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

D11.6 Virtual Centre of Excellence Development

Document identifier:	**APARSEN**-REP-D11_6-01-1_4
Due Date:	2014-12-31
Submission Date:	2015-02-13
Work package:	WP11
Partners:	All other partners
WP Lead Partner:	APA
Document status	Released
URN	urn:nbn:de:101-2014121619

Abstract: This is the final report in the series on the APARSEN Virtual Centre of Excellence. It provides the latest input relevant to the nature and functioning of the VCoE and a blueprint for its implementation.

Delivery Type	Report
Author(s)	David Giaretta and all partners
Approval	David Giaretta / Simon Lambert (coordinators)
Summary	
Keyword List	Virtual Centre of Excellence
Availability	☒ PUBLIC

Document Status Sheet

Issue	Date	Comment	Author
0.1	2014-06-01	Updates of D11.4	ALL partners
0.2	2013-011-15	Add insights gained from launch event	ALL partners
1.0	2014-12-10	Updated version	Ruben Riestra, David Giaretta, Simon Lambert
1.1	2014-12-11	Updated info on organisational issues	David Giaretta
1.2	2014-12-30	Clarify the resources needed	David Giaretta
1.3	2105-01-14	Update on market research results	Ruben Riestra
1.4	2015-01-22	Final updates on organisational options, and editorial checks	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.org

Glossary

Business case[1]**:** A business case captures the reasoning as well as the quantifiable and unquantifiable characteristics of a proposed project. The logic of the business case is that, whenever resources such as money or effort are consumed, they should be in support of a specific business need. Business cases are created to help decision-makers ensure that:

- the proposed initiative will have value and relative priority compared to alternative initiatives
- the organization has the capability to deliver the benefits
- the organization's dedicated resources are working on the highest value opportunities
- projects with inter-dependencies are undertaken in the optimum sequence
- the performance of initiatives is monitored objectively based on the objectives and expected benefits laid out in the business case

Customer Value Proposition[2] : consists of the sum total of benefits which a vendor promises a customer will receive in return for the customer's associated payment (or other value-transfer).It is a clearly defined statement that is designed to convince customers that one particular product or service will add more value or better solve a problem than others in its competitive set

Business model[3]**:** A business model describes the rationale of how an organization creates, delivers, and captures value (economic, social, cultural, or other forms of value). The process of business model construction is part of business strategy; business models are used by managers inside companies to explore possibilities for future development. Whenever a business is established, "it either explicitly or implicitly employs a particular business model that describes the architecture of the value creation, delivery, and capture mechanisms employed by the business enterprise. The essence of a business model is that it defines the manner by which the business enterprise delivers value to customers, entices customers to pay for value, and converts those payments to profit: it thus reflects management's hypothesis about what customers want, how they want it, and how an enterprise can organize to best meet those needs, get paid for doing so, and make a profit"[4].

Business Plan: A business plan is a formal statement of a set of business goals, the reasons they are believed attainable, and the plan for reaching those goals. It may also contain background information about the organization or team attempting to reach those goals[5].

Set of documents prepared by a firm's management to summarize its operational and financial objectives for the near future (usually one to three years) and to show how they will be achieved. It serves as a blueprint to guide the firm's policies and strategies, and is continually modified as conditions change and new opportunities and/or threats emerge. When prepared for external audience (lenders, prospective investors) it details the past, present, and forecasted performance of the firm[6].

Economic sustainability[7]**:** The use of various strategies for employing existing resources optimally so that that a responsible and beneficial balance can be achieved over the longer term. Within a business context, economic sustainability involves using the assorted assets of the company efficiently to allow it to continue functioning profitability over time[8].

1 http://en.wikipedia.org/wiki/Business_case
2 http://en.wikipedia.org/wiki/customer_value_proposition
3 http://en.wikipedia.org/wiki/Business_model
4 Teece, David J. Business Models, Business; Strategy and Innovation. Long Range Planning, vol 43; 2010-193.
5 Business Dictionary: http://en.wikipedia.org/wiki/Business_plan
6 http://www.businessdictionary.com/definition/business-plan.html
7 Read more: http://www.businessdictionary.com/definition/econoMic-sustainability.html#ixzz2ZuZz4Vyt
8 Business Dictionary: http://www.businessdictionary.com/definition/econoMic-sustainability.html

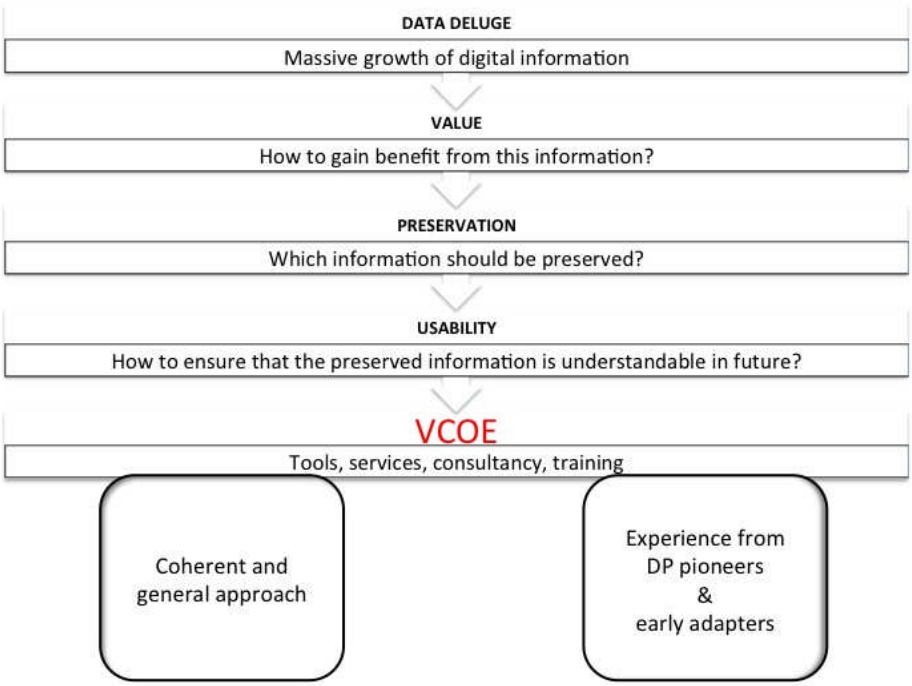
EXECUTIVE SUMMARY

The APARSEN Network of Excellence has the aim of defragmenting the landscape of digital preservation in Europe and leading to a Virtual Centre of Excellence, founded on a common vision of digital preservation.

The landscape of digital preservation is influenced by political, economic, socio-cultural and technological factors. Obviously the main driver is technological, starting with the global deluge of digitally encoded information. But the other factors all play important roles in shaping the demand and the value of this information to give benefit to society at large, businesses and individuals. Rapid changes in technology, environment and tacit knowledge mean we must preserve the data to have a chance to realise these benefits, and in particular to ensure, as far as possible, that it is possible to create new uses for existing information and can be combined automatically to create new information and hence new value.

