Internationalization & Localization

SWEN-444

The Basics

- **Locale –** set of **linguistic** and **cultural parameters** associated with a **geographic region**

- **Localization** - the **process** whereby the software, documentation, and user interface are **adapted** to suit the needs of different world markets and users segments – a **locale**

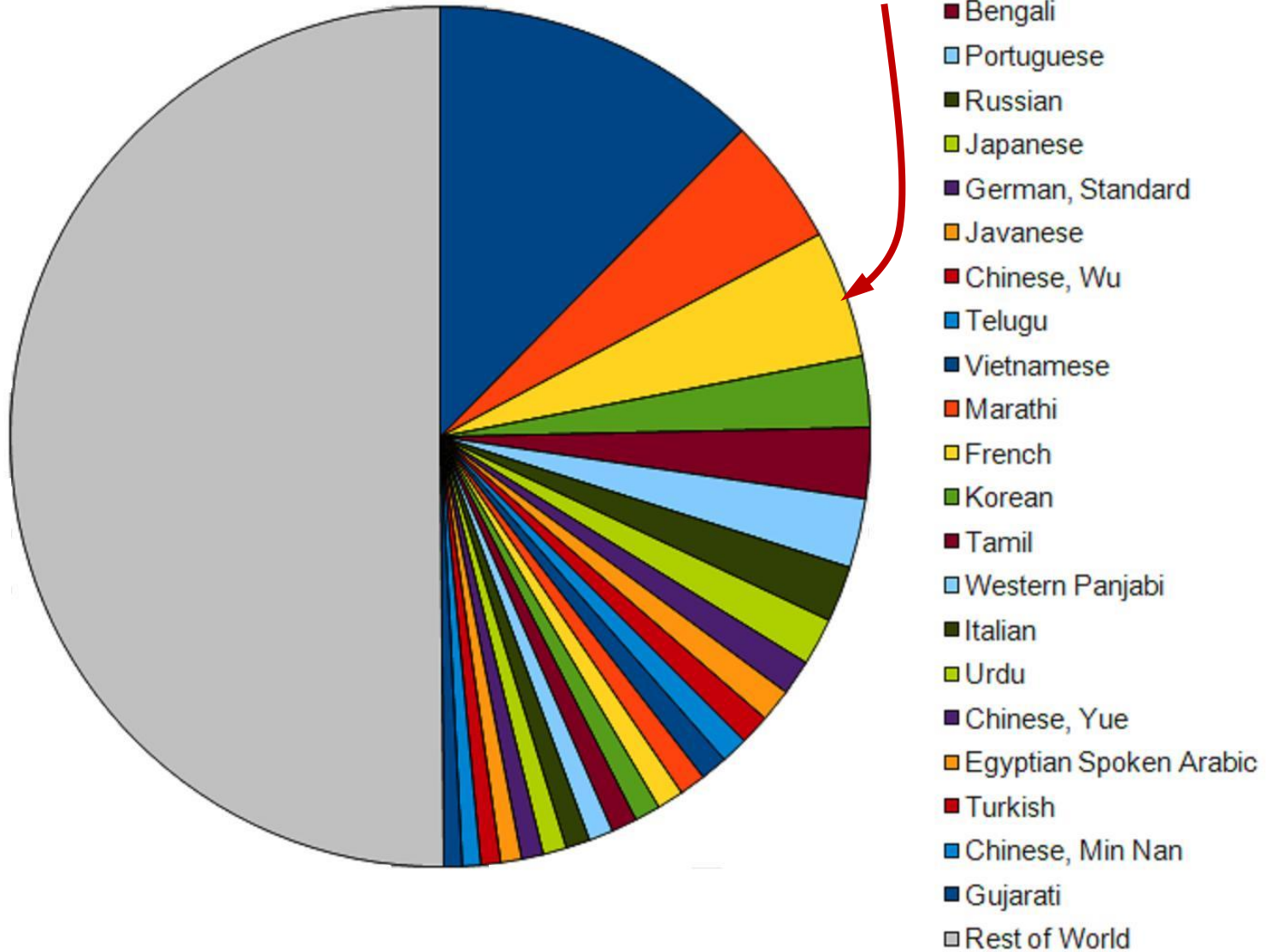- A single country may have multiple locales –example?

