

**Pierluigi Feliciati**

University of Macerata, Italy

**Alessandro Alfier**

IBC - Regione Emilia-Romagna, Italy

## **ARCHIVES ON THE WEB AND USERS EXPECTATIONS: TOWARDS A CONVERGENCE WITH DIGITAL LIBRARIES**

**Abstract:** Digital environments aimed at offering services on historical archives should put together domain standards with typical features of digital libraries. The current tendency is often to adopt two separated methodological approaches: standards-compliant archival descriptions (multi-level, multiple-relations based, providing access just by provenance expressed by original titles, material-centric, careless of the interoperability with general search engines), or digital technologies (horizontal, providing access by subject, sensitive to the “aboutness”, user-centric, focused on item-level descriptions based on Dublin Core-like models). Some issues are coming more and more to surface: how could a web archival environment face users expectations and what content (and how) may be offered to cultural aggregators? How can move archives on line from the plain delivery of cultural information to the facilitation of a cultural experience? This paper examines most of those issues, the necessary difference between input and output to ensure high quality contents, the upside-down changing role of the professional mediation, the peculiarities of web displays for archival descriptions. Moreover, studies of end-users are crucial to guarantee the profitable investment, although the first purpose of the investment should be end-users’ satisfaction. The authors look into those topics by reviewing the international state of the art, and in light of their coordination of a huge user study project as well. It was launched in order to test the prototype of *Una Città per gli Archivi*, a portal disseminating the results of a several-year project of preservation and valorization of documentary funds of the city of Bologna (Italy). Besides paper archives, the project included photographic and graphic materials, sound and audiovisual recordings.

**Keywords:** archives on the Web, digital library, user studies, cultural Web quality, archival mediation

### **Introduction**

Within the context of digital resources regarding historical archives, from here on called “archives on line” to distinguish them from digital-born records environments, A. Sundqvist recently noted that the “general knowledge of a user behavior is a mixture of common sense, presumptions and prejudice” [22]. What methods of research could help to understand users’ points of view in the domain of archives on line, and in general, of digital libraries? Most frequently, it is addressed the evaluation of existing environments, i.e. studying how users react to an existing service with a strong emphasis on usability. Rarely, studies are done before a digital library is developed, or while it is in development, even if they have been recently completed in the framework of the Europeana activities [6].

Anyway, an increased interest and attention to users’ expectations is evident. Although the domain of Digital Libraries has not been crystallized sufficiently to offer consistent models and recommendations, it proves to be satisfactory. As user-centric approach is so common in the declared goals of projects, it rarely corresponds to the support of effective user studies; this is even more evident as regards archives on line, when information access quality seems to be in discussion even

more. A participant of a focus group organized and moderated by Wendy Duff and Penka Stoyanova in 1997, complained about his/her frustration with archival web services:

“If [the goal of archival description is] making material accessible to people, then don't build these barriers, between the ordinary people and the material. The other image I've used is ... I have needle work, and it looks very patterned on one side; you turn it over, it's a crazy quilt. So it makes sense to the archivists in terms of what they have to do in order to turn the material [over] to the people? But don't turn [its] crazy quilt side out because, on your side it looks neatly printed - but to us all that's spread on the back doesn't make any sense at all. So you have to turn it around and write it in language that we can understand”[10].

These words upset every archivist by putting strong evidence on distance between archival information producers and suppliers and end users; it is surprising if we compare it to the traditional role of archivists as mediators between documentation and users, source of their professional and ethic legitimating.

In this paper we take as a starting point that from high-quality delivery of information about cultural object, the cultural industries could not help but moving towards facilitation of cultural experiences. To support this shift, we will briefly present the present issues for the quality offering of archival information on web platforms based not only on *input* standards. Moreover, user studies are evidently crucial to better finalize archives on line and ensure good archival displays – based not only on domain (input) standards) but also on the contribution of a web design, human-computer interaction studies and cognitive psychology. We conclude with the presentation of an Italian archival project, *Città per gli Archivi*, based in the city of Bologna: its mission, the main solutions choose to face the finding aids effectiveness, and the huge user-centric study organized for evaluating and assessing the archival portal prototype.

