Standards & Best Practices for the World Wide Web !

Richard Ishida
W3C Internationalization Activity

Objectives

- Look at recent developments and current issue at the W3C with regard to standards and best practices that make the World Wide Web worldwide.
- Introduce a new initiative, funded by the European Commission, to investigate, provide information about, and recommend work on standards and best practices for the multilingual web.
- Suggest ways for you to get involved.

Outline

- About the W3C
- Recent developments
- Requirements
- Current discussions
- Social context
- Best practices
- MutlingualWeb project
- Getting involved

About the W3C

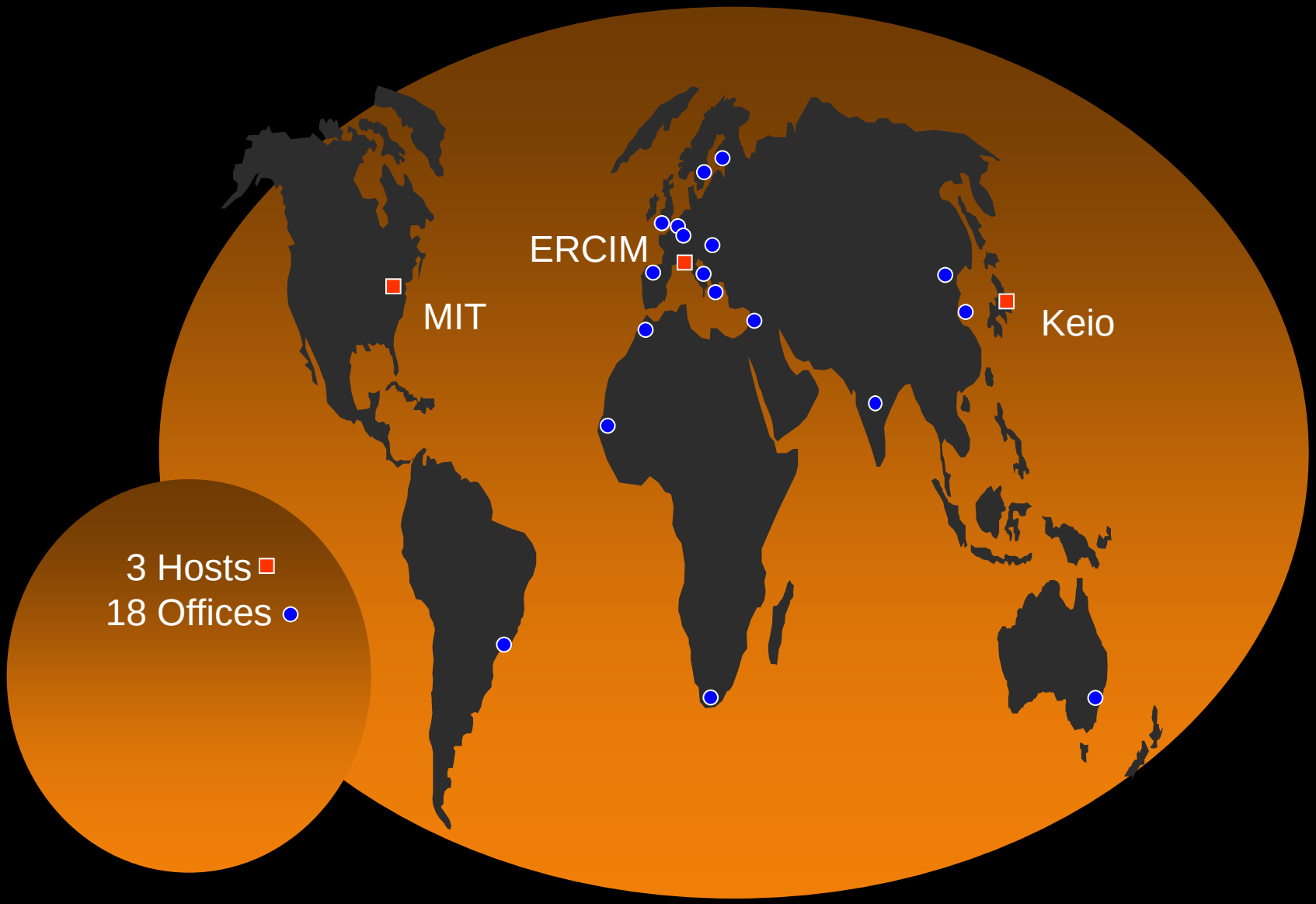
About the Consortium

1994: World Wide Web Consortium created and still led by:
Sir Tim Berners-Lee, inventor of the WWW.

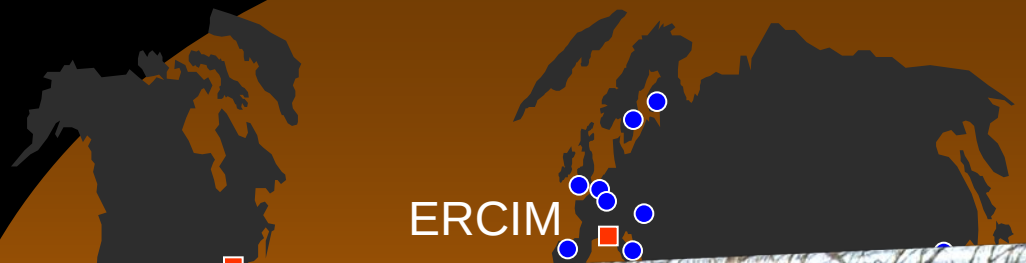
Mission:
Lead the
technical evolution of the Web
and ensure its
interoperability

Keywords:
consensus and vendor neutrality

About the Consortium

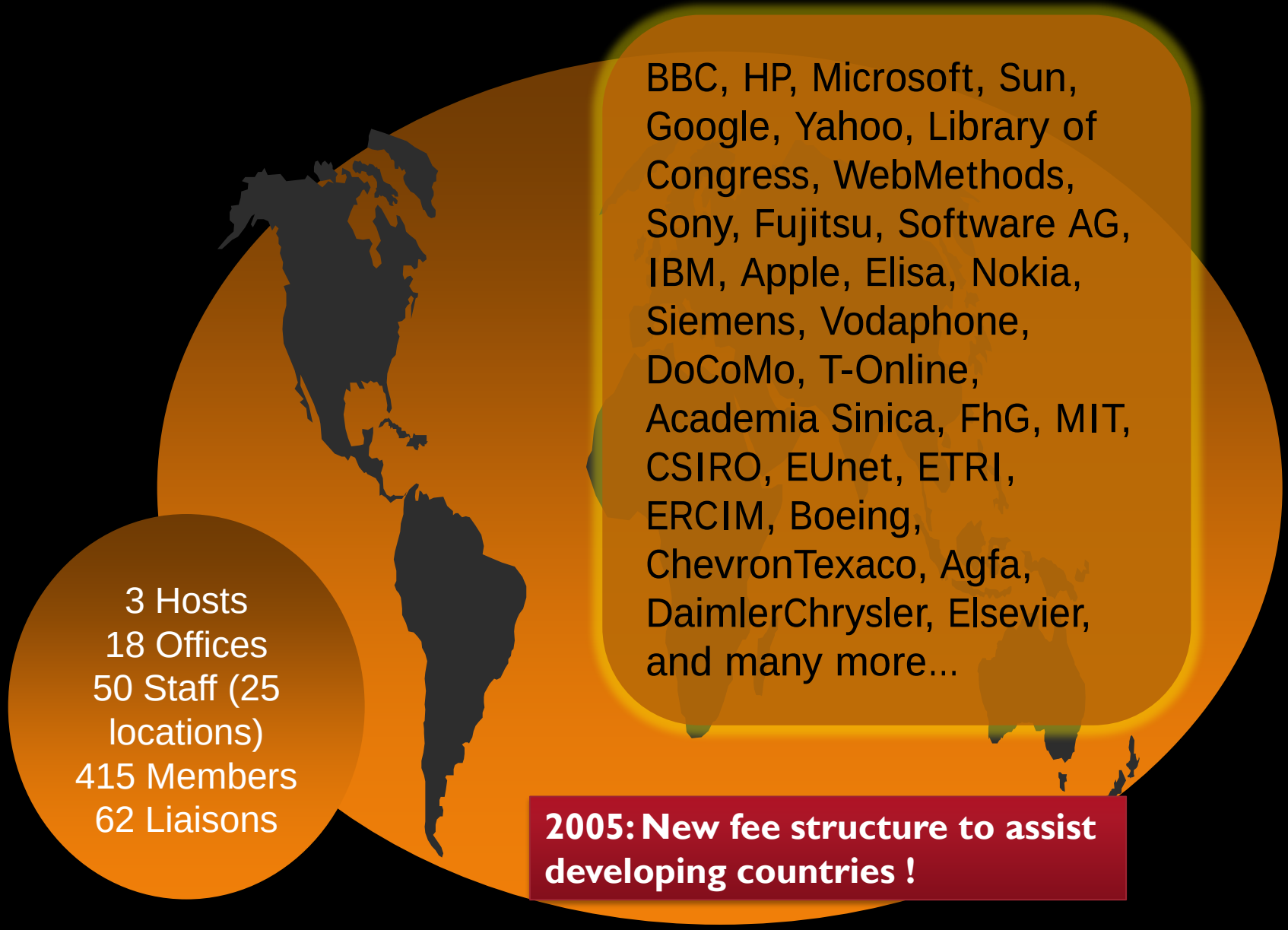


About the Consortium



3 Hosts
18 Offices
50 Staff (25 locations)

About the Consortium



BBC, HP, Microsoft, Sun, Google, Yahoo, Library of Congress, WebMethods, Sony, Fujitsu, Software AG, IBM, Apple, Elisa, Nokia, Siemens, Vodaphone, DoCoMo, T-Online, Academia Sinica, FhG, MIT, CSIRO, EUnet, ETRI, ERCIM, Boeing, ChevronTexaco, Agfa, DaimlerChrysler, Elsevier, and many more...

3 Hosts
18 Offices
50 Staff (25 locations)
415 Members
62 Liaisons

2005: New fee structure to assist developing countries !

About the Consortium

Accountable to the global public !

Technology & Society

- eGovernment
- Patent Policy
- Privacy
- Security

Technology development

- Fonts
- Graphics (SVG, WebCGM)
- HTML
- Math
- Mobile Web
- Multimodal Interaction
- Rich Web Client (WebApps)
- Semantic Web
- Style (CSS)
- Synchronized Multimedia
- Ubiquitous Web Apps
- Video in the Web
- Voice Browser
- Web Services
- Xforms
- XML

Web for All

- Internationalization
- WAI International Program Office
- WAI Technical Activity

22 Activities

50 Working Grp
13 Interest Grp
5 Coordination Groups

Internationalization Activity

- Help W3C Working Groups understand issues and build in requirements relating to worldwide support for Web technologies
- Check specifications in Working Draft stage for internationalization issues and develop best practices for specification developers
- Liaise with other standards organizations to develop support for the international Web
- Evangelize the need to consider multiple languages and scripts when developing Web technologies of any kind
- Help users of Web technology understand what's available to them and how to use it by developing best practices and other resources

عالمگیر ویب کو حقیقی طور پر عالمگیر بنانا

ᑖᕐᓱᒪ ᐃᑭᐊᖅᑭᕕᒃ ᓱᓚᕐᔪᐊᓕᒫᒥᒃ ᓈᕆᑎᑉᐸᑉ.

"Дүниежүзілік торды" нағыз дүниежүзілік етеміз!

वर्ल्ड वाईड वेबलाई यथार्थमै विश्वव्यापी बनाउने !

የዓለም አቀፉን ድር በእውነት አለም አቀፍ ማድረግ!

Κάνοντας τον Παγκόσμιο Ιστό πραγματικά Παγκόσμιο

ਵਰਡ ਵਾਈਡ ਵੈਬ ਨੂੰ ਵਾਕਈ ਵਿਸ਼ਵ-ਵਿਆਪੀ ਬਨਾਉਣਾ !

缔造真正全球通行的万维网

ワールド・ワイド・ウェッブを世界中に広げましょう

ធ្វើឲ្យវើលវ៉ាយវ៉ែបមានទូទាំងពិភពលោកពិត្រាកដមែន!

전세계의 월드 와이드 웹으로 만들기!

Gwneud y we fyd-eang yn wirioneddol fyd-eang!

ליצור מהרשת רשת כלל עולמית באמת!

การทำให้ World Wide Web แพร่หลายไปทั่วโลกอย่างแท้จริง

འཇམ་གྱིང་ཡོངས་འབྲེལ་འདི་ དངོས་འབད་རང་ འཇམ་གྱིང་ཡོངས་ལུ་ཁྱབ་རྒྱུགས་པ་བཟོ་བ།

Recent developments

Recent developments

Unicode

جعل شبكة الويب العالميّة عالميّة حقًّا!

وب جہانی را بہ درستی جہانی سازیم!

عالمگیر ویب کو حقیقی طور پر عالمگیر بنانا

Համաշխարհային ցանցն իրոք համաշխարհային դարձնելը

ᑖᑦᓱᒪ ᐃᑭᐊᖅᑭᕕᒃ ᓯᓚᕐᔪᐊᓕᒫᒥᒃ ᓈᕆᑎᑉᐸ.

"Дүниежүзілік торды" нағыз дүниежүзілік етеміз!

वर्ल्ड वाईड वेबलाई यथार्थमै विश्वव्यापी बनाउने !

የዓለም አቀፉን ድር በእውነት አለም አቀፍ ማድረግ!

Κάνοντας τον Παγκόσμιο Ιστό πραγματικά Παγκόσμιο

ਵਰਡ ਵਾਈਡ ਵੈਬ ਨੂੰ ਵਾਕਈ ਵਿਸ਼ਵ-ਵਿਆਪੀ ਬਨਾਉਣਾ !