What is Culture?

"…the customary beliefs, social forms, and material traits of a racial, religious, or social group. *also* **:** the characteristic features of everyday existence (such as diversions or a way of life) shared by people in a place or time" Merriam-Webster

Why bother? – everyone speaks English

Top 25 World Languages

- Chinese, Mandarin
- Spanish
- English
- Hindi
- Bengali
- Portuguese
- Russian
- Japanese
- German, Standard
- Javanese
- Chinese, Wu
- Telugu
- Vietnamese
- Marathi
- French
- Korean
- Tamil
- Western Panjabi
- Italian
- Urdu
- Chinese, Yue
- Egyptian Spoken Arabic
- Turkish
- Chinese, Min Nan
- Gujarati
- Rest of World

English

Case Study - Facebook

- In 2007 Facebook was English only
- Mission – connect every person on the planet
- Translation tool – crowd source native speakers to translate the site (voting)
- Today – 111 languages with 40 more in process
- 1 billion of 1.7 billion Facebook users speak a language other than English
- Emotional impact of pride (and preservation) of culture

Internationalization

- Internationalization - the process of developing a software product whose core design does not make assumptions based on a locale

- Software design techniques:
  - **Abstraction** – separate code needed to support a locale
  - **Message files** – separate localizable text from source code; one file for each locale
    - 16 bit Unicode to support all languages
  - **Formatting libraries** – numbers, dates, …
  - **Replaceable UI elements** - fonts, colors, images, icons

Software Engineer Internationalization Responsibilities

- Discover international (cultural) **requirements**
  - Identify and understand locales
  - Different markets may require slight but important variations in product functionality
  - Validate suitability of the localized application for the targeted markets
- Verify correctness and consistency in **translations**
- Ensure translation neutrality to accommodate **multiple locales** within individual markets
- Extend usability **testing** to global population

Let's Explore Various Localization Facets

- Calendars
- Time and date formats
- Numeric superstitions
- Interpersonal interaction and gestures

- Religion conventions
- Numbers
- Symbols
- Text and language
- Color

Calendars

- There are many different calendars in use throughout the world

  – Some based on the relative movements of the moon

  – Some based on the relative movements of the sun

- The Gregorian calendar, also known as the Western or Christian calendar, is the internationally accepted civil calendar

A Sampling of Calendars

Japan	Japanese Solar	• The Gregorian calendar written in Japanese characters • Includes an era name in addition to a numeric year • The era name of a date is derived from the reigning emperor
China	Chinese	• A luni-solar calendar, elements from lunar and solar calendar • The year starts on the second new moon after the winter solstice • The counting of years and year cycles is complex
Buddhist Countries	Buddhist Era	• Differs from one country to another along with the recognized birth date of the Buddha • Thailand's calendar counts its years from January 1st, 543 B.C
Korea	Tangun Era	• Tangun was the legendary founder of the first Korean kingdom • Counts years from 2333 BC
Arabic Countries	Hijri	• Counts its years from the Gregorian year AD 622 • Uses 12 lunar months • Has a year of 353, 354, or 355 days

Time and Date Formats

- There are many variations on how dates and time are formatted
  - In the **United States** the date format is **mm/dd/yy**
  - In **Europe** the date format is **dd/mm/yy**
  - In **Japan** the date format is **yy/mm/dd**
    - date format of 2/3/10 is not untypical
    - relates to the year of the current emperor's reign
- Calculations of date and time need to consider that the **first two digits** of a date value **may not be the month**
- **USA - time** format is predominantly **12 hour**
  - Punctuated by AM or PM for before and after midday
- **Europe** - time format is predominantly **24 hour** (military time)
  - In some European Countries AM and PM are not understood

