

# HTML5 i18n: A report from the front line



Richard Ishida  
W3C Internationalization Activity

Jirka Kosek  
University of Economics, Prague

# Outline

W3C Internationalization Activity

Bidirectional text

Ruby

Translate flag

# About the Consortium

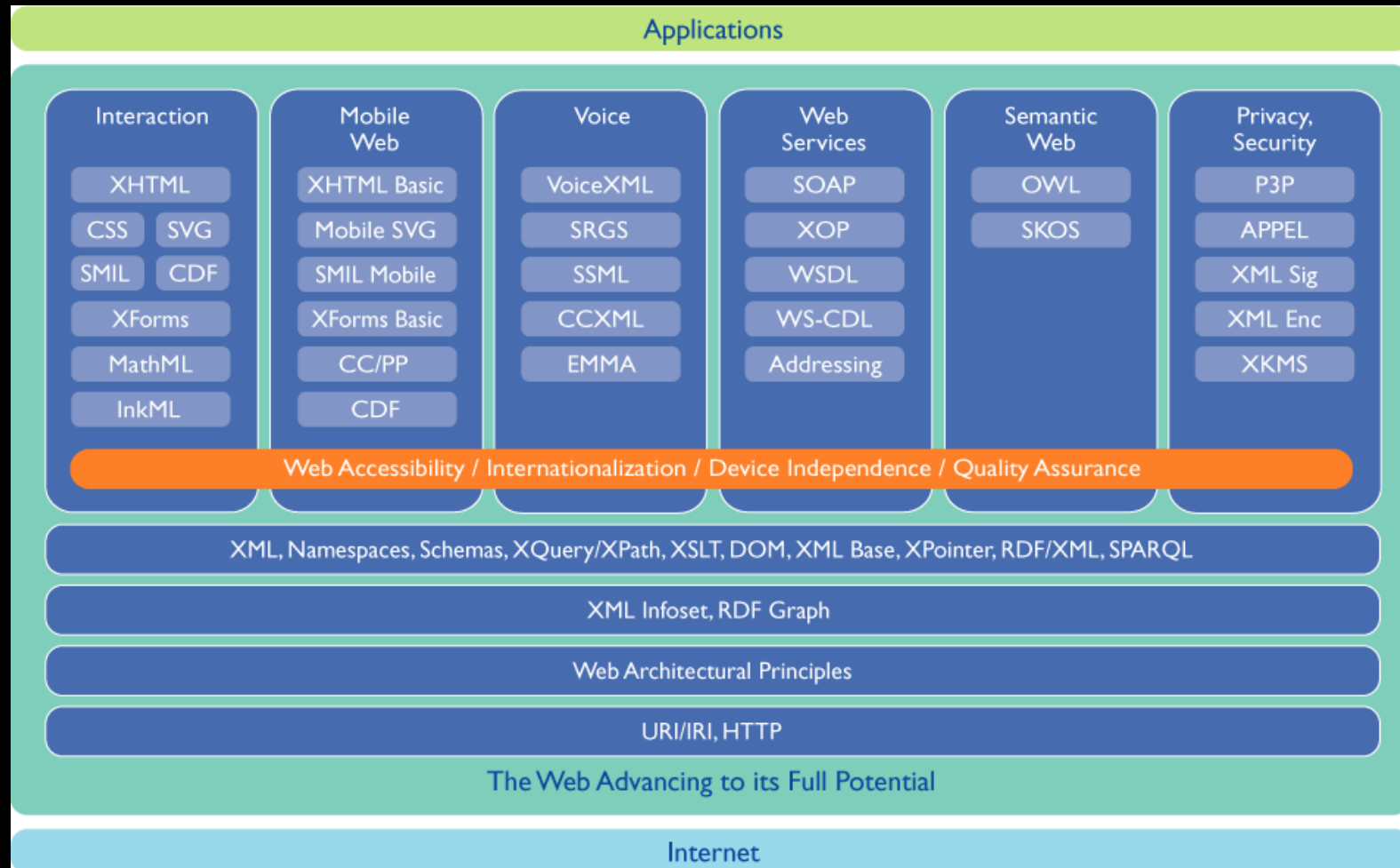


1994: World Wide Web Consortium created and still led by:  
**Sir Tim Berners-Lee**,  
inventor of the WWW.