缔造真正全球通行的万维网

ליצור מהרשת רשת כלל עולמית באמת!

ˈmeɪkɪŋ ðə wɜːld waɪd wɛb ˈtruːlɪ ˈwɜːldˈwaɪd

ワールド・ワイド・ウェッブを世界中に広げましょう

ធ្វើឲ្យវើលវ៉ាយវ៉េបមានទូទាំងពិភពលោកពិត្រាកដមែន!

전세계의 월드 와이드 웹으로 만들기!

Gwneud y we fyd-eang yn wirioneddol fyd-eang!

การทำให้ World Wide Web แพร่หลายไปทั่วโลกอย่างแท้จริง

འཇཱམ་གླིང་ཡོངས་འབྲེལ་འདི་ ངོ་མ་འབད་རང་ འཇཱམ་གླིང་ཡོངས་ལུ་ཁྱབ་ཚུགས་པ་བཟོ་བ།

Recent developments

Unicode

جعل شبكة الويب العالميّة عالميّة حقًّا!

وب جہانی را بہ درستی جہانی سازیم!

عالمگیر ویب کو حقیقی طور پر عالمگیر بنانا

Համաշխարհային ցանցն իրոք համաշխարհային դարձնելը

ᑖᑦᓱᒪ ᐃᑭᐊᖅᑭᕕᒃ ᓱᓚᕐᔪᐊᓕᒫᒥ

"Дүниежүзілік торды" нағыз

वर्ल्ड वाईड वेबलाई यथार्थमै

የዓለም አቀፉን ድር በእውነት አለ

Κάνοντας τον Παγκόσμιο Ιστό πραγματικά Παγκόσμιο

ਵਰਡ ਵਾਈਡ ਵੈਬ ਨੂੰ ਵਾਕਈ ਵਿਸ਼ਵ-ਵਿਆਪੀ ਬਨਾਉਣਾ !

缔造真正全球通行的万维网

ליצור מהרשת רשת כלל עולמית באמת!

ˈmeɪkɪŋ ðə wɜːld waɪd wɛb ˈtruːlɪ ˈwɜːldˈwaɪd

ワールド・ワイド・ウェッブを世界中に広げましょう

ធ្វើឲ្យវ៉ែលវ៉ាយវ៉េបមានទូទាំងពិភពលោកពិត្រាកដមែន!

전세계의 월드 와이드 웹으로 만들기!

Gwneud y we fyd-eang yn wirioneddol fyd-eang!

การทำให้ World Wide Web แพร่หลายไปทั่วโลกอย่างแท้จริง

འཇམ་གླིང་ཡོངས་འབྲེལ་འདི་ དངོས་འབད་རང་ འཇམ་གླིང་ཡོངས་ལུ་ཁྱབ་ཚུགས་པ་བཟོ་བ།

"The Path W3C follows to making text on the Web truly global is Unicode."

Tim Berners-Lee

Unicode

<h2>რჩეული ფოტოსურათი</h1>

<p>ჭიამაია (Coccinellidae), ხოჭოების ოჯახს ეკუთვნის. აქვს ამოზურცული, მომრგვალო ან ოვალური სხეული. ზურგზე ღია ფონზე შავი ლაქები აყრია, იშვიათად

...

Recent developments

Unicode

<h2>
ფოტოსურათი</h

<p><a href="/wiki/ჭიამ
class="mw-redirect">ჭია
ხოჭოების ოჯახს ეკუ
მომრგვალო ან ოვალუ
ფონზე შავი ლაქები ა
...

W3C Recommendation

W3C

Extensible Markup Language (XML) 1.0 (Fifth Edition)

W3C Recommendation 26 November 2008

This version:
http://www.w3.org/TR/2008/REC-xml-20081126/
Latest version:
http://www.w3.org/TR/xml/
Previous versions:
http://www.w3.org/TR/2008/PER-xml-20080205/
http://www.w3.org/TR/2006/REC-xml-20060816/
Editors:
Tim Bray, Textuality and Netscape <tbray@textuality.com>
Jean Paoli, Microsoft <jeanpa@microsoft.com>
C. M. Sperberg-McQueen, W3C <cmsmcq@w3.org>
Eve Maler, Sun Microsystems, Inc. <eve.maler@east.sun.com>
François Yergeau

Please refer to the **errata** for this document, which may include some normative corrections.

The previous errata for this document, are also available.

See also **translations**.

This document is also available in these non-normative formats: XML and XHTML with color-coded revision indicators.

Copyright © 2008 W3C® (MIT, ERCIM, Keio), All Rights Reserved. W3C liability, trademark and document use rules apply.

Abstract

The Extensible Markup Language (XML) is a subset of SGML that is completely described in this document. Its goal is to enable generic SGML to be served, received, and processed on the Web in the way that is now possible with HTML. XML has been designed for ease of implementation and for interoperability with both SGML and HTML.

Status of this Document

This section describes the status of this document at the time of its publication. Other documents may supersede this document. A list of current W3C publications and the latest revision of this technical report can be found in the W3C technical reports index at http://www.w3.org/TR/.

This document specifies a syntax created by subsetting an existing, widely used international text processing standard (Standard Generalized Markup Language, ISO 8879:1986(E) as amended and corrected) for use on the World Wide

Normalization

NFD

NFC

Ízelítőül

Ha a világ beszélni akarna, Unicode-ul szólalna meg. Regisztráljon már most a Tizedik Nemzetközi Unicode Konferenciára, melyet 1997. március 10-12-én rendeznek Meinz-ban, Németországban. Ezen a konferencián az iparág több neves szakértője is résztvesz. Ízelítőül a témákból: a világháló és a Unicode nemzetközisítése és lokalizálása, a Unicode alkalmazása működő rendszerekben és alkalmazásokban, szövegelrendezésnél, és többnyelvű számítógépeken.

Web resource identifiers

http://JP納豆.例.jp/dir1/引き割り.html

Scheme

Domain name

Path

IDN

xn--jp-cd2fp15c.xn--fsq.jp

Web resource identifiers

السعودية
امارات
مصر

Al-Saudiah
Emarat
Misr

http://وزارة-الأتصالات.مصر

IDN

Web resource identifiers

http://JP納豆.例.jp/dir1/引き割り.html

Scheme

Domain name

Path

IRI

/dir1/%E5%BC%95%E3%81%8D%E5%89%B2%E3%82%8A.html

Language tags

Before: RFC 3066

language – region

- ISO 639 language codes
- ISO 3166 country codes

en

en-GB

en-scouse

Language tags: BCP 47

Now: BCP 47 (includes RFC 5646)

language (extlang) script region variant extension private_use

hi

az-Cyrl

zh-Hans

es-419

sl-IT-rozaj-njiva-1994

- nearly 8,000 subtags available
- subtags available only from new IANA registry (based on ISO and UN codes)
- only language subtag required

Recent developments

Bidirectional text support

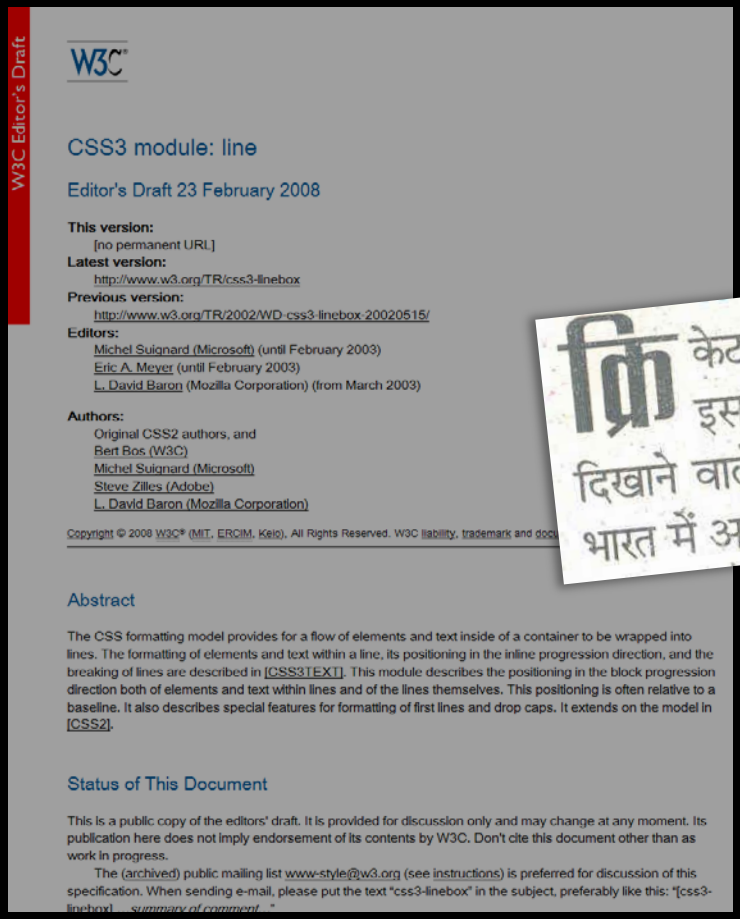
W3C نشاط التدويل،

W3C نشاط التدويل،

```
<description dir="rtl">W3C نشاط التدويل،</description>
```

Recent developments

CSS3



W3C Editor's Draft

W3C

CSS3 module: line

Editor's Draft 23 February 2008

This version:
[no permanent URL]
Latest version:
http://www.w3.org/TR/css3-linebox
Previous version:
http://www.w3.org/TR/2002/WD-css3-linebox-20020515/
Editors:
Michel Suignard (Microsoft) (until February 2003)
Eric A. Meyer (until February 2003)
L. David Baron (Mozilla Corporation) (from March 2003)

Authors:
Original CSS2 authors, and
Bert Bos (W3C)
Michel Suignard (Microsoft)
Steve Zilles (Adobe)
L. David Baron (Mozilla Corporation)

Copyright © 2008 W3C® (MIT, ERCIM, Keio). All Rights Reserved. W3C liability, trademark and docu

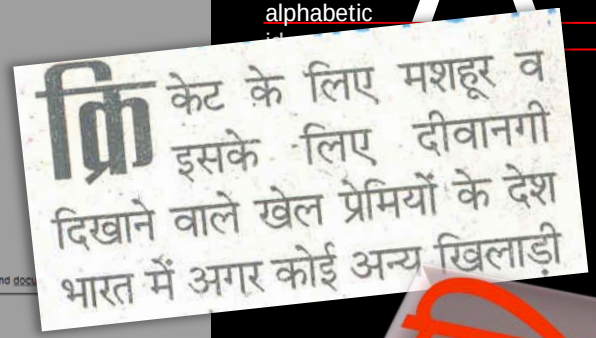
Abstract

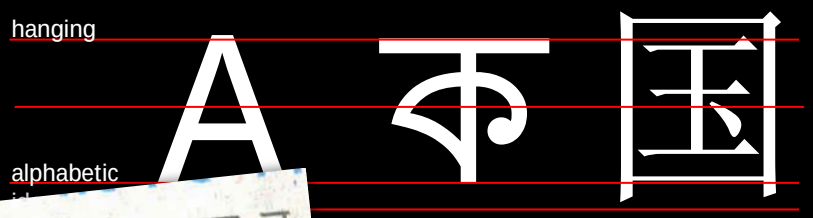
The CSS formatting model provides for a flow of elements and text inside of a container to be wrapped into lines. The formatting of elements and text within a line, its positioning in the inline progression direction, and the breaking of lines are described in [CSS3TEXT]. This module describes the positioning in the block progression direction both of elements and text within lines and of the lines themselves. This positioning is often relative to a baseline. It also describes special features for formatting of first lines and drop caps. It extends on the model in [CSS2].

Status of This Document

This is a public copy of the editors' draft. It is provided for discussion only and may change at any moment. Its publication here does not imply endorsement of its contents by W3C. Don't cite this document other than as work in progress.

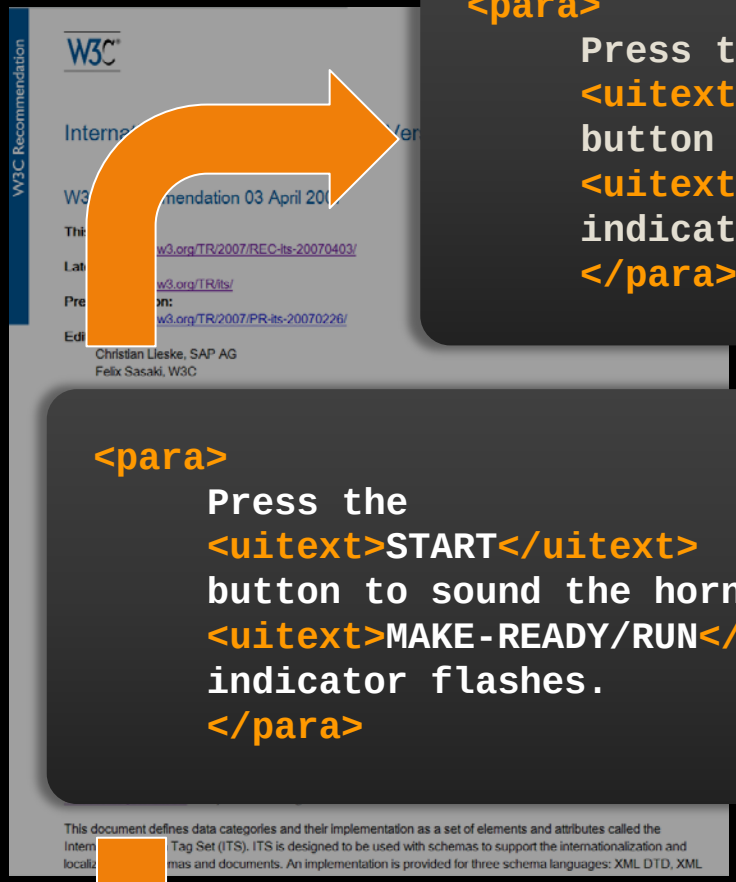
The (archived) public mailing list www-style@w3.org (see instructions) is preferred for discussion of this specification. When sending e-mail, please put the text "css3-linebox" in the subject, preferably like this: "[css3-linebox] ... summary of comment ... "





Implementers of user agents need to be prodded by the public to support the developing marketplace !

