ISSUES TO BE ADDRESSED FOR TRANSFORMING A DIGITAL LIBRARY APPLICATION FOR EXPERTS INTO ONE FOR FINAL USERS

M. Agosti a, *, L. Benfante a, M. Manfioletti a, N. Orio b, C. Ponchia b

a Department of Information Engineering, University of Padua, Via Gradenigo 6/a, 35131 Padua, Italy - (agosti,benfante,manfioletti)@dei.unipd.it
b Department of Cultural Heritage, University of Padua, Piazza Capitaniato 7, 35139 Padua, Italy - nicola.orio@unipd.it,ponchiachiara1@gmail.com

ABSTRACT:

This paper reports on the effort we made in adapting and opening a specialist tool, focused on illumination and designed purposely for scholars and researchers, in order to be suitable also for the general public. We describe the ongoing process we are conducting: the adaptation and the improvement of the IPSA digital archive using the results we collected after several sessions of user interviews, following suggestions of both scholars and simple users. We discuss user studies dynamics, that we consider as a loop-interaction, and the consequences that they entail upon the system design.

1. INTRODUCTION

Currently one of the most significant goals for curators and providers of digital cultural heritage collections is to increase the engagement of users and communities with digital humanities collections, and possibly to gain the interest of new user categories. To accomplish this difficult task, it is essential to draw a precise profile of the different kinds of user and to study new ways of requirements elicitation, to address the exact needs and expectations of the final users.

We carried out the research work presented in this paper using, as a case study, the *Imaginum Patavinae Scientiae Archivum* (IPSA) digital archive. IPSA makes available to scholars a collection of illuminated manuscripts (Mariani Canova, 2002), which includes botanical codices produced mainly in the Veneto region during the XIV and XV centuries (IPSA website). The IPSA archive was initially designed for a specialist public of scholars and researchers (Agosti (et al.), 2003) with specific interest in manuscripts and illuminations.

The effort to address the needs of a wider variety of users is one of the main goals of the CULTURA project (CULTURA Project website), for which IPSA has been selected as one of the representative case studies. Hence it was decided to open the collection to other categories of users, such as non-domain professional researchers, student communities and the general public. This new task required the identification of the needs, wishes and preferences of these new categories in order to define the required changes and improvements to IPSA (Agosti (et al.), 2011). To this end, we carried out an evaluation with different categories of users, comparing the requirements of professional researchers with the ones of interested general public (at the moment represented by students in related disciplines).

The efforts made to open the IPSA digital collection of illuminated manuscripts to a wider public are bringing about useful reflections and suggestions that can contribute to this field of investigation.

Section 2 briefly reports on the characteristics of interest of the digital archive that constitutes the case study. As can be seen in Section 3, interactions with both master students and professional researchers had been carried out, in order to receive valid inputs from different user categories. The suggestions received from the first interactions had been suddenly incorporated in the software application which manages the digital archive, and then tested in another evaluation opportunities, creating a fruitful loop of user interaction described in Section 4. Section 5 highlights additional issues that we faced in opening up IPSA and that constitute results that can be of general interest for curators and providers of cultural heritage collections, and Section 6 discusses on the process of extending a digital resource created for domain experts to the one for other categories of users, including the general public; the process requires a number of steps that are introduced and that represent a novel result, in fact, to our knowledge, no similar effort has been documented in the literature making no feasible for us to compare the results of our studies to other ones.

2. CHARACTERISTICS OF INTEREST OF THE DIGITAL ARCHIVE

The digital archive includes astrological and botanical manuscripts belonging to some of the most important libraries in Europe and the world.
The archive has been initially designed to meet the requirements of scholars in history of miniature that are mostly interested in analyzing images, their style, their elements and possible relations with other images belonging to different manuscripts. Because of that, the metadata included in the archive are both a description of each of those manuscripts together with

---

* Corresponding author.

information on the content and the provenance of each digital image of each manuscript. That makes clear that the "content" of interest for the scholars is constituted by the images. In fact the users are interested in the images ("content") rather than in the texts of the manuscripts. Further information on initial requirements of IPSA together with details of the characteristics of the interface and available functions of the software application initially designed and developed are reported in (Agosti, M. (et al.), 2012).