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



# W3C technology stack



# Internationalization (i18n) Core Working Group

جعل شبكة الويب العالمية عالمية حقًا!

Համաշխարհային ցանցն իրոք համաշխարհային  
ՀՀԻԼ ΔΡΔΨΡΔΨ ςԸΨΔԸԼԸΨ Δ.Ը.Ո.Ն.Վ.

"Дүниежүзілік торды" нағыз дүниежүзілік етем:  
वर्ल्ड वाईड वेबलाई यथार्थमै वश्वव्यापी बनाउने  
የዓለም አቀፉን ድር በእውነት አለም አቀፍ ማድረግ!

Κάνοντας τον Παγκόσμιο Ιστό πραγματικά Παγκόσμιο  
ਵਰਡ ਵਾਈਡ ਵੈਬ ਨੂੰ ਵਾਕਈ ਵਿਸ਼ਵ-ਵਿਆਪੀ ਬਣਾਉਣਾ !

締造真正全球通行的万维网

ליצור מהרשת רשת כלל עולמית באמת!

'meikɪŋ ðə wɜːld waɪd web 'truːli 'wɜːld'waɪd

ワールド・ワイド・ウェブを世界中に広げま  
ធ្វើឱ្យវិទ្យុសាស្ត្រពិភពលោកពិបាកដំឡើង  
전세계의 월드 와이드 웹으로 만들기!

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

การทำให้ World Wide Web แพร่หลายไปทั่วโลกอย่างแท้จริง  
အိန္ဒိယ-အိန္ဒိယ-အိန္ဒိယ-အိန္ဒိယ-အိန္ဒိယ-အိန္ဒိယ-အိန္ဒိယ-အိန္ဒိယ-အိန္ဒိယ-အိန္ဒိယ

- 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 (Unicode, IETF, etc)

- Help users of Web technology understand what's available to them and how to use it by developing best practices and other resources

# MultilingualWeb-LT Working Group

جعل شبكة الويب العالمية عالمية حقاً!

Համաշխարհային ցանցն իրոք համաշխարհային  
ՀԵՐԼ ΔΡΑΦΟΔΟ ΡΕΣΕΛΟΓΓΟ ΔΕΛΟΝ<.

"Дүниежүзілік торды" нағыз дүниежүзілік етем:  
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 ਬਣਾਉਣਾ !

缔造真正全球通行的万维网



- Define meta-data for web content (mainly HTML5) and "deep Web" content, for example a CMS or XML files from which HTML pages are generated, that facilitates its interaction with multilingual technologies and localization processes.
- Demonstrate interoperable implementations.

# Internationalization Interest Group

جعل شبكة الويب العالمية عالمية حقًا!

Համաշխարհային ցանցն իրոք համաշխարհային  
ՀՀԻԼ ΔΡΔΨΡΔΨ ϮԸΨΔԸԼԸΨ Δ.Ը.ՈՒՀ.

"Дүниежүзілік торды" нағыз дүниежүзілік  
वर्ल्ड वाईड वेबलाई यथार्थमै वश्वव्यापी ब

የዓለም አቀፉን ድር በእውነት አለም አቀፍ ማድረፉ

Κάνοντας τον Παγκόσμιο Ιστό πραγματικά Πα  
ਵਰਡ ਵਾਈਡ ਵૈਬ ੁੰ ਵਾਕਈ ਵਿਸ਼ਵ-ਵਿਆਪੀ ਬਣਾਉਣਾ !

缔造真正全球通行的万维网

ליצור מהרשת רשת כלל עולמית באמת!

'meikɪŋ ðə wɜːld waɪd web 'truːli 'wɜːld'wa

ワールド・ワイド・ウェブを世界中に広

ធ្វើឲ្យវិទ្យុសាស្ត្រពិភពលោកពិបាក

전세계의 월드 와이드 웹으로 만들기!

