A, F 3
W12	U+000A	W14	U+000A, F 3
W13	U+000A	W15	U+000A, F 3
W14	U+000A	W16	U+000A, F 3
W15	U+000A	W17	U+000A, F 3
W16	U+000A	W18	U+000A, F 3
W17	U+000A	W19	U+000A, F 3
W18	U+000A	W20	U+000A, F 3
W19	U+000A	W21	U+000A, F 3
W20	U+000A	W22	U+000A, F 3
W21	U+000A	W23	U+000A, F 3
W22	U+000A	W24	U+000A, F 3
W23	U+000A	W25	U+000A, F 3
W24	U+000A	W26	U+000A, F 3
W25	U+000A	W27	U+000A, F 3
W26	U+000A	W28	U+000A, F 3
W27	U+000A	W29	U+000A, F 3
W28	U+000A	W30	U+000A, F 3
W29	U+000A	W31	U+000A, F 3
W30	U+000A	W32	U+000A, F 3
W31	U+000A	W33	U+000A, F 3
W32	U+000A	W34	U+000A, F 3
W33	U+000A	W35	U+000A, F 3
W34	U+000A	W36	U+000A, F 3
W35	U+000A	W37	U+000A, F 3
W36	U+000A	W38	U+000A, F 3
W37	U+000A	W39	U+000A, F 3
W38	U+000A	W40	U+000A, F 3
W39	U+000A	W41	U+000A, F 3
W40	U+000A	W42	U+000A, F 3
W41	U+000A	W43	U+000A, F 3
W42	U+000A	W44	U+000A, F 3
W43	U+000A	W45	U+000A, F 3
W44	U+000A	W46	U+000A, F 3
W45	U+000A	W47	U+000A, F 3
W46	U+000A	W48	U+000A, F 3
W47	U+000A	W49	U+000A, F 3
W48	U+000A	W50	U+000A, F 3
W49	U+000A	W51	U+000A, F 3
W50	U+000A	W52	U+000A, F 3
W51	U+000A	W53	U+000A, F 3
W52	U+000A	W54	U+000A, F 3
W53	U+000A	W55	U+000A, F 3
W54	U+000A	W56	U+000A, F 3
W55	U+000A	W57	U+000A, F 3
W56	U+000A	W58	U+000A, F 3
W57	U+000A	W59	U+000A, F 3
W58	U+000A	W60	U+000A, F 3
W59	U+000A	W61	U+000A, F 3
W60	U+000A	W62	U+000A, F 3
W61	U+000A	W63	U+000A, F 3
W62	U+000A	W64	U+000A, F 3
W63	U+000A	W65	U+000A, F 3
W64	U+000A	W66	U+000A, F 3
W65	U+000A	W67	U+000A, F 3
W66	U+000A	W68	U+000A, F 3
W67	U+000A	W69	U+000A, F 3
W68	U+000A	W70	U+000A, F 3
W69	U+000A	W71	U+000A, F 3
W70	U+000A	W72	U+000A, F 3
W71	U+000A	W73	U+000A, F 3
W72	U+000A	W74	U+000A, F 3
W73	U+000A	W75	U+000A, F 3
W74	U+000A	W76	U+000A, F 3
W75	U+000A	W77	U+000A, F 3
W76	U+000A	W78	U+000A, F 3
W77	U+000A	W79	U+000A, F 3
W78	U+000A		

Copyright © 2022 HCL Technologies Limited

HCL

Cascading Style Sheet

Internationalization & localization

Internationalization

W3C

MOBILE WEB

Cascading Style Sheet

E-Government

- Indentation of character



Challenges :

- Implementation of CSS standards developed by W3C regarding Indian languages
- Standards however need to be provided to those developing CSS so that by default user could have the facility to use bulleting in his own Indic languages.

Roadmap :

Initiation of study for CSS 2.0 specifications with respect to requirements for 5 Indian languages : Hindi, Banala, Punjabi, Kannada, Tamil

E-Government

Internationalization & localization

Internationalization

W3C

MOBILE WEB

Cascading Style Sheet

E-Government

Issues :

- *Some applications are Completely in English.*
- *Some applications have static content in local language but forms in English.*
- *Some applications are multi-lingual but only in limited languages (e.g., English and only one local language) .*

Use cases :

Use Case for Land Records

- *As all type of data related to land, available in these land records*
- *They are used for various planning processes although the manual maintenance of this land record does hinder in effective collation and analysis of the data contained in them.*

E-Government

Internationalization & localization

Internationalization

W3C

MOBILE WEB

Cascading Style Sheet

E-Government

Challenges :

- *enable e-governance applications in Indian languages.*
- *Compliance with W3C standards.*

Roadmap :

- *Major E-Gov Application in Indian languages need to be studied for improving access to Government through better use of the web.*

Target languages : Hindi, Bangla, Marathi ,Telugu, Tamil

Copyright © 2022 HCL Technologies Limited

HCL

WEB ACCESSIBILITY INITIATIVE

Internationalization & localization

Internationalization

W3C

MOBILE WEB

Cascading Style Sheet

E-Government

Challenges :

- Make Web content accessible to people with disabilities w.r.t Indian languages
- WCAG 2.0 Guidelines for success criteria vis-a-vis selected recommendations relevant to Indian context

Initiative in India :

- “Guidelines for Indian Government websites” by NIC , Govt. of India
- STQC Implementing WCAG 2.0 Accessibility through Website Quality Certification
- Centre for Internet and Society developing authorized translation of WCAG 2.0 Guidelines

Roadmap :

- Meet WCAG 2.0 guidelines & techniques w.r.t Indian languages
- Initiation with Hindi, Bangla, Marathi, Telugu, Tamil

Copyright © 2022 HCL Technologies Limited

HCL

Challenges

Internationalization & localization

➤ Multiplicity of Languages

Internationalization

➤ Evolution of Orthography

W3C

➤ Lack of Standardization

MOBILE WEB

➤ Enabling Mobile web in Indian languages

Cascading Style Sheet

➤ Initiation of study for a W3C recommendations with respect to requirements for Indian languages

E-Government

➤ Adoption of W3C standards in terms of Internationalization

➤ Indian Websites should be fully W3C Complaint

➤ E-Gov Application in Indian languages need to be studied for better use of the web.