Users requirements for personalised virtual digital libraries

Roberto Caldelli[1], Cezary Mazurek[2], Paolo Mazzanti[1], Tomasz Parkoła[2] and MarcinWerła[2]

[1]Media Integration and Communication Centre Firenze, Italy
[2]Poznan Supercomputing and Networking Center, Poland

Abstract: This paper presents results of the survey which was performed in the frame of the European project named ENRICH (ECP-2006-DILI-510049), funded under the eContent*Plus* programme. The aim of the survey was to acquire users' preferences regarding various aspects of two essential digital library functions: individual collections (static and dynamic) and individual virtual documents. Overall number of the survey responses reached 459 responses gathered from digital library users from 12 European countries. The image of users' preferences formed on the basis of responses lead to the definition of requirements for the creation of personalised virtual digital libraries. In this paper we present adopted procedure for performing the survey, analysis of the responses and final conclusions regarding personalised virtual digital libraries.

Keywords: digital library, personalised virtual digital library, individual collection, individual virtual document, ENRICH project, digital objects aggregation, virtual research environment.

1. Introduction

The European project ENRICH, funded under the eContent*Plus* programme, aims at creation of the virtual research environment for study of historical cultural heritage documents gathered from various European cultural institutions. The project builds on the Manuscriptorium portal, initially developed by the National Library of the Czech Republic and AiP Beroun in scope of Memoria programme described in Knoll (2004). Such an environment needs a clear set of requirements regarding particular features it will provide. In case of user personalization area of the ENRICH project there are two major features to be focused on: individual static and dynamic collections and individual virtual documents. It was decided that in case of these two major features the requirements will be specified by the end-users which are to be the final beneficiary and most interested party. A survey was conducted to elicit requirements.

The survey prepared within the ENRICH project was composed of five basic parts including: an explanation of the survey form structure and rationale, basic questions about the respondent and detailed questions about particular functions of the two personalization features given under analysis. Additionally, ENRICH partners translated the text of the survey into their native languages to receive more results and with better accuracy. The survey was filled in by over 450 respondents from 12 European countries participating in the ENRICH project. This gave a reasonable amount of data to be analysed and interpreted with respect to the final requirements. The percentage of filled surveys by respondents from particular countries is presented on Figure 1.
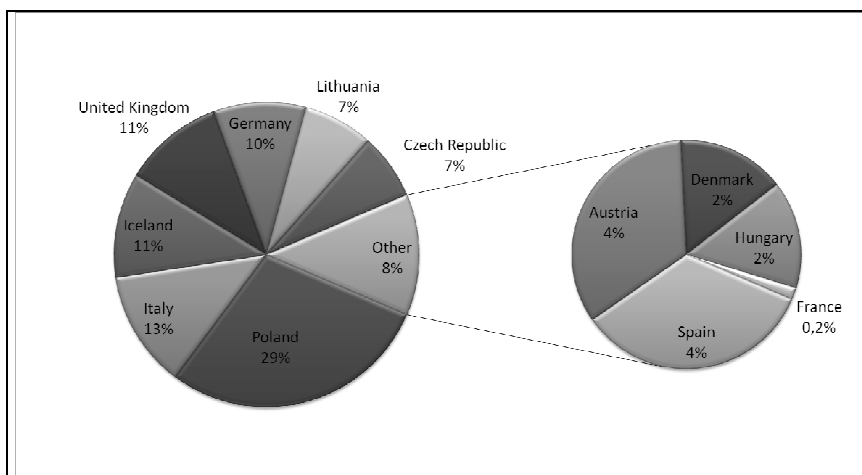
Figure 1 Percentage of filled surveys by respondents from particular countries.

This article gives an overview on the scope of the survey, results analysis and final conclusions; in particular section 2 describes the survey structure, section 3 presents the most important analysis and section 4 includes final conclusions and advices regarding analysed features.

2. The questionnaire structure

According to the MINERVA EC (2003) *Handbook for quality in cultural websites: improving quality for citizens*, "A user is a professional person or not, a specialist or not, who casually or with specific aims, occasionally or systematically uses the Cultural Web Application. User identity is extremely variable depending on cultural profile, aspirations for cultural growth, professional aims and even momentary curiosity" (see also in MINERVA EC (2008)). Therefore, a quality website in general or a quality digital library in particular, must be user-centred, "taking into account the needs of users, ensuring relevance and ease of use through responding to evaluation and feedback" (5th MINERVA Quality Principle, in MINERVA EC (2003-2004)).