Short Date Formats

Short Date format	Short Date Sample	Century	Day Leading	Month Leading
Albania				
yyyy-MM-dd	1994-09-08	Yes	Yes	Yes
Argentina (Spanish)				
dd/MM/yy	08/08/94	No	Yes	Yes
Australia (English)				
dd/MM/yy	08/08/94	No	Yes	Yes
Austria (German)				
d.M.yyyy	8/9/1994	Yes	No	No
Belgium (Flemish)				
dd/MM/yy	08/09/94	No	Yes	Yes
Belgium (French)				
dd/MM/yy	08/09/94	No	Yes	Yes
Bolivia (Spanish)				
dd/MM/yy	08/09/94	No	Yes	Yes
Brazil (Portuguese)				
dd/MM/yy	08/09/94	No	Yes	Yes
Bulgaria				
yyyy-MM-dd	1994-IX-08	Yes	Yes	Yes
Canada (English)				
dd/MM/yy	08/09/94	No	Yes	Yes
Canada (French)				
yyyy-MM-dd	1994-09-08	Yes	Yes	Yes
Chile (Spanish)				
dd/MM/yy	08/09/94	No	Yes	Yes
Colombia (Spanish)				
dd/MM/yy	08/08/94	No	Yes	Yes
Costa Rica (Spanish)				
dd/MM/yy	08/09/94	No	Yes	Yes
Croatia				
yyyy-MM-dd	1994-09-08	Yes	Yes	Yes
Czech Republic				
yyyy-MM-dd	1994-09-08	Yes	Yes	Yes
Denmark				
dd-MM-yy	08-09-94	No	Yes	Yes
Dominican Republic (Spanish)				
dd/MM/yy	08/09/94	No	Yes	Yes
Ecuador (Spanish)				
dd/MM/yy	08/09/94	No	Yes	Yes
Estonia				
dd-MM-yyyy	08-09-1994	Yes	Yes	Yes
Finland				
d.M.yyyy	8.9.1994	Yes	No	No
France				
dd/MM/yyyy	08/09/1994	Yes	Yes	Yes

Long Date Formats

Australia (English)	dddd,d MMMM yyyy
Austria (German)	dddd,dd. MMMM yyyy
Belgium (Flemish)	dddd d MMMM yyyy
Belgium (French)	dddd d MMMM yyyy
Brazil (Portuguese)	dddd,d' de 'MMMM' de 'yyyy
Canada (English)	dddd d MMMM yyyy
Canada (French)	d MMMM,yyyy
Czech Republic	dd MMMM yyyy
Denmark	d. MMMM yyyy
Finland	d. MMMMta 'yyyy
France	dddd d MMMM yyyy
Germany (German)	dddd,d. MMMM yyyy
Hungary	yyyy. MMMM d.
Ireland (English)	dddd,d MMMM yyyy
Italy	dddd d MMMM yyyy
Japan	'yyyy'' 'M'' 'd'' 'WW'
Liechtenstein (German)	dddd,d. MMMM yyyy
Luxembourg (French)	dddd d MMMM yyyy
Luxembourg (German)	dddd,d. MMMM yyyy
Mexico (Spanish)	dddd d' de 'MMMM' de 'yyyy
Netherlands	dddd d MMMM yyyy
New Zealand (English)	dddd,d MMMM yyyy
Norway (Bokmal)	d. MMMM yyyy
Norway (Nynorsk)	d. MMMM yyyy
Poland	d MMMM yyyy
Portugal (Portuguese)	dddd,d' de 'MMMM' de 'yyyy
PRC (Simplified Chinese)	'yyyy' 'M' 'd' 'WW'
Russia	d MMMM yyyy '.
South Korea (Korean)	'yyyy'' 'M'' 'd'' 'WW'
Spain (Modern)	dddd d' de 'MMMM' de 'yyyy
Spain (Traditional)	dddd d' de 'MMMM' de 'yyyy
Sweden	' den 'd MMMM yyyy
Switzerland (French)	dddd,d. MMMM yyyy
Switzerland (German)	dddd,d. MMMM yyyy
Switzerland (Italian)	dddd,d. MMMM yyyy
Taiwan (Traditional Chinese)	dddd,MMMM dd,yyyy
United Kingdom (English)	dd MMMM yyyy
United States of America (English)	dddd,MMMM dd,yyyy

Label Date and Time Fields

- In the United States, a date like 5/2/41 means May 2, 1941; in much of the rest of the world it means February 5, 1941. Globalization cannot work with such ambiguity. Make clear what is meant, as below.

