Practical & Cultural Issues
in Web Design

Richard Ishida
W3C Internationalization Activity Lead

Objectives

W3C®

After this tutorial you should be able to say:

- ▶ identify a number of practices that affect easy localization of your Web pages

- ▶ why you need to think about localization earlier than people typically expect

- ▶ that you need to learn more about internationalization

slide 2

Terminology

W3C®

Localization

The **adaptation** of a product, application or document content to meet the language, cultural and other requirements of a specific target market.

Internationalization

The **design and development** of a product, application or document content that **enables** easy localization for target audiences that vary in culture, region, or language.

slide 3

Localization without internationalization can be very hard. This presentation will use examples to make that point, and stress the value of considering internationalization as an integral part of the design and development activity – not an afterthought left to the 'localization folks'.

Separating content & presentation

W3C®

Separating content & presentation
Linguistic issues
Cultural issues
Internationalization recap
Getting help

slide 4

This slide shows some bare HTML content, displayed using the browser's default styling.
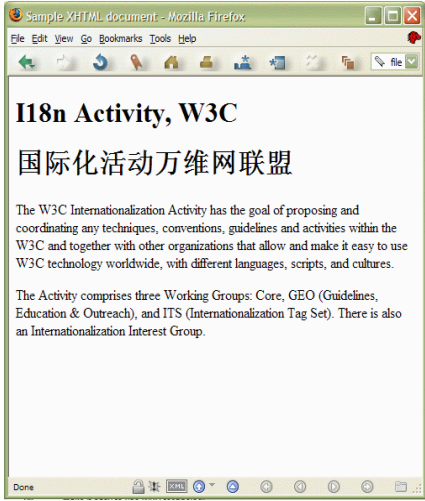
Separating content & presentation
The basics

W3C

Content (XHTML)

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
    "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">

<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en" lang="en">
  <head>
      <meta http-equiv="Content-Type" content="text/html; charset=utf-8"/>
      <title>About the W3C I18n Activity</title>
      <style type="text/css" src="mystyling.css" />
      </head>

  <body>
      <h1>I18n Activity, W3C</h1>
      <div class="international-text" xml:lang="zh-Hans"
              lang="zh-Hans">国际化活动万维网联盟</h1>
      <div class="description">
          <p>The W3C Internationalization Activity has the goal of proposing
              and coordinating any techniques, conventions, guidelines and
              activities within the W3C and together with other organizations
              that allow and make it easy to use W3C technology worldwide,
              with different languages, scripts, and cultures.</p>
          <p>The Activity comprises three Working Groups: Core, GEO
              (Guidelines, Education & Outreach), and ITS (Internationalization
              Tag Set). There is also an Internationalization Interest Group.</p>
          </div>
      </body>
  </html>
```

Presentation (CSS)