### 1. Archives on line and the lack of convergence towards digital libraries

A good starting point for our arguments frame could be the principle that “output is not input”, to quote Wendy Scheir:

“Especially now, when within a decade it is estimated that 60-100 percent of all processed collections in the United States will have on line finding aids, it is essential to establish clear distinction between input and output, even while acknowledging that description and presentation are inextricably intertwined [...] If we do not listen well to all of our users, we will be less able to tune finding aids to a delicate pitch: to meet user needs while sustaining the highest standard demanded by the profession for describing and representing archival collections” [21].

There is a reminder of the precise demarcation between the standards controlling the increasingly sophisticated archival descriptive technique and the guidelines that should decide the best display for end users. It is just the lack of awareness for this distinction that seems to be the basic factor in determining the unsatisfactory performances of on line archival finding aids.

The common assumption seems to be that compliance with standards is not only necessary but also sufficient for good transmission of primary sources to web users, as though such standards had an extensive force of supporting both input and output. At the same time, it is still deeply rooted in the everyday practice of archivists that the transition from analog to digital environment does not require any re-thinking of archival finding aids.

We should consider firstly that archival standards are literally, both for the methodological standard ISAD(G) [14] and for the data structure standard EAD<sup>1</sup>, "output neutral". Particularly interesting is the case of the latter standard *de facto*, whose output neutrality has been largely ignored by the average perception of archivist community. It has been too easily believed that any data structure based on this XML schema could be of use also as an output format, confusing input records with their output display. On the contrary, sector literature has emphasized how EAD compliant displays are not capable by themselves of ensuring the use quality of archival information from web users point of view, confirming that standards compliance is a necessary condition, but not sufficient for usability archival data. In this sense, it is interesting to quote a comment of Elizabeth Yakel, pointing up how EAD compliant finding aids represent, till to-day, barriers for web users:

“Encoded archival description (EAD) has become part of the archival vocabulary. On most occasions, archivists refer to EAD as a data structure for sharing information about collections. Yet, EAD is a boundary object that must not only mediate between archivists and their user communities, but must also facilitate a convergence between the user and the archival content [...] The continuing existence and evolution of EAD is dependent on not only broad implementation in the archival profession, but also widespread acceptance and employment by groups of users. In other words, EAD finding aids must become boundary spanners, and not barriers, in the research process” [25].

In order to produce EAD compliant finding aids not working as web users hurdle, but as effective “boundary spanners”- i.e. able to mediate between input and output and so providing appropriate keys to understand web archival descriptions – it is necessary that archival information displays are designed by integrating the principles of EAD data structure with different disciplinary contributions (cognitive science, human-computer interaction, web design). Joyce Chapman takes up the same position when she observes:

“as archivists work to refine finding aid web displays, it is important to remember that the goals of descriptive standards should not and do not include display. User needs and the principles of web design, rather than the structure of the EAD XML standard, should dictate how finding aids are displayed” [3].

Up to now in the traditional analogue environment, the “decoding” of technical archival descriptions was ensured by archivists through the reference service. In fact, when archival description is closed in its disciplinary edges, the mediation rule of archivists – who are indeed the authors of the descriptions to be decoded – consists in their face-to-face interaction with end users. For this reason:

“we [archivists] express a great deal of fairly standardized information about the content of our collections in what is essentially coded form. We aren’t providing our current users or a much larger number of potential users, with anything resembling a code book” [18].

The lack of decoding support is not sustainable when archival finding aids land on the web, which inhibits the archival mediation at least in its traditional form. Therefore, it must be the archival information itself to support decoding in digital environments. And such sweeping change of logic interaction among archivist, end user and finding aids inevitably involves the need for a deep re-thinking of access resources to the archives.

By now, North-American archivists explicitly argue about the necessary “re-engineering”, “re-conceptualizing” of finding aids, achievable with an inductive method: building high use quality archival displays basing on the contributions of cognitive psychology, human-computer-interaction, web design and in the end checking them through usability tests and user studies.

---

1 Encoded Archival Description, see <http://www.loc.gov/rr/ead/>.

As regards this challenge, Elizabeth Yakel [24] notes that web has certainly increased the access to archives. But at the same time, the new digital context – undermining the traditional mediation performed by the archival reference service – has compromised the accessibility, the use quality of archival descriptions. Thus, end users continue to use, as much as possible, any available form of human mediation to access to archives, because it offers a valid alternative to the scanty usability of on line archival resources. This is threatened by a variety of factors including inter-indexer inconsistency, difficulties in retrieving archival descriptions from general search engines, reduced understanding of data as expressed only in accordance with the internal logic of archival descriptive technique. Yakel shows a basic challenge to ensure usable archival data also in present digital transition, including an effective communication among users and between archivists and users: embedding in archival finding aids tools that are able to re-propose human mediation, such as virtual reference, recommender / reputation systems, visitor awareness services and user generated content services (social tagging, social bookmarking, folksonomies, wikis, commenting).