Internationalization Tag Set




```
<para>
      Press the
      <uitext translate="no">START</uitext>
      button to sound the horn. The
      <uitext translate="no">MAKE-READY/RUN</uitext>
      indicator flashes.
      </para>
```

```
<para>
      Press the
      <uitext>START</uitext>
      button to sound the horn. The
      <uitext>MAKE-READY/RUN</uitext>
      indicator flashes.
      </para>
```

- supported by some translation tools – linked with XLIFF
- being applied by specifications at W3C

```
<its:rules ... its:version="1.0">
    <its:translateRule selector="//uitext" translate="no"/>
    </its:rules>
```

Requirements

Developing requirements

International typography

Developing requirements

International typography

Arabic mathematics

W3C Interest Group Note

W3C

Arabic mathematical notation

W3C Interest Group Note 31 January 2006

This version:
http://www.w3.org/TR/2006/NOTE-arabic-math-2006

Latest version:
http://www.w3.org/TR/arabic-math/

Previous version:
This is the first version

Editors:
Azzeddine Lazrek, with Mustapha Eddahibi and Kha
Morocco
Bruce R. Miller, National Institute of Standards and T

This document is also available in these non-normative for

Copyright ©2006 W3C® (MIT, ERCIM, Keio), All Rights Reserved. W3

$$f(x) = \begin{cases} \sum_{i=1}^{s} x^i & \text{if } x < 0 \\ \int_1^s x^i \, dx & \text{if } x \in S \\ \tan \pi & \text{otherwise (with } \pi \simeq 3.141) \end{cases}$$

$$ت(س) = \begin{cases} \text{مجـ}_{ب=١}^{ص} س^{ب} & \text{إذا كان } س < ٠ \\ \int_{١}^{ص} س^{ب} \, ء س & \text{إذا كان } س \in م \\ \text{ظا } \pi & \text{غير ذلك (مع } \pi \simeq ٣٫١٤١) \end{cases}$$

Abstract

This Note analyzes potential problems with the use of MathML for the presentation of mathematics in the notations customarily used with Arabic, and related languages. The goal is to clarify avoidable implementation details that hinder such presentation, as well as to uncover genuine limitations i
limitations in the MathML specification may require extensions in future versions

Status of this Document

This section describes the status of this document at the time of its publication. supersede this document. A list of current W3C publications and the latest revis be found in the W3C technical reports index at http://www.w3.org/TR/.

This Note is a self-contained discussion of Arabic mathematical notation in Math the handling of Arabic mathematical presentation using MathML 2 Recommend and suggests extensions for a future revision.

This Note has been written by participants in the Math Interest Group (W3C me W3C Math activity. Please direct comments and report errors in this document t list with a public archive.

Publication as a Interest Group Note does not imply endorsement by the W3C document and may be updated, replaced or obsoleted by other documents at any time. It is inappropriate to cite this document as other than work in progress.

Developing requirements

Speech Synthesis Markup Language

W3C Proposed Recommendation

W3C

Speech Synthesis Markup Lan

W3C Proposed Recommendation 23 F

This version:
http://www.w3.org/TR/2010/PR-speech-synthesis11-20100223/
Latest version:
http://www.w3.org/TR/speech-synthesis11/
Previous version:
http://www.w3.org/TR/2009/CR-speech-synthesis11-20090827/

Editors:
Daniel C. Burnett, Voxeo (formerly of Vocalocity and Nuance)
双志伟 (Zhi Wei Shuang), IBM
Authors:
Paolo Baggia, Loquendo
Paul Bagshaw, France Telecom
Michael Bodell, Microsoft
胡方 (Hu Fang), Invited Expert (until 20 October 2009)
黄德智 (De Zhi Huang), France Telecom
黄力行 (Lixing Huang), Chinese Academy of Sciences
康永国 (Yongguo Kang), Panasonic Corporation (until 5 December 2007)
楼晓雁 (Lou Xiaoyan), Toshiba
Scott McGlashan, HP
蒙美玲 (Helen Meng), Chinese University of Hong Kong (until 29 July 2009)
陶建华 (Jianhua Tao), Chinese Academy of Sciences
王霞 (Wang Xia), Nokia (until 30 October 2006)
夏海荣 (Xia Hairong), Panasonic Corporation (until 2 August 2006)
吴志勇 (Zhiyong Wu), Chinese University of Hong Kong (until 29 July 2009)
严峻 (Yan Jun), iFLYTEK

Copyright © 2010 W3C® (MIT, ERCIM, Keio), All Rights Reserved. W3C liability, trademark and document use rules a

Abstract

The Voice Browser Working Group has sought to develop standards to enable access to the Web using spoken interaction. The Speech Synthesis Markup Language Specification is one of these standards and is designed to provide a rich, XML-based markup language for assisting the generation of synthetic speech in Web and other applications. The essential role of the markup language is to provide authors of synthesizable content a standard way to control aspects of speech such as pronunciation, volume, pitch, rate, etc. across different synthesis-capable platforms.

這一晚會如常舉行

這一|晚會|如常|舉行

這一|晚會|如|常|舉行

這一晚|會|如常|舉行

Developing requirements

Speech Synthesis Markup Language

W3C Proposed Recommendation

W3C

Speech Synthesis Markup Language (SSML) Version 1.1

W3C Proposed Recommendation 23 February 2010

This version:
http://www.w3.org/TR/2010/PR-speech-synthesis11-20100223/
Latest version:
http://www.w3.org/TR/speech-synthesis11/
Previous version:
http://www.w3.org/TR/2009/CR-speech-synthesis11-20090827/

Editors:
Daniel C. Burnett, Voxeo (formerly of Vocalocity and Nuance)
双志伟 (Zhi Wei Shuang), IBM
Authors:
Paolo Baggia, Loquendo
Paul Bagshaw, France Telecom
Michael Bodell, Microsoft
胡方 (Hu Fang), Invited Expert (until 20 October 2009)
黄德智 (De Zhi Huang), France Telecom
黄力行 (Lixing Huang), Chinese Academy of Sciences
康永国 (Yongguo Kang), Panasonic Corporation (until 5 December 2007)
楼晓雁 (Lou Xiaoyan), Toshiba
Scott McGlashan, HP
蒙美玲 (Helen Meng), Chinese University of Hong Kong (until 29 July 2009)
陶建华 (Jianhua Tao), Chinese Academy of Sciences
王霞 (Wang Xia), Nokia (until 30 October 2006)
夏海荣 (Xia Hairong), Panasonic Corporation (until 2 August 2006)
吴志勇 (Zhiyong Wu), Chinese University of Hong Kong (until 29 July 2009)
严峻 (Yan Jun), iFLYTEK

Copyright © 2010 W3C® (MIT, ERCIM, Keio), All Rights Reserved. W3C liability, trademark and document use rules apply.

Abstract

The Voice Browser Working Group has sought to develop standards to enable access to the Web using spoken interaction. The Speech Synthesis Markup Language Specification is one of these standards and is designed to provide a rich, XML-based markup language for assisting the generation of synthetic speech in Web and other applications. The essential role of the markup language is to provide authors of synthesizable content a standard way to control aspects of speech such as pronunciation, volume, pitch, rate, etc. across different synthesis-capable platforms.

Workshop on Conversational Applications

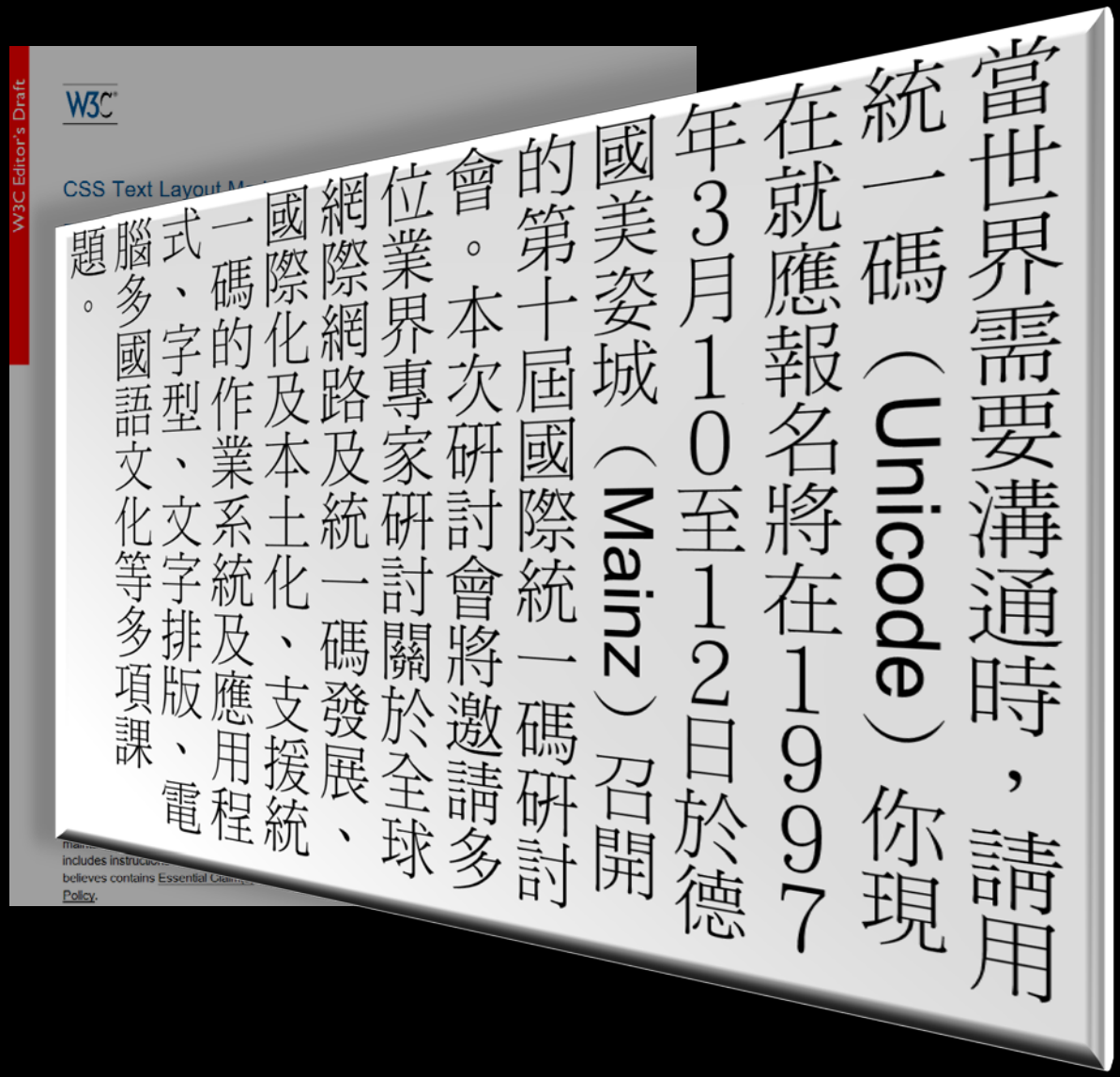
Use Cases and Requirements for New Models of Human Language to Support Mobile Conversational Systems