Enter date: (mm/dd/yyyy) [_____]	March 10, 2005
Date of Birth: [mm/dd/yyyy] (eg.03/10/1970) required	10 March 2005
Not needed after: Month [03 ▾] Day [10 ▾] Year [2005 ▾]	10-Mar-2005

- Use a **24-hour clock** instead of A.M. and P.M.
- **Identify time zone**, e.g., GMT (Greenwich Mean Time (default)).

Software Engineering
Rochester Institute
of Technology

Differing Requirements - Numeric Superstitions

- **Lucky numbers**
  - 3, 8, 168, 518, 888 in traditional Chinese belief
  - 7 in most countries in the world
  - 8 in Hong Kong, Taiwan and Korea
  - 9 in Thailand
- **Unlucky numbers**
  - 2, 514 in traditional Chinese belief
  - 4 is an unlucky number in Hong Kong, Korea and Taiwan
  - 4 in China and Japan (suggests death)
  - 7 in East and West Africa
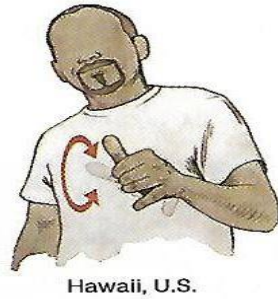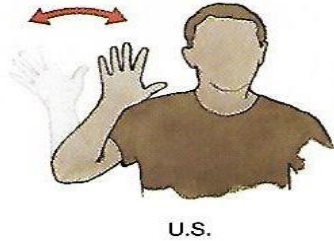 - 13 in most countries in the world
  - 6 in Thailand
- **Others**
  - 666 symbolizes evil to many Christians

Interpersonal Interaction

- The "OK" sign in the USA is an obscene gesture in Brazil.
- Waving the entire hand means
  - "Goodbye" in the UK
  - "No" in Japan
  - "Come here" in Peru
- In Asia and some African countries one gives gifts with both hands
- Direct eye contact means honesty and candor in Western Europe
  - In some Asian and African cultures direct eye contact suggests rudeness

Different motions are used to convey the same meaning, depending on culture and geography. But exceptions abound, even within a country.
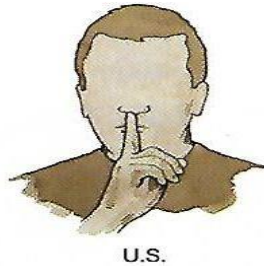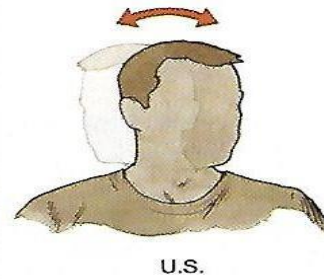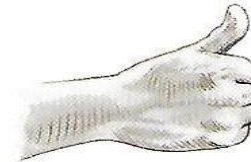
Hello

U.S.

Hawaii, U.S.

India

Come!

U.S.

Turkey

Don't tell

U.S.

U.K.

Look!

U.S.

Indonesia

No

U.S.

Greece
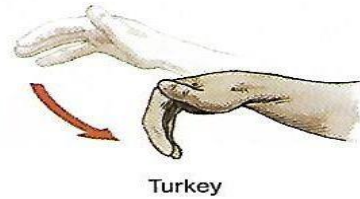
These gestures may look the same, but their meanings vary greatly based on how and where they're used.

Country	Meaning
U.S.	Rock on!
Brazil	Good luck!
Italy	Vulgar insult

Country	Meaning
U.S.	Great!
Germany	One
Japan	Five
Afghanistan	Vulgar insult

Country	Meaning
U.S.	OK
Japan	Pay me
France	Worthless
Russia	Vulgar insult

ART: JOE MCKENDR

16

Differing Requirements - Problematic Gestures

• Blinking the eye	Hong Kong, Taiwan
• Backslapping	India, some European countries
• Prolonged eye contact	Asian cultures
• Sticking the tongue out	Many cultures
• Touching someone's head	Fiji, Indonesia, Thailand, Singapore
• Folded arms	Fiji, Finland
• Slapping arm, fist raised	Most European countries
• Closed fist salute	Many countries
• "Stop" gesture	Greece, Nigeria
• Hands on the hips	Argentina
• Slapping fist	Chile, Italy
• Crossed fingers	Paraguay
• 1st and 4th fingers extended	Many countries
• Pointing with the index finger	Belgium, most of Asia

