Case Studies: Terminology and Software Localization

Terminology Summer School - Cologne
06 - 10 July 2009

Klaus-Dirk Schmitz
Institute for Information Management
Faculty 03
University of Applied Sciences Cologne
klaus.schmitz@fh-koeln.de
Internationalization (I18N)

Developing a product or service in such a way that it will be easy to adapt it to other markets (languages and cultures)

- Goal: eliminate the need to redevelop, reprogram or recompile the original product / content
- Carried out by the development group in conjunction with the localization people
Localization (LI0N)

Adapting a product or service to a local or regional market

- Goal: Appropriate linguistic and cultural aspects
- Performed by translators, localizers, language engineers
Deutsch

Biotopverlust *sub. m. s.*


D-Quelle Schaefer
Kontext Bei der Bewertung der Umweltauswirkungen bestehen

 Specify criteria for the filter. Filter affects display and exporting.
I18N and L10N

Localization

Internationalization
I18N and L10N: Goal

- I18N and L10N comprise the whole of the effort involved in enabling a product.
- I18N is “stuff” you have to do once.
- L10N is “stuff” you have to do over and over again.
- The more stuff you push into I18N out of L10N, the less complicated & expensive the process becomes.
Locale

- Geographically a region of a particular market (e.g. the Italian speaking area of Switzerland)

- In IT a set of information and codes associated with a specific region
  - character set, writing system, keyboard layout
  - format for date, time, currency, units of measurement
  - format of addresses
  - etc.
Software localization: what + how
Localizing documentation

- Localizing printed documentation (handbook, manual, installation guide, packaging material etc.) is similar to “traditional” translation of LSP texts.

- **Tools:**
  - editor
  - word processor
  - DTP software
  - translation memory
  - terminology system
Die folgenden Hotels liegen verkehrsgünstig (wenige Autominuten oder direkte Busverbindung) zur Universität.

The following hotels are located only a few minutes from the university.

Von den beiden Hotels in Dudweiler kann man die Universität auch zu Fuß in etwa 15 bis 20 Minuten erreichen.

Bitte geben Sie bei der Zimmerbestellung an, daß Sie an einem Seminar der Universität teilnehmen. Dies verschafft Ihnen bei einigen Hotels Sonderkonditionen. Preisangaben sind ohne Gewähr.

City-Hotel, Richard-Wagner-Str. 67, 66111 Saarbrücken, Tel.: 0681 / 34088

Domizil Leidinger, Mainzer Str. 10 - 12, 66111 Saarbrücken, Tel.: 0681 / 38011

Hotel Schloßkrug, Schmollerstr. 14, 66111 Saarbrücken, Tel.: 0681 / 36735
Localizing online help

- When localizing online help, you have to care about formatting information (HLP, CHM, ...), document structure and links.

- Tools:
  - editor (HTML)
  - word processor (RTF)
  - translation memory
  - terminology system
  - localization software
Einführung in Corel CATALYST™


Corel CATALYST umfasst hochentwickelte Funktionen zur raschen Integration und Validierung Ihrer lokalisierten Software nach der Übersetzung. Because of the revolutionary "no-compile" localization strategy allowed by Corel CATALYST, you can extract an executable file and test the software without having to recompile source code each time. This means that your software build times and Quality Assurance (QA) test cycles are significantly reduced.

Using Corel CATALYST will streamline the way you localize software products and facilitate timely delivery of localized software products to the marketplace.

For more information see the following:

(Button „JumpID("catalyst.hlp","Corel CATALYST features")") Corel CATALYST features
(Button „JumpID("catalyst.hlp","About Corel Corporation")") About Corel Corporation
(Button „JumpID("catalyst.hlp","Corel technical support services")") Corel technical support services

Related Topics

* Welcome to Corel CATALYST
* Welcome_to_Corel_CATALYST: 001
* Übersicht: Einführung in Corel CATALYST
Localizing websites

- When localizing websites (and HTML help), you have to care about formatting information (HTML), document structure, links and multimedia elements.