June 18-19, NJ, USA

Current discussions

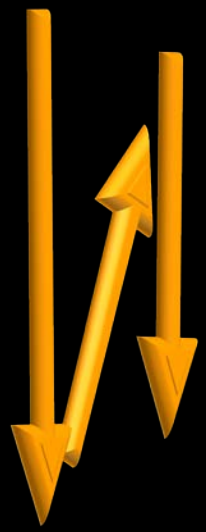
Current discussions

Vertical text



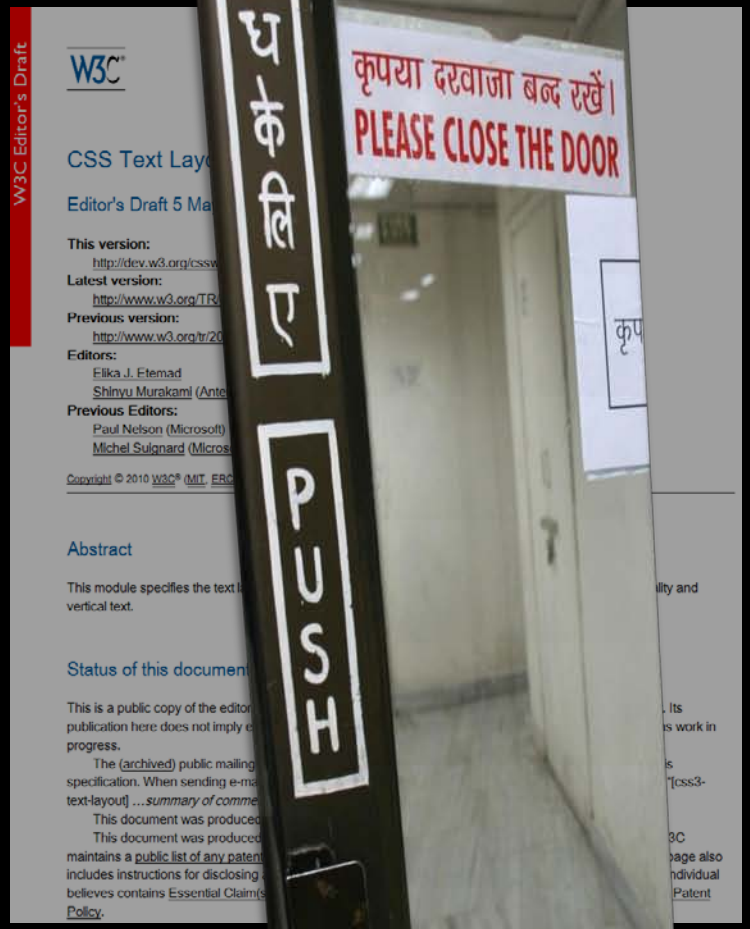
Current discussions

Vertical text



Current discussions

Vertical text



日 34개地自體 철폐 검토

외국인 공무원임용제한 國籍조항

日 월드컵 유치신청 15개도시

「공동개최」불구 경기개최 추진

Current discussions

Web fonts

തൊണ്ണൂറ

ជាតូបនីយកម្ម

کو حقیقی طور پر عالمگیر بنانا

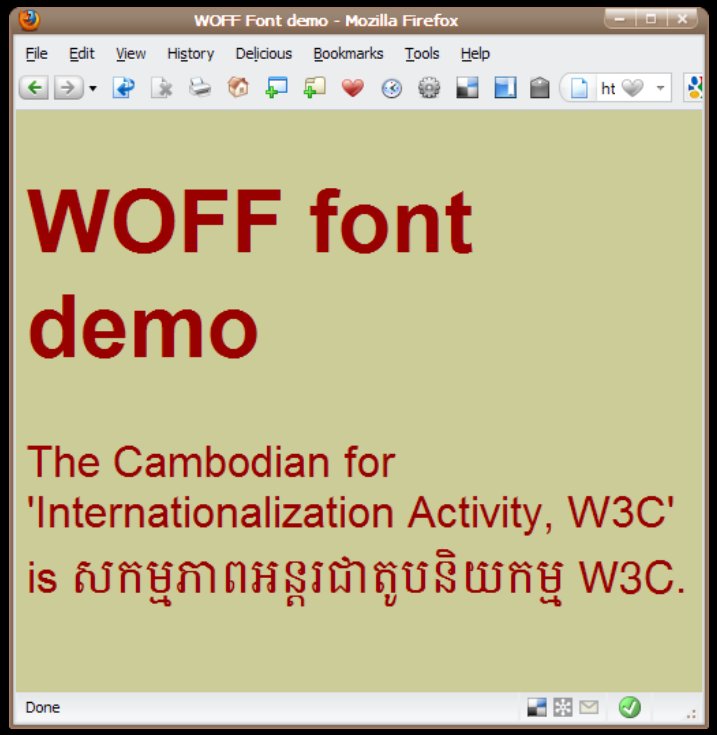
ལ་སྤྱིར་བསྒྱུར་བའི་ལས་དོན།

Current discussions

Web fonts

```
@font-face {
  font-family: 'battambang-woff';
  font-style: normal;
  font-weight: normal;
  src: url(fonts/khmerosbbang.woff);
  }

:lang(kh) {
  font-family: 'battambang-woff';
  font-size: 100%;
  }
```

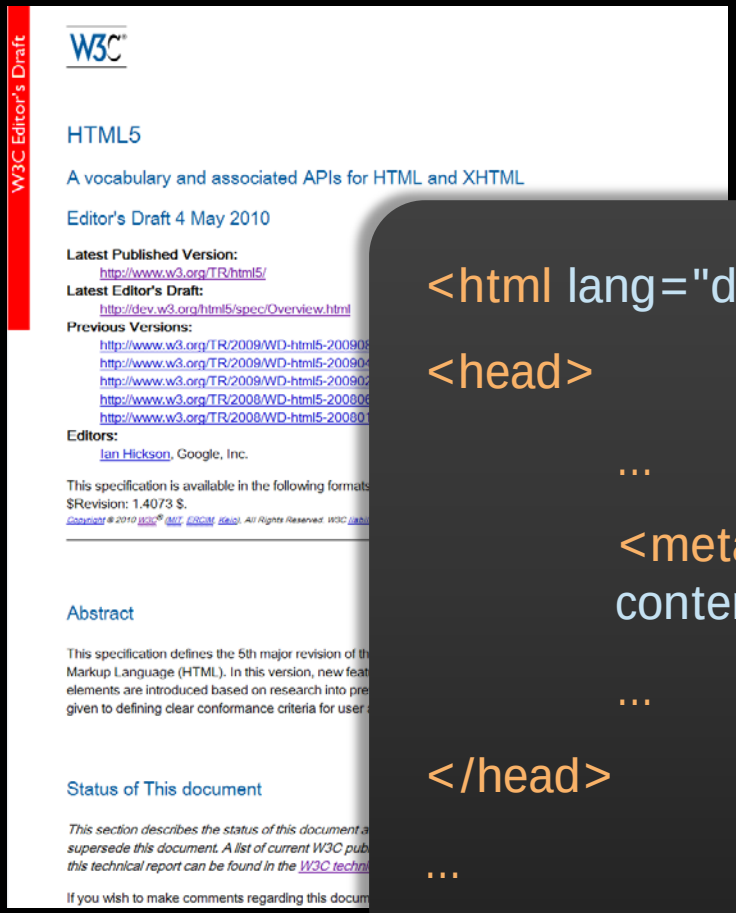


Issues

- Rendering detail for complex fonts.
- Subsetting capability may be needed.
- Can only be used for fonts with an appropriate licence.

Current discussions

Language declarations

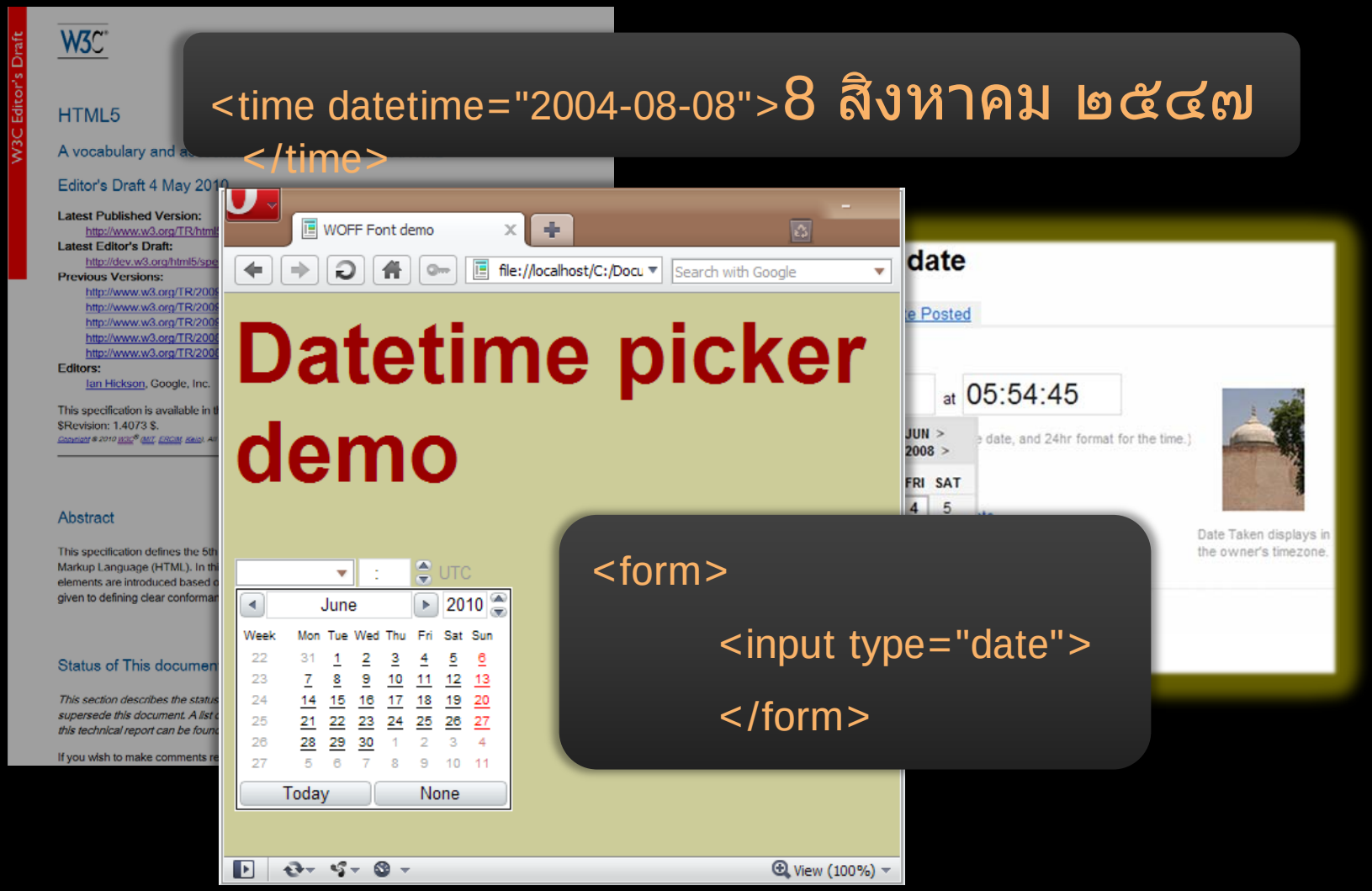



```
<html lang="de">
<head>
        ...
        <meta http-equiv="Content-Language"
        content="de">
        ...
</head>
...
```

Current discussions

Date and time

```
<time datetime="2004-08-08">8 สิงหาคม ๒๕๔๗
</time>
```




```
<form>
    <input type="date">
</form>
```

Current discussions

Ruby annotation

W3C Recommendation

W3C

Ruby An

W3C Recor
2008)

This version:
http://www
(ZIP archi
Latest version
http://www
Previous vers
http://www
Editors:
Marcin SA
Michel SU
Masayasu
Martin DU
Tex TEXI
(See Ackr

Copyright ©1998-2
licensing rules appl

Abstract

"Ruby" are shor
pronunciation o
an XHTML mod

Status of Th

This section des
supersede this

This document
endorsed by the
reference mater
Recommendati
This enhances

W3C Editor's Draft

W3C

CSS3 R

W3C Worl

This version:
http://de
Latest version
http://ww
Previous versi
http://ww
Editors:
Richard
Former editors
Paul Ne
Michel S
Marcin S

Copyright © 201
licensing rules ap

Abstract

"Ruby" are sh
pronunciation
associated wi
[RUBY].

Status of

See open issu

This docume

This section d
supersede thi
can be found

Publication as
document and
inappropriate

W3C Editor's Draft

W3C

HTML5

A vocabulary and associated APIs for HTML and XHTML

Editor's Draft 4 May 2010

Latest Published Version:
http://www.w3.org/TR/html5/
Latest Editor's Draft:
http://dev.w3.org/html5/spec/Overview.html
Previous Versions:
http://www.w3.org/TR/2009/WD-html5-20090825/
http://www.w3.org/TR/2009/WD-html5-20090423/
http://www.w3.org/TR/2009/WD-html5-20090212/
http://www.w3.org/TR/2008/WD-html5-20080610/
http://www.w3.org/TR/2008/WD-html5-20080122/
Editors:
Ian Hickson, Google, Inc.

This specification is available in the following formats: single page HTML, multipage HTML. This is revision $Revision: 1.4073 $.

Copyright © 2010 W3C® (MIT, ERCIM, Keio). All Rights Reserved. W3C liability, trademark and document use rules apply.

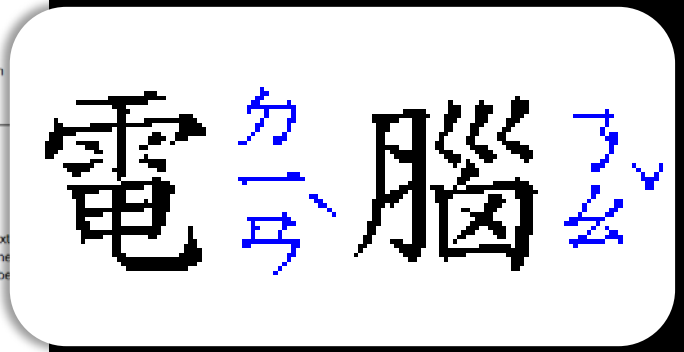
Abstract

This specification defines the 5th major revision of the core language of the World Wide Web: the Hypertext Markup Language (HTML). In this version, new features are introduced to help Web application authors, ne elements are introduced based on research into prevailing authoring practices, and special attention has be given to defining clear conformance criteria for user agents in an effort to improve interoperability.