Differing Requirements - Religious References

• The Christian Cross	Christianity
• The Menorah	Judaism
• Star of David	Judaism
• The Crescent	Islam
• Buddha	Buddhism
• Pagoda, Dagoba, Stupa	Buddhism
• Wheel	Buddhism
• Lingam	Hinduism
• Pentacle	Paganism
• Torii	Shintoism

Differing Requirements - Sacred Animals and Plants

- Cows   Hinduism, Buddhism
- Monkeys   Hinduism
- Serpents   Hinduism
- Lotus flowers   Buddhism
- Chrysanthemum flowers   Japanese belief

Numbers

- **Punctuation**
  - In France: 1.234,56
  - In the United States: 1,234.56
- In expressing **currency**, include:
  - The country
  - The numeric amount
  - The appropriate currency symbol
- In giving **sizes**, state **units**:
  - Inches, feet, pounds, quarts, etc.—the English system (which is not used in England)
  - Meters, centimeters, grams, liters, etc.—the metric system (which is used in England)
- **Addresses**
  - Forcing international users to supply a state and ZIP code is confusing and/or insulting
- **Phone numbers**
  - Regional and country conventions for country code – area/zone code – subscriber number

International Signs

5 Germany – no public urination
6 France – no unleashed dogs
7 Ireland – sudden drop off
8 Australia – speeding endangers cassowaries
9 US – beware of RV mirror
10 France – ski lift how-to
11 Canada – don't eat shellfish
12 Jamaica – speed bump
13 Canada – log in water may shift in a storm
14 Brazil – some bus seats reserved for obese riders
15 Austria – no sledding
16 Switzerland – skiers beware
17 South Africa – unauthorized vendors can't sell food
18 Canada – avalanche danger
19 U.K – elderly crossing
20 Cambodia – toilet etiquette

Doug Lansky Exhibition

Emojis

Other International Symbols [Standards]

- Orientation – direction, coordinate system
- Communication – e.g., traffic signs
- Science and mathematics
- Accessibility

Language Dependent Text Considerations

- **Direction**
  - **Top to bottom characters**, **right to left lines** (Korean, Japanese, Chinese ideograms)
  - **Right to left characters**, **top to bottom lines** (Arabic, Hebrew)
  - **Left to right characters, top to bottom lines** (Western style)
- **Space, word delimiters, and alignment – word size** (German especially)
- **Diacriticals** provide meaning; E versus È
- **Collating sequences** – different alphabets impact sorting