Gwneud y we fyd-eang yn wirioneddol fyd-ea

การทำให้ World Wide Web แพร่หลายไปทั่วโลกอย่าง  
འཛམ་གླིང་ཡོངས་འབྲེལ་འདི་ རྩ་མཐུན་རང་ འཛམ་གླིང་

• [www-international@w3.org](mailto:www-international@w3.org)

• [public-ir8n-bidi@w3.org](mailto:public-ir8n-bidi@w3.org)

• [public-ir8n-cjk@w3.org](mailto:public-ir8n-cjk@w3.org)

• [public-ir8n-indic@w3.org](mailto:public-ir8n-indic@w3.org)

• [public-ir8n-its@w3.org](mailto:public-ir8n-its@w3.org)

• [public-iri@w3.org](mailto:public-iri@w3.org)

Bidi



Bidi

Embedded text

TOP RATED RESTAURANTS

Aroma - 3 reviews

★★★★★

פיצה סגולה - 5 reviews

★★★★★

Roma פיצה סגולה - 3 reviews

★★★★★

TOP RATED RESTAURANTS

Aroma - 3 reviews

★★★★★

5 - פיצה סגולה reviews

★★★★★

Roma פיצה סגולה - 3 reviews

★★★★★

Bidi

Embedded text



```
<p><bdi>Aroma</bdi> - 3 reviews</p>
```

```
<p><bdi>PURPLE PIZZA</bdi> - 5 reviews</p>
```

```
<p><span class="name" dir="auto">Aroma</span> - 3 reviews</p>
```

```
<p><span class="name" dir="auto">PURPLE PIZZA</span> - 5 reviews</p>
```

Ruby

Ruby  
What is it?

にっぽん でん き  
日本 電気

N E C

# Ruby

## Requirements for Japanese Layout



[ [contents](#) ]

### Requirements for Japanese Text Layout

W3C Working Group Note 4 June 2009

#### This version:

<http://www.w3.org/TR/2009/NOTE-jlreq-20090604/>

#### Latest version:

<http://www.w3.org/TR/jlreq/>

#### Previous version:

<http://www.w3.org/TR/2008/WD-jlreq-20081015/>

#### Editors:

Yasuhiro Anan (阿南 康宏), Microsoft  
Hiroyuki Chiba (千葉 弘幸), Invited Expert  
Junsaburo Edamoto (枝本 順三郎), Invited Expert  
Richard Ishida, W3C  
Keiichi Ishino (石野 恵一郎), Antenna House  
Tatsuo Kobayashi (小林 龍生), JustSystems  
Toshi Kobayashi (小林 敏), Invited Expert  
Kenzo Onozawa (小野澤 賢三), Invited Expert  
Felix Sasaki, University of Applied Sciences Potsdam

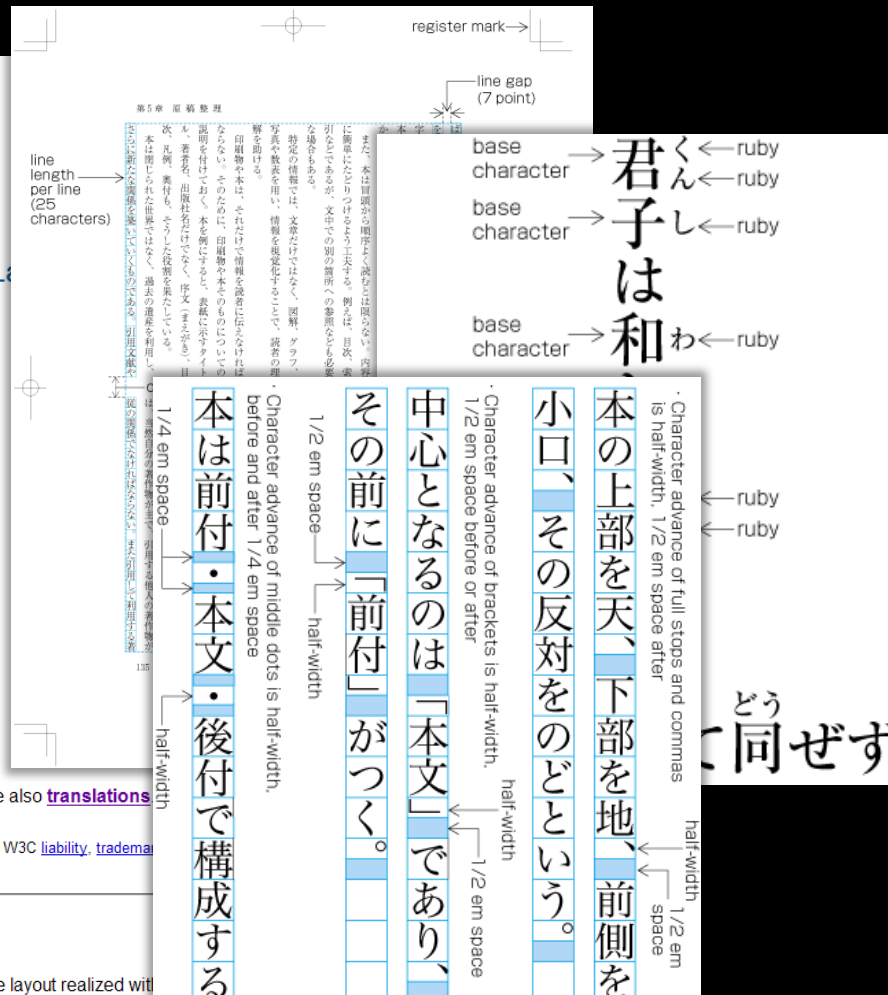
Please refer to the [errata](#) for this document.