```
body {
          background: white;
          color: black;
          font-family: serif;
          font-size: 1em;
          }
h1 {
          font-size: 240%;
          }
div.international-text {
          font-family: MingLiu, sans-serif;
          font-size: 240%;
          }
p {
          margin-top: 1em;
          }
```

slide 6

The HTML is shown on the left. There is no presentational information in the HTML – which is as it should be.

To the right is some CSS code that applies styling to the HTML.

Each of these windows shows EXACTLY the same HTML file. The changes made to the CSS file produced three very different presentations of that basic content.

This is particularly useful for changing the presentational aspects of a site or group of pages. You typically only need to edit a single CSS file, rather than editing all the code of each HTML file.

This can also be beneficial for localization, since typographic approaches, colors, etc, may need to be changed for different locales. Making such changes in the CSS is much easier than adapting the HTML.

Separating content & presentation
The basics

W3C®

I18n Activity, W3C

The W3C Internationalization Activity has the goal of proposing and coordinating any techniques, conventions, guidelines and activities within the W3C and together with other organizations that allow and make it easy to use W3C technology worldwide

OK   Options

slide 8

Remember, also, that the Mobile Web is becoming increasingly important these days – and may be especially so in developing countries in the future. This means that content needs to be adapted to fit on handheld devices with smaller screens.

Again, this would ideally be achieved by styling the content, rather than writing a completely separate Web.

You should not make assumptions, when creating content, that you know what it will look like when finally displayed. These days, it may well be displayed in a number of different formats.

Separating content & presentation
International issues

W3C®

- ◆ problems of resolution to support bold and italics in small CJK characters on-screen

- ◆ different ways of emphasizing text in Japanese (wakiten & amikake)

これは日本語です。

これは日本語です。

Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 9

Here are some ways in which typographic differences may appear between language versions of the same content.

International issues

W3C®

- ◆ problems of resolution to support bold and italics in small CJK characters on-screen

- ◆ different ways of emphasizing text in Japanese (wakiten & amikake)

- ◆ no upper- vs. lower-case distinction in most non-Latin scripts

- ◆ no convention of distinguishing between proportional and mono-spaced fonts for some scripts

slide 10

Separating content & presentation
Practical implications

W3C®

Making the World Wide Web *worldwide*.

<p>Making the World Wide Web <i>worldwide</i></p>

Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 11

You should try to remove all presentational constructs from your content. For example, use of <i> tags shows that you are assuming that the text will be italicized. Because ideographic text doesn't support italicizations well in small font sizes, you could be causing problems for localization.

Practical implications

W3C®

Making the World Wide Web *worldwide*.

`<p>Making the World Wide Web <em>worldwide</em></p>`

✔

Copyright © 2005 W3C (MIT, ERCIM, Keio)   slide 12

Not only is it better for localization to express the idea or semantics in the content, and leave the presentation to the style sheet, it will also improve your original text by making you more aware of what you are actually doing.

Separating content & presentation
Practical implications

W3C

See the **System Administrator Guide** for an example of re-use.

`<p>See the <span class="bold">System Administrator Guide</span> for an example of re-use.</p>`

Copyright © 2005 W3C (MIT, ERCIM, Keio)   slide 13

The same applies to document conventions such as representation of referenced resources. When using class annotations or microformats, don't describe the expected presentational rendering, describe the function of the text.

Separating content & presentation
Practical implications

W3C®

See the **System Administrator Guide** for an example of re-use.

<p>See the System Administrator Guide for an example of re-use.</p>

doctitle
chaptertitle
inputsequence
etc.

slide 14

Setting document direction
How to mark up the document

W3C®

Ch◆teau de La Napoule

Uploaded on June 9, 2005
by r12a

ADD NOTE · SEND TO GROUP · ADD TO SET · BLOG THIS · ALL SIZES · ORDER PRINTS · ROTATE · DELETE ✗

— **r12a's photostream**

1177
photos
View as
slideshow

← more · browse · more →

Tags
- 0506-cote-dazur [x]
- Mandelieu-La Napoule [x]
- Alpes-Maritimes [x]
- France [x]

Add a tag

Additional Information
- © All rights reserved (change)
- Taken with a Fujifilm FinePix S7000.
  More properties
- Taken on June 6, 2005 (edit)
- See different sizes
- Viewed 92 times. (Not including you)
- Edit title, description, and tags
- NEW Replace this photo

slide 15

Linguistic issues

W3C®

Separating content & presentation
Linguistic issues
Cultural issues
Internationalization recap
Getting help

slide 16

Linguistic issues
Syntax differences

W3C®

They are speaking to her from my new house.

Están hablándole desde mi casa nueva.

私の新しい家から彼女と話しています。

تكلّموها من بيتي الجديد

slide 17

This slide shows the same idea expressed in multiple languages. Within each translation of the sentence, the number of words is different, and the order of those words changes.

Agreement

W3C

On

Connectado	-os -a -as
Encendido	-os -a -as
Activado	-os -a -as

slide 18

A word in one language may map onto more than one word in another, as seen in the Spanish translations for 'on'. In Spanish it depends what is on.

In addition, Spanish uses different endings according to whether the thing that is on is masculine, feminine, singular or plural.

Linguistic issues
Plurality

W3C®

Number of pages	Russian word for 'page'
1	страниц**а**
2-4	страниц**ы**
5-10	страниц
11-20 (irregular)	страниц
21	страниц**а**
22-24	страниц**ы**
25-30	страниц
> 30	Repeat pattern of endings for 1 to 10.

slide 19

Plurality is much more complicated in Arabic and Russian than it is in English. The shape of the Russian word for 'page' changes according to the number in a (mostly) cyclical way.

Text expansion

W3C®

Input processing
features

Eingabeverarbeitungsfunktionen

slide 20

Text in English and Chinese is typically much more concise than text in other languages. In some languages, such as Finnish and German, this tendency is worse in short English noun phrases that become a single word in translation.

Linguistic issues
Sentence-like composite messages

W3C®

The < > has been disabled.

printer   stacker

stapler options

slide 21

In this example, the developer has tried to save memory by re-using part of a common sentence. Unfortunately, because of the effects of rules about agreement between gender and number in many languages, this becomes an untranslatable phrase. The developer needs to be aware of the likely impact on translatability of such things.

Linguistic issues
Topic-comment composite messages

W3C

Printer: ⊙ enabled
Stacker: ⊙ enabled
Stapler options: ○ enabled

Imprimante: ⊙ activée
Module de reception: ⊙ activé
Options d'agrafage: ○ activées

http://www.w3.org/International/articles/composite-messages/

Copyright © 2005 W3C (MIT, ERCIM, Keio)   slide 22

Composite messages in topic-comment arrangements typically cause no issues for translation (as long as you don't re-use the comment string in multiple contexts).

Text string re-use

W3C®