Language Dependent Text Considerations

- ## Text direction but also layout is mirrored

Right-to-left (below)
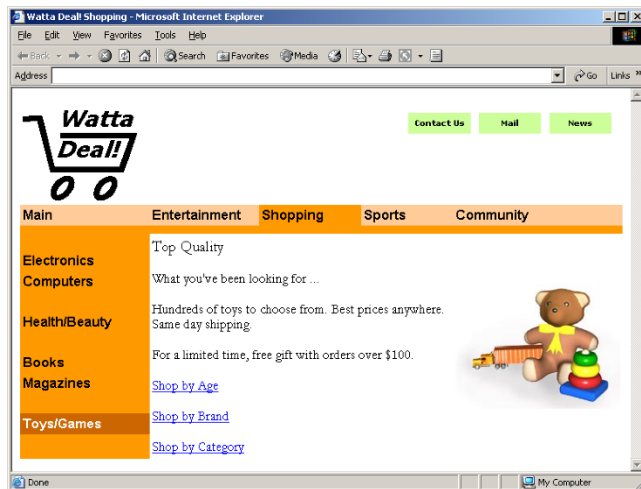
Left-to-right (above)
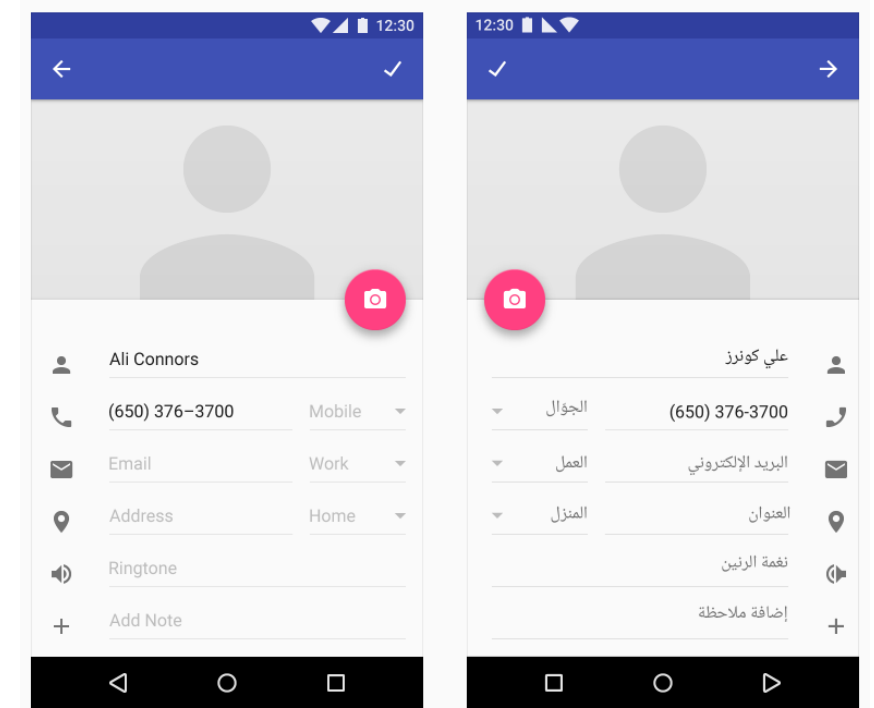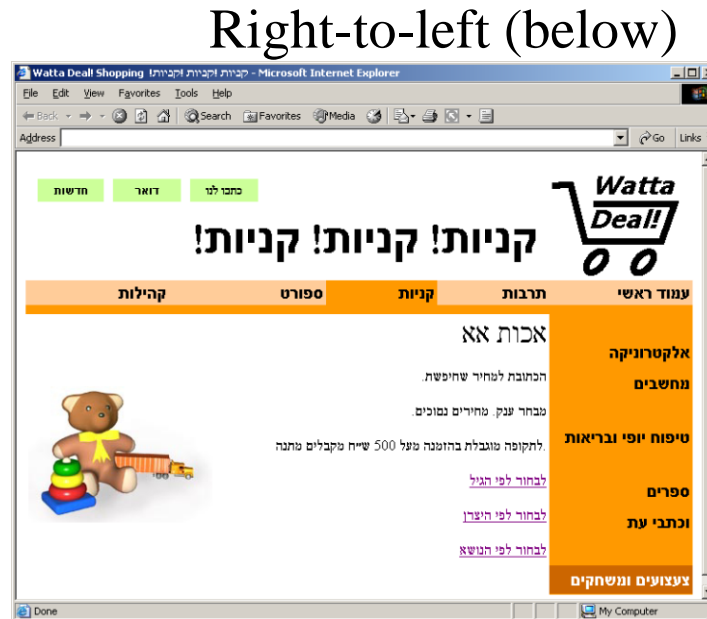
Language Dependent Text Considerations

- **Word Size**
- German word "Systemsteuerung" is nearly twice as long as its English-language equivalent, "settings"

Rechtsschutzversicherungsgesellschaften

Language Dependent Text Considerations

- Sorting

English	Slovak	Danish
Århus	Århus	Cleveland
Chlmec	Cleveland	Chlmec
Cleveland	Houston	Houston
Houston	Chlmec	Zürich
Zürich	Zürich	Århus

Language Dependent Text Considerations

- Example: UI Text

  - {$name} + "added" + {$number} +
  -       "person to his circles"

- Problems:
-    Concatenation (order can be different…)
-    Plural Agreement (if $number > 1)
-    Gender Agreement (if $name is female)
-    Non-Externalized Strings

Text Translation

- Recommendation: use **human translators** exclusively if possible
  - Use **semiautomatic translation** with **human oversight**
- **Avoid parochial references** that may mean little to a person who is not familiar with the subject
  - Jargon and slang, puns, which translate poorly
  - Be wary of metaphors, such as sports
- **Machine translation** may work better for **technical topics**, such as medical reports, that have a **standard vocabulary** that can be built into the translation dictionary

