The Multilingual Web: Latest developments at the W3C/IETF

Richard Ishida W3C Internationalization Activity

Outline

About the W3C

Standards support for the multilingual Web The changing social context Best practices for the multilingual Web Getting involved



About the W3C

Ţ

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

Device Independence Web Accessibility eGovt lidgets ADDS XPointer Web Namespaces SPARQL Geolocation emant

About the Consortium

Accountable to the global public !

Technology & Society eGovernment Patent Policy Privacy Security

22 Activities 50 Working Grp 13 Interest Grp 5 Coordination Groups

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

About the Consortium Internationalization Activity

- Help W3C Working Groups understand issues and build in requirements relating to worldwide support for Web technologies
- Liaise with other standards organizations to develop support for the international Web
- Help users of Web technology understand what's available to them and how to use it by developing best practices and other resources

عالم گيرويب كو حقيقي طورير عالم گيرينا نا

Ċ^{_} ל_ ∆ףס^{_} ףמי לב^{_} לסרוֹך הי הנח<ל. "Дүниежүзілік торды" нағыз дүниежүзілік етеміз! वर्ल्ड वाईड वेबलाई यथार्थमै विश्वव्यापी बनाउने ! የዓለም አቀፉን ድር በስውንት አለም አቀፋ 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 ធ្វើឲ្យវើលវ៉ាយវ៉េបមានទូទាំងពិភពលោកពិប្រាកដមែន! 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 コヨコ



Standards support for the multilingual Web



وب جهانی را بهدرستی جهانی سازیم! The Path W3C follows" مالمكيرويب كو تقيقي طور ير عالمكير بنايا Ruduzhuphujhu guligli hpnp huduzhuphu jhu nunšlitin CCZL ΔΡΔ[®]PA^b ZCSZ CCL C to making text on the "Дүниежүзілік торды" нағыз Web truly global is वर्ल्ड वाईड वेबलाई यथार्थमै विश्वव የዓለም ከቀትን ድር በከውንት ከለያ Unicode." Κάνοντας τον Παγκόσμιο Ιστό ਵਰਡ ਵਾਈਡ ਵੈਬ ਨੂੰ ਵਾਕਈ ਵਿਸ਼ਵ-ਵਿਆਪੀ ਬਨਾਉਣਾ ! **Tim Berners-Lee** 缔造真正全球通行的万维网 ליצור מהרשת רשת כלל עולמית באמת! meikin ðə ws:ld waid web 'tru:li 'ws:ld'waid ワールド・ワイド・ウェッブを世界中に広げましょう ធ្វើឲ្យវើលវ៉ាយវ៉េបមានទូទាំងពិភពលោកពិប្រាកដមែន! 전세계의 월드 와이드 웹으로 만들기!

Gwneud y we fyd-eang yn wirioneddol fyd-eang! การทำให้ World Wide Web แพร่หลายไปทั่วโลกอย่างแท้จริง ฉยัม ฏิราพิรมาฉริงาร์ รับเวลรารรา ฉยัม ฏิราพิรมาญ เฏิราสู์ภุมมาร์รับ



Standards support Unicode

جعل شبكة الويب العالميّة عالميّة حقَّا! وب جهانى را بەدرستى جهانى سازيم! عالمگيرويب كو تقيقى طور پر عالمگير بنانا

Յամաշխարհային ցանցն իրոբ համաշխարհային դարձ Ե՛֊ՀԼ ՃՔՎ℠Քል՝ Հ֊՟ՎՎ–՟ԼՐ՝ ൎ֊ՂՈՙՀ.

"Дүниежүзілік торды" нағыз дүниежүзілік етеміз! वर्ल्ड वाईड वेबलाई यथार्थमै विश्वव्यापी बनाउने ! የንለም ከቀትን ድር በከውንት ከለም ከቀፋ 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 !אמר באמת

meikin ðə ws:ld waid web 'tru:li 'ws:ld'waid

ワールド・ワイド・ウェッブを世界中に広げましょう ធ្វើឲ្យវើលវ៉ាយវ៉េបមានទូទាំងពិភពលោកពិប្រាកដមែន! 전세계의 월드 와이드 웹으로 만들기!

Gwneud y we fyd-eang yn wirioneddol fyd-eang! การทำให้ World Wide Web แพร่หลายไปทั่วโลกอย่างแท้จริง ฉยัม ฏิราพิรมาฉริณาฉริการ์ มาฉารารราก ฉยัม ฏิราพิรมาญาฏราสู์มาณารรักๆ

ASCII UTF-8	Other		
UTF-8	ASCII		
	UTF-8		

Unicode on the Web



<h2> ფოტოსურათი</h

ჭია bოჭოების ოჯახს ეკუთ მომრგვალო ან ოვალუ ფონზე შავი ლაქები ავ

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 <<u>tbray@textuality.com></u> Jean Paoli, Microsoft <<u>eanpa@microsoft.com></u> C. M. Sperberg-McQueen, W3C <<u>crmsmcq@w3.org></u> Eve Maler, Sun Microsystems, Inc. <<u>eve.maler@east.sun.com></u> Francois 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 <u>W3C</u> 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

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

technical reports index at http://www.w3.org/TR/.

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 $\underline{W3C}$

OLARDS OF LINS DOCUMENT



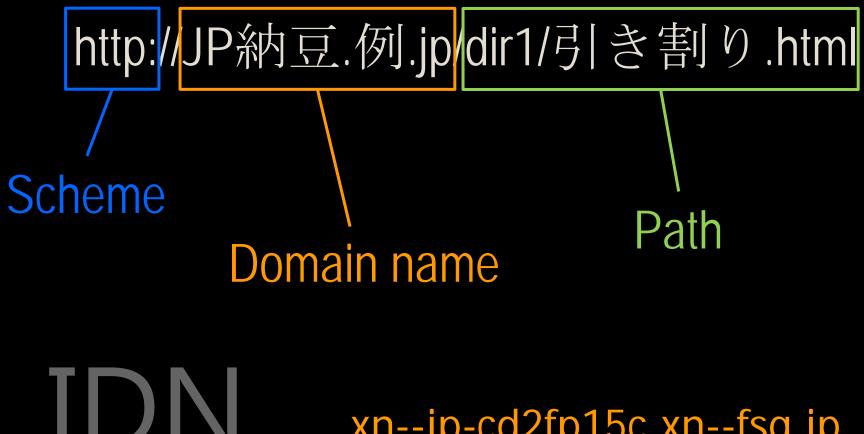
Standards support Unicode normalization

NFDl´zeli´to´´u¨lNFCÍzelítőül

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.