The value proposition of the Virtual Centre of Excellence is based on two pillars established by the APARSEN project:

- a coherent approach based on value, and which can be applied to any kind of digital objects;
- the combined experience of the digital preservation pioneers both in the research field and, more importantly, as worldwide earliest adopters of digital preservation practices.



This report studies the factors that impact on the needs and opportunities for digital preservation, including the results of a market research exercise conducted during the last year of the APARSEN project, which provides valuable quantitative information about potential demand in different sectors and potentially unsatisfied needs. An analysis of 'competitors' is also provided—that is, of other options for provision of services, tools, training and consultancy, plus opportunities and threats for the VCoE, and approaches to business modelling. All this is the *blueprint* for the VCoE. Finally a number of options for implementation of the blueprint are outlined, based on different roles that organisations may play in the VCoE.

CONTENTS

1 INTRODUCTION 7

2 THE LANDSCAPE FOR VCOE CREATION AND OPERATION 9

2.1 THE CONTEXT FOR DIGITAL PRESERVATION SERVICES 9
2.1.1 Dimensions of the context 9
2.1.2 The political dimension 13
2.1.3 The economic dimension 14
2.1.4 The social dimension 15
2.1.5 The technological dimension 15
2.2 POTENTIAL DEMAND FOR VCOE OFFERINGS 17
2.2.1 Scientific and Research Institutions (SRI) 17
2.2.2 Memory Institutions 22
2.2.3 Industry 25
2.3 FORMULATION OF THE USAGE BUSINESS CASE 33
2.3.1 Patterns in usage of digital (preserved or to be preserved) content offered by VCoE potential 33
2.3.2 Particular aspects of functioning of different market segments (cultural heritage and industry) 35
2.3.3 Peculiarities of criteria, acquisition mechanisms and supplier selection 37
2.4 DETECTION OF UNSATISFIED NEEDS/COMMERCIAL OPPORTUNITIES 38
2.5 FORESEEABLE EVOLUTION, GENERAL FORECASTS ON DEMAND DEVELOPMENT AND MARKET SIZING 40
2.6 PRIORITISATION OF MARKET SEGMENTS TO BE ADDRESSED 41

3 VCOE POTENTIAL COMPETITORS 42

3.1 SERVICES AND TOOLS 42
3.2 TRAINING 57
3.3 CONSULTANCY 58

4 THE BUSINESS CHALLENGE FOR THE VCOE 59

4.1 OPPORTUNITIES AND THREATS 59
4.2 OPPORTUNITIES GENERATED FROM DEMAND SIDE OF MARKET 60
4.3 OPPORTUNITIES GENERATED IN SUPPLY SIDE FAILURES 62
4.4 WHY THE VCOE COULD OUTPERFORM CURRENT SUPPLIERS 62
4.5 WHY DEMAND (AT LEAST AT PART OF IT) WILL ACTUALLY CHOOSE THE VCOE 62

5 THE VCOE BUSINESS DESIGN 63

5.1 BUSINESS MODELLING 63
5.1.1 Target markets, value propositions and revenue streams 63
5.1.2 Customer relationships 66
5.1.3 Key activities 67
5.1.4 Key resources 67

6 ORGANISATIONAL OPTIONS: IMPLEMENTING THE BLUEPRINT 69

6.1 GENERAL PRINCIPLES 69
6.2 ROLES IN THE VCOE 69
6.3 EXAMPLES OF VCOE IMPLEMENTATION OPTIONS 71
6.3.1 APA as core organisation 71
6.3.2 Giaretta Associates as core organisation 71
6.3.3 Italian Centre of Excellence as national instance 72
6.3.4 The Research Data Alliance 73

ANNEX 1. SERVICES PORTFOLIO CONFIGURATION 76

I. CONSULTANCY 76
II. SOFTWARE PRODUCTS AND SOFTWARE SUPPORT AND UPGRADES 84
III. TRAINING COURSES 88
IV. OTHER SERVICES/ PRODUCTS/ TOOLS 90

1 INTRODUCTION

Work Package 11 of the APARSEN project is entitled 'Common vision', and essentially had two strands throughout the lifetime of the project. One was concerned with creating the common vision of digital preservation (DP) itself, integrating the results emerging from the many individual focussed work packages.

The second strand was concerned with establishing the Virtual Centre of Excellence (VCoE) envisaged as a lasting outcome of the project. This deliverable provides the last update on the blueprint for the VCoE by gathering and presenting the necessary information on supply and demand, aspects of the overall landscape, opportunities and threats, and options for the VCoE itself.

There have been previous deliverables in the series relating to the VCoE.

D11.2 Virtual Centre of Excellence development: progress to Year 2 focussed on the early stages of the process, drawing on the experience of other Centres of Excellence/Centres of Competence and selecting a business modelling method to be applied. Some possible types of offering of the VCoE were identified (such as consultancy, training, collaboration broker, …). The emphasis was on the process of characterising the nature of the VCoE.

D11.4 Virtual Centre of Excellence development: progress to Year 3 presented an initial analysis of the context in which the VCoE would operate (political, economic, social and technological) and began to sketch the potential sources of demand. It reported on the application of the business modelling method through a series of workshops, looking at value propositions with respect to customer segments and potential revenue streams. It presented a financial plan with initial forecasts for income and costs, assuming that the VCoE was implemented as a membership organisation performing brokerage of services to its members as well as a certain amount of consultancy, training etc. offered by the central organisation itself. The deliverable also included an annex listing all the preliminary offerings of APARSEN partners that could be made available through the VCoE, under the headings consultancy/software products and support/training courses/other services.

This document sets out the blueprint for the VCoE and the processes to establish and consolidate it as a force in transforming digital preservation into a mature component of digital assets long term management practices.

Results shown in this document come from several sources:

- rounds of internal discussions among APARSEN team members
- data and analyses arising from literature review;
- participation of APARSEN team members in third party organised events (e.g. RDA, EGI, iPRES, ANIDP, EUDAT, MER)
- dedicated interviews with digital preservation stakeholders officers, mainly in the context of performing APARSEN sustainability tasks.

A market research exercise was conducted in the final year of the project which contributed strongly to the results. This exercise provided input on the analysis of the demand needs and market potential and on testing the plausibility of the services offering and ultimately the entire business model, prior to being fully deployed by the VCoE in the next few years.

The presentation of results is structured in two main blocks in this deliverable:

- The description and diagnosis of the environment of digital preservation as the context for the creation of the VCoE and source of inputs for the blueprint of the entire VCoE endeavour
- An initial view of the strategic configuration of the VCoE and the operational and financial planning needed for the first years of operation.

The first block (concentrated in chapters 2, 3 and 4) contains the overview of current state of results enabling:

- description of the overall context for the market development potential for digital preservation services
- characterisation of the potential demand that justifies the creation and operation of the VCoE
- overview of the competition (players and their offerings) that the VCoE should take into account
- identification of the main challenges (combination of opportunities and threats) to be addressed in the design and operation of the VCoE, to improve its chances of becoming a self-sustainable endeavour.

