

Digital Preservation in Europe

Strategic Plans, Research Outputs and Future Implementation. The Weak Role of the Archival Institutions

Maria Guercio

University of Rome Sapienza

Abstract

The European investments for digital preservation in the last decade have been large and persistent but not able to support, at the moment, an accepted common vision, general services and adequate infrastructures. The author investigates this weakness by analyzing the nature of the European funding programs in the field and the weak role played by memory institutions, specifically those with archival competences and knowledge, in the European research activities. A specific attention will be dedicated to the relevance for successful results of concepts such authenticity and the role played by a consistent terminology.

Author

Full professor in archival science and electronic records management at the University of Rome Sapienza (coordinator for research projects at Digilab Centre), she cooperates with the State authority for ICT to define Italian legislation for ERMS and manages ERM training programs for the government school for public administration. partner in many international projects for digital preservation (ERPANET, DELOS, CASPAR) director of InterPARES Team Italy (1999-2012), she is co-leading the investigation on digital authenticity for the European project APARSEN (2011-2014). She is part of the steering committee of Section for Archival Education and component of Programme Committee of International Council on Archives. Author of many articles and manuals in the field, in 2009 she has been the winner of Emmet Leahy Award for information and records management.

1. A Premise

The subject here discussed presents a high level of complexity due to the dynamic nature of the technological innovation, the increasing role of the organizational aspects and the interdisciplinary character of the research projects dedicated to digital preservation. To face these difficulties an accurate in-depth and contextualized approach is required. Even if this effort cannot be analyzed in detail on this occasion, this conference and its setting offer an extraordinary opportunity for discussing the main promising but also controversial questions in this field. In Vancouver, thanks to UBC's continuing research work on digital preservation, an incredibly productive international environment has been developed in the last twenty years with stimulating results, such as a common terminology, robust conceptual frameworks and significant occasions for international and cross-domain comparisons and

advanced educational programs. The UBC's projects – this is an important aspect of their success strictly related to their special original nature – had a very clear archival focus but had and have also the capability of involving other communities with an interdisciplinary approach. Moreover, the questions here proposed for debate have to face new scenarios and new organizational environments which are still undefined: large investments have already been spent or, at least, planned. The main innovative and challenging question is related to the fact that the global financial crisis not only implies new economic difficulties for the research environment but also makes evident the major responsibilities of universities and scientific center, recently defined as *entrepreneurs of innovation* able to “balance the need to articulate a broad strategic vision with the need to execute the day-to-day activities that translate the vision into reality”¹. Therefore, the topics here proposed for investigation cannot be ignored, not only for professional and technical reasons, but also and mainly with reference to the centrality in the global society of “research labs, classrooms, and innovations centers, where big ideas are hatched and subsequently translated into reality”.²

Similar to UBC research, many European research initiatives have been largely funded in the same field of digitization and digital preservation in the last decade, but with minor success. They have not been able to maintain a similar level of continuity and, in particular, a comparable level of such global influence. The role and the relevance of the European effort in this research area are undeniable, but they have not been able to re-create (with the partial exception of ERPANET)³ the InterPARES atmosphere of international cooperation and a similar original and authoritative contribution to the research in one scientific domain (in case of InterPARES, the preservation of authentic digital records) and to the enlargement of the interdisciplinary cooperation boundaries. Are these exceptional outcomes due only to the unquestionable quality the researchers involved in the project possess or is something not convincing in the direction and strategies followed in the European funding system in this specific sector? Other successful projects (which are not strictly related to the digital preservation, but connected to the digital heritage) can be mentioned because of their similar global influential capacity as InterPARES: Dublin Core Initiative, the OAIS model, Encoded Archival Description standard, digital curation paradigm, audit framework for trusted digital repositories are just some of the best known products, developed in one sector or thanks to the strict coordination of archival and librarian scholars and professionals and transformed into general and basic elements and tools for building enhanced infrastructures. None of them has been implemented thanks to European funds. Besides, many or all of them have been developed on a domain basis, with a strong commitment of one or two scientific and professional communities. If this did not happen by chance, is it possible to identify and replicate successful conditions as a framework for developing an efficient research environment or, at least, to avoid negative consequences if the requirements are not respected?

These questions are simple and demanding at the same time. Specifically, they concern the identification and the analysis of the main obstacles weakening the relevant investments of the European Commission in the area of digitization and digital preservation and the role played by a cross-domain strategy as factor influencing and partially determining the fragmentation of intermediate and final results

¹ Holden Thorp and Buck Goldstein, *Engines of Innovation. The Entrepreneurial University in the Twenty-First Century*, Chapel Hill: The University of North Carolina Press, 2010: 8.

² *Ibidem*, p. 2.

³ “ERPANET - [Electronic Resource Preservation And Access Network](http://www.erpanet.org)”, Accessed August 22, 2012, www.erpanet.org.

of the financed projects. Some assumptions are of course required as a basis for a sustainable line of reasoning.

A first provisional assumption is that the level of success in this complex and dynamic research area was and is linked to the capacity of investigating the domain specific questions with an open approach, based on general principles, solid and consistent concepts and rigorous methods. This means that a common methodological ground is necessary and has to be accepted as fundamental by the research community. The existing scientific traditions, of course, have to be assessed and updated. Well designed and contextualized dictionaries have to be discussed and approved. For this specific reason, a long and cumbersome series of activities are necessary before an investigation program could be operative and fruitful.

A second, still provisional, assumption concerns the interdisciplinary nature of the research. Once more, a solid interdisciplinary approach requires highly qualified specialization. The integration of competences can be fruitfully supported only as a second step after the recognition and the comparison of the research values inside each scientific domain.