Status of This document

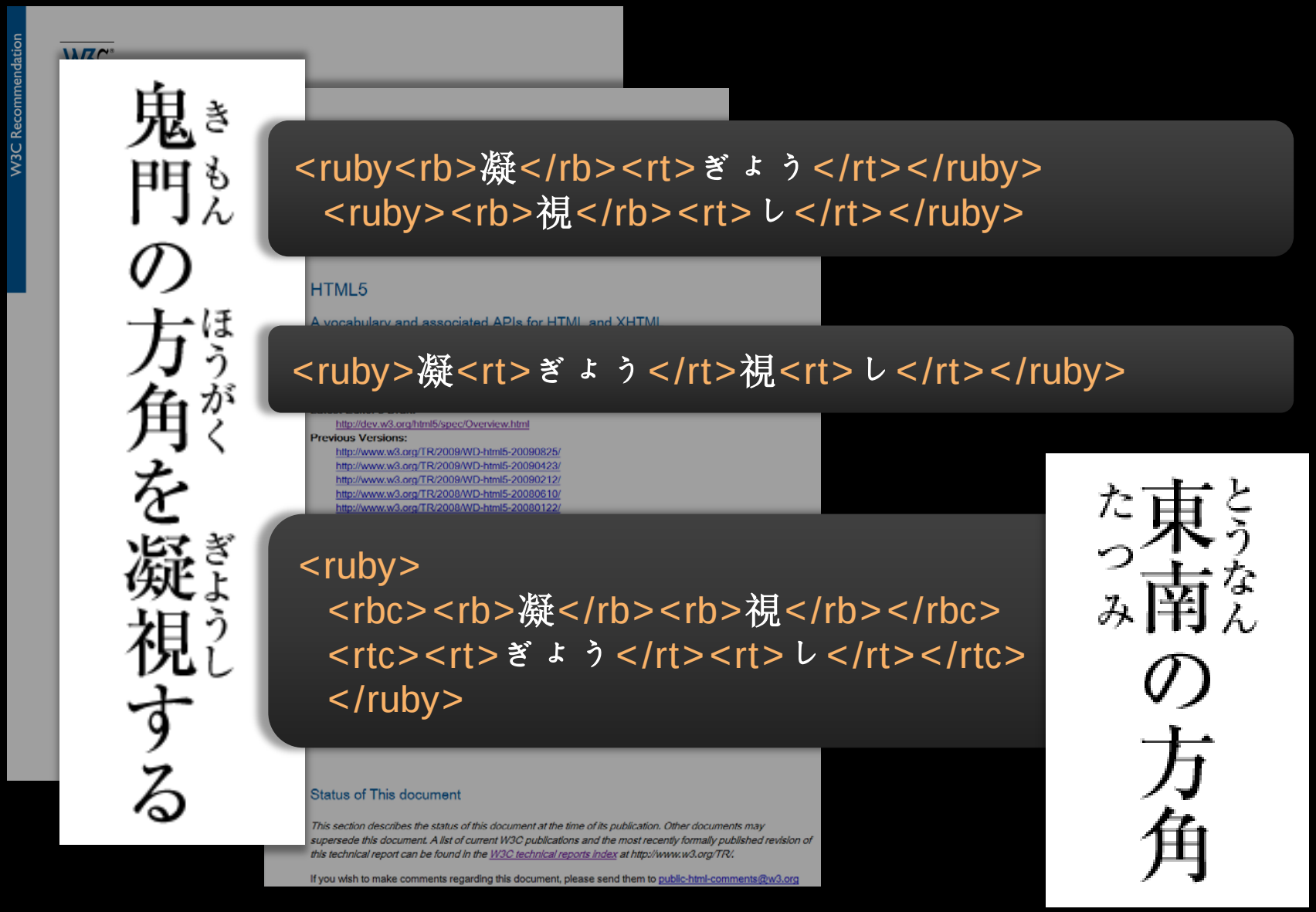
This section describes the status of this document at the time of its publication. Other documents may supersede this document. A list of current W3C publications and the most recently formally published revision of this technical report can be found in the W3C technical reports index at http://www.w3.org/TR/.

If you wish to make comments regarding this document, please send them to public-html-comments@w3.org



Current discussions

Ruby annotation




```
<ruby<rb>凝</rb><rt>ぎょう</rt></ruby>
  <ruby><rb>視</rb><rt>し</rt></ruby>
```

```
<ruby>凝<rt>ぎょう</rt>視<rt>し</rt></ruby>
```

```
<ruby>
  <rbc><rb>凝</rb><rb>視</rb></rbc>
  <rtc><rt>ぎょう</rt><rt>し</rt></rtc>
  </ruby>
```

Current discussions

Key Events

W3C Editor's Draft

W3C

Document Object Model (DOM) Level 3 Ever
Specification

W3C Editor's Draft 21 April 2010

This version:
http://dev.w3.org/2006/webapi/DOM-Level-3-Events/html/DOM3-Events.html?rev=1.131
Latest stable version:
http://www.w3.org/TR/DOM-Level-3-Events
Previous version:
http://dev.w3.org/cvsweb/~checkout~/2006/webapi/DOM-Level-3-Events/html/DOM3-Events.html?rev=1.130
Editor's Draft:
http://dev.w3.org/2006/webapi/DOM-Level-3-Events/html/DOM3-Events.

Editors:
Doug Schepers, *W3C*
Björn Höhrmann, *Invited Expert (until December 2007)*
Philippe Le Hégaret, *W3C (until November 2003)*
Tom Pixley, *Netscape Communications Corporation (until July 2002)*

Copyright © 2009 W3C® (MIT, ERCIM, Keio), All Rights Reserved. W3C liability, trademark an use rules apply.

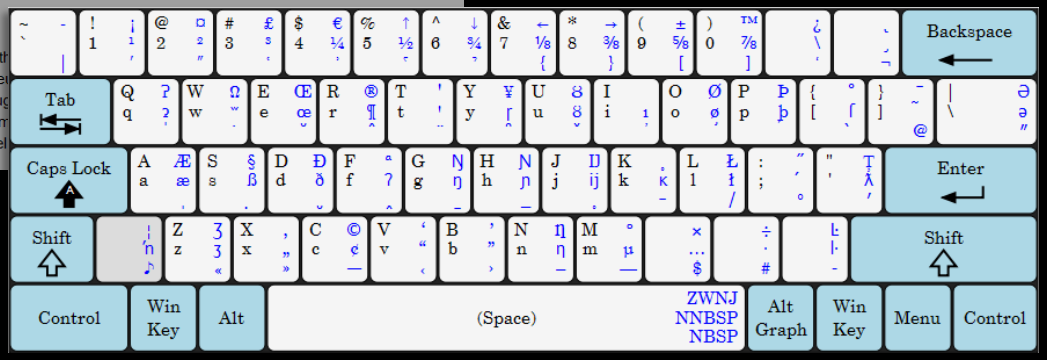
Abstract

This specification defines th
platform- and language-ne
describes event flow throug
for each event. The Docum
Object Model Events Level

SEARCH | MOBILE | ACCESSIBILITY | I18N | TYPOGRAPHY

Internationalization Quicktips

- Use Unicode wherever possible for content, databases, etc. Always declare the encoding of content.
- Use characters rather than escapes (e.g. á á or á) whenever you can.
- Declare the language of documents and indicate internal language changes.



Social context

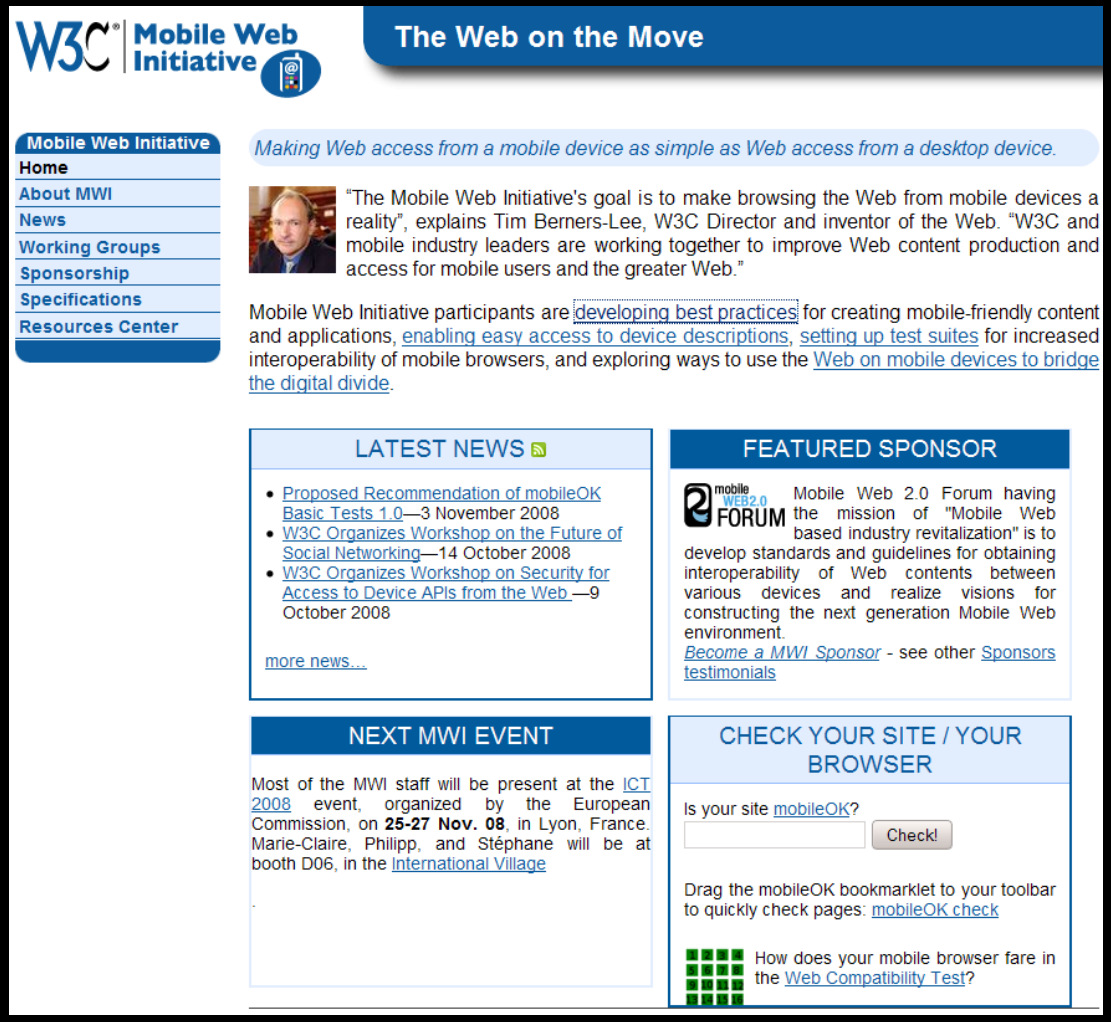
The rise of the Mobile Web

- "In China … over 73m people, or 29% of all internet users in the country, use mobile phones to get online."
- "The number of pages viewed in June by 14m users of [Opera] software was over 3 billion, a 300% increase on a year earlier. The fastest growth was in developing countries including Russia, Indonesia, India and South Africa."

Economist.com, Sept. 2008

The social context

Mobile Web Initiative



Best practices

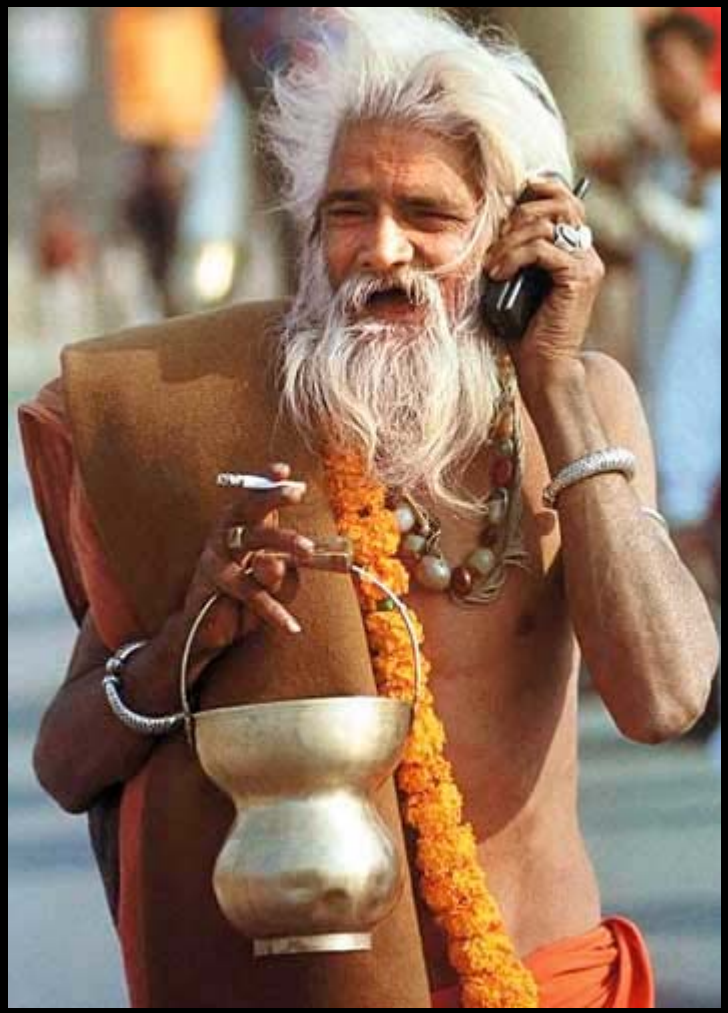
Device descriptions

Test suites

mobileOK checker

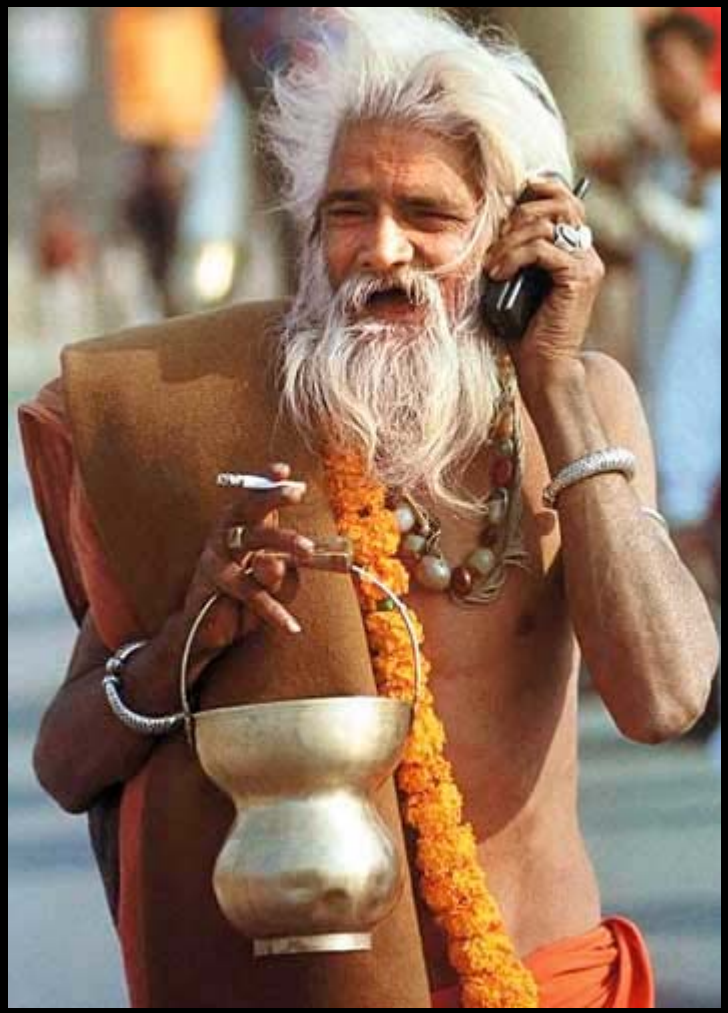
The social context

Mobile Web for Developing Society (MW4D)



Track the social impact of the mobile web in the developing world, to ensure that the web's technical standards evolve to serve this rapidly emerging constituency.

