THE MULTIMATCH PROJECT: MULTILINGUAL/MULTIMEDIA ACCESS TO CULTURAL HERITAGE ON THE WEB

Jennifer Marlow, Paul Clough, Neil Ireson,
University of Sheffield, UK

Juan Manuel Cigarrán Recuero, Javier Artiles,
Universidad Nacional de Educación a Distancia, Spain

Franca Debole,
ISTI-CNR, Italy

http://www.multimatch.eu

Abstract

The EU-funded MultiMatch project aims to overcome language boundary, media and distribution problems currently affecting access to online cultural heritage material. Partners are developing a vertical search engine able to harvest heterogeneous information from distributed sources and present it in a synthesized manner. To design such a system, user requirements were initially gathered and then translated into specific design features to ensure that the search engine developed was consistent with user needs. This paper presents these user requirements, the initial design of the MultiMatch system, and technical discussion of the system architecture and components used to turn these design implications into a working interactive prototype. Following this, we discuss user evaluation and present results from an initial user study. These are being used, in addition to other inputs, to drive the functionality and design of the final system.

Keywords: User-centered design, cultural heritage online, search engine, task analysis, user-centred interface design

1. INTRODUCTION

For most organisations, providing access to information on their collection of available resources is of critical importance, either to provide competitive advantage (allowing the customer to more easily find what they want), or when it is a fundamental reason for the organisation, as is the case in the cultural heritage (CH) domain for libraries, museums, archives or other repositories. The digitization of information on the internet enables CH institutions to provide wider access to their resources by removing the barrier of physical location. To satisfy some information needs it may be necessary for an individual to access resources distributed over a number of different institutions around the world. Although a meta-search engine that collated material retrieved from different sources could be a useful first step, the problems of cross-language and cross-media search would remain.

The MultiMatch project aims to address some of these issues by developing an information retrieval system able to harvest heterogeneous information from distributed sources and present them in a coherent and synthesized manner. Although several studies have investigated how users access CH information, little is known about the impact of cross language/media search or the use of aggregated resources.

In general terms, the MultiMatch system incorporates data mapped directly from CH content providers' databases and
identifies further relevant material via a focused crawl of CH institutions' websites. The retrieved information is enriched using automatic semantic annotation and translation processes and the resultant metadata ingested into the MultiMatch data model. This process allows the user to retrieve information from the MultiMatch system, including relevant text, image and video resources, regardless of the source and target languages used to write the query and/or describe the resources. Finally, the system o

rganises search results in an integrated, user-friendly manner, allowing the users to access, interpret and interact with the information retrieved, and, if necessary, enables refinement of the query to better meet their requirements.

This paper addresses the design of the first interactive prototype of the MultiMatch system, in accordance with a typical user-centered design cycle (Rubin, 1994; Faulkner, 2000; Preece et al., 2002). In MultiMatch, the development cycle consists of: (1) needs assessment and task analysis; (2) preliminary design using low-fidelity prototypes; (3) design and development of interactive prototype; (4) heuristic evaluation and redesign; and (5) user evaluation. This paper begins with a literature review and discussion of user requirements gathering in the context of MultiMatch. Next, the system architecture is described, followed by a report of initial evaluation results. Finally, the functionalities envisioned for the final system are presented.

2. USER REQUIREMENTS GATHERING

2.1 Previous work

Past research on gathering requirements for professionals in CH has varied in terms of the methodologies used, information gathered, and audiences surveyed. Some studies were conducted in the context of evaluating specific websites or systems (Khoon et al., 2002; Economou, 2002; Smith et al., 2005); whilst others were more general and exploratory (Sexton et al., 2004; HEIRNET, 2002). Further work has also addressed the needs of professional or expert user, as opposed to the more casual or general user. For example, in the context of European research projects, there have been analyses of multilingual journalists in CLARITY (Petrelli et al., 2004) and broadcast professionals in VideoActive (O'Dwyer, 2007). Approaches used by these have varied, but commonalities included the drafting of user profiles/scenarios influenced by responses to questionnaires and interviews.

Other research has been carried out with academic users in the humanities domain, focusing on topics ranging from information sources used (Brown et al., 2006) to user's search behaviour (Frost et al., 2000). Several studies have examined the nature of search queries in the CH field, both for general information (Cunningham et al, 2004) and for images (Pask, 2005; Choi & Rasmussen, 2003; Collins, 1998; Chen, 2001). These all had similar findings: namely that people, places, time periods, and subjects were popular topics of search in the CH domain. Whilst findings from the aforementioned studies provide useful background knowledge, none of them compares and contrasts the needs of a diverse group of individuals across job boundaries. Additionally, most tend to focus on the way that users interact with just a single type of media (text, images or video).