The theoretical implications behind these two points are very complex and cannot be here investigated in detail. In this specific context, the attention is *simply* dedicated to illustrate the negative consequences (in terms of lack of generally recognized results because these assumptions have been ignored) of the European ‘political’ strategy dedicated to support digitization of cultural heritage and its preservation by facing the technological innovation. This strategy was initially defined in the period 2005-2007 and only partially updated in 2011⁴ with the goal of financing mainly (if not solely) interdisciplinary projects explicitly focusing on digital convergence. But the nature and the content of this interdisciplinary approach and the meaning and consequences of the digital convergence have never been defined and the convergent inclusion has excluded many relevant scientific communities and domains. The archival sector, in particular, even if its digital heritage is one of the most challenged by present and future technological innovations, has been substantially ignored by the European investments. One of the reasons is related to the underestimation of the potential negative potential impact and consequences of this marginalization (not only for the domain itself) in terms of advanced solutions for the research and for services implementation.

As a matter of fact, the limited contributions made available to the research activities of the European archival community (as will be further illustrated) have prevented its institutions from participating efficiently in the investigation regarding the main innovative topics. Therefore, the great European documentary traditions, diversified but also consistent in their basic common principles and methods, could not have been recognized and exploited in Europe and internationally with at least two relevant limits for the global (not only European) research. Firstly, it has been impossible to develop at the European level a strategy toward a systematic and standardized approach for electronic recordkeeping and preservation systems⁵. Secondly, this delay has prevented the archival and record management

⁴ The relevant European reports on these aspects are: “i 2010 Digital Libraries Initiative” (2005), Commission Recommendation on Digitisation and Online Accessibility of Cultural Material and Digital Preservation (2006 – updated October 27, 2011), “Member States’ Expert Group on Digitisation and Digital Preservation (2007), “Communication from the Commission: ‘Scientific Information in the Digital Age: Access, Dissemination and Preservation’ (2007), Recommendation on Digitisation and Digital Preservation (October 2011).

⁵ The case of European recommendations for electronic record management systems, known as MoReq (MoReq1, MoReq 2 and MoReq2010), is emblematic: the specifications have required 10 years to be approved and to develop certification process but the final version of MoReq2010 is still under development, services for certification are not in place: and the European organizations at the moment are not compliant with them. See M. Guercio, “MoReq1,

community from providing effectively (that is as a recognized and authoritative network) its contribution to the research in the specific field of digital preservation and to the definition of digitization standards and parameters. The low level of qualified convergence and satisfactory level of interoperability of digital library initiatives (including Europeana, the EU digital library) provides undeniable evidence for these limits.

Since its first steps, and especially after 2007, the European strategies for the digital heritage have marked the centrality of actions aimed at promoting the cultural contents convergence, by overcoming quite always the specificities in the research investments and favoring (and selecting) quite only projects able to guarantee the largest participation of cultural institutions of the widest and most variegated nature and provenance as possible: as mentioned, the focus of all the European recommendations for digitization and digital preservation was and is on convergence of domains for accessibility and for interoperability. The quantity but not the quality is crucial for this strategy. At a matter of fact, the explicit political goal in 2007 and, even more, in 2011 was and is to ensure a critical mass of contents for the European digital library, called Europeana by:

[...] setting up or reinforcing *national aggregators bringing content from different domains into Europeana*, and contributing to *cross-border aggregators in specific domains or for specific topics*, which may bring about economies of scale, ensuring the use of *common digitisation standards* defined by Europeana in collaboration with the cultural institutions in order to *achieve interoperability of the digitised material at European level*, as well as the systematic use of permanent identifiers.⁶

The challenges involved in these new services risk to underestimate the role of the cultural heritage scholarships and knowledge differentiation and specialization, which were and are at the basis of the European quality. “Convergence against integration” can appear as an easy and rapid shortcut for dealing with the digital convergence processes, but not necessarily will provide efficient solutions. In our specific area, this general policy had the consequence, among other things, of pushing the research institutions (and of course also the institutions of memory) to present projects characterized by a high degree of heterogeneity of purposes, participants and research components and activities. Even if nobody in any domain explicitly contested and contrasted this perspective and this orientation, the final result has been the marginalization of all the sectors not really involved in digital access and dissemination of published and native open public materials (that is digital resources immediately available on the web), such in the case of e-health records, or in the private business environment and in general in the recordkeeping systems).

Of course, as a consequence, any possibility for scientific investigation in each domain has been sacrificed to ensure multiplicity and diversity of actors, of contents and vocabularies and, more recently, to increase the quantity of information available and not the quality of its representation in terms of contextualized relations (that is with reference to the digital objects intelligibility and the capacity to assess over time their integrity and authenticity, to trust both digital contents and their relations). Even

MoRe2 e MoReq2010: raccomandazioni e prove tecniche di certificazione per la gestione informatica dei documenti”, in *Archivi*, 2012, 1: 7-32.

⁶ Commission recommendation of 27.10.2011 on the digitisation and online accessibility of cultural material and digital preservation, Accessed August 21, 2012,

http://ec.europa.eu/information_society/activities/digital_libraries/doc/recommendation/recom28nov_all_versions/en.pdf.

more, the illusion that the complexities involved in the digital preservation research could find at the end exhaustive answers in the multidisciplinary comparison has influenced the whole research sector: the partnerships could not be developed on the basis of the specific qualifications of the institutions to involve. The domain diversity of the projects partnerships was and still is an obliged requirement for successful funding. It has to be said that the institutions of memory (mainly the archival organizations) have shown here their limited capacity for developing strategic alliances (relevant not only for funding and for exploiting finance channels but mainly to create strong and permanent institutional interconnections. In the specific archival sector, another negative factor has been the substitutive role of coordination assigned, on an unclear basis and with insufficient commitment, to organizations like DLM Forum Foundation⁷. This association, born in 1994 as “an inter-disciplinary cooperative effort led by the EU member states and the European Commission”, became in 2002 an independent body and, in 2010, a non-profit foundation in the area of “archive, records, document and information lifecycle management throughout Europe”. More and more the Foundation has seen the increasing interests and direct participation of software companies and has diminished its original potential nature of a neutral network based on institutional values. Even if recognized by the European Commission, the organization has never been strongly involved in the European research projects. The role assigned by the Commission to the European Archival Group which (will be analyzed more deeply in the second part of this presentation) had a similar marginal nature for the opposite reason: it has been supported for coordinating archival actions among the European institutions but it has never been included in the core research projects dedicated to the technological innovation, specifically preservation and digitization. The only exception is the APENET and now APEX projects and the portal dedicated to make the archival data available on the web and cooperate for ensuring interoperability with Europeana.

