

Internacionalització

Interfícies d'usuari

Aleix Llusà Serra

Enginyeria de Sistemes TIC
Universitat Politècnica de Catalunya
<http://epsem.upc.edu>

30 de març de 2022

Usuaris que parlen idiomes diferents i tenen convencions culturals diferents

- Internacionalització (i18n): Preparar la interfície per a ser traduïda
- Localització (l10n): Adaptar a un idioma i cultura particular
 - Traduir paraules i interpretar conceptes
 - Convencions culturals
 - Estàndards nacionals
 - Context
- Local (locale): llengua/regió particular
 - IANA Language Subtag Registry + IETF BCP 47:
language-extlang-script-region-variant-extension
 - ca ca-valencia (ca-AD? ca-ES? ca-occidental? ca-principat? ca-illes?)
 - oc oc-aranes
 - pt-BR pt-PT en-GB en-US

Tot el text que sigui visible

- Nivell de components: `<p>Bon dia</p>`
- Nivell d'atributs: ``
- Nivell dinàmic: `variable = 'Bon' + 'dia'`
- Nivell de píxel:



- Missatges d'error

Sistema d'escriptura alfabètic

- Ciríl·lic Кири́лица
- Llatí
- Grec Ελληνικά
- Aràbic أَلْعَرَبِيَّةُ (abjad)
- japonès katakana カタカナ (sil·labari)
- xinès hànzì 漢字, japonès kanji 漢字 (logogrames)

- Esquerra a dreta, dreta a esquerra, dalt a baix
- Capgirar tota l'orientació: alineació a la dreta, botó de menú a l'altra banda, etc.
- Barreja de direccions en un mateix text. P.ex. nombres aràbics sempre esquerra a dreta
- També en l'edició de text
- Implementació de bidireccionalitat (Bidi): direcció base + direcció particular de components

- Llibreries d'strings ordenen segons taula de codificació (primer majúscules, després minúscules, etc.)
- Cal saber ordenar paraules segons les regles de cada idioma
- P.ex. llengües europees: solen ordenar sense tenir en compte les lletres modificades
- Però: en espanyol fins al 2010 els dígrafs “ch” i “ll” eren considerats lletres i afectaven l'ordenació; en canvi la lletra modificada “ñ” segueix sent una lletra i afecta l'ordenació
- Plurals: qüestió complicada
 - “Ha arribat 1 usuari”; “Han arribat 0 usuaris”
 - Llengües amb conceptes estesos de plural dual/trial/quadral

- Nombres: 12 345,67 12,345.67 12.345,67
- Dates: 31/03/2017 03/31/2017
- Unitats: 10 10ft 20°C 72°F

- Blanc: casaments (occident), funerals (orient)
- Vermell: casaments (orient)
- Verd: semàfor (occident)
- Blau: semàfor (orient)

- Poden no ser reconeixibles en totes les cultures. P.ex: símbol d'un endoll o símbol d'una bústia
- Jocs visuals: metàfores visuals amb polisèmies de les paraules. Desaconsellables perquè no funcionen en traduir.

- Separar els missatges localitzables del codi en fitxers a part per a facilitar la traducció (gettext)
- Dissenyar els missatges perquè no calgui llegir el codi (traducció per part de no programadors)
- Codificar les parts dinàmiques: “Han arribat o usuaris” “o users have arribed”
- Implementació d'estructura separada de la presentació. P.ex: HTML+CSS
 - Poder canviar icones i imatges
 - Poder canviar fonts per tenir els glifs propis de l'idioma
 - Poder canviar colors
- Unicode
 - 136000 caràcters
 - Caràcter: a Unicode unitat d'informació (grafema, símbol, síl·labi, també caràcters de control)
 - Codificació: representació d'un caràcter, típicament mitjançant bytes. UTF-8 UTF-16 UTF-32
 - Font: com s'han de dibuixar els caràcters a la pantalla
 - Glif: Visualització particular d'un caràcter en una font