In order to realize a high quality European Digital Library, really user-centred, based on his/her needs, tasks and behaviours, the Enrich Questionnaire was a strategic method of investigation of the users requirements for personalised virtual digital libraries. The survey is globally composed by 23 structured multiple questions**,** easy and quick to be completed.

The access to the questionnaire consists of a specific on-line webpage that presents the "*Digital library readers demands/needs survey*" in many European languages for the different users, to ease questionnaire understanding also to people not very familiar with English. Primarily, in order to fill in the questionnaire, the user can select his/her specific country in a reference table for the choice of the language. The introduction is structured to explain the questionnaire as a part of the European ENRICH project and the user is invited to take time to complete the survey and notified that this will help the ENRICH project to address the issues s/he is interested in and ensure that the developed technology meets his/her needs.

The first part of the survey consists of 3 questions, specifically created and formulated in order to collect information about the user experience. In particular, information about the type of respondent as well as his/her

2

experience, that may help to better interpret the survey results. Firstly, it is asked which group of digital library users the user belongs to, in order to identify the user in the following macro categories: researcher, teacher, student and other. Secondly, it is required information about the level of user's experience with digital libraries: a regular user, occasional regular user, a rare user. Finally, it is analyzed the level of user's experience with Manuscriptorium digital library: regular user, occasional user, rare user, never used before.

The second part of the survey, questions 4-23, concerns with individual collections (IC) and virtual document (VD). In order to design the functionalities on the basis of the user needs, the questions ask how the user, as a digital library reader, considers the static and dynamic IC and VD (necessary, useful, neutral, not useful, redundant) and where s/he would need to keep them for further usage (digital library reader account locally, personal computer's hard drive, not need to save IC/VD for further usage). Another important aspect of the personalisation is to grant an interaction of the user with the digital library environment, for this reason it is asked if the user is interested about the possibility to receive a notification (e.g. e-mail notice or RSS feed) about new documents in his/her dynamic IC and VD, to have personal notes about any of his/her IC/VD, or files with additional information to any of his/her IC/VD. The goal of the questionnaire is also to prepare the basic architecture of the search inside the Digital Library, so it is asked if the user is interested to have the possibility to search his/her IC/VD and if s/he is interested to share (not edit) own IC/VD with other readers. In this field is also useful to know who can see the shared IC/VD and is asked to whom the user prefers to make visible his/her individual shared collection and virtual document. User can select one of the following answers: all readers and I do not need the possibility to hide any collection/document, all readers and I have the possibility to hide/unhide any collection/document, selected readers only. Other specific questions are about the possibility to give a copy of the IC/VD to any other reader and what information to share in regard to the IC/VD (the documents list only, the documents list notes and the additional information attached, the documents list and, depending on his/her decision, notes and additional information).

It is necessary to provide to the users the capability to interact with the digital library and by this survey they can say how the services should be improved, so that their expectations can be better met. That is why a question analyzes the importance of the media types within the virtual digital library. User can select one of the following answers for images, texts, audio and video: necessary, useful, neutral, not useful. At the end of the questionnaire, the user has the possibility to write additional comments in a dedicated form, that allows the user to provide further interesting indications on the matter that s/he thinks are not completely covered by the survey. S/He can also write the e-mail address, to receive a copy of the final report with the interpretation of the survey results. In order to give a service user centred, usable and accessible, it is basic to remind that below the reference table of the questionnaire there was an additional link for accessing "the digital library readers demands/needs survey statistics", where the users can directly see and check the survey results, divided by Country, Number of filled surveys, Number of distinct IPs.

3. Analysis of results

The survey was basically directed to researchers, teachers and students. Researchers form the primary group of users, interested in new functionality connected with research environment being built within the ENRICH project.

For researchers, the digital content available in the research environment will need appropriately adjusted features, which will allow performing specific research activities. Teachers may not only be interested in finding and using educational materials already available in the system, but also in using the functions allowing them to prepare course-focused materials for their students. Students will benefit from all the materials (e.g. prepared by their teachers) and will have the opportunity to use the environment to raise their skills and enhance knowledge. Naturally, beside these three basic groups, other interested users were also invited to fill in the survey. These are usually users interested in discovering cultural heritage documents such as manuscripts, early printed books, archival papers, etc. and are mainly amateurs, which would like to take a look at the most impressive documents and discover some interesting facts.