It is surprising to observe that the international scenario shows less archival marginality than the European context, even if (or because?) the work is done on the voluntary basis and the resources are very limited in comparison with the large amount of European investments. It includes the work done under the umbrella of the International Council on Archives, or the international research projects such as InterPARES and the projects on trusted digital repositories, but also the working groups of ISO committees and subcommittees on archival and record management (TC 46 SC 11). The archival Europe supported at various levels (mainly with direct participation) all these efforts thanks to the commitment of single institutions, but these actions and their strategic objectives have been substantially ignored by the EU funding bodies, as the available information on the last decade funded projects clearly illustrates:

- With exclusive reference to the research on digital preservation for the period 2006-2011, fifteen projects have been financed under ICT programs (the most important in the field here investigated); they have been funded with 86 million euros with the specific aims of creating technology solutions and innovative methods to keep *digital resources available and useable over time*,⁸

⁷ The recommendations of the European Council of November 14, 2005 were limited to ask the initiative of the Commission on the following aspects: “further development of European interdisciplinary co-operation on electronic documents and archives, updating and extension of the requirements for setting up electronic document and archive management systems, such as MoReq and a continuation of the DLM Forum conferences on electronic documents and archives”. For the history of FLM Forum Foundation see http://www.dlmforum.eu/index.php?option=com_content&view=article&id=13&Itemid=15&lang=en. Accessed August 22, 2012.

⁸ Pat Manson, *Stakeholders' Workshop on eArchiving* (Luxembourg 24 February 2012).

- In 2012 other 30 projects have been approved for the amount of other 30 million euros;
- None of the funded projects had and has an archival leadership (academic or institutional);
- The institutions with a specific and recognizable archival competence involved in these projects are very few (5-6), while the European partners involved are hundreds: as far as I know, Institution for Archival Studies at the University of Urbino (Caspar⁹), the National Archives of the Netherlands, the National Archives of England, Wales and the United Kingdom, the Swiss Federal Archives (Planets),¹⁰ the National Archives of Sweden and the National Archives of Estonia (Protage).¹¹

In contrast with the recent evolution, the first European strategy for digital preservation was, at least apparently, differently oriented. Only 4 partners ran the ERPANET project, a network of excellence active in the period 1999-2003. Three of them had archival competences (the National Archives of Netherlands, the National Archives of Suisse, and the Institution for Archival Studies at the University of Urbino). The principal coordinator, Seamus Ross, director of HATII at the University of Glasgow, was at that time a close observer of progress made by the documentary disciplines in the field. The same partners (only the National Archives of Suisse were not included) successfully collaborated in the DELOS¹² project within the same work package dedicated to the digital preservation and developed promising ideas which were further implemented by following projects such as PLANETS¹³ and DPE¹⁴ (among other results, a set of procedures to document self-auditing with reference to the requirements for digital repositories trustworthiness, the systematic analysis of file formats for preservation, the creation of test-bed environments).

In the recent projects the archival dimension of the research progressively became marginal and sporadic. Other domains have been able to develop their influence and assume a strategic role in cooperation initiatives such as the library sector with reference to the European digital library Europeana¹⁵ or the network of government cultural institutions which still play a coordinating role thanks to a continuing series of EU funded projects (Minerva and Minerva eC¹⁶, Michael¹⁷ and, more recently, Athena¹⁸ and Linked Heritage¹⁹). Also the audiovisual institutions were strongly supported in Europe with dedicated programs able to create a healthy chain of continuity (Presto, Prestospace²⁰ and Prestoprime).²¹

⁹ “Caspar - Cultural, artistic and scientific knowledge for preservation, access and retrieval”, 2002-2006, Accessed August 20, 2012, <http://www.casparpreserves.eu>.

¹⁰ “Planets - Preservation and long-term access through networked services”, 2002-2006, Accessed August 20, 2012, www.planets-project.eu.

¹¹ “Protage”, 2010-2012, Accessed August 20, 2012, www.protage.eu/project.html.

¹² “Delos Network of Excellence on Digital Libraries”, 2004-2008, Accessed August 20, 2012, www.delos.info/.

¹³ Planets - Preservation and long-term access through networked services, Accessed August 22, 2012, www.casparpreserves.eu e www.planets-project.eu.

¹⁴ “DPE - DigitalPreservationEurope”, Accessed August 22, 2012, www.digitalpreservationeurope.eu.

¹⁵ “Europeana”, Accessed August 20, 2012, www.europeana.eu/portal.

¹⁶ “Minerva” and “Minerva eC”, Accessed August 20, 2012, <http://www.minervaeurope.org/>.

¹⁷ “Michael”, Accessed August 20, 2012, <http://www.michael-culture.org/it/home>.

¹⁸ “Athena – Access to Cultural Heritage Network Across Europe”, 2008-2011, Accessed August 20, 2012, <http://www.michael-culture.org/it/home>.

¹⁹ “Linked Heritage”, 2011-2013, Accessed August 20, 2012, <http://www.linkedheritage.eu/>.

