

# 新たな段階を迎える W3C 国際化活動 - ORF 2004

## by Martin Dürst

### Table of contents

- . 新たな段階を迎える W3C 国際化活動
- . Goal and Overview
- . What is Internationalization
- . Internationalized Resource Identifiers (IRIs)
- . Character Model
- . Web Services Internationalization: The Problem
- . Web Services Internationalization: Patterns
- . Web Services Internationalization: Progress
- . GEO Work
- . Internationalization Tag Set
- . New Activity Structure
- . Conclusions

## 新たな段階を迎える W3C 国際化活動

---

ORF 2004, November 24, Roppongi, Tokyo, Japan

Martin J. Dürst <[duerst@w3.org](mailto:duerst@w3.org)>

W3C *Internationalization Activity Lead*

*慶應義塾大学大学院政策・メディア研究科 助教授*

## Goal and Overview

---

Goal: Update on very recent developments

Overview:

- . What is Internationalization
- . Work being completed: IRIs, Character Model
- . Ongoing work: Web Services, GEO
- . New work: Internationalization Tag Set

## What is Internationalization

---

Software Internationalization: enabling the use of a technology with any language, script, and culture

Web Internationalization: not just single language/script/culture, but also mixed

Localization: actual configuration or adaptation of technology or content to a particular language and culture

I18N = I (18 characters) N = Internationali(s|z)atioN

Universal Access is one of the three goals of W3C

Internationalization is part of Universal Access (together with Web Accessibility, Device Independence,...)

Example: Translations

## Internationalized Resource Identifiers (IRIs)

---

- . Allowing Web addresses to contain a wide variety of characters
- . Well-defined conversion from IRIs to URIs (using UTF-8 and %HH-encoding)
- . Allow to use Internationalized Domain Names (IDN) in Web addresses
- . Current specification: *draft-duerst-iri-10.txt*, see *editing page* for updates
- . IETF process, W3C contributions
- . Specification status: Very close to being approved by the IESG
- . Related: Character Model: Resource Identifiers

Example 1, IDN: <http://納豆.w3.mag.keio.ac.jp/> (*fallback*)

Example 2, filename: <http://www.w3.org/International/articles/idn-and-iri/JP納豆/引き割り納豆.html> (*fallback*)

## Character Model

---

Main goal, very simple: Avoid mojibake

Audience: Mainly writers of specifications

Basic solution:

- . Label encodings
- . Convert to *Unicode*/ISO 10646 for processing

*Character Model for the World Wide Web 1.0: Fundamentals* published as Proposed Recommendation on Monday

*Character Model for the World Wide Web 1.0: Resource Identifiers* published as Candidate Recommendation on Monday

*Character Model for the World Wide Web 1.0: Normalization* is remaining work

## Web Services Internationalization: The Problem

---

- . Web Services defined simply:  
Using Web technology for machine-to-machine interaction
- . Problem: In actual use, users and their preferences cannot be ignored

## Web Services Internationalization: Patterns

---

- . Locale Neutral
- . Data Driven
- . Service Determined
- . Client Influenced

See *Web Services Conference talk* and *Usage Scenarios*

## Web Services Internationalization: Progress

---

### Accomplishments:

- . published *Web Services Internationalization Usage Scenarios* as a Working Group Note in July
- . published *Requirements for the Internationalization of Web Services* as a Working Group Note last week

### New work (proposed):

- . Specification for Web Services Internationalization (SOAP, WSDL)
- . Specification for Language and Locale Identifiers for the WWW

## GEO Work

---

GEO = Guidelines, Education, Outreach

Ongoing work to help people understand and use Web Internationalization features better

Producing Techniques documents, FAQs, articles, tutorials, and tests

Example: *Overview*, *FAQ on form input*

## Internationalization Tag Set

---

Problem: Different XML formats use different solutions for I18N issues (unfortunately, the most widespread approach is to not do anything)

Solution: Tag set for Internationalization/Localization

Tag examples:

- . `<span>` for internationalization attributes such as `xml:lang`
- . `<ruby>`
- . attributes for bidirectional embedding/override (for Arabic/Hebrew)
- . tag for unencoded characters
- . tag for glyph variants
- . translatability flag
- . notes to translators

## New Activity Structure

---

(not yet approved by W3C Members)

- . Internationalization Core Working Group (I18N Core WG)
  - Continue Reviews
  - Finish Character Model
  - Create Web Services Internationalization specification(s)
- . Guidelines, Education & Outreach Working Group (GEO WG)
- . Internationalization Tag Set Working Group (ITS WG)
- . Internationalization Interest Group (I18N IG)

## Conclusions

---

- . Enjoy internationalized Web technology
- . Help us internalizing the Web even better

Questions? Answers!