

**EDC-22071 KOM2002/27938**

Groupware for Distributed Content Production  
KOM2002

**First report of stepwise refinement experience**

Security (distribution level)	Public
Contractual date of delivery	30.08.2003
Actual date of delivery	10.11.2003 (postponed)
Deliverable number	D 4.2.
Deliverable name	First report of stepwise refinement experience
Type	R
Status & version	1.0
Number of pages	18
WP contributing to the deliverable	WP4
WP / Task responsible	Martin Winkler (FH NON)
Other contributors	Jacob Palme (KTH)
Author(s)	Martin Winkler
EC Project Officer	Kimmo Rossi
Keywords	Stepwise refinement, content evaluation
Abstract (for dissemination)	This deliverable describes the ongoing process of analyzing the special interests and demands of our webpage visitors in the first 3 month of the public service. We use a stepwise refinement approach to review and adapt our existing content.



## Table of content

<a href="#"><u>1. Executive summary</u></a> .....	5
<a href="#"><u>2. Introduction</u></a> .....	5
<a href="#"><u>2.1. Description of step-wise refinement</u></a> .....	5
<a href="#"><u>2.2. Goals of step-wise refinement</u></a> .....	6
<a href="#"><u>3. Ask-the-Expert service</u></a> .....	7
<a href="#"><u>3.1. Answering a question</u></a> .....	7
<a href="#"><u>3.2. Topics of interest in the Ask-the-Expert section</u></a> .....	8
<a href="#"><u>3.3. Irrelevant questions to the expert team</u></a> .....	10
<a href="#"><u>3.4. Emergency events / critical questions</u></a> .....	10
<a href="#"><u>3.5. Privacy issues</u></a> .....	10
<a href="#"><u>4. Natural language question answering system</u></a> .....	12
<a href="#"><u>4.1. Introduction</u></a> .....	12
<a href="#"><u>4.2. Topics of interest of the questions to question-answering system (Q/A)</u></a> .....	13
<a href="#"><u>4.3. Stepwise refinement of the question-answering system</u></a> .....	14
<a href="#"><u>4.3. Emergency events/ critical questions</u></a> .....	15
<a href="#"><u>4.4. Privacy issues</u></a> .....	16
<a href="#"><u>5. Direct feedback by the users using email, phone, fax or other options</u></a> .....	17
<a href="#"><u>6. Consequences of the stepwise-refinement analysis</u></a> .....	18



## **1. Executive summary**

This deliverable gives a description of the topics of interests of our users (demand) and the present status of our coverage of different topics with our provided content. We evaluated the areas of special relevance for our users by a careful evaluation of the question and feedback posted to our Ask-the-Expert team in the course of the first three months of public service of the webpage [www.web4health.info](http://www.web4health.info).

Questions of a wide range of psychological and psychiatric problems were asked to the psychotherapists of the different European countries. During the first 3 months more than 250 individual questions were considered and answered. These questions and answers will also be included in our present database of frequently asked questions (“FAQs”).

We also made a first analysis of the usage of our unique natural language question-answering system for this deliverable. We achieved an unexpected high traffic with more than 100000 visitors and nearly 2 millions “hits” in the first months of our public service. So only a representative number of questions posted to the natural language question-answering system could be analysed. It is very interesting to see that the topics of interests according to the content of the questions to our computer natural language question answering system have very similar content compared to the questions to the human expert team. Less than 3 percent of all users of our webpage rated our existing content of the individual FAQs. So we cannot provide a representative analysis of the user rating of the content at the present time. We will provide further user evaluation to get a deeper analysis of the user satisfaction of our content.

The interests of our webpage visitors are very well represented in our existing database of FAQs. Comparing the demands of the users of the natural language question answering system and the interests of the people who asked a question to our expert team we can demonstrate that we meet these interests.

This deliverable was postponed to cover a representative time after the start of our public service in July 2003. Due to the unexpected high traffic on our webpage with more than 1.9 million hits we had to adapt to this enormous demand and adapt the process of stepwise refinement in the ongoing project to this high number of visitors.

In the next phase of the process of stepwise-refinement we will evaluate the individual FAQs to identify information which should be improved or changed. These results will be presented at the final report of stepwise refinement.