²⁰ “Presto”, 1999-2002, Accessed August 20, 2012, <http://presto.joanneum.ac.at>; “PrestoSpace”, 2004-2007, Accessed August 20, 2012, www.prestospace.org.

²¹ “Prestoprime”, Accessed August 20, 2012, www.prestoprime.org.

How effective is the cooperation in these cases is a question assessed only when the single project is evaluated by the Commission in the course of the funding process. At the moment there are no mechanisms in place (but only political initiatives) to develop on more general basis assessments of research areas and/or financing lines, or to evaluate and solve critical questions like the lack of interactions and cooperation or integration among sectors, even if explicitly recognized by official documents.

It is the case of the official critical remarks (published in the final report of the project) about the low level of inter-sector cooperation verified and recorded by the archivists involved in the APANET project when a *convergence* process was activated to create compliance and interoperability between the descriptive systems and standards of archival sector and those already developed by Europeana:

This is the final report of the APENet WP3 team in which it describes the interoperability efforts it undertook throughout the whole APENet project period as well as the difficulties it came across and the ways it found to deal with them [...]. The overall conclusion of this report is that the APENet WP3 team has succeeded in making the technical interoperability fully operational, so in fact has done the job it was supposed to do, but could have accomplished more in case the Europeana team would have been more cooperative. The reason for this is that – despite a lot of efforts from the APENet WP3 team – the interoperability on a strategic level has been a one way communication most of the time during the entire APENet project period, plus the fact that on a technical level Europeana was much like a moving target and at the end of the APENet project even out of reach for the APENet WP3 team.²²

The fact is that the European archival institutions, whose historical custodial function could have supported, with the knowledge and experience accumulated for centuries, the understanding and the translation in the technological environment of crucial concepts such as digital trust, reliability, accuracy and authenticity and identify implementation methods and tools in the field of digital preservation, have never acted as main characters or leading protagonists. They obtained merely *protection* but not recognition as crucial players for central challenges. They have not yet created an effective European network able to support research and initiatives related to digitization processes and digital preservation. The DLM Forum Foundation²³ – as has been already observed– lacks the institutional status for playing this role. The only initiative, lightly supported by the European Commission and previously mentioned, is the *European Archives Group*²⁴ but its attention is mainly dedicated to the traditional and still complex challenges of the analogue archives.

In the specific area of European research for digital preservation, a general and relevant critical question concerns the brief funding period always granted to the projects, no matter how promising the initiatives are: according to this policy, each research project should be able to survive and become sustainable in only 3-4 years of financial support and should be able to develop its own research center for strategic future programs. In other scientific sectors (as in the so-called *basic research*) European support

²² APENET, “Final Interoperability Report Deliverable D.3.2”, Accessed August 21, 2012, www.apenet.eu.

²³ “DLM Forum”, Accessed August 20, 2012, <http://www.dlmforum.eu>.

²⁴ “European Archives Group – EAG”, Accessed August 18, 2012, http://ec.europa.eu/transparency/archival_policy/eur_arch_group/index_en.htm. The working group is based on the participation of National archives directors. The main activities have concerned the creation of a European archival portal (Apenet, now Apex), the database of European archival legislation Euronomos (in cooperation with the ICA European Branch EURBICA), the definition of guidelines on the thefts of archival materials and other studies. Relevant goals and actions, but not sufficient for building a strategy and a solid and persistent infrastructure.

ensures more continuity of funding and, for this reason, makes success and international recognition more possible.

The fragmentary nature of the European support for digital preservation research – partially mitigated by the not unusual opportunity for an institute and its researchers of being funded for years even if in different projects (as happened in the case of University of Urbino) – has strong and specific political EU reasons, which cannot be easily overcome even if in the European Research Framework new perspectives have been identified and could be able to create a radical and positive evolution in the near future. It is moreover clear that many other factors contributed to maintain the present situation, specifically characterized by isolated projects and by difficulties in creating fruitful interconnections and supporting an effective strategy:

- The market and the stakeholder interests are relevant, variegated and conflicting and, also for this reason, the differentiation of the investments is prevailing in the evaluation process;
- Archival institutions, which are involved (even if they are not numerous) in the European projects, have not been able to develop a common strategic leadership and promote a coordinating action; the other institutions which have not been involved do not seem even aware of their exclusion in the last decade of European funding.

With particular reference to the last point and to the obstacles that have further limited the effective cooperation in Europe among archival institutions, many elements have to be considered. They include:

- The institutional resistance and reserve about working on single projects without insurance for continuity, while (or because) they are facing other dramatic, still open and unsolved challenges (such as the documentary proliferation and the complex problems related to e-government initiatives both in the legislation and in the national governmental implementation);
- The difficulty of and the delay in conjugating the institutional activity for protection and custody of archival sources with the scientific investigation and in cooperating with the research initiatives conducted by the academic institutions;
- A traditional autonomous attitude attested in the past by the delay in supporting standardization processes;
- A passive acceptance of the European policies established in the period 2005-2007 focused on principles of convergence and interoperability of the digital cultural heritage;
- The underestimation at European and at national level of the problems related to the preservation of digital archives and to the need of building an early and consistent network of custodial institutions, which are authoritative, reliable, certified, trustworthy and neutral. It should be sufficient to remember the very recent proposal of the European Commission to modify the regulation related to the European historical

archives²⁵ with reference to the removal of the obligation of transfer to the central archival repository in Florence for the European institutions when the records are digital: in this case the creators, according to this new proposed rule²⁶ and against any consolidated archival principle, will maintain their complete responsibility for ever (“the originating institutions will remain responsible for the long-term preservation of their digital archives”). Because no other rule is defined with respect to this responsibility, the Commission proposal seems to completely ignore the risks of losing control over the quality and the integrity of the records when they are no longer active (and relevant for the creators) and are not yet (and perhaps will never be) transferred to dedicated *trusted* archival repositories. As stressed by Angelika Menne Haritz, when the proposal was discussed at the EAG meeting in October 2011, because “digital archives are not fixed”, only the custody in archival repository assigned to third neutral parties is able to guarantee that “these archives will not change after they are opened to the public”²⁷. The decision of the European Commission is in patent contradiction with the main results of the research projects funded by the Commission itself, explicitly in favor of the creation of certified institutions dedicated to the digital preservation as the essential requirement for long-term preservation.