Similar projects focusing on the aggregation, presentation, and semantic navigation of online CH information include eCHASE (Sinclair et al., 2005) and MultimediaN (Schreiber et al., 2006). These systems were identified as part of an initial competitor analysis that influenced the overall design process and also helped reveal how MultiMatch sets itself apart from similar projects. This differentiation occurs in MultiMatch's overall combination of functionalities: it enables both multilingual and multimedia search of domain-specific, semantically-enriched CH material provided by reputable archives and crawled from the web.

2.2 User Requirements Gathering

One hundred person-to-person interviews were conducted with domain experts (educational, tourism and CH professionals) in order to collect their opinions and needs. The interviews were conducted mainly in face-to-face mode using a questionnaire, and backed up by a set of scenarios and a vision document in order to give the respondents an idea of the proposed system functionality. In addition to the interviews, analysis of selected log files from online portals provided by project partners WIN D (a major Italian internet service provider) and Fratelli Alinari[1] i (a large historical photographic archive), and examination of the results of previous user studies in the CH domain occurred. For further information see (Minelli et al., 2007). Responses by potential users (both experts and general users) yielded a very large set of requirements. Summarising briefly, we can say that the main findings were the following:

- CH professionals do use the Internet widely and as part of their daily work routine, but they currently depend largely on generic search engines to find the information they need.

-

- They would like full capabilities for multimedia retrieval (i.e. images and video as well as text) but, in most cases, are only accustomed to executing text searches.

- Their main focus is about works of art and their creators. They would like access to all associated information, such as critical reviews, information on exhibitions, different versions of same document.

- They tend to be frustrated by the volumes of information available on the same subject and would find information filtering, clustering and aggregation functionalities very useful.

-

---

[1] Alinari (http://www.alinari.com) is a leading Italian provider of historic photographs, and a MultiMatch partner [site accessed: 25/01/08]

-
- They tend to restrict their searches to their own language plus English, thus missing information only available in other languages.
- If multilingual search was available, they would like to have the results associated with descriptive snippets in their own language (preferably) or English (optionally).

To further explore the needs of expert users in more detail and to learn more about their typical tasks, contextual interviews were conducted with fifteen CH professionals at three different institutions. Seven interviewees were academics (university professors in the fields of arts and heritage management, architecture, archaeology, French literature, and history) from the University of Sheffield (UK), four were image professionals from Fratelli Alinari (IT), who worked for the photographic archive, and four were video professionals from the Netherlands Institute for Sound and Vision (the Dutch national television archive), who regularly searched through various motion picture archives. This provided a variety of perspectives relating to the CH field and expanded upon past studies by addressing a broader range of tasks and topic areas. Table 1 illustrates some typical tasks (scenarios) for the different cultural user groups. These were developed, with the help of the expert users, to assist with the design and evaluation of MultiMatch.

User type	Task	Media and languages involved
CH professional	Searching for video footage on Pier Paolo Pasolini, needs to gather background information on who he was	Text, Images, Video English, Dutch
CH professional	Looking for images of (non-famous) people drinking coffee that capture a certain emotion	Images English, Italian
Academic	Preparing a presentation on Don Quixote and how it has influenced the arts	Text, Images, Video, Audio (?) English only
Cultural tourist/General user	Planning a visit to Turin, wants to know about museums to visit, what can see whilst there	Text, Images, Audio (podcasts) English, Italian

Table 1: Revised scenarios based on comments and observations from interviews

The interview questions focused on topics including a general overview of the interviewee's work, commonly used tools, use of the Internet, use of multimedia material, concrete examples of past search processes, ways of improving the current search process, and language issues relating to information access. Interviews were recorded and subsequently transcribed to facilitate detailed analysis and categorization of responses. Each comment made was classified into a category based on whether it

related to typical tasks, multimedia use, search topics, tools employed, the search process, general issues, or language-related issues.