Standards support Web resource identifiers



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



Standards support Web resource identifiers

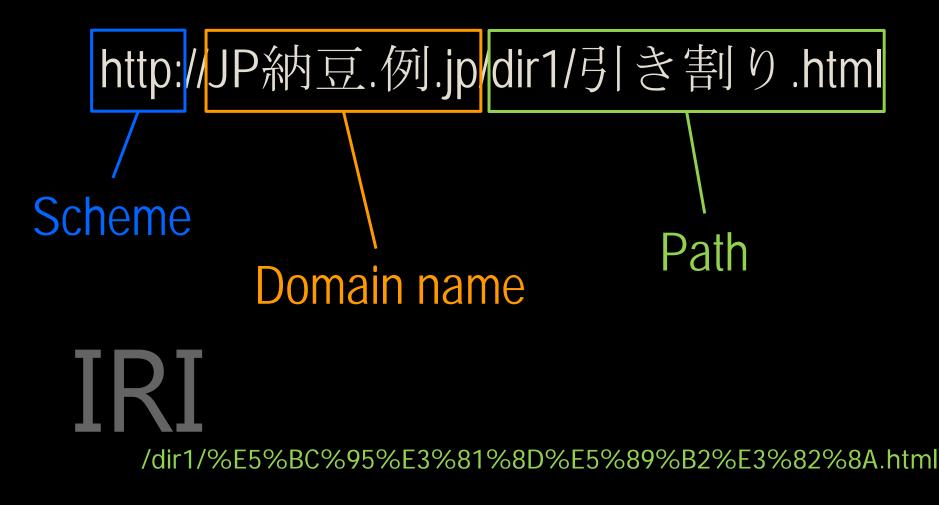


وزارة-الأتصالات مصر// http://





Standards support Web resource identifiers





Standards support Language tags

Before: RFC 3066

language – region

- ISO 639 language codes
- ISO 3166 country codes

en-GB

en

en-scouse



Standards support Language tags: BCP 47

Now: BCP 47 (includes RFC 5646)

language script region variant extension private_use (extlang)

hi

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

az-Cyrl zh-Hans es-419

sl-IT-rozaj-njiva-1994



Standards support Key Events

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 W3C⁹ (MIT, ERCIM, Keio), All Rights Reserved. W3C liability, trademark at use rules apply.

Internationalization Quicktips

ACCESSIBILITY

118N

TYPOGRAPHY

MOBILE

 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.

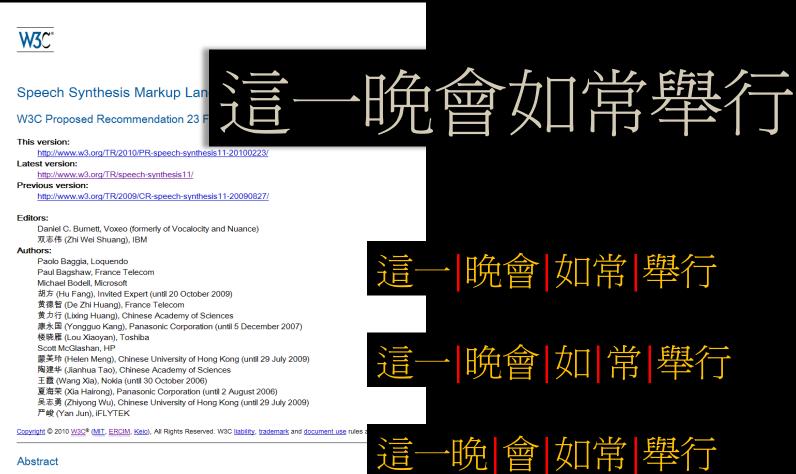
 Declare the language of documents and indicate internal language changes.

A	bs	tra	ct
1			

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

_																											
ies th	ĩ T	! 1	i ,	@ 2	D 2 "	# 3	£ s	\$ 4	€ ¼	% 5	↑ 1⁄2 5	^ 6	↓ \$⁄4 ?	& 7	+ 1⁄8 {	* 8	→ 3/8 }) 9	± 5%8 [) 0	тм 7⁄8]		ن \ י			Ba	ckspace
e-nei hrouç bcum evel	Tab		Q q	2	W	Ω 	E e	œ ĕ		® ¶	T t	1	Y y	¥ ŗ	U u	8 8 *	I i	1	0 •	ø	P p	Þ þ	{ [, ,	}]	- ~ @	
	Caps Lo	ock	A a	. A s				D d	Ð ð v	f	° (0 ? (2 ^			H h	N Л –	J	IJ ij	K k	к -	L 1	Ł ł /	; '	, •	, X		E	nter
	$\stackrel{\rm Shift}{\mathbf{\Delta}}$		in t	Zz		5 -	X x	· ·	C c	¢	· · · ·		B b		N i n	- I	M n	ο μ		× \$		÷ • #		L - -		Shi	ft \
	Contr	ol		Win Key		Alt							(Spac	e)				Ν	ZWN NBS NBS	P	Alt Grap	_	Win Key	М	Ienu	Control

Developing requirements Speech Synthesis Markup Language



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.

different synthesis-capable platforms.

content a standard way to control aspects of speech such as pronunciation, volume, pitch, rate, etc. across

Developing requirements Speech Synthesis Markup Language

Workshop on Conversational Applications

Speech Synthesis Marku,

W3C Proposed Recommendation 23 February 2010

This version:

W3C*

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), France Telecom 黄永国 (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.

different synthesis-capable platforms.

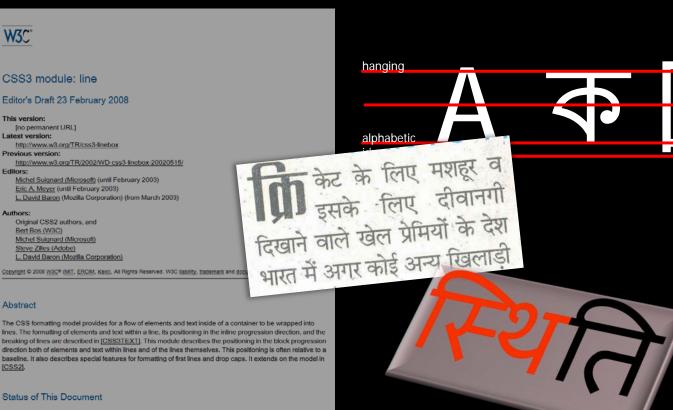
content a standard way to control aspects of speech such as pronunciation, volume, pitch, rate, etc. across