MW4D expected outcomes

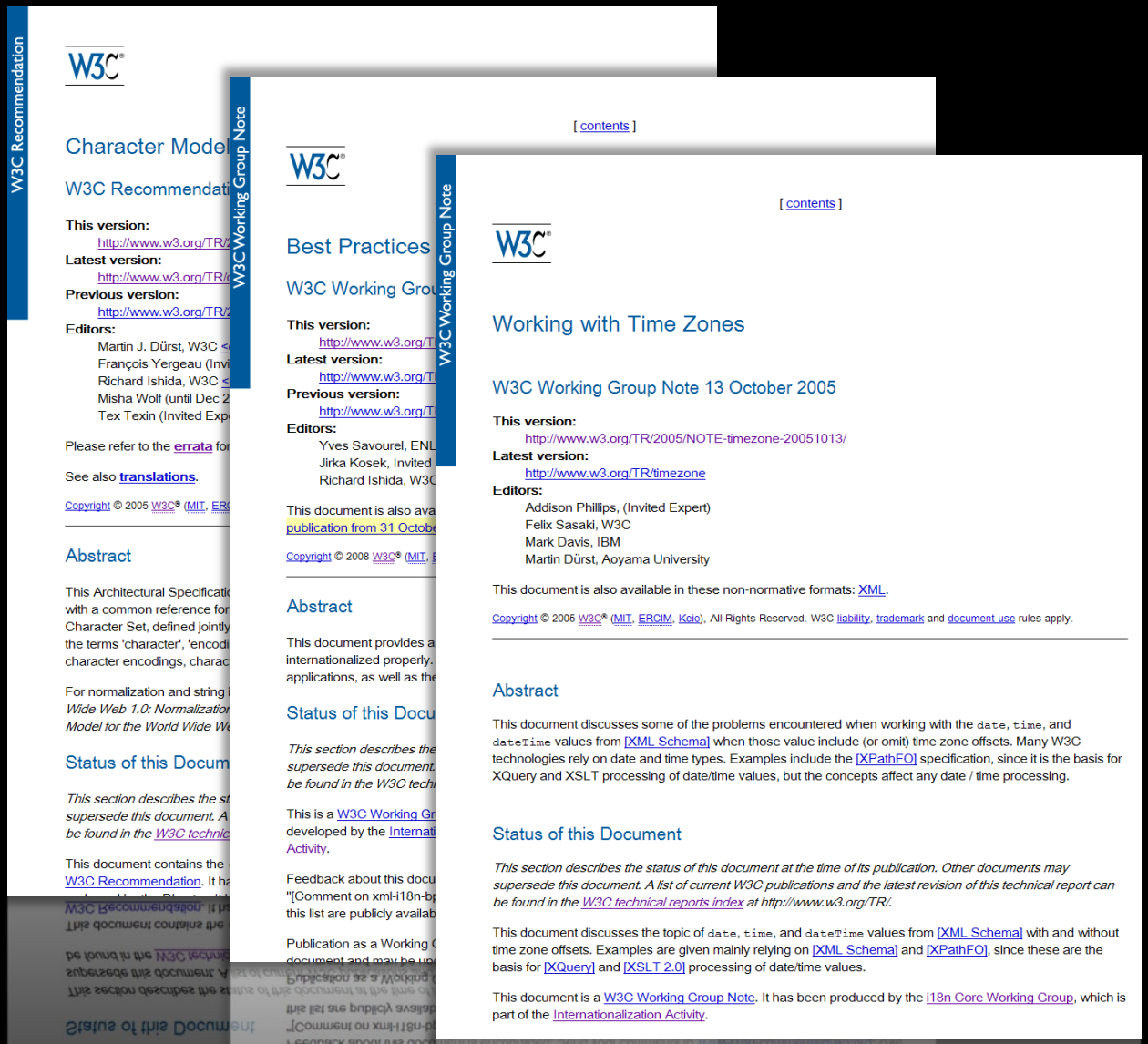


- A Handbook about how to deploy social-oriented content and applications on mobile
- A roadmap identifying the key actions to launch to lower the barriers for deploying social services, and for accessing it
- A set of resources