The second block (distributed in chapters 5 and 6) comprises the definition of the components of the VCoE business model, and some options available for implementing it.

2 THE LANDSCAPE FOR VCOE CREATION AND OPERATION

2.1 THE CONTEXT FOR DIGITAL PRESERVATION SERVICES

2.1.1 Dimensions of the context

This section addresses the *Political, Economic, Socio-cultural, and Technological factors*[9], i.e. the macro-environment variables[10] or *Market Shaping Forces* that provide the context for organisations to require digital preservation related services and be able to allocate resources for their acquisition. The value of this analysis rests on the fact that the future evolution in the key trends shapes the scenario for VCoE future sustainability.

Political conditions and governmental policy not only influence macroeconomic conditions; regulations and the political realm also affect organisations both directly and indirectly. Economic conditions affect how easy or difficult it is to be successful and profitable because they affect availability of capital, cost and demand. Increasing globalisation[11] presents a challenge to compete against products with world-class quality and financial strength, as well as to enter the markets of developing economies through joint ventures or partnerships:

- **Government Policy**: European policies tend to foster the use of connected ICT and legacy infrastructures, looking beyond the recent economic crisis to the need for preventing and overcoming similar conditions in the future through improved regulation and visibility.
- **Intellectual property rights, legal deposit of publications, and statutory requirements**: Legal issues that have impact on DP vary from country to country (including among EU countries) regarding copyright and other intellectual property rights (IPR) such as moral rights; access and security of content, as electronic materials which could be easily copied and re-distributed and/or statutory requirements.
- **Legal compliance**: The legal environment in which Enterprises operate affect positively the potential adoption of digital preservation technologies and big data technologies by firms as new regulatory policies will necessitate changes to IT systems and link legacy systems
- **New policies for sharing research data fully, openly and in timely manner**: Legal issues which impact on the operations of libraries, archives and other repositories vary from country to country (including among EU countries)
- **Political conditions and governmental policies influence the context for DP technology adoption**: The research landscape is going through big changes, with new and pressing policies from governments and funders to share data fully, openly, and in a timely manner

Economic conditions and government policy both influence a number of other macro environmental factors, such as how capital markets determine the conditions for alternative types of funding for organisations. They tend to be subject to government controls, and they will be guided by the prevailing economic conditions. Also, economic macro-environment includes job growth or unemployment. The labour market reflects the availability of specific skills at national and regional levels; this is affected by training, which is influenced by governmental action.

Economic trends affecting VCoE operations include:

- **Price driven services and competition for budgets**: Organizations and Governments are under a new normal: budget constraints, do-more-with-less while competing in a hypercompetitive and hyper-connected market.

[9] The PEST analysis. Aguilar, Francis. *Scanning the Business Environment*. New York: MacMillan.

[10] In the sense of expressing a lack of managerial control of such variables by DP market players

[11] See Friedman impact of flattening of the globe and growth of pace of globalisation. Friedman,Thomas L. "The Dell Theory of Conflict Prevention.": Barclay Barrios. Boston Bedford, St. Martins, 2008. Also, globalisation index calculated by the Swiss think tank KOF. http://globalisation.kof.ethz.ch/

- **Data driven economy and society**: The data-driven economy will stimulate research and innovation on data while leading to more business opportunities, productivity growth, and increased competitiveness in data across the whole economy
- **Service-based economy**: The growing trend toward a service-based economy is providing the path for knowledge-based business development as technological advances transform services; high growth has been experienced in the high-value-added "knowledge" sectors, such as biotech, software, and ICT.
- **Regional competitiveness**: Europe's knowledge base remains strong but needs a more strategic focus. The EU is facing increasing world competition, in particular at the higher end of global value chains.
- **Internationalisation, mergers and acquisitions**: Mergers and acquisitions are strategies that enable enterprises to thrive and ultimately excel within a competitive landscape, meet today's marketplace requirements and position the company in other geographical markets for future growth and sustainability

The **socio-cultural environment** comprises the demand, its attitudes and tastes, which depend on social dynamics and new cultural models[12].

Socio-cultural, behavioural transformation is influenced by the generalized shift from a pre-digital to a digital world, and the development of the new social dynamics and cultural models of the Information Society:

- ***Social disruption and change in user behaviour***: Cultural globalisation is supported and fostered by the Web, as the Internet is changing everyday life for 1.5 billion people worldwide, with their socio-economic standing influencing how they use the Web. Some of the new personality traits, values, attitudes, interests, or lifestyles include:
- ***The tendency to create active communities***: The dynamic growth of digital content includes user-created content generated by the rapid growth of Web 2.0 social networks, which have become commonplace, and the related phenomenon of collaborative creation (co-creation) of content
    - A gravitation toward social media sites where they can participate in discussions, share experiences and get involved in cultural conversations
    - A desire to be in control of their own lives, and a contentedness with complexity
    - A desire to work in more creative industries and be less restricted by rigid social structures.
- ***Social media as a primary method of communication and creative expression***: The dynamic growth of digital content includes user-created content generated by the rapid growth of Web 2.0 social networks, which have become commonplace, and the related phenomenon of collaborative creation (co-creation) of content. End-users are becoming active participants as well as information providers (about 230 million tweets per day and 100 terabytes of data uploaded daily to Facebook).
- ***Retention disruption***: There are several misconceptions regarding retention of information or data: people can jeopardise preservation schemes, particularly in relation with non-structured information (e.g. emails deleted right after reading) without understanding the value of keeping the information accessible; while on the opposite side, many people believe that "Digital media lasts forever"
- ***Increasing need of digital resources, consumption models and storage services***: Increasing usage of digital resources (e-books, digital audio, digital images, emails, video games, etc.)

[12] Alvermann D. (2004) 'Media, information communication technologies, and youth literacies - A cultural studies perspective' American Behavioral Scientist 48, 1, pp.78-83

generates need for digitalization or digital born resources to cover their consumption; as well as new digital "access" platforms and business models (e-lending, e-commerce, etc.).

- ***Globalisation and ppenness of knowledge for tackling social challenges***: Globalisation also has impact in science where the share of scientific outputs is growing rapidly, trespassing boundaries and going at a cross-discipline level. In addition increasing conditions for complying with open access policies to make content and data easily findable and fully usable by communities is heading towards solving major societal challenges such as climate change, resource scarcity, and demographic shifts.

Technology is part of the organisation and the industry, and it is also used for the creation of competitive advantage[13]. Effective knowledge management is crucial for enterprises and institutions, enabling them to identify, classify and preserve corporate memory and protect corporate assets, to assure sustainability and improve prospects for future success. Overflow of digital information as well as obsolescence of hardware, software and formats are concerns for creation of new DP solutions, which not only could create new industries whose products or services might be provided, but also provide improvements on manufacturing and service industries.