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

June 18-19, NJ, USA



CSS3

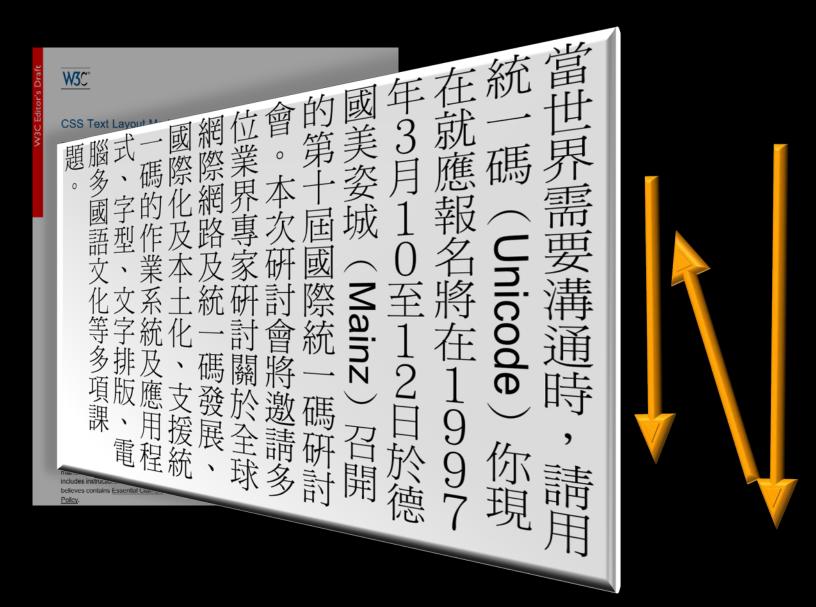


This is a public copy of the editors' draft. It is provided for discussion only and may change at any moment. Its publication here does not imply endorsement of its contents by W3C. Don't cite this document other than as work in progress.

The (archived) public mailing list <u>www-style@w3.org</u> (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: "[css3linebox]

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

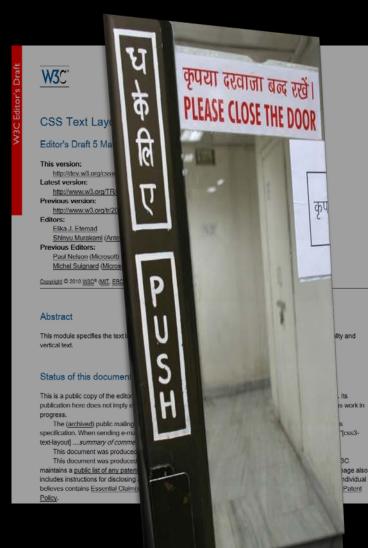








Standards support Vertical text





Standards support Ruby annotation

W3 C [*]		
Ruby And		
This version:	CSS3 R	W3C"
http://www (ZIP archi Latest version http://www Previous versi	This version:	HTML5
http://www Editors: Marcin SA <u>Michel SU</u>	http://ww Previous versi <u>http://ww</u> Editors: Richard	A vocabulary and associated APIs for HTML and XHTML Editor's Draft 4 May 2010 Latest Published Version:
<u>Masayası</u> <u>Martin DÜ</u> <u>Tex TEXI</u> (See <u>Ackr</u>	Former editors Paul Ne <u>Michel S</u> <u>Marcin S</u>	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/
Copyright ©1998-2(licensing rules appl	Copyright © 2010 licensing rules ar	http://www.w3.org/TR/2009/WD-html5-20090423/ http://www.w3.org/TR/2009/WD-html5-20090212/ http://www.w3.org/TR/2009/WD-html5-20080610/ http://www.w3.org/TR/2008/WD-html5-20080122/ Editors:
"Ruby" are shor pronunciation or an XHTML mod	Abstract "Ruby" are sh pronunciation associated wit	Ian Hickson, Google, Inc. This specification is available in the following formats: <u>single page HTML</u> , <u>multipage HTML</u> . This is SRevision: 1.4073 S. <u>Countart # 2010 USC[®] MUT, EDDM Kelo</u> , All Rights Reserved. HOC <u>Deblity</u> : <u>preserved</u> and <u>document use</u> noise apply:
Status of Th	<u>(RUBY</u>). Status of ⁻	Abstract
supersede this a This document l endorsed by the	See <u>open issu</u> This docume	This specification defines the 5th major revision of the core language of the World Wide Web: the I Markup Language (HTML). In this version, new features are introduced to help Web application ar elements are introduced based on research into prevailing authoring practices, and special attentit
reference mater Recommendation This enhances t	This section a supersede thi. can be found Publication as	eiements are introduced based on research into prevaiing autoring practices, and special arening given to defining clear conformance criteria for user agents in an effort to improve interoperability.
	Publication as	Otation of This designment

Status of This document

document and

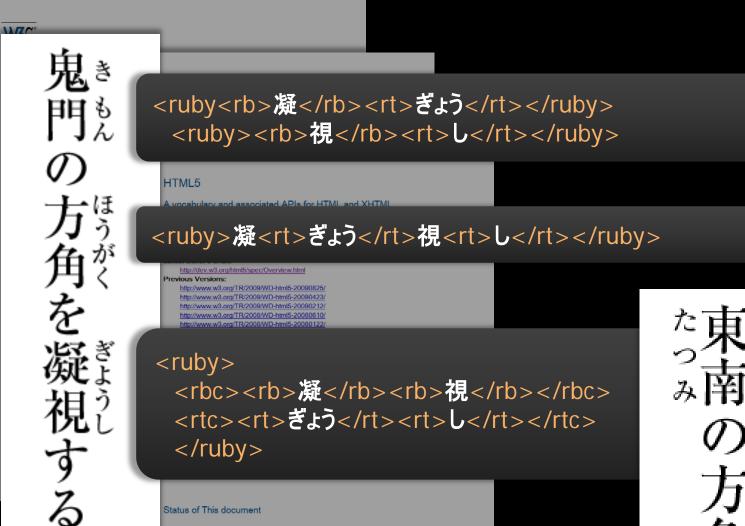
This section describes the status of this document at the time of its publication. Other documents may supersede this document. A list of current W3C publications and the most recently formally published revision of this technical report can be found in the W3C technical reports index at http://www.w3.org/TR?.

