

Project no. 269977

APARSEN
Alliance for Permanent Access to the Records of Science
Network

Instrument: Network of Excellence

Thematic Priority: ICT 6-4.1 – Digital Libraries and Digital Preservation

D43.1 SURVEY FOR THE
ASSESSMENT OF TRAINING
MATERIAL/ASSESSMENT OF DIGITAL
CURATION REQUIREMENTS

Document identifier:	APARSEN-REP-D43_1-01-4_1
Due Date:	2012-02-29
Submission Date:	2012-12-11
Work package:	WP43
Partners:	DPC, FRD
WP Lead Partner:	DPC
Document status	Released
URN	urn:nbn:de:101-20140516297

Abstract: This document reports on the preliminary assessment of training materials and the data curation requirements evidenced through research by the APARSEN Network of Excellence. It makes specific recommendations for the development of training courses.

Delivery Type Report**Author(s)** Chiara Cirinnà (FRD), Sharon McMeekin (DPC), William Kilbride (DPC)**Approval****Summary****Keyword List** Training Courses, Vocational Training**Availability** ☒ Public**Document Status Sheet**

Issue	Date	Comment	Author
V0.1	11-Nov-2011	Preliminary draft	Chiara Cirinnà
V1.0	15-Jan-2012	Complete draft	Chiara Cirinnà
V1.1	16-Feb-2012	Consistency checks	Chiara Cirinnà
V1.2	22-Feb-2012	Revisions, edits and additions	William Kilbride
V1.3	25-Feb-2012	Revisions, edits and additions	Sharon McMeekin
V1.4	27-Feb-2012	Revisions and additions (Ch 1 & 4)	William Kilbride
V1.5	28-Feb-2012	Revisions and additions (Ch 5 & 6)	Sharon McMeekin
V1.6	28-Feb-2012	Revisions (Ch 6)	William Kilbride
V1.7	29-Feb-2012	Embed reviewer's comments	Holger Brocks / William Kilbride
V1.8	29-Feb-2012	Embed reviewer's comments	Sabine Schrimpf / William Kilbride
V2.0	29-Feb-2012	Prepared for submission	William Kilbride
V2.1	29-Feb-2012	Embed Project Manager's comment	David Giaretta / William Kilbride
V2.2	16-Apr-2012	Revisions (Ch 4)	William Kilbride
V2.3	19-Apr-2012	Revisions and Additions	Sharon McMeekin
V2.4	20-Apr-2012	Top-to-bottom review	William Kilbride
V2.5	22-Apr-2012	Structural edit and new conclusion to Ch 3	William Kilbride
V2.6	23-Apr-2012	Reference checks	William Kilbride
V2.7	26-Apr-2012	Revisions and Additions (Ch 1 & 6)	Sharon McMeekin
V2.8	26-Apr-2012	Minor edits and proof check	William Kilbride
V2.9	27-Apr-2012	Incorporate review comments	David Giaretta, Holger Brocks, Sharon McMeekin
V3.0	29-Apr-2012	Minor format update	David Giaretta
V3.1	26-Oct-2012	Section 3.1 updated with extended survey results	Chiara Cirinnà

V3.2	4-Nov-2012	Revised Sections 2.2, 6.1 (including tables 8,9 and 10) and 7	William Kilbride
V3.3	19-Nov-2012	Included Appendix C	William Kilbride
V3.4	21-Nov-2012	Proof check and update ToC	William Kilbride
V3.5	05-Dec-2012	Additions to Ch 3 and Appendix B	Chiara Cirinnà
V3.6	06-Dec-2012	Additions to Ch 6, Appendix C and copy edit	Sharon McMeekin
V4.0	10-Dec-2012	Preparation for Submission	Sharon McMeekin
V4.1	10- Dec-2012	Checked for final submission	William Kilbride, Simon Lambert

Project information

Project acronym:	APARSEN
Project full title:	Alliance for Permanent Access to the Records of Science Network
Proposal/Contract no.:	269977

Project coordinator: Simon Lambert/David Giaretta

Address:	STFC, Rutherford Appleton Laboratory Chilton, Didcot, Oxon OX11 0QX, UK
Phone:	+44 1235 446235
Fax:	+44 1235 446362
Mobile:	+44 (0) 7770326304
E-mail:	simon.lambert@stfc.ac.uk / david.giaretta@stfc.ac.uk

CONTENT

EXECUTIVE SUMMARY	7
1 SETTING THE SCENE: DIGITAL PRESERVATION RESEARCH AND PRACTICE	8
1.1 UNDERSTANDING TRAINING NEEDS.....	8
1.2 EVOLVING DEFINITIONS OF AN EMERGING FIELD	10
1.3 DIGITAL PRESERVATION: DIVERSE SKILLS AND DIVERSE SECTORS.....	11
1.3.1 Research Communities.....	11
1.3.2 Archives, Libraries and Museums (ALM).....	12
1.3.3 Business and Industry.....	14
1.3.4 Government Agencies and the Public Sector.....	16
1.4 THE EVOLUTION OF NEW JOB PROFILES.....	17
1.5 CONCLUSIONS	19
2 EDUCATION AND TRAINING IN DIGITAL PRESERVATION.....	20
2.1 EUROPEAN POLICY FRAMEWORK FOR EDUCATION AND TRAINING IN DIGITAL PRESERVATION	20
2.1.1 The Bologna Process (1999 – present).....	20
2.1.2 Copenhagen Process (2002) - On Vocational Education and Training	22
2.1.3 Lisbon Strategy and Europe 2020	22
2.2 DIGITAL PRESERVATION TRAINING INITIATIVES	23
2.3 CONTENTS OF DIGITAL PRESERVATION CURRICULA	27
2.4 NEED TO IDENTIFY GAPS AND CHALLENGES	28
3 ANALYSIS OF CURRENT PROVISION IN DIGITAL PRESERVATION TRAINING.....	30
3.1 CURRENT PROVISION.....	30
3.1.1 Objectives of the research	30
3.1.2 Key Findings	32
3.1.3 Summary of the main findings of the analysis of current provision	40
3.2 PARTICIPANTS' VIEWS OF TRAINING	43
3.2.1 Evaluation of the Digital Preservation 'Road Shows' 2009-2011	43
3.2.2 Evaluation of DPC Expert Briefings 2009-12.....	47
3.3 DIGICURV FOCUS GROUPS ON TRAINING NEEDS	49
3.4 ANALYSIS OF CURRENT TRAINING PROVISION - CONCLUSIONS	51
4 PROFESSIONAL STANDARDS AND DIGITAL PRESERVATION	53
4.1 INTRODUCTION.....	53
4.2 STANDARDS IN DIGITAL PRESERVATION: AN HISTORICAL OVERVIEW.....	54
4.3 STANDARDS AND REPOSITORY OPERATIONS	57
4.4 INDUSTRY AND SECTOR SPECIFIC STANDARDS FOR DIGITAL PRESERVATION	58
4.5 DATA LIFECYCLES AND EXTRA-REPOSITORY STANDARDS	59
4.6 OTHER EMERGING STANDARDS AND TRAINING.....	61
4.7 PROFESSIONAL PRACTICE IN DIGITAL PRESERVATION: CONCLUSIONS FOR TRAINING	61
5 ANALYSIS OF REQUIRED DIGITAL PRESERVATION COMPETENCIES AND SKILLS (TRAINING REQUIREMENTS) AS HIGHLIGHTED BY PEER REVIEW	63
5.1 METHODOLOGY USED TO ANALYSE THE INTERVIEWS	64
5.2 ANALYSIS OF THE INTERVIEWS	65
5.2.1 Interviews with Reviewers	66
5.2.2 Interviews with the Repository Team.....	69
5.3 SKILLS, COMPETENCES AND TRAINING NEEDS.....	71
5.3.1 Identification of Senior Management Skills, Competences and Training Needs.....	71
5.3.2 Identification of Operational Management Skills, Competences and Training Needs	73
5.3.3 Identification of Operational Staff Skills, Competences and Training Needs	75
5.4 TRAINING NEEDS HIGHLIGHTED BY PEER REVIEW – CONCLUSIONS.....	76
6 IDENTIFIED TRAINING NEEDS AND LINKS TO APARSEN RESEARCH AND INTEGRATION ACTIVITIES	77

6.1	IDENTIFIED TOPICS AND RECOMMENDATIONS FOR TRAINING	77
6.2	POTENTIAL TRAINING PROVISION FROM APARSEN	82
7	CONCLUSIONS AND RECOMMENDATIONS.....	85
	REFERENCES	87
	ANNEX A: INTERVIEW STRUCTURES	92
	ANNEX B: LIST OF ANALYSED TRAINING INITIATIVES.....	96
	TRAINING COURSES	96
	WORKSHOPS	99
	TUTORIALS	104
	ONLINE COURSES.....	105
	ANNEX C – PROPOSED APARSEN TRAINING SCHEDULE 2013/14	107

EXECUTIVE SUMMARY

The aim of this report is to present the findings of the background research required by APARSEN work package 43 (WP43) “Training Courses” in order to set the objectives and strategies for the development of training courses for digital preservation practitioners within the Network of Excellence. The aim of the background research was to draw together a comprehensive picture of the digital preservation training that is currently available and to identify the training needs of practitioners working within the field. It is essential that the training developed in this work package meets the highest standards of content, delivery and relevance as this training provision will mature into a core component of the Virtual Centre of Excellence that will evolve out of APARSEN.

This report is Deliverable 43.1 “Survey of training material/ assessment of digital curation requirements.” and has five main sections providing context for and describing the results of the research undertaken. The first section offers some practical definitions and attempts to draw boundaries around the scope of the problem and the evolution of digital preservation as a professional practice. The second section describes the state of art: analysing the policy framework for training in digital preservation, introducing relevant initiatives and starting to identify the gaps that are apparent even at this high policy level. Section Three takes this into more detail. It analyses 134 training courses offered in the last 30 months in 14 different countries, and it looks at the impact of training as perceived by practitioners, drawing out the gaps in the current provision.

In the next two sections the report moves on to identify training requirements through the lens of particular activities and practical experiences from within the APARSEN Network of Excellence. These sections examine the increasingly important issues of emerging professional practice and repository standards. Section Four provides the context for this discussion by examining the most recent professional standards in preservation, thus informing training through desk-based analysis of the community’s formal definitions of good or best practice. Section Five then goes on to describe the results of interviews with experts who have undertaken audits recently as part of the development of audit and certification processes within the APARSEN work package “Peer Review and 3rd Party Certification of Repositories”. These interviews help to identify training requirements for digital preservation practitioners based on perceptions of the strengths and weaknesses of current practice.

Section Six brings together this research identifying the need for more advanced training on digital preservation as well as focused training on topics such as authenticity and provenance, annotation and data quality, preservation planning, sustainability and establishing trust. These topics are then related to APARSEN work packages to clearly identify how the research and integration activities can support training development. This section also includes a number of recommendations relating to the delivery of training courses that APARSEN will incorporate into the planning and development of the training programme. These recommendations include the need for more practical examples and exercises and for training targeted at specific sectors or professional roles.

Readers of the report may find it advantageous to concentrate on Section Six which most clearly articulates the findings of this research and makes specific recommendations for training based on those findings.

Therefore, the overall goal of this report is **to determine the main priorities framing and guiding training initiatives in digital preservation, providing recommendations for the training modules to be developed in the course of the APARSEN Network of Excellence.**

1 SETTING THE SCENE: DIGITAL PRESERVATION RESEARCH AND PRACTICE

1.1 UNDERSTANDING TRAINING NEEDS

This report seeks to document and understand training needs in digital preservation in order to inform decisions about the targeted investment in training which the APARSEN Network of Excellence will make. In order to better understand training needs, the authors have developed a number of inter-connected analyses based on a variety of desk-based research, survey and interviews. Clear recommendations for the training output of APARSEN follow from this.

The purpose of the APARSEN Network of Excellence is to reduce the emerging fragmentation that is present within the rapidly growing digital preservation community. It provides a forum and a vehicle for the exchange of ideas and is working towards the development of a common vision for digital preservation research. Ultimately this common vision, through the broad range of activities undertaken during the APARSEN project, will evolve into the foundations of a Virtual Centre of Excellence (VCoE) for digital preservation. Ultimately the VCoE will offer services ranging from a standardised vocabulary for digital preservation (see the beginnings of the Digital Preservation Glossary: <http://www.alliancepermanentaccess.org/index.php/knowledge-base/member-resources/digital-preservation-glossary/>) to consultancy on issues such as audit and certification. Training provision will be a key component of the services that will be offered by the VCoE and the main aim of WP43 will be to establish the core courses to be offered by developing training that is relevant, informative and high-quality.

The purpose of WP43 is to ‘define courses and collect, or create where necessary, training materials for digital preservation practitioners.... A number of workshops and events will be run during the lifetime of the project but the aim will be to make these self-funding by the end of APARSEN.’ There are also a range of other outputs from APARSEN which will be directly relevant to the development and delivery of training courses, including the common vision and ‘Virtual Centre of Excellence’ (WP11), the Communications Plan (D44.1), work on ‘Peer Review’ of repositories (WP33), and the co-ordination of standards (WP13). As will become clear, other APARSEN activities and outputs, including research on identifiers and citability (WP22), authenticity and provenance (WP24), storage solutions (WP23), research on authenticity and provenance (WP24), intelligibility (WP25) annotation (WP26), scalability (WP27) and many more will be able contribute content relevant to training.

WP43 has benefited from close relations with other work packages in the Integration and Spreading Excellence Streams that have similar and complementary goals and will continue to do so. This includes: collaboration, coordination and sharing of content and experience with WP15 (Internal Workshops, Symposia and Events) and WP41 (External Workshops, Symposia and Events); and alignment of messages with and use of the communications channels established by WP44 and WP45. Alignment with WP11 (Common Vision) will also be a key factor, ensuring that the training courses developed reflect the values and objectives of the Network of Excellence, and ultimately the VCoE. Finally there will be a close relationship with WP42 (Formal Qualifications) which, among other aims, is working towards the development of curricula for higher education courses and frameworks for continuing professional development (CPD). WP42 and WP43 have aligned their objectives to ensure that the training courses developed fit within the curricula and CPD frameworks. The two work packages will also collaborate on the development of a shared online resource center for training courses and materials.

It is worth noting that WP43 is not a research workpackage. This report should be understood as an analytical synthesis to support informed and practical action. There is no question that a more substantial research report in this field would be possible but it would be inefficient given the range of initiatives already surveying this domain such as the DigCurV project (see Engelhardt Strathmann and McCadden 2012, Karvelyte et al 2012). More importantly a more detailed analysis would reduce the effort available for focused delivery. D43.1 examines and extends existing fragmented research and it makes a specific and valuable contribution which is broader and more useful than would be possible in a project which was only about researching the problem. It provides recommendations which will inform APARSEN, enable the spreading of excellence, and which may be adopted by others outside the Consortium. These include specific recommendations on the range of topics which are needed, the preferred teaching methods and the needs of different audiences (see Section Six). The fundamental objective of the work package remains not research into training needs, but provision of training. This is in line with the objective for WP43, is congruent with the Description of Work, and is based on the assumption that the APARSEN seeks to make most efficient and effective use of existing research.

The preliminary assessment which D43.1 offers is necessary because, in the words of the description of work, there are ‘...an increasing amount of training materials being produced in the area of digital preservation, and there is much overlap. This task... foresees a preliminary survey aimed to identify and describe the instances of best practices, specific courses, training curriculum, and approaches used for continuing professional development in the communities addressed by the project. To design training courses effectively tailored on the user community requirements....’ In completing this work a range of institutions have generously shared their educational initiatives and, by remaining aware of the policy framework, the constraints of timetables and other restrictions, and the range of needs we are able to make practical and informed recommendations about the contents of new training resources which when delivered will equate to more than 30 hours of classroom teaching.

This opening section sets the scene and seeks to clarify definitional issues, noting the rapid evolution of digital preservation practice, the range of different terms and activities which imply or encompass digital preservation and the range of communities engaged. Section Two sets APARSEN’s plans for digital preservation training in a wider context of academic and vocational training in Europe and further afield. Section Three presents and scrutinizes the current provision of digital preservation training, firstly by surveying 134 courses offered in the last 30 months in 14 different countries, secondly by examining what participants think of particular elements of this provision, and thirdly by examining what expertise practitioners believe is required of colleagues and staff. Section Four examines training needs from the perspective of emerging professional standards and how these standards provide a conceptual framework for configuring professional practice. Section Five examines current professional practice in more immediate terms by reviewing a series of ‘test audits’ to derive lessons about strengths and weaknesses and in this way help to identify areas where training needs may be most acute. Each of these analyses is useful in its own terms but each is also partial: but by reviewing the state of the art through multiple perspectives, the report is able to provide a more nuanced and complete view of training needs. It enables broad, empirically-grounded recommendations to be made which can subsequently be implemented. Section Six then brings together the findings of the previous sections as part of an analysis of the outputs of APARSEN to identify how the Network of Excellence can develop training to best meet the needs identified. Final recommendations are presented in the report’s Conclusion.

Three appendices are included. Appendix A describes the methodology of the interviews used in section 5, Appendix B provides the extensive list of training courses analysed in Section 3. Appendix C shows provides the proposed APARSEN Training Schedule for 2013-14, illustrating how APARSEN is responding to this analysis in the immediate future

The main audience of this report is the APARSEN Network of Excellence, in particular the participants of WP43 and also WP42 on 'Formal Qualifications'. However the report is intended as a public deliverable and will be of interest to a range of parties, including those developing training materials (such as within other EC-funded digital preservation research as well as in the United States); practitioners with career development requirements in digital preservation and their managers; parallel and complementary initiatives in the development of academic qualifications for digital preservation; and those with an interest in capacity building in archives, libraries, records management and information technology.

1.2 EVOLVING DEFINITIONS OF AN EMERGING FIELD

The world is in the midst of a general-purpose technological revolution. Although this revolution has taken many names, there is little doubt that it is a technological revolution or a new techno-economic paradigm, brought about by a set of new Information and Communication Technologies (ICT). The past three decades have witnessed decisive advances in the production, management, distribution and use of digital resources. Impacts of ICT are most obvious in accessing information and knowledge, with dramatic increase in the power and speed to access, process, adapt and organize information. This has had significant and practical impacts on learning, innovation, knowledge creation and distribution.

The pervasiveness of digital technologies and their growing importance in social, economic, scientific and cultural transactions has led to the awareness that technology – and in particular data - needs to be carefully managed to ensure stability and continued access. A number of high profile examples have raised awareness about the material impact that results from data loss (see Waller and Sharpe 2006 8-9). Moreover, rapid changes in technologies, threats of obsolescence and the complex dependencies of digital resources compel us to more rapid intervention than would be required for analogue collections, especially for collections that are unique (e.g. Condrón et al 1999). Digital preservation has emerged as a new field concerned with ensuring the long-term viability of digital resources.

In the early days, digital preservation tended to emphasize the dependencies of digital material on the technology that was used to create and operate it. In this sense digital preservation was defined as “a broad range of activities designed to extend the usable life of machine-readable computer files and protect them from media failure, physical loss, and obsolescence” (Cornell University Library, 2003). This implied a relatively passive state, where material would be stored safely in inaccessible “dark archives” to ensure integrity and authenticity.

This perspective has evolved rapidly. More recently emphasis has shifted to more dynamic perspectives which enable accretion and the ongoing release of value – ensuring that digital material is managed throughout its lifecycle, remaining accessible and usable. Beagrie and Jones (2001) described digital preservation as “all of the actions required to maintain access to digital materials beyond the limits of media failure or technological change,” shifting emphasis onto access and use.

A range of other terms have also recently come into use which further emphasizes the management and preservation of digital material (Abbott 2008). For example Peter Burnhill at the launch of the *Digital Curation* Centre argued that “[curation...] brings together concerns about longevity - digital preservation - and added value activities -” (Burnhill, 2003). This term is often used loosely to encompass a number of streams of activity, including the production, use and dissemination of resources: “it encompasses a set of activities aiming at the production of high quality, dependable digital assets; their organization, archiving and long-term preservation; and the generation of added value from digital assets by means of resource-based knowledge elicitation” (Dallas et al. 2009). Buneman and Heiko (Buneman et al, 2009) use the term “curated database” to describe databases that are populated and updated with a great deal of human effort through the consultation, verification, and

aggregation of existing sources, and the interpretation of new raw data. Other terms, such as ‘digital continuity’ have also been used (The National Archives, 2011).

Although there is scope for nuanced debate about the meanings of these terms, and although other terms like ‘digital archiving’, ‘digital legacy’, ‘permanent access’ have also been used, there is a shared set of concerns about ensuring long-term access to data with all the relevant technical and managerial metadata that these may require. So for the purpose of this report, the term digital preservation, digital continuity and digital curation are used synonymously.

The APARSEN Network of Excellence has a stated goal to deliver a common vision and a conceptual model for digital preservation shared by all the sectors and partners represented in the consortium. Digital preservation is defined as “the act of maintaining information, independently understandable by a Designated Community, and with evidence supporting its authenticity, over the long-term” (OAIS definition¹).

If a more accessible definition is needed, then we propose the one offered by PARSE.Insight in 2009, “digital preservation denotes the process of storing digital information in such a way that it remains accessible, understandable and usable over the long-term (usually 5, 10, 50 or more years). This means that data needs to be specifically curated and enriched with extra information (metadata) to remain useful. For example: where did the data come from? How has it been stored? Which file formats have been used? Has the file been migrated and what might have changed? What special terminology or other information is needed to interpret and use the data?”

1.3 DIGITAL PRESERVATION: DIVERSE SKILLS AND DIVERSE SECTORS

In the last decade or so, digital preservation research agenda have included and prioritized all the digital object maintenance practices required for long-term access, starting from the pre-creation design and planning of resources, to provision of access over long time scales (Lee, 2009).

Digital preservation research agenda often emphasize the multi-disciplinarity of a field which is at the intersection of computer science, records management, library and information sciences, museology, economics, and social sciences (Ross and Hedstrom 2003). Consequently digital preservation is characterized by a close interaction of disciplines, which transform and inform each other. As industry and commerce have come to rely on digital resources, so digital preservation has come to engage a wide range of commercial stakeholders each keen to exploit the lasting value of the collections they own and manage. These diverse communities share a common set of challenges. More important for the purposes of this report, this creates a particular training challenge. The diversity of sectors and skills implies a diversity of skills, training needs and expectations.

1.3.1 Research Communities

Research communities are subject to a range of influences which create the need for preservation skills and training. For example, the *data types* used by researchers are evolving from static works of limited volume to dynamic multimedia resources, network-based data (web sites, e-mail, chat history), databases, 3D models, and software applications, which pose significant challenges in terms of

¹ ISO 14721:2003 - Space data and information transfer systems – Open archival information system – Reference model - OAIS

usability and understandability and personalization: complex digital objects contain core content and ancillary value-added features, interactive documents provide users with the opportunity to set preferences, dynamic objects are updated on an arbitrary basis.

Increased accessibility of scholarly outputs is promoted and occasionally mandated by open access policies and therefore institutional repository initiatives need to consider digital preservation issues in order for research outcomes to be sustainable and accessible over time. Accessibility allows for re-analysis of existing data, it may stimulate interdisciplinary collaborations and the advancement of science: according to the final survey report of the PARSE.Insight project, these are the most important drivers for the preservation of research data. But at the same time *new problems and challenges* related to the cross-disciplinary use of research data are emerging: sharing data through digital archives poses new legal issues and sometimes generates distrust in the capability of digital archives to properly handle research data; while the main *threats* are associated to access and use restrictions (e.g. Digital Rights Management) which may not be respected in the future, or the loss of ability to identify the location of data, and the perception of the risk that the ones we trust to look after the digital holdings may let us down.

To give a more detailed picture of the scenario, the PARSE.Insight report differentiated between different disciplines. It noted that most disciplines agree that the influence that the lack of sustainable hardware and software or support may have on preservation is considerable. The humanities researchers seem mostly concerned with the threat that future users may be unable to understand the data. Researchers from the agriculture & nutrition disciplines and medicine disciplines are most concerned with the loss of evidence due to uncertain origin and authenticity of the data. Sustainability is also a major concern among the researchers. Many—especially socio-cultural and social sciences researchers—consider the possibility that organizations or projects may cease to exist a major threat to the preservation of digital research data (PARSE.Insight, 2009).

A final policy consideration is particularly important: the responsibilities regarding long-term preservation are in many cases not clear. The effectiveness of data archiving and preservation policies vary to a large extent per research discipline and depend on the willingness of researchers and the availability of repositories. This perception urged the authors to affirm that in some research disciplines there seem not to be any repositories and the profession of data curator is simply *non-existent* (PARSE.Insight, 2010).

This analysis of Research Communities allows the identification of a number of clear training requirements for those within this sector who seek to engage with digital preservation. Although there will always be a need for generic training for those new to the issues there are also specific topics that will be of particular relevance to the sector. These topics can be summarized as: preservation planning and services (particularly for complex digital objects); representation information; information for use particularly in relation to discovery and access, annotation and data quality; accessibility; sustainability; policy and governance; and enabling and supporting novel uses of data.

1.3.2 Archives, Libraries and Museums (ALM)

Archives, libraries, museums and cognate institutions can be considered as a distinct group: preservation has always been part of their core mission. They are often called *memory institutions* (Dempsey et al., 2000) to put emphasis on their social role. They have a mandate to collect, preserve and communicate knowledge from one generation to another, primarily for research, education, cultural and legal purposes. Charged with giving access to and shaping cultural heritage, memory

institutions are sometimes characterized as storehouses, to be used for many different purposes, from education to commerce, from research to entertainment. These institutions have similar roles as part of a supportive educational and research infrastructure.

Despite these similarities, traditionally libraries, archives and museums have occupied different roles in social and informational space. The strategies they have adopted to interact with their users, and the organization and interpretation of their collections, differ and shape the definition of 'education' in and for each of these settings. In a change from the past, when customers were assumed to come to them with an implicit knowledge of an intrinsic value, most institutions are now looking at outreach and other ways to be more relevant to their communities and their customers' daily lives (O'Neill, 2006): at present all three types of institutions are moving towards providing access to their collections online. But peculiarities remain (Trant, 2009), at least in audience (user community), use of resources and professional roles and values.

Libraries are encyclopaedic collections that provide access to diverse knowledge for local, national, specialist or academic communities primarily from print collections. The audiences for public libraries are very broad, but the clientele of any particular library can be very diverse. A traditional library might be understood as a repository of published material that exists in multiple copies (most often printed), where catalogues and subject classifications are designed to identify relevant volumes and cluster similar works, thus enabling searching and browsing. Although roles are changing, the librarian would traditionally be seen as an enabler of the research process, with the task of acquiring and managing the collection, and specifically supporting use, especially in the discovery stage by helping users in identifying the relevant works, finding specific volumes or providing some kind of access to materials. Research in a library takes place in a largely unmediated manner.

Archives might typically be described as providing the long-term memory of larger agencies such as companies, universities, hospitals, schools, courts or entire jurisdictions like nations or territorial governments. Users of archives generally have specific problems or use in mind (for example to identify the facts surrounding a specific transaction). If done for historical purposes, researchers approach an archival collection to 'browse' within the confines of a known collection. Fonds are tied directly to organizational contexts, as recognized in the principle of provenance. They are maintained separately in the order given by their creators. Archival materials constitute the unique evidence of the transactions of organizations and institutions; documents do not circulate as books do in a library. To access an archive encompasses often a preliminary consultation with an archivist to help identify relevant collections and to establish where in those fonds specific types of material may exist. Many archival collections belonging to corporations or institutions remain in the custody of the originating institution to be used (sometimes exclusively) by staff while large public or academic archives support a broad range of professional and public researchers.

Museums might be described as subject-based collections of exceptional or noteworthy objects or specimens. The majority of museum visitors seek cultural, aesthetic, educational and entertainment value from the collection while only a relatively small number visit for research purposes. Unique artefacts, usually characterized by preciousness and rarity, need to be first preserved and protected, and secondly interpreted and presented. Usually a visit to a museum collection is a highly mediated experience: visitors are strongly guided through a highly selected fraction of the collection. The exhibition space is assembled according to a curatorial narrative and sequenced to support an argument or illustrate a theme. By and large only a small fraction of the collection is ever exhibited and the public is seldom able to access the majority of the collection. Within a gallery, didactic educational materials provide context and meaning for the works chosen, offering an interpretation and explanation for the visitor. Labels both identify artefacts or specimens and explain their relevance to a

particular context. Museum educators encourage encounters with artefacts, times, periods and places previously unknown to the visitor, or to raise themes and topics which have particular resonance with the visitors. Engagement with the collection and its interpretation are generally encouraged and it is not unusual for museums to fabricate surrogates of collection items in order to facilitate object handling by large numbers of visitors. Museum information resources are created to support the study and interpretation of the collection and entire facilities have been constructed in recent years to support sophisticated and extended access to entire collections. Professionals taking care of museum collections are cataloguers, curators, or subject specialists. Large museums often have both libraries (to support research) and archives which document institutional history, trace the provenance and conservation of objects and provide support to interpretation.

With the advent of the digital age, ALM have been re-thinking their services and revising their user models. The digital representation of objects might be different in relation to the final user needs and discipline, metadata may vary according to the different requirements and standards adopted. New challenges arise around the concepts of objects authenticity, provenance, and integrity. Digital asset management and preservation are therefore a common concern for libraries, archives and museums.

As with Research Communities it is possible to extract from this analysis specific digital preservation training needs for this sector beyond the requirement for general introductory training. Topics that would be of interest to representatives from the Archives, Libraries and Museums sector would likely include: policy and governance; legislation and regulatory frameworks; digital rights management; representation information (discovery and access); user analysis (including identification and analysis of the designated community); preservation planning and services; and provenance and authenticity.

1.3.3 Business and Industry

Commerce and industry derive enormous economic advantage from their use of digital resources. Although they do not share the strategic requirement to preserve associated with memory institutions, the business interests associated with creation and dissemination of digital resources mean that there are significant, sector-specific inducements for business and industry to manage its digital collections for the long-term. These might be categorized into four overlapping groups: businesses which exploit intellectual property through publishing and distribution; businesses with a regulatory requirement to ensure long-term access; businesses that manage, facilitate or underwrite long-term legal and contractual undertakings; and business that provide digital preservation solutions.

The first group is typically represented by publishers, but also includes design engineering, the music industry, film and broadcast media. They have significant amounts of intellectual property from which they derive on-going revenues and royalties. The complex supply chains in these industries mean that revenues may in fact be distributed to a large number of stakeholders who each therefore have a vested interest in preservation. This brings with it the opportunity to sustain flows of revenue for the long-term, and the intrinsic reasons why this sector – a significant part of the economy – has a strategic reason to engage with digital preservation.