General technological trends include falling prices and improved performance of ICT products and services, the improved ability to store at lower cost and share information in standard formats, the vastly increased numbers of people online, and businesses routinely using the Web in increasingly knowledge-driven, customer-centric organizations.

Organizations are living under a new scenario:

Access to reliable information (preserved and current) has a central role in most organizations.

The business and cultural risk of missing or using wrong information is critical.

Organizations need fluid access to data and content that is growing exponentially.

Access and analysis of data and information (currently and in the future requires contextual information (entities, locations, relationships and concepts) that ensures their integrity, trust and reliability.

Four big trends affect ICT and drive 80% of its growth: mobile, cloud, social, and Big Data and Analytics technologies. Big Data market is foreseen to grow from $3.2 billion in 2010 to $16.9 billion in 2015. This represents a compound annual growth rate (CAGR) of 39.4% or about seven times that of the overall information and communication technology (ICT) market.

In addition the following trends are also of high impact:

- **Exponential growth of content and formats**: According to IDC[14] the information overload and the high cost of finding the right information are significant current issues. IDC expects that the "digital universe" will grow to 2.7ZB in 2012 to 8ZB in 2015 (almost 3 times in 3 years). Analogously, the Institute for the Future has reported that over the next decade, the infosphere will grow 40 times of today information levels, by 2020[15].

- **Scientific Data deluge**: Exponential growth for scientific data generation is happening and expected to continuo happening. Examples of e-Science data generated from sensors, satellites, high-performance computer simulations could be[16]: The UK e-Science programme analysing sensor data generated by Rolls Royce aero-engines estimates that there are around 100,000 Rolls Royce engines currently in service. Each trans-Atlantic flight made by each engine, for example, generates about a Gigabyte of data per engine from pressure, temperature and vibration sensors

[13] ICT for competitiveness and innovation. http://ec.europa.eu/enterprise/ict/index_en.htm

[14] IDC Market Analysis for Bid Data Technology and services (2012-2015 forecast). March 2012 Volume I.

[15] Information Intensification: The automation of awareness. 2012 Forecast. The Institute for the future. 2012. SR-1473. www.iftf.org

[16] Hey, Tony; Trefethen Anne. **The Data Deluge: An e-Science Perspective;** UK e-Science Core Programme. http://users.ecs.soton.ac.uk/ajgh/DataDeluge(final).pdf

- **'Everything as a Service' and Cloud Computing**: Serviced based economy, drives for flexible and dynamic composition of services, processes, interfaces, reports and electronic forms, so as for customised serviced based solutions. Thus trends such as the following are emerging: Software as Service (SaaS), Service-Oriented Architecture (SOA), and Computing as a service or "Cloud Computing" storage as a service.
- **Storage and Data Management**: Storage demands and capabilities are increasing very rapidly, applications with requirements for petabyte storage are emerging in pharmaceutical research, manufacturing simulations and testing, medical imaging or digital media. According to our findings data volumes grow around 40% to 50% every year
- **Network Connectivity and Mobility**: the network connectivity boosts the increasing usage of mobile devices and objects which at the same time become very important. This supports interconnecting or exchanging resources in libraries, archives or Scientific Research Organisation
- **Security, Trust and integrity**: Security and trust challenges are coupled with the ability of how data will be managed over time as well as providing safe decentralised environments for large and heterogeneous number of users (probably in indifferent geographic regions) and content integrity, especially when new trends such as cloud computing are emerging.

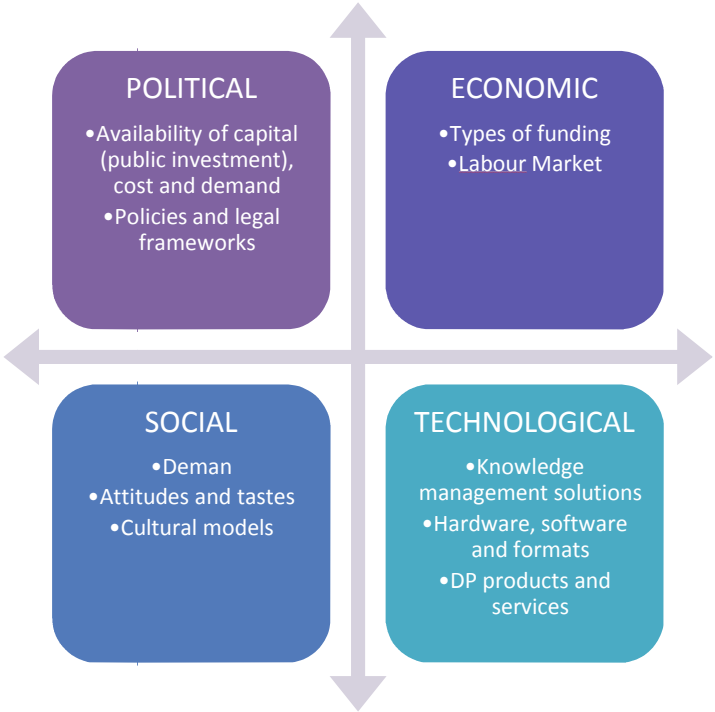


Figure 1: DP market shaping factors

The following paragraphs provide an overview of the competitive context for the VCOE creation, drilling across the four mentioned dimensions:

2.1.2 The political dimension

Political impact: Politicians' attitudes to the digital world are changing especially when they are increasingly faced with economic difficulties. Nationally[17] and internationally[18] one sees a demand for pay-back on public investment in research, through making the results, in terms of public availability both to journal publications as well as to the data which underpins them.

Commissioner Kroes has declared "Data is the new gold"[19], to underline the value of data. Nevertheless it is relevant to point out that gold is valuable because it is rare and does not combine with other elements, whereas data is valuable because there is so much and it is more valuable when it is combined.

This value indirectly drives the need for digital preservation because:

i. pay-back can take time and
ii. pay-back from data usually comes from combining data obtained at different times from different domains[20].

Despite all the above there remain legal impediments to digital preservation as described next.

<u>European level legal framework</u>

Currently there is a lack of comprehensive, coherent, legislation about long term preservation of scientific data. Existing legal frameworks for data preservation are concerned with:

1. privacy of personal data;
2. telecommunications data, which is available as a result of data retention obligations;
3. the Cybercrime Convention (the Budapest Convention on Cybercrime, http://conventions.coe.int/Treaty/EN/Treaties/html/185.htm)

The European policy concerned with scientific data preservation is almost all embodied in documents such as "Riding the wave, How Europe can gain from the rising tide of scientific data" available at http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf, which addresses the opportunities and identifies a number of important issues. Nevertheless those documents have not been translated into a legislation framework.

<u>National legal framework</u>

The numerous national levels do not have specific legal frameworks concerned with scientific data preservation except for ones relevant to the police (anticrime) domain as per national frameworks and relationships with international agreements (e.g. http://www.cepol.europa.eu).