From this point of view the archival institutions, the absence of archival competences and knowledge in the European research for digital preservation determined, among other consequences, that many research questions, which are still relevant and critical for the domain, have not yet found an adequate and systematic analysis and risk remaining unanswered (and sometimes also unexpressed). Only some of them can be here listed:

- How should one apply OAIS in the archival domain?
- How should one transform the archival requirements for information representation and the specific references to contexts and complex relations into the digital library standards without losing the intelligibility of the digital objects made available on the web?
- How should one develop services able to document the custodial history in the transfer processes?
- How and when should one document the authenticity evidence and who would be responsible for its maintenance?

These questions (and many others) are urgent and cannot find adequate answers if the institutions are left alone. It is not by chance that recently, thanks to the request of some European governments and,

²⁵ “For a Council regulation amending, Reg. (EEC/Euratom) No 354/83, as regards the deposit of the historical archives of the institutions at the European University Institute in Florence”, Accessed at August 24, 2012, <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:DKEY=687809:EN:NOT>.

²⁶ The annex related to the provision for the deposit of the historical archives of the European Union institutions makes only explicit that “in the case of digital archives, the EUI shall have permanent access to the documents in such a way as to allow it to fulfill its obligation to make the historical archives accessible to the public from a single location and to promote their consultation”, *ibidem*.

²⁷ European Archives Group – EAG, Minutes, accessed August 18, 2012, http://ec.europa.eu/transparency/archival_policy/docs/eag/cr/111007_minutes_en.pdf.

specifically, of Slovenia, a specific action concerning *long-term data archiving* has been finally proposed to the European Commission (*Proposal for support of eArchiving in CIP ICTPSP WP12*) and partially accepted. The focus is – at the moment – limited to the transfer of custody for digital archives disposed for long-term preservation as proposed by Jože Škofljanec (Archives of the Republic of Slovenia) at the first meeting. This meeting was held in Luxembourg in February 2012 with the participation of many archival stakeholders and experts for digital preservation and the proposal was presented with the following motivations: “the level of digital preservation is very different among member states; (limited) practical experience is available; pan-European standardisation is taking place in the field of RM (MoReq) but not regarding ingest procedures“. The Slovenia proposal was promptly supported by all the archival institutions of many European countries (among others Austria, Denmark, Estonia, Finland, Italy, Hungary, Latvia, Netherlands, Norway, Portugal, Spain, Sweden, UK) and the experts involved in this internal workshop. The project obtained the recognition of its relevance when the government representatives met some months later. The basic request concerned the increase of efforts and support for developing guidelines and common tools (that is, fewer projects and more coordinating actions, programs and working groups). Specifically, the final report made explicit the need for “common set of guidelines and tools for transfer and ingest processes and formats in line with European interoperability initiatives“ to make available to the member States “simple pan-European access tools and e-services – strengthen reuse“. A pilot project is now planned in the form of a call for a proposal for “developing and piloting eArchiving service solutions, the necessary technical infrastructure, policies and processes to support the preparation and transfer of content and its ingest to repositories, ensuring its authenticity, provenance and integrity over time as well as its continued availability and usability” – in synthesis for implementing “scalable sustainable practical eArchiving services”²⁸. Of course, the recognition of the critical value of this topic does not imply a correct archival solution. The risk of delegating the main preservation problems (such as authenticity and integrity) to technological mechanisms is always around the corner, as was made evident by the Italian legislation approved in 2004²⁹.

The methodological contribution of the archival knowledge for digital preservation research

From the methodological point of view, the most relevant open question concerns the concrete and specific content of the contribution that the archival domain should be able to provide to the research dedicated to the digital preservation. The answer can be found by analyzing the results of the European researches and, specifically, their main weaknesses. Of course, there are no parameters for identifying such deficiencies and no metrics available to support such analysis. Evidence for this risk can be found in the APARSEN deliverable for work package on authenticity and interoperability³⁰. On this occasion the researchers had to revise dozens of deliverables and reports published by recent European projects dedicated to digital preservation and were able to verify how limited at the moment the investments (internal to the projects) are for the definition of a project’s stable vocabulary and for a systematic

²⁸ Cfr “Orientations for CIP ICT PSP WP2013” (internal document) and in particular the point dedicated to “Scalable sustainable practical eArchiving services”.

²⁹ See Maria Guercio, *Archivistica informatica*, Roma: Carocci, 2010, chapter 5.

³⁰ APARSEN Project: Deliverable 24.1. Report on Authenticity and Plan for Interoperable Authenticity Evaluation System, 2012, Accessed August 22, 2012, http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2012/04/APARSEN-REP-D24_1-01-2_3.pdf.

conceptual framework³¹ and how difficult is to orient users without such tools which be able to function as *protocols of understanding*.

On the contrary, both the terminology and the conceptual framework are fixed ideas, even an obsession, for the archival community, specifically when they operate at the international level. There is no doubt about the capacity of the documentary disciplines and, in particular, the archival science, to provide effective contribution to this activity, specifically for the creation of a theoretical outline. The international experience matured by the archival community on this subject is long lasting and well qualified (as it will be later explained).