Despite the differences in job roles and areas of expertise, many common patterns emerged across all groups of interviewees. Some of these confirmed the findings from the earlier questionnaires, as mentioned previously, while others emerged as specific needs of a particular group as more information was learned; for example, video professionals also had unique needs that did not necessarily apply to other user groups. An updated summary of similar and different needs, along with their design implications, can be seen in Table 2. Many of these needs formed the basis for the functional specification of the first working prototype system. However, due to time constraints, others were not initially implemented but will be included in the second and final prototype.

Relevant Groups	Need	Design Implication
Academic Image Video	1. Look for the same material on different sites (either because it couldn't be found or because it needed to be confirmed)	1. Aggregate multimedia material dispersed across the Web and fuse results in a single place Employ a common ontology to facilitate the exploration of content coming from diverse origins in a unified and organized way. The system will create a new, unique metadata scheme to create common links among the content whilst respecting the original native metadata formats.
Academic Image Video	2. Search and browse (typically search for a topic and browse through results)	2. Provide support for both types of behaviour. Organize and sub-divide large results sets into semantically related clusters to avoid duplication of results and to facilitate ease of finding the desired information when the query is ambiguous.
Academic Image Video	3. Have support for a variety of searches relating to people, subjects, time, and places (who, what, when, and where), particularly queries involving two or more of these aspects	3. Let users navigate semantic relationships between multiple categories (via faceted browsing or some other method). Categories/facets should be logical, intuitive, and correspond to the main classes of searches. Utilize specialized thesauri relating to names of people and places to support linguistic variations in these areas
Academic Image Video	4. Search for items conveying a feeling, mood, style, or other aspect that cannot be easily conveyed verbally	4. Facilitate content-based retrieval, query by example, and multimodal querying
Academic Image Video	5. Consult authoritative, quality information sources	5. Provide information about provenance and facilitate the filtering of results by domain (e.g., exclude Wikipedia results)

Academic Image Video	6. Be aware of copyright issues	6. Provide information about copyright; institute a log-in system so that users must register and accept a copyright policy in order to access non public domain material
Academic Image Video	7. Conduct searches in unknown languages	7. Give the option of performing and inspecting automatic query translation, thus broadening the coverage of a search and eliminating the need to rely on dictionaries to manually translate terms Use various thesauri relating to areas in which multiple variations of words exist across languages (e.g. the Getty ULAN for artist names and TGN for place names). Useful for queries whose spelling varies across languages (e.g. Raphael in English, Raffaello in Italian)
Academic Image Video	8. Have assistance in browsing documents or results written in foreign languages	8. Offer the option of automatically translating foreign-language summaries or documents (although the degree and nature of support needed might depend on the media type being searched for)
Video	9. Get a quick overview of a video's content	9. Provide both textual descriptions and keyframe storyboards to use as summaries. Allow users to retrieve relevant portions of a video clip based on automatic speech recognition transcripts and visual metaphor tools.
Video	10. Find places in a video where certain words are spoken (e.g., a famous quotation)	10. Enable keyword search of video transcript and the ability to jump into relevant parts of the video

Table 2: Summary of needs and associated design implications.

These design specifications were incorporated into a series of low fidelity mockups, which were converted into a working interactive prototype. Many of them are present in the first prototype, whilst others will not be implemented until the final version of the system.

3. ARCHITECTURE

As mentioned before, the first prototype system was intended to establish the system architecture and provide a basic initial set of services to evaluate. Needs 1, 5, 7, 9, and 10 are addressed in the first system by providing aggregated multimedia result sets for a query. The content of the collection has been specifically crawled from reputable and relevant sites on the Web, or donated by CH partners. Also included in the first prototype is a cross-language search functionality that offers automatic query translation (which can facilitate searching for both bilingual and monolingual users). Finally, the various media types can be interacted with in a more sophisticated way; these specialized searches are intended to address the needs of specialized professional users.

The system is being developed using a service-orientated architecture (SOA), with co-operating but independent components working together to provide a complete system. . MultiMatch uses a centralised index where metadata for cultural heritageCH objects (texts, images sounds and videos) is crawled from on-line sources (e.g. culturally-specific web pages, blogs and wikis), provided by the cultural institutions themselves, or harvested from OAI-compliant sources. The index is searched centrally, but the resources themselves can be located anywhere on the Internet. The current prototype system consists of four co-operating but independent sub-systems as shown in Figure 1.

Fig. 1: Components in the MultiMatch system.

