All About Internationalization

▌ Overview

If you plan to offer your product in other languages, "internationalization" is a critical element in your overall development process.

Internationalization is the first step of a two-step process. It consists in generalizing your product to support as many languages as required. The second step - localization - consists in adapting the product to meet the needs of different languages and cultures. Internationalization reduces cost and time-to-market by making localization easier and avoiding work duplication.

Unfortunately, the complexity of internationalization is often underestimated, resulting in missed deadlines and cost overruns. This complexity arises primarily from the large number of issues involved and from their inherent pervasiveness: language issues can potentially affect every system, every component, every third-party tool, every line of code, every document, every help file, every test script, and every business process involved in releasing your product to the world.

Our workshop is designed to deliver maximum value in a minimum amount of time. The class is hosted at the location of your choice, either at your site or at a remote conference facility to minimize distractions. It is now also available in a self-paced eLearning format allowing organizations to train all new and old staff, at any time, at any place, at any pace.

▌ Target Audience

- The first day is intended for all audiences: Software Architects/Managers, Software Engineers, Test Engineers, Tech Writing, Product Managers, Solution Managers, User Experience, Release Management, Customer Support, Directors, and any other stakeholders. Technical concepts are presented in a non-technical manner using visually intuitive and compelling graphics.

- The remaining time is targeted more for technical staff, but much of it remains understandable by all. Software architects, managers and engineers, as well as product managers should attend.

▌ Benefits

This workshop will prepare you for all aspects of an internationalization project. You will know the issues, you will know the pitfalls and you will know the solutions. The workshop will provide you with a clear understanding of industry best practices, how to apply them and what their benefits are.

All attending stakeholders within your organization will share a *common terminology* which is the most critical ingredient to enable clarity of design, good communication and true understanding.

▌ Duration

The agenda described below is for a 3-day session. The first part (chapters 1 to 6) can also be used as an independent introductory module for all stakeholders. In its eLearning format, the course lasts roughly 12.5 hours, and can comfortably be taken in 10 sessions of 60-90 minutes each.

▌ Pre-requisites: none

▌ Agenda

PART 1- for all audiences: executives, product managers, developers, testers, support…

1. The Problem Domain: What Are We Up Against?

The characteristics of different languages and writing systems are presented visually. Particular attention is paid to the requirements of Asian and Arabic languages. Differences in cultural conventions are also highlighted. This chapter tells us WHAT issues programmers, webmasters and managers should be concerned with (and often are not aware of).

- The globalization dimension; languages and scripts
- Asian & Middle-Eastern languages
- Visual presentation of linguistic & cultural differences that impact development
- Text processing: searching, sorting, transcoding, etc.
- Cultural formats: date, time, calendars, currency, etc.

2. Key Globalization Concepts

Internationalization and Localization. Translation and what translators need. Levels of internationalization complexity. This chapter tells us WHY we should be concerned with the issues presented above, i.e. WHY we should prepare for translation.

- Globalization, Internationalization, Localization, Translation
- Internationalization complexity: dimensions and levels
- Ensuring translation quality
- Maintaining a global product: localizing the next version!
- Optimizing localization: glossaries & translation memories

3. A Visual Model of Internationalization

An animated graphical model that illustrates the three basic tasks of internationalization: resource externalization, cultural generalization and preparation of the localization kit. Each of the three tasks is explored in more detail. This chapter shows us HOW, in general terms applicable to both software and Web sites, to prepare for translation

- Graphical model of core internationalization tasks
- Resource externalization
- Cultural generalization
- Localization kit generation and strategy

4. Managing Internationalization

From the business case to customer support, through requirements, architecture & analysis, development and testing. Avoiding the pitfalls; managing risk.

- Business case and ROI
- Multilingual requirements
- Risk management: the "baseline and compare" strategy
- Evaluation & planning;
- Roles and responsibilities
- Common myths and pitfalls

5. Do's and Don'ts

Numerous practical examples from real projects, visually illustrated.

- Dimensions of evil: concatenation, alignment, abbreviation
- Pluralization and message formatting
- Indexes and sorted lists
- Clueless Manipulation

6. Testing internationalization

A brief and visual overview of internationalization testing issues for all (programmers, managers, etc.)

- A model of multilingual testing
- Testing resource externalization & pseudo-translation
- Testing cultural generalization & regression tests and pseudo-translation
- Testing the localization kit/infrastructure and localization process
- Managing multilingual testing: platform coverage and bug workflow

PART 2 - for technical staff: architects, programmers, web developers, test script writers…

7. Coded Character Sets: Past, Present and Future

- Single-byte character sets
- Multi-byte Asian character sets on Windows and Unix
- Unicode: a 21-bit character set

8. Unicode: Character Set & Standard

- The 10 Unicode design principles
- The Unicode standard, annexes and technical reports
- UCD: Unicode Character Database
- Unicode resources: tools, books, etc.

9. Unicode Representation

- Standard encodings: UTF-8, UTF-16, UTF-32
- Encoding size comparisons
- Serialization & the Byte-Order Mark (BOM)
- Compression schemes: SCSU & BOCU-1
- Data stability: Unicode normalization
- The four Unicode normalization forms

10. Unicode Implementation

- Unicode books, tools and libraries
- Transcoding and its complexity for bidirectional languages
- Text Boundaries: Grapheme Clusters, Words, Lines, Sentences
- Collation-Based Processes, Multilingual Collation Standards
- Data structures for efficient Unicode processing

11. Locales

- Locale identification: BCP 47
- Locale architecture (locale vs. language and country)
- Locale negotiation
- Complexities of language selection & suggested strategies
- Practical advice

12. Resources

- Resource organization & naming strategies
- The localization process
    - inventory resource types & simplify
    - create extraction and merging tools
    - pseudo-translation
- Practical advice

13. Input and Output

- Input methods and reading fields
- Fonts & text rendering
- Pan-Unicode fonts, font linking, Noto font
- Font attributes
- Debugging tips

▋ Handouts

Onsite attendees will receive a 650+ page booklet, with ample room for notes, complete with table of contents and glossary. The booklet is designed as a practical easy-to-use reference "book" for regular use during an internationalization project.

eLearning attendees will receive a PDF certification document if they pass the quizzes. The glossary mentioned above is available in HTML format on the eLearning site while the booklets can be purchased separately.

About our Instructor – Pierre Cadieux

Pierre Cadieux is a veteran with over 35 years' experience in internationalization of software, Web sites and mobile devices. He has taught internationalization at the Université de Montréal. Pierre has been technology editor for the LISA newsletter, VP Technology at ALIS and director of technology at Bowne Global Solutions.

At ALIS, Pierre pioneered the transparent handling of Arabic and Hebrew languages and created the core bi-directional technology licensed by Microsoft.

As Director of Localization Technology at Bowne Global Solutions, he carried out research and analysis on multilingual Web sites and published the first generic model of Globalization Management Systems.

Additionally, Pierre holds a B. Sc. and M. Sc. in Computer Science.