

File name:

KOM2002-translating-v1.doc

EDC-22071 KOM2002/27938
Groupware for Distributed Content Production KOM2002

D 3.1: Translating Web4Health to a new Language

Security (distribution level)	Public
Contractual date of	January 2003
delivery	
Actual date of delivery	24 Dec 2003
Deliverable number	Attachment A to Report D3.1
Deliverable name	Translating Web4Health to a new Language
Туре	Report, Public
Status & version	1.0
Number of pages	11
WP contributing to the	WP3
deliverable	
WP / Task responsible	FH NON
Other contributors	КТН
Author(s)	Jacob Palme
EC Project Officer	Kimmo Rossi
Keywords	Translation, Localization
Abstract	This deliverable gives an easy to use description of how
(for dissemination)	to translate the Web4Health web site to a new language.

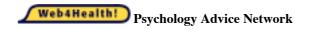


Table of Contents

1	Executive Summary	2
2	Stages of Translation	3
3	Account	3
4	Translating of HTML Documents	5
5	Stage I: Translating the User Interface and Subject Tree	6
	5.1 Translation of the User Interface	6
	5.2 Translation of the Subject Tree	6
6	Stage II: Translation of the FAQs	7
7	Stage III: Natural Language Question Answering	10
8	References	11

1 Executive Summary

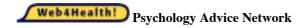
The intention of this manual is to give an easy to use description of the work needed to translate the Web4Health web site to a new language.

This translation can be done in three stages, where each stage leaves a usable system in itself:

Stage I: Translation of the user interface and the subject tree. After this stage, users can see commands and menus in the new language, but the actual texts of the FAQs will be in English.

Stage II: Translation of FAQs. After this stage, users can find FAQs using the subject tree in their own language.

Stage III: After this stage, users can ask questions and find answers from the natural-language question-answering system in their own language.



2 Stages of Translation

Stage	Description	User services	Approximate cost for translating organization for 300 FAQs	Approximate cost for KTH for 300 FAQs
I	Translation of user interface and subject tree.	Users can browse in their native language, but will see the FAQs in English	2-4 weeks	1-2 weeks
II	Translation of FAQs.	Users can read FAQs in their native language.	2-4 months	1-2 weeks
III	Translation of natural-language question-answering system.	Users can ask questions in their native language.	1-3 months	2-4 weeks

3 Account

To use KOM2002, you need an account. You get an account by connecting to http://cmc.dsv.su.se/eu/kom/. You will then see the following page:



Push the Register button. You will then find the following page:

Register in EU-KOM

Fill in the form and click on the "Submit" button below. TIP: Use the TAB button to jump between the fields.

		-	
Your name:		Type your na for example	me,
Choose a short, unique login id:		Mary Middle Choose a login	ı id,
Your email address:		for example m	arym
Choose a password:			
Repeat the password:			
O Male	O Female		(not required)
☐ I want a newsletter		‡	How often do you want to be reminded when there is news fo you in KOM. Once a
Default for text input:	Plain Text 🔾 Smart	: O HTML O	day is recommended.
Command Language:	Give language 💠	all langua	ges are not fully supported
Preferred Languages:			anguages in the format: sv,en,de
Personal homepage (no	t required):	Not important	
http:// Give language \$ Short	presentation of yourse	or required): Plain Text 🔾
Smart O HTML O	presentation of yourse	ii (iiot requireu	,. Hair toke o
Submit Preview/Save [Delete	Language cod de = German en = English el = Greek it = Italian nl = Dutch sv = Swedish	es:

After you have registered, write an e-mail to cmc@dsv.su.se and ask us to give you privileges as an expert in our project. We will then admit you to the internal pages for our project.

Push the Submit button

4 Translating of HTML Documents

Many of the documents to be translated are in the HTML format. The HTML format combines plain text with markup. Markup is commands to structure the text and format the layout. *To translate HTML*, *you do not really have to know the meaning of the different HTML commands*. All you need to recognize is what is plain text (to be translated) and what is HTML commands (not to be translated except in a few seldom used cases).

Examples of HTML code, the part which is to be translated is marked with a border in the table below:

HTML code What the user sees Jack and Jill went up the hill. Jack fell Jack and Jill went up the hill. down, and Jill came tumbling after. Jack fell down, and Jill came tumbling Humpty Dumpty sat on a wall. Humpty after. Dumpty fell down. And all the kings horses and all the kings men could not put Humpty Dumpty sat on a wall. Humpty Dumpty fell down. And all the kings Humpty Dumpty together again. horses and all the kings men could not put Humpty Dumpty together again. <q> Jack and Jill went up the hill. Jack fell Jack and Jill went up the hill. down, and Jill came tumbling after. More. Jack fell down, and Jill came tumbling after. More. Dumpty fell down. And all the kings horses and all the kings men could not put Humpty Dumpty sat on a wall. Humpty Humpty Dumpty together again. Dumpty fell down. And all the kings horses and all the kings men could not put Humpty Dumpty together again. <01> 1. First choice. First choice. 2. Second choice. Second choice. <A HREF="ed-causes-overview.htm"</pre> Overview Slim ideals; links class=linklist>Overview <A HREF="ed-causes-ideals-links.htm</pre> class=linklist">Slim ideals; Links <img src="/images/small-ball.gif"</pre> width="6" HEIGHT="6">

Most pages contain only simple HTML. If you find that a page contains too complex HTML, you might get an HTML expert to help with translation of that particular page.