With reference to the conceptual framework, someone could object that digital preservation systems already follow a standardized reference model, which is at the same time both consolidated and flexible, that is the OAIS model. This model is very important and generally appreciated both when general rules have to be developed (see the case of Italian legislation under approval³²) and when a digital repository has to be created and implemented (as in the case of Regione Toscana which is presented at the conference by Ilaria Pescini paper) or reconsidered (again, as in the case of e-health standard for Italian government, UniSInCRO³³). This model is today and will remain for a long-time the benchmark for digital preservation projects, but (also because of its nature necessarily open, flexible and abstract), it is unable to meet more specific requirements. The critical aspects regard not only the sectorial, domain-dependent vocabularies, but also the conceptual structure needed for sustaining in the daily life of the digital resources the implementation of the main categories of Preservation Descriptive Information (provenance, reference, context and fixity). These categories are very often difficult or impossible to handle separately. A robust methodology is required to avoid ambiguities and overlapping³⁴. It is not by chance if the just published release of OAIS, has reinforced its relationships with some archival concepts, such authenticity, even if it developed in a peculiar manner³⁵.

In particular, the authenticity is a very crucial subject, increasingly recognized for its centrality among the concepts commonly used and referred to by the communication, information and knowledge society. The information society by its nature directly works on the creation and narration of social and

³¹ It is frequent to meet new and specific glossaries, created for each deliverable and developed by simple accumulation of terms of various provenances. Even the compliance with recognized standards is not always required or pursued. Of course, building and sharing vocabularies is a very complex and ambitious task: it is never a question of collecting terms, but it implies a strict control of the relations and, more important, their contextualization. The agreement among co-authors is also a very demanding work.

³²“Regole tecniche di conservazione di documenti digitali”, <http://www.digitpa.gov.it/sites/default/files/Bozza%20-%20Regole%20tecniche%20conservazione.pdf>, accessed September 5, 2012.

³³ Uni 11386:2010 – “Supporto all'interoperabilità nella conservazione e nel recupero degli oggetti digitali (SInCRO)”, Accessed August 22, 2012, <http://webstore.uni.com/unistore/public/searchproducts>

³⁴ To make these concepts practically applicable in preservation cross-domains environment, a double exercise is required. First of all it is necessary to translate/accommodate them into the OAIS framework, specifically into the information package relevant for handling digital content preservation, the Preservation Description Information which includes Reference, Context, Provenance and Fixity Information. Of course, this ‘translation’ is not automatically extensible: an interpretation is required based on the definition of resource typologies and linked to specific domains knowledge. Some contradictions among OAIS PDI categories still require to be solved, as in case of the reference information: for archival records this concept implies identifiers of various levels (persistent identifiers, but also registry, classification code) and includes context information like archival bond while OAIS defines context as a separate area of PDI (the relationships of the Content Information to its environment, i.e. why it was created, how it relates to other Content Information objects, etc.).

³⁵ CCSDS: Reference Model for an Archival Information System – OAIS. Draft Recommended Standard, 650.0-P-1.1 (Pink Book), Issue 1.1, 2009, Accessed August 22, 2012, <http://public.ccsds.org/sites/cwe/rids/Lists/CCSDS%206500P11/CCSDSAgency.aspx>.

individual identities and, for this reason, requires tools, procedures and fundamentally robust concepts for entrusting and documenting their authenticity specifically when facing the digital world challenges. Moreover, the authenticity in the sense of identity of the resources and their integrity has not by chance been developed by the documentary disciplines. It has been creatively and thoroughly analyzed by Paola Carucci³⁶ with reference to the contemporary records and investigated for the digital environment by Luciana Duranti and other InterPARES researchers.³⁷ The knowledge and the experience of archival institutions and scholars are in this respect undoubtedly central and the concept of authenticity is undoubtedly central for the research on digital preservation and for building measures and tools for trusted digital repositories of any type.

Another sector where the archival contribution is substantial and unavoidable – as, at the end, was also by the information scientists involved in the research environment – concerns the use of principles, concepts and tools borrowed from the *record management*. This also includes recommendations and standards developed in the course of the last decade by ISO committee on archives and record management and by ICA. Among other things, they concern the recognition that the digital preservation, intended as dynamic process for digital continuity, can be sustainable and can be handled only on the basis of information lifecycle management or in the form of digital continuity because of standardized and well controlled vocabularies consistent with the functional model which governs the information phases and layers. From this point of view OAIS is the essential reference model, but – as previously underlined – is not self-supporting when consistency of contents is required.

With reference to the global research on digital preservation, conceptual frameworks, well defined vocabularies, contextualized concepts should be further developed and discussed if a real and effective interdisciplinary approach would be supported for improving the general investigation methodology. They are fundamental requirements for any successful project, specifically when the sector is undetermined and in continuing transformation as the digital environment is. As was underlined years ago by Seamus Ross in his keynote speech at the Budapest conference on digital libraries, the digital contents in any domain „require knowledge of its context of creation, and they demand evidence of its provenance. These are processes to which archives respond well because they have developed an appropriate theoretical framework and have operationalised it in repository design, management and use over at least three centuries. The archival framework meets requirements surrounding the production, management, selection, dissemination, preservation and curation needs of information. It also supports a layering of services from repository services at the foundation to user services at upper levels”³⁸. As been noted, this archival framework has been further developed and standardized in the last decade both with reference to the vocabulary and to the functional model by the theory of electronic recordkeeping and the standardization processes developed in this sector, thanks to the cooperation of a large international community based on ISO and ICA initiatives, has provided a significant contribution to support the collection of relevant information for integrity and identity of digital resources in the creation and

³⁶ Paola Carucci, *Il documento contemporaneo*, Roma: Carocci editore, 1987.