This challenge involves an exacting change in the nature of archives on line, in comparison with archival finding aids in analogue environments: no longer static tools, hardly upgradeable, designed with an asymmetric top-down logic by “memory technicians” for few end users (“sacerdotal approach”) [11], but on the contrary dynamic environments, conceived with an information core represented by the data structured and managed by archivists and informative integrations performed through the network of relationships among any possible end user and between them and archivists (“secular approach”). In this sense, we could have to accept the “shared authority” [8] in archival description activities, but also the finding aids transformation into “information social phenomena”, according to a new scientific-informational paradigm, as advocated for archival science by Fernanda Ribeiro [19].

Archives on line bring into question the traditional archival mediation also because inside World Wide Web another powerful protagonist lives, the “user agent”. More and more frequently end users have access to the on line archival information not directly, but through an automatic mediation supplied by robots, spiders, crawlers, harvesters. According to the most recent surveys, in the U.S.A. approximately 90% of access to on line archival resources occurs starting from the general search engines [17, 20]. And their conditioning power is so overwhelming that it has produced the so-called phenomenon of “hidden collections” [20, 12]: archival finding aids whose descriptions (published online) are actually “invisible” to end users, because their contents or descriptions are not ready to be indexed by search engines. For such cases, the access to web archival data is assured only to end users having a strong previous background. The publishing of archival resources into the deep web layers (data bases invisible to web agents) ask their re-conceptualization, out of the web pages we conceive as unique environments of access.

Web users, following this argument, may be roughly classified into two categories: the “browsers”, who adopt search strategies typical of the so-called “berry pickers”<sup>2</sup> and the “searchers” [11], fishermen of single units of information. The latter type is particularly the case of the youngest segment of the population – the so-called “Google generation”<sup>3</sup> – strongly influenced by the general search engines when retrieving information [6, 7]. These use profiles should not be ignored while

---

2 The concept of *berrypicking* relating to the browsing technique is due to Marcia J. Bates [1].

3 See the project commissioned by the *British library* and the *Joint information systems committee* at the *University college London*: the results have been published in January 2008 with the title *Information behaviour of the researcher of the future: a ciber briefing paper*, [http://www.jisc.ac.uk/media/documents/programmes/reppres/gg\\_final\\_keynote\\_11012008.pdf](http://www.jisc.ac.uk/media/documents/programmes/reppres/gg_final_keynote_11012008.pdf).

planning and developing our archives on line. We don't have not to favour only the browsing behaviour, represented by the hierarchical structure of the archival description associated with traditional archival search forms. It is necessary to invest our energies so that searchers can be provided with efficient methods of information recovery. The production and structuring of metadata could facilitate the resource indexing and harvesting by web search engines and culture aggregators; an internal information retrieval system of the resource could allow recovering adequately analytic and subject based data; the propensity to a "meta-dation" may determine a better decoding of item levels. To conclude this topic, the integration of the two methods of accessing to information will support the typical cases when different use approaches overlap and cross one another, even in the same user.

In the process of re-engineering and re-conceptualizing of archival finding aids, a particularly interesting term of comparison is represented by the digital libraries, when they represent resources aimed at offering an information service to final users; such services are designed according to the known needs of a specific community. One of the main definitions of DL reads:

A potentially virtual organization that comprehensively collects, manages and preserves for the long depth of time a rich digital content, and offers to its targeted user communities specialized functionality on that content, of defined quality and according to comprehensive codified policies [5].