<u>Intellectual property rights</u>

The length of copyright has been increasing, and the investment in and economic value of IPR has also increased. The commercial need to protect IPR can become more important than other considerations about the cultural heritage preservation. The needs of memory institutions for legal exceptions to undertake archiving are often overlooked or not sufficiently understood. Legislation concerned with IPR is defined while differences between countries are in place.

---

[17] For example RCUK Policy http://www.rcuk.ac.uk/documents/documents/RCUKOpenAccessPolicy.pdf
[18] EU Open Access
http://ec.europa.eu/research/science-society/index.cfm?fuseaction=public.topic&id=1294&lang=1
[19] Opening Remarks, Press Conference on Open Data Strategy
Brussels, 12th December 2011
http://europa.eu/rapid/press-release_SPEECH-11-872_en.htm
[20] Ref. ODE (Opportunities for Data Exchanged) Project:
http://www.alliancepermanentaccess.org/index.php/community/current-projects/ode/

Almost all policies and implementations relevant to the scientific data preservation are ***delegated to the relevant scientific organisations***, their different levels of committees and boards, and associated with specific agreements between communities and groups. Generally speaking, there are some practical solutions and guidelines but there is no "common set of rules".

Of course, this is a weak point in the design, implementation and provision of services concerned with long term data preservation because the absence of a common rules framework producse a kind of 'Tower of Babel' in the solutions landscape, forcing a great deal of effort to be spent on interoperability and interfaces. Unless these legal problems are solved then this will seriously impede true long term preservation in several areas.

2.1.3 The economic dimension

Economic Impact: A knowledge economy may refer to either the economy of producing and management of knowledge, or, more usually, a knowledge-based economy by which is meant the use of knowledge to produce economic benefits. We use the latter meaning.

Knowledge penetrates all fields of industry; it is at the base of innovation, effectiveness and therefore of the competitive advantage that all economic organisations seek. This proves the importance of the role that knowledge has in current economy. Some of the relevant economic drivers for digital preservation could be: the need to protect investments and the opportunities of future reuse to reach specific organisational objectives as well as the retention of the business efficiency. In fact, most of the data represents a strategic heritage of each organisation, in terms of internal procedures, transactions, acts, and process and product know-how, which is used for creating new products or services. Effective and affordable digital preservation strategies and systems will transform archives into digital assets to be re-used. This is particularly true in the various research domains. Digital preservation process may be fundamental in R&D phases, where the speed of development of a new product is crucial for determining the success or the failure of a new product. This happens usually in the pharmaceutical industry where the time for a new drug to reach the market is one critical factor of profitability.

However, increasingly we see that commercial concerns see the advantage of making their information open in order to encourage serendipitous discoveries (http://www.genengnews.com/insight-and-intelligenceand153/big-pharma-s-open-innovation-initiatives-zoom-in-on-discovery/77899468/). The requirement here is that the digital information is both accessible and also understandable/usable. The records management and archive worlds have much in common with the preservation world. One key difference has been the stress put on data lifecycle and the need to delete unnecessary digital holdings, driven initially by the cost of storage but increasingly by the cost of legal discovery[21]. Nevertheless even in the records management world what was referred to as "eTrash" (i.e. to be deleted) is being looked at as potential "eTreasure" (i.e. worth Mining). Even the costs of legal discovery are being reduced through the use of "predictive coding", using computer techniques to replace much of the effort which has been needed from expensive humans (lawyers). At the moment this computerisation deals essentially with text Mining but one can foresee this being applied to data.

The other side of economic benefits are the economic costs of preservation, and perhaps more importantly, the long term commitment to sustain these costs – which no organisation can realistically provide. One way around this, as proposed by the Blue Ribbon Task Force[22], is that instead of commitment to preservation into the indefinite future, one should speak about buying an option for the future. In other words one does what is necessary now, especially in terms of capturing knowledge, so that a decision may be taken at some defined future point as to whether to:

[21] MER 2013 Conference http://www.merconference.com/details/schedule.php

[22] Sustainable Economics for a Digital Planet: Ensuring Long-Term Access to Digital Information. Final Report of the Blue Ribbon Task Force on Sustainable Digital Preservation and Access. February 2010 http://brtf.sdsc.edu/biblio/BRTF_Final_Report.pdf

- continue to preserve for the next period or
- hand over to some other body or
- allow to gradually "rot" or
- delete.

2.1.4 The social dimension

Social Impact: The last decades have seen an extremely rapid development in the computerization of society. The phenomenon is so massive, and changes so influential in the development of our civilization, that we now speak about the Information Society. The volume and range of information produced is expanding more and more. Now, digital publications in many countries complemented, but have not replaced, traditional publication. This increase in both traditional and digitally encoded information is challenging national institutions, particularly national libraries, to introduce new solutions in order to manage information and knowledge. At the same time, many of the traditional filtering and editing roles of publishers are disappearing as the Web opens publishing to individuals and organisations. For this reason, it seems necessary to find new institutional bodies that will deal with appraising and selecting records to preserve. Currently, many libraries are undertaking such commitment. This exponential increase in information is not confined to publishing; *it applies even more to data in the academic and research sectors, particularly in the sciences.* Publications are now only one aspect of popular culture and the cultural record. Film, television, and the World Wide Web define an increasing part of our culture. Mechanisms to consider new areas of collection development and future research needs may be required as part of any national scheme.

Activities increasingly take place on a global scale and outside the traditional national frameworks, in general, for all sectors, and also for digital preservation. With the development of international publishers who can deliver their digital publications from anywhere, the role of archiving in a national context is less clear. Moreover, the growth of the Web and the international activity it empowers to transcend national boundaries. In similar way, globalization also applies to developments in hardware and software. Information technology (IT) companies and market trends operate on a global scale and apply to many different sectors. This means that there is more substantial common ground between institutions internationally and across sectors and greater potential for and benefits from international collaboration. The increasing access to information by members of society also places the onus on applications and in particular mobile apps to be able to deal with data from various sources[23]. The growth of Linked Open Data helps in this. This provides challenges in areas of automated use of data and of preservation of links[24].

2.1.5 The technological dimension

Technological Impact: rapid technological changes are considered something "normal" today and in general a technological evolution is expected almost every month. This means a constant and continuous production and introduction in the market of new software, new hardware as well as new file types, new semantics and so on. A direct consequence of this is the fast obsolescence of hardware, software and media obsolescence. A less obvious consequence is the evolution of terminology and tacit knowledge.

The Internet, the explosive deployment and consumption growth of Broadband infrastructure and information services hubs (e.g. Google), and the dramatic shift in the ways and methods in which information is created are among the critical factors explaining the need for DP. There is an explosive growth in the volume of digital contents generated on an annual basis, a multiplication of their sources

[23] APPLIFISH http://www.i-marine.eu/Content/Multimedia.aspx?id=5ffcfcdf-df2e-4960-b09e-810be78fe047

[24] PRELIDA http://www.prelida.eu/

(authors and production devices) and a proliferation of specific technical formats. Nowadays, a large quantity of information exists in a diversity of digital forms: emails, blogs, books, enterprise records, legal documents, papers, videos, pictures, etc., being created in different scenarios from a wide range of authors, an unprecedented scope of the usage. This creates additional management burdens due to the increasing heterogeneity of digital content and assets.