The results from the first part of the survey which was used to obtain basic information about the respondents show that the final requirements should be reliable and representative for the target users as majority of respondents was familiar with digital libraries and also convergent with survey target users group. It also appeared that the researchers are the most experienced group of respondents with digital libraries. On the other hand students are the least experienced respondents group.

The most important parts of the survey queried users about individual collections and individual virtual documents features. Users were given a detailed explanation of these two terms in order to interpret the questions appropriately.

Individual collections allow any reader to create and maintain his/her own set of documents within his/her profile and to share it with other readers. Documents grouped in individual collections are always easily available to the reader for instant usage. There were two types of individual collections given under analysis:

- static individual collections – created by readers by adding directly selected documents;
- dynamic individual collections – created by readers by specifying criteria for the collection membership. For example, a dynamic individual collection may be specified as all the documents created before 1450 and related to alchemy. As the content of the digital library changes, the content of the dynamic collection may also change.

More information about the implementation of such collections in distributed environments can be found in Mazurek (2006).

Virtual documents are documents created manually with parts (e.g. selected pages) of other digital library documents. For example a reader can build a virtual document demonstrating the art of illumination for a selected period (e.g. from the most representative parts of illuminated documents for this period). Additional information may be also attached to the virtual document by its creator.

The basic question in regard to individual collections was about the necessity of this feature. Overwhelming majority of respondents stated that both dynamic and static individual collections would be either useful or necessary when working with digital library. For individual documents the initiating question investigated the importance of the media type to use for individual documents. There were four possible answers: text, image, video, audio. It appeared that text media type is the most necessary media type to work on. The image media type situated just after the text. Video and audio media types received very similar marks – useful and neutral, but not necessary. Figure 2 depictures

responses concerning the necessity of individual collections and Figure 3 the importance of the media types for virtual documents.
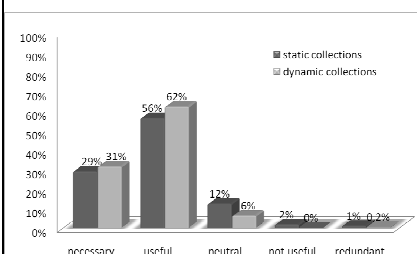

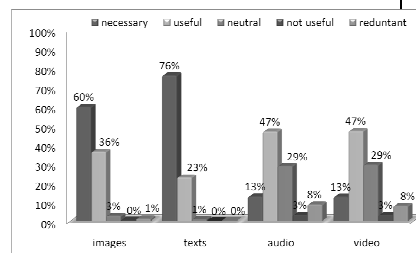Figure 2 The necessity for individual collections feature

Figure 3 The importance of the media types for virtual documents feature

Second part of questions concerned interaction of the user with the digital library environment; in particular respondents were asked about the storage place for individual collections and documents, changes notification for dynamic individual collections and search mechanism for static individual collections. Both for individual collections and documents respondents preferred to store information about them in the on-line facilities (e.g. reader account) rather than in the local storage facilities (e.g. hard drive). Changes notifications in the dynamic individual collection (to which changes notifications may only apply) were evaluated as a useful function. In this case group of users identified as students was the one which mostly voted for this function. This is probably because of a higher popularity of notification tools (e.g. RSS feeds) among young people. The last, but a very important function evaluated within the interaction part of the survey concerned the search mechanism. It appeared that it was rated by most of respondents as necessary. Major part of remaining respondents recognised this function as useful. It is interesting to note that this function was evaluated as the most necessary in respect all functions connected with individual collections. Therefore the results prove that search function is a very important function connected with individual collections. We may also indirectly conclude that users are willing to create individual collections containing many documents therefore they require search functionality for them.