```php
<?php
function getStatus (device) {
        if ($device.status != 0) { return 'On'; }
        else { return 'Off'; }
        }

?>
…
<p>Printer: <?php print getStatus($printer); ?></p>
<p>Stacker: <?php print getStatus($stacker); ?></p>
<p>Stapler options: <?php print getStatus($soptions); ?></p>
```

http://www.w3.org/International/articles/text-reuse/

slide 23

This shows an example where text-string re-use will cause significant localization problems, due to the word mapping and agreement features of many languages.

The code prevents the use of multiple translations for the words 'on' and 'off'.

Linguistic issues
Variable variable positions

W3C®

There were %d spelling mistakes in file %s.

Datei %s enthält %d Rechtschreibfehler.

printf("There were %d spelling mistakes in file %s.",
currentpage, totalpages)   ✗

printf("There were %1\$d spelling mistakes in file %2\$s .",
currentpage, totalpages)   ✓

printf("Datei %2\$s enthält %1\$d Rechtschreibfehler.",
currentpage, totalpages)

slide 24

This is an example of syntax differences affecting development techniques.

The order of variables needs to be different between English and German versions. Unless you are using slightly more advance techniques in PHP, you will prevent this possibility and seriously affect translatability.

English and Chinese text usually expand when translated. You should consider the potential impact of this on page design, and either allow text to flow into larger areas, or leave expansion space.

For example, putting labels beside form fields is often likely to cause expansion space problems. This issue can often be avoided by allowing text to expand above the field, instead.

Cultural issues

W3C®

Separating content & presentation
Linguistic issues
Cultural issues
Internationalization recap
Getting help

Formatting & reading data
Symbolism & examples
Color
Differing approaches
International text

slide 26

Setting document direction

W3C®

Formatting & reading data

slide 27

Formatting & reading data

W3C®

03/02/01

2.3.01

2001-III-2

2001年3月2日

Date:* ☐ ☐ ☐

slide 28

The expected format of data such as dates and times varies around the world. You should internationalize your deliverable so that it respects local customs.

Much more difficult than displaying a date correctly, however, is recognizing a date supplied by a user if you don't clearly signal the expected format.

Cultural issues
Formatting & reading data

W3C®

Chinese	Chang Wen Yun
Malay	Isa (bin) Aman
Indian	M. Thiruselvan
	Thiruselvan s/o Manickavasagam

slide 29

People's name formats can also be very different around the world. It is important to consider how you will ask for people's names in form design, but also how you will capture the information in your database.

Cultural issues
Formatting & reading data

W3C®

Россия
г. Пермь 614055
ул. Крупской 93-82
Селивановой Юлии

Country: United States
First name:
Last name:
Address:
City:
State: AZ
Zip code:
Telephone: (   )
Application date:

slide 30

Be careful about assuming what others' name and address formats will be.
Also think about how you will store the names and addresses in the database.
For example, do you really need to split out street number? How will you