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





Standards support Ruby annotation



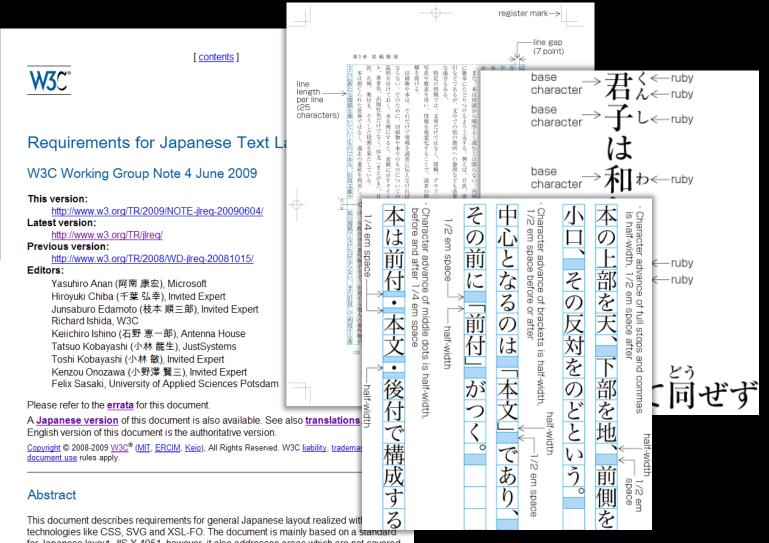
This section describes the status of this document at the time of its publication. Other documents may supersede this document. A list of current W3C publications and the most recently formally published revision of this technical report can be found in the <u>W3C technical reports index</u> at http://www.w3.org/TR?

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

Group Note

W3C Working

Developing requirements Requirements for Japanese Layout



for Japanese lavout, JIS X 4051, however, it also addresses areas which are not covered by JIS X 4051.



Standards support Web fonts \mathbf{O} 67552 0K



Standards support 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.



Standards support Language declarations

W3C*

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-html5-20090 http://www.w3.org/TR/2009/WD-html5-20090 http://www.w3.org/TR/2009/WD-html5-20090 http://www.w3.org/TR/2009/WD-html5-20090 http://www.w3.org/TR/2009/WD-html5-20090

lan Hickson, Google, Inc.

This specification is available in the following formats \$Revision: 1.4073 \$. Contract # 2010 USC [®] AUT, ERCAR, Keip, AU Rights Reserved, W3C (Int.)

Abstract

This specification defines the 6th major revision of th Markup Language (HTML). In this version, new feat elements are introduced based on research into pre given to defining clear conformance criteria for user r

Status of This document

This section describes the status of this document a supersede this document. A list of current W3C pub this technical report can be found in the <u>W3C technic</u>

If you wish to make comments regarding this docum

<html lang="de">

. . .

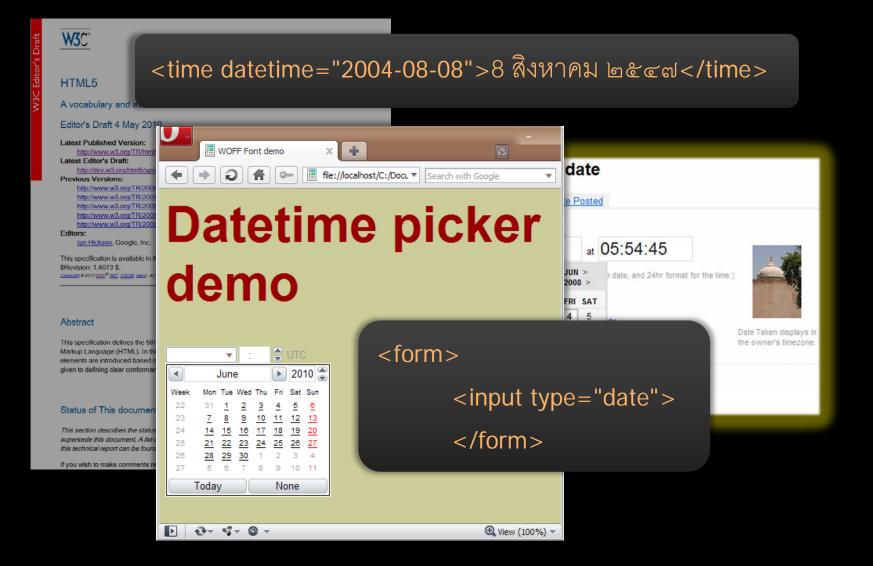
<head>

</head>

<meta http-equiv="Content-Language" content="de">



Standards support Date and time





Standards support Bidirectional text support

ب Wime التدويل، نشاط التدويل، W3C

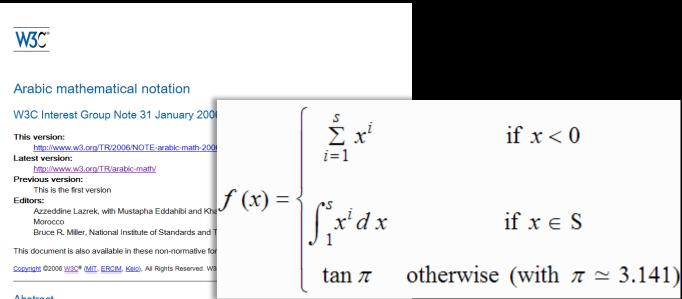
<description dir="rtl">W3C ، نشاط التدويل، description dir="rtl">w3C /description

Developing requirements Augmenting bidi support in HTML5 & CSS

[contents]									
₩3C°		TOP I	RATED	TOP RATED					
Additional Requirements for B W3C Working Draft 4 March 2010	Bidi in HTML	$\frac{\text{RESTAURANTS}}{\text{Aroma - 3 reviews}}$		RESTAURANT	S				
This version: http://www.w3.org/TR/2010/WD-html-bidi-201 Latest version:				Aroma - 3 reviews					
http://www.w3.org/TR/html-bidi/ Editor:	Translate text, webpages and documents								
Aharon Lanin, Google Additional Contributors:	Enter text or a webpage URL, or upload a document.								
Adil Allawi, Technical Director, Diwan Softwa Matitiahu Allouche, Bidi Architect, IBM Uri Bernstein, Google Douglas Davidson, Apple Mark Davis, Senior I18n Architect, Google; P	Internationaliz	bpages and documen	its						
Martin J. Dürst, W3C I18n Interest Group Ch Asmus Freytag, President, ASMUS, Inc.	Translate fror	m: Enc En	Enter text or a webpage URL, or upload a document.						
Richard Ishida, I18n Lead, W3C Shanjian Li, Google Mohamed Mohie, IBM Jeremy Moskovich, Google Shachar Shemesh, Lingnu Open Source Con Gaal Yahas, Google	Translate into			W3C	خابطه لسانی عد میت ،				
Copyright © 2007-2010 W3C [®] (MIT, ERCIM, Keio), All Rights I document use rules apply.	Reserved. W3C <u>liability,</u> j		ranslate from		T'ate				
Abstract Authoring a web app that needs to support both right to take as input and display both left-to-right and right number of challenges that make it an especially labor	nt-to-left data, usually	ght interfaces, or	ranslate into:						

number of challenges that make it an especially laborious and bug-prone task. Some of these are due to browser bugs, but some can be traced to a gap in the specification of

Developing requirements Arabic mathematics



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 <u>W3C technical reports index</u> at http://www.w3.org/TR/.