It is never easy to move cross-sector from memory institutions, traditionally closed in their jargons and specialties: However we believe that the lack of genuine convergence between archives on line and digital libraries has to be crossed, and to build a bridge based on the common sensibility for use quality. The following table summarizes the main differences:

**Table 1:** Differences between digital libraries and on line archives

Digital libraries	Archives on line
▪ User centric	▪ Material centric
▪ Providing access by subject	▪ Providing access just by provenance
▪ Focused on aboutness	▪ Focused on hierarchy
▪ Defined by a simple structure (horizontal / net)	▪ Defined by a multilevel structure
▪ Focused on item level and digital reproductions	▪ Focused on general, related descriptions
▪ Sensitive to human users and web agents	▪ Sensitive to human users and rarely interoperable
▪ Use of simple standards (i.e. Dublin Core)	▪ Use of complex standards (i.e. EAD and EAC)

## 2. Matching with final users

If we approach the concept of quality for a digital service, it could be defined by using the standard definition: "the capability of the software product to enable specified users to achieve specified goals with effectiveness, productivity, safety and satisfaction in specified contexts of use" [15]. If we agree that a DL is composed of three core elements, the content it makes available, the services it provides, and the users, so the activities of understanding the needs, expectations, and perception of users, whenever these needs are explicit or not, are core to the delivery of a good service [13].

This approach seems even more crucial for archives on line, considering what is exposed above. In the North-American context the quality of use of archival information is no more presupposed neither

defined in abstract terms, but concretely tested through some techniques measuring the effective interaction experiences. It was possible to consolidate a heritage of user study results and data, despite of the variety of measuring techniques of the audience and the different resources analyzed. These different studies [4, 2, 3, 25] converge on some basic issues in the effective interaction of final users against archives on line. Their results concern mostly four different areas:

- The archival terminology. All studies concur that a too technical language used within the on line archival finding aids represents a barrier for users. In particular, the most specific archival terms are not immediately understandable<sup>4</sup>. The barrier generated by the archival jargon represents an obstacle not only for the comprehension of the descriptions, but also for the use of the extended search functions, whose utilization is forbidden to users since they go with complicated and ambiguous labels deduced just from the archival language. It is therefore inevitable and absolutely necessary to “decode” the archival jargon, translating it into a more intuitive and common language.
- The hierarchical structure. It was widely noticed a sensible difficulty for users to move in the multilevel hierarchy typical of archival descriptions, even if in some cases the most inexperienced subjects have shown an ability to learn the structural nature of archival resources<sup>5</sup>. The ambiguity of the hierarchy leads users to choose preferably the search functions to retrieve information, rather than browsing through the descriptive levels. This drive-back effect seems to be due to the different perspective of end users and archivists towards archival finding aids: if users are interested in what archives is related to (in the “aboutness”, in content datum), archivists concentrate on what composes archives, on internal relations among descriptive items, on the structural datum, following essentially the principle that records in an archives are not records *about* an activity, but records *of* an activity [9].
- The searching tools. The main studies show the critical use of the search functions typically provided by archives on line. The examined users accustomed to OPAC tend to think that the query methods are identical for archival finding aids and that the latter are modeled like bibliographical descriptions. Another part of users seem to have difficulty in choosing terms or search parameters, often using the default values proposed by the system: this inability precludes any refining of results. The studied search behaviors include also controlled dictionaries, and results are not fully concordant. They show a tendency of users to carry out subject-based queries, obviously combined with the above explained preference for aboutness. Nevertheless, some studies related to the limited use of the controlled dictionaries within some information systems give opposite results, due perhaps to users misunderstanding of the access keys purposes and their connections with descriptions. Other critical aspects could be found in the criteria of presentation of the results provided by information retrieval engines (ranking): archivists have traditionally a non-evaluative attitude about archives documentation, tending to build up their information systems according to a “neutral order” of any list (for example a chronological order). On the other hand final users – accustomed to the general search engines – expect the search results presented according to a semantic relevance rank (such as in Google).
- Contents visualization. Carried out studies do not provide any unequivocal position about this aspect. There are some contradictory signs relating to users common preference for minimal

---

4 It is sufficient quoting the term “fonds”, the core concept of archival profession, which sounds often ambiguous or even incomprehensible to users [10].

5 This evidence could indicate that the hierarchical structure is not difficult in itself to be understood, but that its easiness depends mostly on the display.

descriptions rather than from detailed and analytical ones. Some appreciations have been noted after the use of displays presenting short narrative texts linking to more detailed information.

All the studies quoted above referred on archives on line, to check users' satisfaction against services. Nevertheless, user analysis could be very useful as well in the early stages of a project (prototype phase). Evaluation, when conducted while the DL is still under development, is called *formative*: part of the data gathered during this stage will focus on the content and the functionalities of the DL and how these correspond to the users' needs [23].

Those needs could be checked by asking users directly who can provide feedback that helps in the recognition of flaws in the interaction design. Moreover, expert opinions can provide assistance when the DL prototype has been refined, giving their insights on the quality features of the system.