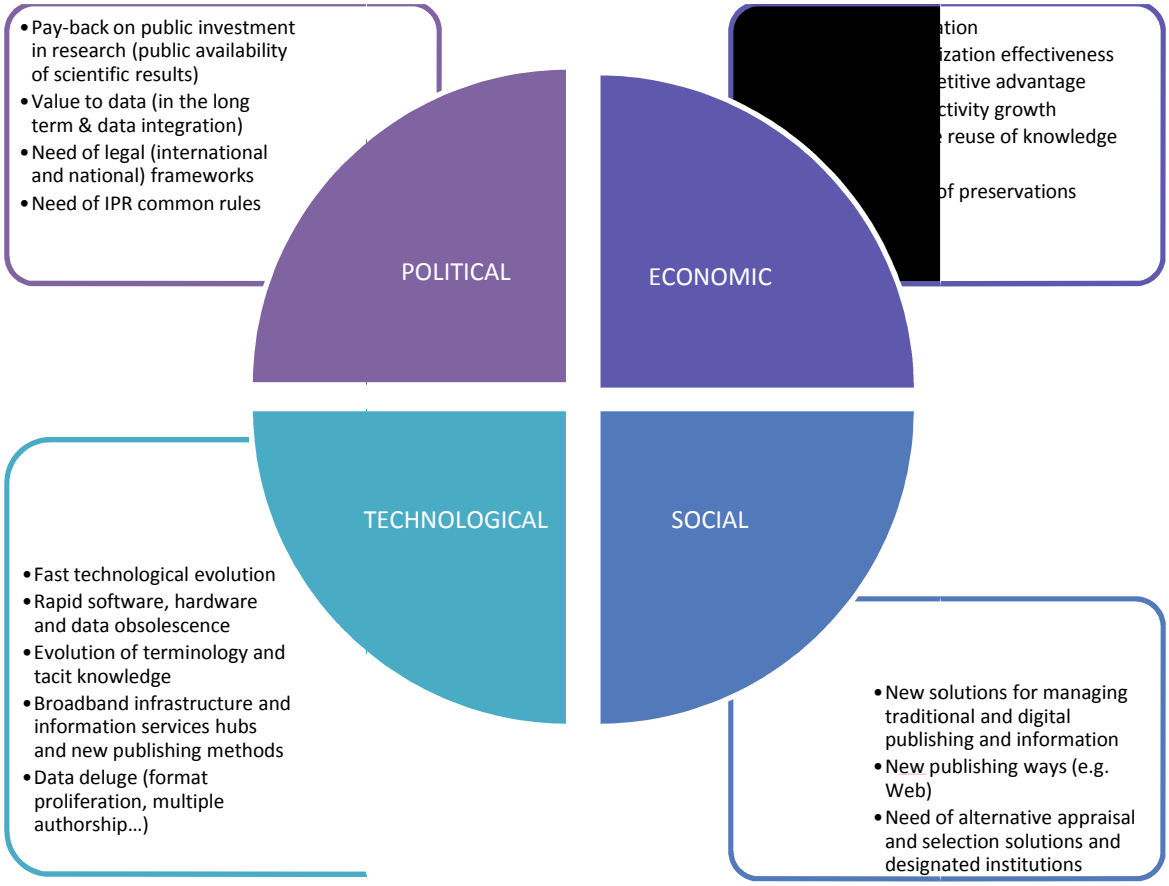


Figure 2: Impact Analysis

Looking into the future, the progressive penetration of the Information Society into each and all aspect of social and economic activities, allows drawing the vision of a wider role of DP in the years to come.

2.2 POTENTIAL DEMAND FOR VCOE OFFERINGS

This section provides information and analytical elements to define the existence of business opportunities for offering a range of DP related services. In particular, the kind of DP products and services offering envisioned to be provided through an ecosystem of trusted suppliers that would include the APARSEN VCoE. In this regard, in the following subsections we provide an overview of:

a) The need for DP practices
b) The business case for products and services supporting the implementation of DP practices

This section covers the main groups that represent market players which potentially could be considered as part of the demand to be addressed including: Scientific Research, Cultural Heritage/Memory Institutions (museums, libraries and archives) and Industry.

2.2.1 Scientific and Research Institutions (SRI)

The SRI sector comprises two types of institutions: Government Research Institutions and University Research Institutions.

- Government Research Institutions comprise organisation such as:
  - European institutions: European Organization for Nuclear Research (CERN)
  - At national level usually central research institutions or laboratories attached to different ministries:
    - Central Science institutions, agencies or laboratories that depend from Environment, Food and Rural Affairs: e.g. Centre National de la Recherche Scientifique (CNRS), the Central Science Laboratory, the Centre for Environment, Fisheries and Aquaculture Science or the Veterinary Laboratories Agency, or Geological/Ecological National Councils
    - Central Science institutions, agencies or laboratories: Department of Health (central health institutions or National Radiological Protection laboratories), Central Science institutions, agencies or laboratories from Ministry of Defence such as the Atomic Weapons agencies/establishments or Weapon and Defiance Science and Technology Laboratories.
- University research institutions, mainly large science oriented and research centres that may specialize in basic research or may be oriented to applied research; these includes natural science and social science research. Some European institutions in this area are: The University of Amsterdam, Ecole Polytechnique Federale de Lausanne, University of Cambridge, University of Edinburgh, University of Freiburg, Danmarks Tekniske Universitet or KU Leuven.

According to the Joint Research Centre report[25] there are 24743 Research Universities and 19068 SRI that participated in the 6th Research Framework Program (FP6). According to the European Research Ranking[26], the number of SRI in Europe is larger, as this body reports that at least 31100 SRI in Europe have participated in funded projects.

These organisations gather almost 0.9 million full time equivalent (FTE) researchers in 2009, with 4% of annual growth and employed in public laboratories[27]. To improve conditions and access to finance for research and innovation, and support the creation of growth and jobs in Europe in the year 2020;

[25] Henriques, Luisa; Schoen, Antoine; Pontikakis, Dimitrios . Europe's top research universities in FP6: scope and drivers of participation. JRC-IPTS. 2009. EUR 24006 EN

[26] http://researchranking.org/index.php?action=partner&p=deb

[27] Deloitte. DG Research and Innovation Researchers. Researchers' Report 2012. European Commision 2012.http://ec.europa.eu/euraxess/pdf/research_policies/121003_The_Researchers_Report_2012_FINAL_REPORT.pdf.

Europe aims at achieving the target of spending 3% of EU GDP on R&D by this year; which in turn could create 3.7 million jobs and increase annual GDP by close to EUR 800 billion by 2025. To reach this target, Europe additionally needs at least 1 million researchers; while challenging an increase in its ageing population.