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

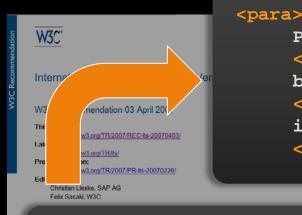
This Note has been written by participants in the <u>Math Interest Group</u> (W3C me <u>W3C Math activity</u>. Please direct comments and report errors in this document t list with a public <u>archive</u>.

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

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

list with a public archive.

Standards support Internationalization Tag Set



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

This document defines data categories and their implementation as a set of elements and attributes called the Intern Tag Set (ITS). ITS is designed to be used with schemas to support the internationalization and localiz mas and documents. An implementation is provided for three schema languages: XML DTD, XML.



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

The changing social context

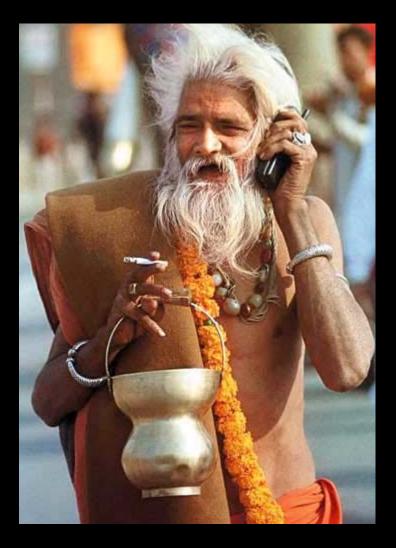
Social context The rise of the Mobile Web



- "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

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.



Best practices for the multilingual Web

Best practices Capturing guidance for spec developers

<u>W3C</u>		
Pote		[contents]
Zharacter Model کے	W3C"	
W3C Recommendati	<u>vvJ</u>	[contents]
This version:	2	
http://www.w3.org/TR/2	Best Practices	
Latest version:	Dest Flactices 2	
http://www.w3.org/TR/c	W00.00 11 0	
Previous version:	W3C Working Grouई	
http://www.w3.org/TR/2	8	
Editors:	This version:	Working with Time Zones
Martin J. Dürst, W3C ≤	http://www.w3.org/T	
François Yergeau (Invi	Latest version:	
Richard Ishida, W3C <	http://www.w3.org/Tl Previous version:	W3C Working Group Note 13 October 2005
Misha Wolf (until Dec 2	http://www.w3.org/Tl	
Tex Texin (Invited Exp	Editors:	This version:
Please refer to the errata for	Yves Savourel, ENL	http://www.w3.org/TR/2005/NOTE-timezone-20051013/
Thease refer to the citata lo	Jirka Kosek, Invited	Latest version:
See also translations.	Richard Ishida, W3C	http://www.w3.org/TR/timezone
		Editors:
Copyright © 2005 W3C® (MIT, ER	This document is also ava	Addison Phillips, (Invited Expert)
	publication from 31 Octobe	Felix Sasaki, W3C
Abstract	Copyright © 2008 W3C [®] (MIT, E	Mark Davis, IBM
Abstract		Martin Dürst, Aoyama University
This Architectural Specification		This document is also available in these non-normative formats: XML.
with a common reference for	Abstract	
Character Set, defined jointly		Copyright © 2005 W3C [®] (MIT, ERCIM, Keio), All Rights Reserved. W3C liability, trademark and document use rules apply.
the terms 'character', 'encodi	This document provides a	
character encodings, charac	internationalized properly.	
	applications, as well as the	
For normalization and string i		Abstract
Wide Web 1.0: Normalization	Status of this Docu	
Model for the World Wide We		This document discusses some of the problems encountered when working with the date, time, and
	This section describes the	dateTime values from [XML Schema] when those value include (or omit) time zone offsets. Many W3C technologies rely on date and time types. Examples include the [XPathFO] specification, since it is the basis for
Status of this Docum	supersede this document.	XQuery and XSLT processing of date/time values, but the concepts affect any date / time processing.
	be found in the W3C techi	Aquery and ASET processing of date anne values, but the concepts affect any date rune processing.
This section describes the st	THE REPORT OF	
supersede this document. A	This is a W3C Working Gr	
be found in the <u>W3C technic</u>	developed by the Internati Activity.	Status of this Document
This document contains the	veavity.	This section describes the status of this document at the time of its publication. Other documents may
W3C Recommendation. It ha	Feedback about this docu	supersede this document. A list of current W3C publications and the latest revision of this technical report can
	"[Comment on xml-i18n-br	be found in the W3C technical reports index at http://www.w3.org/TR/.
	this list are publicly availab	be loand in the <u>recented incontention at how</u> at http://www.wolorg/invi
W3C Recommendation. It ha		This document discusses the topic of date, time, and dateTime values from [XML Schema] with and without
This document contains the	Publication as a Working (time zone offsets. Examples are given mainly relying on [XML Schema] and [XPathFO], since these are the
	document and may be uny	basis for [XQuery] and [XSLT 2.0] processing of date/time values.

This document is a W3C Working Group Note. It has been produced by the i18n Core Working Group, which is part of the Internationalization Activity.

n xml-i18n-bp W	VD]" in the subject line of your email, followed by a brief subject. The archives for	
out this docume	ent is encouraged. Send your comments to www-i18n-comments@w3.org. Use	
fers [CharlR]].	part of the Internationalization Activity.	
includes and the second		O FUNNING

markup for bidirectional text

Normalization

working with case sensitivity

more information about date & time

Publication as a Working This section describes the s this list are publicly avail Feedback abo

Best practices

\cap	r C
$\mathbf{C}\mathbf{S}$	

	UA IE
18n) Activity	version 8
	OS XP,
s News Groups About	OS Unis renderer 1.62
	date 2010
	Alphabetic script (Georgian)
	Alphabetic script (Armenian)
€ ∋ • 🔮 😹 😂 🍪 📮 📮 💗 🤅	Complex scripts no (Khmer)
ltr context, direction rtl, unicode-b	Complex scripts no (Hindi)
	Shaping scripts no (Arabic)
Characters should be i	Shaping scripts no s (Urdu)
	Mixed scripts no (Thai)
ط التدويل" i18n activity	Mixed scripts no s (Tibetan)
	Mixed
ط التدويل" i18n activity	scripts (Myanmar)
	Assertion:
	Link to tests: V
	EOT fonts
Accession in a LTD content if	EUTIONIS
	UA IE
	version 8
	OS XP,
= means you should check th	OS Unis
same, except for font differenc	renderer 1.62
	date 2010 Alphabetic
	script yes (Georgian)
	Alphabetic script yes
	(Armenian)
Prov.	Complex
Done	_
XHTML1.0***** XHTML5 XHTML1.1	
	 Ison test: Itr context, direction rtl; Ele Edit View Higtory Delicious Bookmarks Ele Edit View Higtory Delicious Bookmarks Itr context, direction rtl, unicode-b Itr context, direction rtl, unicode-b Characters should be i Characters should be i Itan activity "Lize Laboration direction direction in a LTR context, if element containing mixed dire that element will not be in the means you should check th