A **Japanese version** of this document is also available. See also [translations](#).  
English version of this document is the authoritative version.

Copyright © 2008-2009 W3C® ([MIT](#), [ERCIM](#), [Keio](#)). All Rights Reserved. W3C [liability](#), [trademark](#), [document use](#) rules apply.

#### Abstract

This document describes requirements for general Japanese layout realized with technologies like CSS, SVG and XSL-FO. The document is mainly based on a standard for Japanese layout, JIS X 4051, however, it also addresses areas which are not covered by JIS X 4051.



# Ruby HTML5 innovations

## XHTML

```
<ruby>  
<rb>日本</rb>  
<rt>にっぽん</rt>  
</ruby>  
<ruby>  
<rb>電</rb>  
<rt>でん</rt>  
</ruby>  
<ruby>  
<rb>気</rb>  
<rt>き</rt>  
</ruby>
```

## HTML5

```
<ruby>  
日本  
<rt>にっぽん</rt>  
電  
<rt>でん</rt>  
気  
<rt>き</rt>  
</ruby>
```

Use case  
Ruby base styling

<ruby>

**B**

<rt>a</rt>

**B**

<rt>a</rt>

<ruby>

和文との混植こんしよくに用いる欧字

## Use case Ruby base styling

<ruby>

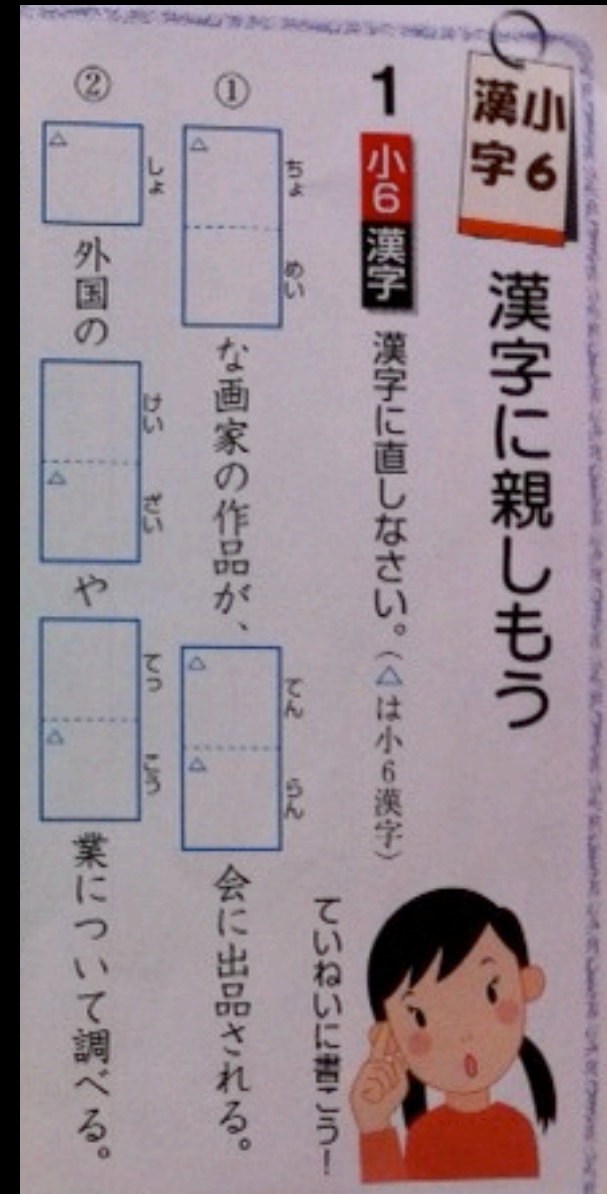
<span>**B**</span>

<rt>a</rt>

**B**

<rt>a</rt>

</ruby>





Use case

## Ruby base styling

- Use of span works.
- If used a lot, less obvious that this is better than rb.
- But allows styling within a base unit of multiple characters.
- Only works where span is present, so mass conversion for accessibility is not possible – but then neither does rb help if optional.
- Possible alternative, implicit rb element.

今日けふは会議かいぎの初日しよにちです。  
きょうはかいぎのしよ  
にちです。

Use case  
Fallback

<ruby>

**B**

<rp>(</rp>

<rt>a</rt>

<rp>)</rp>

**B**

<rp>(</rp>

<rt>a</rt>

<rp>)</rp>

<ruby>

今日は会議かいぎの初日です。  
今日は会（かい）議  
（ぎ）の初日です。  
今日は会議（かいぎ）  
の初日です。

Use case  
Fallback

<ruby>

**B**

<rp>(</rp>

<rt>a</rt>

<rp>)</rp>

**B**

<rp>(</rp>

<rt>a</rt>

<rp>)</rp>

<ruby>

今日は会議かいぎの初日です。  
今日は会（かい）議  
（ぎ）の初日です。  
今日は会議（かいぎ）  
の初日です。