Cultural Differences and Use of Language

- Americans, as a rule, prefer direct statements: "Buy Now!"
  - That would be quite rude in some cultures, such as Japan
- Therefore the translator must be a person who has lived in the country of the target language, and preferably be a native speaker

Getting it right does matter

Google Maps in Sweden
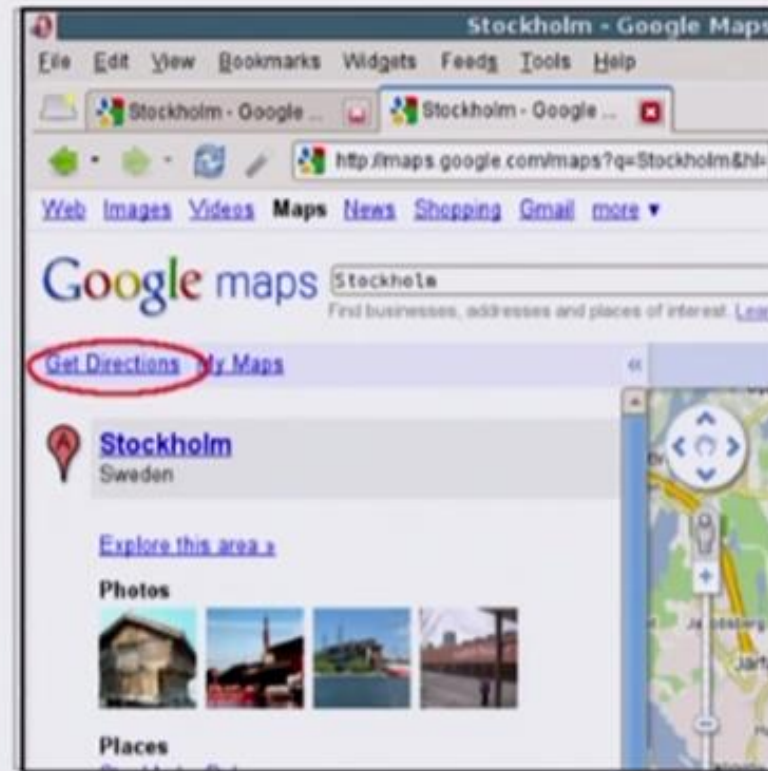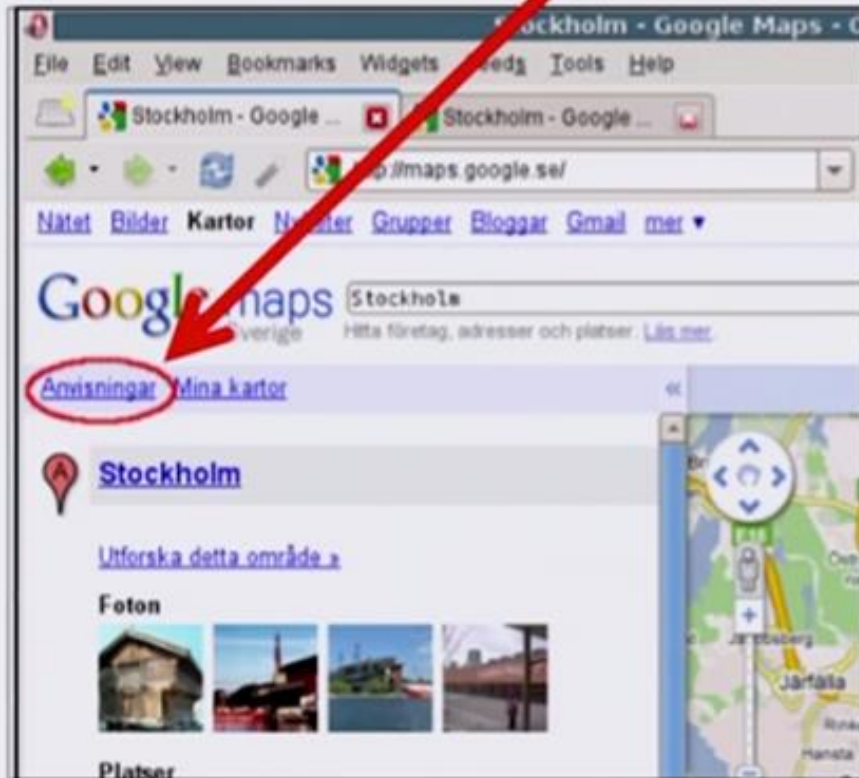
Bad Translations Have Many Origins

- In an Acapulco hotel: "The manager has personally passed all the water served here."