		XHTML1.0 ^{×m1}	XHTMLS	XHTML
		XHTML1.0 ^{xml}	XHTML5	XHTML

🕑 Test results: Webfonts - Mozilla Firefox								
Ejle Edit View Higtory Deljicious Bookmarks Tools Help								
€∋)- 🔐 🕼 🗧	è 🌀 📮 🕯	şi 🥪 🔞 (🕘 🖬 📃 i	🕋 📄 http://	/www.w3.org/Int	ernational/te 🤍 👻 🚼 🛪 Goog	le 🔎
	WOFF for	nts						~
		IE	IE	Firefox	Chrome	Opera	Safari	
	version	8	7	3.6.3	4.1.249.1045	9.64	4.0.5	
1	OS	XP, SP3						
	OS renderer	Uniscribe 1.626.5756.0	Uniscribe 1.626.5756.0	Uniscribe 1.626.5756.0	Uniscribe 1.626.5756.0	Uniscribe 1.626.5756.0	Uniscribe 1.626.5756.0	
	date	20100427	20100427	20100427	20100427	20100427	20100427	
	Alphabetic script (Georgian)	no 🗸	no 🗸	yes	no 🗸	no 🖌	no 🗸	
	Alphabetic script (Armenian)	no 🗸	no 🗸	yes	no 🗸	no 🗸	no 🗸	
	Complex scripts (Khmer)	no 🗸	no 🗸	yes	no 🗸	no 🖌	no 🖌	
	Complex scripts (Hindi)	no 🗸	no 🖌	yes	no 🖌	no 🖌	no 🖌	
	Shaping scripts (Arabic)	no 🗸	no 🖌	yes	no 🖌	no 🖌	no 🖌	
	Shaping scripts (Urdu)	no 🗸	no 🗸	yes	no 🗸	no 🗸	no 🗸	
	Mixed scripts (Thai)	no 🗸	no 🗸	yes	no 🗸	no 🗸	no 🗸	
	Mixed scripts (Tibetan)	no 🗸	no 🖌	yes	no 🖌	no 🖌	no 🖌	0
	Mixed scripts (Myanmar)	no 🗸	no 🗸	yes	no 🖌	no 🖌	no 🗸	
	Assertion:							

WOFF fonts

IE Firefox Chrome Opera Safari 7 4.0.5 3.6.3 4.1.249.1045 9.64 P, SP3 XP, SP3 XP, SP3 XP, SP3 XP, SP3 XP, SP3 niscribe Uniscribe Uniscribe Uniscribe Uniscribe Uniscribe 626.5756.0 1.626.5756.0 1.626.5756.0 1.626.5756.0 1.626.5756.0 20100427 20100427 100427 20100427 20100427 20100427 no 🧹

(≥ ≥ ≥ 5

X X X V



	(a) W3C Internationalization (I18n) Ad	ctivity - Mozilla Firefox		s x)
	Ele Edit View History Delicious Bookmarks Tools Help			
	💽 🌛 - 🖹 🔄 🏷 📮 📮 🤎 🎯 🔛 📰 🗎 🗎 http://www.w3.org/Inte	ernational/	ରି 💚 🝷 🚼 ଟ Google	٩
	W3 C [•] Internationalization (I18n) Activity		I18n site search:	
	Making the World Wide Web world wide!		RSS Feeds 🔝	1
	Home Resources Techniques Topics News Groups About			
Articles	Home page	Site links	i18n resources	
	The W3C Internationalization (I18n) Activity works with W3C working groups	Getting Started	Authoring HTML & CSS	
	and liaises with other organizations to make it possible to use Web technologies with different languages, scripts, and cultures. From this page	⊞ Tasks	Authoring SVG	0
	you can find articles and other resources about Web internationalization,	E Topics	 Authoring XML Developing specifications 	
	and information about the groups that make up the Activity.	■ Resource types	Setting up a server	
Tutoriale	More Recent highlights	≣ Reviews	 Developing schemas Using the Web 	
Tutorials		■ Mail archives	V doing the Web	/
	 30 April 10 Text direction tests updated 27 April 10 First translation into Hindi 	■ Aggregated data	Quick links	
	 27 April 10 Webfonts tests updated to include WOFF 	Translation	Planet	U
	14 April 10 Language declaration tests updated	News by category	Specifications	
http://www.ww2.org/l	ntornotion conference	News archives	 Articles etc. 	
http://www.w3.org/l			 Tests 	
J	All news		 MultilingualWeb 	
	24 May 2010		I @webi18n tweets	
	Talk slides: W3C India Office opening			
	On 6th May Richard Ishida gave a talk entitled Internationalization & its challenges at the Technology, Standards and Internationalization		About the Activity	4
Tests	Conference that was run to mark the opening of the W3C India Office in New D		Groups: Core, ITS, IG, JLTF	
IESIS	The talk describes the work of the Internationalization Activity at the W3C, and Asia can contribute to that work.	d ways in which people in South	Mission, Contacts Activity Statement	
	₽ Edit		Deutlainedal	
	Categories: Talks, w3cWebDesign		Participate!	4
		30 April 2010	Join a Working Group Review a W3C specification	
T U	Text direction tests updated The HTML and CSS text direction tests in the Internationalization Activity test	suite have been undated as	Translate a specification or	
Talks	follows:	suite nave been updated as	page Subscribe to the Interest Group	
	 HTML5 and XHTML5 were added as new test formats tests now use strict DOCTYPEs 		list	J
	 reference graphics in the vertical text tests were improved 		Search for news	1 H
	the order of tests related to browser chrome was changed			1 H
	In addition, a new results page was created for the tests about vertical text.			
Tools	Other changes behind the scenes include gathering of all tests into a single PH longer used). Code in the PHP file was also streamlined.	P file (generatentmi.php is no	 All Words Some Word 	
10013	[search key: results-vertical-text]		Entire phrase	
	🖉 edit		Search	
	Categories: Highlight, Test, w3cWebDesign, w3cWebUserAgents		Admin	. II
	New terms left as late 10 at	28 April 2010	Admin	4 H
Deviewe	New translation into Hindi Thanks to the Bhavatmaj Seth, the following getting-started article has been t	ranslated into Hindi.	 Admin Profile (rishida) 	- 11
Reviews	वेव पर भाषा (Language on the Web)		 Logout (rishida) 	~
	Done		- - - - - - - - - -	
	Done			12
	बेब पर भाषा (Language on the Web)		Logout (nshida)	
	Thanks to the Bhavatmaj Seth, the following getting-started article has been t	ranslated into Hindi.	Profile (rishida)	

New translation into m



Articles

Tutorials

Technical notes

Tests

Talks

Tools

Reviews

W3C^{*} Internationalization (I18n) Activity Making the World Wide Web world wide!