³⁷ See, among other, Luciana Duranti, Terry Eastwood, Heather MacNeil, “Preservation of the Integrity of Electronic Records”, Dordrecht: Kluwer Academic Publishers, 2002: Chapter 1; Luciana Duranti, Heather MacNeil, “The Preservation of the Integrity of Electronic Records: An Overview of the UBC-MAS Research Project”, in *Archivaria* 42 (1997), 2: 46–67; Heather MacNeil, *Trusting Records. Legal, Historical and Diplomatic Perspectives*. Kluwer Academic Publishers, Dordrecht, 2000.

³⁸ Seamus Ross, “Digital Preservation, Archival Science and Methodological Foundations for Digital Libraries” (keynote speech presented at ECDL Budapest 2007): 8, Accessed August 18, 2012, http://www.ecdl2007.org/Keynote_ECDL2007_SROSS.pdf.

preservation processes. In general, this effort (which is not necessarily linear, but on the contrary is very often conflicting and contradictory) is still operative and implies working groups at national and international level and thorough knowledge from each domain.

With specific reference to the authenticity, a rich archival literature has been developed, based on the InterPARES research and its *template for analysis*. Nevertheless, not many European projects have used this reference structure. The only exceptions have been CASPAR and now APARSEN³⁹ which have integrated, thanks to the presence of archivists involved in the research and of course with the support of experts of OAIS functional model, the template elements into OAIS. The other European projects, even those that had the authenticity management as part of their crucial requirements, have not proposed original solutions and have normally ignored the question and its definition. This lack of interest is not easy to understand and even more is difficult to justify, specifically because today terms such *trust* and *reliability*, at the centre of any research and any debate on digital preservation, acquire their intelligibility and resonance, if and when they can be concretely measured against the capacity to verify or, at least, presume the authenticity of the digital resources to be preserved. The same question is at the basis of any digital certification process of digital repository⁴⁰. The question of authenticity and how the research projects are or not normally able and/or willing to handle it can be emblematic of the present tendency to overcome the critical factors by ignoring them, instead of sustaining the effort to identify the nature of the problems, the complexity of the actions involved and the related responsibilities on the basis of the operational contexts.

The need for consistent and reliable vocabularies is again central (but of course cannot be handled as an accumulation of terms). They make the research results intelligible, measurable and exchangeable and allow for the continuity of the investments and the research work done. As already has been said, the archival community had a strongly commitment in this area. The promotion of a consistent archival terminology has been, not by chance, a key point of the initial program for ICA when launched by the director of US National Archives, Solon J. Buck, in 1946 and still is at the center of ICA attention.⁴¹

³⁹ This analysis has been discussed by the author in the deliverable D24.1 of APARSEN project: Silvio Salza, Mariella Guercio, Monica Grossi, Stefan Pröll, Christos Stroumboulis, Yannis Tzitzikas, Martin Doerr, Giorgos Flouris, "D24.1 Report on Authenticity and Plan for Interoperable Authenticity Evaluation System", April 2012, accessed August 18, 2012, http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2012/04/APARSEN-REP-D24_1-01-2_3.pdf.

⁴⁰ See on this aspect APARSEN Project: Deliverable 24.2. Implementation and testing of an authenticity protocol on a specific domain, 2012, Accessed August 22, 2012, http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2012/04/APARSEN-REP-D24_2-01-2_2.pdf.

⁴¹ The Buck roll call, titled *A Proposed Archives Program for the United Nations Educational, Scientific and Cultural Organization* (September 1946) is published as annex in the article of Oliver W. Holmes, *Towards an International Archives Program and Council, 1945-1950*, in «American Archivist», 39 (1976), 3: 296-299. See on this aspect Giulia Barrera, "Il villaggio globale degli archivisti. Organizzazioni internazionali e forme di cooperazione tra gli archivi del mondo", in *Archivistica*, Linda Giuva and Maria Guercio eds, Roma: Carocci (in press). The ICA has dedicated time and financial resources to develop and update its dictionaries. The first dictionary has been published in 1964: *Elsevier's lexicon of archive terminology: French, English, German, Spanish, Italian, Dutch*, Compiled and arranged on a systematic basis by a Committee of the International Council on archives. Amsterdam - London - New York : Elsevier, 1964. A second series of tools has been published in many languages: *Dictionary of Archival Terminology: English and French, with Equivalents in Dutch, German, Italian, Russian, and Spanish = Dictionnaire de terminologie archivistique*, edited by Peter Walne, 2nd rev. ed. (1 ed 1984) München - New York: K.G. Saur, 1988 (ICA handbooks series, 7). A third edition completely renewed with contextualized definitions developed in 2004 was only published on the web by Angelika Menne-Haritz, Accessed August 20, 2012, <http://www.staff.uni-marburg.de/~mennehar/datii/intro.htm>.

The future of the European action: new opportunities within the Common Strategy Framework

The European bodies are certainly aware of the limits and of the challenges still open at the present state of their investments and which are here only partially listed and discussed. One of the main qualities of their policies is the capacity to continuously update their high-level evaluation processes. More complex is the transformation of the intermediate decision level, specifically when the challenges involved are even difficult to be defined and to be designed for future strategies. An important (even if not decisive) step in this process was the EU workshop held in Luxembourg (Shape New Visions for EU-Research On Digital Preservation, May 4-5, 2011)⁴², with the participation of many European and North American researchers and experts in the field, invited for defining the new vision for EU research on digital preservation. The final report recognized that:

“There was not yet a corpus of best practice to refer to. These communities were not, however, well connected, and had not yet engaged widely beyond public memory institutions and research bodies. There also appeared to be lack of dialogue between digital preservation specialists and those who produce and curate data”.⁴³

The results of this workshop, by recognizing (even if partially) some of these needs, suggested:

- Shifting the attention from the preservation of data to the preservation of knowledge: this shift would be “be able to demonstrate the authenticity of preserved data”,
- Focusing more on the integration of digital preservation into lifecycle management of digital contents as proposed a decade ago by the record management and archival science scholars.