The manuscripts that are of interest for the CULTURA project are the 10 botanical manuscripts that are included in IPSA, because of that, the details and information on those manuscripts have been summarized in Table 1 where the author, the title, the call number, and the number of images of each manuscript are reported.

Author	Title	Call Number	Number of Images
Anonymous	Tractatus de Herbis	Paris, Bibliothèque de l'Ecole des Beaux-Arts, Masson 116	338
Nicolò Roccabonella	Liber de Simplicibus	Venezia, Biblioteca Nazionale Marciana, Lat. VI. 59	983
Macer Floridus; Serapiom; Guarnerino da Padova	De viribus herbarum; Aggregator; Herbarius	Bergamo, Biblioteca Civica Angelo Mai, MA. 592	185
Pseudo-Apuleio	De herbarum virtutibus	Padova, Biblioteca dell'Orto Botanico, Ar.26.n.1283	200
Pseudo-Apuleio	Herbarius	Firenze, Biblioteca Medicea Laurenziana, Plut. 73.41	89
Pseudo-Apuleio	Herbarius	London, British Library, Cotton Vitellius C. III	84
Pseudo-Apuleio	Herbarius	Montecassino, Archivio della Badia, Casin. 97	46
Pseudo-Apuleio	Herbarius	Wien, Osterreichische Nationalbibliothek, Codex Vidombonensis 93	132
Serapiom	Liber Agregà	London, British Library, Egerton 2020	53

Table 1. Botanical Manuscripts included in IPSA

3. USER STUDIES

Two parallel evaluations have been carried out on the IPSA digital archive. Although there were large differences between the user groups in the interest of the digital archive content and in the motivation for the interaction, the user studies highlight some common outcomes regarding the difficulties of user-driven development, which appear to be more relevant in the case of cultural heritage applications.

User requirements were elicited in two ways. Firstly, thorough interviews with professional researchers were held, both with the domain professional researchers involved in the creation of IPSA from the very beginning, and with non-domain researchers expert in the field of History of Medieval Art but not acquainted with the IPSA collection and the History of Illumination in general. All the interviews were held on an individual basis.

With regard to the student community, the organization of different trials on a group basis with IPSA was thought to be the best solution for obtaining useful input and information. The trials were held in two different moments: the first time the trial was carried out with two groups of university students in October and November 2011, the second time was in April 2012 with a different cohort of university student. The IPSA trials turned out to be a fruitful way to have a useful feedback from the students, and they helped to improve the application and its functions; the results of the trial we conducted in 2011 are mainly reported in (Ponchia, 2012). In the following we report on the additional insights we gained with trial and interviews conducted in 2012.

3.1 Professional Users

For the interaction with the professional users it was decided to involve three researchers in History of Art specialized in different domains. Because IPSA was created purposely for specialists in History of Illumination, two of the professional users chosen for the interview were scholars expert in this research area. The first one, Federica Toniolo, is Associate Professor of History of Medieval Art and History of Illumination at the University of Padua. In her career she thoroughly studied the Italian Medieval and Renaissance Illumination, particularly the Ferrara production. Her more recent studies are about illuminated manuscripts and incunabula made in Northern Italy in the XIV and XV centuries, especially choir books, and she is currently involved in the cataloguing of the illuminated medieval manuscripts of the Veneto libraries. The second researcher in History of Illumination is Silvia Fumian, Research Fellow at the University of Padua, whose research interests are about illuminated manuscripts made in Veneto in the late XIV and early XV centuries, and in Ferrara and Mantua in the XV century.

In addition, since it is essential to engage with professional users, from different domains, it was decided also to involve a researcher in History of Medieval Painting, Zuleika Murat, currently a third year Ph.D. Fellow at the University of Padua, with a thesis on the XIV century Paduan painter Guariento di Arpo. The aim of her study is to analyse the painter's catalogue, and reconstruct the connection between paintings and their original settings and functions, the typology of scattered panels, and the influence the patrons had on the choice of iconography and eventually of style. It was thought that, thanks to her different field of specialization, she could have different expectations, needs and wishes about the digital system and its functions, and that the interaction with her could foster a new original reflection on how to improve IPSA.