Each of the sub-systems provides a specific functionality optimised for the CH domain. That is, they are designed to specifically support the gathering, information extraction, indexing, searching, and display of European CH objects, rather than more general material. However, the basic functionality is to a degree domain-independent, so that components can be ported to other domains in the future without having to perform extensive revision. The architecture of the current prototype consists of four sub-systems discussed in detail below.

3.1 Information Indexing and Extraction

This
 sub-system ingests data into the MultiMatch system, either mapped from CH databases or using focused crawling to locate web pages of interest to the MultiMatch from which relevant data is extracted (e.g. artist names and artwork titles). The data is transformed into metadata in a format defined within the MultiMatch project (Ireson and Oomen, 2007)), the linguistic data is translated into the target system languages before the metadata is indexed and passed into the metadata repository.

The content used in the first prototype system comprises data from a variety of sources ranging from structured data from the CH partners, through semi-structured data (e.g. Wikipedia articles), to unstructured textual data from crawled web pages, and additional audio files (Table 3). This offers varying levels of difficulty with respect to processing the data, and additional difficulties in providing a unified view of the content indexed by the MultiMatch system.

Origin	Alinari	Beeld and Geluid	Biblioteca Virtual Miguel de Cervantes	Wikipedia	White list crawl	University of Amsterdam Audio Corpus	The European Library
Quantity	5,000 stills (jpg)	50 videos (mpeg-1)	9,000 texts	65,000 articles	40,000 pages	20 hours audio	1.6 million records
Metadata dimensions	Proprietary Dublin Core	Proprietary Dublin Core	Proprietary Dublin Core	Proprietary format	Proprietary format	Proprietary format	OAI-DC

Table 3: Content indexed in the first MultiMatch prototype.

MultiMatch aims to produce a complete metadata record for the digital object, allowing users to interpret the wealth of CH information by presenting objects, not as isolated individual items but as situated, richly connected entities. Figure 2 shows an overall view of the data model used to represent the CH domain in MultiMatch and Figure 3 shows the interconnection between the main entities in the model. Such a representation allows the user to explore the information space related to a given object.

Fig. 2: MultiMatch data model

Fig. 3: MultiMatch Entity Relationships

Due to the fact that the MultiMatch ingests information from multifarious sources and applies automatic enrichment, the representation maintains the source of the information (both the original data source and the information extraction process). This information is necessary both for system development purposes (to evaluate and debug the information gathering process) and for certain CH users, who require an "audit" trail so that the provenance of the information is explicit, enabling them to determine their level of trust in the information provided (Need 5).

3.2 Searchable Metadata Repository

The searchable metadata repository provides a service for storing and retrieving MultiMatch cultural objects (images, video, text, audio, etc) and the associated metadata documents generated by the information extraction sub-system. The repository is optimised for querying the specific types of data that are contained within the metadata records, whether this is text, numeric, temporal, geographic, image, etc. It also caches each MultiMatch object with its metadata record for fast retrieval by the Enhanced Information Retrieval module. Storage, versioning and serving the digital objects is handled by the metadata repository content cache, which is based upon the Milos Multimedia Media server (Amato et al., 2004).

3.3 Enhanced Information Retrieval

The enhanced information retrieval sub-system provides a set of four services offering a sophisticated multilingual CH-oriented search facility. This sub-system provides an external interface to MultiMatch functionality, such that organizations or users can develop their own custom-built applications. The four main services are:

1. **MultiMatch Search Service**: this provides a stateless search interface, allowing the client to specify what types of objects to search, how many to return and which items in the result set should be returned. The results of the search can be either a brief snippet such as might be provided on a results summary screen; or the full records including all metadata and digital content.
2.

3. **MultiMatch Query Translation Service**: this translates the query terms between English, Spanish, Dutch and Italian. MultiMatch is developing components for both document and query translation and procedures for matching one against the other. Much effort is being dedicated to the building of domain-specific multilingual resources catering for the terminology adopted in the CH domain (Jones et al., 2007). Both dictionary-lookup and Machine Translation services have been developed.
4.

5. **The MultiMatch Query Expansion Service**: this provides three mechanisms for query expansion that can be used to refine the users search experience. It offers the blind relevance feedback which expands the query without user intervention, relevance feedback using user selected items as seeds for new terms, and a thesaurus service that provides cognate terms.
6.

7. **MultiMatch Browse Service**: this service allows navigation through indexed content using any of the metadata schema indexes known to MultiMatch, e.g. browsing for names, titles, dates, locations, etc. While browsing functionalities will not be present in the first prototype, this service is a part of the final system's architecture.