Scholarly publishers are a particular example of this business requirement for preservation. Again the PARSE.Insight survey report gives quite a clear picture of how the digital preservation topic is perceived by scholarly publishers and how their role could evolve. Research publication, and in particular the research journal has changed considerably in the last decade, bearing multimedia, data-rich formats: and the way research articles are created, published and accessed is rapidly changing. This has an impact on the stages a manuscript passes as well as the versions of a manuscript/article

that are created. Also the awareness of the role supported by scholarly publishers has also emerged in the PARSE.Insight survey (2009): the majority of scholarly publishers (large and small) see themselves as the first group to carry the responsibility for the preservation of journals, followed by National Libraries and third party service providers like Portico and CLOCKSS. PARSE.Insight suggested that scholarly publishers want research articles to be preserved alongside illustrative materials (sound, images, and videos) as an integral part of the research. Around the half of those surveyed thought that research data should be preserved, even if they think responsibility lies generally with the researchers or their institutes. It is interesting to note that in the Brussels Declaration, STM publishers expressed their belief that “raw research data should be made freely available to all researchers. Publishers encourage the public posting of the raw data outputs of research. Sets or sub-sets of data that are submitted with a paper to a journal should wherever possible be made freely accessible to other scholars”². When publishers have been asked about the main threats to digital preservation, the majority reported to fear the sustainability of data when the current custodian of the data ceases to exist in the future. This opens to further reflections about the growth of their role in the near future.

The second group of industrial or commercial interests does not come to preservation for the intrinsic value of the collections but because of an external need to ensure compliance with legal and other regulatory requirements. Stereotypically this includes sectors like aerospace, manufacturing, banking, private health and welfare providers, civil engineering, nuclear power and processing and environmental management. The regulations for these sectors mean that, with varying degrees of quality and different time-frames, companies within these industries may be required to provide access or deposit records over the long-term. They are unlikely to make profit from the records directly but it could be costly in terms of reputation and cash-flow if they did not show some interest in digital preservation. A number of industrial sectors straddle the first two groups: for example pharmaceutical companies and geo-prospection companies have a variety of regulatory requirements for preservation, but data also presents an advantage for future exploitation.

A third and less visible commercial interest in preservation can be found in those agencies that manage long-term contracts on behalf of others. Legal companies have for many years offered a range of escrow services to support contracts. The principle behind escrow is that a legal third party is written into a contract to look after a resource in case a service provider is unable to supply a service any longer: if the service provider is liquidated the users can turn to the escrow service instead. This principle underpins a range of ‘post-cancellation access’ services in publishing, but is not limited to this and software escrow is now also common. For escrow to be effective it requires the resource to be viable once released, and this in turn creates a requirement for the escrow service to engage in digital preservation.

A fourth sector completes the commercial interest in digital preservation. A range of commercial operations have entered the market in recent years precisely to offer digital preservation services, whether as software vendors (like Tessella or Ex Libris) or as service providers (like Hanzo or the Internet Archive). The business case of such agencies depends on their ability to demonstrate competence in digital preservation. The TIMBUS project is currently trying to estimate the size of the digital preservation market-place in which these companies operate.

Generic introductory training would likely be of use to those engaging with digital preservation within Business and Industry but it would also be easy to conclude that the most useful training would be more closely tailored to a particular industry’s standards, common workflows and regulatory

² http://www.stm-assoc.org/2007_11_01_Brussels_Declaration.pdf

framework. From this analysis it is also possible to extract the suggestion of training needs relating to topics such as: digital rights management; legislation and regulatory frameworks; preservation planning and services (with a particular focus on complex objects); sustainability; representation information, information to support use such as annotation and data quality; and enabling novel uses of data to obtain value.

1.3.4 Government Agencies and the Public Sector

Public bodies hold a variety of materials and contents, including demographic, economic, cultural and historical data. The opportunity to document public sector information and preserve these data in the long-term has increased in recent years, improving security, social relationships, and institutions' and citizens' rights. This development, both in qualitative and quantitative terms, has proved possible with the rapid progress in ICT. The public sector is a big producer and collector of a large variety of data/information and content. Moreover, public content can be more widely spread to reap important economic and social benefits, promoting transparency and enabling new kinds of enterprise.

Combining these resources with ICT capabilities allows the realisation of a variety of innovative services and products. In (OECD, 2006), two main technological developments have been declared responsible for having changed and shaped the role of public sector information and content: i) technologies that enable the digitisation of public resources and ii) the deployment of broadband technologies as a means of rapid dissemination.

These developments are crucial for the public sector content diffusion. Once digitized, information and content become manageable and transportable. Moreover IT applications help increase transparency in public administrations, thanks to the definition of short timeframes in developing practices and to the opportunities given to citizens to directly verify the state of their practices. The innovations of content digitisation and Internet dissemination bring along new opportunities but also challenges for the public sector regarding information management and content preservation. In fact, government agencies are often legally obliged to keep records over decades for purposes of accountability, continuity of operations and organizational memory. All of these institutions have a vital interest in affordable preservation methods and systems to protect their core institutional records. Even if such records are not permanently archived, many have to be kept long enough (often more than 50 years) for there to be concern about the impacts of changing technologies. Therefore a topic of extreme relevance is preservation of authentic records. Management of digital records produced by public bodies poses challenges of preserving authentic and trustworthy information.

As with the previous three sectors there is no doubt that introductory digital preservation would be useful to practitioners within this sector but there are a number of other key topics for Government Agencies and the Public Sector that could be covered by more focused training. This would include: policy and governance; preservation planning and services; authenticity; establishing trust, and understanding the value of data for novel uses.

This brief consideration of the different sectors involved in digital preservation illustrates three themes which need to be born in mind in the development of training materials. Firstly, as we have seen there are different reasons why agencies are interested in preservation. Sectors have different expectations of what they may need to achieve and so will have diverse needs from training. Secondly, sectors are at different levels of maturity in responding to the challenge. In some cases it may be necessary to pitch courses at an introductory level and for others it will be possible to present more advanced materials. In other cases – such as scholarly publishing – the supply chains may be so evolved that

specific elements of the digital preservation challenge may be relevant. Thirdly, the extent of engagement with digital preservation is changing quickly. To be most effective, the digital preservation community – and the training it provides – needs to respond adroitly to these sector specific requirements.

1.4 THE EVOLUTION OF NEW JOB PROFILES

The last two decades have also seen rapid changes in the workforce. On one hand, developments of digital applications have accelerated the processes of fragmentation of existing professional roles in libraries, archives and museums as well as within academic research. On the other hand, individuals in different domains have increasingly come to recognize that there are a common set of challenges and opportunities that they share: the growth in digital objects requiring long-term preservation is increasing the demand for new skills and competences and the emergence of new jobs. This is important for digital preservation training as it implies that there may, or indeed should, be a range of job titles and professional skills already available in any digital preservation class room. To a great extent the task of training is about training staff who are already experts – and in many cases leaders - in their own domains.

For example, in 2005 the US National Science Foundation's National Science Board [NSF, 2005] called for the creation of “data scientists”, “the information and computer scientists, database and software engineers and programmers, disciplinary experts, curators and expert annotators, librarians, archivists, and others, who are crucial to the successful management of a digital data collection”. This definition makes evident the complexity of emerging professional profiles that embrace people with pure research roles, computer scientists or members of the library community. It implies a breakdown in the traditional boundaries between the library, archive and research lab.

A parallel message was articulated in 2006 by Sharpe and Waller who reflected that a whole professional infrastructure has been built up around the preservation of paper documents whereas digital preservation is very much in its infancy. They concluded that ‘the new discipline of digital preservation needs to be supported’ and called for digital preservation training to form part of the professional training for conservators, archivists and librarians (Sharpe and Waller 2006, 31). In 2009, the Digital Preservation Coalition which commissioned this research explicitly included ‘workforce development’ as one of the key objectives of its strategic plan (DPC 2009) and established a grant programme to facilitate access to training.

In 2008 a report to JISC (Swan and Brown, 2008) articulated and clarified different roles in preservation – at least as far as academic research is concerned – following a data life cycle approach. They identified four roles: data creator, data scientist, data manager, and data librarians:

- **Data creators or data authors:** are researchers with domain expertise who produce data. They have a high level of expertise in handling, manipulating and using data, gained through experience and as a result of need or personal interest.

- **Data scientists:** work where the research is carried out – or, in the case of data centre personnel, in close collaboration with the creators of the data – and conduct a number of functions including, in many cases, being data creators too. In origin and training they may be domain experts, computer scientists or information technologists and their career development may have required them to assimilate skills from a discipline from which they did not originate. Some data scientists report that an important part of their role is to be a ‘translator’, communicating the needs of the data creators to

data managers and working with the data managers to ensure that data are stored and accessible in a usable way.

- **Data managers:** are the computer scientists, information technologists or information scientists responsible for the computing facilities, storage, continuing access and preservation of data. They liaise with data scientists, ensuring that the right technological facilities are available. Some data managers describe their role as data ‘plumber’, piping data from one place to another, ensuring data flows operate properly and that valuable data are not lost.

- **Data librarians:** are trained and specialize in the curation, preservation and archiving of data. Originally, the term data librarian seemed to be confined to librarians dealing with social science data, but the title now encompasses people with data skills in all disciplines. It is a particularly important area as institutions begin to develop digital repositories for the collection and curation of their research outputs. Datasets are part of those outputs, an institutional repository is a natural home for them and the repository is usually in the care of the library.

Swan and Brown also remark that these terms are not used precisely within job descriptions. People whom they would define as data scientists are in jobs called data manager or data specialist. They note that also the boundaries between these roles are currently uncertain, but the classification remains important because it outlines the roles involved in almost all the sectors defined above.

This question of digital preservation skills is not just an academic one however. In 2010 a research by the consulting firm Gartner predicted the rapid emergence of the role of ‘Digital Archivist’, estimating 15% of businesses would employ someone in this role by 2012 (in 2009 it was 1%). Moreover it was reported that “organizations need staff with different skills from the ones they were originally hired for. These are not IT people as organizations know them. Staying relevant in this changing environment will require a new way of thinking about organizational models and staffing in IT projects”. (Gartner, 2010)

The European Commission has included this issue among the objectives of Europe 2020 the *Agenda for New Skills and Jobs*, where the urgency of defining vocational and educational training to meet present and future needs for new professional profiles in Europe is stressed, in order to adapt to changing labour market needs (European Commission 2010).

These are only a few examples of recent reports that have proposed new kinds of professional roles with digital preservation competencies in the workforce. The pace of change is challenging both to institutions, that are experiencing difficulties in recruiting individuals with appropriate skills and competences, and to hired staff that need vocational, in situ training in digital preservation in order to manage the transition to digital collections management. The pace of change also presents a challenge to training institutions in terms of keeping pace to fill the need.

Although there are a handful of accredited graduate-level academic programs and professional training and education programs available there is still a debate on how to address many challenges. Some of the most important are related to the wide range of activities that must be performed and to the different contexts that are subject to digital preservation.

The topic of ‘convergence’ has been subject of much discussion, especially in the museum, archive and library communities, because of the emerging similarities between these types of cultural heritage

institutions, most evident today in their on-line activities. But very often this term has been used more to indicate a practical and managerial cooperation, rather than real common practices, interdisciplinary curricular contents, common and shared knowledge.

The development of a professional community is strictly connected with the quality of its educational and training infrastructure: if it is adequately developed and sufficiently open to change, it will be able to play a relevant role to ensure a qualified level of professionals in diverse contexts. Work within APARSEN WP42 (Formal Qualifications) will help advance this development as it contributes to the creation of recommended curricula for graduate-level courses and continuing professional development.

1.5 CONCLUSIONS

Many if not all sectors have undergone significant changes in the last two decades as a consequence of and in response to the remarkable increase in accessing, processing, and organizing information. This has had a profound effect on learning, innovation, knowledge creation and distribution.

The need for preservation of key digital resources has been recognized in a large number of sectors even if the reasons for preservation vary from sector to sector. However capacity – the skills needed to provide this are not widely dispersed. Information professionals from diverse communities are increasingly responsible for taking care of digital objects across their full life: the activities required to fulfil these responsibilities can occur within various professional, disciplinary, institutional, or organizational contexts. All the while the digital domain continues to expand in scale, complexity and importance. Job profiles are rapidly changing and almost every sector is affected by this.

The development of a professional cadre of staff skilled in preservation is closely connected with the quality of the educational and training infrastructure which the community can make available: if it is adequately developed and sufficiently open to change, it will be able to play a relevant role to ensure a qualified level of professionals in diverse contexts.

As we shall see, there have been considerable efforts to provide training. Even so, the rapid pace of research has made it hard for training and best practice to remain connected. However, we can now begin to understand in more detail the impact which this training has had, both in the words of the participants and their managers, and the in the quality of the preservation services which this training has helped to form. By examining the impact of the current provision of training alongside the quality of the current provision of preservation services we can more closely model the emerging training requirements of a rapidly growing community and in this way concentrate our own efforts on delivering the most cost-effective and most impactful training resources.

2 EDUCATION AND TRAINING IN DIGITAL PRESERVATION

The aim of this section is to provide context for the development of training within APARSEN by describing the landscape within which it will exist. This will include descriptions of relevant European Policy Frameworks, other digital preservation training initiatives, requirements for digital preservation training in relation to continuing professional development, and issues to be considered when identifying needs and gaps that exist in the current training infrastructure.

It should be remembered that APARSEN is examining training from two complementary perspectives. In Work Package 43 the primary concern is to sketch out and meet the need for vocational courses which in turn can draw upon and influence formal qualifications. Work Package 42 (Formal Qualifications) is addressing the inclusion of Digital Preservation theory and practice into higher education curricula and the development and delivery of a portfolio of courses and e-learning material. By considering the current state of training at a policy level this section sketches out the context for the recommendations that follow. But we also make a practical contribution to the development of Formal Qualifications and agreement about how such a curriculum may need to be constituted.

2.1 EUROPEAN POLICY FRAMEWORK FOR EDUCATION AND TRAINING IN DIGITAL PRESERVATION

Before analysing European digital preservation training, it is necessary to examine conditions for higher education and vocational training to understand the context in which it must sit. A clear definition about the difference between higher education and vocational training has already been provided by the Digital Preservation Europe (DPE):

“The first is higher education – formal studies awarded by academic degree that are provided by universities and/or other higher education schools. The second is vocational training – studies of different degree of formality that are aimed at provision of knowledge and skills necessary to apply in professional practice. Vocational training can be provided by different players, including universities, colleges, professional associations etc.” (DPE, 2006)

Both types of training are, however, inherently linked, with one often leading on from the other, and thus an understanding of both contexts is essential when meeting training needs. The following sub-sections provide the background information required to understand the importance of and connections between these contexts.

2.1.1 The Bologna Process (1999 – present)

Some of the most important changes in European higher education of the last decade have been brought about in result of the Bologna Declaration. Within this declaration it is recommended that the educational system in Europe should be harmonized on the basis of the students’ mobility:

“Promotion of mobility by overcoming obstacles to the effective exercise of free movement with particular attention to:
- for students, access to study and training opportunities and to related services

- for teachers, researchers and administrative staff, recognition and valorisation of periods spent in a European context researching, teaching and training, without prejudicing their statutory rights.”³

As a result of this call Ministers announced the creation of two-cycle degrees at European universities. Consequently the standardization of structure and dissemination of credits as well as the harmonization of teaching methods in higher education was initiated.

Two years after this constitutive announcement the Bologna Process continued in Prague. In the Communiqué of this second Bologna Process meeting, the Ministers encouraged the establishment of Lifelong Learning as an ‘essential element’ of higher education.⁴

In 2003, when the European Ministers of Education met again in Berlin, the importance of the training of young researchers was highlighted:

“[...] Ministers consider it necessary to go beyond the present focus on two main cycles of higher education to include the doctoral level as third cycle in the Bologna Process. They emphasize the importance of research and research training and the promotion of interdisciplinarity in maintaining and improving the quality of higher education and in enhancing the competitiveness of European higher education more generally. Ministers call for increased mobility at the doctoral and postdoctoral levels and encourage the institutions concerned to increase their cooperation in doctoral studies and the training of young researchers.”⁵

With the Bologna Process in Bergen in 2005 new goals in European higher education were approved and the need to develop the professional skills was stressed. Specifically mentioned is the priority to enhance lifelong learning and vocational training in Europe:

“We underline the importance of ensuring complementarity between the overarching framework for the EHEA and the proposed broader framework for qualifications for lifelong learning encompassing general education as well as vocational education and training as now being developed within the European Union as well as among participating countries.”⁶

The importance of flexible vocational training paths was pointed out with the Bologna Process in Leuven 2009:

³ The Bologna Declaration of June 1999 ‘Joint declaration of the European Ministers of Education’

⁴ Towards the European higher education area ‘Communique of the meeting of European Ministers in charge of Higher Education’

⁵ Communique of the Conference of Ministers responsible for Higher Education in Berlin on 19 September 2003 ‘Realising the European Higher Education Area’

⁶ Communique of the Conference of Ministers Responsible for Higher Education, Bergen, 19-20 May 2005 ‘The European Higher Education Area – Achieving the Goals’

“Widening participation shall also be achieved through lifelong learning as an integral part of our education systems. Lifelong learning is subject to the principle of public responsibility. The accessibility, quality of provision and transparency of information shall be assured. Lifelong learning involves obtaining qualifications, extending knowledge and understanding, gaining new skills and competences or enriching personal growth. Lifelong learning implies that qualifications may be obtained through flexible learning paths, including part-time studies, as well as work-based routes.”⁷

In the last twelve years the Bologna Process has prepared the ground for vocational training and the issues identified should be considered when developing any kind of training. The international corporation of the Ministers of Education sets the basic requirements for a transferable credit system, mobility in study and training and flexible learning structures for vocational training.

Overall the Bologna process does place more stress on issues relating to academic education, but the issues mentioned above are important when developing vocational training. More crucial for the vocational training area are the Copenhagen Process and the Lisbon Strategy which will be discussed in the next two sub-sections.

2.1.2 Copenhagen Process (2002) - On Vocational Education and Training

The Copenhagen Process brought together the European Ministers of Vocational Education and Training as well as the European Commission to strengthen the alignment of the higher education area for professional and marketplace skills.

“The ministers responsible for vocational education and training and the European Commission have confirmed the necessity to undertake the objectives and priorities for actions set out in this declaration and to participate in the framework for an enhanced cooperation in vocational education and training, including the social partners. A meeting in two years’ time will be held to review progress and give advice on priorities and strategies.”⁸

The DPE (DigitalPreservationEurope) summarized the two main points in their ‘Outline of Training Principles and Objectives’ (DPE, 2006). First the programme activity of the Copenhagen Process provides the focus for the “creation of a common space for vocational education and training in Europe” and in a second step it defines the “skills and competencies adequate for [the] contemporary labour market and set[s] objectives to raise the profile of vocational education and training, quality and transparency of educational provisions and promote European co-operation in this area”.

2.1.3 Lisbon Strategy and Europe 2020

The subsequent programme of the successful Lisbon Strategy for Growth and Jobs is called Europe 2020. It was defined in 2010 as the continuation of the Europe 2010 programme establishing a vision for vocational education and training:

⁷ Leuven, 2009

⁸ Declaration of the European Ministers of Vocational Education and Training, and the European Commission, http://ec.europa.eu/education/pdf/doc125_en.pdf

“At EU level, the Commission will work:

- [...] To give a strong impetus to the strategic framework for cooperation in education and training involving all stakeholders. This should notably result in the implementation of life-long learning principles (in cooperation with Member States, social partners, experts) including through flexible learning pathways between different education and training sectors and levels and by reinforcing the attractiveness of vocational education and training. Social partners at European level should be consulted in view of developing an initiative of their own in this area;

- [...] To ensure that the competences required to engage in further learning and the labour market are acquired and recognized throughout general, vocational, higher and adult education and to develop a common language and operational tool for education/training and work: a European Skills, Competences and Occupations framework (ESCO).

[...] At national level, Member States will need:

- [...] To ensure that the competences required to engage in further learning and the labour market are acquired and recognized throughout general, vocational, higher and adult education, including non-formal and informal learning.”⁹

This strategy, along with the Bologna and Copenhagen Processes, present a number of issues that must be considered when developing training in terms how the training fits within an individual’s continuing professional development, particularly the links between such courses and higher education, its relation to established competencies within the relevant field and the appropriate modes of delivery for such training.

2.2 DIGITAL PRESERVATION TRAINING INITIATIVES

Higher Education infrastructure in Europe, supported with high-level political engagement will benefit from these changes. Although they only represent a small fraction of the total vision, they are likely to have a specific and beneficial impact on vocational training in digital preservation. Along with the many initiatives supporting and establishing digital preservation within organizations, the needs of the digital preservation workforce are growing. This is particularly demonstrated by the establishment of organizations with special interest in digital preservation. As early as 2002 the Digital Preservation Coalition in the UK was founded with an interest in this area and in 2005 the first International Digital Curation Conference was held and the first Digital Preservation Training Programme was delivered. Since then number of projects developing training for digital preservation has grown substantially.

Early European projects in digital preservation – like ERPANET – included specific training activities. More recently, the 6th Framework Programme (2006-2009) of the European Commission, tools and methodologies for digital preservation were developed among projects like Digital Preservation Europe (DPE) project¹⁰, the Cultural, Artistic and Scientific knowledge for Preservation, Access and Retrieval (CASPAR) project¹¹, the Preservation and Long-term Access through NETworked Services (PLANETS) project¹² and the Securing Heritage Access through Multivalent Archiving project (SHAMAN)¹³. The DPE project presented a report on the main priorities in framing and guiding

⁹ http://ec.europa.eu/archives/growthandjobs_2009/pdf/complet_en.pdf

¹⁰ <http://www.digitalpreservationeurope.eu>

¹¹ <http://www.casparpreserves.eu>

¹² <http://www.planets-project.eu>

¹³ <http://www.shaman-ip.eu/> but see also <http://www.digitalpreservationeurope.eu/preservation->

education and training initiatives in digital preservation including recommendations on course and curricula design in both university and continuing professional development settings (DPE, 2006). A range of activities were brought together under the 'We Preserve' umbrella to help maximize impact. (For a more detailed analysis of research on digital preservation within EC funded and co-funded projects see Strodl Petron and Rauber 2011)

As well having a specific commitment to training about their own tools and services these projects have also provided informal user testing and which therefore have a feedback into the development of tools also. They also illustrate the evolution of training materials organisationally and in terms of their subject matter. Two particularly significant examples of this are the PLANETS project and the KEEP project.

The primary goal of PLANETS was to build practical services and tools to help ensure long-term access to digital cultural and scientific assets. The project hosted a series of training events called 'Digital Preservation the PLANETS Way' across Europe between 2009 and 2010 and a range of training materials were produced. Although these focussed on PLANETS tools and products, the breadth of the project's goals means that these provide an overview of issues of general relevance, such as preservation planning and validation as well as fundamental issues of file management and obsolescence. PLANETS training materials remain available¹⁴, and the project's evolution into the not-for-profit Open Planets Foundation means that the community which provided them remains active with a series of events and useful resources such as webinars and wikis which supplement the PLANETS training materials¹⁵. Moreover the Open Planets Foundation is active in the EC-funded SCAPE project and is therefore hosting a new series of training events on the scalable preservation services. Again these are focussed on the SCAPE outcomes but the issues they address – like file format identification – are generalizable. Therefore, the PLANETS project has created training materials of general utility, and these initial resources are embedded and sustained through an ongoing series of initiatives like OPF and SCAPE which have a training function of their own.

A second example of good-practice in training from an EC-funded project is the 'Keeping Emulation Environments Portable (KEEP) Project. KEEP is more recent than PLANETS and took a different approach to preservation by developing emulation services to enable accurate rendering of both static and dynamic digital objects: text, sound, and image files; multimedia documents, websites, databases and videogames. KEEP organised a range of training events across Europe in 2011 and 12 and the resulting training materials remain available¹⁶. As with PLANETS, the training material highlights the project's own outputs but is relevant across a range of different use cases so they are useful even for those not directly interested in KEEP technologies. Because the KEEP project had a special interest in computer games, they are particularly useful for issues of software preservation.

A final example of good practice in digital preservation training from an EC-funded activity can be found in the PRESTO series of projects¹⁷. PRESTO is important in its own right as a targeted series of actions on the preservation of audiovisual content, and thus it has a direct connection to the core themes of APARSEN. PRESTO has also been through a series of iterations. The first PRESTO project was co-ordinated by the BBC as early as 2002. A series of substantive projects followed which

training-materials/files/shaman.pdf

¹⁴ <http://www.planets-project.eu/training-materials/>

¹⁵ <http://openplanetsfoundation.org/>

¹⁶ <http://www.keep-project.eu/ezpub2/index.php?eng/KEEP-Training-Materials>

¹⁷ <http://www.prestoprime.org/training/index.en.html>

acted on the needs identified in the initial survey - PrestoSpace and PrestoPrime. Training materials were made available under PrestoPrime and all of the project outcomes are now embodied as the PrestoCentre Foundation, a multi-national, multi-stakeholder agency which includes training as part of its appeal to members. So as well as providing relevant specialist training, the PRESTO series of projects provide a model for the evolution of APARSEN and the role that training may play in a virtual centre of excellence.

A more recent batch of research grants from the EC means that others are now active in this field and that the range of training materials from preservation research will continue to grow. As well as the SCAPE project mentioned above, the ENSURE Project¹⁸ which is examining commercial use cases in preservation is scheduled to offer its first training in 2013; the TIMBUS Project¹⁹ which is examining the preservation of business processes is hosting its first full training event at the time of writing; the SCIDEP-ES Project²⁰, an Earth Sciences data preservation project will provide online and face to face training which it hopes to make relevant to many sectors. These are simply a selection to illustrate the point that EC-funded research in digital preservation has made a significant amount of training material available and it will continue to do so during the lifetime of APARSEN.

At a more local and national level initiatives like Nestor and NCDD have provided fora for knowledge exchange and training among a growing number of staff engaged in preservation, while a number of degree programmes have started to offer digital preservation within their formal curriculum. Amongst these, the Lulea University in Sweden has developed an entire two year Masters programme on 'Digital Curation' which can now be studied in situ or remotely by distance learning.

Nonetheless, the requirements for educational and vocational training for digital curators have continued to evolve.

Euroguide LIS (2004)²¹ generally classified and described certain levels of qualification for information professionals from the assistant to expert level. Identifying the assignment/tasks within each level, the Euroguide LIS overview helps (new) professionals in the information business to orientate. The European Curriculum 'reflections on library and information science education' (2005) displays certain traditional and modern roles in the information sector, which should be taken in consideration as the foundation for the digital curation profession. Furthermore they stressed the importance of a balance between theory and practice:

"The issue of theory vs. practice and of academics prospective vs. vocational education was one of the first to arise in the Group's discussion. Using the words of Ton de Bruyn: we have to consider the integration between the architect and the builder, to stress that we have to build a palace and if we want that this palace will be strong and effective, we need both. The reflective practitioner approach seems to be the best example of a perfect balance, but this issue is really controversial."²²

¹⁸ <http://ensure-fp7-plone.fe.up.pt/site>

¹⁹ <http://timbusproject.net/>

²⁰ <http://www.scidip-es.eu/>

²¹ <http://www.certidoc.net/en/euref2-english.pdf>

²² http://www.library.utt.ro/LIS_Bologna.pdf

A significant step forward in terms of regularly available digital preservation training came in 2005 with the first iteration of the Digital Preservation Training Programme (DPTP) by the University of London Computer Centre. It is an introductory course designed for all those working in institutional information management who are grappling with fundamental issues of digital preservation.

“It provides the skills and knowledge necessary for institutions to combine organizational and technological perspectives, and devise an appropriate response to the challenges that digital preservation needs present.”²³

Other projects such as ‘Preserving Access to Our Digital Future: Building an International Digital Curation Curriculum’ (DigCCurr Project (2006-2009) and its extension (2008-2012) aimed to enhance the status of the digital curation profession. Above all the projects intended to raise awareness of the need for digital curators.²⁴

The third project of ‘International Research on Permanent Authentic Records in Electronic Systems (InterPARES 3, 2007-2011) has also been investigating training needs, aiming to emphasize:

“[...] what skills professionals will require to conduct such operations. On this basis, teaching modules will be developed for in-house training programs, continuing education workshops, and academic curricula that will provide professionals with the competence not only to preserve over the long-term society's documentary heritage in digital form, but also to ensure the accountability of organizations and institutions through the protection of the accuracy and authenticity of the digital information they produce.”²⁵

Another project investigating similar issues is Digital Curator Vocational Education Europe (DigCurV, started in 2011), a ‘project funded by the European Commission’s Leonardo da Vinci programme to establish a curriculum framework for vocational training in digital curation’²⁶. The DigCurV network aims to support and develop the training perspectives for digital curators in libraries, archives, museums and other cultural heritage institutions. Such skills are, of course, crucial for digital preservation and the resulting long-term management of all kinds of digital collections. Main activities of the project include the identification and analysis of existing training opportunities and methodologies, a special survey on training needs to identify key skills and competences for an initial curriculum and framework for digital curators.

The International Digital Curation Education and Action (IDEA) Working Group²⁷ is a result of the increasing discussions on the digital curator curriculum but an international level. In the IDEA mission statement are embedded the following goals:

- Describe, promote and contextualize current training and education offerings

²³ <http://www.dptp.org>

²⁴ <http://ils.unc.edu/digccurr/aboutI.html>

²⁵ <http://ils.unc.edu/digccurr/aboutI.html>

²⁶ <http://www.digcur-education.org/>

²⁷ <http://ideaworkgroup.org/index.html>

- Identify and exploit collaborative training and education opportunities
- Maximize interdisciplinary training and education opportunities
- Develop a shared digital preservation training infrastructure to enable reuse of training and education materials
- Ensure synergy and complementarity between emerging curation and preservation education programmes with professional development training courses²⁸

Finally, in 2011 a session was pronounced dedicated to ‘Education for Digital Curation’²⁹ at the World Library and Information Conference (WLIC) of the International Federation of Library Association (IFLA). This session was held to identify new and traditional competencies needed in a digital curation curriculum.

These projects have provided significant amounts of information and analysis on the topic of training for those undertaking digital preservation activities but often focused more on the skills required for those entering the profession rather than the on-going development of staff in line with the increased maturity of existing repositories and their need for continued improvement.

2.3 CONTENTS OF DIGITAL PRESERVATION CURRICULA

Continuing professional development requires an individual commitment that should start to be developed during university studies. Individual commitment and capacity to learn is developed through a set of teaching methods and practices. Therefore, curricula should be designed by collaborative teams of researchers and practitioners from various fields (Guercio, 2005, Hallam, 2006).

When identifying a new type of information professional, higher and vocational training for digital preservation will require a standardized overview of mandatory skills and competencies. It certainly cannot be managed by simply exchanging the name of the profession. The digital curator should be capable of understanding and accomplishing all of the responsibilities within the information lifecycle. Curation in the digital age calls especially for technical expertise and knowledge of information lifecycles as well as of expected future roles. Within the DigCCurr-Project Liz Madden in 2007 defined digital curators as ‘a shepherd of data as it transitions from one stage of the digital life cycle to the next’ (Madden, 2007).

