Internationalization, Globalization, Localization
(Terminology and Software Localization)

Klaus-Dirk Schmitz
Cologne University of Applied Sciences
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
Globalization (G11N)

Activities performed for the purpose of marketing a (software) product in regional markets

- Goal: Global marketing that accounts for economic and legal factors
- Focus on marketing; total enterprise solutions and top management support
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
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.

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

(EZ: DM 125,-; DZ: DM 170,-)

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

(EZ: DM 145,- - 175,-; DZ: DM 180,- - DM 200,-)

Hotel Schloßkrug, Schnellerstr. 14, 66111 Saarbrücken, Tel.: 0681 / 36735

(EZ: DM 50,- - DM 65,-; DZ: DM 90,- - 120,-)

Drücken Sie F1, um die Helfefunktion aufzurufen.
Localizing online help

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

• Tools:
  – editor (HTML)
  – word processor (RTF)
  – translation memory
  – terminology system
  – localization software
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/Intranet so that everyone in your company can access it - wherever they are based, anywhere in the world. On-line, of course.

Terminology Access over the Internet/Intranet

- 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/Intranet so that everyone in your company can access it - wherever they are based, anywhere in the world. On-line, 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**

**TX**

source text

**user**

**developer**

**EXE**

program

0101 1001 0100
1110 1010 0110
0010 1011

**resource files**

**compiler**

**linker**

**binary files**

Do while I < 100
I = I + 1
Print "Zähler =", I
Wend

LION tool

- editor
- localizer/translator
- LION tool
- **TX**
- source text
- **user**
- **developer**
- compiler
- linker
- EXE
- program
- binary files

---

- editor
- localizer/translator
- LION tool
- **TX**
- source text
- **user**
- **developer**
- compiler
- linker
- EXE
- program
- binary files

---

Do while I < 100
I = I + 1
Print "Zähler =", I
Wend

---

LION tool

- editor
- localizer/translator
- LION tool
- **TX**
- source text
- **user**
- **developer**
- compiler
- linker
- EXE
- program
- binary files

---

Do while I < 100
I = I + 1
Print "Zähler =", I
Wend
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 LION workflow

workflow „simple“, no Δ, but late on the international market
Software development and LION workflow

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

$\rightarrow$ Terminology needed very early!
I18N and L10N 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 content development

• Terminology must be defined accurately and used consistently at least within:
  – one document
  – one product
  – one company or organization

• Only one term for each concept (avoid synonyms !)

• Only one concept for each term (avoid homonyms !)
Terminology for software users

- Selection and creation of terms:
  - transparency
  - consistency
  - appropriateness
  - linguistic economy
  - derivability
  - linguistic correctness
  - preference for native language
Can you image that both user(goup)s will use and understand the same term for a given concept?
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
Symbols & icons (e.g. software)
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**
Terminology: menu

• Which type of “terms” do we find in software user interfaces?

• And which “terms” shall we include in our termbase?
Terminology: dialog boxes
Some theoretical considerations!

- Which “terms” appear in software user interfaces?
- What is a term and what is a concept?
- Which concepts/terms should be managed in a termbase?
- Should / could we define software concept as we have learned?
- What about the context?
- Do we need specific data categories and data modeling?
• 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 2003 ?
  – Print… in Word 97, in Word 2003 and in Word 2013 !?
  – Print… in Word and in Excel !
Drucken in Word95, Word 2003, Word 2013!
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 / desirable to (intensionally) define GUI concepts?

- Is there a difference between such a definition and a traditional one? (File vs. File)
Example: ruler

- A **ruler** is an **instrument** used in geometry, technical drawing and engineering/building to *measure distances and/or to rule straight lines*. (Wikipedia 17.11.09)

material or immaterial? definition still valid?
Example: Tools

- **Tool** - A mechanical *device* intended to make a task easier. (Wiktionary)
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!
Context of GUI terms
Context of GUI terms

- **Insert**
  - Draw Table
  - Insert
  - Delete
  - Select
  - Merge Cells
  - Split Cells…
  - Split Table
  - Table AutoFormat…
  - AutoFit
  - Heading Rows Repeat
  - Insert Multidimensional Cell…
  - Convert

- **Paste**
  - Paste as Hyperlink
Context of GUI terms
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)
Data categories for LION (Dandelion-Project)

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

### 1. Administration Information Section

#### 1.1 Administration Record

<table>
<thead>
<tr>
<th>Identifier</th>
<th>Version</th>
<th>Registration Status</th>
<th>Administration status</th>
<th>Justification</th>
<th>Origin</th>
<th>Effective Date</th>
<th>Change Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>termLocation</td>
<td>1.0</td>
<td>private</td>
<td>private</td>
<td>DC developed in the Dandelion project to document types of string applications in localization. Used in some software localization environments.</td>
<td>Dandelion</td>
<td>2006-07-11</td>
<td>Added tag to value domain</td>
</tr>
</tbody>
</table>