3.4 User Interface

The User Interface sub-system provides a set of client services allowing development of user interfaces that offers both simple (free text) and advanced (metadata based) search, browsing of the content, as well as stored personalisation, search history, and preferences features. MultiMatch provides a default browser-based interface for design and test purposes, enabling the user to interact with content indexed by the MultiMatch system, and using functionalities developed for search and browse. The interface is responsible for collecting user queries, presenting results and allowing users to enhance and refine their information need. Users are able to retrieve cultural objects through different two different search modes:

- **Free Text Search.** This search mode is similar to any general purpose search engine (e.g. Google), with the difference that (1) MultiMatch will provide more precise results since the information indexed is selected from sources containing CH data, and (2) MultiMatch will provide support for multilingual search. This means that the user can formulate queries in a given language and retrieve results in one or all languages covered by the prototype (according to his/her preferences). Figure 4 shows an example of how multilingual search is integrated into a free text search action through a "translation wizard". In this case the user has queried about "still life" but wants to retrieve information in Spanish. The query is translated into Spanish and launches the retrieval process. Once the retrieval has been performed, the system shows the user the translation, giving him the opportunity to change it if necessary by adding, removing or writing new translation terms. In this case, the system has also detected "still life" as a phrase so it suggests the Spanish term "bodegón" as the translation.

Fig. 4: Illustration of cross-language search functionality

- **Metadata Search.** From the results of the expert users' survey, we can conclude that, on average, CH professionals tend to classify searches for information about creators (authors, artists, sculptors, composers, etc.) and creations (works of art and masterpieces) as their most common search tasks. The use of metadata search is designed to meet user needs 2 and 3 by providing an alternative way of searching for a topic (based on metadata fields such as creator or creation) or browsing through results (e.g. to find other artworks made by the same creator of a given result).
-

A key objective is to provide a system that can be easily adapted to different user needs. For this reason, MultiMatch searches can be performed at two main levels of interaction in the current prototype:

- **Default search mode.** The simplest search mode is the default MultiMatch search level. This is provided for generic users with a limited knowledge of MultiMatch system capabilities, or with very general search needs. In this case, no assumption is made on the user query, and MultiMatch retrieves information from all indexed material. This interaction level, thus, involves the retrieval of web pages, images and videos related to the query. Figure 5 shows this behaviour for the current prototype. For the query "pasolini" the system retrieves all available resources matching the query and shows them to the user in an overview mode. In this case there

are images and web pages available in English, Spanish, and Italian, with the addition of videos in Dutch.

Fig. 5: Illustration of default search

Specialised search mode. Users with a more precise knowledge of MultiMatch system functionality, and with specific search needs, may use one of the specialized interaction levels available. These allow the user to query MultiMatch specific search services (for instance, video search, image search, etc.) and retrieve all the relevant information available via the selected search service. In this way, MultiMatch will include standalone image, video and metadata-based searches, each with its own search fields, display and refinement options. The image search gives as a result a set of images, ranked according to some relevant criteria, which will give access to image thumbnails, original images, and to the image relevance feedback functionality, centered around visual attributes (such as texture and colors) and realized using the capabilities of GIFT, a research effort at the Vision Group at the CUI (computer science center) of the University of Geneva (http://www.gnu.org/software/gift/).

The video search service provides the user with a video list, ranked and retrieved using text retrieval techniques based on spoken audio transcriptions, and it offers also the possibility of retrieving only a specific segment of a given video clip relevant to a specific query. In order to return only the audiovisual information of interest, a streaming video server has been installed on the MultiMatch integration server.

Figure 6 shows the specialized video interface where the user has queried for Dutch videos about Picasso". The system shows the set of retrieved videos and also gives access to a set of specific video tools where the user can manipulate a video for playback. These specialized functionalities were designed to meet the needs of video professionals (see needs 9 and 10 of Table 2). They allow a user to view keyframes for a given video, start playback at a desired point by clicking on the relevant keyframe, and also to search the ASR transcript for words of interest.

Fig. 6: Screenshot of specialized video playback

4. EVALUATION

In accordance with the user-centred design methodology being adopted, on completion of the first interactive prototype (designed in accordance with user needs and requirements), it was necessary to evaluate it with both expert and general users. Overall, the evaluation consisted of three parts, each of which yielded feedback on various aspects of the system. These three exercises included the distribution of questionnaires associated with a system demo, an internal evaluation, and a larger scale user testing.