The interviews took place on May 2012 and were carried out on a group basis. The interaction followed a three-fold structure:

- Short introduction to present the results of the work done.
- Presentation of the new IPSA version and its functions, such as the new image search procedure in a manuscript catalogue file and the new way of creating a link between two illuminations.

- Open discussion.

It is important to note that professional researchers are already using a number of digital resources for their studies, including a prior version of the IPSA collection. The analysis of requirements could then be carried out also comparing IPSA functions with the ones of systems already available on the Web.

3.2 Master Students

Since the IPSA collection is meant to be open to general users due to its involvement in the CULTURA project, trials with other categories of users than professional ones are required in order to obtain input and information about the needs of the general public. For this reason we chose to present the IPSA system to a group of students who are specialized in other fields of the humanities, for the purpose of receiving direct comments and feedback from a significative subset of general users.

Therefore we organized two trial sessions in which students were asked to use the IPSA archive pursuing specific tasks we assigned. Every task was designed in order to verify the different user dynamics and to understand which kind of obstacles a general user can find in using a specialist tool. After the two sessions we organized a focus group during which students could discuss the whole experience underlining both positive and negative aspects of the IPSA digital archive and suggesting modifications to improve the environment, the browsing experience and the research function.

The IPSA trial that was conducted in April 2012 was developed specifically for a group composed by 25 master students majoring in History and Management of Archival and Bibliographic Heritage and in Modern Languages who were attending the course on Database and Internet held by Maristella Agosti, professor of Computer Science at the University of Padua. We chose such a sample because we considered that users who have specialization in a different field of the humanities than history of art can provide valuable contributions in the improvement of the whole environment, since even if they do not have particular interests in illumination or history of art in general, nevertheless they have a thorough knowledge and specific expertise which allow them to notice details who may have been ignored before by both art historians and computer scientists; interviewing and interacting with such a subset of users, a whole new perspective was given, allowing a significative improvement of the entire system thanks to the disclosure of new points of view.

The trial was composed by two parts that took place in April 2012 within two weeks. In each part of the trial students had to accomplish some tasks that were designed purposely to interact with the IPSA system using its specific functionalities. The trial was structured in three units:

1. **Introduction** to the IPSA archive: in this unit a specialist (a PhD student in History of art for the first trial, and a linguist for the second trial) explained some technical characteristics of the IPSA system, in order to clarify some crucial aspects such as specific terminology or normalization issues, and mostly to illustrate how to use the archive properly.
2. Description of the **tasks**: in this unit students were asked to actually use the IPSA digital archive following a series of defined assignments purposely designed in order to experience the most important IPSA functions such as research of images and manuscripts, the possibility of establishing links between two images which are considered mutually related and the browsing tool which allows the user to easily scroll the manuscript pages in order to find the required illumination.
3. Filling of a **questionnaire**, in order to obtain structured feedback: after each trial the students answered an evaluation questionnaire developed by a team of psychologists of the University of Graz who are partners of the CULTURA project (CULTURA Project website: partners). The questionnaire aimed at evaluating the interaction with the system and the acceptance of the users. The questionnaire used, first covers questions on demography, knowledge of technology, in general, and the experience with the digital collection, in particular. This is followed by two short items querying the level of support of individual and community needs/objectives by the digital collection. Subsequently, usability and user acceptance are queried. The use of standardized questionnaires provides the advantage of a sound assessment on a certain evaluation parameter, and the possibility to compare evaluation results even with other studies.

After the two trial sessions we organized a focus group with all the students who took part to the study, in order to have a wider discussion about the IPSA experience and about possible issues or suggestions. The debate about the user experience was very interesting and it highlighted several modifications that can be made in order to improve the system both from a graphic and a technical point of view. The students observations underlined the importance of working on simplification, in order to make a specialist tool more involving for a general public. It is a fact that this kind of audience needs to be charmed and intrigued by an agreeable environment whose functioning is clear and immediate. Simple users generally do not have particular interests in illumination or manuscript but they can find fascinating to browse a beautiful collection if the entire experience is made easy and smooth.