Reviewing former research it can be seen that digital curation has only been a small part of preservation studies in Library and Information Science or Archiving. In a first step education institutions added special topics like digital libraries or archives, digital preservation including techniques as well as information management in the digital age. But these have always been only a part of the whole course. The situation has begun to change and with it the education for the current and future information society must change, as reported by the above mentioned initiatives for enhancing digital curation and its education strategies. Continued improvement of the technical skills provided by these courses is required to prepare students for the evolution and innovation of the digital preservation community.

²⁸ <http://ideaworkgroup.org/index.html>

The minimum basic skills for digital preservation require knowledge about the ‘Information Lifecycle’, for example the ‘DCC Curation Lifecycle Model’, all kinds of ‘repositories’ as well as ‘information ecology’. When providing education on digital preservation cultural competencies also play an important role in maintaining the cultural heritage in an appropriate way (Botticelli, 2011).

Vocational training in digital curation should also include different technical oriented positions and afford the future professional with technical, analytical as well as organizational familiarity and also self-assurance in resolving problems. Liz Madden specified in 2007 ‘Hot Topics for the Digital Curator’ (Madden, 2007):

- Digital production workflow
- Data-in-the-raw: “pre-standards” stage
- Data transfer, transformation, manipulation
- Automation and repeatable processes
- Data flexibility, shareability, sustainability
- Requirements for production tools, storage systems and display applications
- Documentation and institutional memory

All of these topics and issues must be considered by anyone attempting to develop training in digital curation, to make sure that any course covers all of the theoretical and practical issues relevant to the subject at hand.

2.4 NEED TO IDENTIFY GAPS AND CHALLENGES

This discussion of the current framework for digital preservation allows a preliminary analysis of gaps and it enables a series of practical recommendations to turn policy into action. A systematic survey of current provision follows, but some useful conclusions can be drawn from a policy level too.

In almost every source about education and vocational training for digital preservation it is stressed that there needs to be the right *balance between theory and practice*. The challenge to bridge the difference between theory and practice should be managed in overcoming the lack of standards, in addition to discovering best practice and toolsets (Botticelli, 2011). The understanding of archival theory concepts and their best practice models as well as basic knowledge of technical solutions is not enough for forming excellent professionals in digital curation.

The *capability to criticize or frame innovative and traceable solutions* must be incorporated into digital curation vocational training (Guercio, 2011) and also the ‘capability to synthesize and create new knowledge, reflect on and evaluate current practice critically’ (DPE, 2006).

The importance of building a *sustainable course structure* for digital preservation should also be emphasized. With the wide range of technical problems, new software and lifecycle components it is difficult to teach all solution mechanisms for digital curation whilst keeping content up to date.

Other challenges in vocational training for digital preservation have been recognized in relation to choosing and identifying the *right genres and documents of literature*. The discussion as to whether basic knowledge should come more from the tradition of library or archival science has yet to be

settled (Botticelli, 2011). In fact, teaching essential competencies from the different contexts and the interchange of knowledge between archival, library and museum communities is highly desirable and already practiced. Added to this is that the ‘vision of the more technically informed digital curator is becoming increasingly popular among educators and scholars (DPE, 2006). The challenge here is again to achieve a balance between potentially conflicting issues while providing a well-grounded overview of standard literature and concepts for digital curation.

Above all there is the need to address the *technical challenge* in digital preservation curricula. As already mentioned, digital preservation also includes, besides the traditional skills required for information management, a great variety of technical assignments, which consequently indicate another challenge of an information overload with respect *multidisciplinary* topics:

“Establishing new principles related to technological preservation environment – computer science and related disciplines (e.g. computer engineering, human computer interaction etc.), and finally to those that shape social, cultural and organizational contexts of creation and use and curation of digital material (e.g. information management, art media studies etc.). The potential danger in designing multidisciplinary curricula is immature approach that expands the range of subjects without a clear objective of studies integrating disciplines into meaningful whole” (DPE, 2006).

On the whole when assessing digital preservation vocational training requirements, the needed training activities can be identified as being within the following areas (DigCCurr project): ‘*technical infrastructure, digital data processing, metadata maintenance, interactions, strategies, priority validation and administrative functions*’ as well as ‘*labour market requirements in the digital curation field and appropriate content*’ (Guercio, 2011) and *teaching methods for education and vocational training* (DPE, 2006).

As the developments of the political and educational landscapes demonstrate, digital preservation has become a crucial part of information management but consistent, mature training frameworks remain elusive. These issues will need to be addressed if we wish to aid in the development of a well skilled and knowledgeable workforce to deal with the issues surrounding digital preservation.

3 ANALYSIS OF CURRENT PROVISION IN DIGITAL PRESERVATION TRAINING

This section examines the current provision of digital preservation training in Europe and seeks to identify where gaps might exist. In developing this examination we have attempted to use every technique available to us in order to maximize results and calibrate conclusions. This research has been undertaken in three parts: firstly by surveying as far as possible the complete range of digital preservation training activities undertaken between 1st January 2010 and 30th June 2012 in order to understand the range of topics and the extents of audiences; secondly by synthesizing participants' own evaluation of a cross section of recent training activities in order that their voices can be heard; and thirdly by recounting the emerging results of a series of focus groups which have examined what stakeholders - practitioners and their managers - think about training provision.

Each iteration of this research draws on smaller and smaller samples and each iteration provides subtly different insights. The findings of this survey provide a detailed and current assessment of the provision of digital preservation topics and their integration within a broader thematic context of education. This approach enables a genuine range of views to surface without having to recourse to a wide but shallow questionnaire of training needs. Such questionnaires are increasingly problematic given the 'survey fatigue' articulated on numerous occasions by likely participants, the very great effort which is required to obtain a representative sample, and the desire to make best possible use of available resources. The findings here will be further tested in sections four and five where we examine considerations of quality, thus testing the extent to which the practice that these training resources point towards can be considered best practice.

3.1 CURRENT PROVISION

3.1.1 Objectives of the research

This section describes the results of an investigation on what is currently available in vocational training courses and programmes for digital preservation in Europe. Focus has been on European courses but the main US initiatives have also been included, considering the frequency of initiatives dealing with the digital preservation subjects. The data was collected through desk-based research in October 2011, with some supplementary material added to the 1st January 2012 and a further update in October 2012.

The training courses that have been taken into account are about the management and storage of digital objects in the long-term. Only training events where digital preservation (considered broadly) was the main theme were considered, excluding those where digital preservation was a minor issue among other relevant topics. Sessions which covered issues relating to audit and certification were also identified in support of WP33.

The research has taken into account all the courses that have been organized, including the number of times that the same course was delivered in the same region or country by the same organization. The statistics therefore take into consideration both the content of courses and the frequency with which it is covered in order to provide a real estimation of the user expectations towards the topics addressed. In Annex B some of the courses have the same url: this is due to the type of web page used by the organizer to advertise training events. In these cases the name of the location has been added to better identify each training initiative.

The research has taken into account all the sectors and audiences that reflect the diversity of the project consortium, without any restriction. This ensures the accuracy and completeness of the survey data.

In order to describe the current scenario and its main tendencies, the research examined all available training activities between 1st January 2010 and 30th June 2012. The survey was conducted using event lists, bibliographies, and websites of international and national digital preservation initiatives. Data was only collected for those events that provided a documented description of the following fields:

Organizer / Country / Title of training event / Description of the covered topics / Learning objectives / Date / URL

Moreover these data were collected when available:

Prior knowledge or experience / Reference standards and guidelines / Reference sector / Target audience / Sessions on audit

For the scope of the research, only training courses, workshops, tutorials and on-line courses were considered.

Training courses are defined as formal thematic training events with clearly defined goals, learning outcomes, teaching methods and training modules that sometimes confirm qualifications and skills by an appropriate certificate.

Workshops are defined as short training events covering some facet of knowledge or skill; participants may require pre-defined skills and may get a certificate; the speaker's goal is to impart knowledge of the topic, and he or she typically uses a combination of lecture, visual aids, interaction with participants, and hands-on exercises.

Online tutorials are more interactive and specific than a workshop. Depending on the context a tutorial can take one of many forms, ranging from a set of instructions to complete a task to an interactive problem solving session. From the research findings, only one web-based tutorial has been found: it was in the form of presentations of content, with examples, broken up into discrete modules or sections, with screen recording, written documents (either online or downloadable), and audio files.

Online courses cover “webinars”, (Web-based Seminars) in the form of presentations, lectures, workshops or seminars transmitted over the Web. Webinars may be collaborative and include polling and question & answer sessions to allow full participation between the audience and the presenter. Depending upon the provider, webinars may provide hidden or anonymous participant functionality, making participants unaware of other participants in the same meeting.

The limits of this research must also be presented. First the survey gathered data only if available on the Internet: only events available on the Web were investigated. Second, the survey has excluded the results which did not provide detailed information on the required topics. On the other hand, the

research has expanded the investigation by including digital preservation courses in languages other than English, thus achieving a good coverage of European initiatives.

The research has also been assisted by early access to the results of parallel research conducted by the Digital Curator Vocational Education Europe (DigCurV)³⁰ project: a document was delivered in January 2012 presenting the findings of a survey carried out by Vilnius University Library (DigCurV, 2012) which identified existing training opportunities in digital curation and long-term preservation available for digital curators.

3.1.2 Key Findings

The research identified 134 separate digital preservation training initiatives, comprising 44 training courses, 69 workshops, 1 tutorial and 20 on-line courses. These are listed in Annex B.

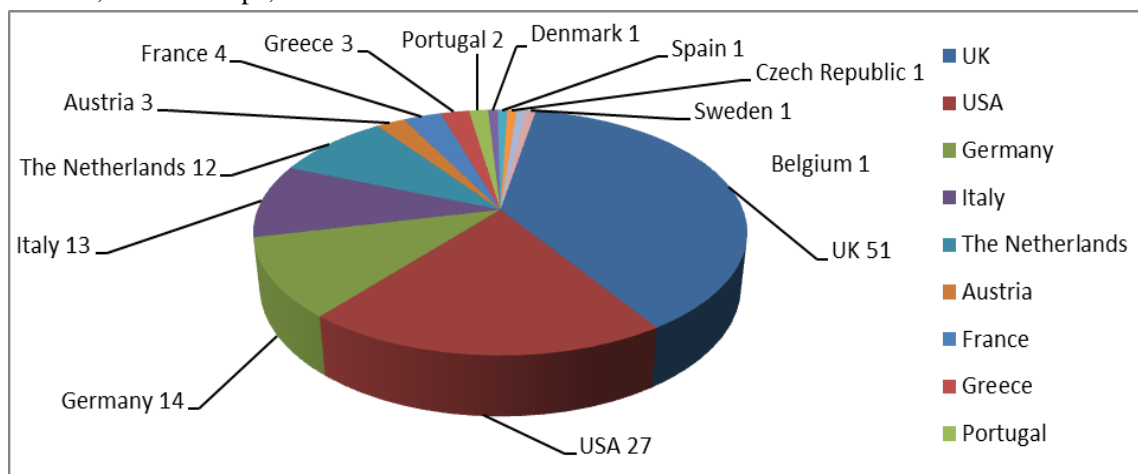


Figure 1: Location of training initiatives

The research has shown a high prevalence of training initiatives in UK (51) and USA (27); a good coverage of initiatives has been registered in Germany, Italy and The Netherlands where courses are often conducted in local language. Very few initiatives have been registered in other EU countries. The extent of provision in the UK is particularly striking.

³⁰ Digital Curator Vocational Education Europe (DigCurV), <http://www.digcur-education.org/eng/About>

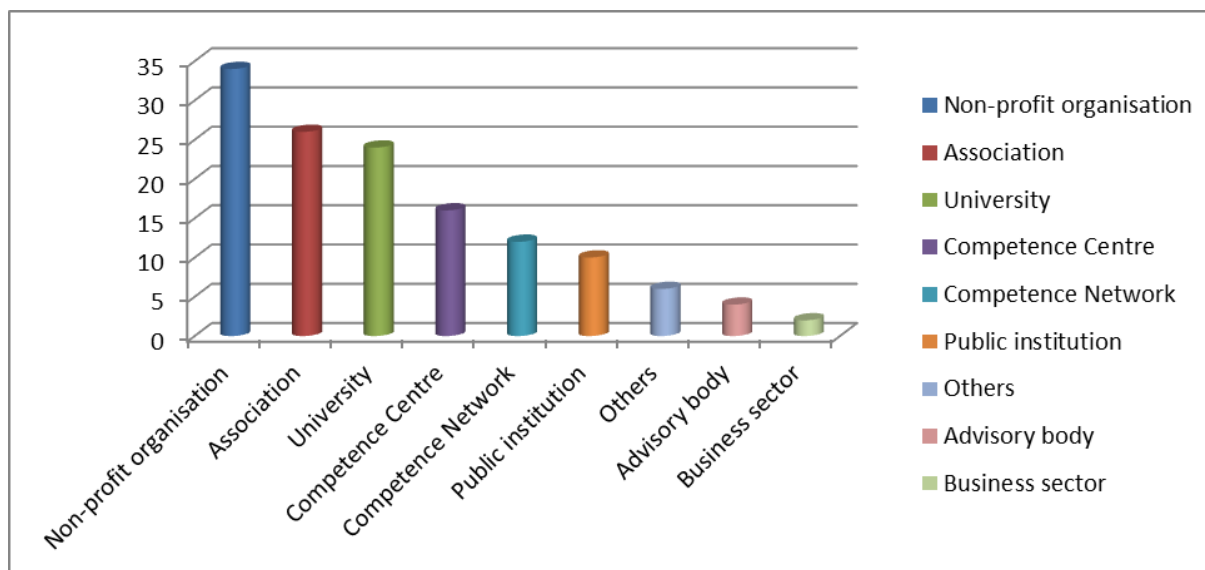


Figure 2: Type of organizer

The types of institutions that organized the events were quite heterogeneous. The majority of organizers are non-profit organizations (34), associations (26), universities (24), competence centres (16), competence networks (12) and public institutions (10). Fewer results came from advisory bodies (4) and business sectors (2).

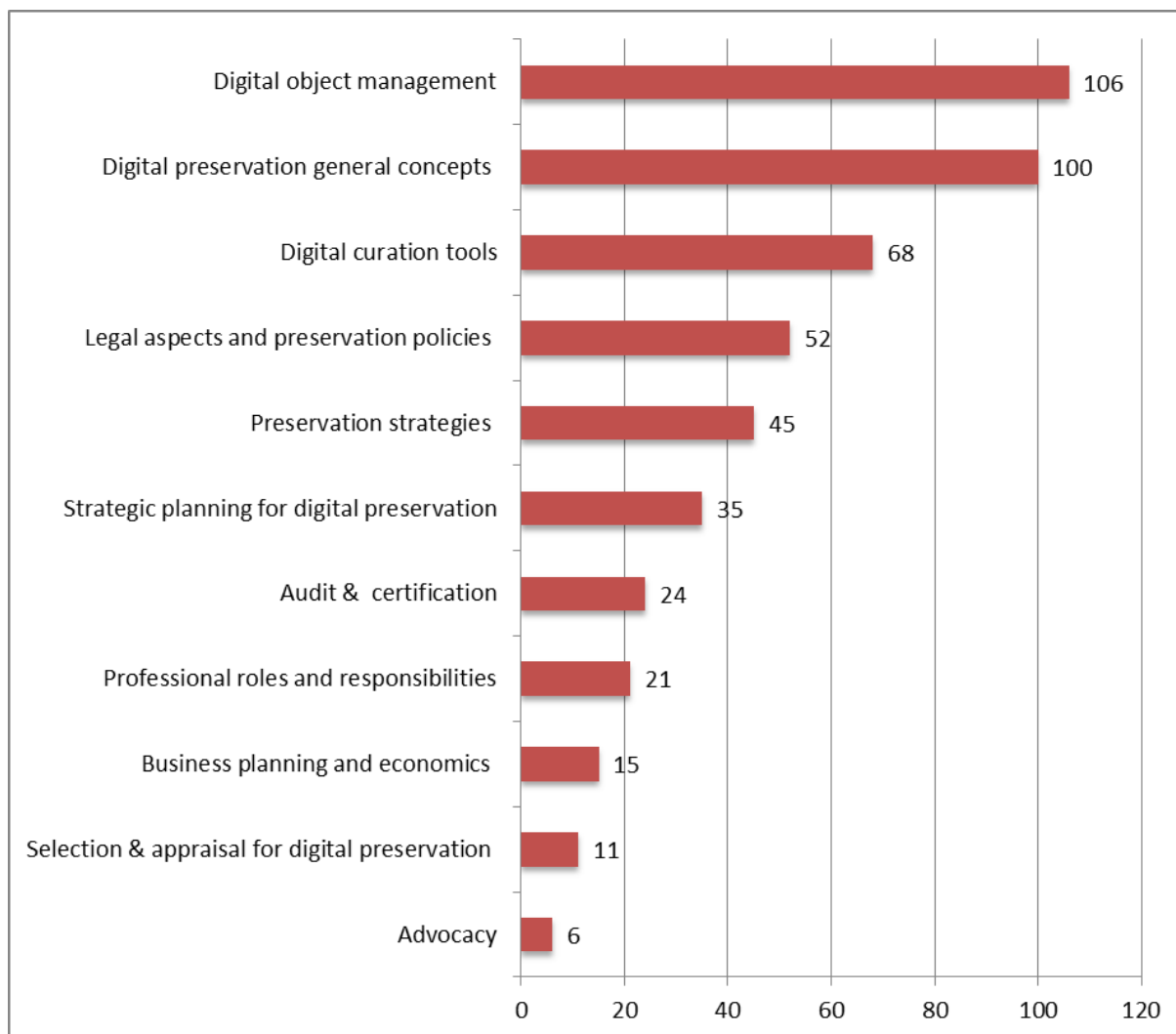


Figure 3: Key topics covered in training events

An analysis of the themes to which the events were dedicated was performed. Training events sometimes addressed a unique topic but in many cases (mainly in training courses) dealt with more than one issue. For this reason the total amount of proposed topics is higher than the total amount of the events. This analysis distinguishes the most important digital preservation topics as they are perceived by the event organizers. The keywords listed in Figure 3 emerged from the analysis of the objectives covered by the training initiatives: the research has grouped the specific topics into more general categories which correspond to the most addressed themes in digital preservation training.

Findings show that events fall mainly into these topics:

- digital object management (106);
- digital preservation general concepts (100);
- digital curation tools (68);
- legal aspects and preservation policies (52);

Other topics were also well-covered:

- preservation strategies (45),
- strategic planning for digital preservation (35);
- audit and certification (24);
- the definition of new professional roles and responsibilities (21)

The following issues were covered by fewer training events:

- business planning and economics for digital preservation (15),
- the topic of selection and appraisal (11);
- advocacy (6).

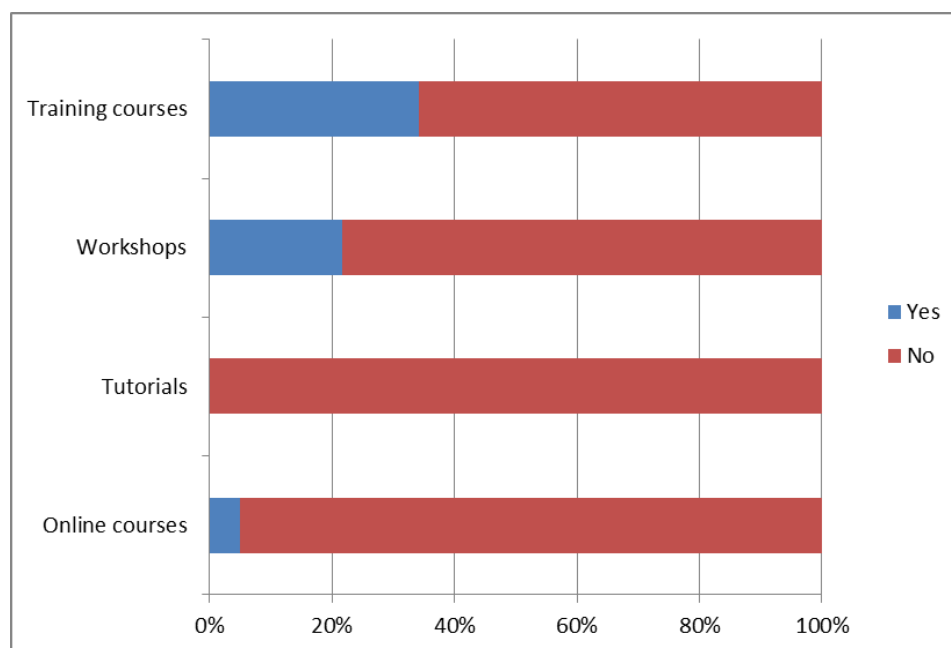


Figure 4: Prior knowledge required

Prior knowledge is required in 34% of training courses, in 22% of workshops, 5% of online courses and in the only one online tutorial. In most cases only a basic understanding of digital preservation concepts is required; three courses specifically require technical knowledge of the main standards (ISO:14721:2003, Trustworthy Repositories Audit and Certification (TRAC) Checklist, nestor Handbook for Digital Preservation) and seven courses require a good knowledge of digital preservation main topics.

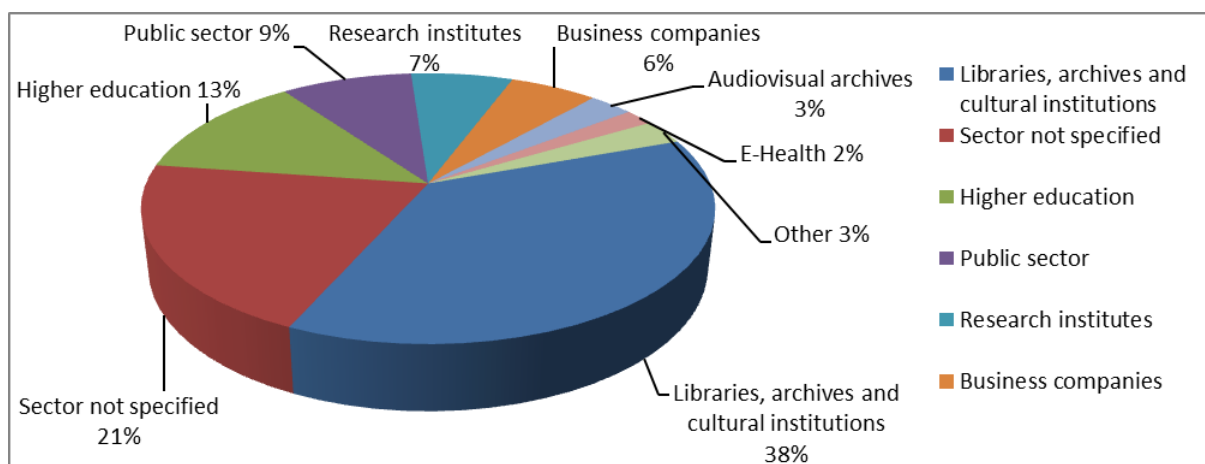


Figure 5: Target sector

Many training events did not specify the target sector of their audience (21%). An overwhelming majority of courses that did identify an audience were addressed to libraries, archives and cultural institutions (38%), higher education (13%), and the public sector (9%). A smaller percentage was addressed to research institutes (7%), business companies (6%), audiovisual archives (3%) and e-health (2%); the remaining few initiatives specified as their target sector computer forensics, software art, e-journals and CAD (Computer Aided Design) users.

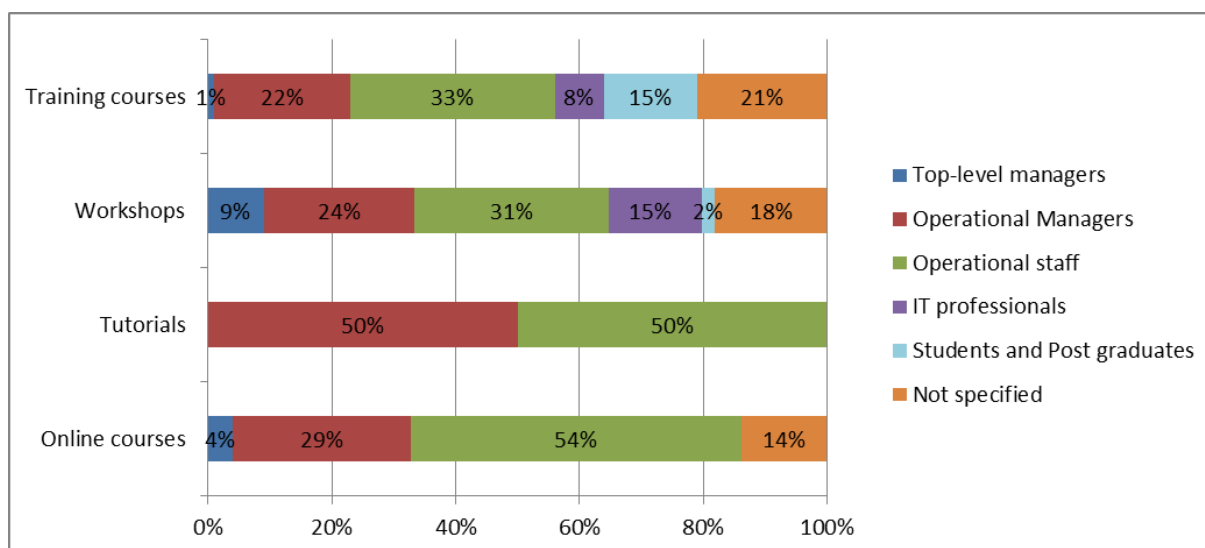


Figure 6: Target audience

As well as the sectors to which the training events are addressed, it was possible also derive the categories of professionals which have most opportunities to improve their knowledge in digital curation. The professional categories have been identified as the following:

- top-level (senior) managers: indicate those people who have the overall responsibility for the organization, the funding, and the strategic planning;
- operational managers: people responsible for running the archive of the organization; they are responsible for the main functions and processes of the repository;

- operational staff: people responsible for the effective functioning of the archive (hands-on staff)
- IT professionals: they are IT experts, responsible for digital curation technical developments within the MLA, the government, e-science and business sectors
- Students and post graduate students: some training initiatives include students at different levels.

The majority of training courses (figure 6) were addressed to operational staff (33%) and operational managers (22%). A significant percentage of courses did not have a specified target audience (21%). A smaller number of courses are addressed to students (15%) and IT experts (8%), and very few to top-level managers.

Workshops are mainly targeted at operational staff (31%) and operational managers (24%), followed by IT professionals (15%). Several workshops had no a target audience (18%). Very few workshops were addressed to top-level managers and students. The tutorial was addressed both to operational managers and operational staff. Online courses are mainly tailored to operational staff (54%) and operational managers (29%). A small percentage of these initiatives are addressed to top-level managers (4%).

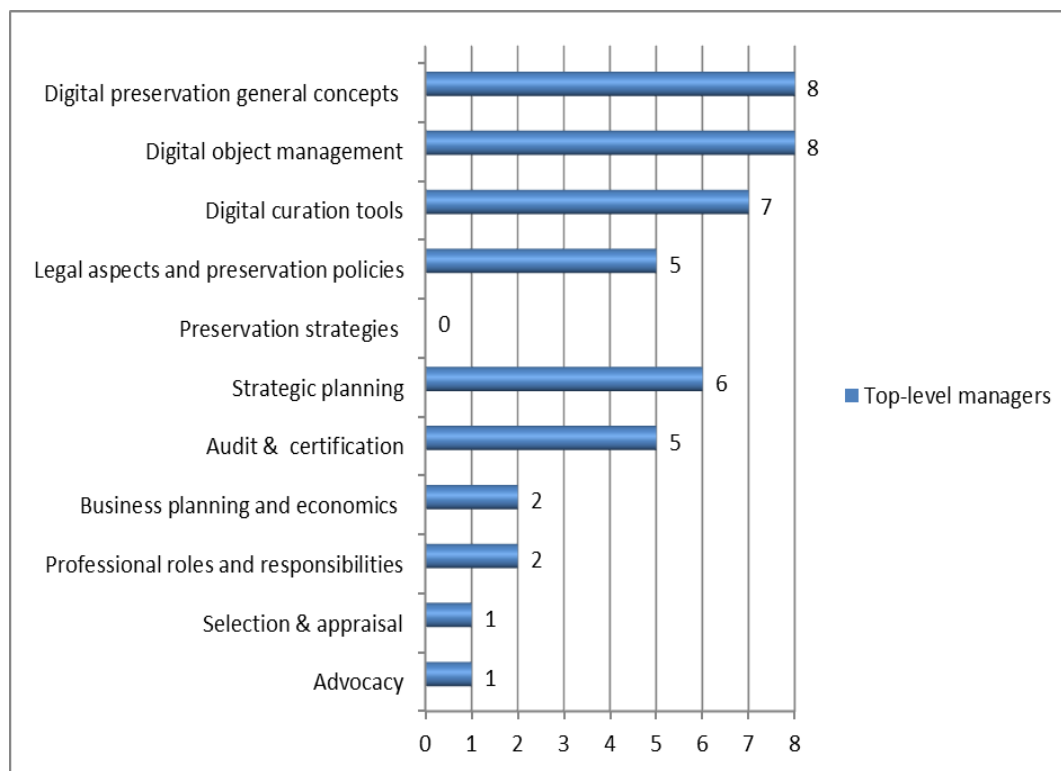


Figure 7: Key topics addressed to top-level managers

The total number of initiatives aimed at top-level management is numerically very small but a cross analysis of the data related to the key topics covered by training initiatives and the target audiences to which they are addressed, it can be seen that top-level managers mostly have training opportunities to improve their knowledge of digital preservation concepts, digital object management and digital curation tools. Other topics were less often covered, like strategic planning and audit and certification. The most interesting information from this analysis is that top-level managers have no opportunities to improve their knowledge on advocacy and few opportunities to deal with the topic of professional roles and responsibilities, business planning and economics, and preservation strategies.