- Tools:
  - editor (HTML, tag)
  - translation memory
  - terminology system
Have you ever wished you could produce a company dictionary, easily accessible at all times by all your colleagues? Now you can - with MultiTerm Web Interface, the TRADOS product for terminology access over the Internet. Just put your Company Dictionary on the Internet so that everyone in your company can access it - wherever they are based, anywhere in the world. Online, of course.

- **General Information**
  - Why Terminology? Because crystal clear communications with customers - and also in-house - are vital.

- **Creating & Distributing Terminology**
  - The MultiTerm Web Interface is the ideal complement for MultiTerm 95 Plus!. You create and manage your terminology with MultiTerm 95 Plus, and use the Web Interface to make it accessible over the Internet or Intranet.

- **Technical Information**
  - What are the technical requirements for using the MultiTerm Web Interface?

- **Sample On-line Databases**
  - Want to see the MultiTerm Web Interface in action? Just visit some sample on-line databases, and see for yourself what the MultiTerm Web Interface can do for you.

- **Ordering Information**
  - Interested? Here, you find information on how to order the MultiTerm Web Interface and related products.
Localizing software

- **editor**
- **localizer/translator**
- **LION tool**

**resource files**
- **TXT**
- **source text**
  - `Do while I < 100
    I = I + 1
    Print "Zähler =", I
    Wend`

**compiler**
- **linker**

**binary files**
- **EXE**
- **program**
  - 0101 1001 0100 1110 1010 0110 0010 1011
Localizing software

- When localizing software, you have to differentiate between the localization of resource files (RC) and the localization of executable files (EXE, DLL, ...). Different treatment of menus, dialog boxes, tool tips, (error) messages etc.

- **Tools:**
  - editor (RC, tag)
  - localization software
  - (translation memory)
  - **terminology system**
Terminology for software localization

- The selection of appropriate terminology with well defined concepts as well as the correct and consistent usage of terms are pre-conditions for successful software development and software localization:
  - very often innovative domains and topics
  - large projects with many persons involved
  - consistent terminology through all parts of the software product
  - pre-defined terminology (operating system, related software, laws and regulation, corporate language)
Software development and localization workflow

workflow „simple“, no $\Delta$, but late on the int. market
Software development and localization workflow

workflow complex, huge $\Delta$, very early LION, „simship“

→ Terminology needed very early!
Internationalization and localization aspects in content development

- Content development taking into consideration the process of internationalization and localization has to take care about e.g.:
  - terminology, symbols
  - (illustrations, character sets, fonts, colors, typography, style, formats etc.)

- More aspects are discussed in:
  - Schmitz/Wahle (2000): Softwarelokalisierung
  - Reineke/Schmitz (2005): Einführung in die Softwarelokalisierung
Terminology for software users

- Selection and creation of terms:
  - transparency
  - consistency
  - appropriateness
  - linguistic economy
  - derivability
  - linguistic correctness
  - preference for native language
Transparency

- The concept designated by the term can be inferred without a definition

- The meaning of the term is visible by:
  - morphological motivation
    - page setup, error message
    - data network identification code*
    - network printing device setup*
  - semantic motivation
    - mouse, worm, virus, infected file, vulnerability
    - firewall*
Transparency: fire wall
Transparency: fire wall

Hamilton Network

MAIN

INTERNET

OUTSIDE

Firewall

Hamilton Community can access any services on the Internet except for well-known vulnerabilities.

Off-campus users can only access approved Hamilton services (e-mail, web servers, SSS, Blackboard, Citrix, departmental servers, remote desktop, etc.)

Protected from on-campus and off-campus access.

Datatel Server

Basic Hamilton Firewall Network Diagram
January, 2004

K.-D. Schmitz, IIM, FH Köln
Transparency: fire wall
Consistency

The term is used consistently in all parts of the product (e.g. software products):

- menus, dialog boxes, error messages
- online help, tutorial
- documentation, packaging material, website

- return key, input key, enter key, line feed, carriage return, acknowledge key etc.