Last update: 24 Dec 2003

If you look at the HTML code, you can see that it consists of plain text and HTML commands. HTML commands start with "<" and end with ">". The main principle for translating HTML is that only the plain text, not the HTML commands, are translated. Thus, in the examples above, "ed-causes-overview.htm" and "small-ball.gif" are *not* translated.

There are some exceptions from this rule. For example, in <INPUT name=Submit type=submit value="Submit question" class="FormSubmit">>, the string "Submit question" needs to be translated. Instead of learning which these exceptions are, easiest is to first not translate them, and then look at what the user sees. If any text in English is seen by users, then this needs to be translated in the HTML code. It is better to first translate too little than too much. If, for example, you wrongly translated "name=Submit" to another language, the resulting page would not work correctly.

Another exception is "alt=" and "title=". These HTML constructs are used to give text which is shown to blind people, and also to ordinary users when they move the cursor on top of an image. They should be translated, but mostly occur in templates, not very often in the body of FAQs.

Here are some examples where there is a border around the text which should be translated:

```
<IMG SRC="http://web4health.info/images/phobia-treat-snake-3.jpg" WIDTH="340"
HEIGHT="431" ALIGN="LEFT" alt="Picture of patient handling a snake">

<IMG alt="Should not be visible" border=0 height=2 src="/images/uppsvart.gif"
width=15>

<a class="tabNotChosen" title="Advanced search"
href="http://web4health.info/en/advanced-search.html"></a>
```

5 Stage I: Translating the User Interface and Subject Tree

5.1 Translation of the User Interface

The user interface consists of a number of pages and page templates in HTML format, for example the *home page*, the *advanced search page* and the *templates for the FAQ answer pages*. Since these pages contain quite complex HTML, it may be better to ask a person with some knowledge of HTML to translate them.

5.2 Translation of the Subject Tree

The subject tree is a set of linked pages, such that users can find all or most answers by only clicking on links in the subject tree pages. After translation of the subject tree, users will see the subject tree in the new language, but when they click on the link to an FAQ which has not been translated, this FAQ will be shown to them in English.

Do not translate the subject tree in the same way as ordinary FAQs (described in chapter 6 on page 7). It is easier to translate the subject tree by asking KTH for a copy of the file *all-menus.doc* and use Microsoft Word to translate the HTML in this file. In this file, every link is a separate line, keep it that way after translation.

When you are ready, send the file to KTH for entering into the Web4Health data base.

6 Stage II: Translation of the FAQs

Warning: On a Macintosh, we recommend that you do not use Microsoft Internet Explorer for this. Use instead for example the Safari or Mozilla browswers.

To translate the FAQs, go with a web browser to the following web address:

http://web4health.info/eu/kom/9;login

You may then have to enter your login name and password first. You will then see a page containing the following text:

Create FAQ table 1 Help



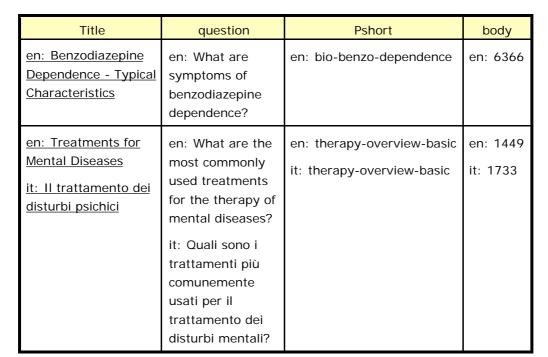
If you are to translate from English to Italian, we recommend that you change the settings in this page as follows, before clicking on List all.

Create FAQ table 1 Help



When you click on List all, you will then get a long table, part of which looks like this:

Web4Health!