generate a Russian or Japanese address that goes from general to specific
from top to bottom?

Cultural issues
Formatting & reading data

W3C®

- Numeric formats
- Calendars & date formats
- Address formats
- Name conventions
- Time formats
- Currency formats
- List separators
- Telephone numbers
- etc....

slide 31

Setting document direction

W3C
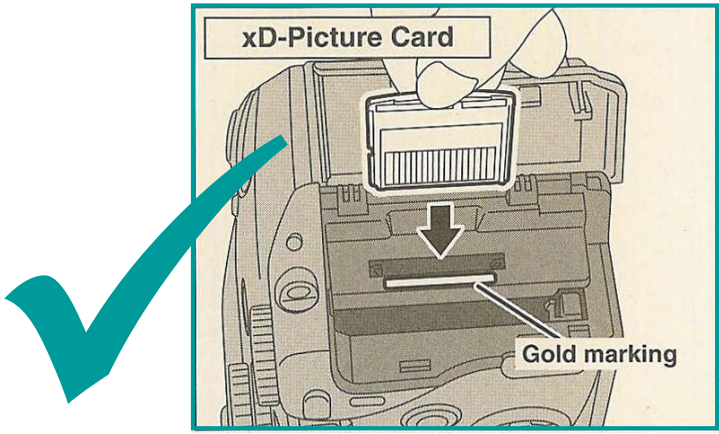
Symbolism & examples

slide 32

Cultural issues
Symbolism & examples

W3C®

xD-Picture Card

Gold marking

Copyright © 2005 W3C (MIT, ERCIM, Keio)   slide 33

Symbolism can differ from place to place. For example the check mark means *incorrect* in some places around the world.

Ensure that you do not give the wrong message through your use of colors, symbolism, examples, etc.

Symbolism can differ from place to place. For example the check mark means *incorrect* in some places around the world.

Cultural issues
Symbolism & examples

W3C

Copyright © 2005 W3C (MIT, ERCIM, Keio)   slide 35

Graphics may need to be changed if they don't reflect the local culture of certain places.

Cultural issues
Symbolism & examples

W3C®

slide 36

Body language and gestures are particularly dangerous. Each of these symbols can give offense in one part of the world or another.

Setting document direction
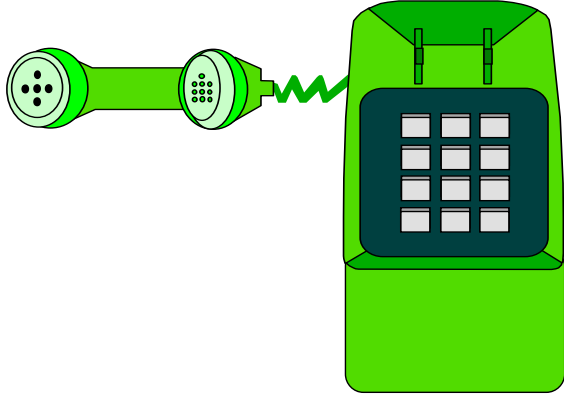
W3C

Color

slide 37

Color

W3C®
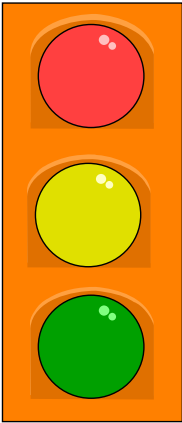
slide 38

Color can have strong subliminal suggestiveness due to conventional uses in particular cultures.

Also, color names can change from culture to culture. Germans, for example, don't have a strong concept of the color 'purple'.

Color also has different connotations in different parts of the world.

Cultural issues
Color

slide 41

W3C®

Differing approaches

slide 42
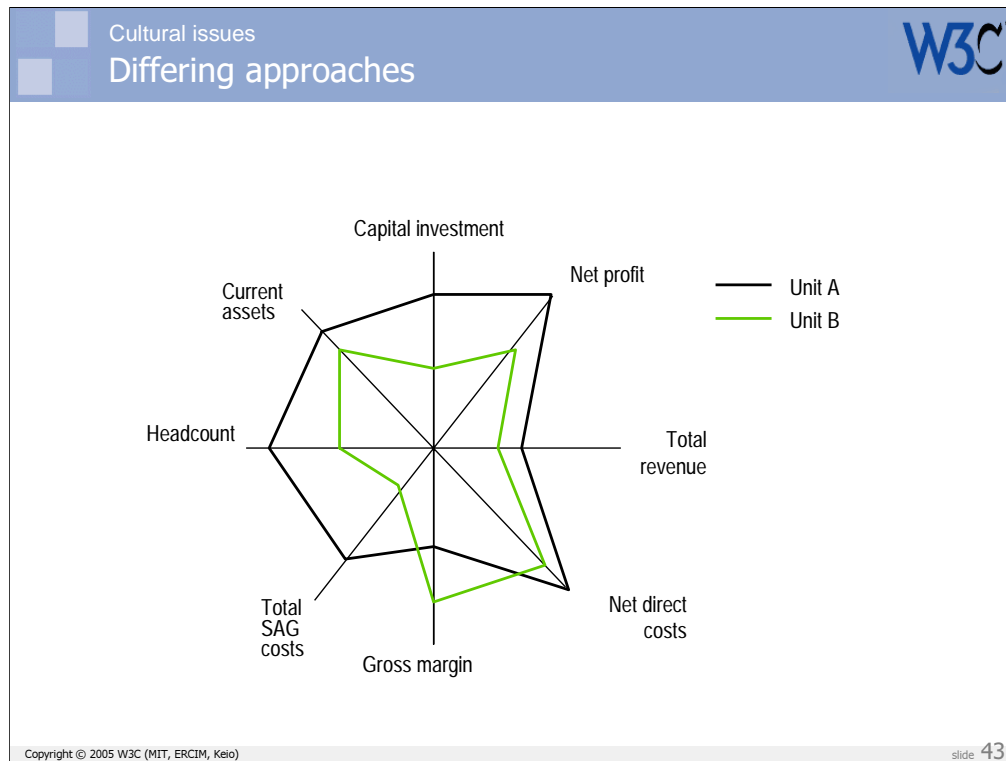
Then you need to be aware that people in different parts of the world may do things in different ways. For example, the radar chart was such a common way of representing comparative data in Japan that, when Lotus 1-2-3 was launched in that area they had to reengineer it to add that.

Cultural issues
Differing approaches

W3C®

"... one Latin American teacher recently complained to me that the US-manufactured and well-translated educational software currently being used in his country's primary schools presupposed 'solitary problem solvers', whereas his culture stressed collective problem-solving."

Kenneth Keniston,
Language International, May 1996

slide 44

Considerations of this kind require you to make big decisions at the very start of the development phase about how to proceed. Otherwise you could waste a lot of time and energy producing something that doesn't meet your customer's needs.

This and the following slides show how Yahoo adapts its categorizations to reflect the preoccupations of various different countries.

Setting document direction

W3C®

International text

slide 48

Cultural issues
International text

W3C®