### **2. Introduction**

#### **2.1. Description of step-wise refinement**

The term “step-wise refinement” is a common description of successive optimization in computer science:

Successive definition and improvement by breaking the original problem into smaller subproblems and then applying the same process to each subproblem

While this basic principle is applied to a broad aspects of continuously improvement within our project KOM 2002 this deliverable refers to the main tasks of workpackage W4 to improve the quality of the existing content of the webpage by analyzing the user feedback to our existing webpage with a careful analysis of the questions of the Ask-the-Expert service and the questions addressed to the natural language question-answering system.

While the webpage evaluation (Deliverable D7.1 Evaluation plan) offers the systematical investigation of the access, usability, transparency and overall quality of the webpage with statistical data, stepwise-refinement refers to the demands and needs and experiences of the user with our webpage and the options for further development of the webcontent.

## **2.2. Goals of step-wise refinement**

To adapt to the demands of the users of our webpage we try to provide psychotherapeutic information according high standards of medical webpages (see Deliverable D 2.1. Content plan). For the process of stepwise refinement the monitoring of the individual usage of our webpage gives relevant information about specific areas of interest and strenghts of the existing content as well as possible topics that need a more detailed consideration in the course of the project.

There are a wide variety of all kinds of webpages in different languages available, which cover special information or try to offer different information in the area of mental health. Considering our recent resources and our technical opportunities we do not plan to copy these webprojectes, but try to establish a new type of content access to the users. The main focus will be the unique option to get answers to common (or individualized) questions in the area of psychology and psychotherapy using the multilingual natural-language question-answering system. We offer the options provide a multilingual database of these questions and answers and easy access to these information. More important, the users do not need to be aware of the specialized terms but can do a natural-language search and will get a cluster of relevant information connected to their individual question. This kind of individual patient information is a unique feature of our project.

So we decided to start with a set of commonly asked questions and answers, which were send to the medical experts during the time span prior the project and during the course of Workpackage 1 and Workpackage 2. We offer more than 500 of these question-answering FAQ at the start of the public service of our webpage and will include more according to the future questions and demand of our users. So we do not expect to cover all possible aspects of mental health at the beginning or meet all needs of all people worldwide. But we already achieve a good response to the most relevant questions of the users. This shows the future potential of a buisness solution with a larger group of medical and psychological experts which can use the system for psychoeducation.

This deliverable describes first experiences of the interests of the users of our webpages by analyzing the main topics of interests of questions addressed to our natural language question answering system and the human expert team.

### **3. Ask-the-Expert service**

#### **3.1. Answering a question**

One of the most popular options of interaction between our users and the medical experts is the Ask-the-Expert-option.

The official opening of our webpage was in July 2003. It usually takes at least 2 to 3 months until a new webpage is listed in the major search-engines. We got our first questions at the 26.7.2003. For this report the questions of the month August to October have been analyzed. The average number of question to the Ask-the-expert-section is subject to changes. At the beginning of the service we received about 2-5 questions per day. We notice an ongoing increase of the number of individual users and questions in the Ask-the-Expert section. Access to the Ask-the-Expert area is free of any charge.

To answer a question of a user one of the medical experts marks the question of the user in our groupware. So only one of the expert team will need to reply to the individual question. Any expert is free to provide an additional answer so a user might get different opinions from different members of the psychotherapeutical expert team.