**Key**: 1823  
**PID**: [http://www.isocat.org/dataset/DC-1823](http://www.isocat.org/dataset/DC-1823)  
**Type**: complex/closed  
**Owner**: Wright, Sue Ellen  
**Scope**: public
Specific datCats for localization

### 2. Description Section

<table>
<thead>
<tr>
<th>Profile</th>
<th>Terminology</th>
</tr>
</thead>
</table>

#### 2.1 Data Element Name Section

<table>
<thead>
<tr>
<th>Data Element Name</th>
<th>term location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
<td>Dandelion</td>
</tr>
</tbody>
</table>

#### 2.2 English Language Section

<table>
<thead>
<tr>
<th>Language</th>
<th>English (en)</th>
</tr>
</thead>
</table>
| 2.2.2 Definition Section
| Definition       | Type of graphic feature on a GUI associated with a text string defined as a term in a localization-oriented terminology database. |
| Source           | SEW          |
| 2.2.2.1 Note Section
| Note             | Localizers use L10n information as a contextual reference when translating strings in resource files. |
| 2.2.3 Example Section
| Example          | Examples of Menu items for the File menu in IE include: /New/, /Open/, /Edit/, /Save/, /Save as/, etc. |
| Source           | MS           |
| 2.2.4 Explanation Section
| Explanation      | Any value of /term location/ represents a location in the corpus where the term frequently occurs, such as a user interface object (in software), a packaging element, a component in an industrial process, and so forth. |
| Source           | TBX-Basic    |

#### 2.3 French Language Section

#### 2.4 German Language Section
### 3. Conceptual Domain

<table>
<thead>
<tr>
<th>Data Type</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>string</td>
<td>/checkBox/ (check box)</td>
</tr>
<tr>
<td></td>
<td>/comboBox/ (combo box)</td>
</tr>
<tr>
<td></td>
<td>/comboBoxElement/ (combo box element)</td>
</tr>
<tr>
<td></td>
<td>/dialogBox/ (dialog box)</td>
</tr>
<tr>
<td></td>
<td>/groupBox/ (group box)</td>
</tr>
<tr>
<td></td>
<td>/informativeMessage/ (informative message)</td>
</tr>
<tr>
<td></td>
<td>/interactiveMessage/ (interactive message)</td>
</tr>
<tr>
<td></td>
<td>/menuItem/ (menu item)</td>
</tr>
<tr>
<td></td>
<td>/progressBar/ (progress bar)</td>
</tr>
<tr>
<td></td>
<td>/pushButton/ (push button)</td>
</tr>
<tr>
<td></td>
<td>/radioButton/ (radio button)</td>
</tr>
<tr>
<td></td>
<td>/slider/</td>
</tr>
<tr>
<td></td>
<td>/spinBox/ (spin box)</td>
</tr>
<tr>
<td></td>
<td>/tab/</td>
</tr>
<tr>
<td></td>
<td>/tableText/ (table text)</td>
</tr>
<tr>
<td></td>
<td>/textBox/ (text box)</td>
</tr>
<tr>
<td></td>
<td>/toolTip/ (tool tip)</td>
</tr>
<tr>
<td></td>
<td>/userDefinedType/ (user defined type)</td>
</tr>
</tbody>
</table>
Data modeling for GUI terminology

Termbase Wizard - Step 1 of 5

Termbase Definition

Specify whether you wish to create a termbase definition using one of the predefined termbase templates, load an existing termbase definition or create a new termbase definition from scratch.

- Create a new termbase definition from scratch
- Use a predefined termbase template
- Bilingual glossary
- Multilingual termbase
- Software: Graphical User Interface
- Term Lifecycle Management (TLM)

- Use an existing termbase as a template

Copy catalog objects to the new termbase
Data modeling for GUI terminology

Termbase Wizard - Step 4 of 5

Descriptive Fields
Create descriptive fields for your termbase entries. Use the Properties dialog box to specify the type of data each field may contain. The default data type for all fields is text.

Field label:

Description:

Available descriptive fields:
- LION-Type
  - Part of Speech
  - Linguistic Type
  - Hotkey
  - Special Features
  - Context
  - Source
  - Shortcut
  - Development Environment
  - Parameter
  - Project

Properties - LION-Type

Data type:
Picklist

Picklist:
- Button description
- Check box
- Combo box
- Combo box content
- Dialog box
- Dialog text
- Error message
- Group box
- Menu item
- Progress bar
- Pushbutton

Options:
- Keep history
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