Best practices

Best practices

Capturing guidance for spec developers



markup for bidirectional text

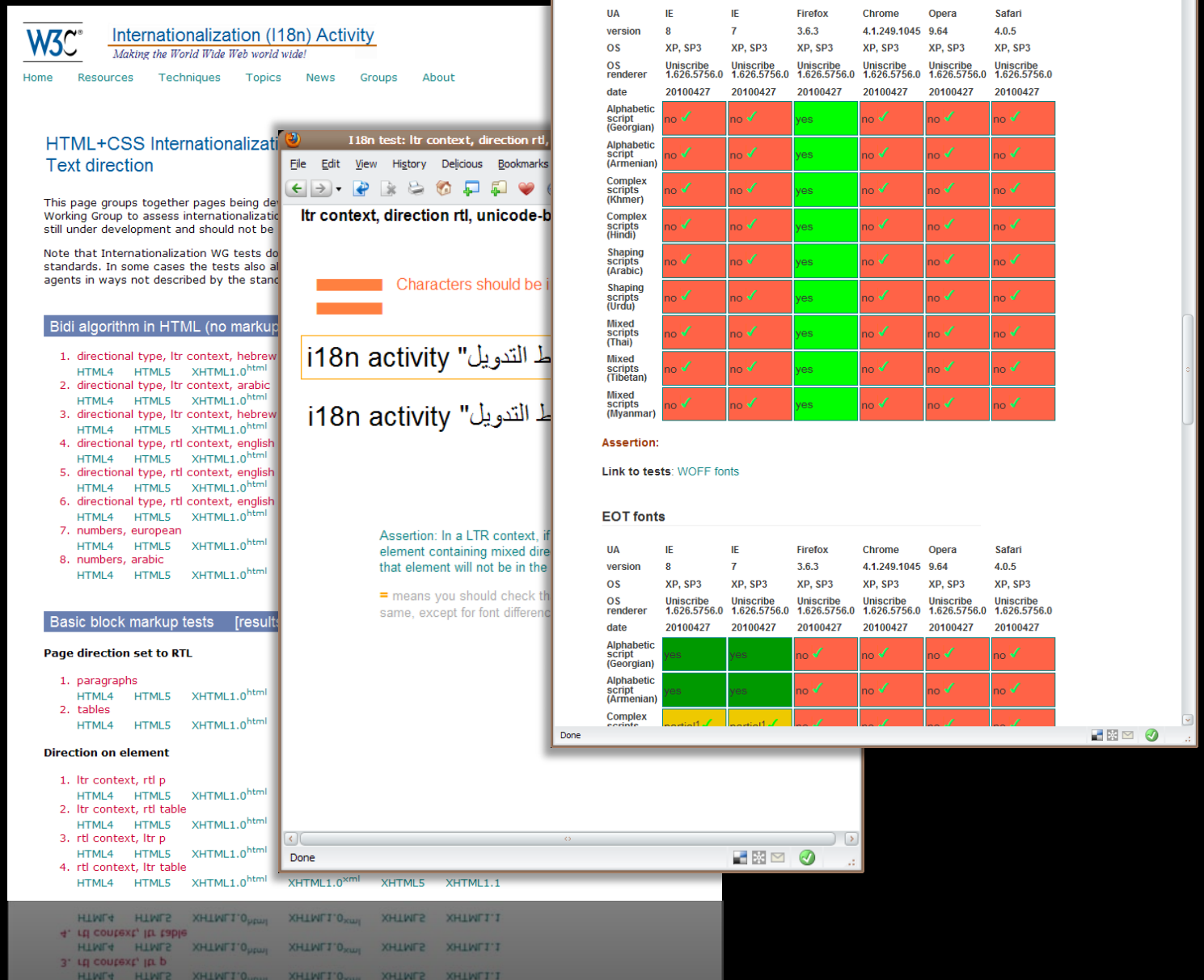
Normalization

working with case sensitivity

more information about date & time

Best practices

Tests



Best practices

I18n resources

Articles

Tutorials

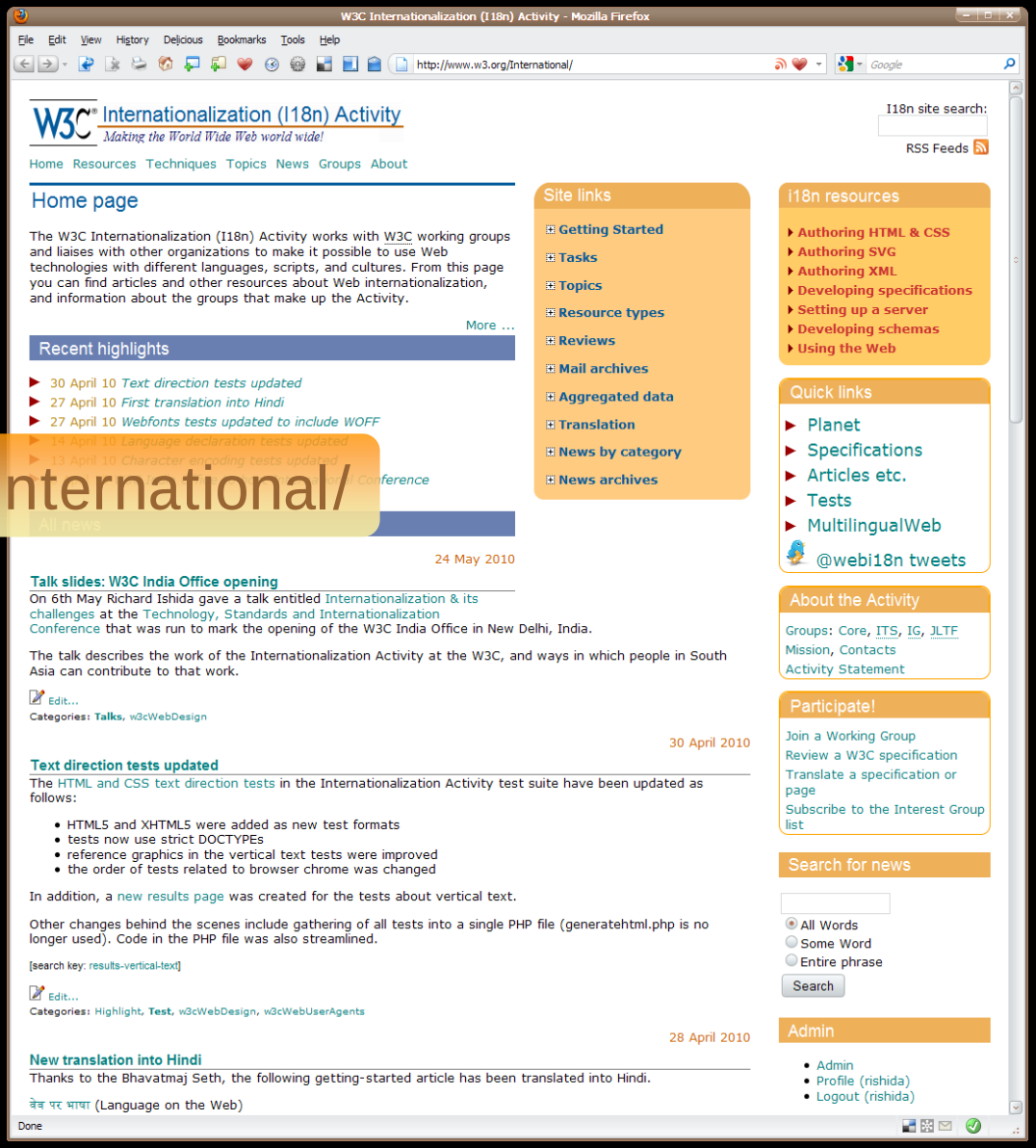
Tests

Talks

Tools

Reviews

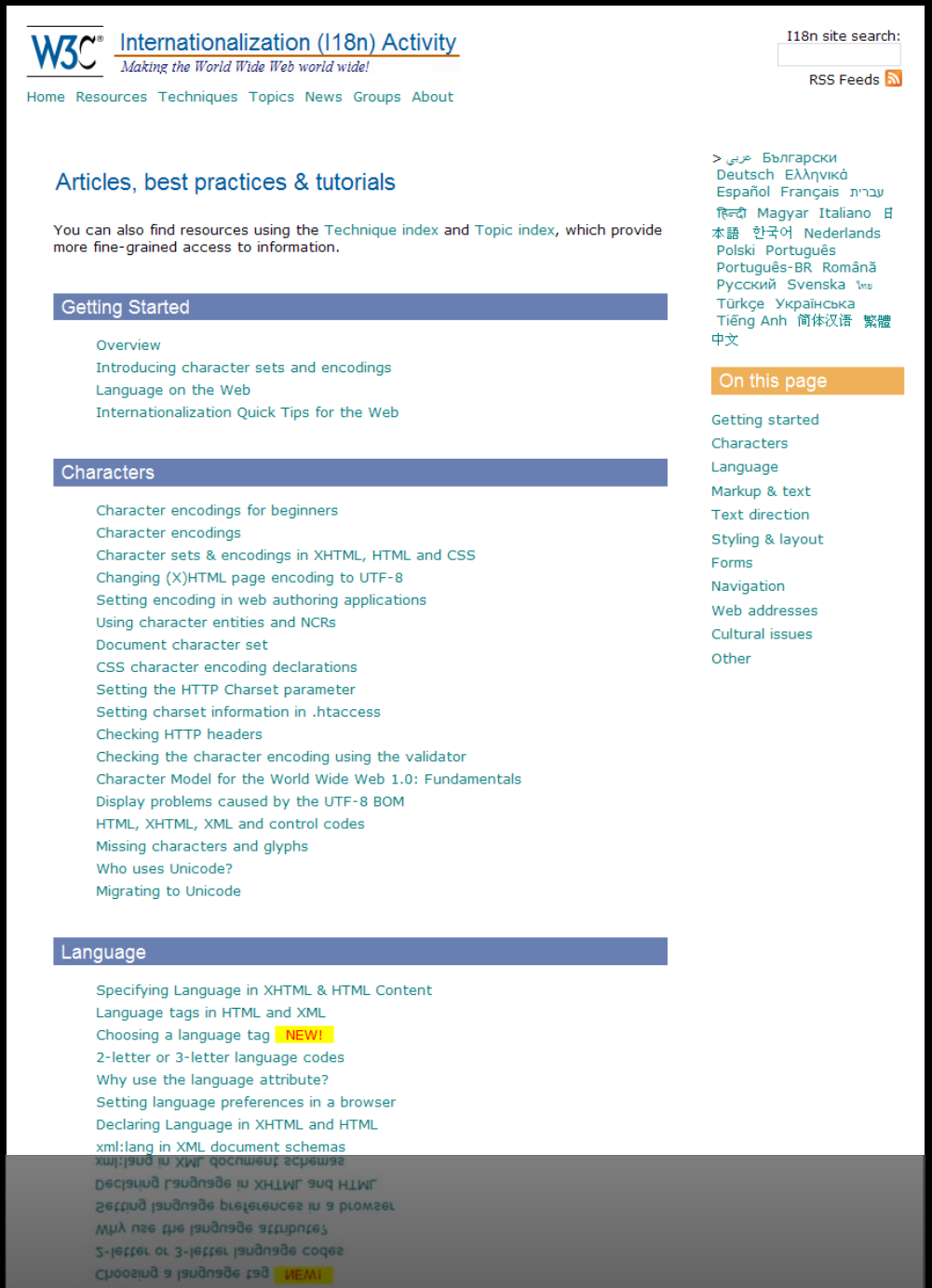
http://www.w3.org/International/



Best practices

I18n resources

- Articles
- Tutorials
- Technical notes
- Tests
- Talks
- Tools
- Reviews



W3C Internationalization (I18n) Activity

Making the World Wide Web world wide!

I18n site search:

RSS Feeds

Home Resources Techniques Topics News Groups About

Articles, best practices & tutorials

You can also find resources using the Technique index and Topic index, which provide more fine-grained access to information.

Getting Started

Overview
Introducing character sets and encodings
Language on the Web
Internationalization Quick Tips for the Web

Characters

Character encodings for beginners
Character encodings
Character sets & encodings in XHTML, HTML and CSS
Changing (X)HTML page encoding to UTF-8
Setting encoding in web authoring applications
Using character entities and NCRs
Document character set
CSS character encoding declarations
Setting the HTTP Charset parameter
Setting charset information in .htaccess
Checking HTTP headers
Checking the character encoding using the validator
Character Model for the World Wide Web 1.0: Fundamentals
Display problems caused by the UTF-8 BOM
HTML, XHTML, XML and control codes
Missing characters and glyphs
Who uses Unicode?
Migrating to Unicode

Language

Specifying Language in XHTML & HTML Content
Language tags in HTML and XML
Choosing a language tag NEW!
2-letter or 3-letter language codes
Why use the language attribute?
Setting language preferences in a browser
Declaring Language in XHTML and HTML
xml:lang in XML document schemas

> عربي Български
Deutsch Ελληνικά
Español Français עברית
हिन्दी Magyar Italiano 日本語 한국어 Nederlands
Polski Português
Português-BR Română
Русский Svenska ไทย
Türkçe Українська
Tiếng Anh 简体汉语 繁體中文

On this page

Getting started
Characters
Language
Markup & text
Text direction
Styling & layout
Forms
Navigation
Web addresses
Cultural issues
Other

Best practices

I18n resources

W3C I18n Site Index - Mozilla Firefox

File Edit View History Delicious Bookmarks Tools Help

http://www.w3.org/International/technique-index

Internationalization (I18n) Activity

Making the World Wide Web world wide!

Home Resources Techniques Topics News Groups About

I18n site search:

RSS Feeds

W3C I18n technique index

Select the task you want help with from this page to find resources on the W3C Internationalization site among its articles, tutorials, tests and techniques documents.

General task

Select a general task...

- Using the Web
- Creating HTML & CSS
- Authoring SVG
- Authoring XML
- Setting up a server
- Developing specifications
- Developing schemas

Author: Richard Ishida, W3C.

Content last changed 2009-03-12 16:23 GMT

Copyright © 2009 W3C® (MIT, ERCIM, Keio), All Rights Reserved. W3C liability, trademark, document use and software licensing rules apply. Your interactions with this site are in accordance with our public and Member privacy statements.

See also

The topic index organizes links to resources by keywords, rather than tasks (like the index in a book).

The resources by type page lists resources by type (eg. articles, tools, mail archives, etc).

Current status

This index is still a work in progress. It doesn't yet point to all resources on the site. The content will also continually grow and change as resources are added to the site.

http://www.w3.org/International/technique-index?topic=htmlauth

Best practices

I18n resources

W3C I18n Site Index - Mozilla Firefox

File Edit View History Delicious Bookmarks Tools Help

http://www.w3.org/International/technique-index?topic=htmlauth

Internationalization (I18n) Activity

Making the World Wide Web world wide!

Home Resources Techniques Topics News Groups About

I18n site search:

RSS Feeds

W3C I18n technique index

Select the task you want help with from this page to find resources on the W3C Internationalization site among its articles, tutorials, tests and techniques documents.

General task	Creating HTML and CSS

Select a general task...

- Using the Web
- Creating HTML & CSS
- Authoring SVG
- Authoring XML
- Setting up a server
- Developing specifications
- Developing schemas

Select a topic area...

- Characters
- Language
- Markup & text
- Text direction
- Styling & layout
- Forms
- Navigation
- Cultural issues
- Troubleshooting

Author: Richard Ishida, W3C.

Content last changed 2009-03-12 16:23 GMT

Copyright © 2009 W3C® (MIT, ERCIM, Keio), All Rights Reserved. W3C liability, trademark, document use and software licensing rules apply. Your interactions with this site are in accordance with our public and Member privacy statements.

See also

The topic index organizes links to resources by keywords, rather than tasks (like the index in a book).

The resources by type page lists resources by type (eg. articles, tools, mail archives, etc).

Current status

This index is still a work in progress. It doesn't yet point to all resources on the site. The content will also continually grow and change as resources are added to the site.

http://www.w3.org/International/technique-index?topic=htmlauth

Best practices

I18n resources

W3C I18n Site Index - Mozilla Firefox

File Edit View History Delicious Bookmarks Tools Help

http://www.w3.org/International/technique-index?topic=htmlauth

Internationalization (I18n) Activity

Making the World Wide Web world wide!

Home Resources Techniques Topics News Groups About

I18n site search:

RSS Feeds

W3C I18n technique index

Select the task you want help with from this page to find resources on the W3C Internationalization site among its articles, tutorials, tests and techniques documents.

General task	Creating HTML and CSS	Language
Select a general task...	Select a topic area...	Select a link...
Using the Web	Characters	Getting started
Creating HTML & CSS	Language	Using attributes to declare language
Authoring SVG	Markup & text	Declaring metadata about the language of the intended audience
Authoring XML	Text direction	Declaring language for documents aimed at speakers of more than one language
Setting up a server	Styling & layout	Choosing language values
Developing specifications	Forms	Identifying in-document language changes
Developing schemas	Navigation	Indicating the language of a link destination
	Cultural issues	Styling by language
	Troubleshooting	Using Accept-Language for locale setting

See also

The topic index organizes links to resources by keywords, rather than tasks (like the index in a book).

The resources by type page lists resources by type (eg. articles, tools, mail archives, etc).

Current status

This index is still a work in progress. It doesn't yet point to all resources on the site. The content will also continually grow and change as resources are added to the site.

Author: Richard Ishida, W3C.

Content last changed 2009-03-12 16:23 GMT

Copyright © 2009 W3C® (MIT, ERCIM, Keio), All Rights Reserved. W3C liability, trademark, document use and software licensing rules apply. Your interactions with this site are in accordance with our public and Member privacy statements.

http://www.w3.org/International/technique-index?topic=htmlauth

Best practices

I18n resources

W3C I18N (X)HTML & CSS Techniques - Mozilla Firefox

File Edit View History Delicious Bookmarks Tools Help

http://www.w3.org/International/techniques/authoring-html#langvalues

Choosing language values

Recommendations

- Follow the guidelines in the IETF's BCP 47 for language attribute values.
- Use the shortest possible language tag values.
- Where possible, use the codes zh-Hans and zh-Hant to refer to Simplified and Traditional Chinese, respectively.
- Use the subtag zxx when the text is *known to be* not in any language.
- If the XML format you are using supports it, use xml:lang="", otherwise use xml:lang="und" when the language is undetermined *and you have to label it*.

How to's

- Choosing a Language Tag
  Which language tag is right for me? How do I choose language and other subtags? Covers all the subtag types in the latest version of BCP47. W3C article.
- Language tags in HTML and XML
  A simple overview of the syntax for language tags in BCP 47. W3C article.
- How to choose language values
  In W3C techniques document, Specifying Language in XHTML and HTML Content.
- Specifying language tag values
  In W3C tutorial, Declaring Language in XHTML and HTML.
- Tagging text with no language
  How do I use language markup in HTML or XML content when I don't know the language, or the content is non-linguistic? W3C article.
- Two-letter or three-letter language codes
  Should I use two-letter or three-letter ISO language codes in language tags? W3C article.

Particularly useful links

- IANA Language Subtag Registry
  This is the official location where you will find all subtags available for use in language tags.
- Language Subtag Lookup tool
  User friendly interface to IANA's language tag registry by Richard Ishida. Provides for checking of subtags as well as lookup. Up-to-date with latest version of BCP 47.
- Internet-Draft: BCP 47
  Points to a document containing both RFC 5646 (Tags for the Identification of Languages) and RFC 4647 (Matching Language Tags)
- RFC 5646 Tags for the Identification of Languages
  The specification that describes language tag syntax.
- RFC 4647 Matching of Language Tags
  The specification that describes alternative ways of matching language tags.

Choose another topic
Go to the techniques index

Identifying in-document

Recommendations

- Use the lang and/or xml:lang attributes around text to indicate any

Done

Best practices

I18n resources

Choosing a language tag - Mozilla Firefox

File Edit View History Delicious Bookmarks Tools Help

http://www.w3.org/International/questions/qa-choosing-language-tags

W3C Internationalization

> Español

Home Resources Techniques Topics News Groups About

Choosing a Language Tag

on this page: question - background - answer - by the way - further reading

Intended audience: XHTML/HTML coders (using editors or scripting), script developers (PHP, JSP, etc.), CSS coders, schema developers (DTDs, XML Schema, RelaxNG, etc.), XSLT developers, Web project managers, and anyone who needs guidance on how to construct language tags.

question

Which language tag is right for me? How do I choose language and other subtags?

[Skip to the answer]

background

In HTML and XML documents a **language tag** is used to indicate the language of content.

A language tag is composed of one or more **subtags** separated by hyphens. Subtags can be of various types.

Language tag syntax is defined by the IETF's BCP 47. In the past it was necessary to consult lists of codes in various ISO standards to find the right subtags, but now you only need to look in the IANA Language Subtag Registry. We will describe the new registry below.

This article provides advice on how to choose the components of a language tag. For an overview of the concepts defined in BCP 47, see Language tags in HTML and XML.

Addison Phillips and Mark Davis, authors of BCP 47, provided guidance during the writing of this article.

answer

Accessing the subtag registry

All the subtags you will need to create a language tag are found in one place, the **IANA Language Subtag Registry**. The registry is a long text file, containing nearly 8,000 entries.

The first (and often only) subtag in a language tag always designates a language. It is referred to in BCP 47 as the **primary language subtag**. We will use that term in this document to refer to the subtag that represents a language, to more clearly make the distinction from 'language tag', which refers to the whole thing.

I18n site search:

related links

- Resources
- Techniques
- IANA Language Subtag Registry
- Subtag search tool
- BCP 47

BCP stands for 'Best Current Practice', and is a persistent name for a series of RFCs whose numbers change as they are updated. The latest RFC describing language tag syntax is RFC 5646, Tags for the Identification of Languages, and it obsoletes the older RFCs 4646 3066 and 1766.