SRI Communities and Infrastructures

ESFRI`s roadmap comprises a list of 48 infrastructures; of which ten of them are under implementation phase. ESFRI is committed to reach 60% of the needed research infrastructure by 2015, through new pan-European Research Infrastructures or major upgrades of current ones; 50 projects have been approved from 260 proposals since 2006 to reach current stage[28]. e-Infrastructures like EGI bring together +300 EU driven e-infrastructures sites or organisations.

The following are communities that manage grid-based or networked e-infrastructures:

- The European Grid Infrastructure (EGI) is a federation of independent national and community resource providers, comprising 340 organisations providing computing centres in 34 countries (National Grid Initiatives – NGIs-), two European Intergovernmental Research Organisations (EIROS), which are CERN and EMBL. It provides 373,800 CPUs and storage capacity of 190pB, 180 PB tapes. It provides capabilities 212 Virtual Organisation (VO) that comprise 22,000 researchers and scientists. Each VO has its own rules as to who can join their community and membership may be open to non-European collaborators.
- The Open Science Grid (OSG) comprises 72 institutions including universities, service and resource providers, and national laboratories, as well as computing centres from 115 geographical sites. Multidisciplinary collaborations from disciplines like biology, chemistry, astronomy, and geographic information systems dedicate 300,000 computing-hours per day to the analysis of data from particle colliders. OSG comprises over 25,000 computers with over 43,000 processors. The research performed using the grid's resources have been published in the Journal of Physical Chemistry.
- GÉANT connects over 50 million users (of which 40 million are European from 43 European NRENs) in real time, from 10,000 institutions of 65 countries (40 European) and covering 50,000 km of network infrastructure across Europe, operating at speeds of up to 500Gbp and facilitates the generation of new user communities in emerging areas of science[29]. After 10 years GEANT has been re-evaluated to achieve the Europe 2020 goals and remain at the forefront of the scientific and social developments. “GÉANT 2020”
- The Finnish University and Research Network (FUNET)[30] is an advanced data communications network serving the Finnish research community. It connects about 80 research organizations and over 350 000 users that work over a superfast data communication network.
- Partnership for Advanced Computer in Europe (PRACE)[31]. PRACE looks forward the creation of an HPC ecosystem in Europe integrating services in the entire infrastructure as well as the hardware and software industry, with independent access to HPC-systems. PRACE has 20 members representing 20 European countries. Funded has been secured from 2010 to 2015 (€400 million from France, Germany, Italy, Spain; +€70M from EC FP7 for preparatory and implementation and also a new funding decision for €100 million from the Netherlands is

[28] ESFRI. European Research Infrastructures with global impact report. 2012. http://www.copori.eu/_media/ESFRI_Brochure_210912_lowres.pdf

[29] Geant Homepage: http://www.geant.net/Pages/default.aspx; and Cavalli, Valentino. GÉANT: Enabling European Collaboration presentation. May 2013. http://www.ceenet.org/wp-content/uploads/2013/05/G%C3%89ANT-presentation-CEEUSER.pdf

[30] http://www.csc.fi/english/institutions/funet/about/customers/index_html

[31] Homepage: http://www.prace-ri.eu/documents/

expected. PRACE targets Tier 0 (Petaflops), reaching an accumulated capability of more than10 PF by 2013. Through PRACE execution of core hours are granted to projects; for example for the project Quantum Chromo Dynamics, which objective was the determination of the equation of state of QCD 63,000,000 core hours were granted to this project from 08-2010 to 11-2011.

Repositories

These institutions manage digital content mainly through two types of infrastructures/organisational instruments: the *Institutional Repositories* (*IRs*) and the *World Data Systems* (*WDSs*).

IRs could be open access or non-open access. Publication process is the same in both cases. Open access provides free of charge access to scientific information (peer reviewed articles and scientific data) increasing the scale of dissemination, exchange of information or collaboration, accelerates scientific discovery, enables new forms of data-intensive research and allows industry to reach scientific research results.

In the ***OpenDOAR directory*** 1200 IR[32] in Europe with more than 74.000 resources have been identified. From a geographical point of view, 80% of the repositories are located in 10 EU countries

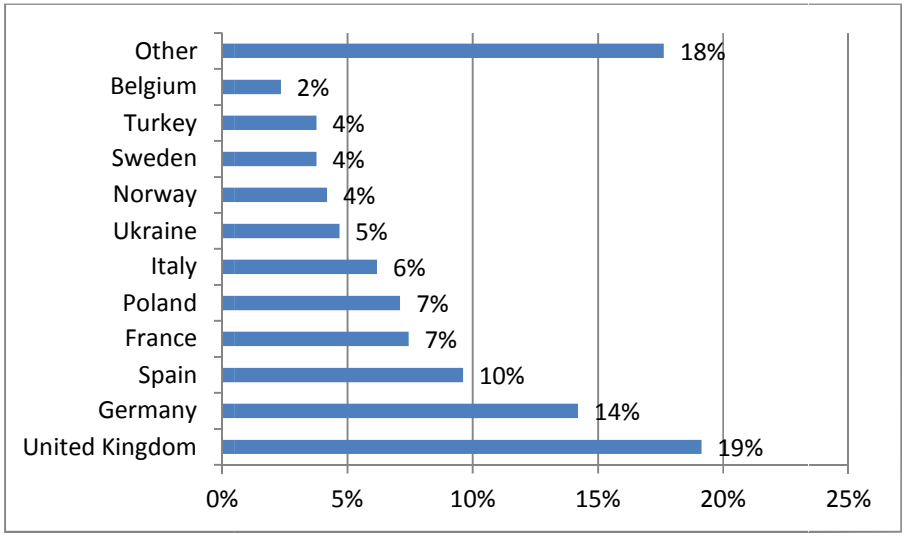


Figure 3:Distribution of open repositories by country

OpenDOAR provides tools for preservation policies are part to the governance. The Policies Tool provides structure to create five types of policies: Metadata Policy, Data Policy, Content Policy, Submission Policy and Preservation Policy. All elements of the OpenDOAR project are meant to promote an Open Access standard, complementing the OAI standard[33].

The "Ranking Web of World repositories"[34] which aims to provides a list of mainly research-oriented repositories arranged according a composite index derived from their web presence and the web impact of their contents, data obtained from the major commercial search engines, lists 746 repositories in Europe, of which 650 are in the EU out of a total of 1654 in the world.

[32] http://www.opendoar.org/. OpenDOAR is an authoritative directory of academic open access repositories
[33] http://www.dcc.ac.uk/resources/external/opendoar#sthash.uLIxTRdE.dpuf
[34] http://repositories.webometrics.info/en

"Repository 66"[35] lists 3,045 repositories in the world; of which 1,161 are located in Europe. This repository has grown 46% from 2008 to 2014[36.] The numbers of Open Access Journals have almost doubled from 2009 to 2012 (from 4360 to 8115 according the DOAJ).