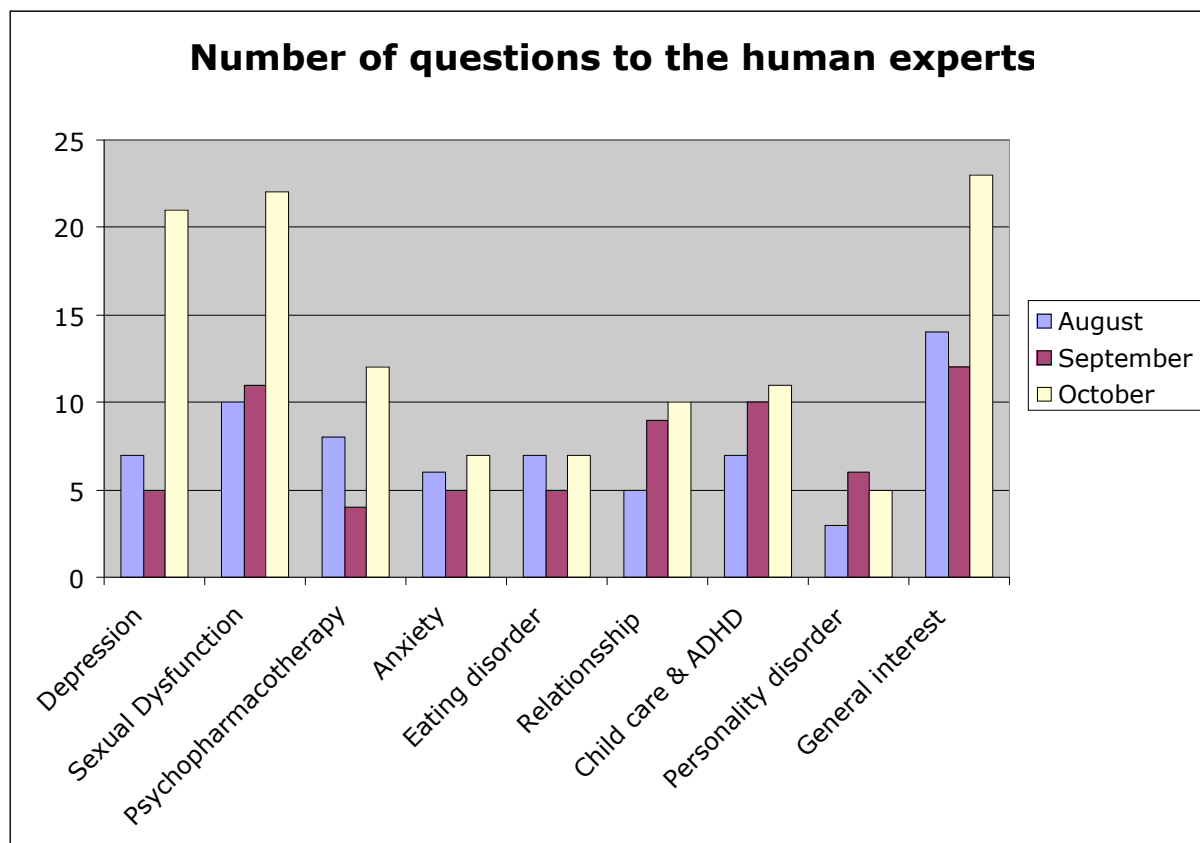
Usually we have to consider a time span of at least 24 to 48 h until an answer will be provided. We have to write an appropriate answer (sometimes after a further evaluation of recent literature for sources or additional information) and provide a translation for the user. Most experts will write an answer in the native language and provide a translation themselves to shorten the time-span until the answer can be provided. They might use the machine translation tool provided by our partner Systran for this purpose or translate the text themselves without any technological help. We have to accept minor restrictions of the quality of translations with some spelling errors or grammar problems to provide an answer within a time no longer than usually 7 working days. Most answers were given within 48 hours. No answer was posted after 10 days.

### 3.2. Topics of interest in the Ask-the-Expert section

We receive questions concerning very different psychotherapeutically issues.

The table gives a summary of some of the most relevant topics of interest of users in the Ask-the-Expert section of KOM2002 (webpage [www.web4health.info](http://www.web4health.info)). Most of the questions do refer to two or more topics, so only the main topic of the question is considered for this summary.