3.3 Closed Loop Development

A frequent interaction with users during their evaluation of IPSA allowed a very fast evolution of some features as consequence of the user difficulties we observed. During the evaluations the principal causes of problems were:

- Low performance on rendering the pages.
- Difficulties on maintaining the context during the navigation.
- Difficulties on understanding some elements of the interface.

These observations on a very closed evaluation-development loop produced a number of enhancements on the IPSA system and interface that were made directly available to users during the second round of evaluation.

First of all, the list of images of a work evolved from a simple wall of all images, very heavy to render and to explore, to a partial wall of twenty images, very fast to render. The user can request more blocks of images clicking on a button that activates an AJAX request, and it is still very fast to render, without losing the current context. As a (currently) last step, on the top of the wall has been added a textual drop-down list of all images of the work, ordered by page; selecting one element from the list, the user can jump to the block containing the thumbnail of the desired image, as depicted in Figure 1. At this stage of interaction with the IPSA application, the user can request to examine a specific image; the result of this request is the presentation of all the metadata of the image together with the possibility of seeing the complete image as in Figure 2.
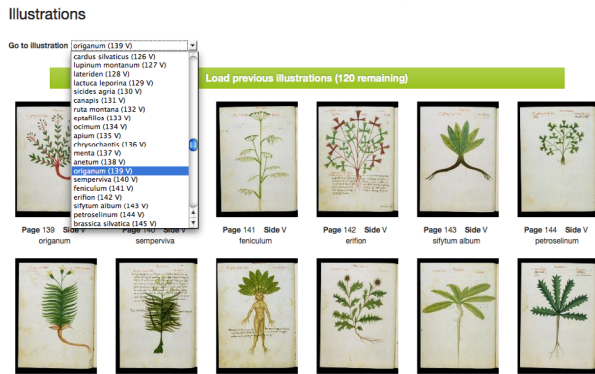
Figure 1: The textual drop-down list of all images of the work, ordered by page, and the block containing the thumbnail of the desired image are shown.
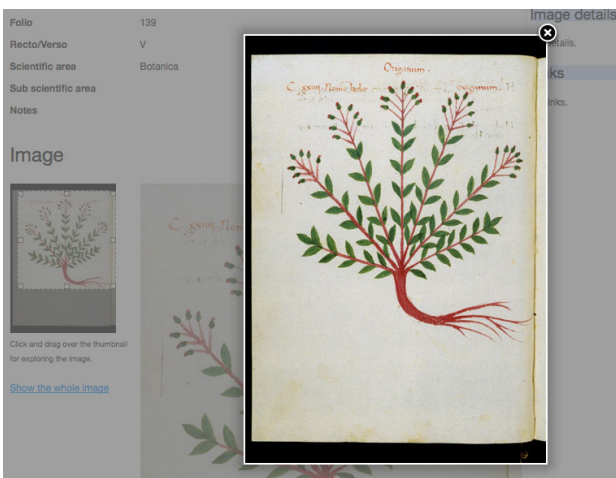

Figure 2: Presentation of all the metadata of the image of interest together with the presentation of the complete image.

For reducing the loss of context when the user navigates from the page of a work, to a page of an image, and back to the work, the position in the wall of images is maintained during all the navigation in the same work and its images. Using the experience acquired on the page of a work, almost the same interface has been implemented for the wall of images of the result of a search of images.