- update, hotfix, update rollup, security pack, private fix etc.
Consistency: terms

Matthias Heyn, SDL, tekom-Jahrestagung 2005

K.-D. Schmitz, IIM, FH Köln
Appropriateness

- Appropriateness means that terms:
  - have to be familiar to the user (localization!)
  - don’t cause confusion or insecurity
  - have no negative connotations (neutral, politically correct)
  - express installation (only components needed)
    network installation (all components)
  - system error, severe error, fatal error, user error etc.
  - master/slave, web designers’ bible, knowledge nugget etc.
Can you imagine that both user(group)s will use and understand the same term for a given concept?
Appropriateness
Oh come on—how fatal can it be?
Other features

■ Linguistic economy
  ■ *Ultrakurzwellenüberreichweitenfernsehrichtfunkverbindung*

■ Derivability
  ■ *medicinal plant vs. herb --> herbal, herbalist, ...*

■ Linguistic correctness
  ■ *aktualisieren vs. updaten, geupdated, upgedatet, ...*

■ Preference for native language
  ■ *Startseite vs. Homepage*
Symbols & icons

- Symbols and icons are very often used to optimize communication with the user as well as to optimize localization efforts.

- Although symbols are “language-free”, they are very much dependant on the “cultural environment” of the user.

- Symbols are very often taken from common sphere of life and use allusions (concrete vs. abstract), homonymy or even homophony.
Symbols & icons (e.g. software)

EN: trash
DE: Papierkorb

EN: clipboard
DE: Zwischenablage

EN: home
DE: Anfang / Startseite

EN: table (spreadsheet)
DE: Tabelle

EN: paws = pause (in a game)
DE: Tatzen ≠ Pause
Symbole

Which symbol you would choose for “incoming”?
Tools for software localization

- Consistent terminology needed across all components of a software product, therefore **terminology management systems**
- (internally + externally) repetitive texts allow for translation memory systems
- Localization of the program itself requires specific localization software
- Complexity and tide schedules require project management tools
Tools for software localization

- Established interaction between terminology management and translation memory software
- Cooperation between terminology management and localization tools still lacking
- Established principles of terminological data modeling are not always adequate for software localization terminology solutions
  ⇒ Specific solution needed
Specific terminology solution needed!

- Which additional data categories are needed for software localization terminology management?
  - Is it helpful to know, if the term appears in a menu, in a dialog box, in an error message, in an online help file or in the manual?
  - Is it helpful to know, if e.g. the term within a dialog box refers to a *check button*, *push button* or *radio button*?

  - *Activate Task List* = *Aktiviert die Aufgabenliste*  
    *Aufgabenliste aktivieren*  
    *Aktivieren Sie die Aufgabenliste*  
    (status line / dialog box / user request)
GUI terms
GUI terms

[Image of a software interface with selected terms highlighted in red]
GUI terms

Klicken Sie, um Notizen hinzuzufügen

Ein Element auf dieser Folie auswählen und "Effekt hinzufügen" anklicken, um die Animation hinzuzufügen.
Research needed!

- What is a terminological unit?
- What is a concept?
  - Format
  - Print
  - Save as ...
  - Copy file
  - Date and Time
  - Convert into Adobe PDF and send via e-mail
  - Unexpected error in application program
  - File %f could not be opened
  - etc.
GUI concepts

- Concept orientation:
  one concept = one terminological entry

- The GUI terms shown are managed within the software (and the localization tool) as one unique unit (with an unique ID)
  → one entry → one concept ?!

- When is it necessary to identify different concepts?
  - *Print...* in my and your Word 2007 ?
  - *Print...* in Word 97 and in Word 2007 !?
  - *Print...* in Word and in Excel !

K.-D. Schmitz, IIM, FH Köln
Definition of GUI concepts

Definition: representation of a concept by a descriptive statement which serves to differentiate it from related concepts (ISO 1087-1)

Intensional definition: definition which describes the intension of a concept by stating the superordinate concept and the delimiting characteristics (“ideal” def.)

Is it possible / necessary / desireable to (intensionally) define GUI concepts?

Is there a difference between such a definition and a traditional one? (Datei vs. )
Context of GUI terms

- Context: text which illustrates a concept or the use of a designation (ISO 1087-1)
- In „traditional“ terminology management, the context (example) shows the linguistic usage of a term.