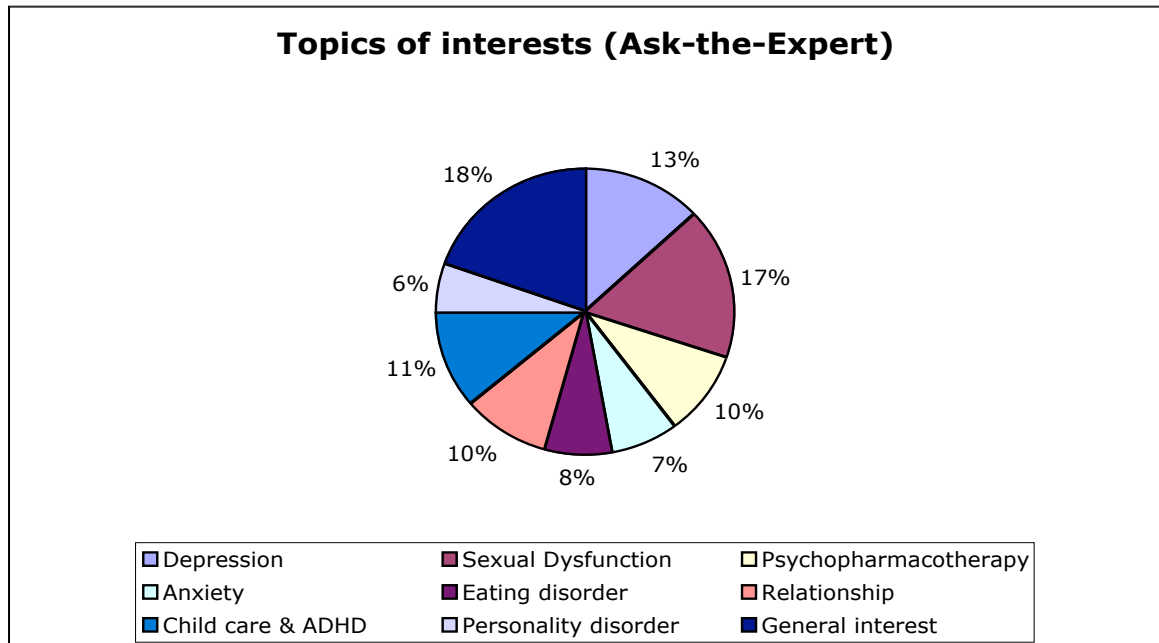
Topic	Aug. 2003	Sept 2003	Oct 2003	Total
Depression	7	5	21	33
Sexual problem/ abuse	10	11	22	43
Psychopharmacotherapy	8	4	12	24
Anxiety disorder	6	5	7	18
Eating disorder	7	5	7	19
Relationship	5	9	10	24
Child care / ADHD	7	10	11	28
Personality disorder	3	6	5	14
General question	14	12	23	49
<b>Total</b>	<b>67</b>	<b>67</b>	<b>128</b>	<b>252</b>



The areas of interest represent the whole spectrum of psychotherapy. Most of the major topics are already covered by our areas of content. So we think that our existing content plan (Deliverable D2.1) fits to the demand of our actual users.



The actual number of questions for the different topics is subject to changes. This will also be influenced by the ranking our webpage in the search engines and the special interests of clients. So the choice of topics might not represent the average interest of all internet users at the present time. We would expect more questions related to Depression and / or Anxiety disorders.



This might be partly influenced by the fact that there are a lot of other well known webpages with information about depression or anxiety disorders available on the internet.

While this option is very popular for the readers of the webpage with a rather high daily access of this forum of the webpage, the actual number of questions is limited if you consider the high traffic of our webpage in the first months of free access via search engines. We had about 17000 individual visitors with more than 350000 “hits” in August. In September there was a marked increase of visitors (42663) and hits (more than 747000), mainly influenced by a single link by a popular American radioshow. In October we had about 626678 hits at our main webpage and about 40300 visitors.

Most of our visitors used the option to find relevant information with the directory of content and the common subject trees. About 8000 questions were given to the natural language question answering system in September. The number of questions addressed directly to the expert team was rather small. So only a very small percentage of the users had the need to ask a questions that was not yet covered by our natural-language question-answering system.

However, this service is popular to our users and we have an increasing interest of our visitors to ask individual questions. So we think, that the user tested the question-answering system during the first weeks of the public service and start now to look for more individual support by the expert team. This can be seen by the increasing number of people who register for the services of our webpage and the increasing number of question sended to our expert team in October.

### **Irrelevant questions to the expert team**

Since our personal resources are limited we did not promise to answer all questions in the Ask-the-Expert area. So we have to make a selection of relevant topics that are suited to the goals of our project.

We also received some questions, which do not refer to our major topic psychotherapy and public health information. Some of the users had medical questions (related to surgery, special aspects of inflammatory disease or blood in semen) or wanted a specific treatment advice related to the individual medical problem. We included a remark at our webpage interface and the Ask-the-expert section that we cannot answer these questions. However, we still get 2 or 3 of those medical questions daily. We monitor these questions but will not provide an answer.