To directly involve users in the prototype step, it could be useful to launch a survey (e.g. using on-line questionnaires), but it has to be put in evidence how it will be suitable only if a designed community to address is already available. On the other hand, a more complex and useful method is to organize as much moderated focus groups as possible, and to try to address particular users who work with web service and are interested in its future development. Finally, the expert study could involve both experts on content and functionality, to gather advanced qualitative data on performance, usability and expected success of our project.

After such prototype testing activities, it is usually possible to define an implementation policy based on obtained recommendations, and the researcher may return to previous stages of development and adopt the studies' findings, thus altering the status of the prototype.

### 3. The *Città per gli Archivi* project

The project "Una Città per gli Archivi" started in 2007; it was supported by two banking foundations (Fondazione del Monte di Bologna e Ravenna and Fondazione Cassa di Risparmio in Bologna). Its purpose was to valorize the most interesting archives of the modern and contemporary history of Bologna, preserving the collective memory and providing the city community with the necessary tools to experience that memory.

The project had recourse to the scientific advice of several experts on archival science and History research, and to the active participation of the state, regional and local institutions operating in the field of the conservation and the cultural promotion of archives: the Soprintendenza archivistica per l'Emilia-Romagna<sup>6</sup>, the State Archives of Bologna<sup>7</sup>, the Istituto per i beni ambientali culturali naturali dell'Emilia-Romagna (IBC)<sup>8</sup>, the Gramsci foundation for Emilia-Romagna<sup>9</sup>, the Provincia di Bologna<sup>10</sup>, the Biblioteca comunale dell'Archiginnasio<sup>11</sup>.

They all have contributed to draw the development strategy of the project by:

- selecting archives with a conservative risk profile and archives that were safely maintained by conservation institutes, but had not any finding aid in order to enable users to consult them;
- differentiating archives as much as possible with a clear specification of their nature, their typology and their consistency. As a consequence, the project has concerned not only paper

---

6 See <http://www.sa-ero.archivi.beniculturali.it/>.

7 See <http://www.archiviodistatobologna.it/>.

8 See <http://www.ibc.regione.emilia-romagna.it/>.

9 See <http://www.iger.org/>.

10 See <http://www.provincia.bologna.it/archivistorico/Engine/RAServePG.php>.

11 See <http://www.archiginnasio.it/>.

archives, but also photographic, sound and audiovisual archives and collections of graphic materials, produced by different kinds of subjects: physical persons, families, corporate bodies (cultural associations, school institutes, academies, trade-unions and political organizations, charitable and hospital institutions, public administrations and military bodies).

After some interventions whose purpose was the documents physical conservation, in 2008 the project has increased the activity of archives description by adopting *xDams*<sup>12</sup>, the platform of document management provided by the software house Regesta.com. *xDams* is a native XML application which has been developed completely for the web and provides an input record that is fully consistent with the international archival standards. This application is characterized by a set of application profiles of EAD, corresponding to the different archives typologies: paper archives, photographic, sound and audiovisual archives and collections of graphic materials. Up to now the use of the *xDams* platform has allowed to inventory almost 200 archives and to produce a collection of 300.000 documents descriptive records and authority records, thanks to the work of 82 archivists.

#### 4. The re-engineering of archival finding aids in *Una Città per gli Archivi*

In 2010, a new working group has been established inside the project. This new team has been composed by archivists, computer scientists and communication experts, whose function is to coordinate the development of a portal for the on line publication of the archival materials elaborated with *xDams* software. The web portal will probably be publicly launched by the beginning of 2013. Its functioning is ensured by an articulated technological infrastructure based on several different components.

Firstly, database of *xDams* represents the core information source of the portal, providing all archival descriptions encoded in native XML format and the related authority lists of access points (archives creators, persons, families and places quoted in archival descriptions, documents authors, archival descriptions subjects).

Moreover, the archival descriptions and the access points are available on the web portal by web services, working in a layer between the content management system *BEedita* and the *xDams* environment. This information retrieval process is mediated by the search engine *Cogito*, produced by Expert Systems<sup>13</sup>. *Cogito* is a useful platform based on algorithms of natural language processing which gives to the end users a possibility to carry out not only simple searches for keywords, but also advanced search queries in natural language. In fact, *Cogito* is able to analyze natural language as regards morphology, grammar, logic and semantics. Thus, the information retrieval engine deals and processes synonymity and polysemy.