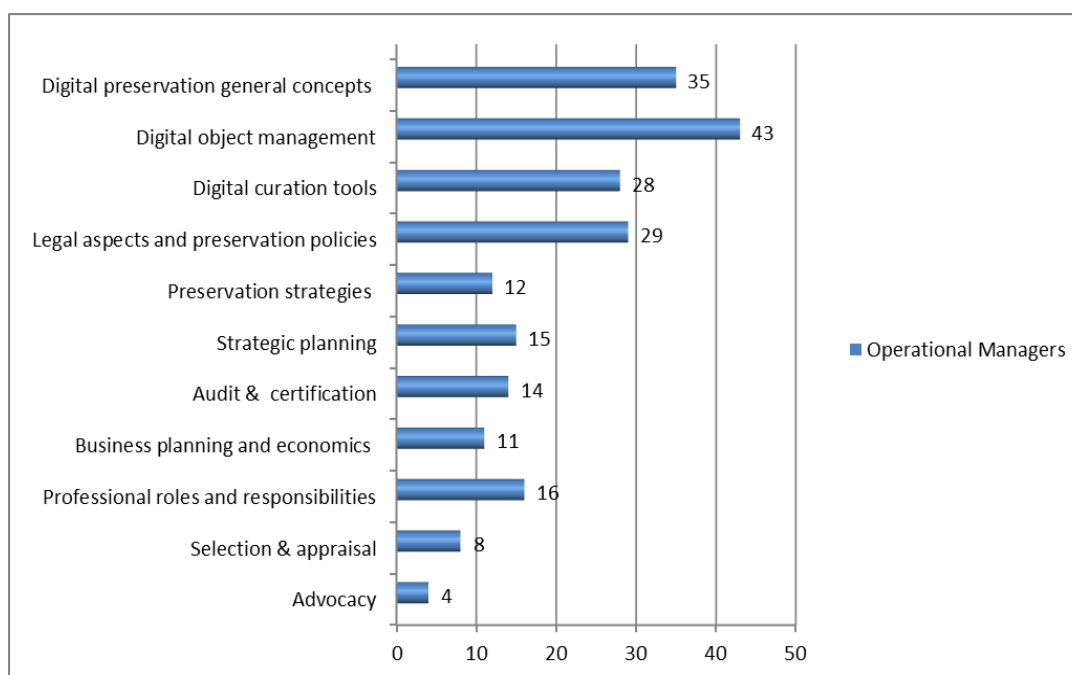


Figure 8: Key topics addressed to operational managers

The same analysis has also been performed for the category of operational managers: the most covered topics refer to digital object management and tools, general concepts of digital preservation, legal aspects and policies, followed by strategic planning. The topics related to professional roles, audit and certification and preservation strategies are less well covered. From the chart it is easy to notice that advocacy, preservation strategies and professional roles and responsibilities are scarcely addressed.

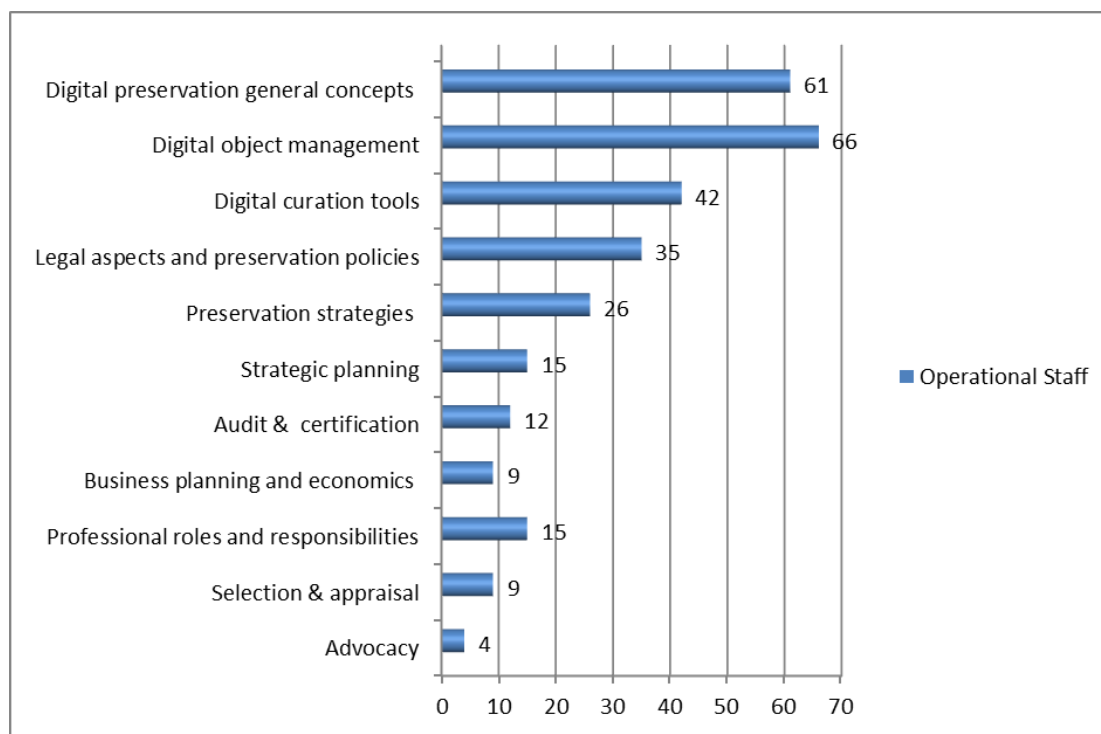


Figure 9: Key topics addressed to operational staff

Many training events directed to operational staff discuss digital object management (66) and digital preservation general concepts (61). A smaller number of events deal with digital curation tools (42), legal aspects (35), preservation strategies (26) and strategic planning (23). Very few courses were organized around the remaining topics.

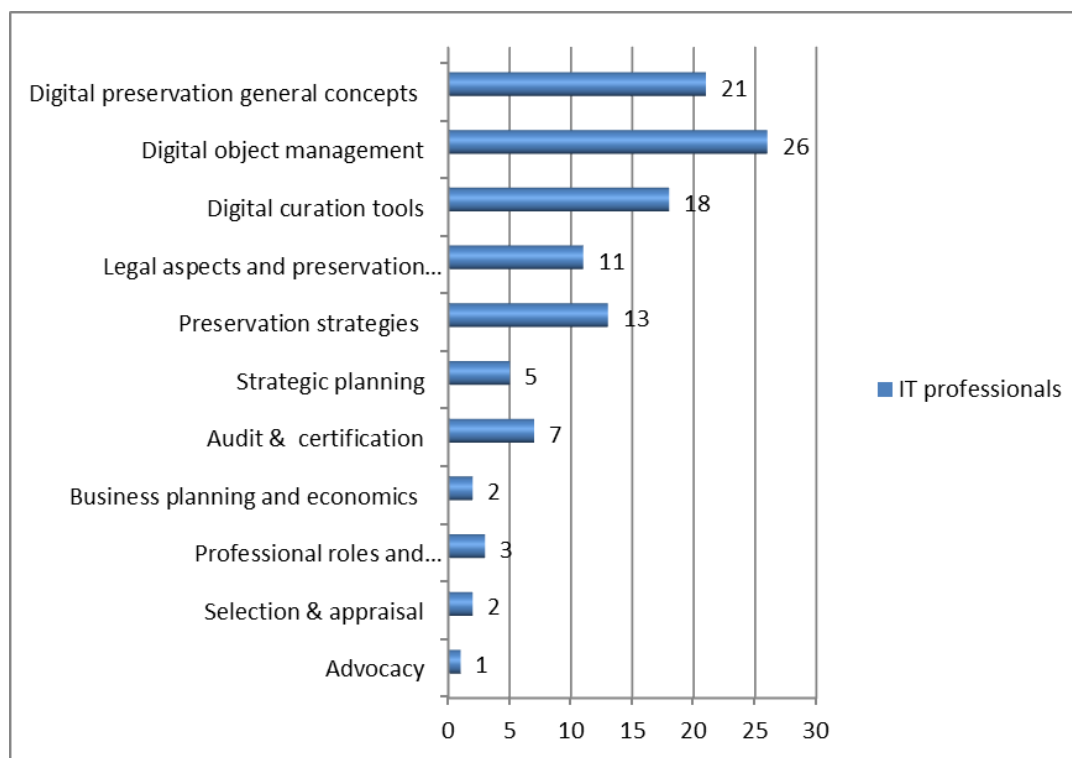


Figure 10: Key topics addressed to IT professionals

IT professionals are not in the centre of attention of training events' organizers, and the number of events is numerically small. The majority of issues covered are related to digital object management and digital preservation general concepts and tools.

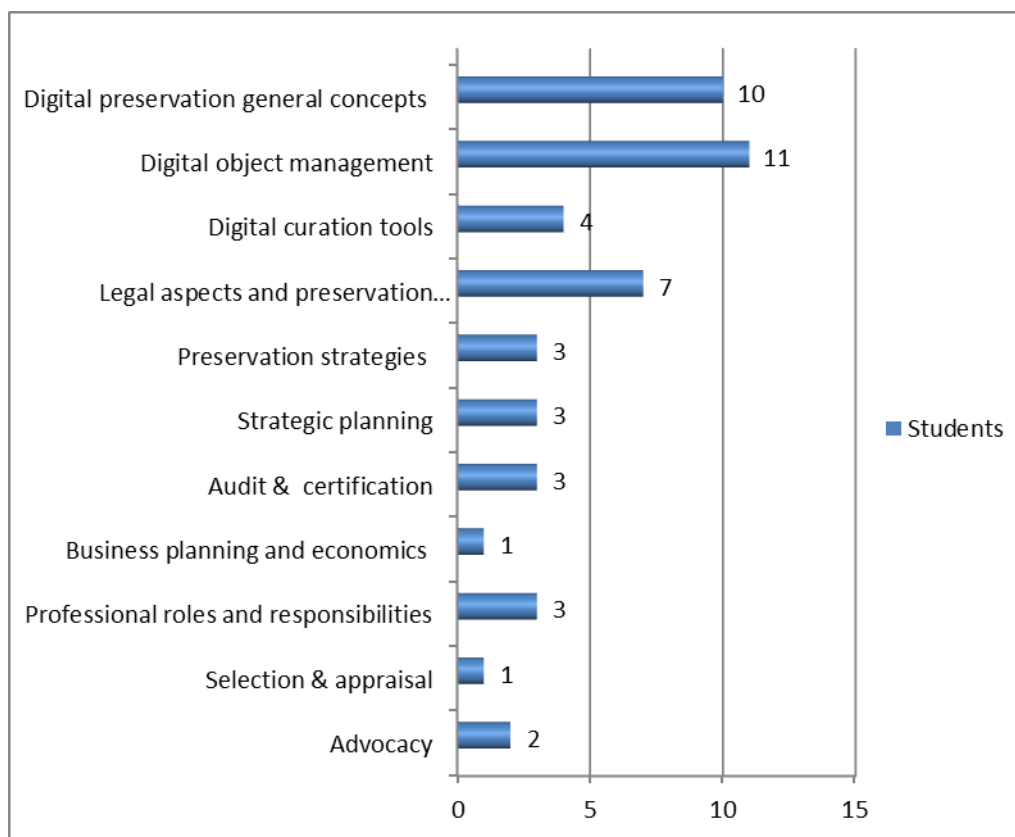


Figure 11: Key topics addressed to students

The analysis of training events was focussed on initiatives addressed mainly to professional profiles at all levels. Nevertheless many events included students within their target audience. For this reason they have been included in this analysis, even if specific events directed exclusively to students in this research have not been considered. As with other categories training for students provided good coverage of general concepts (10), digital object management (11) and legal aspects (7) but there were, at most, only a few courses available on the other topics.

3.1.3 Summary of the main findings of the analysis of current provision

The research has taken advantage of the results of an online survey conducted by the Digital Curator Vocational Education Europe (DigCurV) project: the DigCurV report, which was delivered in January 2012, presented the findings of an online survey which was carried out by Vilnius University Library for the project (DigCurV, 2012). The aim was to identify existing training opportunities in digital curation and long-term preservation available for digital curators. This research will take into account only data which is comparable for the purpose of the present analysis.

The initiatives on training were mainly located in UK and USA; a good coverage of initiatives has been registered in Germany, Italy and The Netherlands but few initiatives held in other EU countries. A further investigation showed the large diversity of institutions, organizations and centres providing

courses on digital curation and long-term preservation. The majority of organizers are non-profit organizations and associations, although university and competence centres are also delivering a good number of courses.

Amongst the various topics addressed by training courses, the most important themes in digital preservation as they are perceived by event organizers were:

- *digital object management*: these sessions were focussed on providing an overview on a particular type of digital object, considering a specific digital object lifecycle towards its preservation for long-term access;
- *digital preservation general concepts*: these events are aimed at raising awareness on the key challenges in digital preservation, on the nature and variety of born-digital information resources and the necessity to ensure long-term sustainability of digitization initiatives, identifying the main objectives and principles of digital preservation;
- *digital curation tools*: several training initiatives were aimed at the presentation of practical tools for managing digital objects over time (DRAMBORA, DAF, PLATO, DMP);
- *legal aspects and preservation policies*: a relevant number of initiatives also tackled legal aspects associated with preservation policies like protection of intellectual property rights (IPR), data protection, privacy and confidentiality issues, licensing and contractual agreements with right holders, rights management practices in institutions implementing digital preservation etc.

Other topics well covered by the training initiatives were:

- *preservation strategies*, aimed at giving an overview of the strategies that will dictate how the digital data will be kept usable and accessible through any technological changes that might occur;
- *strategic planning for digital preservation*: about setting the fundamental aims, choosing the most appropriate goals towards those aims, and fulfilling both over time;
- *audit and certification*, which included not only topics like audit but also self-assessment, risk management, certification of repositories;

The issues that were not in the centre of attention of training events were:

- *roles and responsibilities*, about the characterization of core professional principles for digital curation and the range of activities and roles that should be considered when planning and implementing a repository aimed at long-term preservation;
- *business planning and economics*, which refers to rational financial decisions and sustainable business models;
- *selection and appraisal*, the process of the development of criteria referring to the value, volume, scope, properties, types, and preservation time of materials, and;
- *advocacy*, intended as initiatives aimed at showing how digital preservation activities could produce benefits and positive impacts on the organization if integrated into the information management.

The DigCurV survey shows similar results with respect to the topics covered by training initiatives: the main digital preservation general concepts, digital curation the tools and object management (technical issues) and legal aspects are the most covered topics. In this research, unlike the DigCurV survey, standards are not significantly addressed.

The majority of the courses did not require prior knowledge, but, when requested, in most cases only a basic understanding of digital preservation concepts was required; three courses specifically require technical knowledge of the main standards (ISO:14721:2003, Trustworthy Repositories Audit and

Certification (TRAC) Checklist, nestor Handbook for Digital Preservation) and seven courses require a good knowledge of digital preservation. The DigCurV survey has obtained similar results (see Engelhardt, Strathmann and McCadden 2012).

Almost one fifth of training courses did not specify the target sector; this could suggest that many initiatives are aimed at introducing an undifferentiated target of users to digital preservation general concepts; this assertion is supported by the fact that the majority of topics are general purpose and related to the management of digital objects. An overwhelming majority of courses were addressed to libraries, archives and cultural institutions working in the field of digital curation. It can be concluded that there is a demand for continuing professional development in these sectors in particular: indeed these professionals need to be regularly updated on the developments of constantly changing environments. A significant number of courses were also addressed to higher education which is perhaps evidence that preservation of research is an increasingly important step in the research process. Also the public sector has been addressed by a number of initiatives: this result may be the consequence of the dematerialisation process as the direct result both of the gradual increase in computerised records management in public and private administrative and of the replacement of conventional administrative records with electronic records.

The analysis of target audiences allows for the identification of the target for each type of training initiative. Relevant to the questions considered here is that senior managers mostly have access to workshops and online courses, operational managers equally to all the types of training, and operational staff mostly have access to online courses, followed by training courses and workshops.

Finally the analysis concentrated on the evaluation of how the key topics related to digital preservation were addressed to different target audiences. For the research objectives, the most interesting results were the following:

- Senior managers have mostly access to training initiatives on general digital preservation concepts, digital object management and digital curation tools, while there is a critical need for specific courses to increase awareness about the importance of digital curation (advocacy) and its preservation strategies, to understand which are the professional roles and the responsibilities needed for long-term preservation, and to address the importance of business planning and economics in digital preservation;
- The analysis of current training initiatives for operational managers reveals that the aspects most covered are related to digital object management and tools, to the general concepts of digital preservation, legal aspects and strategic planning. There is need for advocacy on digital curation and the professional roles involved and, among the technical issues, the strategies for preservation are not addressed; it is striking that audit and certification issues were so scarcely addressed among managers;
- Operational staff can improve their knowledge mainly through training events which discuss digital preservation general concepts and digital object management. A smaller number of events deal with digital curation tools, legal aspects and the strategies for preservation. Very few courses were organized around the remaining topics.

Information professionals are increasingly responsible for taking care of digital objects across their full life, from strategic planning to business management, from pre-creation design and planning to provision of access, potentially over long time scales. The activities required to fulfil these responsibilities can occur within various professional and organizational contexts. The findings of the analysis performed through this survey will be exploited in the final section along with further evidence to identify gaps in current training provision. These gaps will then be considered with reference to APARSEN outputs to help shape the project's training plan for 2013/14.

It is worth observing the research here has been undertaken to support the immediate goals of APARSEN and in particular WP43. However it is likely that other projects in the digital preservation domain (e.g. SHAMAN, PLANETS, CASPAR, DPE and WePreserve) have carried out similar gap

analyses in the past, and that others (ENSURE, SCAPE, Blog4Ever, SciDipES, LOTAR and TIMBUS) may wish to use this analysis.

3.2 PARTICIPANTS' VIEWS OF TRAINING

Understanding current provision is helpful in establishing gaps and it helps to direct resources in the development of new training materials. However it does not provide a strong sense of the strengths and weaknesses of the different approaches nor does it provide a strong basis for understanding the impact which any given course may have. This section turns its attention to the evaluation which participants have provided for two specific sets of training activities – namely the range of training activities provided in the last three years through the DPC.

The DPC is a not-for-profit membership organization whose mission is to 'ensure our digital memory is available tomorrow'. It draws its members from across a variety of sectors including memory institutions and research agencies and is an advocate for change in the adoption and development of digital preservation policies, tools and techniques. In 2009 the DPC initiated a new strategic plan which included the explicit strategic goal of 'Empowering and developing our members' workforces', noting that "All of our members operate in a competitive and dynamic knowledge environment where roles and responsibilities of all staff are constantly changing. It is crucial, therefore, that staff remain informed of, trained in and confident with the new developments and tools which are released and made available. This is particularly important when existing staff are retrained to embrace a new skills set. The DPC will seek to address this issue by facilitating training and support activities and creating practitioner-focused material and events." (DPC 2009)

This strategic objective was tackled through the provision of three interconnected training programmes aimed at different levels which provide an insight into the effectiveness of training provision: two series of 'Digital Preservation Roadshows' which offered a one day introductory overviews of the topic; a series of one-day-long thematic specialist briefings and conferences; and a grant programme allowing practitioners to attend specialist and intensive third party training. This third element – the 'leadership programme' – is evaluated by third party course providers and is not available to this study. However, the DPC routinely asks participants to evaluate the events that it organizes so an analysis of the DPC's 'road shows' and the 'expert briefings' can be presented. The findings help inform the direction and configuration of future training activities.

3.2.1 Evaluation of the Digital Preservation 'Road Shows' 2009-2011

Encouraged by targeted research on the training needs of local government archivists in the UK (Boyle et al 2008) the DPC, the National Archives and the Society of Archivists established a plan for a series of one-day low cost digital preservation training days across the UK and Ireland. The subsequent programme – dubbed the 'Digital Preservation Road Show' – offered a mix of theoretical and practical lessons in digital preservation aimed primarily at archivists but open to students, records managers and museum curators. The Road Show made no assumptions about prior knowledge, but carefully selected speakers on the basis of their hands-on experience in digital preservation and their ability to interpret complex digital preservation requirements for non-technical but otherwise highly expert staff. The programme travelled over 1000 miles and engaged around 400 separate archivists with around 80 different presentations in eight different cities (Kilbride and Todd 2010).

Analysis of the evaluation of these events is methodologically challenging. Each event had subtle variation in the feedback pro-forma and in any case it would be unwise to draw any statistical conclusions from the self-selecting responses that such a process elicits. Nonetheless, feedback from these events presented some consistent messages.

The events proved unfailingly popular: events in Manchester, Gloucester and London were fully booked and so would the events in Aberystwyth, Edinburgh and Dublin except that larger venues were available for these. Such was the demand that an additional event was added in Cardiff at short notice and a further invitational event followed at Cambridge University shortly thereafter. These events were aimed at a non-expert audience of archivists who self-identified as having a need to understand digital preservation. The events clearly appealed to an audience that had not been reached by other initiatives. Moreover, respondents were consistently satisfied that the level of information presented was appropriate. So it is safe to say that there is a market for basic, easy to access digital preservation training.

In terms of content the road shows made a particular virtue of emphasizing practical solutions with a mix of case studies and scenarios. This was an evident success. Sample comments include the following comments received from participants in Gloucester, London and York.

- **When asked ‘What were the best aspects of the presentations’ participants said:**
- common sense approach to problems in case studies. Idea of getting started and giving it a try
- case studies gave useful real examples and inspiration knowing others had faced the same issues
- some were very practical in nature & gave hope of being able to do something
- case studies v. useful
- the quality of presentation was very good. Case study of WYJAS was the best.
- those case studies which matched most closely my own situation.
- information about tools available
- advice from people who have experienced digital handover projects
- naming the free tools available and the practical steps to take
- first hand case studies
- the use of freeware tools/demonstration
- practical approaches with case studies and concrete examples
- one of the strengths of this course was how, through a particular language and through concrete examples, it equipped one with surprisingly clear ideas about what digital curation does and doesn't involve. I came away from the course with the impression that while difficult, digital curation is feasible and more importantly perhaps, I came away with definite ideas on how to adapt what I had heard on the course to the digital scenarios I am trying to deal with

These same audiences were asked to comment on what had not worked in the roadshows and what was not useful to them. Here are some typical comments received:

- **When asked ‘What was not useful?’, participants responded:**
- some (presentations) were a little too detailed for the purposes of the roadshow

- a lot of sitting & listening! Could have done with more audience engagement
- needed a pc that everyone in a breakout group could see, not just a little laptop!
- more practical sessions, eg using pronom and droid
- I found some (presentations) too technical for me (sic) level of knowledge, but they gave sources for further learning
- need to address the needs of delegates with a very low level of expertise in this field and help to de-mystify the subject by offering practical solutions that can easily be applied in the workplace. Less acronyms & organizational blurb – more plain speaking.
- A projector would have been useful in the checksum session, Also it didn't go far enough in that the question was not resolved as to what to do next when you find files are not identified.
- May we have another series which takes each of the steps in turn eg identifying what you are being offered (I realize you couldn't cover all 5000 plus formats, but the more common ones), how to capture it, how to store it and how to make it available to the public. I realize there is no one answer to each of these, but to come away from the course with a basic checklist of issues to consider when dealing with digital records and their preservation would be useful

Taken in the round, these comments from participants at the Digital Preservation Roadshow 2009-10 tend to emphasize three themes consistently:

1. These audiences welcomed practical, case-study based training that matched their needs over theoretical knowledge or tools and services that were beyond their level of knowledge or irrelevant in their places of work or lacked practical application
2. These audiences wanted practical interaction with preservation processes, including trying out the tools for themselves and working with live materials to do practical problem solving that can apply to their own setting.
3. These audiences were by and large keen to make a start with preservation. Although they were sometimes daunted by the scale of the problems they faced they wanted to make progress. They were not concerned to have a complete theoretical overview of preservation first.

Encouraged by the popularity of the series, in 2011 the DPC established a second, shorter road show series with the British Library Preservation Advisory Centre with the express title 'Getting Started in Digital Preservation'. This programme provided introductory and accessible advice about digital preservation though it was intended for an audience of librarians rather than archivists. The programme for this second series was designed explicitly in response to the evaluation of the first and so was delivered as a set of short practical exercises on risk assessment and preservation planning, interspersed with practical case studies of how different agencies have embedded digital preservation into their workflows. This workshop was taken to four cities between January and April 2011 and attracted just over 150 attendees.

As with the first series, demand was greater than capacity. The London event in February 2011 sold out so quickly that it was moved to a larger venue to accommodate additional delegates. Once again the style and content of the presentations seems to have been consistent with the needs of the audience – when asked if they were satisfied with the level of information, 82% said they were satisfied or very satisfied by the level of detail available.

Numerous comments were received back about what the audience liked and disliked. Here are some typical responses from the workshops in Glasgow and London.

- **When asked to comment on ‘what went well’, the audiences responded:**
- Excellent presentation of the issues and the practical tools available to overcome the issues surrounding digital preservation.
- The variety of speakers and case studies.
- Highlight the issues. Give a wide (if not complete) view of what is being done in terms of digital preservation.
- Explaining complicated processes simply. Provided simple, practical steps. Upbeat approach to explaining subject.
- The practical sessions and case studies were useful.
- Very good, strong overview. Lively and varied - good speakers with good knowledge of subject and 'our' issues with it.
- I think the use of a very practical example to explain the risks was quite superb.
- Practical guidance on how to approach this issue.
- Presentations were excellent. Good response to issues and incorporation of comments from floor.
- Practical work. Case study - helped to see how to apply some of the tools and processes to our project work.

The audiences were asked if there was anything else they would like to have seen, the audience emphasized their desire for practical work. Typical comments include:

- **When asked if there was anything else they would like to have seen, participants said:**
- Run through of more practicalities that were only briefly mentioned at the end of the talk. PLATO ... would have been good to run our Church Collection example through some of the steps.
- List of recommended preservation and access file formats for typical file types e.g. images, word processing files, audio files etc.
- Maybe more guidance on kind of digital metadata we need to capture??
- How to preserve different formats. What and how to migrate. What metadata you need to record for different formats.
- Perhaps some technical detail/advice. Perhaps more detail re planning process.
- More information on options available when digitising different formats.
- More discussion on types of media, problems, likely life cycle before replacement.
- As an introduction at 'entry level' it was fine. I'd like to do a more practical session in the future i.e. lab session.

It is methodologically difficult to analyse the results from both of these programmes in any great detail because the responses are fragmentary and the audiences self-selecting. Moreover, because the audiences were mostly drawn from the UK and Ireland we should be aware that this is unlikely to be entirely representative of needs and experiences across Europe. Nonetheless, this combined evaluation provides an honest range of views derived from determined effort to provide introductory training in digital preservation, reaching just short of 700 professionals in a range of archive and library functions, so it would be remiss not to pay it some attention.

Four consistent messages can be discerned from the evaluation of the Digital Preservation Roadshows which might usefully inform future training activities:

1. There is a great demand for training from staff already engaged in library and archive settings, and in particular for accessible introductory material.

2, Participants at these two ‘road show’ series have consistently prioritized practical experience over theoretical knowledge, in particular through case studies and worked examples.

3, Participants at these two ‘road show’ series have frequently requested practical exercises within training, learning through doing rather than passively listening to information.

4, Participants at these two ‘road show’ series have shown willingness, even urgency, to make pragmatic progress in the preservation of digital collections. They favour small parcels of practical advice which is ‘good enough’ over comprehensive theoretical overviews and inaccessible research questions.

3.2.2 Evaluation of DPC Expert Briefings 2009-12

Alongside the entry level ‘road show’ series, DPC offers a much more focussed series of expert briefings. These events provide a concise overview of specific topics in digital preservation which have been identified by members of the Coalition as being problem issues of general concern. There have been 14 such events since the last DPC strategic plan was introduced in 2009. Two of these turned into major international conferences and six of them have contributed to substantive Technology Watch Reports. The complete list is included below in Table 1.

Table 1: List of DPC Expert Briefings

Title of event	Date	Nature of event	Number attending
Missing Links – The Enduring Web	July 2009	Co-sponsored conference	119
File formats for Digital Preservation	November 2009	Invitational briefing	13
Ejournals are for ever	April 2010	Co-sponsored conference	80
Designed to last: preserving computer aided design	July 2010	Invitational briefing	25
Jpeg2000 for the practitioner	November 2010	Co-sponsored conference	40
Preserving Digital Art	March 2011	Invitational briefing	36
Preserving digital sound and vision	April 2011	Invitational briefing	40
Data for life: digital preservation in medical science	May 2011	Co-sponsored workshop	43
Archival principles in the digital age	June 2011	Co-sponsored workshop	29
Digital forensics for preservation	June 2011	Invitational briefing	37
Preserving email	July 2011	Invitational briefing	44
The future of the past of the web	October 2011	Co-sponsored conference	106
Intellectual property rights and digital preservation	November 2011	Invitational briefing	25
Trust and preservation for e-Journal content	January 2012	Invitational briefing	36

Although different in topic, scale and delivery, these events have many features in common. They are aimed at a well-defined set of practitioners from within the Digital Preservation Coalition who are

already working in digital preservation roles. In this sense the audience might be characterized as an expert community which has advanced skills in one or other area of digital preservation but which is in need of filling gaps in their knowledge, or in need of updating existing expertise in response to some change in circumstances. Each of these events blended overviews and case study presentations with some element of research-reporting and horizon-scanning. All of them involved extended periods of discussion. In most cases attendance was actively monitored to encourage an informal and accessible format. Each event was planned with a minimum and maximum attendance threshold: with the exception of the large co-sponsored conferences, the lower threshold was set at 12 and the upper threshold was set at 40 with a target attendance of 25. As can be seen, five reached or exceeded the upper threshold, and only one came in below target. No events have been cancelled through lack of interest but in many cases applicants have been turned away because events are over-subscribed, or venues have had to change to cope with greater than expected demand.

The topics for these events are also worth noting. As a general rule, these events have tended to prioritize themes not tackled elsewhere in the digital preservation community. Scientific research data, for example, are core themes of the annual Alliance for Permanent Access conference or the International Digital Curation Conference, while emerging tools and trends in preservation technologies are well presented at conferences like iPRES. By prioritising themes directly relevant to collection management and repository operations, while inviting case studies and research overviews, the DPC briefing days sit at an intersection of research and practice.

It is methodologically difficult to undertake a systematic evaluation of these events based on the feedback received on the day alone. However a number of observations are worth making. Sample responses from the Preserving Digital Sound and Vision and Digital Forensics for Preservation give a sense as to why many of the participants attended.

- *When asked to give their reasons for attending, participants responded*
- Expand knowledge; hear latest thinking, case studies.
- Practical knowledge; to catch up with progress of tools and services I know about from previous projects; to inform my thinking on current projects.
- Learn more about AV side of digital preservation.
- Information gathering for local audio/visual storage and delivery project
- Gather information about preservation and digitisation of AV material for a new project.
- To find out more about what is going on in other organizations regarding preserving and providing access to AV files.
- To keep up to date with developments in this field.
- To understand issues of digital preservation of AV material.
- Research regarding Digital Asset Management and time-based media conservation.
- To see where our own techniques crossed over with forensics and to see where we can apply techniques.
- General digital preservation awareness and info on new tools and techniques we might use
- Want to use forensics in our new system
- Understanding concepts. Gaining familiarity with state of the art. Planning for future work in relation to University archives.

It is harder to establish a pattern of what works and what does not because the events were quite different. But when asked 'what did we do well', participants responded with comments such as:

- The two tools presentations helped me to make the jump from theory to practice and were incredibly useful to see
- I liked the fact that it was tools-focused.
- Presenters obviously very knowledgeable. Discussion very well facilitated.
- Well structured, friendly event, kept to time.
- Good range of speakers, particularly being able to attract those working at the forefront of their field.
- Event ran very smoothly and to time. Good pace, maintained interest
- Varied in level of discussion - scary but useful! Awareness. Excellent demos from presenters.
- Discussion session at end was very useful.
- Excellent networking opportunities.
- Organization of it, chairing of it, getting some good speakers together
- Provide a range of sessions addressing digital preservation as an end-to-end process.