Some medical students or journalists asked for highly specialized statistical data (e.g. “number of eating disorders in New Zealand”). These questions are not of any common interest to the users in Europe, so we decided to ignore topics that are not closely related to psychotherapy.

#### **Emergency events / critical questions**

During the development of the content of the project KOM2002 we put very much attention to any possible emergencies or questions and requests of any client in acute danger (e.g. suicidal threats, intoxication, aggressive behaviors). However we had no critical event in the first months after going public. We will provide a permanent monitoring of any possible acute questions and have included information for any case of emergency.

#### **3.5. Privacy issues**

Special care is provided to handle privacy issues. Many users write a personal email-address in their question. We do not publish these addresses to protect the privacy of the users. Individual email-response will be provided for registered users. No additional personal data is stored.

## **4. Natural language question answering system**

### **4.1. Introduction**

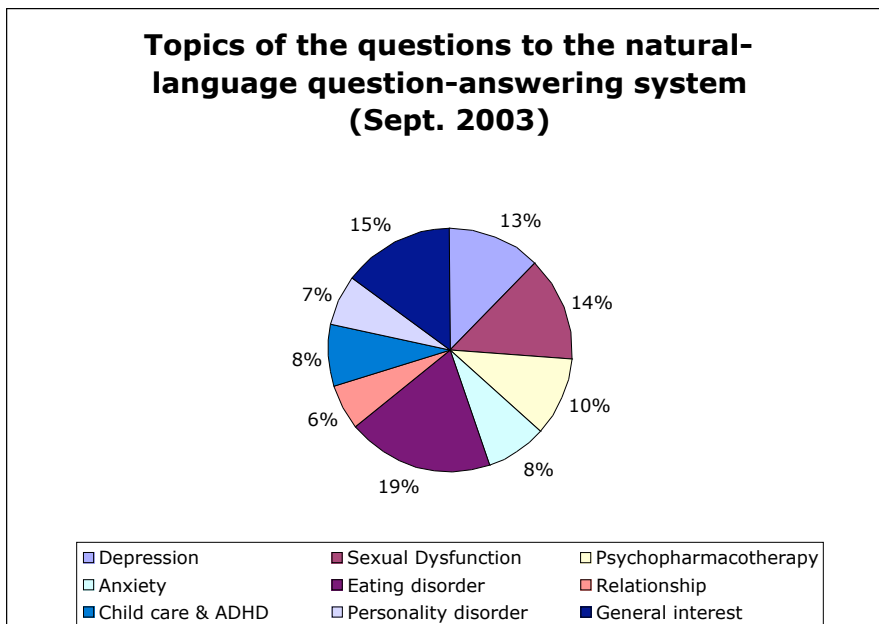
The question-answering technology (Q/A) is a new option to present information to the public. At the beginning of the online service many people were not used to the new technology of the question-answering system using complete sentences or questions to get a search result. So we provided additional subject trees and additional menus for the users. To achieve a higher use of the specific question-answering system we decided to include a new option for the users to get an additional question by the question-answering system if the user has only found the individual FAQ via external search engine or our subject trees. So a higher number of users will try the unique features of the question-answering technology.

The groupware technology offers a logfile of the usage of the question-answering system. So we get a report when no answer was found to a specific question of the individual user or when the system found no appropriate answers in the pool of FAQs. We will also provide the opportunity to have a user rating for each individual question to identify questions that do not meet the expectations or personal needs of the users. Another option to rate the user interest is the time a user spends for reading a specific text. Short reading time might be an indicator for low interest or insufficient information. At the present stage of the project we do not want to present statistical data of this evaluation. Only a very small percentage (less than 3 percent of all users) gave a rating to the individual answers. So we do not have representative data concerning this topic at the present time.

We are going to optimize the analysis tools for the higher traffic of our webpage and will present data related to the stepwise refinement process using the logfile information in our final report of stepwise refinement.

#### 4.2. Topics of interest of the questions to question-answering system (Q/A)

To evaluate the interests of the clients using the natural language question-answering system we made a deep analysis of all questions posted to the question-answering engine in September 2003. More than 8000 entries were analysed. For this first analysis we excluded all entries which were obviously caused by a computer entry, double entries of users and topics not related to mental health issues. It is very interesting to notice, that the main interests of the users are similar to the topics posted to the human experts. But we can also identify a subgroup of clients interested in eating disorders (19% of all questions), who use the special opportunities of the Q/A-search engine to find individualized patient information. These users tend to ask more questions to a specific topic (e.g. “what are the causes of bulimia nervosa”).