- GUI terms don’t have a linguistic context!
- But the “situational” or “functional” context is important!
Kontext von Software-Benennungen
Kontext von Software-Benennungen
Data modeling for GUI terminology

- **Dandelion Project**
  Data Modeling and Data Exchange for Software Localization (DandDEL10N)
  - Cologne University of Applied Sciences (IIM)
  - June 2005 – May 2007

- **Objective:**
  - Optimizing the tools used for software localization (mainly terminology management systems)
  - Proposal for LION-specific ISO 12620 data categories
  - Proposal for adjusting TBX (and XLIFF) to LION needs
Data categories for LION

- LION-Type
- Development environment
- Project
- String ID
- LION-Context
- Source
- Term (e.g. Öffnen)
- Full form (e.g. Ö&ffnen... *tab* Strg+O)
- Linguistic type
- Part of speech
- Parameter
- Hotkey
- Shortcut
Specific datCats for localization

- **LION-type** (for Windows XP):
  - **Menu type:** Menu item (MI)
  - **Dialog box type:** Dialog box (DB), Group box (GB), Text box (TB), Combo box (CB), Combo box element (CBE), Check box (CHB), Tab (T), Push button (PB), Radio button (RB), Spin box (SB), Progress bar (PRB), Slider (SL)
  - **Message type:** Informative message (IM), Interactive message (IAM), ToolTip (TT)
  - **Table:** Table text (TAT)
  - **User defined type (UDT)**
LION-type: example

- **Menu type**: Menu item (MI)
LION-type: example

- Dialog box (DB)
LION-type: example

- **Group box (GB)**

```
Additional information

General comment:  
Title:             
Author:           
```

- **Text box (TB)**

```
General comment:  
Title:             
Author:           
```
LION-type: example

- Combo box (CB)
- Combo box element (CBE)
- Check box (CHB)
LION-type: example

- **Tab (T)**
  - ![Tab Example Image]

- **ToolTip (TT)**
  - ![ToolTip Example Image]
LION-type: example

- Push button (PB)
- Radio button (RB)
- Spin box (SB)
LION-type: example

- **Informative message (IM)**

- **Interactive message (IAM)**
## Extension of ISO 12620: new datCats for TBX

<table>
<thead>
<tr>
<th>Name</th>
<th>ISO 12620</th>
<th>Representation</th>
<th>dataType</th>
<th>Manda-tory?</th>
<th>Level</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Localization type</td>
<td>n/a</td>
<td><code>&lt;descrip type=&quot;lionType&quot;&gt;picklist box&lt;/descrip&gt;</code></td>
<td>picklist: Menu item, Dialog box, Group box, Text box, Combo box, Combo box element, Check box, Tab, Push button, Radio button, Spin box, Progress bar, Slider, Informative Message, Interactive message, ToolTip, Table text, User defined type</td>
<td>N</td>
<td>T</td>
<td>This data category needs to be added to TBX. The picklist values are from the Dandelion EU project. (Institute for Information Management, University of Applied Sciences Cologne).</td>
</tr>
</tbody>
</table>
Predefined termbase template for software localization in MultiTerm 2007
Data modeling for GUI terminology
Predefined termbase template for software localization in MultiTerm 2007
Conclusion I

- The creation of transparent, motivated, appropriate and consistent terminology is a pre-condition for the ease-of-use of software products and the empowerment of end-users.

- Therefore, terminology theory, working methods and applications are important for software development and localization.

- Diligent terminology (management) does not only help end-users but also documentation and localization experts.
Conclusion II

- Only a **proactive** and **multi-disciplinary** terminology approach will really help.

- **proactive:**
  when concepts are created and named

- **multi-disciplinary:**
  programmers, interface designers, technical writers, web designers, marketing experts, hotline people, localizers etc.
Conclusion III

- Specific terminology data modelling (incl. datCats) is required for software localization.

- Interfaces and data interchange features are needed between software development platforms, localization tools and terminology management systems.
  (Adaptation of TBX and XLIFF)

- Research in the field of general terminology theory is necessary: what are objects and concepts in software user interfaces?
Thank you for your attention

Prof. Dr. Klaus-Dirk Schmitz
Fachhochschule Köln
Fakultät 03 - ITMK/IIM
Mainzer Str. 5
50678 Köln
klaus.schmitz@fh-koeln.de