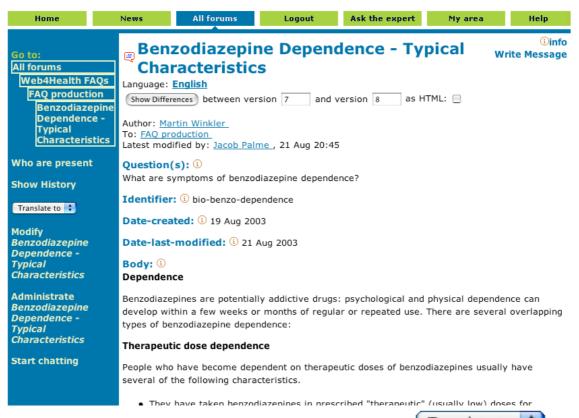
In the two rows shown above, the first row shows an FAQ which had not yet been translated to Italian when the table was produced, and the second row shows an FAQ which had been translated to Italian. The numbers, for example en: 1449 and it: 1733, are the number of characters of the body. If the body is less than 300, the whole text is shown instead of the number.

In order to decide whether to translate an FAQ to a new language, click on the value in the Title.

Tip: Since generation of the table takes a long time, I recommend that you open the FAQ in a new window or a new tab. You get this effect by holding down a special key while clicking on the link. This special key varies between browsers, it might be the SHIFT or the CTRL or on the Macintosh command (策) key. On some browsers you can set this key in the Preferences/Settings/Options window.

If you do this, you can get back to the table again by just closing the new window or tab, and do not have to perform the time-consuming generation of the table of FAQs over and over again.

When you click on the title of an FAQ in the table, you will get a page which begins something like this:



If you decide to translate this to a new language, pull down the **Translate to** menu to the language you want to translate to. You will then get a window which looks like this:

You can only change the information in the rightmost column (Italian in the example), do not attempt to change the leftmost (English) column in this window.

The rightmost window is sometimes filled with a Systran translation, which you can use as a basis if you so wish. If you do not want to use the Systran translation, you can remove it by clicking in the text box, giving the *Select All* command, and then pushing the *Delete* key.

If you so wish, you can click in the English text, give the Select All and Copy commands, then click in the target text box, give the Select All and Paste commands. You will then get all the HTML markup correctly, and need only replace the English text in the target text box with the translation.

For stage II, you need only translate the following text boxes:

- Title
- Question(s)
- Body
- Author
- Sources, references
- Meta-description

• (Meta-keywords¹)

Leave the other boxes as they are.

Note: For FAQs in the subject tree (such FAQ have usually the word "Link" or "Menu" in their name) it may be easier to translate them as described in chapter 5.2 on page 6.

When you are ready with translating an FAQ, push either the

Submit or the

Submit and export everywhere buttons.

The Submit button will save the translation, but not yet show it to ordinary users.

You can modify a partly-finished translation with either the *Translate* (in the English page) or the *Modify* command (in the target language page). *Translate* will show the English and target language text side by side, *Modify* will only show the target language text.

The Submit and export everywhere button will save the translation and show it to ordinary users. The result of this button will contain a list like the following:

Export Results:

/WWW/en/answers/bio-benzo-dependence.htm

http://faq.web4health.info/w4h-faq-en/faq.exe?bio-benzo-dependence

/WWW/en/answers/print/bio-benzo-dependence.htm

/WWW/en/answers/source/bio-benzo-dependence.htm

/WWW/en/KOMFAQ/bio-benzo-dependence.htm

/WWW/en/KOMFAQ-text/bio-benzo-dependence.htm

/WWW/en/answers/examples/bio-benzo-dependence.htm

In this list, you can click on the different rows, to see how this FAQ will look to ordinary users.

7 Stage III: Natural Language Question Answering

For this stage, you need to translate the following fields

- Required
- Optional
- Priority
- Forbidden

How to fill in these fields is described in D 2.2B at http://web4health.info/documentation/D2-2b-classification.pdf

You also have to translate the synonym table, which defines terms beginning with \$, for example \$causes.

Note: Do not translate words beginning with \$, for example \$causes. Translate words not with only one word, but with all synonyms in a wide sense to this word. For information on how to fill in these fields

Attachment A to Report D3.1

¹ Not so important. Only used by some search engines as additional information when people search for pages. Not used by Google.

Advice: It is important that you test the classification afterwards, by writing questions and see the result in the natural-language question-answering system on page http://web4health.info/test/ and the pages in other languages linked from this page.

8 References

D5.1B	Dreamweaver templates for Web4Health web site design, by Jacob Palme. KOM2002 deliverable D 5.1B.
	http://web4health.info/documentation/health-info-form.pdf
D 5.1C	Naming of Web4Health Informational Pages, by Jacob Palme, KOM 2002 deliverable D 5.1C.
	http://web4health.info/documentation/KOM2002-page-naming.pdf
D 2.2B	Natural Language Question Answering System - Classification Manual
	http://web4health.info/documentation/D2-2b-classification.pdf
D 5.1A	KOM2002 Groupware - Use in Web4Health
	http://web4health.info/documentation/KOM2002-groupware.pdf