### **4.3. Stepwise refinement of the question-answering system**

The logfiles of the question answering system are screened by the medical coordinator on a daily schedule. We try to identify common interests of the users not covered by the existing FAQs to look for relevant possible topics for new FAQ to be included in the system. At the present status of the project we are mainly interested to improve the technological functions of the question-answering system and to improve the usability of the webpage.

The natural-language question-answering system will itself report an error in its log file if the question contained an unknown word or was not enough similar to any of the existing FAQ templates. The system in this way found that 85 % of all questions were successful and 15 % not successful. This is not a fully reliable indication of whether the answers were satisfactory to the user, to find this, more user testing need to be done.

According to our experience in the process of stepwise refinement the main reasons for an non-successful report in the logfile are :

#### **Questions not within the scope of our project**

Many users tried the natural question-answering system with a lot of different questions to test the response of this technology. Obviously many users tried many somatic questions or topics, which are not covered in our database of existing FAQs. Many users try to repeat these kind of questions with a slightly different choice of words. We have written a clear instruction that no pure somatic or medical questions will be answered in the question-answering system, but we still get many of these question. About 30 to 40% of the logfile entries are caused by this kind of questions not within the scope of the project. The approach to improve the existing FAQ database was a new FAQ with a general remark about the main topics of our webpage if users ask for common medical problems. Using the stepwise refinement approach we screen the questions of the users for typical medical expression not covered by the existing psychological topics. If the system identifies such typical question the user will get a feedback that we do not provide this information and a link to another possible source of information.

#### **Spelling errors or wrong terminology**

While we cover a lot of everyday terminology for psychological or psychiatric expressions in our synonym list we still have a significant number of spelling errors or bad terminology of users of the question-answering system. Many user correct their mistake themselves and ask a correct question later. But we will test a new software tool to handle the most frequent spelling errors to provide a higher number of sufficient answers to the user. Another option of stepwise refinement of the system is to improve the lists of common synonyms for the indexing and classification of the FAQs. .

### **Use of single search terms**

Many of the users of the question-answering system tried to use single word search terms. We provide a distinct option for this search option but the full potential of the technology will be achieved by a complete natural language question. Unspecific terms will not give specific answers to the users. This is a rather common problem for question-answering technologies. Very often the user will start testing the options of a question-answering system with a single term. Then the average user will try more complex search options.

Some users will not at all use the question-answering system but used external search engines like Google (TM) or external links to a specific FAQ. To help these users to use our existing question-answering technology we offer an additional option to use the question-answering system with a question closely related to the single FAQ. So the user will get additional answer closely related to his interests.

### **Answer existed, but was not found by the system**

The indexing and classification of the FAQs is very important to offer a good answer to the users. If we use a very specific indexing approach the user will only get a very specific answer. Our actual approach offers a cluster of relevant answers to the user. There are a set of answers closely related to the question and an additional section with related information concerning the topic. At the present time of the project almost all English FAQs have been indexed and should be found by the question-answering system. The webpages of the other languages have not yet been fully translated. All existing FAQs and some FAQs have not yet been included in the system. This will be finished in the course of WP4.

We will continue to improve the indexing system of the natural-language question system. This will be done by further testing and adaption of the existing indexing using the stepwise refinement process. If we include new FAQs about a specific topic an adaption of the indexing of some other FAQs has to be done to consider the more specific new FAQs.

### **User not satisfied with the answer**

At the present time we cannot provide further statistical data about the user satisfaction with the answers. The major problem to evaluate the user satisfaction at the moment is caused by the small number of people who give a rating to the individual FAQ. Less than 3 percent of the readers of the FAQs gave a rating for the individual quality of this answer. Any analysis would not be representative.

### **4.3. Emergency events/ critical questions**

The system detected 2 questions which contain the words “should I kill myself?”. We provide a special answer for these kind of questions with further information where to find professional help. However, the analysis of the logfile shows that the user tried several unusual questions and might have tested the response of the computer system to such extreme questions.