Third part of questions concerned individual collection and document notes functionality and was investigating not only the possibility to textually describe particular individual element, but also attach to it additional items (e.g. files) which are related to the source material. There were questions asking respondents for their opinion about textual personal notes connected with: an individual collection, a whole document and part of a document. There were also questions about necessity for attaching personal files to: individual collection and whole document. Results for all these questions are presented on Figure 4 and 5. According to respondents, possibility to have personal notes and personal files connected with individual collections is useful. When comparing results for these two functions it appears that personal notes are more important for users than personal files. When analysing gathered results in regard to individual documents it appears that respondents created by their answers a kind of priority list for three investigated functions. The most needed function is the one that allows users to add notes about the whole individual document, the second one is a function for adding notes to particular parts of document, and the less required function is the one that allows users to attach files with additional information to individual virtual document. The results are

quite natural because they prioritise functions from the most simple one (notes to the whole document) to the most advanced one (attaching files to documents).
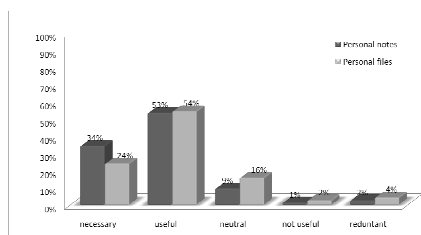

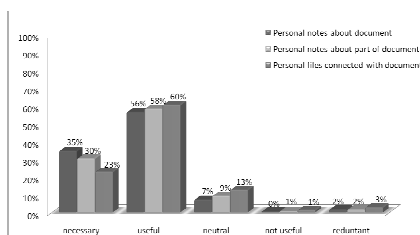Figure 4 Personal notes and files necessity for individual collections

Figure 5 Personal notes and files necessity for individual documents

The last set of questions investigated community building functionalities such as sharing individual collections and documents with other users, controlling visibility of particular parts of shared content and even giving a copy of particular collection or document to other users. Functionality which would allow users to share individual collections and documents with other users was evaluated by most of respondents as useful. Around one-tenth of respondents stated that these are necessary therefore these functions are not mandatory and should be introduced only when other, more important functions are ready. Three additional questions were investigating specific aspects of the functionality for sharing individual collections and documents.

First aspect was connected with the visibility of shared collections and documents to other users (results are depicted on Figure 6 and 7). It appeared that most of respondents would prefer to choose which individual collections and documents will be available to all other users and which will be hidden from them. Majority of reminder respondents would prefer more restrictive access to their individual collections and documents (grant access to selected users only). Additional analysis of relations between the question concerning sharing function and the question about visibility of collections and documents show that the more respondents were sceptic about sharing their individual elements the more they wanted to secure their elements from other users (by selecting visibility option giving them more advanced possibilities). This fact suggests that responses are quite consistent therefore more reliable.
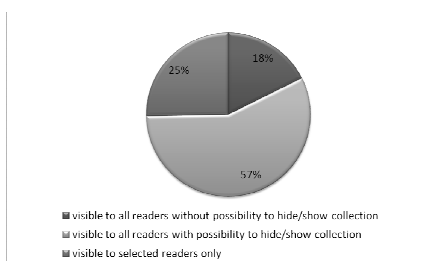

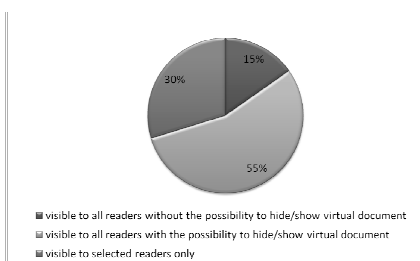Figure 6 Results for visibility function for shared individual collections

Figure 7 Results for visibility function for shared individual documents

Second aspect concerned the issue of additional elements (notes, attached files) which should be shared with other users (results are depicted on Figure 9 and 10). Both for individual collections and documents most of respondents selected the option to have the full configuration possibilities regarding the

elements to share. Only minor part of respondents would share the document and additional information without any constraints.
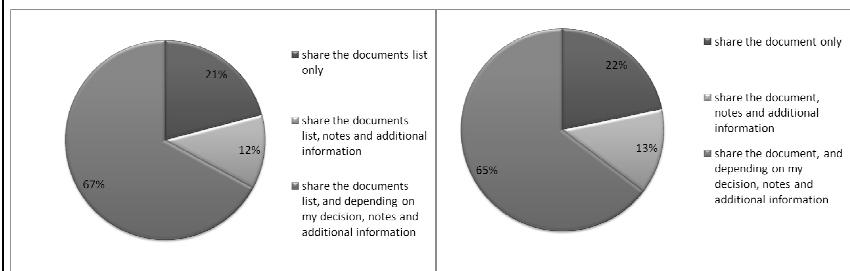