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

I18n site search:

RSS Feeds 🔝

> حربي Deutsch Eλληνικό Español Français 노고 ਪ੍ਰਕਾਰ Italiano 田 本語 한국어 Nederlands Português-BR Română Русский Svenska אויי Türkçe Українська Tiếng Anh 简体汉语 繁體 中文

On this page

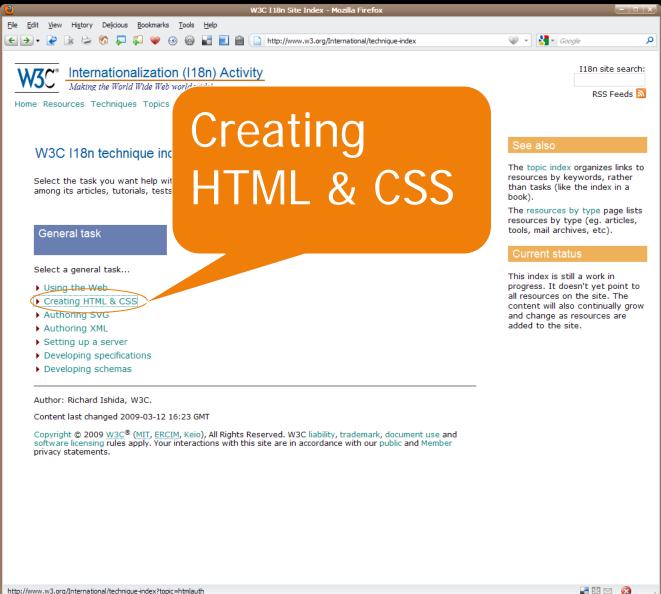
Getting started Characters Language Markup & text Text direction Styling & layout Forms Navigation Web addresses Cultural issues Other

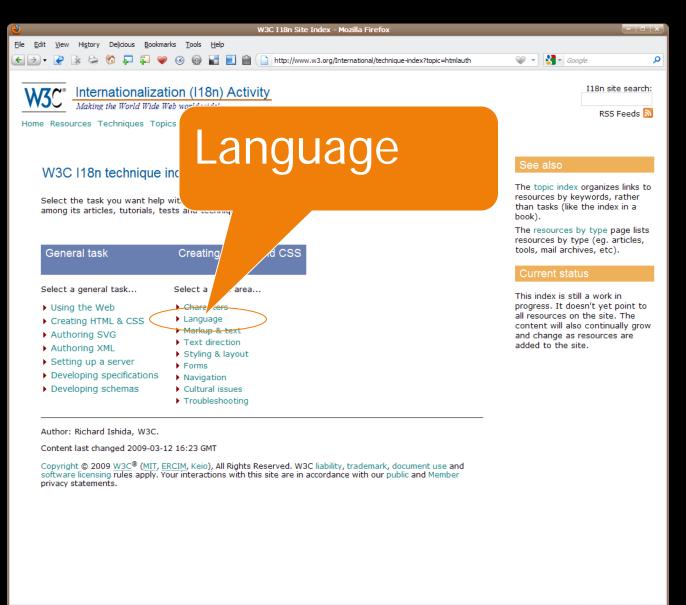
Language

Specifying Language in XHTML & HTML Content Language tags in HTML and XML Choosing a language tag <u>NEW!</u> 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