#### **4.4. Privacy issues**

Using the questions-answering system we have no further personal information about the person. However we might be able to identify the provider of the client, if we were forced to do so by officials or in case of an extreme emergency. We provide a privacy statement to inform our users of the webpage about this issue. We think the we provide all possible efforts to protect the privacy of our webpage users.

## **5. Direct feedback by the users using email, phone, fax or other options**

At the present time we received only one email by a user who was not satisfied with the specific information we provided about tinnitus aurium (chronic ear noise). He asked us for better coverage of this topic. We will try to use this feedback to include more FAQs in the near future. Other options of user feedback (phone, fax, visits to our offices) have not been used by the users.

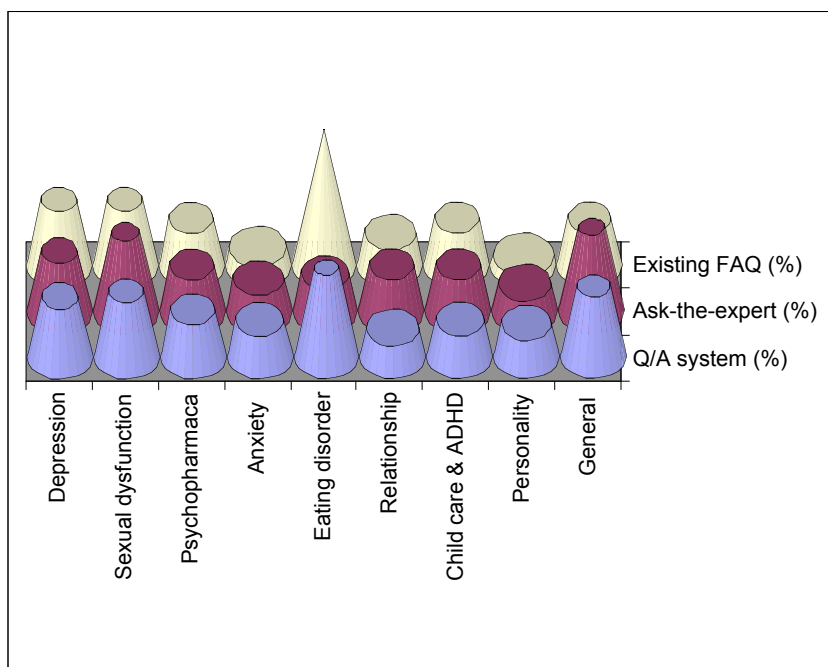


## 6. Consequences of the stepwise-refinement analysis

At the present status of the project we are already able to meet the basic interests of a broad range of our clients. Our existing content plan which was developed according to our own clinical experiences and expectations of possible user interest (see deliverable D2.1) meets the actual requests of our webpage users very well. Our existing content seems to be very well adapted to the questions sended to the Q/A system and the human expert team.

Topics	Q/A system (%)	Ask-the-expert (%)	Existing FAQ (%)
Depression	13	13	14
Sexual dysfunction	14	17	14
Psychopharmaca	10	10	10
Anxiety	8	7	4
Eating disorder	19	8	29
Relationship	6	10	7
Child care & ADHD	8	10	10
Personality	7	6	2
General	15	18	10

This can be demonstrated if we finally compare the different topics of interests with the questions posted to the computer question-answering system and the requests for information sended to our human expert team with the existing content of FAQs.



We have an excellent coverage about eating disorders (more than 200 FAQs), so most users do not have to ask a question to the human expert team. But this content area is also very popular to ask a question to the question-answering system (Q/A). This demonstrates the

options to use such a technical system if a reasonable amount and quality of content is available.

We have not yet provided very much content about anxiety disorder or personality disorder. This was mainly influenced by the need to set priorities and a prior analysis of existing webpages for these areas of content. We will provide improve our existing content for these areas in the course of the project.

The present status of existing FAQs for our project will be optimized in the course of the next months. We will include a reasonable number of the questions and answers to our expert team in the dynamic database for the natural language question answering system. At the present time new FAQs are included nearly daily.

The process of stepwise refinement will be continued for the whole duration of our project.

We try to get more user feedback for the individual FAQ evaluation and the user evaluation with an online questionnaire at the present time. But we are also aware of the fact that people with mental health problems might be very sensitive and try to avoid to participate in any further evaluation process.