This points to a demand for specialized, targeted and practical training among the emerging cohort of already-expert digital preservation practitioners. The DPC's expert briefing programme clearly shows that there is a demand for training that sits somewhere between the research agenda and day-to-day operations of repository management: training which can inform repository practice and gives early warning of emerging trends.

Thus, it is reasonable to conclude that training needs in digital preservation are not simply about novices and new entrants. On the contrary, there is a distinct requirement for the continuing development of existing staff – even those who might be considered expert. This need might be described as integrating practical repository operations with new and emerging tools, and ensuring that staff, policies and procedures remain responsive.

3.3 DIGCURV FOCUS GROUPS ON TRAINING NEEDS

As mentioned above in section 2.2, DigCurV is an EC-funded project working towards developing a curriculum framework for digital curation training. Two APARSEN partners, FRD and the DPC, are involved in the DigCurV project, although the DPC plays only a minor role. This connection has allowed collaboration between the two projects towards a common goal, the identification of training needs relating to digital preservation, although both projects are interested in the information for different purposes: DigCurV to guide the development of their curriculum framework and APARSEN to identify topics for the training courses it will create.

As part of the development of their framework DigCurV held a number of focus groups across Europe, encouraging practitioners and managers (drawn mainly from cultural heritage institutions) to share their opinions and ideas about training needs for those working within digital preservation. As the information to be gathered would also be of interest to APARSEN it was agreed that the project would provide some support for this process. APARSEN partners collaborated on four of these focus groups: FRD, as part of their commitment to DigCurV, organized focus groups in Milan (13th October 2011, 11 participants) and Florence (28th November 2011, 8 participants) and the DPC, representing APARSEN, joined with the University of Glasgow's HATII department to host two focus groups in London on 9th November 2011 (5 and 3 participants respectively). Access was also been provided to results from a 5th focus group in Dublin (6th October 2011, 10 participants). Early results from these focus groups provide interesting new information on the subject of training for digital preservation and clearly identified several topics in interest that were consistently highlighted during discussions.

Each focus group worked through a series of pre-set questions aimed at provoking discussions covering topics from the identification of the challenges of digital curation, through the competencies and roles required to meet these challenges, to finally identifying the training needs that would help staff to meet them. Of the recurring themes around the issues relating to digital curation the most prominent included the need to:

- Improve communication between technical and non-technical staff through a shared vocabulary.
- Raise awareness of digital curation widely across organizations.
- Better integrate physical and digital collections.
- Ensure staff possessed domain knowledge **and** digital curation skills.
- Develop methodologies for keeping on top of rapidly changing technologies.
- Provide non-technical staff with a basic understanding of information technology concepts such a database structures.
- Develop mechanisms for making informed decisions about standards and systems to be implemented.
- Maximize access and understand users.
- Make informed decisions about selecting digital collections material for preservation.
- Embed process in the normal workflows of the organization.
- Better understand the legal issues surrounding digital curation particularly in relation to negotiating for collections and Copyright/IPR.

Working from their list of issues the groups identified training needs in relation to each of the following topics:

- The Digital Curation Lifecycle
- Advocacy
- Metadata (Representation Information – for discovery, access and preservation)
- Negotiation with Depositors
- Copyright and IPR
- Selection and Appraisal
- User Analysis and Engagement
- Change Management
- Succession Planning
- Requirements Analysis
- File Format Identification
- Quality Assurance
- Risk Mitigation

The list of topics drawn out from the discussions confirm many of the conclusions from the analysis in the preceding sections. Those working in digital preservation are looking for training on issues covering the whole preservation lifecycle with a particular emphasis on practical solutions and issues such as selection and appraisal and advocacy. There was also a greater emphasis on subjects such as user analysis and engagement than has previously been identified.

Finally discussions around potential formats for training suggested that there was a tension between what was seen as preferred and what was actually realistic. Overall participants were relatively negative about training courses of a week or shorter that tried to cover all aspects of digital curation. Those wanting to gain an understanding of a broad range of digital curation issues thought that longer modular courses were more appropriate with a “blended” mix of in person and online/distance learning, but that this would always need to be reinforced by gaining practical experience. Participants concluded that shorter courses could only be expected to cover a single or narrower range of topics if they were to be successful. For all types of training it was generally agreed that training courses targeted at the different roles within digital curation would prove useful, that training courses should be taught by practitioners, and that content should be regularly reviewed (at least annually) to remain relevant.

These conclusions were counterpointed by discussions highlighting the difficulties in gaining access to training, usually due to an inability to gain funding to cover the training costs or to negotiate the time away from work. The irregularity of the availability of digital preservation training was also mentioned as a serious issue.

3.4 ANALYSIS OF CURRENT TRAINING PROVISION - CONCLUSIONS

This section has examined current provision in three ways: it has surveyed the many initiatives that have offered digital preservation training in Europe between 2010 and 2012; it has synthesized participants’ own evaluations of a cross section of recent training activities; and it has reviewed what a number of key stakeholders – practitioners and their managers - think about training provision. Each iteration has provided subtly different insights but consistent messages have appeared.

Firstly it is worth celebrating the active development of digital preservation training and the increasing penetration this is having, especially in the archive and library community. Some of the communities introduced in section one, including the commercial and scientific research communities seem under-represented in these activities. Some topics, including introductory ‘general concepts’, ‘digital object management’ and ‘tools’ seem well represented, while more advanced topics like ‘preservation strategies’, ‘strategic planning and ‘selection and appraisal’ are less well provided. Also, some elements of the digital preservation workforce are better served than others: senior managers and IT staff are less well provided for than operational staff.

The views of participants offer some clear messages about current provision to refine the findings of the survey and help us to anticipate the preferred learning styles of participants. There is a clear demand for introductory material for new entrants, and whilst the survey suggested that operational staff were well provided for, evidence shows that this group seeks advanced or forward-looking training on specialist topics so that they can respond more confidently to new or emerging challenges. Both groups have a preference for a blend of practical and experience based learning over didactic or theoretical teaching, and both groups point to a willingness – sometimes an urgency - to ‘get started’, learning from their own mistakes rather than waiting for a canonical body of theoretical good practice.

In terms of specific topics, practitioners and their managers have articulated a long list of views on the specialist topics which they think are most important. These include the ‘lifecycle’ models, advocacy, metadata, negotiation with depositors, legal aspects of preservation, selection and appraisal, user-engagement, organizational change, succession planning, requirements analysis, file format identification, quality assurance and risk management. It is probably beyond the resources of APARSEN to supply this entire ‘wish list’, but thoughtful investment and careful analysis should make it possible to supply the need in some of these topics.

The findings here will be further tested in sections four and five where we examine considerations of quality, thus testing the extent to which these aspirations point towards the community's expectations of best practice or genuine and recordable weaknesses in repository operations.

4 PROFESSIONAL STANDARDS AND DIGITAL PRESERVATION

4.1 INTRODUCTION

Sections two and three have considered different aspects of training needs in digital preservation: section two examined the strategic fit between digital preservation training and more general vocational training policies; section three examined the current provision through a variety of means and drew conclusions about how this might be expanded. This section, combined with section five, offers a different but complementary perspective. Training is not an end in itself: the ultimate measure of success of any training is the extent to which professional practice is improved. Thus, the next two sections examine emerging professional practice in digital preservation. This section makes recommendations about training based on an analysis of standards in preservation and how these have emerged in the last decade or so. Section five will examine strengths and weaknesses within operational preservation facilities in order to make recommendations about training which will lead to material improvements in digital preservation services.

The term standard is used broadly in this section, representing any form of emerging professional discourse or professional practice which has been codified to convey best practice. Some of these (such as OAIS, or the ISO 27000 series) are standards in a strict sense of having been adopted by a standards agency. Others (such as PREMIS or MoReq2010) are practical standards and specifications which summarize professional practice, though may not have been adopted by a standards agency. There are also a series of tools (like DRAMBORA or TRAC) which help repositories model standards from other domains (such as the ISO 31000 family of standards on risk assessment). Standards in digital preservation continue to emerge and, because of a perception of fragmentation, APARSEN (WP13) is undertaking a specific action to co-ordinate common standards and help users decide which standards are appropriate to their own context. Readers may wish to refer to D13.1 as a compendium and index which supports this section.

The underlying assumption of this section is that training needs to be responsive to what might be considered an emerging 'professionalism' or professional discourse around preservation that is most obviously expressed in a number of standards that have emerged in the last decade. As has already been reported from participants in workshops, there is a tension between the abstract formalism of the standards, and the contexts in which standards might be deployed. This is a problem which the standards themselves have had to consider. Moreover uncertainty and fluidity as to the authoritative sources of professional practice mean that there is a subtle negotiation between recognizing the broad range of standards, selecting the ones which are relevant then applying them. Participants at workshops have clearly articulated the desire, in some cases urgency to 'get started', assuming that compliance with standards is a longer term goal rather than a short-term requirement.

Three interconnected observations are offered about the role of standards in digital preservation training:

- that knowledge of key standards and how they might be applied is desirable to maximize the impact of lessons already learned and to reduce the possibility of fragmentation;
- that standards in digital preservation cannot be understood in isolation: the standards themselves cross reference and support each other; and their application 'always-already' assumes some adaptation to meet local circumstances;
- that standards for digital preservation operate at broadly four conceptual levels: repository-wide standards (which pertain to the architecture of an entire repository); repository-operation standards (which pertain to specific functions within a repository); industry-wide standards (which pertain to specific genres of repository); and extra-repository standards (which pertain

to data lifecycles before ingest or after access). The rapid expansion of the field means that each of these groups is in flux.

4.2 STANDARDS IN DIGITAL PRESERVATION: AN HISTORICAL OVERVIEW³¹

In 1996 a joint working group of the Commission on Preservation & Access and the Research Libraries Group Task Force on Archiving of Digital Information recognized the development of national – and, increasingly, international – systems of digital repositories that would be responsible for the long-term access to the world’s social, economic, cultural, and intellectual heritage in digital form (CPA / RLG 1996). This work was progressed in March 2000 when RLG and OCLC (Online Computer Library Center) started a collaboration to establish attributes of a digital repository for research organizations, building on and incorporating the emerging international standard of the Reference Model for an Open Archival Information System (OAIS). They went on to define “a trusted digital repository” as one whose mission is to provide reliable, long-term access to managed digital resources to its designated community, now and in the future (RLG/OCLC, 2002).

Also in 2002, the Consultative Committee for Space Data Systems released the first complete version of OAIS which was recognized as an ISO standard in 2003 (ISO 14721:2003). An OAIS is “an archive, consisting of an organization of people and systems that has accepted the responsibility to preserve information and make it available for a Designated Community”. An ‘OAIS archive’ could be distinguished from other uses of the term ‘archive’ by the way that it accepted and responded to a series of specific responsibilities. OAIS provided for the first time a systematic framework for understanding and implementing the archival concepts needed for long-term digital information preservation and access, and for describing and comparing architectures and operations of existing and future archives.

OAIS defined roles, processes and methods for long-term preservation, but it delegated the development of other standards by providing a “Road map for development of related standards” (ISO 14721:2003, 1.5). This identified topics which could be developed by other bodies so long as their actions were coordinated in order to minimize incompatibilities and efforts. Examples include:

- standard(s) for the interfaces between OAIS type archives;
- standard(s) for the submission (ingest) methodology used by an archive;
- standard(s) for the submission (ingest) of digital data sources to the archive;
- standard(s) for the migration of information across media and formats;
- standard(s) for recommended archival practices; and
- standard(s) for accreditation of archives.

These are not part of the OAIS per se but were considered essential to the development of an OAIS in practice. The last two of these – ‘recommended practices’ and ‘accreditation’ – are particularly relevant to subsequent developments.

³¹ This historical survey of standards in digital preservation is not intended to be exhaustive, nor does it describe the standards in particular detail. The former is available from APARSEN D13.1 (Report on Standardization Activities) which gives an extended list of standards currently relevant to digital preservation with bibliographical details which enables a more detailed examination.

Measuring compliance with the OAIS reference model has been the subject of considerable effort not least because it was incorporated among the attributes by of a Trusted Digital Repository by RLG/OCLC in 2002. This framework of attributes was designed to accommodate different situations and institutional responsibilities while providing a basis for shared expectations. The attributes they identified reflected the expert community's best thinking at the time:

1. Compliance with the *Reference Model for an Open Archival Information System (OAIS)*
2. Administrative responsibility
3. Organizational viability
4. Financial sustainability
5. Technological and procedural suitability
6. System security
7. Procedural accountability

Although they are somewhat generic and uneven, these seven themes have proven to be a neat summary of the different issues that have faced the digital preservation community in establishing the viability of their services. Some services, such as AHDS, developed obvious strengths in relation to criteria 1, 5, 6, and 7 but struggled with 2, 3 and 4. Others, such as national libraries and archives were robust in terms of their administrative responsibility but have had to work hard to develop technological and procedural suitability.

It would be wrong to attribute a causal relationship, but the RLG/ NARA task force also made seven specific recommendations for targeted research and development which do in fact characterize a great deal of the research in the decade that followed:

- *Develop a framework and process to support the certification of digital repositories.*
- *Research and create tools to identify the attributes of digital materials that must be preserved.*
- *Research and develop models for cooperative repository networks and services.*
- *Design and develop systems for the unique, persistent identification of digital objects that expressly support long-term preservation.*
- *Investigate and disseminate information about the complex relationship between digital preservation and intellectual property rights.*
- *Determine which technical strategies best provide for continuing access to digital resources.*
- *Define the minimal-level metadata required to manage digital information for the long-term. Develop tools to automatically generate and/or extract as much of the required metadata as possible.*

In May 2005 RLG continued this work by releasing a draft 'Audit Checklist for Certifying Digital Repositories' (RLG-OCLC, 2002) which developed criteria to "identify digital repositories capable of reliably storing, migrating, and providing access to digital collections." It was presented in four sections: *Organization* (governance, staffing, policies and procedures, financial sustainability and contracts and other obligations); *Repository functions* (the range of repository preservation responsibilities including ingest, archival storage, description, metadata, access, and preservation strategies); *Designated Community* (the creators and users of content and the capacity of the repository to meet their needs); and *Technologies and technical infrastructure* (security, software and hardware,

and similar issues that enable digital preservation.) The checklist implied that implementation may take different forms to serve different needs, while higher level policy could be assessed more easily.

In February 2007, the Center for Research Libraries and OCLC released version 1 of Trustworthy Repositories Audit & Certification: Criteria and Checklist (CRL, 2007 referred to hereafter as TRAC). This developed further the draft RLG/NARA checklist and was written to be applicable to any kind of digital repository or archive. It listed 86 criteria in three sections: Organizational infrastructure; Digital object management; and Technologies, technical infrastructure, and security.

Using the terminology of “threats” and “risks”, TRAC introduced the risk assessment process as the main tool for assessing information security management systems, a connection which has remained a persistent theme within preservation planning, and which points to a family of standards around risk assessment and management (now ISO 31000:2009). TRAC also made an explicit link to standards for data security, quoting ISO 17799:2000 “Information technology - Code of practice for information security management” (now 27001:2005 and 27002:2005) as the reference standard for Information Security Management. Its requirements for information security seek data security compliance to a very granular level. In specific terms it requires that management systematically examine an organization's information security risks; design and implement a coherent and comprehensive suite of information security controls and/or other forms of risk treatment; and adopt an overarching management process to ensure that the information security controls continue to meet the organization's information security needs. Information security is defined within the standard as a ‘C-I-A triad’: Confidentiality, Integrity and Availability.

The TRAC checklist has recently been developed into a full standard, *Audit and Certification of Trustworthy Digital Repositories*. In September 2011, CCSDS released the new recommended practice on “Audit and certification of trustworthy digital repositories” which is due to be published as ISO 16363. The overlap with TRAC is significant but, while TRAC was mainly directed to self-audits of digital libraries, ISO 16363 is designed to form the basis for an external audit process of all types of repositories, from cultural to science to commercial. A trial implementation of ISO16363 has been funded through APARSEN (see D33.1B, “Report on Peer Review of Digital Repositories”), and lessons learned from the experience are presented in section five.

TRAC’s 86 criteria and its expectation of conformance to cognate standards in risk management and information security makes compliance challenging. A range of other standards have also emerged which attempt to support repository improvement through a self-assessment and audit.

At the same time as CRL and OCLC were developing TRAC, the DIN Standards Committee in Germany adopted the DIN 31644 standard - a set of criteria that define standardized requirements for the setup and management of digital archives, based on recommendations from a working group of nestor. The nestor Catalogues of Criteria for Trusted Digital Repositories (nestor 2006 and 2009) define a set of criteria necessary to demonstrate trustworthiness and to prepare repositories for certification. Further developments in collaboration with international initiatives led in January 2007 to the formulation of 10 core requirements for trusted repositories.

DIN 31644 (published 2012) “Information and documentation- Criteria for trustworthy digital archives” is intended for use by all institutions that have the aim to preserve information in digital form. The standard consists of a list of 34 requirements structured in 3 parts: organization; management of intellectual entities and their representations, and infrastructure and security.

Appendixes with examples of digital archives and best practices for each requirement, as well as literature, complete the standard.

Finally, in 2008 the DANS – Data Archiving and Networked Services – established by the Royal Netherlands Academy of Arts and Sciences (KNAW) and supported also by the Netherlands Organization for Scientific Research (NWO), released the Data Seal of Approval (DSA, 2008). This is a self-assessment process for research archives which contains a total of sixteen guidelines for the application and verification of quality aspects with respect to the creation, storage and (re-)use of digital research data in the social sciences and humanities. In the assessment, the organizations are asked to consider how these guidelines relate to the repository and how they have been implemented. The assessment reflects the current situation of the repository in a transparent and open manner. It is renewed every year through a modification procedure, and in this way it supports on-going peer-review and practical, attainable service improvement. A number of repositories have since been ‘awarded’ the Data Seal of Approval.³²

The apparent proliferation of repository standards is a potential barrier to participation. Consequently the European Commission has hosted a series of meetings to discuss and agree on the possibility of a European-wide approach to repository standards. As a result, there is now a Memorandum of Understanding to define a European Framework for Audit and Certification of Digital Repositories.

4.3 STANDARDS AND REPOSITORY OPERATIONS

The standards described so far have been at a high level, pertaining to the operation of repositories in general and assessing their competence. It is worth noting, however, that communities of practice and standards have emerged around a series of issues within the operation of repositories, such as issues of preservation metadata, persistent identifiers, authenticity, provenance, annotation, preservation planning and file format characterisation.

Four of these topics – persistent identifiers, authenticity and provenance, and annotation and data quality – have been the subject of extensive reports by APARSEN already and therefore readers are recommended to examine reports D22.1 (Persistent Identifier Framework), D24.1 (Authenticity and Plan for Interoperable Authenticity Evaluation System), D24.2 (Implementation and Testing of an Authenticity Protocol on a Specific Domain) and D26.1 (Report and Strategy on Annotation, Reputation and Data Quality). A fifth topic – File Format Characterisation – has recently been explored extensively by the EC-funded SCAPE project (Van der Kniff and Wilson 2011). These will not be discussed in detail below except to note that training needs to take account of the emerging standards in these fields. However the topics of preservation metadata and preservation planning deserve particular attention because they impact on all other preservation actions.

Although OAIS had a comprehensive information model, it offered no specific mechanism to encapsulate preservation metadata. So in 2003 OCLC and RLG established the PREMIS working group in order to define implementable, core preservation metadata, with guidelines/recommendations for management and use. In May 2005, PREMIS released the Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group which included the PREMIS Data Dictionary version 1.0 (version 2.1 is currently available, 2011), a comprehensive, practical resource for implementing preservation metadata in digital archiving systems. The PREMIS working group has

³² List of repositories with the Data Seal of Approval – see website
<http://assessment.datasealofapproval.org/seals/>

evaluated, among others, some important characteristics necessary for long-term preservation which need to be captured alongside the objects themselves (PREMIS, 2008). It is important to distinguish PREMIS metadata from descriptive or resource discovery metadata. It exists to support typical digital preservation actions including managing viability, renderability, understandability, authenticity and identity. It assumes metadata will be auto-generated and that other suitable descriptive, technical and packaging metadata standards will be used in conjunction with it. It defines the relationship between five digital preservation activities or entities which can be gathered and retained at a file level if necessary: agents, rights, events, objects and intellectual entities.

PREMIS is frequently quoted in parenthesis with METS – the Metadata Encoding and Transmission Standard. METS is an XML schema designed to store and structure all the metadata associated with a digital object. METS files have four major constituents: an inventory of the files associated with the digital object; administrative metadata; descriptive metadata; and a structural map of how the components are fitted together to create a digital object. Metadata and data associated with a digital object may be either stored internally within the METS file, or held externally and referenced from it. The content of each section is not prescribed by METS itself: any XML data or metadata may be used; however, METS does recommend a number of schemas. The flexibility of METS implies that its practical implementation can be very flexible as well: any system capable of handling XML documents can be used to create, store and deliver METS-based metadata. METS Profiles can be used to document a particular METS implementation within a project. As METS was designed to act as OAIS Archival Information ‘no conceptual leap is required to fit METS into the OAIS landscape’ (Gartner and Lavoie 2005, 3).

Another part of OAIS which has seen considerable development in the last few years are the elements pertaining to preservation planning. Preservation planning occupies a significant proportion of the OAIS functional model and while elements of the planning process and the relationships between them are described, there is no methodological statement of how to write a preservation plan. Consequently a number of planning tools and approaches have been developed to support the development and validation of preservation plans. The sharing of preservation plans is recognized not only as a way to improve the transparency of preservation decisions but also as a way to reduce the barriers to preservation. The PLANETS project in particular, developing ideas from the DELOS Network of Excellence, offered a four part methodology for the writing of preservation plans which started with the definition of requirements, the evaluation of alternatives, considering the results of experiments, and followed by writing the plan itself (Becker et al 2008). Subsequent work has turned this methodology into an iterative planning tool called PLATO and a library of existing plans³³.

4.4 INDUSTRY AND SECTOR SPECIFIC STANDARDS FOR DIGITAL PRESERVATION

As will be obvious from the previous sections, the main contributors to digital preservation standards have been either the scientific research community with particular needs to retain data, or memory institutions with particular responsibilities for preservation. More recently, specific industries have also become active in the development of preservation standards and particular types of content and use cases have emerged that overlap and extend standards within the scientific research community or memory institutions. WP13 has the task of tracking and co-ordinating common standards for APARSEN and the authors of this report have based it partly on a draft interim report which attempts to create a comprehensive baseline of preservation standards, so readers may wish to consult this (D13.1(i)). The same report will become an exemplary reference resource for training. In this section, we pick two examples of industry specific approaches to digital preservation standards, partly for their

³³ <http://www.ifs.tuwien.ac.at/dp/plato/intro.html>

likely relevance to a range of use cases, but also because they point to an inherent dilemma about 'professional practice' in digital preservation.

Audio visual materials present a special case for digital preservation because the media on which analogue recordings are held are subject to inevitable, and at times rapid, degradation. Consequently, it has become common place for the preservation of analogue recordings to become a question of digitization, which in turn creates a digital preservation challenge (Wright 2012). Actions to preserve digital content are a corollary to secure access to all sound and vision collections. Recommendations for audio recordings and video recordings exist under the auspices of International Association of Sound and Audio-visual Archives (such as IASA 2009), while a range of industry bodies and content holders including the BBC, RAI, ORF and INA have formed the PrestoCentre to progress research and development of preservation standards in this field.

For different reasons, but with similar effects, the aerospace industry has particular requirements in product lifecycle management and information exchange which have given rise to a series of industry wide initiatives with the goal of standardising approaches. The origins of this work are about aligning and sharing CAD drawings for engineering under the auspices of an industry wide membership body called PROSTEP which in turn created the ISO 10303 'Standard for Exchange of Product Model Data'. This has developed into an industry standard for the creation and preservation of engineering data called LOTAR, (ASD-STAN 2007). LOTAR combines the aspirations of long-term access with the need for standardized, readily understood and completely reliable geometries. It is not inconsistent or incompatible with OAIS nor with many of the other standards mentioned in sections 4.2 or 4.3, but because it has a strong industrial focus and because it fits within a data exchange protocol so important to the industry, it is likely to have a degree of precedence over more generic preservation standards. Aerospace engineers are more likely to encounter OAIS through LOTAR than the other way round.

The development of LOTAR, and IASA-TC06 are instructive. There is considerable credit in digital preservation standards being embedded in sector-specific standards since this will greatly assist their adoption. It presents a challenge to co-ordination of standards (being tackled by APARSEN WP13) in the longer term as there is a clear risk of fragmentation.

It also presents a dilemma to training providers: whether in the long run it is preferable to teach digital preservation as a distinct set of skills towards the creation of a specific profession, or it is better folded it into existing professional training. As section three shows there is still considerable demand for digital preservation training, and by helping to disseminate the state of the art, training is likely to assist in the co-ordination and development of common standards across the industrial sector that interest WP13.

4.5 DATA LIFECYCLES AND EXTRA-REPOSITORY STANDARDS

As noted at the start of this report, the definition of digital preservation has varied through time and context. Other terms like 'Digital Curation' and 'Digital Continuity' have also been used to refer to the need for data to be long lived. In addition, experience has shown that the effort associated with preservation is 'front-loaded' – that preservation is most effective when planned at an early stage and that much of the work required to run an effective repository has to happen outside of the repository. In the short-term this is most obvious in the negotiations that happen between the repository and producers; in the long-term it is most obvious in the work that is needed to track the needs and implicit knowledge of the designated user community and their use of data.

It is outside the scope of this report to try to encompass all of the standards that apply to data outside the repository and the functions that may be required in order to track a designated community. However, a series of 'data lifecycle' models exist which provide a model for how repository staff may approach these tasks. Both of the examples quoted below are from the academic research community where such topics are relatively well developed. They are of unequal relevance to other sectors but they provide a useful template which repository managers may wish to extend or refine to other contexts.

Research funders in the United Kingdom have, for some time, required that applicants undertake an explicit process of data management planning prior to the approval of grants. This approach was trialled by the Arts and Humanities Research Board (the precursor to the Arts and Humanities Research Council) in 1999 when it asked for applicants to its newly established grant programmes account for Project Management, Data Development Methods, Infrastructural Support and Preservation and Access (James 2003). The process even included an assessment from the repositories that would ultimately be responsible for the data. Although policy has evolved since then, the basic concept of requiring data management planning from grant applicants is instructive. It requires that designs for preservation are written into projects before the data are created – indeed before the project is even properly committed.

This approach – intervention and review by the repository before data creation – is not practical in every circumstance, and even when it is enforced it can still mean complicated negotiations between producers and repository administrators. But it indicates ways in which repository managers might be able to influence standards outside of the repository. This is likely to be possible in agencies where the repository is closely aligned with the executive that funds a project, or where the repository and the executive are one and the same and therefore able to require compliance with standards set by the repository.

A second approach to how a repository can attempt to influence standards outside repository can be seen in research data lifecycle planning. The Digital Curation Centre has adapted classic archival theory, noting that any document or digital object goes through a series of iterative stages from 'conceptualisation', through creation, selection and appraisal, ingestion, preservation, storage, access use and re-use, and transformation to disposal (Higgins 2008). This model complements rather than restates a number of standards (such as OAIS but also ISO 15489-2:2001 on Documentation and Records Management) and nor is it the only research data lifecycle published in the last decade (UKDA 2009). But by attempting to map the actions that have given rise to data, it helps repository managers align their expectations of data with creators and users. The lifecycle is not intended to be detailed and the practical meaning of the terms vary greatly for different sectors. But it can be used to help develop some kind of shared consciousness between creators, users and repositories and if deployed in detail can begin to map and give early warning of any number of processes and transformations applied to data. All of this helps repository managers plan for ingest and provides additional descriptive metadata which may in turn also facilitate use.

As with planning for repository development there is also a potential benefit in utilising such a lifecycle when planning training courses on digital preservation. If the aim of the course is to provide participants with as complete a set of digital preservation skills as possible then there is no doubt merit in structuring training around the stages of the lifecycle, thus ensuring that all issues relating to the ongoing preservation of digital objects are covered.

4.6 OTHER EMERGING STANDARDS AND TRAINING

One of the problems with provision of training in digital preservation is the rapid development of the field. The need for a standards observatory function has already been noted in APARSEN (WP13) and will not be repeated here. However there are a number of particularly active areas of standards development that may need to be considered in the short-term by those considering training.

OAIS was adopted as a standard by ISO in 2003 is therefore almost a decade old. A review of the standard was initiated in 2008 and a series of recommendations made for changes to the standards (e.g. DPC/DCC 2009). Although early drafts of the 2008 update to OAIS did not point to radical changes, it will clearly be necessary to align training based on OAIS to accommodate these changes. Moreover, the five-year review of the 2008 review will be required to begin almost at the same time as the new standard is published.

The Storage Network Industry Association has also begun to make progress on the development of a series of standards which could have a major impact on digital preservation in the next decade. SNIA has a working group on long-term data retention which has responsibility for both physical and logical preservation, and whose mission included the creation of reference architectures, services and interfaces for preservation. Moreover this same group has established a small working party on training so that SNIA members can become more closely involved in preservation. In addition, a working group on Cloud Storage is likely to become particularly influential in relation to preservation, assuming that preservation is offered increasingly as a service in the cloud. Cloud architectures change how an organization view repositories and how they access services to manage them. For example, it is unclear how one would measure the success of a 'trusted digital repository' that was based in a cloud provider. SNIA is an industry body which involves the major vendors of software and hardware so potentially has very great influence over the deployment of preservation architectures. Consequently it would be highly desirable for APARSEN (via WP13) to form an understanding with SNIA, and in particular (via WP43) to make a connection with the SNIA Long-term retention training group.

4.7 PROFESSIONAL PRACTICE IN DIGITAL PRESERVATION: CONCLUSIONS FOR TRAINING

This section has attempted to illustrate the most significant developments in professional standards and professional expectations in digital preservation in the last decade or so. It is appropriate that training courses pay attention to these standards, advise participants of their existence and encourage professional development by sharing good practice of intelligent conformance.

Standards for repository operations, especially OAIS, can seem generic and therefore difficult to appreciate or implement directly. It is striking how the various standards about conformance to OAIS – such as TRAC – need to balance the generic requirements of a reference standard and operations specific to each repository. Of necessity such standards are relevant to all staff across a repository, including senior management. Moreover because they touch on all aspects of repository operations an overview of such standards is likely to provide a useful, if somewhat theoretical overview of the entire operation of a repository. Consequently it seems sensible to recommend that training should blend theoretical understanding of standards with practical application in the field.