Use case  
Fallback

```
<ruby>  
<rb>B</rb>  
<rb>B</rb>  
<rp>(</rp>  
<rt>a</rt>  
<rt>a</rt>  
<rp>)</rp>  
</ruby>
```

今日は会議かいぎの初日です。  
今日は会（かい）議  
（ぎ）の初日です。  
今日は会議（かいぎ）  
の初日です。

## Use case Fallback

- Should we go to a column-based, rather than row-based approach? This would involve introducing rb.
- Is there some other way of producing a fallback?
- Can we reach a situation quickly where fallbacks are not needed?

今日は会議の初日です。  
今日は会（かい）議（ぎ）の初日です。

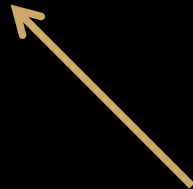
Translate flag

Translate flag

## Use cases

- some parts of texts should not be translated
- nowadays translation is usually automated thus there is a need for machine readable information about translatability

<p>Click the Resume button on the Status Display or the CONTINUE button on the printer panel.</p>



CONTINUE is a label on a hardware button and should not be translated as labels on hardware buttons are not going to be localized.

## Translate flag HTML5 solution

- `translate="no"` attribute can be added to any HTML element to indicate that content of element should not be translated

`<p>Click the Resume button on the Status Display or the CONTINUE button on the printer panel.</p>`



`<p>Click the Resume button on the Status Display or the <span translate="no">CONTINUE</span> button on the printer panel.</p>`



Translate flag

## Additional information

- translate attribute was recently added to HTML5 by effort started at the MLW Pisa workshop in 2011
- it is available in many other XML based formats thanks to ITS (Internationalization Tag Set)
- translate attribute can have **yes** and **no** values which can be overridden
- already supported by on-line translation services from Microsoft and Google
- already supported by content formats like DITA or DocBook

Getting involved...

# 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/>