Anyway, the heart of the CPA web portal is *Bedita* (Semantic Content Management Framework)<sup>14</sup>, provided by Chia Lab srl<sup>15</sup> and Channel Web srl<sup>16</sup>. *BEdita* is a platform created with open source components, providing a high-level content management system for the publication contents coming both from *xDams* (archival descriptions and authority records, both encoded in XML) and from the media server. *BEdita* guarantees also the generation of new contents (news, events, editorial

---

12 See <http://www.regesta.com/cosa-e-xdams/>.

13 See <http://www.expertsystem.net/page.asp?id=1515&idd=200>.

14 *Bedita* (<http://www.bedita.com/>) is available under two licensing scheme: an open source version released under *Affero General Public License 3*; a proprietary version, released commercially by the copyright owners, with advanced modules and the right to create derivative commercial softwares.

15 See <http://www.chialab.it/>.

16 See <http://www.channelweb.it/>.

contributions, virtual exhibitions) and the enhancement of user generated information resources, according to the typical models of web 2.0 (social bookmarking and social tagging).

The information available on the web portal include not only archival description and authority records created in *xDams*, but also many digital reproduction: image files reproducing iconographic and visual documents (photos, graphic papers, posters and placards), textual documents, multimedia files that provide in a digital format the sound and audiovisual documents originally produced in analog media. These digital objects are accessible to end users through the communication between *BEdita* and a media server providing necessary storage and streaming services, and of course, all the descriptive, administrative, and structural metadata regarding digital objects. “Una città per gli archivi” has adopted the *Metadata encoding and transmission standard* (METS) scheme, maintained by the Library of Congress<sup>17</sup>. This choice allows on the one hand to encode metadata ensuring that digital objects can be deposited and properly managed by specialized repositories created for long-term conservation; on the other hand METS provides metadata ready for the exchange of digital objects with other information systems, where they can be accessed, redirecting to the portal “Una città per gli archivi” for the complete completion of descriptions and media objects.

As regards the problematic issues resulting from the North American user studies quoted above, the planning of the portal “Una città per gli archivi” has led to some solutions, implemented in the portal prototype.

To face archival terminology barriers, the choice was not to use, as far as possible, archival jargon, both for the labels of archival descriptions and for search tools; on the contrary, the decision was to adopt terms of common sense and, as far as possible, universally known.

To help users with the hierarchal structure, the portal provides browsing features as intuitive as possible. Moreover, some alternative tools to hierarchical browsing were implemented, such as authority files of access points and a specific ontology dedicated to modern and contemporary history of Bologna. Its grid of instances can be browsed by users, requesting for each one the related archival descriptions.

To finalize the search functions supporting an immediate approach of users towards the documents, instead of the archivists approach towards documentation, the information retrieval engine was based on an algorithm of natural language, allowing the management of synonymies and the ranking of results based on a their semantic relevance, taking in consideration mainly the general search engines logics.

Finally, as regards the content visualization, a double approach has been chosen. From the ICT side, the re-embedding of metadata into text strings as close as possible to natural language was provided with the CMS supporting automatic retrieving and extraction of metadata. From the archival side of the medal, the project gave preference to structured texts, to facilitate users' decoding even in case of long contents, and excluded the use of archival abbreviations and terms. Moreover, as much as possible, archival descriptions have been supplied with the digital reproductions of the described documents: when the system will work at its full capacity, it will be populated by around 200.000 digital objects, each will be precisely contextualized by a set of metadata composing the relevant archival description.

---

17 See <http://www.loc.gov/standards/mets/>.

## 5. User studies for CPA portal prototype development

The CPA Portal is presently in a prototype phase; consequently it is not possible yet to examine the real experience of users using its contents and functionalities. Nevertheless, user studies could be crucial to guarantee this platform quality, collecting data and producing useful recommendations for present release finalization and for future developments. A multiple methodology has been adopted with different study methods applied, and a huge amount of qualitative data was collected.

Firstly a line of study involving all the internal staff of the project was organized. For the quality archival descriptions, the competence of archivists is crucial, but rarely those specialized content providers are directly involved in the development of output environments. The CPA study of users and functionalities resulted in two special focus groups, inviting together archival staff and portal developers. The sessions were scheduled at the beginning and at the end of the users' study activities. Those meetings were moderated according to a special protocol, mixing typical focus group's method with that of a moderated brainstorming, putting archivists versus engineers.