These two conclusions seem to confirm what was previously observed with reference to the areas where the contributions from the archival community could have been more than productive. The workshop provided many other opportunities for discussing all the open questions in the field, even if many aspects were not debated at all, specifically those of strategic value. For instance, the report in any way did not include, in the list of missing actions, the creation of reference centers or networks for the definition of rules, services, procedures, tools and also methods, even if similar ideas are present in projects like

⁴² “Shape New Visions for EU-Research On Digital Preservation”, Luxembourg, May 4-5, 2011, Accessed August 20, 2012, http://cordis.europa.eu/fp7/ict/telearn-digicult/digicult-future-digital-preservation_en.html.

⁴³ Clive S.G. Billness, “Report on the Proceedings of the Workshop The Future of the Past. Shaping New Visions for EU-Research On Digital Preservation”, organized by Cultural Heritage and Technology Enhanced Learning, Luxembourg, May 4-5, 2012. Accessed August 19, 2012, http://cordis.europa.eu/fp7/ict/telearn-digicult/future-of-the-past-summary_en.pdf. See also the report presented at the workshop, Stephan Strodl, Petar Petrov, Andreas Rauber, “Research on Digital Preservation within Projects co-funded by the European Union in the ICT Programme, 2011”, Accessed August 20, 2012, http://cordis.europa.eu/fp7/ict/telearn-digicult/report-research-digital-preservation_en.pdf. The report presents an overview of the past and present European research projects. It is interesting to note that it is recognized that in the first phase (Delos and ERPANET) a significant effort was developed - under the influence of the library and archival communities - for the “establishment of common problems, definitions, terminology and concepts” (p. 52), while in a second phase the attention was (and is) concentrated on non-traditional objects and on more technical aspects, with less attention to the development of common bases. By illustrating the outputs and topics of the funded European projects, the report makes evident, even if not explicitly, the fragmentation of the whole picture. It is useful to stress that only two archival experts were invited to the workshop and the awareness that this absence could have produced a negative impact on the definition of a more complete strategy was not at all expressed in that occasion.

APARSEN (Virtual Center of Competence)⁴⁴ or in PLANETS (Open Planets Foundation)⁴⁵. The eArchiving pilot proposal (see above) can also be interpreted as a first recognition of this weakness, in this case specifically referred to the archival domain. It will be the responsibility of the European archival institutions to transform this tepid credit opening into a real opportunity and an occasion for collecting ideas and proposals and creating the basis of an effective network or of a virtual center of archival competence.

Another important initiative, which should attract the future attention of the stakeholders, is the new European financial infrastructure proposed for the period 2014-2020, specifically with reference to the definition of a Common Strategy Framework for innovation and research and to the future plans included in the program called Horizon 2020⁴⁶. In relation to the aim of “Connecting Europe Facility (CEF) - cross-border infrastructures in energy, transport and ICT to strengthen the internal market” the new program defines the need to reinforce with adequate research resources “*competence centers on digitisation and preservation of digital cultural heritage*”. The European and international networks for cooperation are here indicated as strategic tools. These changes could provide new possibilities for overcoming the old critical difficulties but the challenges (which were not faced in the past as has already been indicated) will be even more complex in the near future and will require great capacity for cooperation, clear ideas and qualified alliances. The present is characterized both by the instability of the technological innovation and the complexity of the creation of digital contents, but also by the continuing growth in the development of digital data which is “outstripping the rate of growth in data storage technologies”. This uncertainty can be faced or at least limited with contextualized definitions of solid basic concepts and with a persistent coordination among qualified institutions, possibly characterized by common goals, even if the dynamic nature of the digital environment will never stop in provoking new conflicts and creating new uncertainties. The digital environment imposes (not only in case of digital heritage) undefined boundaries and *seems* to make irrelevant the traditional specificities, but implementations in the real world does not accept a high percentage of ambiguity as has been increasingly recognized in the recent debate on postmodernism decline.⁴⁷ Good methods, good concepts, consistent vocabulary cannot be provided when the knowledge is not well developed and openly discussed and if the boundaries are still unclear or even completely unknown. These challenges cannot be faced by any research sector involved in the digital preservation without a contemporary double investment both in strengthening its own domain knowledge and competences and in supporting adequate and qualified cooperation processes.

Of course, many of the questions here proposed do not concern only the European institutions and, naturally, are not confined to the archival world. These demanding and global challenges involve the whole international scientific community at large, specifically scholars and professionals involved in the research and in the implementation actions. For this reason, it is more than probable that our academic and professional communities will not waste any possible present and future opportunity for

⁴⁴ “Aparsen - Alliance for permanent access to the records of science network”, Accessed August 22, 2012, www.alliancepermanentaccess.org/current-projects/aparsen.

⁴⁵ “Open Planets Foundation”, Accessed August 22, 2012, www.openplanetsfoundation.org.

⁴⁶ “Elements for a Common Strategic Framework 2014 to 2020”, Accessed August 20, 2012, http://ec.europa.eu/regional_policy/what/future/index_en.cfm#2 and “Horizon 2020. The EU Framework Programme for Research and Innovation”, Accessed August 20, 2012, http://ec.europa.eu/research/horizon2020/index_en.cfm.

⁴⁷ E. Docx, Postmodernism is dead. In Prospect, 20 July 2011, <http://www.prospectmagazine.co.uk/magazine/postmodernism-is-dead-va-exhibition-age-of-authenticism>.

strengthening their cooperation. But to reinforce their will and their reasoning capacities, it is important to recognize that not all the efforts already made are in the right direction and that it is essential to recover, with a stronger will and more intellectual consistency, the precious scholarly assets accumulated in our scientific studies and technical practices.