The system demo was carried out at a large education conference that was attended by many people whose work involved CH in some way. After viewing a guided demo of the first prototype system and its functionalities, they filled out a questionnaire asking them to rate the usefulness of the various features. In total, 20 people responded and the top-rated functionalities were as follows:

Functionality	Mean Rating (1=not useful; 5=very useful)

Advanced search based on metadata fields	4.50
Search for video fragments in specialized video search	4.25
Term suggestion to perform text relevance feedback	4.20
Image similarity search to perform visual relevance feedback	4.18
Different interaction levels (overview vs. specialized)	4.13

Table 4: Top 5 rated functionalities of first prototype

The second stage occurred prior to conducting the main user testing. An internal pilot testing of the evaluation was carried out with project partners. Realistic scenarios of use drove the evaluation; users were given a series of search tasks to complete using MultiMatch, which were based on typical behaviours mentioned in the requirements gathering phase (see Table 2). The tasks guided the users through the system and highlighted the various specialized media functionalities. After completing these tasks, users filled in questionnaires focusing on system satisfaction and usability.

This highlighted important system bugs and also provided some input into further development of the system. Many related to minor design or technical issues, but the following suggested changes were mentioned:

- Users wished to be able to use the cross-language functionality to search across all languages, not just language pairs
- Users requested further support in reading cross-language search results in languages they did not understand

The large-scale user evaluation is currently in progress at the time of writing. Overall, over 40 expert and general users are expected to complete the tasks and offer their feedback on the system. These comments will then be used to help guide the improved design of the second prototype. Initial results indicate that, once again, users are pleased with the cross-language searching functionality. However, they did have some trouble dealing with material in languages they could not speak. The evaluation concluded with a question asking users the usefulness of a variety of features that could be included in the second prototype. The features rated as most useful included (in order): document translation, exploring relationships between artists, and viewing artist information on a timeline. This supports the inclusion of the features mentioned in the following section.

5. FINAL VISION

New features in the final prototype system will take two forms: improvements made in response to user evaluation comments, and additions relating to functionalities based on user needs that were not able to be included in the first prototype due to time and technical constraints.

With regards to this first category, the final system will address language issues by including an enhanced "translation wizard" for cross-language searching purposes and providing document translation (as requested above and in Need 8 in Table 2).

The new system features will build upon the first prototype and address further needs by providing added functionalities to enhance interaction with the content. These enhancements will primarily be related to the areas of collection browsing (Need 2) through a faceted browsing feature (Need 3). This will provide a more directed search and exploration of the MultiMatch collection. By facilitating faceted browsing of author and artwork-related material (a user can search using more than one category as criteria; for example, find all American male artists from the 20th century). Both expert and general users signalled this as a useful way of enhancing access to information beyond currently available methods. In addition to enabling the faceted browsing of artists and their artworks (Figure 7), alternative display modes are available to more naturally present geographic and temporal information, so that the users can also explore artist birthplaces on a map (Figure 8) and view timelines putting different artists' lives into context (Figure 9). Again, these myriad ways of exploring the content will support the varied requirements of an individual as mentioned in Need 3.

Fig. 7: Sample faceted browsing for American male artists in the 20th century

Fig. 8: Sample map depicting artists' birthplaces (using Google Maps)

Fig. 9: Sample artists timeline view

6. CONCLUSION

This paper has described the design process for the first prototype system of a targeted search engine for the cultural heritage domain. The process began with an analysis of user requirements which then fed into a list of functional specifications which are being implemented as a series of prototypes. The working prototype was evaluated with typical users, which then yielded recommendations and improvements for the next cycle of redesign. It is expected that the system will provide added benefit to individuals searching for cultural heritage material online through the provision of aggregation, specialized multimedia interaction facilities, and the potential of exploring semantic relationships between creators and their creations. Beyond the context of MultiMatch, the work presented here reveals how an understanding of cultural heritage professional user needs can also be of use to others developing information systems for this type of audience.

ACKNOWLEDGEMENTS

REFERENCES

Amato, G., Gennaro, C., Savino, P., and Rabitti, F. (2004). Milos: A Multimedia Content Management System for Digital Library Applications. In Research and Advanced Technology for Digital Libraries, ECDL 2004, Bath, U.K., Proceedings, Springer 14-26.