Then, five focus groups were organized in order to collect precious qualitative information about the potential target users' needs and satisfaction against CPA portal prototype. Altogether there were more than 60 people involved in the project: 15 high school students (aged 16-19), 16 university students in two different sessions (one in Macerata and one in Bologna, aged 23-29), 16 people classified usually as general public (aged 39-63, including high school teachers, administrative staff, an architect, one web master, three retired persons), and 16 archivists and cultural heritage professionals (aged 29-52, including librarians and experts in cinema history). All focus groups were based on the same detailed protocol, during each session of 100 minutes, participants could ask questions in six sections:

1. Demographic and ICT skills questionnaire;
2. prototype brief presentation;
3. First impressions and general discussion;
4. Execution of five tasks of use the prototype, whose results were recorded asking some questions for each task;
5. Advanced impressions and discussion;
6. Final discussion.

Finally, an expert study was scheduled for the next autumn, when the prototype will be almost ready to be published, and its functionalities will be thoroughly evaluated. As mentioned above, the usability of evaluation made by experts should be implemented in a preliminary version of the project, in order to consider all scheduled features and most of the contents [16]. Seven experts have been already involved in the project - 3 archivists (2 university professors and director of the City Archive), 3 digital librarians and metadata experts, and an ICT university professor. The protocol adopted will focus on performance, functionality, usability issues, terminology, and the quality of content.

## 6. First conclusions

Archives on line could be considered as special Digital Libraries: they share most characteristics and functionalities: they open unusual issues, especially to meet the expectations of users and guarantee their satisfaction. The web environment undermines the traditional cultural mediation between archivists and users, first of all by distinguishing the web output from the *encoded* input - it is *decoded* and clear. The archives professionals should offer to users centered displays of their contents, matching their descriptive techniques with human-computer interaction studies and cognitive psychology, checking their prototypes by applying activities of evaluation and testing. In this sense,

users' studies are crucial to finalize the publication environments, and to ensure a (really) good quality. The prototype stage of the project offers - when possible - a good occasion for evaluation and assessment of activities based on users' involvement.

The Italian CPA project shows that user studies are not too expensive; they require an expert (and neutral) coordination, a good organizational staff and a network of people and institutions.

### Bibliographic references<sup>18</sup>

- [1] Bates 1989 = M. J. BATES, *The design of browsing and berrypicking techniques for the online search interface*, in »Online review«, 13 (5), 1989, pp. 407-424, <http://pages.gseis.ucla.edu/faculty/bates/berrypicking.html>
- [2] Chapman 2009 = J. C. CHAPMAN, *What would users do? An empirical analysis of user interaction with online finding aids*, master's paper submitted to the faculty of the School of information and library science of the University of North Carolina, Chapel Hill, 2009, pp. 75.
- [3] Chapman 2010 = J. C. CHAPMAN, *Observing users: an empirical analysis of user interaction with online finding aids*, in »Journal of archival organization«, 8, 2010, pp. 4-30.
- [4] Daniels e Yakel 2010 = M. G. DANIELS, E. YAKEL, *Seek and you may find: successful search in online finding aid systems*, in »American archivist«, 73, 2010, pp. 535-568.
- [5] DL.org 2011 = L. CANDELA, G. ATHANASOPOULOS, D. CASTELLI, K. EL RAHEB, P. INNOCENTI, Y. IOANNIDIS, A. KATIFORI, A. NIKA, G. VULLO, S. ROSS, *The digital library reference model*, DL.org, 2011, pp. 273, <http://bscw.research-infrastructures.eu/pub/bscw.cgi/d222816/D3.2b%20Digital%20Library%20Reference%20Model.pdf>
- [6] Dobrev et Al. 2010A = M. DOBREVA, E. MCCULLOCH, D. BIRRELL, P. FELICIATI, I. RUTHVEN, J. SYKES, Y. ÜNAL, *User and Functional Testing. Final report*. 2010, <http://version1.europeana.eu/web/europeana-project/documents>
- [7] Dobrev et Al. 2010B = M. DOBREVA, E. MCCULLOCH, D. BIRRELL, P. FELICIATI, I. RUTHVEN, J. SYKES, Y. ÜNAL, *Digital Natives and Specialised Digital Libraries: a Study of Europeana Users*, in S. KURBANOGU et al., *Technological Convergence and Social Networks in Information Management*, Springer-Verlag, Berlin-Heidelberg, 2010, pp. 45-60.
- [8] Duff, Harris 2002 = W. DUFF, V. HARRIS, *Stories and names: archival description as narrating records and constructing meanings*, in »Archival science«, 2/3-4, 2002, pp. 263-285.
- [9] Duff, Johnson 2002 = W. DUFF, C. A. JOHNSON, *Accidentally found on purpose: information-seeking behavior of historians in archives*, in »Library quarterly«, 72 (4), 2002, pp. 472-496.
- [10] Duff, Stoyanova 1998 = W. DUFF e P. STOYANOVA, *Transforming the crazy quilt: archival displays from users' point of view*, in »Archivaria«, 45, 1998, pp. 44-79.
- [11] Feliciati 2007 = P. FELICIATI, *Dall'inventario alla descrizione degli archivi in ambiente digitale: si possono offrire agli utenti risorse efficaci?*, in "E-laborare il sapere nell'era digitale" conference, Montevarchi (Italy), 22-23 November 2007, <http://hdl.handle.net/10760/11159>
- [12] Feliciati, Valacchi 2010 = P. FELICIATI, F. VALACCHI, *To be or to appear? The hidden archives in the digital age*, in »History of education & children's literature«, vol. V, n. 2, pp. 453-476.
- [13] Griffiths 2012 = J. R. GRIFFITHS, Questionnaires, interviews and focus groups as means for user engagement with evaluation of digital libraries, in *Users Studies for Digital Library Development*, M. Dobrev, A. O'Dwyer and P. Feliciati (eds.), Facet Publishing, London, 2012, pp.65-73.
- [14] ICA 2000 = INTERNATIONAL COUNCIL ON ARCHIVES, *ISAD (G): General international standard archival description, second edition*, Ottawa, 2000, <http://www.icacds.org.uk/eng/ISAD%28G%29.pdf>
- [15] ISO 2001 = ISO/IEC 9126-1:2001. *Software engineering — Product quality*. Part 1: Quality model.