The outcome was a heavy defeat for the new House Democractic leadership. Ways and means committee chairman Dan Rostenkowski looked particularly inept. He took a vacillating, off-again, on-again, stand and eventually lost control of his committee. Majority Leader Richard Gephardt, failed to recognize the strength of the drive for a capital-gains cut. Finally, but too late, he helped draft the Democratic alternative - which combined deductible IRA contributions for everybody with an increase in the tax rate on people with incomes over $200,000. The capital gains cut, he declared, was "designed to keep Leona Helmsley's dream alive - that only little people pay taxes." Republicans retorted, in effect, There you go again, proposing a tax *increase*. "Fish gotta swim, birds gotta fly, and Democrats gotta raise taxes" was the way one G.O.P. quip put it.

slide 49

The text in blue is difficult for a British or non-English speaking person to understand. Sometimes because of references to things that don't exist in other cultures, sometimes because of culture-specific knowledge.

Cultural issues
International text

W3C

Montre qui tu es !

The 5 R's

Nothing sucks like an Electrolux!
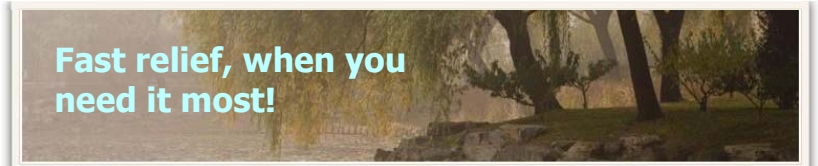
slide 50

You should check your text for things that might be interpreted differently in another culture.

When dealing with graphics, consider how to deal with text. Ideally the text
will be overlaid on a graphic, rather than embedded in it. If the text is within
the graphic, try to ensure that you develop it in layers, with text on a separate
layer, so that when it comes to translation the text can be easily removed and
replaced over complicated backgrounds.

Internationalization recap

W3C®

Separating content & presentation
Linguistic issues
Cultural issues
Internationalization recap
Getting help

slide 52

Internationalization recap

W3C®

Internationalization means:

- using a Quality approach to reduce the overall cost and time to market/release of multinational deliverables
- *designing* into the deliverable or product an internationalized base, and a modular and easily adaptable architecture
- not always doing extra work – maybe just working in a better way

slide 53

Getting help

W3C®

Separating content & presentation
Linguistic issues
Cultural issues
Internationalization recap
Getting help

slide 54

Getting help
GEO (Guidelines, Education & Outreach) WG

W3C®

How do I ...

- Ensure that XHTML forms return data in the right encoding?
- Make my Urdu, Arabic or Hebrew text display correctly?
- Declare language and encoding for XML documents?
- Order XSL output according to French rules?
- Approach the creation of multilingual documents in HTML?
- Help users navigate to the right localized page?
- Ensure the table I'm about to write has all the right i18n features?
- etc

slide 55

A range of helpful materials are available from the Internationalization home page.

We are building indices to help you locate the information more easily.

An example of an article.

Getting help
Getting involved

W3C

Get involved:

*   visit the I18n Activity Home Page

*   join a W3C Internationalization Working Group, or the Interest Group (www-international@w3.org)

*   offer to help with reviews, or provide local knowledge for other WGs

*   provide translations of W3C specifications or articles

*   take advantage of the i18n-readiness of W3C technology

slide 59

Getting help

And finally...

W3C

- ◆ this is your Web – not the W3C's – if something isn't right, get involved to fix it

slide 60