The process of the creation of a link between images was very difficult to understand because of the simple, mostly textual, interface available. This operation is a multi-step process in which the user, starting from an image, can navigate and use all IPSA features for searching the ending-point image of the link. During all the process, on the top of the page, is showed a box with the status of the operation. Following the observations of the user, this box has been enlarged and now it includes the thumbnails of the selected images, some help text, and large very explicative buttons for completing the link, or for deleting the operation. Even the navigation has been modified a little bit: now after the starting of the link creation, i.e. after selecting the first image, the user is moved to the advanced search for images, with the form totally empty, and still no results, as shown in Figure 3, where the query - "rosa alba" - has been inserted in the form but the system has not been asked yet to process it. This has been evaluated by the users as a very comfortable, understandable and useful type of interaction for this situation. When the user selects the image to link, among the ones retrieved through the advanced search, the system copies the image on the screen, and gives the possibility to the user to complete the link, as it is shown in Figure 4.
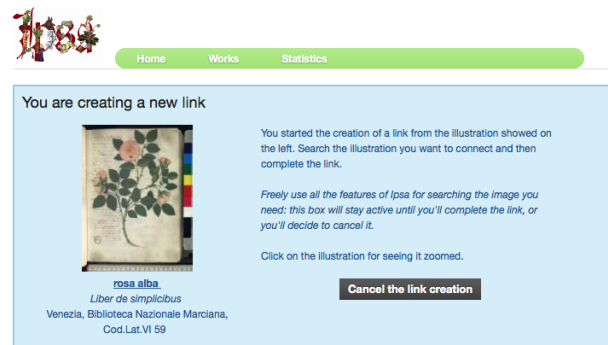
Figure 3: The user selects an image from which to create a link, and he can conduct an advanced search to find the related one.
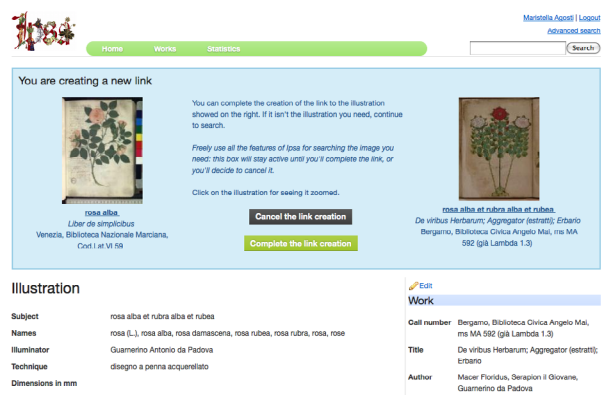

Figure 4: The presentation to the user of the image to relate to the initial one.

The choice of modifying the user interface while carrying out the evaluation showed to be successful. Students felt that their comments and suggestions were taken into account, and where increasingly motivated in interacting more deeply with the IPSA digital archive. Nonetheless, it was crucial to incorporate the input given by the students without losing contact with the professional researchers needs. That is why it was decided to carry out a new in-depth interaction with professional users. We gave to the IPSA user studies a loop shape. The circle we created starts from the professional users, it goes through the general public filter and it comes back to the original target. Thanks to this passages the IPSA archive takes in account both needs of researchers and simple users, creating an environment that can be used as a professional resource and at the same time can be browsed by general users as a catalogue of beautiful images.

4. CONSIDERATIONS ON USERS INTERACTION

The choice of implementing the required changes while carrying out the evaluation had an additional motivation. In fact, one of the most striking thing that came out from the discussion results, is that for both groups of users, it was difficult to give suggestions on a large scale.

In the case of the professional researchers, although the interview aimed at fostering their reflection on their general research needs, that CULTURA wants to address with new specific tools, the focus was on improvements of minor importance. For example, Fumian, Murat and Toniolo said that the links presented in the catalogue file of an illumination

should be shown in a different way: actually, it is difficult to distinguish the manuscript library and call number from the text annotated by the user. So the interviewees suggested that the manuscript call number should be justified, while the text annotated by the user should be a flush left italic text, preceded with the heading "Notes" and written in a different colour (e.g. dark green). Another example is that the three researchers paid a lot of attention to the written instructions, and suggested many changes to the text, often small refinements (e.g. "Start a new link" should be replace by "Create a new link"). It emerges clearly that they felt more self-confident in giving small advices concerning details of minor relevance, as if it was easier for them to imagine little improvements rather than picture big changes and new functions that could address their research needs.

During the interviews, we thus stimulate reflections of general requirements, asking general questions such as:

- Which kind of functions do you think can help you in your research?
- How could technology most meaningfully assist you in your research process?
- What kind of information and/or visualisations should the research environment make accessible in order to support your research process?

The answers were quite general. It was also noted that when the three interviewees tried to suggest improvements and additions to IPSA they took inspiration from tools they already are confident with, such as ARTstor (ARTstor website). For example, Toniolo said that she would like to have the possibility to create a folder where she can save her searches and decide which research results can be shared with other users and with whom, a functionality that already exists in ARTstor. This underlines one more time that lack of imagination and influence of previous models can be an obstacle in user requirements elicitation. So the interviews with the professional researchers prompted the need to find and develop new ways of stimulating users' imagination and a deeper reflection on their needs and wishes.