Standards in digital preservation make reference to generic standards in cognate fields such as information security and risk management. It is therefore sensible for digital preservation staff to have a working knowledge of these topics, but it is probably more important for such material to be presented with relevant case studies of their utility in preservation.

Other elements of professional practice pertain to specific roles and functions of a repository, such as metadata management, preservation planning, persistent identifiers, authenticity management, annotation, characterisation and data quality so it is possible to provide a nuanced and detailed approach to specific tasks. Because these areas are relevant to the practical implementation of almost any digital preservation infrastructure, they form an important part of emerging professional practice which can be transferred from one institution to another. However any one of these topics – preservation planning, characterisation of collections, authenticity and provenance monitoring, or metadata creation and management – could very well encompass an entire training programme. So a balance needs to be struck between the detailed knowledge of specific operations and the wider overview of repository operations. It should be noted that APARSEN is in the fortuitous position of commissioning state of the art reports and research in a variety of these fields (e.g. WP22, WP24, and WP26) which create the potential for base line training materials.

Standards for preservation have begun to evolve in industrial settings, and, because repositories cannot sit in isolation, it is necessary for training to consider both the industrial context of training (and adapt to meet this), as well as the way that a repository may seek to influence standards through an entire data lifecycle.

This work of tracking and presenting emerging professionalism of digital preservation is made all the more challenging by the number of stakeholders involved and the need to keep pace with the many new and emerging standards in preservation.

5 ANALYSIS OF REQUIRED DIGITAL PRESERVATION COMPETENCIES AND SKILLS (TRAINING REQUIREMENTS) AS HIGHLIGHTED BY PEER REVIEW

This section contains the results of an investigation conducted through a series of interviews with APARSEN partners and other participants to establish training needs identified during the recent peer reviews carried out in WP33 (Peer Review and 3rd Party Certification of Repositories). This work has been undertaken in light of the Memorandum of Understanding signed to define a European Framework for Audit and Certification of Digital Repositories, where it was agreed that the repositories represented at the meeting and others should perform test audits. Details of this work is provided by the work package's final report (D33.1B, Report on Peer Review of Digital Repositories). The research presented here has also benefited from access to a series of draft reports and papers developed in other work packages of APARSEN.

The purpose of this section is to derive practical recommendations for training by analysing perceptions of strengths and weaknesses of operations in seven different repositories. In this way, APARSEN training will be linked to repository improvement and the operational needs of repository workforces. It should be emphasized that the purpose is *not* to learn about the audit process per se, nor to develop recommendations about the training of auditors.

The European repositories involved were:

- ISO 16363 focus:
 - o Data Archiving and Networked Services (DANS)
 - o Centre Informatique National de l'Enseignement Supérieur: Département Archivage et Diffusion (CINES-DAD)
 - o UK Data Archive (UKDA)
- DIN 31644 focus
 - o Deutsche Nationalbibliothek

All of the above have already attained the standard necessary for the Data Seal of Approval.

In addition three repositories in the USA volunteered to be reviewed following ISO 16363. (Note that these repositories received no EU funding to participate in the audit, though the learning acquired from them is nonetheless relevant for the purposes of APARSEN training). The US repositories were:

- o NASA's National Space Science Data Center (NSSDC)
- o The Socioeconomic Data and Applications Center (SEDAC) at the Center for Earth Science Information (CIESIN)
- o Kentucky Department for Libraries and Archives (KDLA)

The interviews that are the subject of the analysis here were conducted with experts that participated in the reviews, both as staff of the repositories reviewed and as peer reviewers. The main aim of these interviews was to provide a case study based on the reflective process of review that would identify practical training needs within the digital preservation community. For repository managers peer review involves constant monitoring, planning and maintenance as well as conscious actions and

strategy implementation to meet their digital preservation aims. Such reviews may be needed in the different phases of the life of digital objects to provide feedback and suggestions for improvement in relation to preservation activities. The ability to carry out an assessment of training needs of repository staff at different levels is an interesting and useful by-product of this process.

Seven interviews were carried out between November 2011 and January 2012, with reviewers and managers of the repositories they reviewed.

- Simon Lambert – Project Manager – STFC (reviewer)
- Mariella Guercio – Professor – University of Urbino (reviewer)
- David Giarretta – Director - APA (Alliance for Permanent Access) (reviewer)
- Barbara Sierman - Manager Digital Preservation - KB (National Library of the Netherlands) (reviewer)
- Kevin Ashley – Director – DCC (reviewer)
- Ingrid Dillo - DANS - Head Policy Development and Communication - & Henk Koning - Technical Archivist (repository managers)
- CINES - Olivier Rouchon - Head of D.P. Team (repository manager)

The structure, format and questions from the interviews are described in ANNEX B and the methodology used to analyse the result of the interviews is described below. As well as general training needs the interviews information was also provided about the training needs for staff working with the emerging standards for certification. This information does not form part of this analysis but has made available to WP33.

5.1 METHODOLOGY USED TO ANALYSE THE INTERVIEWS

The interviews were conducted in a semi-structured format (Cohen et al, 2008; RECOUP, 2008), with topics that had to be addressed in a particular order. The questions were focussed on the review of digital preservation practice, but with open-ended answers, to give room for clarifications and reasoning about both the questions and the answers.

This section explains the methodology used to perform the analysis of the information provided through the interviews with reviewers and the repository teams respectively.

The interviews allowed the collection of diverse types of information with the objective being to try to gather cross-information from both groups of interviewees about the other. While the interviews with reviewers were larger in number and provided more extensive information, the two interviews with repository managers also provided a significant amount of useful information. The intention had been to interview respondents from three repository teams, but unfortunately only two respondents were available within the window for conducting the interviews.

The questions across each type of interview (reviewer and repository manager) were standardized to allow straightforward comparisons between responses, but the interviews were carried out in a semi-structured format to allow some divergence in answers if interesting and useful information was being communicated.

The methodology adopted to analyse the information collected was based on a qualitative content analysis. The main research questions were:

1. For reviewers:

- What is the point of view of the reviewers about the professional roles, skills and competences of the team responsible for digital preservation of the repositories they audited?
- Which were the most difficult issues faced during the audit and how to prevent them through the delivery of appropriate training?
- Which are the professional competences of auditors and their training needs?

2. For repository team interviews:

- What are the roles, the competences and the contributions of each component of the repository team to the audit process?
- Which were the most difficult issues faced during the audit and how to prevent them through the delivery of appropriate training?
- Which are the professional competences of auditors and their training needs?
- What competences and skills are required after an audit to enable improvement based on results?

The questions include topics more directly focused on the audit process itself and while important to an understanding of the interview process answers to these are not be examined below but have been supplied to the WP33.

The first step in the analysis was to draw together the answers to these questions from the responses available and then draw out the most important conclusions, highlighted in the analysis below using text boxes which are numbered in progressive order.

The next step was to identify for each group a list of required skills corresponding to the conclusions derived from the analysis of the interviews (tables 2, 4, 6). Separate tables have been used to communicate the required competencies and training needs of the three different levels of repository personnel identified: top-level managers, operational managers, and operational staff. This was done as while the repository performance ultimately depends on the performance of its individual employees at all levels: educational approaches and training methods will be most successful if tailored to the specific identified requirements.

The final step was to derive, from the skills identified for each group, the list of corresponding training needs (tables 3, 5, 7). These lists are intended to give a comprehensive view of the required training needs of repository teams at different levels.

5.2 ANALYSIS OF THE INTERVIEWS

In order to clarify the roles involved in the process of the audits, in this document the term “repository DP (Digital Preservation) team” refers to the group of people that was interviewed during the process of the test audit by the auditors. It does not necessarily mean that the “DP team” is a stable group dedicated to the digital preservation in a certain repository, but they are more often individuals in various departments that contribute to the DP management process.

It is worth observing that these interviews are useful insofar as they provide guidance from a particular set of respondents. However, the researchers fully appreciate that a different set of interviewees could

produce a different set of results. Hence any analysis from these interviews needs to be calibrated against the other sources also examined in Sections Two, Three and Four.

5.2.1 Interviews with Reviewers

The first questions (**from 1 to 3**) were directed to acquire general information about the respondent and the audited repositories to gain some context for the review, regarding in particular the sector and the types of digital objects managed by the repository. Those interviewed each took part in one or more of the test audits and with the answers provided during the interview usually referring to all the repositories reviewed. It was decided this was an acceptable approach as the overall aim of the interviews was not the evaluation of difficulties for individual reviews but rather the identification of general training requirements.

The next set of questions (**from 4 to 7**) was about the staff belonging to the repository DP team. These questions were intended to provide the point of view of the reviewers about the team responsible for digital preservation of the audited repository: to understand how the team was composed in terms of professional roles, skills and competences; if the team had all of the competences they needed and reflected the necessary professional roles; what contribution each component made to the review process; and if some roles were missing from the group undertaking the review. Of this group of questions answers to the first two will be considered in this report. From the answers collected, even though it is difficult to come up with generalisations, it is possible to say that in most cases the repository team involved in the review was a good representation of the whole “chain of command”:

- *Senior management*, who had the overall responsibility for (and knowledge of) funding and its limitations, the strategic planning, succession planning, and the position of the repository within the larger organization;
- *Operational management*, composed of people who were responsible for running the archive; they provided information about the main functions of the repository system and explained how the processes were documented and verifiable;
- *Operational staff*, responsible for archival tasks and for the effective functioning of the repository (the day-to-day operations), dealing, for example, with acquisition and ingest of material: they provided useful information about the processes, through step by step procedures.

As well as these three main groups, it must be noted that other professional contributions were considered of great importance:

- *Domain experts* (mainly in science repositories), whose contribution was important because they are in contact with the user communities and they have a deep knowledge of general issues and the stored data
- *IT staff*: they provided in-depth knowledge of IT systems used including storage and security.

A general conclusion that can be reached is that there was no real consistency in relation to the link between job title and actual role across the organizations. Preservation tasks were managed by different members of staff at each repository. One of the purposes of the review is to make sure that all the levels connect together in an appropriate way, so when the senior managers speak about their policies, then auditors can check that those policies are really reflected in what happens day-to-day.

RI – The *senior management* should have team management skills to improve team effectiveness, develop team members, communicate effectively with people inside and outside the team.

R2 - The *senior management* has the overall responsibility for (and knowledge of) funding and its limitations, the strategic planning, succession planning, and the position of the repository within the larger organization.

The interviews also highlighted that the people who have made (or would make) decisions about preservation actions such as migration were largely based in the second level of the staff hierarchy identified above, the *operational management*. It is they who would be responsible for making decisions about issues such as migration of formats and implementation of metadata standards.

R3 – The *operational management* should have the ability to effectively implement standards.

R4 - *Operational management* is composed of people responsible for running the archive.

The reviews also showed that it is very important to also interview the *operational staff*, as they are the people responsible for carrying out preservation tasks including acquisition and ingest, and as such they should be expected to have detailed knowledge of these processes. For example, how are they applying the policies about format checking, anonymization or virus checking?

R5 – The *operational staff* should have the ability to effectively implement standards.

R6 - The *operational staff* are responsible for the day-to-day running of the archive and for the effective functioning of the repository.

One area where some difficulties were experienced during the reviews was in gaining satisfactory answers in relation to the legal aspects of digital preservation. In most cases the reviewers and the senior managers did know about this topic but in others there was insufficient information to accurately assess compliance due to a lack of competences and knowledge within both the audit and repository DP team.

R7 - The repository team should include at least one person who understands the relevant *legal issues*, in terms of country-specific legislation and legal constraints/requirements of the repository domain.

Another interesting observation from a respondent was about the repository DP team: none of the organizations the reviewers visited had what would be described as a coherent ‘digital preservation team’, instead the required digital preservation roles were filled by individuals in various departments who all contributed to the DP management processes, each doing only a small part of what is required to achieve long-term access. This was also often accompanied by a lack of awareness of the work being carried out by those in other departments and of the complete lifecycle management process. Moreover in large organizations, the DP team will likely not work with end users: which raises the question how do they identify the designated community and its needs? And how do they track their changing needs over time? The DP team may not be able to answer these questions fully.

R8 – It would be useful and productive if there was greater cohesion between the individuals working on various aspects of DP and for them to all have a solid understanding of the complete lifecycle of a digital object within their repository.

R9 – The repository team should include a *domain expert* who has a deep knowledge of general issues relating to the stored data, helps identify the designated community and its needs, performs the user requirements research and tracks the changes needs over time.

The reviewers also suggested that it is important during the review to go back to the functional model of OAIS and to make sure that they talk to people whose job roles cover all the *OAIS functional areas*. This breaks down differently in each organization reviewed: in some organizations individual members of staff will only work within 1 or 2 of the functional areas, in others they will work across more of the functions.

R10 - The repository team should include job roles which cover all the OAIS functional areas.

The next questions (**8 and 9**) were about the most difficult issues faced during the review and the ways to prevent them through the delivery of appropriate training. For the most part the information gathered from these questions is not relevant but one interesting conclusion in relation to training can be drawn. One of the most frequent problems reported by respondents was the needed for shared language and a common terminology: during the review, the repository staff often did not understand the question that was being asked by the reviewers. This particularly happened when the questions were related to the OAIS terminology: terms like “designated community” or “representation information” triggered lengthy discussions to establish a common understanding. There is, therefore, a clear need to establish a baseline of understanding about digital preservation and the related terminology amongst those working in the field.

R11 - The repository team at all levels should have knowledge and experience of OAIS concepts and general principles of preservation.

Regarding awareness of the designated community, it was reported that the operational staff who are working on the archive itself (the day-to-day operations) generally have a good awareness of the sources of the data, the ingested data, but are not necessarily aware of whom this data is being preserved for. Repositories should have a clear definition of their designated community in order to understand their commitment to them. Otherwise they will fail in their provision of services. The same issues came into focus when the auditors talked about “provenance” or “representation information”: preserving the bits is not necessarily enough in itself to preserve the information and that it may need to be supplemented with additional representation information. Repositories are often doing that anyway, they are gathering metadata and storing it but are perhaps not aware of how this fits within the basic OAIS concepts.

R12 - The operational staff of the repository team should have an overall view of the repository mission and commitment including the designated community.

R13 – Staff would benefit from training on user analysis.

R14 – Staff would benefit from training on issues relating to provenance.

R15 – Staff would benefit from training on metadata.

In terms of preparation the reviewers suggested that a self-assessment before the formal peer-review process would be useful: this would allow organizations to understand if they have the information available to give to the reviewers, to gather this information together and to do the monitoring between reviews to ensure improvement and continued compliance to the standards they are trying to meet.

R16 – Staff would benefit from training on self-assessment and review.

The other issues identified during the reviews related to insufficient information being made available about funding and succession planning. It was found that senior management were not sufficiently engaged with these topics to provide suitable answers and other staff did not possess enough knowledge or information to meet the needs of the auditors.

R17 – Training for *senior management* aimed at increasing engagement with DP issues would be useful.

R18 – Staff would benefit from training on cost analysis and budgeting.

R19 – Staff would benefit from training on succession planning.

The last questions (**from 10 to 12**) were aimed at evaluating the professional competences of reviewers and defining their training needs and as such answers to these questions are out with the scope of this report.

5.2.2 Interviews with the Repository Team

As in the previous set of interviews, the first questions (**1 to 3**) were structured to provide context and find out general information about the respondent and the audited repositories, in particular the organizational sector and the types of digital objects managed by the repository.

The subsequent questions (**from 4 to 6**), as with the reviewers, were about the roles, competences and contributions of the repository DP team to the review process. In this case the questions were aimed at providing the point of view of the repository team itself, by describing their perception of the issues that had already been mentioned by the auditors. Both of the respondents explained that their organization has already participated in self-assessment and in other types of external audit and that they have a large amount of expertise in this field.

The next question (**7**) was about the fundamental things that a member of the DP team inside the repository should know before participating in a review, in terms of general knowledge, regardless of the specific domain. The respondents considered their teams to be composed of expert people with all the necessary knowledge to cover the audit subjects. This was an interesting contrast to the responses provided by the auditors who had identified some specific gaps in the expertise available to answer their questions, particularly in relation to topics highlighted above.

The next question (**8**) was about the most difficult issues faced during the review. The respondents confirmed what the reviewers already reported: sometimes it was difficult to understand the requests and to interpret review criteria. Problems with communication and terminology were also apparent to the repository team, reinforcing the identified need for the development of a baseline of knowledge about DP across a repository team.

Another aspect that was stressed was that review does not involve only the technical side of repository processes but also the organizational issues and sometimes it was hard to find the right information to meet these requirements, in particular regarding financial and legal issues. For the respondents it was difficult to get access to some confidential financial documents, even though confidentiality agreements were made with all of the repositories. This again suggests that it would be productive for staff at different levels to have a broader understanding of the DP lifecycle and the management processes required to support it, particularly in relation to financial issues.

The next question (9) was about how some of the issues with the audit process could be avoided if appropriate training was provided for repository staff. The respondents identified two kinds of training they might find useful:

- Training courses focusing on the components of the auditing program, covering steps in the process, required documentation and how to implement recommendations;
- Training courses focusing on the issues surrounding advocacy and directed to senior management levels to help them understand the audit process and the positive benefits that can come from it.

R20 – The repository team should have a good knowledge of the audit process

R21 – Training for *senior management* should include content on the importance of audit and certification to the fostering of trust and also capacity enhancement.

The next question (10) was aimed at understanding the point of view of the respondents about the competencies of reviewers and as such is out with the scope of this report.

Questions 11 and 12 aimed at gathering information about the competences and skills required after a review to enable improvement based on the results. For the respondents these were difficult questions to answer as, at the time the interview took place, they didn't yet have the results of the test audits, so complete feedback could not be given. Among the ideas that were shared by the respondents were:

- Training sessions on the formal review process and information on previously audited repositories would be of value (covered above in **R20**);
- Training sessions could include content focused from the point of view of Management, Preservation, IT;
- Practically focused training sessions on general digital preservation issues, with tools and methods would be useful;
- Courses specifically addressed to management, with topics dealing with strategic planning, policies, financial issues etc, that could be short and very focussed, would help inform them of the benefits of audit and digital preservation more generally (covered above in **R17** and **R21**).

R22 – Training sessions are most useful when focused on staff performing similar roles.

R23 – There is a need for more practically focused training courses.

5.3 SKILLS, COMPETENCES AND TRAINING NEEDS

To provide a clearer view of the outputs derived from the analysis of the interviews in the previous section, a list of corresponding skills and competences has been identified below. From this for each list of skills and competences a corresponding table with appropriate training needs has been created. As it was identified that separate training would be desirable for staff at different levels within the organization (**R22**) the tables for repository staff have been split into three groups: senior management, operational management and operational staff.

5.3.1 Identification of Senior Management Skills, Competences and Training Needs

Table 2: Required Skills for Senior Managers

Reference input	Skills and competences
R1	General communication skills
R1	Team management skills
R2	Management and business continuity skills
R2, R18	Strategic planning competencies
R2, R19	Succession planning competencies
R7	General awareness of legal aspects and preservation policy framework
R11, R17	Knowledge of fundamentals of digital preservation
R11, R16, R20, R21	General awareness of audit and certification standards
R21	General awareness of the benefits of audit (advocacy)

Table 3: Training Needs for Senior Managers

Skills and competences	Training needs
General communication skills	General communication techniques
Team management skills	Team management
Management and business continuity skills	Business planning and cost analysis, sustainability, change management and risk mitigation
Strategic planning competencies	Strategic planning
Succession planning competencies	Succession planning
General awareness of legal aspects and preservation policy framework	Policy, regulatory and legal framework
Knowledge of fundamentals of digital preservation	Digital preservation general concepts
General awareness of audit and certification standards	Audit and certification standards
General awareness of the benefits of audit (advocacy)	Advocacy issues

5.3.2 Identification of Operational Management Skills, Competences and Training Needs

Table 4: Required Skills for Operational Managers

Reference input	Skills and competencies
R3, R16, R20	Knowledge and experience of audit and certification standards
R3, R4, R8	Knowledge of fundamentals of digital preservation including lifecycles
R3, R4, R23	Practical experience in digital preservation
R4, R11	Knowledge of digital preservation standards including OAIS
R7	General awareness of legal aspects and preservation policy framework
R9, R13	Knowledge of the repository's designated community and skills required to carry out user analysis
R14, R15	Knowledge of best practice for Ingest including provenance and metadata capture
R18	Knowledge of cost analysis and financial planning
R19	Knowledge of Succession planning

Table 5: Training Needs of Operational Managers

Skills and competences	Training needs
Knowledge and experience of audit and certification standards	Audit and certification standards
Knowledge of fundamentals of digital preservation including lifecycles	General digital preservation concepts
Practical experience in digital preservation	Digital object management Good practice, benchmarking and known problems Preservation strategies and methods
Knowledge of digital preservation standards including OAIS	Digital preservation standards
General awareness of legal aspects and preservation policy framework	Policy, regulatory and legal framework
Knowledge of the repository's designated community and skills required to carry out user analysis	User analysis
Knowledge of best practice for Ingest including provenance and metadata capture	Ingest processes and procedures Provenance and authenticity Capturing representation information
Knowledge of cost analysis and financial planning	Cost analysis and financial planning
Knowledge of succession planning	Succession planning

5.3.3 Identification of Operational Staff Skills, Competences and Training Needs

Table 6: Required Skills of Operational Staff

Reference input	Skills and competences
R5, R16, R20	Knowledge and experience of audit and certification standards and the audit process
R6	Digital object management
R5, R6, R8	Knowledge of fundamentals of digital preservation including lifecycles
R5, R6, R23	In-depth knowledge of methods and practices of digital preservation
R5, R11	Knowledge of digital preservation standards including OAIS
R7	General awareness of legal aspects and preservation policy framework
R9, R13	Knowledge of the repository's designated community and skills required to carry out user analysis
R14, R15	Knowledge of best practice for Ingest including provenance and metadata capture

Table 7: Training Needs of Operational Staff

Skills and competences	Training needs
Knowledge and experience of audit and certification standards and the audit process	Digital preservation standards
Digital object management	Digital object management Good practice, benchmarking and known problems
Knowledge of fundamentals of digital preservation including lifecycles	Digital preservation general concepts
In-depth knowledge of methods and practices of digital preservation	Authenticity and provenance, digital rights management, capture of representation information, Ingest processes and procedures, file format characterisation
Knowledge of digital preservation standards including OAIS	Practical preservation strategies/methods
General awareness of legal aspects and preservation policy framework	Audit and certification standards

5.4 TRAINING NEEDS HIGHLIGHTED BY PEER REVIEW – CONCLUSIONS

As can be seen from the tables above it is possible to derive some clear conclusions about the skills required by repository staff and related training needs as highlighted by APARSEN's peer review process. A number of specific topics have been identified, many of which echo conclusions reached in earlier sections of this report, for example legal aspects of digital preservation and ingest processes and procedures. The call for training with a more practical edge to compliment theory is also reiterated by the interviewees. Some generic topics, such as management and communication skills emerge through the analysis, though it is hard to see how APARSEN could justify spending significant resources on these.

Perhaps the most significant conclusions to be drawn from the answers provided above is the identification of a need to have training tailored to staff working at different levels within a repository hierarchy. While there are some topics repeated between levels they were often mentioned at different times with reference to a particular group. For example, it was felt that it was important for senior managers to understand the broad principles of digital preservation but that operational managers and operational staff understand processes and procedures at a more granular level. For example, in relation to repository enhancement, it was suggested that one of the most important skills for senior managers to have was the ability to act as an advocate for the process within their organization whereas other staff should have more detailed knowledge of practical process.

APARSEN is well placed to meet the training needs of staff at all levels and this identified need for tailored training must be taken into consideration when planning the Network's training output.

6 IDENTIFIED TRAINING NEEDS AND LINKS TO APARSEN RESEARCH AND INTEGRATION ACTIVITIES

The previous sections of this report have deployed a number of different methods to identify training needs relating to digital preservation including analyses of:

- The diverse sectors that are stakeholders in digital preservation;
- Professional development frameworks;
- Initiatives relating to digital preservation training;
- Current training provision;
- Participants' views of training;
- Focus group data relating to job profiles and related training needs;
- Emerging standards and professional practice; and
- Interviews with reviewers and repository managers that participated in peer review as part of APARSEN WP33.

Each analysis produced a number of conclusions relating to gaps in current training provision and skill sets within the field, as well as suggestions for improvement to and best practice for the delivery of training. Some of these conclusions and recommendations have been clearly stated by participants whereas others have been derived from analysis of information resources relating to training or repository improvement. These analyses have largely produced complimentary results and several recurring themes have presented themselves. These included the need for more advanced and practically focused training as well as a repeated desire for training on topics such as legal issues, capture of representation information, cost analysis and digital object lifecycles.

This section will provide a summary of these topics and recommendations before relating these, where possible, to the research and integration activities of APARSEN, clearly identifying where the Network is best placed to fill gaps in training provision.

6.1 IDENTIFIED TOPICS AND RECOMMENDATIONS FOR TRAINING

The analyses in the earlier sections of this report identified a number of topics relating to digital preservation that are either not currently covered by training or covered but only in theory or at a basic level. Most topics arose repeatedly throughout the analyses reinforcing the importance of these topics to stakeholders in digital preservation.

The analyses point to a requirement amongst stakeholders for training covering both the complete digital preservation lifecycle, but at a more advanced level than is currently available, and individual topics of particular interest covered in more detail. In relation to individual topics the gaps identified include requirements for training on each of the topics listed below in Table 8.

Table 8: Identified Training Needs by Topic

Topic	Sub-Sections Highlighting Training Needs													
	1.3.1	1.3.2	1.3.3	1.3.4	3.2.1	3.3	4.2	4.3	4.5	4.6	4.7	5.3.1	5.3.2	5.3.3
Preservation Planning and Services	✓	✓	✓	✓	✓	✓		✓		✓			✓	✓
Discovery and Access (Package Description)	✓	✓			✓	✓							✓	✓
Annotation and Data Quality	✓		✓					✓						
Representation Information and Preservation Description Information					✓	✓		✓					✓	✓
Accessibility (incl. interoperability and usability)	✓													✓
Sustainability	✓		✓		✓							✓		
Digital Preservation Policy	✓	✓		✓								✓	✓	
Enabling Novel Uses of Archived Data and Releasing Value	✓		✓	✓		✓			✓				✓	
Legislation and Regulatory Frameworks		✓	✓									✓	✓	
Digital Rights Management		✓	✓			✓							✓	✓
User Analysis (understanding Designated Communities)		✓				✓							✓	✓
Authenticity and Provenance		✓		✓				✓					✓	✓
Establishing Trust (incl. Peer Review and Audit)				✓			✓			✓		✓	✓	✓
Digital Preservation Lifecycles	✓	✓	✓	✓		✓			✓			✓	✓	✓
Standards for Digital Preservation							✓				✓		✓	✓
Advocacy						✓						✓		
Selection and Appraisal				✓		✓							✓	✓
Change Management						✓						✓		
Succession Planning						✓						✓		✓
Requirements Analysis						✓							✓	
File Format Characterisation						✓		✓					✓	✓
Quality Assurance						✓							✓	✓
Risk Mitigation						✓	✓					✓	✓	
Business Planning and Cost Analysis			✓									✓	✓	

Table 8 does, however, only represent the topics of interest as derived from the analyses in this report and not the gaps that exist in current training provision. To identify the gaps there must be a comparison of these topics with the data presented in section 3.1 and this is detailed below in Table 9.

Table 9: Identification of Training Gaps in Relation to Specific Topics

(Training Coverage - Green: topic well covered, Amber: Some training coverage, Red: limited or no training available)

Topic	Training Coverage	Additional Notes
Preservation Planning and Services		A good number of courses but mostly project related or strategically focused
Discovery and Access (Package Description)		Current training courses tend to focus on ingest and policy and only limited material on access
Annotation and Data Quality		Not specifically covered by current training provision
Representation Information and Preservation Description Information		Previously well covered but only limited coverage during the period analysed: limited to conferences and subsidiary workshops.
Accessibility (incl. interoperability and usability)		Some coverage of integration of tools but broader issues not specifically covered by current training provision.
Sustainability		Strategic planning training tends to focus on the development of new repositories, cost modelling or preservation planning but less on 'business' issues pertinent long-term preservation such as evolving business models. Connected to 'releasing value'
Enabling Novel Uses of Archived Data and Releasing Value		Not obviously covered by current provision.
Digital Preservation Policy		A core component of most general courses
Legislation and Regulatory Frameworks		General concepts well covered – need for training tailored to specific sectors
Digital Rights Management		Covered as an aspect of many training courses but requires local practical focus so hard to recycle
User Analysis (incl. understanding Designated Communities)		Training on the analysis of user behaviour is available but is seldom linked to preservation and not specifically embedded in current training provision

Topic	Training Coverage	Additional Notes
Authenticity and Provenance		The subject of much research and is evident in conference presentations and occasional workshops but little practical training
Establishing Trust (incl. Peer Review and Audit)		Covered by courses relating to individual projects but no regular training programme relating to the newly established standards (i.e. DSA, ISO16363)
Digital Preservation Lifecycles		A core component of most general courses
Standards for Digital Preservation		A core component of most general courses
Advocacy		Covered by a small number of courses in the analysis
Selection and Appraisal		Covered by a small number of courses in analysis
Change Management		On-going organisational management issues are implied in many courses but not addressed directly
Succession Planning		Some coverage in relation to strategic planning
Requirements Analysis		Planning for digital preservation is well covered. Particular requirements analysis techniques may be deemed out with the scope of regular digital preservation training.
File Format Characterisation		Regularly covered in relation to available tools and an active topic in conferences.
Quality Assurance		Some training around QA for digitisation and validation of preservation plans, but on-going quality assurance within preservation facilities are poorly represented.
Risk Mitigation		Some coverage in the past by DRAMBORA, and some anticipated from TIMBUS in relation to Business Continuity management but no current provision
Business Planning and Cost Analysis		Mostly focused on results of particular projects

A comparison Tables 8 and 9 allows a list of priorities for training to be identified. These can be defined as follows:

- Preservation Planning and Services
- Discovery and Access (Package Description)

- Annotation and Data Quality
- Representation Information and Preservation Description Information
- Accessibility (Interoperability and Usability)
- Enabling Novel Uses of Archived Data and Releasing Value
- Sustainability
- Legislation and Regulatory Frameworks
- Digital Rights Management
- User Analysis (incl. Understanding Designated Communities)
- Authenticity and Provenance
- Establishing Trust (incl. Peer Review and Audit)
- Selection and Appraisal
- Change Management
- Succession Planning
- Quality Assurance
- Risk Mitigation
- Business Planning and Cost Analysis

Alongside the topics identified above, a number of recommendations for the development and format of digital preservation training can be elicited from the analyses described by this report. These recommendations can be summarized as follows:

1. There is a need for training that marries together theoretical concepts and practical experience. Such training should be focused around real case-studies as they are particularly useful. Although there are no perfect solutions in digital preservation it is important to get started and learn from the work and experience of others.
2. Participants prefer training that includes practical exercises and worked examples, learning through doing rather than just passively listening to information.
3. It is important to have a mix of courses that cover the broad issues surrounding digital preservation and more focused, specialized courses covering only one or a small number of related issues in more detail.
4. It would be useful to have training targeted at specific audiences, whether that is for a particular sector (i.e. cultural heritage organizations or the aerospace industry) or for staff within a particular role (i.e. senior management, digital curators or IT staff). Such targeted training would allow greater flexibility of content to increase relevance and better meet the needs of the audience.
5. There is a need for continuing professional development within the digital preservation community for staff at all levels, even those considered experts.
6. There is a diversity of needs and the training should take account of the mix of sectors, skills and expectations of participants.
7. Training courses must be responsive to the development of new standards in a field where they develop rapidly and at times organically. It is essential that course content remains current and authoritative.
8. Efforts should be made to embed elements of digital preservation training in courses with a broader information or risk management focus.