---

18 The here cited links have been visited last time on June 15th, 2012.

- [16] Klas 2012 = C.-P. KLAS, *Expert evaluation methods*, in *Users Studies for Digital Library Development*, M. Dobrev, A. O'Dwyer and P. Feliciati (eds.), Facet Publishing, London, 2012.
- [17] Light 2008 = M. LIGHT, *The endangerment of trees*, 2008, pp. 6,  
<http://www.archivists.org/publications/proceedings/EAD@10/Light-EAD@10.pdf>
- [18] Meissner 1997 = D. MEISSNER, *First things first: reengineering finding aids for implementation of EAD*, in »American archivist«, 60, 1997, pp. 372-387.
- [19] Ribeiro 2001 = F. RIBEIRO, *Archival science and changes in the paradigm*, in »Archival science«, 1, 2001, pp. 295-310.
- [20] Schaffner 2009 = J. SCHAFFNER, *The metadata is the interface: better description for better discovery of archives and special collections, synthesized from user studies*, OCLC research, Dublin (Ohio), 2009, pp. 18,  
<http://www.oclc.org/research/publications/library/2009/2009-06.pdf>
- [21] Scheir 2005 = W. SCHEIR, *First entry: report on a qualitative exploratory study of novice user experience with online finding aids*, in »Journal of archival organization«, 3 (4), 2005, pp. 49-85.
- [22] Sundqvist 2007 = A. SUNDQVIST, *The use of records: a literature review*, in »Archives and Social studies: a Journal of Interdisciplinarity Research«, 1 (1), 2007, pp. 623-653.
- [23] Tsakonas 2012 = G. TSAKONAS, *Users within the evaluation of digital libraries*, in *Users Studies for Digital Library Development*, M. Dobrev, A. O'Dwyer and P. Feliciati (eds.), Facet Publishing, London, 2012, pp.51-61.
- [24] Yakel 2003 = E. YAKEL, *Impact of internet-based discovery tools on use and users of archives*, in »Comma«, 2/3, 2003, pp. 191-200.
- [25] Yakel 2004 = E. YAKEL, *Encoded archival description: are finding aids boundary spanners or barriers for users?*, in »Journal of archival organization«, 2 (1/2), 2004, pp. 63-77.

[p.feliciati@gmail.com](mailto:p.feliciati@gmail.com)

[aalfier@regione.emilia-romagna.it](mailto:aalfier@regione.emilia-romagna.it)