The same considerations apply to the student cohort. Most of the written comments collected through the questionnaires were about the interface layout (e.g. the number of characters in the text fields or the positioning of images) and the perceived easiness of use, which improved between the two trials thanks to the closed loop development approach. This tendency of concentrating on details was evident also during the discussion of the focus group. Comments regarded the position on screen of images, their dimension, the length of descriptive texts, and so on. These results are inline with considerations about the "bike-shed effect", which is a well-known drawback in computer software design (Fogel, 2005): there seems to be an inverse proportionality between the relevance of a feature and the time spent discussing it. Again, the choice of implementing directly the required functions was in the direction of reducing this effect, although it is normal that non-expert users concentrate on the interface and its details because they feel they do not have the competences for providing useful comments on the architectural and methodological levels.

Notwithstanding these considerations, there was a major difference between students and professional researchers for one IPSA function. It regarded their experience with the search tools. While specialists do not have problems accessing the digital content, thanks to the their deep knowledge of the domain, students sometimes had difficulties in carrying out search tasks. Difficulties were partially related to their lack of knowledge about illuminated manuscripts, because it is not trivial to understand how search results are ranked, which are the most useful search fields and how to address relevance of retrieved items.

5. ADDITIONAL ISSUES IN OPENING UP IPSA

At present the IPSA data are stored using a PostgreSQL relational database application (PostgreSQL website) and managed using also a Lucene index (Lucene website). In the near future some data will be managed by a documentation management system and by an annotation management system. So IPSA is already a system composed by different heterogeneous subsystems internally integrated.

The involvement in the CULTURA project and the goal of extending the categories of users including the general public, opened the door to various problems of integration with external systems, mainly for allowing such systems to retrieve and use data from IPSA, both metadata and images. The main problem here is not technical, but cultural and administrative.

Most of the images included in the IPSA collection have been obtained with the grant for using them in the IPSA application, but not for publicly distributing them, or for transmitting them to an external system on which we cannot guarantee an appropriate use. It is in fact a problem of copyright on the images, which will be addressed asking the copyright holders an extension of the grants for freely distributing at least a degraded water-marked version of the original image. As a sort of digital rights management (DRM) protection, the water-mark could even include the name of the user, or of the system, that obtained the image from IPSA.

Similar problems of copyright/privacy could apply to the metadata that the IPSA users add to the collection, in form of annotations, links, or other types of content that need intellectual protection. In this case the choice will probably be left to the user itself, that can assign different levels of sharing to his contributions. For example, he could create one or more groups of colleagues or students, and share his contents to the group(s) he prefers. Alternatively he could release his material as public.

Moving to the technical parts, at present IPSA has some interaction points with external systems through the exposition of some web services. Given the heterogeneity of the systems involved, the exposed services needed simplicity and independence from the hardware and software platform, and from the specific programming language. So we decided to implement them as RESTful (REpresentational State Transfer) web services, as any system capable of sending a simple HTTP (Hypertext Transfer Protocol) request to a specific URL (Uniform Resource Locator) could use them (World Wide Web Consortium website). At present services are available for obtaining the metadata of all the works in XML format (World Wide Web Consortium website), the metadata of a single work in XML format, and the thumbnails of a single work bundled in a file in zip format. The names of the thumbnail files inside the zip file are built using the unique identifier of the image that can be retrieved from the corresponding XML metadata.

Each service need HTTP authentication for being accessed. The authentication is necessary for assuring that the user (or system) owns the expected level of privileges for invoking the service and obtaining the requested data. For example, only a user with administrative privileges can obtain the full set of metadata.

As you can see, in this case the administrative problems are more challenging than the technical solutions. When we own the necessary grants on the images, we can adopt a similar approach for having them accessible from external systems. Moreover we could impose limitations on the maximum image resolution, on the specific work, or even on the specific image to the users and systems using the services.