Figure 9 Individual collections sharing function

Figure 8 Individual documents sharing function

Third investigated aspect was about giving a copy of personal collection or document to other users. Obviously, when sharing a collection with the public, registered users cannot edit the content of the element (individual collection or document). The purpose of copying of an element is to create a new instance of the element and to enable a selected user so he/she can start to use it as his/her own. Answers given by most of respondents in regard to this functionality show that it would be useful. When comparing these results with the question about the necessity of sharing individual collections and documents it appeared that for respondents the functionality to give a copy of an individual collection or document is more important than the functionality which allows to share it.
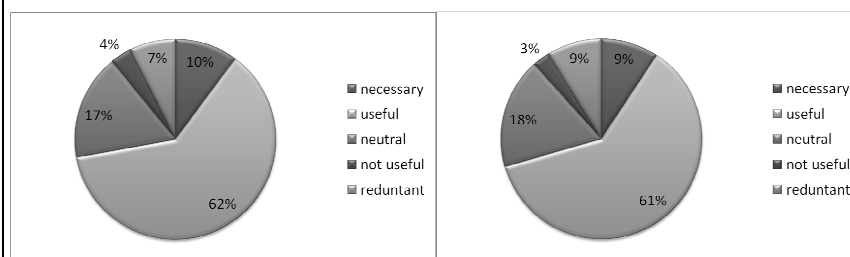

Figure 11 Functionality for giving a copy of individual collections

Figure 10 Functionality for giving a copy of individual documents

4. Conclusions

Results of the survey performed in frame of the ENRICH project gathered opinions of experts, professionals, students and hobbyists from 12 European countries. The analysis based on the results gave reliable overview of the users requirements concerning user personalisation features in digital libraries. The survey investigated two basic personalisation features for digital libraries: static and dynamic individual collections and individual virtual documents. The results, as interpreted in the previous section, proved the correctness of the intentions to implement the new features. Full analysis of the survey is available as the ENRICH project deliverable D 4.1, Caldelli (2008).

At the very basic level the results proved that investigated features for user personalisation in scope of digital libraries are reasonable to be introduced. In particular, further activities should be focused on providing individual collections functionality and virtual documents functionality with a special

reflection for processing and manipulating text documents and providing community building facilities.

The textual form of documents in digital libraries seems to be the most attractive one for research and other types of use. It also appeared that respondents prefer to use on-line services for storing collections and documents they will work on, which is quite natural because of the convenience for the users. Change notifications for dynamic collections and searching mechanism for individual collections also received positive marks from respondents and seem to be useful.

Personal notes and files for individual collections and documents were evaluated as valuable. In case of individual collections respondents preferred personal notes over personal files. In case of individual documents respondents created a priority list where particular functions were ordered as follows (starting from the most required function): personal notes about whole document, personal files connected with part of a document and personal files connected with whole document.

Community building functionalities were evaluated as useful with some sceptic attitude for opening all individual elements to the public. Respondents would rather keep the visibility of their elements in control and grant appropriate rights to selected elements or users only. What is interesting, the functionality to give a copy of individual collections or documents higher marks than the functionality which allows sharing individual collections and documents.

Based on the survey results, the most important subset of investigated functions will be introduced in the virtual research environment which is built in scope of the ENRICH project. Further activities should be also focused on providing the remaining, positively evaluated functions.

Acknowledgements

References

Caldelli, R., Kligiene, R., Mazzanti, P., Parkoła, T., Psohlavec, T., Werla, M., (2008). Definition of requirements for the creation of personalised virtual digital libraries. ENRICH project deliverable D4.1. http://enrich.manuscriptorium.com/files/ENRICH_WP4_D_4_1_final.pdf

Knoll, A. (2004). Digital Access to Old Manuscripts, *Linguistica Computazionale*, 2004, Vol. XX-XXI, Digital Technology and Philological Disciplines, 277 – 286.

Mazurek, C., Stroiński, M., Węglarz, J., Werla, M., (2006). Metadata harvesting in regional digital libraries in PIONIER Network, *Campus-Wide Information Systems,* 2006, Vol. 23, No. 4, 241–253.

MINERVA EC, (2003) *Handbook for quality in cultural websites: improving quality for citizens*, *Identity* http://www.minervaeurope.org/publications/qualitycriteria1_2draft/cap1.htm

MINERVA EC, (2003-2004) *Cultural Website Quality Principles* http://www.minervaeurope.org/publications/tenqualityprinciples.htm

MINERVA EC, (First edition September 2008) *Working Group Handbook on cultural web user interaction* - Quality, Accessibility and Usability http://www.minervaeurope.org/publications/handbookwebusers.htm