Together the topics and recommendations can provide clear objectives for any initiative wishing to develop training courses. For APARSEN the next step is to link these to the competencies available with the Network of Excellence and the research and integration activities it is undertaking.

6.2 POTENTIAL TRAINING PROVISION FROM APARSEN

Taking the outcomes of Section 6.1 it is now possible to look more clearly at how APARSEN is placed to fill the training gaps identified. As stated in Section One, WP43 benefits from close collaboration with a number of other work packages within the Integration and Spreading Excellence Streams that will help facilitate the development and presentation of training. These benefits are, however, more of a practical and procedural nature in relation to the promotion and staging of training courses and it will be assumed that these are implicit and will not be described in detail here. Rather this section will focus on the work with the Research Stream of APARSEN and how the outputs of these work packages can be harnessed to provide content and support for the development of training courses covering the topics highlighted above.

As a Network of Excellence APARSEN offers both a great depth of knowledge and skill relating to digital preservation as well as the opportunity to undertake innovative research and integration activities, bringing together leaders in the field to work collaboratively. It follows, therefore, that APARSEN is well placed to provide training courses that are high quality and provide participants with access to information about systems, standards and best practice at the leading edge of digital preservation work.

As mentioned above two potential types of courses have been identified as desirable through the analyses undertaken. The first being a course that provides an overview of the complete digital preservation lifecycle and the second being courses that focus on one issue (or a small group of related issues). The topics identified above must be considered in relation to both of these types of course.

With regards to the first type of course, APARSEN is well placed to develop a course covering the complete digital preservation lifecycle at a more advanced level than those currently available and with more practically focused training and case studies by drawing on the collective experience of the partners within the Network of Excellence and, potentially, affiliated projects. Many of the partners within the Consortium can not only point to their own digital preservation systems as being innovative and examples of best practice but are also among the leaders in digital preservation within their respective countries and/or fields. Many of the work packages will also be able to directly provide content to such a course and work being carried out by work packages such as WP11 (Common Vision) and WP13 (Coordination of Common Standards) can ensure that training provision from APARSEN remains both relevant and at the forefront of developments.

When considering the second type of training, focused courses examining one particular topic (or a small number of related topics), it is essential to return to the list of topics identified above and compare these with the topics covered by the research and integration work packages within APARSEN. Such a comparison will clearly identify where APARSEN can fill gaps in current training provisions. This comparison is detailed in Table 10 below.

The analyses of this report, particularly those carried out in Sections 3.1.2 and 5.3, have also clearly identified the need for training tailored to particular job roles. For each topic listed in Table 10 potential audiences will also be identified according to the roles defined in Section 5.2.1.

Table 10: Training Topics and Links to APARSEN Research and Integration

Topic	APARSEN Topic Areas and WPs of Particular Interest	Potential Audiences (by role)
Preservation Planning and Services	Sustainability, Usability, WP21 (Preservation Services)	Operational Managers, Operational Staff
Discovery and Access (Package Description)	Access, Usability, WP13 (Coordination of Common Standards)	Operational Staff
Annotation and Data Quality	Trust, Usability, WP26 (Annotation, Reputation and Data Quality), WP13	Operational Managers, Operational Staff
Representation Information and Preservation Description Information	Sustainability, Usability, WP21 (Preservation Services)	Operational Managers, Operational Staff
Accessibility (incl. Interoperability and Usability)	Usability, WP25 (Interoperability and Intelligibility)	Operational Staff
Sustainability	Sustainability	Senior Managers
Legislation and Regulatory Frameworks	WP35 (Data Policies and Governance)	Senior Managers, Operational Managers
Digital Rights Management	Access, WP31 (Digital Rights and Access Management)	Operational Managers, Operational Staff
User Analysis (incl. understanding Designated Communities)	Usability, Access	Operational Managers, Operational Staff
Authenticity and Provenance	Trust, WP24 (Authenticity and Provenance)	Operational Managers, Operational Staff
Establishing Trust (incl. Peer Review and Audit)	Trust, WP33 (Peer Review and 3 rd Party Certification of Repositories)	Senior Managers, Operational Managers, Operational Staff
Selection and Appraisal	WP36 (Business Cases)	Operational Staff
Change Management	Sustainability	Senior Managers
Succession Planning	Sustainability	Senior Managers, Operational Managers
Quality Assurance	Trust, WP24, WP26, WP33	Operational Managers, Operational Staff
Risk Mitigation	Sustainability	Senior Managers
Business Planning and Cost Analysis	Sustainability, WP32 (Cost/Benefit Data Collection and Modelling), WP36 (Business Cases)	Senior Managers, Operational Managers

An examination of Table 10 shows that APARSEN is well placed to fill many of the gaps in current training provision and should thus seek to develop training on a selection of topics from the list above. Ultimately the number of topics selected will reflect what is achievable within the resources available to WP43 and the availability of results at the time of training development. Also, although the topics have been separated here there are a number of complimentary issues, such as preservation planning and services and representation information and preservation descriptions information, which may be drawn together for a training course around a particular theme.

Thus the main aims of the APARSEN training efforts will be to develop two strands of training: one longer training course covering the complete digital preservation lifecycle at a more advanced level than the courses currently offered and with more focus on practical implementations and case studies; and a number of shorter courses focusing on some of the relevant topics identified in Table 9. It will, of course, be possible to share content between the two strands, making the development of courses as efficient and consistent as possible. This approach will satisfy the greatest number of identified requirements from the analyses described within this report.

It is evident from Table 10 that there is a reasonably even spread of coverage of the topics listed in relation to APARSEN's research streams. It would, therefore, seem an easy conclusion to draw that the most efficient use of APARSEN's resources will be to organise the majority of the shorter, more focused courses around the four research streams. Thus the project would provide training on a number of topics from the table above arranged around the themes of 'Trust', 'Sustainability', 'Usability' and 'Access'.

There are also a number of the topics in Table 10 which have been identified as a priority for the Senior Manager audience group. This is an audience that is notoriously difficult to engage due to pressures of time and responsibility, so training that is easily accessible is preferred. APARSEN has a commitment to the delivery of online training and the production of such training materials for those in Senior Management roles could be a key step in fulfilling this commitment.

7 CONCLUSIONS AND RECOMMENDATIONS

The main goal of the deliverable D43.1 is to determine the priorities that will frame and guide training initiatives in digital preservation, in particular providing recommendations for the training modules to be developed in the course of APARSEN, ensuring that our training provision is based on evident need, complements current provision, and is aligned with the outcomes from APARSEN's research work packages. A number of gaps have been identified and a number of challenges can be recognised.

One of the most significant issues that have affected the provision of training in digital preservation is the result of success. Research in digital preservation has moved quickly and so it is difficult for training provision to keep pace with state of the art. The speed of development also makes it hard to reflect on real practical experience which directly affects the practical pedagogy of the sector: there are too few worked examples. This is compounded by the diverse and growing needs of the various sectors with an interest in digital preservation and a degree of fluidity about emerging job profiles. APARSEN needs to respond to this challenge by providing courses that draw from the cutting edge of research but which are also comprehensible by a large and rapidly growing audience with diverse needs.

Although the task is challenging, there are existing resources which APARSEN could use or model as best practice. This includes the European Commission's wider best practice framework for vocational education and higher education (although the latter is more properly the subject of WP42 than WP43). It also includes a range of initiatives co-funded by the European Commission and by a range of national and international collaborations. Consequently, in developing its training output, APARSEN should form strategic relationships with existing providers, modelling itself to best practice in vocational education.

APARSEN has analysed the topics of training currently available, alongside participants' commentary on current training provision and employers' and managers' expectations about the skills they expect their staff to bring. This threefold analysis highlights some clear gaps within the training provision and ancillary expectations from participants about teaching methods which are considered to be successful. Current gaps include: annotation and data quality; accessibility; sustainability of preservation services; releasing value from data; tracking the designated community; managing authenticity; change management in preservation; and risk management in preservation. In addition there is a need for different types of training for different functional levels within an organisation, and participants have clear views about how to fit material to delivery methods.

Standards can be understood as a particular encapsulation of good practice and therefore are relevant to APARSEN's training activities, especially when working across a diverse audience. APARSEN therefore should where possible connect training to standards in order to provide a robust and mutually comprehensible framework for training.

The analysis of current repository practices in Section Five gives a direct understanding of how training can link to repository improvement, and it makes sure that the requests for training are not based solely on what practitioners think they need without a strategic view to what the training will deliver in terms of practical improvement. Moreover, to be effective, APARSEN's training needs to be tailored to an individual's role within a repository.

The conclusions of the previous sections are all brought together in Section Six and analysed in relation to APARSEN's research and integration activities. This provides a map of topics which APARSEN is well placed to fill. APARSEN is uniquely well placed to link research and training in a very particular way.

These conclusions provide clear recommendations for training within APARSEN they can be summarized as follows:

- APARSEN is well placed to link research with training and should seek to do so in order that training remains current in a fast-moving field.
- APARSEN should target training to different job roles in preservation and should ensure that as many different sectors are able to participate.
- APARSEN should build on the work of previous projects with an interest in digital preservation training and should seek to form alliances with existing or emerging training providers.
- APARSEN should offer a mix of in-depth courses covering the complete digital preservation lifecycle and shorter, focused training on a single or small number of related subjects.
- APARSEN's training should be a good mix of theoretical concepts and practical advice and experience with case studies and worked examples being key teaching tools.
- APARSEN's training should go beyond the introductory level covered by much existing training. This will help to start build a framework for continuing professional development in digital preservation.
- APARSEN's training should provide the basis for repository improvement to ensure that courses have practical impact.
- APARSEN's training should be informed by emerging standards as codification of good practice.
- APARSEN training should prioritise themes with little or no current provision, especially those which are connected to APARSEN Research and integration activities, such as: annotation and data quality; accessibility; sustainability of preservation services; releasing value from data; tracking the designated community; managing authenticity; change management in preservation; and risk management in preservation.

The purpose of the APARSEN Network of Excellence is to reduce the emerging fragmentation that is present within the rapidly growing digital preservation community. It provides a forum and a vehicle for the exchange of ideas and is working towards the development of a common vision for digital preservation research. By developing training that meets demonstrable needs, is responsive to the views of participants and has a strategic fit to wider policy and practice, the APARSEN Training Courses will provide a strong foundation to engage an ever larger and ever more diverse community, with mutually comprehensible concepts and models for ensuring permanent access.

REFERENCES

- Abbott, D. (2008), *What is digital curation?* Digital Curation Centre: Edinburgh, UK. Retrieved October 28, 2010, from <http://www.dcc.ac.uk/resources/briefing-papers/introduction-curation/what-digital-curation>.
- ADS-STAN 2007 LOTAR - Long-term Archiving and Retrieval of digital technical product documentation, such as 3D-CAD and PDM data, online at: <http://www.asd-stan.org/Lotar.html>
- Audit and Certification of Trustworthy Digital Repositories, available from <http://public.ccsds.org/publications/archive/652x0m1.pdf> and will also be available from www.iso.ch as ISO 16363.
- Beagrie, N. and Jones, M. (2001) *Preservation and Management of Digital Materials: A Handbook*, The British Library, London, also available online at: <http://www.dpconline.org/advice/preservationhandbook>
- Becker, C Kulovits, H, Rauber, A and Hofman H (2008) Plato: a service-oriented decision support system for preservation planning, in *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries (JCDL'08)*, Pittsburgh, Pennsylvania , June 16-20
- Botticelli, P., Fulton, B., Pearce-Moses, R., Szuter, C, Watters, P., (2011), *Educating Digital Curators: Challenges and Opportunities*, The International Journal of Digital Curation, Issue 2, Vol. 6, available at: <http://www.ijdc.net/index.php/ijdc/article/view/188/253>
- Boyle, F, Eveleigh, A and Needham, H 2008 *Report on the Survey Regarding Digital Preservation in Local Authority Archive Services*, Digital Preservation Coalition, York, online at <http://www.dpconline.org/docs/reports/digpressurvey08.pdf>
- Buneman, P., Müller, H. (2009), *Curating the CIA World Factbook*, in “The International Journal of Digital Curation”, available at: <http://www.ijdc.net/index.php/ijdc/article/viewFile/132/171>
- Burnhill, P. (2004), *Speech for the Launch of the Digital Curation Centre (DCC)*, Edinburgh, 5 November 2004, available at: http://www.nesc.ac.uk/talks/514/Speech_Peter_Burnhill.pdf
- Cohen D., Crabtree B. (2006), *Qualitative Research Guidelines Project*, available at: <http://www.qualres.net/HomeSemi-3629.html>
- Commission on Preservation and Access / Research Libraries Group 1996 *Preserving Digital Information: Report of the Task Force on Archiving of Digital Information*, RLG, online at: <http://www.oclc.org/research/activities/past/rlg/digpresstudy/final-report.pdf> last checked 06/07/2011
- Condon, F et al 1999 *Strategies for digital data: findings and recommendations from digital data in archaeology: a survey of user needs*. York: Archaeology Data Service also available online at: <http://ads.ahds.ac.uk/project/strategies/>
- Cornell University Library (2003), *Digital preservation management: Implementing short-term strategies for long-term problems* [online], available at: http://www.library.cornell.edu/iris/tutorial/dpm/eng_index.html
- CRL, The Center for Research Libraries, OCLC, Online Computer Library Center, (2007), *TRAC, Trustworthy Repositories Audit & Certification: Criteria and Checklist*, available at: <http://www.crl.edu/PDF/trac.pdf>
- Dallas, C., Doorn, P. (2009), *Report on the workshop on digital curation in the human sciences at ECDL 2009*. Corfu, 30 September, 1 October 2009, in «D-Lib Magazine» 15, n.11-12, available at: <http://www.dlib.org/dlib/november09/dallas/11dallas.html>

- DPC 2009 *Strategic Plan 2009-11*, Digital Preservation Coalition, York online at: http://www.dpconline.org/component/docman/doc_download/741-dpcstrategicplan2009-12
- DCC, DPE, NESTOR, CRL, *CORE Requirements for digital archives*, Chicago, <http://www.crl.edu/archiving-preservation>
- DIN 31644: 2012 *Information und Dokumentation - Kriterien für vertrauenswürdige digitale Langzeitarchive*
- Directorate for Science, Technology and Industry - Committee for information, computer and communications policy (2006), *Digital broadband content: public sector information and content*, available at: <http://www.oecd.org/dataoecd/10/22/36481524.pdf>.
- DPE DigitalPreservationEurope (2006), *Outline of training principles and objectives*, Deliverable 2.1, available at: http://www.digitalpreservationeurope.eu/publications/reports/DPE_Outline_training_Deliverable_2_1.pdf
- DSA Data Seal of Approval, (2008), <http://www.datasealofapproval.org>
- Engelhardt, C, Strathmann S and McCadden K (2012) *DigCurv: report and analysis of the Survey of Training Needs*, DigCurv Project online at: <http://www.digcurv-education.org/eng/Resources/Report-and-analysis-on-the-training-needs-survey>
- ERPANET, (2004), *The Role of Audit and Certification in Digital Preservation*, Final report of the workshop held in Antwerpen, 14 - 16 April 2004
- European Commission (2010), *Europe 2020, An agenda for new skills and jobs*, available at: <http://ec.europa.eu/europe2020/tools/flagship-initiatives>
- The European Parliament and Council of the European Union (2009), *Recommendation of the European Parliament and of the council on the establishment of a European Credit System for Vocational Education and Training (ECVET)*, Official Journal of the European Union, 8 July 2009, available at: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2009:155:0011:0018:EN:PDF>
- Gartner, Inc. (NYSE: IT), (2010), *Gartner Identifies Four Information Management Roles IT Departments Need to Remain Effective*, Press Release, available at: <http://www.gartner.com/it/page.jsp?id=1282513>
- Gartner, R and Lavoie, B (2005) *Preservation Metadata*, DPC Technology Watch report 05-01, online at: http://www.dpconline.org/component/docman/doc_download/88-preservation-metadata
- Guercio, M., (2005), *Records and archival education as fundamental requirement for digital society citizenship: challenges to face, risks to govern, knowledge to re-build and implement*, Budapest, DLM Forum. 5-7 October 2005.
- Guercio, M., (2011), *Le discipline del documento e l'innovazione tecnologica nelle iniziative di formazione degli archivisti*, in Digitalia, Anno Vi, Numero 1, available at: <http://digitalia.sbn.it/upload/documenti/DIGIT%2012011.pdf?l=it>
- Hallam, G., Partridge, H., (2006), *Students, teachers and professionals: educating the new information professionals. In Transforming IT education: promoting a culture of excellence*, Santa Rosa, California: Informing Science Press, 2006.
- Higgins, S (2008) *The DCC Curation Lifecycle Model*, in *International Journal of Digital Curation* 3, online at: <http://www.ijdc.net/index.php/ijdc/article/view/69>
- IASA, 2009. Technical Committee, Kevin Bradley, (Ed.) *Guidelines on the Production and Preservation of Digital Audio Objects (IASA-TC 04 Second edition)* Canberra, IASA. <http://www.iasa-web.org/audiopreservation-tc04>

- ISO 10303-1:1994 *Industrial automation systems and integration Product data representation and exchange - Overview and Fundamental Principles*
- ISO 14721:2003 - *Space data and information transfer systems – Open archival information system – Reference model* - OAIS www.iso.org/iso/catalogue_detail.htm?csnumber=24683.
- ISO 14721:2003 - *Space data and information transfer systems – Open archival information system – Reference model* - OAIS, www.iso.org/iso/catalogue_detail.htm?csnumber=24683.
- ISO 15489-1: 2001 *Information and documentation – Records management. Part 1: General.*
- ISO 15489-2: 2001 *Information and documentation – Records management. Part 2: Guidelines.*
- ISO 16363:2011 *Audit and certification of trustworthy digital repositories*
<http://public.ccsds.org/sites/cwe/rids/Lists/CCSDS%206521R1/Attachments/652x1r1.pdf>.
- ISO/DIS 16919, *Requirements for Bodies Providing Audit and Certification of Candidate Trustworthy Digital Repositories*, available at: <http://public.ccsds.org/publications/archive/652x1m1.pdf>, and will also be available from ISO as ISO 16919.
- ISO/DIS 16363, *Space data and information transfer systems – Requirements for Audit and Certification of Trustworthy Digital Repositories*,
<http://public.ccsds.org/sites/cwe/rids/Lists/CCSDS%206521R1/Attachments/652x1r1.pdf>
- ISO/IEC 17021:2011, *Conformity assessment – Requirements for bodies providing audit and certification of management systems.*
- ISO / IEC 17799: 2000 *Information technology - Code of practice for information security management*
- ISO/IEC 19011:2002, *Guidelines for quality and/or environmental management systems auditing.*
- ISO/IEC 27001: 2005 *Information technology - Security techniques - Information security management systems - Requirements*
- ISO/IEC 27002: 2005 *Information technology --Security techniques - Code of practice for information security management*
- ISO 31000: 2009 *Risk Management Principles and Guidance*
- James, H (2003) *AHDS Notes on Writing the AHRC Technical Appendix*, online at: <http://www.ahds.ac.uk/creating/information-papers/writing-appendix/index.htm>
- Karvelyte V, Klingaite N, Kupriene J, Molloy L, Snow K and Gow A (2012) *DigCurV: Report on baseline survey into training opportunities and evaluation framework*, DigiCurV project online at: <http://www.digcur-education.org/eng/content/download/3033/41768/file/D2.1.1%20DigCur-survey-into-training-opportunities.pdf>
- Kilbride, W and Todd, M (2010) *The Digital Preservation Roadshow 2009-10: The Incomplete Diaries of Optimistic Travellers*, in *Ariadne* 62 online at: <http://www.ariadne.ac.uk/issue62/dp-rdshw-rpt/>
- Lee, C. (2009), *Curation Junction, What's your Function? Defining What it Means to Do Digital Curation in Order to Teach It*, in *Proceedings of the 2009 Annual Conference of the Association for Library and Information Science Education (ALISE)*.
- nestor 2006 *Catalogue of Criteria for Trusted Digital Repositories Version 1* (draft for public comment), nestor Working Group Working Group Trusted Repositories Certification, online at: http://files.d-nb.de/nestor/materialien/nestor_mat_08-eng.pdf
- nestor 2009 *Catalogue of Criteria for Trusted Digital Repositories - Version 2*, nestor Working Group Working Group Trusted Repositories Certification, online at: http://files.d-nb.de/nestor/materialien/nestor_mat_08_eng.pdf

- Madden, L. (2007), *Digital Curation at the Library of Congress: Lessons Learned from American Memory and the Archive Ingest & Handling Test*, Library of Congress, DigCCurr project, available at: http://www.ils.unc.edu/digccurr2007/slides/madden_slides_6-2.pdf
- Minutes, Technical Workshop for Audit and Certification of Trusted Repositories, Luxembourg, on 08 July 2010 <http://wiki.digitalrepositoryauditandcertification.org>
- National Science Foundation, National Science Board (2005), *Long-Lived Digital Data Collections Enabling Research and Education in the 21st Century*, available at: <http://www.nsf.gov/pubs/2005/nsb0540/>
- O'Neill, M. (2006) Essentialism, adaptation and justice: towards a new epistemology of museums in *Museum Management and Curatorship* 21, 95-116
- OECD Organization for Economic Co-operation and Development (2006), *Digital broadband content: public sector information and content*, Working Party on the Information Economy.
- PARADE Partnership for Accessing Data in Europe (2009), *Strategy for a European Data Infrastructure*, <http://www.csc.fi/english/pages/parade>
- PARSE.Insight, (2009), *Insight into digital preservation of research output in Europe – Survey Report*, available at: http://www.parse-insight.eu/downloads/PARSE-Insight_D3-4_SurveyReport_final_hq.pdf
- PARSE.Insight, (2010), *Insight into digital preservation of research output in Europe - Insight Report*, available at: http://www.parse-insight.eu/downloads/PARSE-Insight_D3-6_InsightReport.pdf
- PREMIS, (2011), *Data Dictionary for Preservation Metadata*, version 2.1, www.loc.gov/standards/premis/v2/premis-2-1.pdf
- PREMIS, (2008), *PREsevation Metadata: Implementation Strategies*, available at: <http://www.loc.gov/standards/premis>
- RECOUP, (2008), *Semi-structured interviews*. Research Consortium on Educational Outcomes and Poverty, available at: http://manual.recoup.educ.cam.ac.uk/wiki/index.php/Semi-structured_interviews
- RLG/OCLC Working group on digital archive attributes, (2002), *Trusted digital repositories. Attributes and responsibilities. An RLG-OCLC Report*, RLG, Mountain View, CA, available at: <http://www.oclc.org/research/activities/past/rlg/trustedrep/repositories.pdf>
- RLGNARA Task force on digital repository certification, (2005), *Draft Audit Checklist for Certifying Digital Repositories*, available at: <http://web.archive.org/web/20050922172830/http://www.rlg.org/en/pdfs/rlgnara-repositorieschecklist.pdf>
- Rosenthal, D. (2005), *Requirements for Digital Preservation Systems: A Bottom-Up Approach*, D-Lib Magazine 11:11, available at: www.dlib.org/dlib/november05/rosenthal/11rosenthal.html
- Ross, S. and Hedstrom, M. (2003), *Invest to Save: Report and Recommendations of the NSF-DELOS Working Group on Digital Archiving and Preservation*, available at: http://eprints.erpanet.org/94/01/NSF_Delos_WG_Pres_final.pdf
- Strodl, S. Petrov, P. and Rauber, A 2011 *Research on Digital Preservation within projects co-funded by the European Union in the ICT Programme*, online at <http://www.ifs.tuwien.ac.at/~strodl/paper/Report%20-%20Research%20on%20Digital%20Preservation.pdf>
- Swan, A. and Brown, S. (2008), *The skills, role and career structure of data scientists and curators: An assessment of current practice and future needs*, Technical Report, School of Electronics & Computer Science, University of Southampton, available at: <http://eprints.ecs.soton.ac.uk/16675/>

The National Archives, (2011), *Understanding digital continuity*, version 1.2. Retrieved January 19, 2012, from <http://www.nationalarchives.gov.uk/documents/information-management/understanding-digital-continuity.pdf>

Trant, J. (2009), *Emerging Convergence? Thoughts on museums, archives, libraries and professional training*, available at: <http://www.archimuse.com/papers/trantConvergence0908-final.pdf>

UK Data Archive (2009) *Managing and Sharing Research Data: A Best Practice Guide for Researchers*, University of Essex, Colchester

van der Kniff, J and Wilson, C (2011) *Evaluation of Characterisation Tools Part 1: identification*, SCAPE Project, online at: http://www.scape-project.eu/wp-content/uploads/2012/01/SCAPE_PC_WP1_identification21092011.pdf

Waller, M and Sharpe, R 2006 *Mind the Gap: Assessing Digital Preservation Needs in the UK*, Digital Preservation Coalition, York, online at: <http://www.dpconline.org/advocacy/mind-the-gap>

Wright, R (2012) *Preserving Moving Pictures and Sound*, DPC technology Watch report 12-01, online at: <http://www.dpconline.org/advice/technology-watch-reports>

ANNEX A: INTERVIEW STRUCTURES

INTERVIEW FOR AUDITORS:

Abbreviations: d.p. = digital preservation

1. General information about the respondent:

Name: _____

Institution: _____

Job Title: _____

Phone: _____

Email: _____

2. About the organizations you have recently audited:

Organization sector:

- ☐ Academic/research
- ☐ Business
- ☐ Government
- ☐ Public sector
- ☐ Private
- ☐ Other (Please specify) _____
- Country: _____

3. Which types of digital objects are managed by the organization you audited? (Please tick all that apply)

- ☐ Textual documents
- ☐ Raw data
- ☐ Processed data
- ☐ Images
- ☐ Video
- ☐ Audio
- ☐ Databases
- ☐ Web sites/ web pages
- ☐ Software
- ☐ Multimedia
- ☐ Datasets
- ☐ Maps/Geospatial data
- ☐ Compound objects
- ☐ 3D objects
- ☐ Other (please specify) _____

4. Please indicate the responsibilities of the d.p. team inside the repository:

- ☐ Archivist;
- ☐ File formats expert;
- ☐ XML specialist;
- ☐ IT managers;
- ☐ Information Security Manager;
- ☐ Hardware and OS specialists;
- ☐ System administrators;
- ☐ IT developers;
- ☐ Service support;
- ☐ Other (specify) _____

5. Please explain the contribution that the component of the repository d.p. team has made to the audit process:

- ☐ Role 1: [e.g. Archivist]_____

Contribution:_____

- ☐ Role 2: [e.g. System Administrator]_____

Contribution:_____

- ☐ Role 3: _____

Contribution:_____

- ☐ Role 4: _____

Contribution:_____

- ☐ Role 5: _____

Contribution:_____

6. With respect to the interviewed repository d.p. team, who do you think should be interviewed in an “ideal” audit process?

7. Were there any missing competencies that should have been represented in the audited repository team?

8. What were the most difficult issues faced during the audit?

9. In your opinion, which problem identified could be prevented through the delivery of appropriate training on the repository staff?

10. What fundamental professional competences should an auditor have (regardless of the specific domain)?

11. What fundamental things should an auditor know before starting an audit?

12. Is there anything else about your experience of the audit process, particularly in relation to training needs, that you would like to share?

INTERVIEW FOR ORGANIZATIONS THAT WERE AUDITED

Abbreviations: d.p. = digital preservation

1. General information about the respondent

Name: _____

Institution: _____

Job Title: _____

Phone: _____

Email: _____

2. About the organization that was recently audited.

Organization sector:

- ☐ Academic/research
- ☐ Business
- ☐ Government
- ☐ Public sector
- ☐ Private
- ☐ Other (Please specify) _____

Country: _____

3. Which types of digital objects are managed by the organization audited? (Please tick all that apply)

- ☐ Textual documents
- ☐ Raw data
- ☐ Processed data
- ☐ Images
- ☐ Video
- ☐ Audio
- ☐ Databases
- ☐ Web sites/ web pages
- ☐ Software
- ☐ Multimedia
- ☐ Datasets
- ☐ Maps/Geospatial data
- ☐ Compound objects
- ☐ 3D objects
- ☐ Other (please specify) _____

4. Please indicate roles of those who participated in the audit from your organization:

- ☐ Archivist;
- ☐ File formats expert;
- ☐ XML specialist;
- ☐ IT managers;
- ☐ Information Security Manager;

- ☐ Hardware and OS specialists;
- ☐ System administrators;
- ☐ IT developers;
- ☐ Service support;
- ☐ Other (specify) _____

5. Please explain the contribution that each member of the repository d.p. team made to the audit process:

- ☐ Role 1: [e.g. IT manager] _____

Contribution: _____

- ☐ Role 2: [e.g. System Administrator] _____

Contribution: _____

- ☐ Role 3: _____

Contribution: _____

- ☐ Role 4: _____

Contribution: _____

- ☐ Role 5: _____

Contribution: _____

6. Considering the different expertise represented in the repository team involved in the audit, what would you improve or add, in terms of specific competences or skills?

7. What are the fundamental things that a member of the repository d.p. team needs to know before starting an audit (in terms of general knowledge, regardless of the specific domain)?

8. What were the most difficult issues faced during the audit?

9. In your opinion, which of the identified problems could be prevented through the delivery of appropriate training on the staff?

10. What additional competencies or skills do you think it would be useful for auditors to have before undertaking an audit?

11. What competencies or skills do you think are required after an audit to enable improvement based on the results?

12. Is there anything else about your experience of the audit process, particularly in relation to training needs, that you would like to share?

ANNEX B: LIST OF ANALYSED TRAINING INITIATIVES

In Annex B some of the courses have the same url: this is due to the type of web page used by the organizer to advertise training events. In these cases the name of the location has been added to better identify each training initiative.