Providing such services we can expect the following scenario:

1. An external system authenticates on the IPSA services.
2. It requests the list of the works it can access in XML format.
3. Using that list, it chooses or iterates on the work of interest, requiring the metadata of each work.
4. Using that metadata, it can request the bundle of thumbnails, or, in the future, the single water-marked image.

6. DISCUSSION

The process of extending a digital resource created for domain experts to other categories of users, including the general public, requires a number of steps. These steps are related either to additional user requirements that have to be harmonized with the existing system or to technological aspects that have to be addressed to improve the user experience with cultural heritage collections. Figure 5 represents this process as a spiral, going from a specific application developed for a focused group of expert users to a general purpose system able to raise the interest of a larger audience.
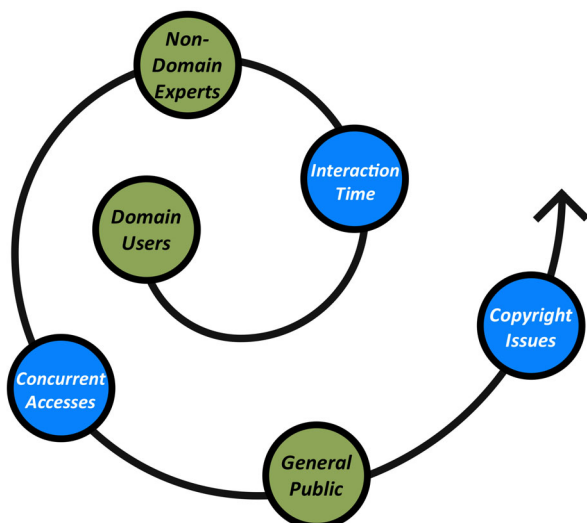

Figure 5: A representation of the different steps of the process to extend cultural heritage resources from domain experts to the general public.

In this paper we describe some aspects of this ongoing process, which we are developing with a user-centred approach. The choice of addressing the technological issues during the user evaluation, providing participants with a prompt feedback on their requirements, helped us to partially overcome the problems of recollecting input on the usability of the IPSA digital archive. Some questions still remain to be addressed, as the ones related to copyright management previously described. The next steps will regard the integration of IPSA digital archive within the CULTURA portal, supported by a continuous validation based on user studies.

References:

Agosti, M. (et al.), 2003. *IPSA: A Digital Archive of Herbals to Support Scientic Research*, T.M.T. Sembok (et al.) (Ed.): Proc. of the Int. Conf. on Asian Digital Libraries, LNCS 2911, Springer-Verlag, Berlin Heidelberg 2010, pp. 253–264.

Agosti, M. (et al.), 2011. *Methods of personalising a collection of images using linking annotations*, Proc. of the First Workshop on Personalised Multilingual Hypertext Retrieval (PMHR 2011), ACM, New York, NY, USA, pp. 10–17.

Agosti, M. (et al.), 2012. A Case Study for the Development of Methods to Improve User Engagement with Digital Cultural Heritage Collections. Grana, C. and Cucchiara, R. (Eds), *Multimedia for Cultural Heritage*, CCIS 247, Springer-Verlag, Berlin Heidelberg, pp. 166-175.

Fogel, K., 2005. *Producing Open Source Software: How to Run a Successful Free Software Project*. O'Reilly, Cambridge, Massachusetts.

Mariani Canova, G., 2002. La cultura universitaria padovana e la nascita del realismo nell'immagine botanica. *Atti e memorie dell'Accademia di Storia della Farmacia*, XX(3), pp.198-212.

Ponchia, C., 2012. *Engaging the User: Elaboration and Execution of Trials with a Database of Illuminated Images*. Taking Care of Legacy and Cultural Heritage Documents – Proc. of the 8th Italian Research Conference on Digital Libraries (in print).

References of websites

ARTstor website, URL: http://www.artstor.org/index.shtml
CULTURA Project website, URL: http://www.cultura-strep.eu/
CULTURA Project website: partners, URL: http://www.cultura-strep.eu/partners/
IPSA website, URL: http://www.IPSA-project.org/
Lucene website, URL:http://lucene.apache.org/
PostgreSQL website, URL: http://www.postgresql.org/
World Wide Web Consortium (W3C) website, URL: http://www.w3.org/