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



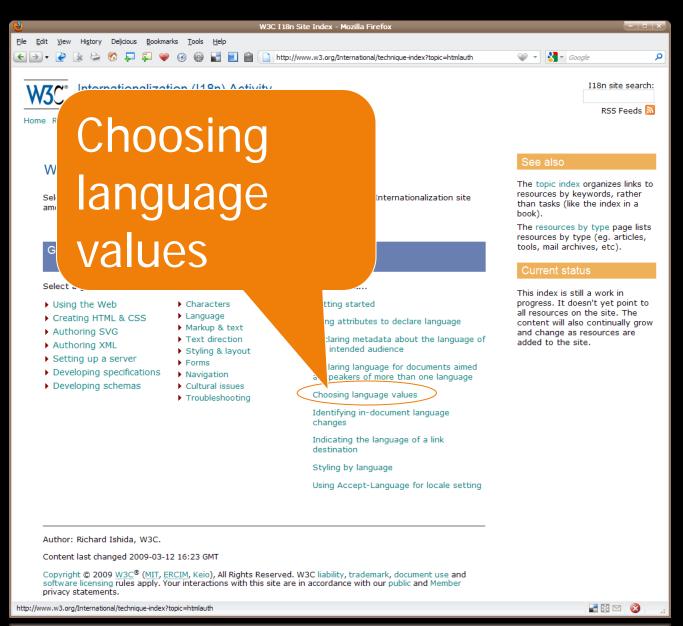




http://www.w3.org/International/technique-index?topic=htmlauth

🖃 🗟 🖂 🔞

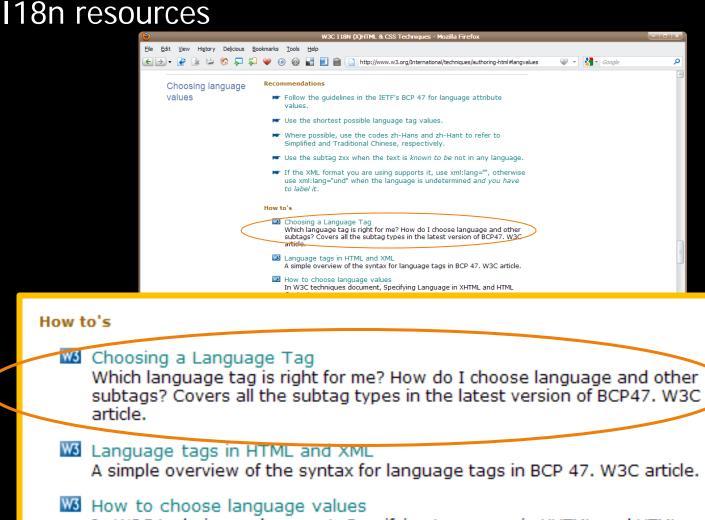




四限网

03

Best practices



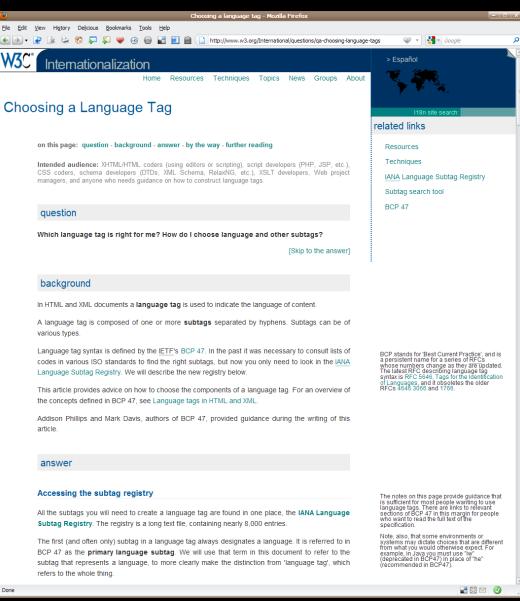
In W3C techniques document, Specifying Language in XHTML and HTML

		Choose another topic Go to the techniques index)		
Done	ldentifying in-document	Recommendations Use the lang and/or xml:lang attributes around text to indicate any 		I	~
<u> </u>					
Done	in-document	Use the lang and/or xml:lang attributes around text to indicate any		0	11 (S)

ο

Best practices

118n resources



refers to the whole thin

subtag that represents a language, to more clearly make the distinction from 'language tag', which

ecommended in BCP47).

E 18 63



Best practices Text expansion





© Anyone can see this photo (edit) optoaded on lan 28, 2007 | Map | Delete 1,414 views / 5 comments

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

P

Best practices Text expansion

Global settings			Acuan Umum		Allgemeine			
Interface			Bahasa			Voreinstellung	en	
language	English	•	Pengantar			Sprache der		
Search	English	_	di Antar Muka	Ingriss	-	Benutzeroberfläche	Englisch	•
language	English	•	Bahasa	ingnoo		Suchsprache	Englisch	•
Number of results	10 💌		Pengantar			Anzahl der		
Save prefe	erences		untuk	Ingrice	-	Ergebnisse	10 🔹	
	oronoos		Penelusuran	Ingliss	¥	Einstellungen spe	ichern	
			Jumlah Hasil					
			Penelusuran	10 🔹				
			Simpan Ac	uan				



Best practices Checker

Deet		 Detailed report 		
Besi	practices	Rovarity Description		
Che	ecker tool		ternational	lization Checker (Prototype only!)
		✓ Address		http://rishida.net/tools/i18nchecker/test.php
		http://rishida.net/tool	s/i18nchecker/t	Run another check
		▼ Results		
http://	/ qa-dev.w3.org/	i <mark>18n-checke</mark>	er/ 2	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 content="text/html;
charset=iso-8859-1" http-equiv="Content-Type"/>
•	Discover	HTML5 meta charset element	None found.	
		Language		Code
2.	Check	<html lang="</td"><td>kk</td><td><html <br="" dir="ltr" lang="kk" xml:lang="to">xmlns="http://www.w3.org/1999/xhtml"></html></td></html>	kk	<html <br="" dir="ltr" lang="kk" xml:lang="to">xmlns="http://www.w3.org/1999/xhtml"></html>
~ •	Check	<html xml:lang="</td"><td>to</td><td><html <br="" dir="ltr" lang="kk" xml:lang="to">xmlns="http://www.w3.org/1999/xhtml"></html></td></html>	to	<html <br="" dir="ltr" lang="kk" xml:lang="to">xmlns="http://www.w3.org/1999/xhtml"></html>
		HTTP Content-Language	ka, ta	Content-Language: ka, ta
		meta content-language element	en,fr,sp	<meta content="en,fr,sp" http-equiv="Content-Language"/>
		Text direction		Code
		Default direction	ltr	<html <br="" dir="ltr" lang="kk" xml:lang="to">xmlns="http://www.w3.org/1999/xhtml"></html>
		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		utf-8;q=0.7,*;q=0.7

The social context Mobile Web Initiative

W3C^{*} Mobile Web

The Web on the Move

Mobile Web Initiative Home

About MWI

News

Working Groups

Sponsorship

Specifications

Resources Center

Making Web access from a mobile device as simple as Web access from a desktop device.

"The Mobile Web Initiative's goal is to make browsing the Web from mobile devices a reality", explains Tim Berners-Lee, W3C Director and inventor of the Web. "W3C and mobile industry leaders are working together to improve Web content production and access for mobile users and the greater Web."

Mobile Web Initiative participants are <u>developing best practices</u> for creating mobile-friendly content and applications, <u>enabling easy access to device descriptions</u>, <u>setting up test suites</u> for increased interoperability of mobile browsers, and exploring ways to use the <u>Web on mobile devices to bridge</u> the digital divide.

Best practices

Device descriptions

Test suites

mobileOK checker

LATEST NEWS FEATURED SPONSOR Proposed Recommendation of mobileOK FORUM MODILE WED 2.0 1 Orall Web based industry revitalization" is to Basic Tests 1.0-3 November 2008 W3C Organizes Workshop on the Future of Social Networking-14 October 2008 develop standards and guidelines for obtaining W3C Organizes Workshop on Security for interoperability of Web contents between Access to Device APIs from the Web -9 various devices and realize visions for October 2008 constructing the next generation Mobile Web environment. Become a MWI Sponsor - see other Sponsors more news... testimonials NEXT MWI EVENT CHECK YOUR SITE / YOUR BROWSER Most of the MWI staff will be present at the ICT 2008 event, organized by the European Is your site mobileOK? Commission, on 25-27 Nov. 08, in Lyon, France. Check! Marie-Claire, Philipp, and Stéphane will be at booth D06, in the International Village Drag the mobileOK bookmarklet to your toolbar to quickly check pages: mobileOK check How does your mobile browser fare in the Web Compatibility Test? 13 14 15 18 111 the Web Compatibility Test? How does your mobile browser fare in to quickly check pages: mobileOK check Drag the mobileOK bookmarklet to your toolbar

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/trtechnology-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. (<u>http://www.org/International/</u>)
- 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 II8n 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/1010-madrid/