Training courses

N.	TITLE OF THE TRAINING COURSE	COURSE PROVIDER	URL
1.	Electronic Records Management (ERM) Certificate Program (Atlanta)	AIIM (Association for Information and Image Management)	http://www.aiim.org/Training/Certificate-Courses/ERM/Course%20Descriptions
2.	Electronic Records Management (ERM) Certificate Program (Silver Spring)	AIIM (Association for Information and Image Management)	http://www.aiim.org/Training/Certificate-Courses/ERM/Course%20Descriptions
3.	Electronic Records Management (ERM) Certificate Program (Phoenix)	AIIM (Association for Information and Image Management)	http://www.aiim.org/Training/Certificate-Courses/ERM/Course%20Descriptions
4.	Electronic Records Management (ERM) Certificate Program (Dallas)	AIIM (Association for Information and Image Management)	http://www.aiim.org/Training/Certificate-Courses/ERM/Course%20Descriptions
5.	Electronic Records Management (ERM) Certificate Program (Chicago)	AIIM (Association for Information and Image Management)	http://www.aiim.org/Training/Certificate-Courses/ERM/Course%20Descriptions
6.	Electronic Records Management (ERM) Certificate Program (San Francisco)	AIIM (Association for Information and Image Management)	http://www.aiim.org/Training/Certificate-Courses/ERM/Course%20Descriptions
7.	Understanding and preserving audio collections	British Library, Preservation Advisory Centre	http://www.bl.uk/blpac/audio071111.html
8.	Writing and using a preservation policy	British Library, Preservation Advisory Centre	http://www.bl.uk/blpac/policy171111.html
9.	Diploma/Certificate in Professional Studies: Archives and Records Management	Centre for Archive Studies - Lucas -University of Liverpool	http://www.liv.ac.uk/lucas/dcpsarm/
10.	Records and information management education for the public sector	Centre for Archive Studies - Lucas -University of Liverpool	http://www.liv.ac.uk/lucas/dcps_rim/
11.	Preservation techniques and methodologies for digital audiovisual works	CIANT (International Centre for Art and New Technologies in Prague, Czech Republic) and TransISTor	http://transistor.ciant.cz/2011/module-i-preservation-techniques-and-methodologies-for-digital-audiovisual-works
12.	Master in Library and Information Science	Cologne University of Applied Sciences	http://malis.fh-koeln.de/fileadmin/templates/down

			load/MALIS_Flyer_english.pdf
13.	FRAME (Future for Restoration of Audiovisual Memory in Europe)	Ina EXPERT, with International Federation of Television Archives (FIAT/IFTA), the European Broadcasting Union (EBU/UER) and MEDIA, a programme of the European Union	http://www.ina-sup.com/en/about-ina-sup/frame-future-restoration-audiovisual-memory-europe
14.	Long-Term Audiovisual Digital Preservation: Strategy, Planning & Tools	Institut National de l'Audiovisuel	http://www.regonline.com/builder/site/default.aspx?EventID=970982
15.	Managing Digital Collections	JISC Digital Media	http://www.jiscdigitalmedia.ac.uk/training/courses/managing-digital-media-collections (link to current page)
16.	Digital Preservation	King's Digital Consultancy Services	http://www.kdcs.kcl.ac.uk/training/information-schedule.html
17.	nestor summer school 2010 "Experiences in Digital Preservation"	nestor	http://nestor.sub.uni-goettingen.de/summer_school_2010/index.php (link broken April 2012)
18.	master's program Conservation of New Media and Digital Information	State Academy of art and Design	http://www.mediaconservation.abk-stuttgart.de/english
19.	The Digital Preservation Training Programme (DPTP) (delivered on four occasions through 2010 and 2011)	ULCC (University of London Computer Centre), working with its partners the Digital Preservation Coalition (DPC) and Cornell University	http://www.dptp.org
20.	Information Management & Preservation (Digital)/(Archives & Records Management)	University of Glasgow	http://www.gla.ac.uk/postgraduate/taught/informationmanagementpreservationdigitalarchivesrecordsmanagement/
21.	DigCCurr Professional Institute: Curation Practices for the Digital Object Lifecycle	University of North Carolina at Chapel Hill	http://ils.unc.edu/digccurr/institute.html
22.	Digital Preservation	University of Urbino	Not available
23.	Preservation of digital archives	University of Rome Sapienza	http://digilab.uniroma1.it/corsi/filiere_dett.aspx?ID=20
24.	Preservation of digital archives	Fondazione Feltrinelli	http://www.fondazionefeltrinelli.it/feltrinelli-cms/cms.view?munu_str=0_4_2&numDoc=877
25.	Formazione, gestione e conservazione di archivi digitali in ambito pubblico e privato	University of Macerata	http://www.unimc.it/af/master/10/formazione-gestione-e-conservazione-di-archivi
26.	Indicizzazione di documenti cartacei, multimediali ed elettronici in ambiente digitale (MIDA)	University of Tor Vergata in Rome and BAICR Sistema Cultura	http://web.uniroma2.it/modules.php?name=Content&section_parent=517
27.	La conservazione a lungo	Soprintendenza archivistica, Istituto	http://www.archiviando.org/forum/

	termine degli archivi digitali: metodologie e progetti	per i Beni artistici, culturali e naturali della Regione Emilia-Romagna, Polo archivistico regionale and Associazione Nazionale Archivistica Italiana, section Emilia-Romagna region	viewtopic.php?f=4&t=765 http://www.archiviando.org/forum/download/file.php?id=419&sid=2e8aec3065d9241dbb531b831286e6b3
28.	Archivum docet - Gestire e valorizzare l'archivio scolastico	Veneto Region, ANAI (Italian Association of Archives)	http://www2.regione.veneto.it/cultura/cms/allegati/Archivi/Archivum-docet-2012-programma.pdf
29.	Conservazione a lungo termine degli Archivi Digitali	State Archive of Florence, Schools of Archives, Palaeography and Diplomatics "Anna Maria Enriques Agnoletti, Fondazione Rinascimento Digitale	http://www.rinascimento-digitale.it/archividigitali.phtml
30.	Gestione e conservazione digitale dei documenti	ANORC, Associazione Nazionale per Operatori e responsabili della conservazione digitale	http://www.anorc.it/documenti/nuova%20brochure%20Master%20Course%20ANORC%202012_HR.pdf
31.	La gestione informatica dei documenti	Scuola Superiore della Pubblica Amministrazione (SSPA)	http://www.sspa.it/?p=4254
32.	Master of Records Management	Archivschule Marburg – Hochschule für Archivwissenschaft	http://www.archivschule.de/master-of-records-management
33.	Data Intelligence 4 Librarians	3TU.Data centre, DANS	http://dataintelligence.3tu.nl/nl/home/
34.	Digitale duurzaamheid: bewaren voor de toekomst	DO Consultancy	http://www.doconsultancy.net/trainingen/digitaal-duurzaam-archiveren.html
35.	Digitaliseren en digitaal archiveren Module 7: Digitaliseren van audiovisuele objecten	FARO, Flemish Interface Centre for Cultural Heritage	http://www.faronet.be/kalender/digitaliseren-en-digitaal-archiveren-module-7-digitaliseren-van-audiovisuele-objecten-beweg
36.	Digital Futures	King's Digital Consultancy Service, King's College London	http://www.kdcs.kcl.ac.uk/digifutures/london.html
37.	Digital Preservation Training Programme	University of London Computer Centre	http://www.dptp.org
38.	Audio and video materials: Passive preservation, digitization and preservation	ONB (National Library of Austria)	http://www.onb.ac.at/brainpool/seminars/kurs23.htm
39.	Pérennisation et communication de l'information numérique	PIN (not for profit organization)	http://pin.association-aristote.fr/doku.php/public/actualites_blog/formation_perennisation_et_communication_de_l_information_numerique_du_12_au_16_novembre_2012
40.	Masters in Digital Libraries	Universidad de Valladolid	http://mbidi.uva.es/ http://www.emagister.com/master-universitario-bibliotecas-digitales-cursos-1093870.htm

41.	Digital preservation of cultural heritage. digital archives	Aristotle University of Thessaloniki – Faculty of Theology -School of Theology	http://www.theo.auth.gr/theo/Lists/Courses/Attachments/105/Syllabus(49).pdf
42.	AccessIT - e-Courses	Central Public Library of Veria in corporation with The Laboratory of Digital Libraries and Electronic Publications - Department of Archives and Library Science- Ionian University	http://accessit.libver.gr
43.	Material Preservation	Ionian University - department of archives and library science	https://e-class.ionio.gr/modules/auth/opencourses.php?fc=11
44.	Multiple courses: Archival and information science	Mid Sweden University	http://www.miun.se/Utbildning/Hitta-din-utbildning/ATLAS-Visa-amne/?AmneId=2&KatalogId=8

Workshops

N.	TITLE OF THE WORKSHOP	WORKSHOP PROVIDER	URL
1.	Preservation Assessment Survey Workshop	British Library, Centre for Conservation	http://www.bl.uk/blpac/pasworkshop2.html
2.	Digital Curation Tools And Methods	CNI-Coalition for Network Information North Carolina	http://www.cni.org/news/curategear-digital-curation-tools-and-methods (link to 2012 workshop)
3.	Digital Preservation Management Workshops	Cornell University	http://www.dpworkshop.org/workshops/fiveday.html (link to 2012 workshop)
4.	Digital Curation 101: How to Manage Research Data - "Institutional Challenges in the Data Decade" (Glasgow, UK)	Digital Curation Center (DCC)	http://www.dcc.ac.uk/dcc.ac.uk/training/dc-101
5.	Digital Curation 101: How to Manage Research Data - "DATUM: Research data management" (Newcastle, UK)	Digital Curation Center (DCC)	http://www.dcc.ac.uk/dcc.ac.uk/training/dc-101
6.	Digital Curation 101: How to Manage Research Data - "Institutional Challenges in the Data Decade" (Sheffield, UK)	Digital Curation Center (DCC)	http://www.dcc.ac.uk/dcc.ac.uk/training/dc-101
7.	Digital Curation 101: How to Manage Research Data - "DC101 Lite" (Cambridge, UK)	Digital Curation Center (DCC)	http://www.dcc.ac.uk/dcc.ac.uk/training/dc-101
8.	Digital Curation 101: How to Manage Research Data - "Digital Curation 101 Lite"	Digital Curation Center (DCC)	http://www.dcc.ac.uk/dcc.ac.uk/training/dc-101

	(Chicago, USA)		
9.	Digital Curation 101: How to Manage Research Data – “Institutional Challenges in the Data Decade” (Bath, UK)	Digital Curation Center (DCC)	http://www.dcc.ac.uk/dcc.ac.uk/training/dc-101
10.	Digital Curation 101: How to Manage Research Data - "DC101 Lite" (London, UK)	Digital Curation Center (DCC)	http://www.dcc.ac.uk/dcc.ac.uk/training/dc-101
11.	Digital Curation 101: How to Manage Research Data - "DC101 Lite" (Oxford)	Digital Curation Center (DCC)	http://www.dcc.ac.uk/dcc.ac.uk/training/dc-101
12.	Tools of the Trade Workshops Manchester	Digital Curation Center (DCC)	http://www.dcc.ac.uk/dcc.ac.uk/training/dc-101
13.	Digital Curation 101: How to Manage Research Data (Cambridge, UK)	Digital Curation Center (DCC) and Cambridge University Library	http://www.dcc.ac.uk/dcc.ac.uk/training/dc-101
14.	Digital Curation 101: How to Manage Research Data (Oxford, UK)	Digital Curation Center (DCC) with the Oxford eResearch Centre, Oxford University Computing Services and the Bodleian Libraries	http://www.dcc.ac.uk/dcc.ac.uk/training/dc-101
15.	Software Art	Humanities Advanced Technology & Information Institute (HATII) based at the University of Glasgow	http://www.dcc.ac.uk/events/external-events/preservation-software-art-symposium
16.	JISC Digital Preservation Benefits Tools Project	London South Bank University	http://www.ukoln.ac.uk/events/i2s2-krds/programme
17.	Preserving Email: Directions and Perspectives	Digital Preservation Coalition (DPC)	http://www.dpconline.org/events
18.	Intellectual Property Rights and Digital Preservation	Digital Preservation Coalition (DPC)	http://www.dpconline.org/events
19.	The Future of File Format Identification: PRONOM and DROID User Consultation	Digital Preservation Coalition (DPC)	http://www.dpconline.org/events
20.	Digital Preservation: What I Wish I Knew Before I Started	Digital Preservation Coalition (DPC)	http://www.dpconline.org/events/previous-events
21.	Digital Preservation for Forensics	Digital Preservation Coalition (DPC)	http://www.dpconline.org/events/previous-events
22.	Revisiting Archival Principles from a digital preservation viewpoint	Digital Preservation Coalition (DPC)	http://www.dpconline.org/events/previous-events
23.	Data for Life: Digital Preservation for Health Studies	Digital Preservation Coalition (DPC)	http://www.dpconline.org/events/previous-events
24.	Preserving Digital Art: Directions and Perspectives	Digital Preservation Coalition (DPC)	http://www.dpconline.org/events/previous-events

25.	Getting Started in Digital Preservation (Cardiff, UK)	Preservation Advisory Centre and Digital Preservation Coalition (DPC)	http://www.dpconline.org/events/previous-events
26.	Getting Started in Digital Preservation (Glasgow, UK)	Preservation Advisory Centre and Digital Preservation Coalition (DPC)	http://www.dpconline.org/events/previous-events
27.	Getting Started in Digital Preservation (London, UK)	Preservation Advisory Centre and Digital Preservation Coalition (DPC)	http://www.dpconline.org/events/previous-events
28.	Decoding the Digital: A Common Language for Preservation	Preservation Advisory Centre and Digital Preservation Coalition (DPC)	http://www.dpconline.org/events/previous-events
29.	Designed to Last: Preserving Computer Aided Design	Digital Preservation Coalition (DPC)	http://www.dpconline.org/events/previous-events
30.	e-Journals are forever?	Digital Preservation Coalition (DPC)	http://www.dpconline.org/events/previous-events
31.	Practical Tools for Digital Preservation: A Hack-a-thon	The Digital Preservation Coalition (DPC), Open Planets Foundation	http://www.dpconline.org/events/details/34-hackathon?xref=33
32.	Gli archivi digitalizzati: tecniche, regole, cautele necessarie per la conservazione e protezione	Consorzio NuovaPA	http://asp.lattanziogroup.eu/sito/schedaseminario.aspx?idnazione=5&idsottogruppo=324
33.	Archivi ibridi: soluzioni operative per un problema del presente	Municipality of Padua	http://www2.regione.veneto.it/cultura/cms/allegati/Archivi/ArchiviIbridi_Programma.pdf
34.	Managing Research data	Royal Netherlands Academy of Arts and Sciences (KNAW)	- Not available -
35.	Data Management and Research Infrastructures	SURF (ICT / Network provider of Dutch Research community)	http://www.surf-academy.nl/archief/event/?id=388
36.	Workshop Digital Forensics	SURF (ICT / network provider for Dutch Research Community)	http://www.surf-academy.nl/archief/event/?id=430
37.	Seminar Sharing Medical Research data	SURF	http://www.surf-academy.nl/archief/event/?id=433 http://www.surf-academy.nl/programma/event/?id=466
38.	Seminar: Research data: wat kan, mag en moet	SURF	http://www.surf-academy.nl/archief/event/?id=413
39.	Seminar: Privacy and the	SURF	http://www.surf-

	Cloud: State of Art		academy.nl/archief/event/?id=464
40.	Persistent Identifier Seminar	Knowledge Exchange	http://www.knowledge-exchange.info/Default.aspx?ID=440
41.	Naar de ontwikkeling van een PI infrastructuur	DANS – KB – SURF foundation	http://www.dans.knaw.nl/content/categorieen/symposia/symposia-archief/workshop-%E2%80%98naar-de-ontwikkeling-van-een-pi-infrastructuur
42.	Digital Durability of GEO-data	DANS and GEONOVUM	http://www.dans.knaw.nl/content/categorieen/symposia/symposia-archief/studiemiddag-%E2%80%98digitale-duurzaamheid-van-geodata%E2%80%99
43.	Data beheer in de praktijk	DANS – and SURF foundation	http://www.dans.knaw.nl/content/categorieen/symposia/symposia-archief/symposium-databeheer-de-praktijk-resultaten-van-het-ca
44.	(R)evolution in the public information: preserve, certify and make accessible	The National Archives of Portugal	http://www.bad.pt/noticia/2011/10/17/seminario-%E2%80%98r-evolucao-da-informacao-publica-preservar-certificar-e-acessibilizar
45.	What I Wish I Knew Before I Started	Digital Preservation Coalition (DPC)	http://www.dpconline.org/events/details/38-studentconference?xref=38
46.	Emulation in Digital Preservation	KEEP Project, University of Portsmouth	http://keep-cardiff.eventbrite.com/ http://www.keep-project.eu/downloads/keep_cardiff.pdf
47.	Preservation Of Complex Objects Symposia (POCOS) on Preservation of Games and Virtual Worlds	Preservation Of Complex Objects Symposia (POCOS) on Preservation of Games and Virtual Worlds	http://www.pocos.org/index.php/pocos-symposia/videogame-environments-a-virtual-worlds
48.	Trust in Post-Cancellation Access Services for E-Journals	Digital Preservation Coalition (DPC)	http://www.dpconline.org/events/previous-events/837-trust-and-e-journals
49.	Digital Preservation and Digital Resilience	Digital Preservation Coalition (DPC)	http://www.dpconline.org/events/details/44-resilience?xref=46
50.	Links That Last: Linked Data, Persistent Identifiers and Digital Preservation	Digital Preservation Coalition (DPC)	http://www.dpconline.org/events/details/47-linksthatlast?xref=49
51.	JISC-British Library DataCite Workshop: Describe,	DataCite, JISC, British Library	http://www.eventbrite.com/event

	disseminate, discover: metadata for effective data citation		t/3765155682
52.	Digitale Langzeitarchivierung technologische, organisatorische und rechtliche Aspekte	ADV Arbeitsgemeinschaft für Datenverarbeitung	http://www.adv.at/veranstaltung/en/20110922_Langzeitarchivierung.htm http://www.adv.at/veranstaltung/en/programme/Langzeitarchivierung_20100916.pdf
53.	Digitale Langzeitarchivierung	Verband Österreichischer Archivarinnen und Archivare	http://kulturgueter.kath-orden.at/termine-service/digitale-langzeitarchivierung
54.	Digital Long-term Preservation at Higher Education Institutions	nestor	http://www.dini.de/veranstaltungen/workshops/langzeitarchivierung
55.	Web Archiving	nestor	http://www.langzeitarchivierung.de/Subsites/nestor/DE/Veranstaltungen/vergangene/Webarchivierung2012.html;jsessionid=95276697D43204B440615903C41040D9.prod-worker4
56.	Nestor Practitioners Day	nestor	http://www.langzeitarchivierung.de/Subsites/nestor/DE/Veranstaltungen/vergangene/nestorPraktikertag2011.html;jsessionid=95276697D43204B440615903C41040D9.prod-worker4
57.	Preservation of Social and Economics Data	nestor	http://www.langzeitarchivierung.de/Subsites/nestor/DE/Veranstaltungen/vergangene/RatSWDGESISnestor.html;jsessionid=D6AAF2491B086ED6C5390448DE4B3C2B.prod-worker4
58.	Security and Trustworthiness in Digital Preservation	nestor	http://www.langzeitarchivierung.de/Subsites/nestor/DE/Veranstaltungen/vergangene/BTag2011.html;jsessionid=D6AAF2491B086ED6C5390448DE4B3C2B.prod-worker4
59.	Introduction to New Approaches in Digital Preservation. Interim Results and Perspectives of the SHAMAN Project	SHAMAN Project and nestor	http://shaman-ip.eu/shaman/sites/default/files/publications/SHAMAN_ISP1_Training_event_1Feb2011.pdf
60.	Long-Term Preservation of Digitised Holdings	nestor, Bayerische Staatsbibliothek, Leibniz-Rechenzentrum and	http://www.babs-muenchen.de/index.html?pconte

		Förderung der Deutschen Forschungsgemeinschaft	nt=babs_workshop&vcontent=info&l=de
61.	Digital Archiving Today – Insights into Practice	nestor	http://www.langzeitarchivierung.de/Subsites/nestor/DE/Veranstaltungen/vergangene/nestorPraktikertag2010.html;jsessionid=50FA2EAF863ACB773E14A7D5088B9CE5.prod-worker4
62.	Self-supporter or All-Inclusive? How to Organize the Submission of Digital Documents	nestor	http://www.langzeitarchivierung.de/Subsites/nestor/DE/Veranstaltungen/vergangene/Selbstversorger.html;jsessionid=D30B82D3509D15F6C25DD388A775DEE4.prod-worker4
63.	DCC Roadshow Northwest England (Salford, UK)	Digital Curation Centre (DCC)	http://www.dcc.ac.uk/events/data-management-roadshows/dcc-roadshow-northwest-england
64.	DCC Roadshow Northeast England (Newcastle upon Tyne, UK)	Digital Curation Centre (DCC)	http://www.dcc.ac.uk/events/data-management-roadshows/dcc-roadshow-northeast
65.	OPF Hackathon - A Practical Approach to Preservation Systems	Open Planets Foundation	http://wiki.opf-labs.org/display/KB/2012-05-21+OPF+Hackathon+-+A+Practical+Approach+to+Preservation+Systems
66.	SPRUCE Digital Preservation Mashup	SPRUCE project	http://www.dpconline.org/events/details/40-SpruceGlasgow?xref=41
67.	Cologne Hackathon	Open Planets Foundation	http://www.openplanetsfoundation.org/events/2011-08-31-opf-hackathon
68.	OPF Hackathon - A Practical Approach to Database Archiving	Open Planets Foundation	http://wiki.opf-labs.org/display/KB/2011-08-31+OPF+Cologne+Hackathon
69.	The KEEP approach to digital preservation	KEEP (Keeping Emulation Environments Portable)	http://www.openplanetsfoundation.org/events/2011-11-29-keep-approach-digital-preservation-29-30-november-2011-rome-0

Tutorials

N.	TITLE OF THE TUTORIAL	TUTORIAL PROVIDER	URL
1.	Digital Preservation Management Tutorial (and Workshops) - Digital Preservation Management: Implementing Short-term Strategies for Long-term Problems	Cornell University	http://www.icpsr.umich.edu/dpm/dpm-ita/index.html

Online courses

N.	TITLE OF THE ONLINE COURSE	COURSE PROVIDER	URL
1.	Digital Preservation Module	AIIM (Association for Information and Image Management)	http://www.aiim.org/training/courses/312
2.	Protecting Future Access Now: Models for Preserving Digitized Books and Other Content at Cultural Heritage Organizations	ALA (American Library Association) ALCTS (Association for Library Collections and Technical Services)	http://www.ala.org/ala/mgrps/divs/alcts/confevents/upcoming/webinar/pres/042711.cfm
3.	Fundamentals of Preservation	ALCTS Association for Library Collections & Technical Services	http://www.classes.ala.org/course/category.php?id=3
4.	Digital Preservation: An Introduction to the Basic Concepts	ALCTS Association for Library Collections & Technical Services (USA)	http://www.ala.org/ala/mgrps/divs/alcts/confevents/upcoming/webinar/pres/101710digital.cfm
5.	Preserving Your Personal Digital Memories	American Library Association (ALA)	http://www.ala.org/ala/mgrps/divs/alcts/confevents/upcoming/webinar/pres/042811.cfm
6.	Steps in a Digital Preservation Workflow	American Library Association (ALA)	http://www.ala.org/ala/mgrps/divs/alcts/confevents/upcoming/webinar/030712.cfm
7.	Auditing Basics for RIM (Records and Information Management) Professionals	ARMA International (a not-for-profit professional association and the authority on managing records and information)	https://www.arma.org/eweb/DynamicPage.aspx?WebKey=35B79A60-E570-40F9-B74E-5AEE5DF46147
8.	Intro to Digital Preservation #1 - Identifying & Selecting Content	Association of Southeastern Research Libraries	http://www.aserl.org/
9.	Intro to Digital Preservation #2 -- Considerations for Storage and Protection of Content	Association of Southeastern Research Libraries	http://www.aserl.org/
10.	Intro to Digital Preservation #3 -- Make Plans to Manage Content and Provide Access Over Time	Association of Southeastern Research Libraries	http://www.aserl.org/
11.	DPE Digital Preservation Video Training	DPE DigitalPreservationEurope	http://www.digitalpreservationeurope.eu/video-training/
12.	Preservation Planning Success Stories - Knowledge Futures: Digital Preservation Planning	Duraspace	http://events.r20.constantcontact.com/register/event?oeidk=a07e5igqrqr69ac8f0e&llr=5iy95gcab
13.	Preservation and Archiving highlights from the Alliance Digital Repository - Knowledge Futures: Digital Preservation Planning	DuraSpace	http://events.r20.constantcontact.com/register/event?oeidk=a07e5igqs2c83b47dfc&llr=5iy95gca b
14.	Digital Preservation: Audio and Video Formats	Infopeople	http://infopeople.org/training/audio-and-video-formats

15.	Digital Preservation: Fundamentals	Infopeople	http://infopeople.org/training/digital-preservation-fundamentals
16.	Digital Preservation: Text and Image Formats	Infopeople	http://infopeople.org/training/audio-and-video-formats
17.	Electronic Records: Preservation Options of PDF	Society of American Archivists	http://saa.archivists.org/Scripts/4Disapi.dll/4DCGI/events/54.html?Action=Conference_Detail&ConfID_W=54&Time=1641838503
18.	Management and Preservation of Digital Records	University of Dundee Centre for Archive and Information Studies	http://www.dundee.ac.uk/cais/cpd/modules_for_cpd.htm
19.	An Introduction to Digital Preservation	WebJunction	http://www.webjunction.org/events/webinars/webinar-archives/-/articles/content/110494344
20.	Planos de Preservação Digital: roadmaps informacionais das organizações – Plans for Digital Preservation: Information Roadmaps of Organisations	The National Archives of Portugal	http://www.bad.pt/noticia/2012/03/29/4o-webinar-bad-planos-de-preservacao-digital-roadmaps-informacionais-das-organizacoes

Annex C – Proposed APARSEN Training Schedule 2013/14

Workshop and Training Course on Trust - February/March 2013

Proposed Location: Dublin, Ireland

Participating Partners (TBC): DPC, CINI, FRD, FORTH, APA, AFPUM

Two day event looking at issues relating to the topic of Trust. Day one will be a workshop featuring speakers from relevant APARSEN work packages covering topics such as persistent identifiers, authenticity and provenance, and audit and certification. Opportunities will be provided throughout the day for feedback and discussion. Day two will feature practical training on the Authenticity framework developed within WP24 and Persistent Identifiers work from WP22. Each training session will last 3 hours and consist of an hour's introductory presentation and then two hour-long 'sub-modules' featuring practical, hands-on training and, potentially, in-depth analysis of a sub-topic.

Timeline	
Aug 2012	Proposal for WP41/WP43 Collaboration on Workshops and Training
Dec 2012 - Jan 2013	Planning of event, development of training materials
Jan 2013	Begin publicity for event
Feb 2013	Final planning, peer review of training materials, submission of workshop presentations
Feb/Mar 2013	Stage event
Mar 2013	Produce feedback report

Launch of Executive Briefing Podcasts - April 2013 (additional material released July and October 2013 and January and April 2014)

Participating Partners (TBC): DPC, KB, ONB, DANS

Envisioned as both an online training resource and an advocacy tool for digital preservation practitioners, this set of podcasts will focus on a number of key issues relating to digital preservation from the point of view of those at an executive level. Topics to be covered will likely include the Value of Digital Preservation, Risk Management, Compliance, and Digital Legacy. Each podcast will last approximately 30 minutes and feature a presentation by a leading member of the Digital Preservation community. The first podcast will be released in April and will be an introduction to basic Digital Preservation concepts. Additional podcasts will be added on a quarterly basis in July and October 2013 and January and April 2014.

Timeline	
Nov 2012 – Jan 2013	Topics for podcasts chosen, planning (investigate possible formats, logistics of recording, choose speakers)
Mar-Apr 2013	Production and release of first podcast
Jun-Jul 2013	Production and release of second podcast
Sep-Oct 2013	Production and launch of third podcast

Advanced Practitioner Course - July 2013

Proposed Location: Glasgow, UK

Participating Partners: DPC, FRD, CINI, others to be confirmed depending on agreed course structure/content.

This week long training course will provide the next level of training to digital preservation practitioners with existing experience in the field and/or have attended an introductory level course such as the DCC's Digital Preservation 101 and ULCC's DPTP courses in the UK. It will aim to present training from the leading edge of digital preservation research and practice, drawing from work within the APARSEN project as well as other European projects and initiatives.

Timeline	
Dec 2012	Meeting with key WP43 partners and external stakeholders to discuss structure of course and potential content.
Jan 2012 - Apr 2013	Planning of event, development of training materials
Mar 2013	Begin publicity for event
May -Jun 2013	Final planning, peer review of training materials, submission of final training materials
Jun/Jul 2013	Stage event
Aug 2013	Produce feedback report

Workshop/Training Course on Sustainability - September/October 2013

Proposed Location: The Hague, Netherlands

Participating Partners (TBC): DANS, BL, UPAT, others to be selected at planning stage.

Two day event looking at issues relating to the topic of Sustainability. Day one will be a workshop featuring speakers from relevant APARSEN work packages covering topics such as storage, preservation, business cases and cost modelling. Opportunities will be provided throughout the day for feedback and discussion. Day two will feature practical training on the work relating to cost modelling in WP32 and business cases in WP36.

Timeline	
Jun-Aug 2013	Planning of event, development of training materials
Jul 2013	Begin publicity for event
Aug-Sep 2013	Final planning, peer review of training materials, submission of workshop presentations
Sep/Oct 2013	Stage event
Mar 2013	Produce feedback report

Workshop/Training Course on Usability – February/March 2014

Proposed Location: TBC

Two day event looking at issues relating to the topic of Usability. Day one will be a workshop featuring speakers from relevant APARSEN work packages covering topics from the Usability topic area as well as guest speakers covering complementary topics. Opportunities will be provided throughout the day for feedback and discussion. Day two will feature practical training on the work on this topic carried out within APARSEN and selected with reference to the outcomes of D43.1.

Advanced Practitioner Course – June/July 2014

Proposed Location: Austria

This week long training course will provide the next level of training to digital preservation practitioners with existing experience in the field and/or have attended an introductory level course such as the DCC's Digital Preservation 101 and ULCC's DPTP courses in the UK. It will aim to present training from the leading edge of digital preservation research and practice, drawing from work within the APARSEN project as well as other European projects and initiatives.

Workshop/Training Course on Access – September/October 2014

Proposed Location: Italy

Two day event looking at issues relating to the topic of Access. Day one will be a workshop featuring speakers from relevant APARSEN work packages covering topics from the Access topic area as well as guest speakers covering complementary topics. Opportunities will be provided throughout the day for feedback and discussion. Day two will feature practical training on the work on this topic carried out within APARSEN and selected with reference to the outcomes of D43.1.