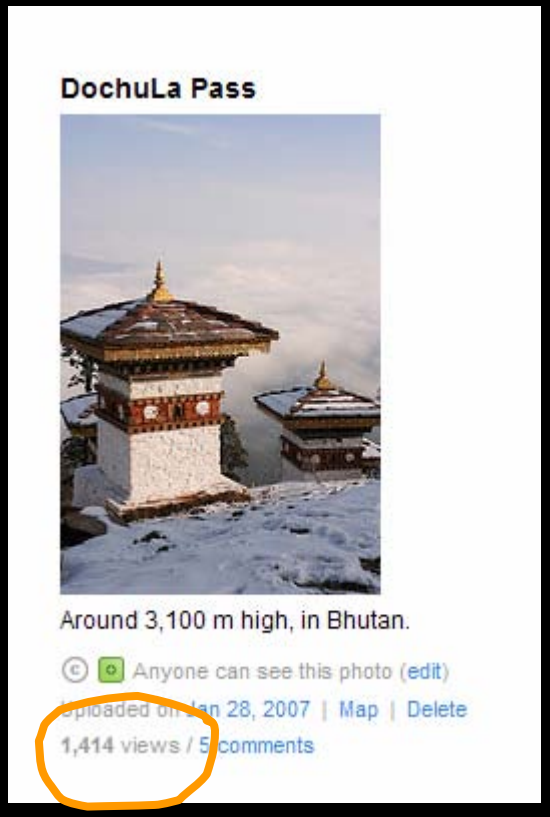
The notes on this page provide guidance that is sufficient for most people wanting to use language tags. There are links to relevant sections of BCP 47 in this margin for people who want to read the full text of the specification.

Note, also, that some environments or systems may dictate choices that are different from what you would otherwise expect. For example, in Java you must use "iw" (deprecated in BCP47) in place of "he" (recommended in BCP47).

Done

Best practices

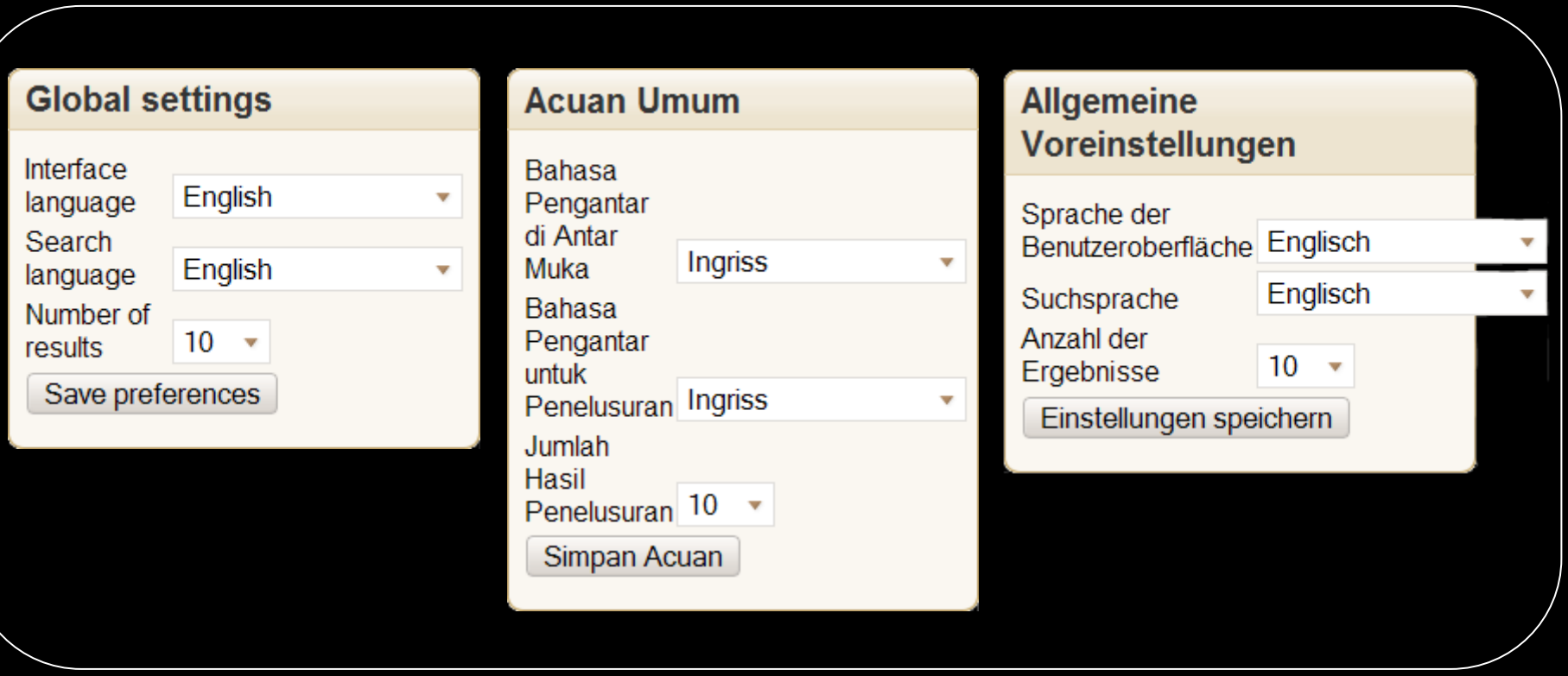
Text expansion



Korean		0.8
English	views	1
Chinese	次檢視	1.2
Portuguese	visualizações	2.6
French	consultations	2.6
German	-mal angesehen	2.8
Italian	visualizzazioni	3

Best practices

Text expansion



Best practices

Checker tool

http:// qa-dev.w3.org/i18n-checker/

1. Discover
2. Check

Detailed report

Severity Description

W3C Internationalization Checker (Prototype only!)

Is your Web site internationalized?

Address http://rishida.net/tools/i18nchecker/test.php

http://rishida.net/tools/i18nchecker/test.php Run another check

Results

2 8 1

XHTML 1.0 :: text/html

Character encoding		Code
HTTP Content-Type	No charset found.	Content-Type: text/html
Byte order mark (BOM)	UTF8	
xml declaration	None found.	
meta charset element	iso-8859-1	<meta http-equiv="Content-Type" content="text/html; charset=iso-8859-1" />
HTML5 meta charset element	None found.	
Language		**Code**
<html lang=	kk	<html lang="kk" xml:lang="to" dir="ltr" xmlns="http://www.w3.org/1999/xhtml">
<html xml:lang=	to	<html lang="kk" xml:lang="to" dir="ltr" xmlns="http://www.w3.org/1999/xhtml">
HTTP Content-Language	ka, ta	Content-Language: ka, ta
meta content-language element	en,fr,sp	<meta http-equiv="Content-Language" content="en,fr,sp" />
Text direction		**Code**
Default direction	ltr	<html lang="kk" xml:lang="to" dir="ltr" xmlns="http://www.w3.org/1999/xhtml">
Class & id names		**Code**
Non-ascii class or id names	8	Show list
Non-NFC class or id names	4	Show list
Request headers		
Accept-Language	en,pt-br;q=0.8,fr;q=0.5,ch;q=0.3	
Accept-Charset	ISO-8859-1,utf-8;q=0.7,*;q=0.7	

↑ TOP

Detailed report

MultilingualWeb project

What is it?

Thematic Network funded by the European Commission through the ICT PSP Grant Agreement No. 250500, and as part of the Competitiveness and Innovation Framework Programme.

Coordinated by W3C/ERCIM.

Runs for 24 months from 1 April 2010.

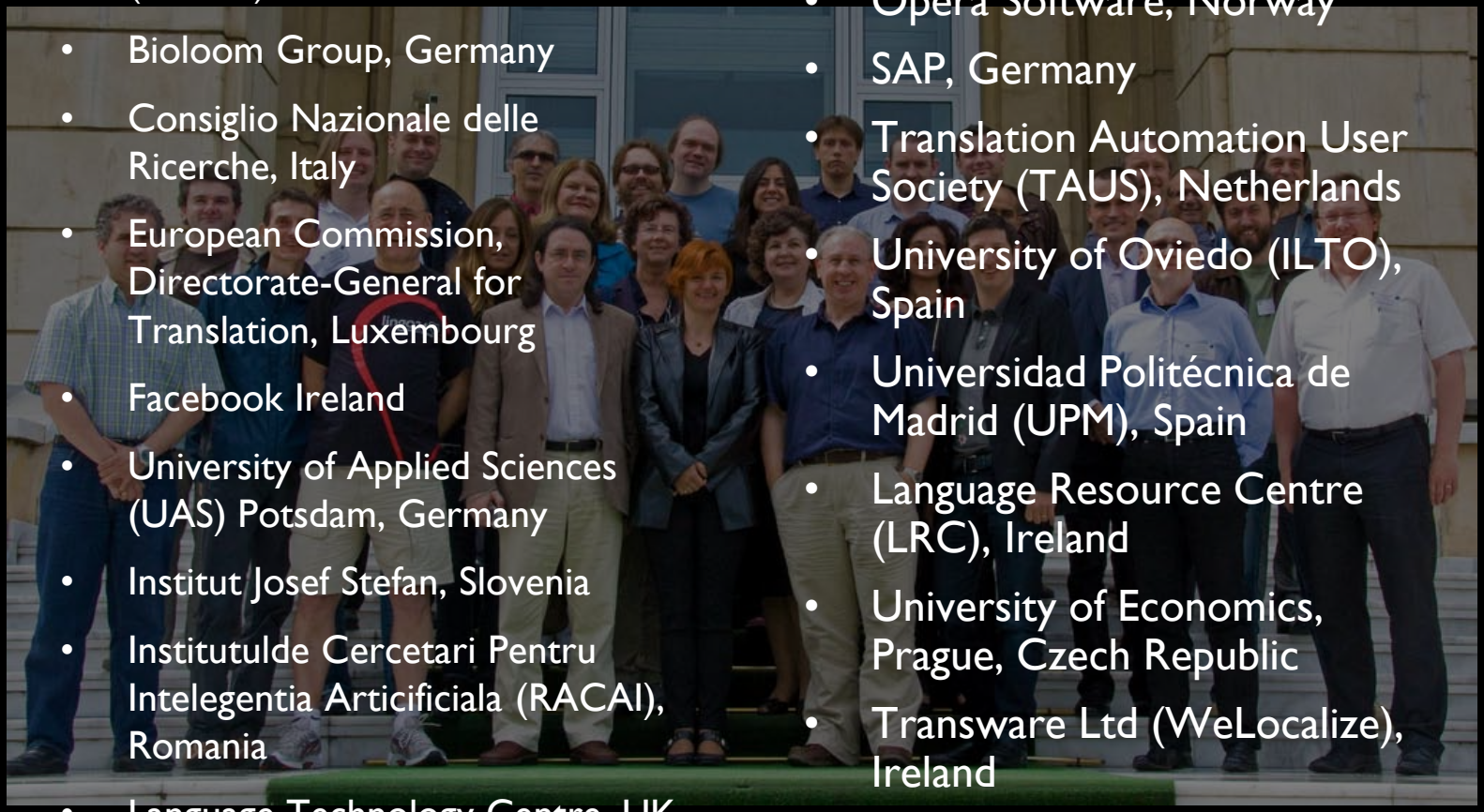
4 public events to examine how the multilingual Web can be improved through standards and best practices.

Objectives

Improve the Multilingual Web by:

- Increasing visibility and use of existing best practices and standards.
- Acting as catalyst for future projects related to standardization, best practices and tools development.
- Creating network links across organizations, scientific disciplines & industry application areas.
- Improving use of standards & best practices in the creation of pages using (X)HTML and CSS by content developers.
- Improving support for multilingual features in Web user agents.

Partners



- Aalto-Korkeakoulusaatio (Finland)
- Bioloom Group, Germany
- Consiglio Nazionale delle Ricerche, Italy
- European Commission, Directorate-General for Translation, Luxembourg
- Facebook Ireland
- University of Applied Sciences (UAS) Potsdam, Germany
- Institut Josef Stefan, Slovenia
- Institutulde Cercetari Pentru Intelegentia Articificiala (RACAI), Romania
- Language Technology Centre, UK
- Lionbridge Belgium
- LISA, Switzerland
- Microsoft Ireland
- Opera Software, Norway
- SAP, Germany
- Translation Automation User Society (TAUS), Netherlands
- University of Oviedo (ILTO), Spain
- Universidad Politécnica de Madrid (UPM), Spain
- Language Resource Centre (LRC), Ireland
- University of Economics, Prague, Czech Republic
- Transware Ltd (WeLocalize), Ireland
- W3C/ERCIM, France (coordination)
- XML-INTL, UK

Events

- Launch event: The Multilingual Web – Where are we?
  5-6 October 2010, Madrid, Spain
- Workshop: Content creation.
  March 2011, Pisa, Italy
- Workshop: Translation tools.
  September 2011, Limerick, Ireland
- Workshop: TBD.
  February 2012, Luxembourg

MultilingualWeb project

Dissemination

http://www.multilingualweb.eu/

@multilingweb

TBA

Getting involved...

Getting involved

- Follow the discussions on the i18n mailing lists (eg. www-international@w3.org), and track other technologies for internationally relevant topics. Follow our RSS feeds and twitter channels (@webi18n and @multilingweb)
- Read and review specifications (http://www.w3.org/TR/tr-technology-drafts) and send comments to the i18n list or direct to the Working Group.
- Discuss local requirements for the Multilingual Web, and if you identify missing features, find ways to coordinate proposals.
- Use features needed for non-Latin script support and push implementers to include more in browsers and authoring tools.

Getting involved

- Review or contribute to development/dissemination of outreach materials, to help others understand how to implement and use international features of the Web.
- Take on board that internationalization is something done by developers and designers – not localizers. Find out how to do it. (http://www.org/International/)
- Use Unicode wherever you can.
- Consider how your content will appear on the Mobile Web.
- Participate in the MultilingualWeb events planned over the coming year and a half.
- Use the I18n Checker (http://qa-dev.w3.org/i18n-checker/) and send ideas for improvements.
- Don't rely on us to do the work for you! We need your help.

The Web needs your help

this is your Web –
not the W3C's

the Web is about
people, not
technology

we need You to
make the Web
worldwide

get involved

Thank you
http://www.w3.org/International/talks/1006-berlin/