Brown, S., Ross, R., Gerrard, D., Greengrass, M., & Bryson, J. (2006). RePAH: Research portals in the Arts and Humanities: A user analysis project. http://repah.dmu.ac.uk/report.

Chen, H. (2001). An analysis of image queries in the field of art history. JASIST, 52(3), 260-273.

Choi, Y., & Rasmussen, E.M. (2003). Searching for Images: The Analysis of Users' Queries for Image Retrieval in American History. JASIST, 54(6), 498-511.

Collins, K. (1998). Providing subject access to images: A study of user queries. The American Archivist, 61, 36-55.

Cunningham, S.J., Bainbridge, D., & Masoodian, M. (2004). How people describe their image information needs: A grounded theory analysis of visual arts queries. JCDL, 47-48.

Economou, M. (2002). National Council on Archives and National Archives Network User Research Group (NANURG): User Evaluation: Report of Findings. http://www.mla.gov.uk/resources/assets/N/nanurg-pdf-4796.pdf.

Faulkner, X. (2000). Usability Engineering. Houndmills: Macmillan.

Frost, C.O., Taylor, B., Noakes, A., Markel, S., Torres, D., & Drabenstott, K.M. (2000). Browse and Search Patterns in a Digital Image Database. Information Retrieval, 1, 287-313.

HEIRNET. (2002). Historic Environment Information Resources Network: Users and their uses of HEIRs. http://britarch.ac.uk/HEIRNET/users.pdf.

Ireson, N. & Oomen, J. (2007). Capturing e-Culture: Metadata in MultiMatch. In Proc. DELOS-MultiMatch workshop, February 2007, Tirrenia, Italy.

Jones, G.J.F., Zhang, Y., Fantino, F., Newman, E., & Debole, F. (2007). Multilingual Search for Cultural Heritage Archives by Combining Multiple Translation Resources. In Proc. of the ACL Workshop on Language Technology for Cultural Heritage Data (LaTeCH 2007), Prague, Czech Republic, June 2007.

Khoon, L., Ramaiah, C., & Foo, S. (2002). The design and development of an online exhibition for heritage information access. In J. Trant and D. Bearman (eds.) Museums & the Web 2002: Proceedings.

Minelli, S., Marlow, J., Clough, P., Cigarran, J., Gonzalo, J., & Oomen, J. (2007). Gathering requirements for multilingual search of audiovisual material in cultural heritage. To appear in Proc. of Workshop on User Centricity – state of the art (16th IST Mobile and Wireless Communications Summit), Budapest, Hungary, 1-5 July 2007.

O'Dwyer, A. (2007). User Profiles and Requirements Plan. Deliverable 6.1 for VideoActive Project (internal document).

Pask, A. (2005). Art Historians' Use of Digital Images: A Usability test of ARTstor. Dissertation at University of North Carolina, Chapel Hill.

Petrelli, D., P. Hansen , M. Beaulieu, M. Sanderson, G. Demetriou, P. Herring. (2004). Observing Users - Designing Clarity: A Case study on the user-centred design of a cross-language retrieval system. JASIST, 55(10), 923-934.

Preece, J., Rogers, Y., and Sharp, H. (2002). Interaction design: Beyond human-computer interaction. New York: Wiley.

Rubin, J. (1994). Handbook of usability testing: how to plan, design, and conduct effective tests. New York: Wiley.

Sexton, A., Turner, C., Yeo, G., & Hockey, S. . (2004). Understanding users: A prerequisite for developing new technologies. Journal of the Society of Archivists, 25(1), 33-49.

Schreiber, G., Amin, A., van Assem, M., de Boer, V., Hardman, L., et al. (2006). MultimediaN: E-Culture Demonstrator. LNCS 4273, 951-958.

Sinclair, P., Lewis, P., Martinez, K., Addis, M., Pillinger, A., & Prideaux, D. (2005). eCHASE: Exploiting cultural heritage using the semantic web. Proceedings of the 4th International Semantic Web Conference, ISWC 2005, Galway.

Smith, R., Howes, D., Shapiro, W., & Witchey, H. (2005). Shaping Pachyderm 2.0 with user requirements. In J. Trant and D. Bearman (eds.) Museums & the Web 2005: Proceedings. Toronto: Archives & Museum Informatics, published March 31, 2005 at http://www.archimuse.com/mw2005/papers/smithR/smithR.htm