- At a Rhodes tailor shop: "Order your summers suit. Because is big rush we will execute customers in strict rotation."

- An American T-shirt maker in Miami printed shirts for the Hispanic market to commemorate the Pope's visit. Instead of "I Saw the Pope" ("el Papa"), the T-shirts read "I Saw the Potato" ("la papa").

- Pepsi's "Come Alive with the Pepsi Generation" slogan translated into Chinese as "Pepsi Brings Your Ancestors Back From the Grave."

- Airline ticket office, Copenhagen: "We take your bags and send them in all directions."

Color

- Localization – differences by culture or commercial context (logo color) in the meaning and emotion of color

- **Different meaning in different cultures**
- **Favorite example:**
  - Brides in western countries wear white
  - In China, white is the color of death
- **Get local expert advice**

Colours In Culture

David McCandless & AlwaysWithHonor.com // v1.0 // Apr 09 // InformationIsBeautiful.net

source: Pantone, ColorMatters & web

A Western / American
B Japanese
C Hindu
D Native American
E Chinese
F Asian
G Eastern European
H Muslim
I African
J South American

1 Anger
2 Art / Creativity
3 Authority
4 Bad Luck
5 Balance
6 Beauty
7 Calm
8 Celebration
9 Children
10 Cold
11 Compassion
12 Courage
13 Cowardice
14 Cruelty
15 Danger
16 Death
17 Decadence
18 Deceit
19 Desire
20 Earthy
21 Energy
22 Erotic
23 Eternity
24 Evil
25 Excitement
26 Family
27 Femininity
28 Fertility
29 Flamboyance
30 Freedom
31 Friendly
32 Fun
33 God
34 Gods
35 Good Luck
36 Gratitude
37 Growth
38 Happiness
39 Healing
40 Healthy
41 Heat
42 Heaven
43 Holiness
44 Illness
45 Insight
46 Intelligence
47 Intuition
48 Religion
49 Jealousy
50 Joy
51 Learning
52 Life
53 Love
54 Loyalty
55 Luxury
56 Marriage
57 Modesty
58 Money
59 Mourning
60 Mystery
61 Nature
62 Passion
63 Peace
64 Penance
65 Power
66 Personal power
67 Purity
68 Radicalism
69 Rational
70 Reliable
71 Repels Evil
72 Respect
73 Royalty
74 Self-cultivation
75 Strength
76 Style
77 Success
78 Trouble
79 Truce
80 Trust
81 Unhappiness
82 Virtue
83 Warmth
84 Wisdom

Grey
Yellow
Silver
Gold

http://www.informationisbeautiful.net/visualizations/colours-in-cultures/

Color

- **Microsoft**

  - When coloring in 800,000 pixels on a map of India, Microsoft colored eight of them a different shade of green to represent the disputed Kashmiri territory.

  - The difference in green meant Kashmir was shown as non-Indian, and the product was promptly banned in India.

  - Microsoft was left to recall all 200,000 of Windows 95 and took a hit to its relations with India.

Evaluate Your Project

- How well does it meet universal usability needs?
- How would you need to change the design to improve it?

References

- Localization Testing - www.sasqag.org/pastmeetings/LocalizabilityTesting.ppt
- Planning a Global Release - download.microsoft.com/download/e/0/1/e012808f-05e1-45d0-8965-1437a58d3466/ARC303.ppt
- Material from: Pat O'Sullivan "A Software Model Approach to Accommodating Cultural Diversity in the Development of Multilingual Applications"
- Dr. O. De Troyer, Localization Slides - wise.vub.ac.be/downloads/courses/**user**aspects/chapter5.ppt