OpenAIRE (Open Access Infrastructure for Research in Europe)[37] has been funded by the EC and supports the implementation of open access in Europe by providing an infrastructure and national helpdesks. OpenAIRE currently identifies over 27.000 FP7 publications, some 9.500 of which are open access and some 16.800 of which are still under embargo. The EC also has implemented open access to research results from

The World Data System (WDS) is an Interdisciplinary Body of the International Council for Science (ICSU) created in 2008, to build on the 50+ year legacy of the World Data Centre (WDC) and Federation of Astronomical and Geophysical data analysis Services. ICSU World Data System promotes universal and equitable access to, and long-term stewardship of, quality-assured scientific data and data services, products, and information covering a broad range of disciplines from the natural and social sciences, and humanities, coordinates trusted scientific data services for the provision (use and preservation of relevant datasets) and it is striving to build worldwide 'communities of excellence' for scientific data services by certifying member organizations.

WDS has 86 Member organizations (July 2014), including 10 Network Members, 3 Partner Members and 17 Associate Members. WDS has strong collaboration with ICSU's Committee on Data for Science and Technology (CODATA).

Model of data production and consumption

Figure 4 shows how data producers and data consumers interact in the scientific production value chain, where value is added to data producers by providing financially services and data, community and user services under an e-infrastructure environment (sharing services); also value is added by outputs provided by other players that complete the value chain (publishers, data holders/repositories, and or distribution of digital objects –libraries-).

Data producers domain (scientists and research workers in their various institutions create data, and who will submit this data to organizations under the e-infrastructure "research community" which provide data, community or curation services (shown in the centre of the image). Data curated within this research community domain will be accessed by researchers and others (consumers), as shown on the right. In this case, e-Science researchers could act as producers and consumers. Also the other consumers (citizens or industrial workers) may be contemporaneous with the data producers, or could be separated from the producers by many years.

[35] Repository 66 - http://maps.repository66.org/
[36] Comparing with research from DPimpact study.
[37] Homepage: https://www.openaire.eu/

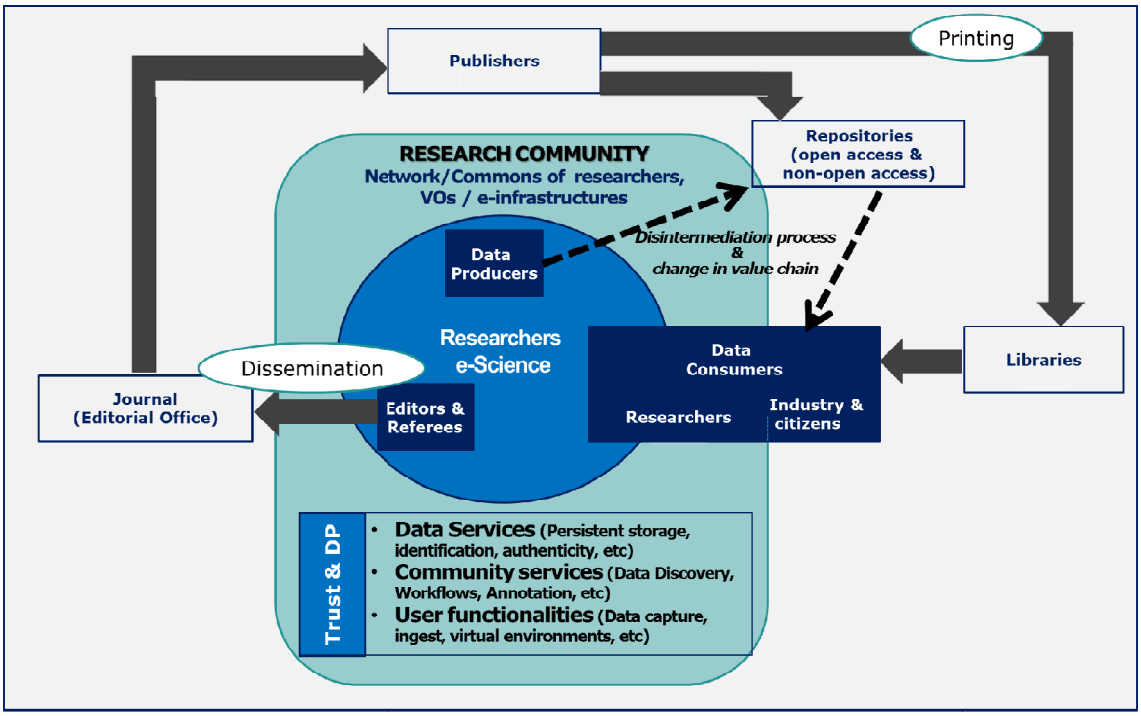


Figure 4: Model of data production and consumption[38]

Increasing e-publication trend

The publishing sector in Europe is a strategic sector for the EU. This sector has a key economic role to play and is vital to the development and preservation of culture, information, education and democracy at large in Europe. This sector is a source of growth and employment, as well as an essential tool for education. 25,400 peer-reviewed journals are published worldwide by 2,000 different publishers; however top 5 publishers account for 26% of the journals[39].

Sales revenue of book publishers of the EU and the EEA in 2011 was approximately €22.8 billion, according to FEP (Federation of European Publishers); the largest markets (by turnover) in 2011 were Germany, UK, France, Spain and Italy. FEP estimates that the publishing industry account for 0.5 % of the GDP across the EU27 Member States, employing around 750.000 jobs in more than 64.000 companies. Book publishing is the largest European cultural industry[40].

Future Internet technologies, cloud computing, smart devices and other developments related to technical migration represent a significant challenge for the publishing sector, as well as an opportunity. **Specifically, opportunities will be discovered as it is understood that digital transition in the sector of publishing is not restricted to e-books but to all the value chain** from authoring, editing, production and printing till marketing and distribution. In 2012 e-books sales surpassed the sales of hard cover books` sales.

The number of digital journals increased dramatically as well as the internet distribution of journals in few years; this shift can be illustrated with the Elsevier case, which claims to have the third largest Internet revenues, behind only AOL-Time Warner and Amazon[41].

In 2011 Academic and professional publications reached 50% of total sales. In the case of the Scientific, Technical and Medical (STM) the market generated from English language STM journal

[38] This model is based on the publishing cycle from the STM (STM report P18) and applying the concept of collaborative data infrastructure in the global e-infrastructure platform presented in the report Riding the Wave. How Europe can gain from the rising tide of scientific data. October 2010. (http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf P31).

[3939] PARSE.Insight. D3.6 p23/51. Reference to the STM Report.

[40] Turin, Enrico; The publishing sector in Europe. TISP. Federation of European Publishers.

[41] Edlin, Aaron; Rubinfeld, Daniel. The exclusión or efficient pricing? The Big Deal Bundling of academic journals.

publishing was almost €7 billion in 2011; about 32% of these revenues came from Europe (USA dominates the global output of research papers). The industry employs 110,000 people globally (40% employed in EU) and 20-30,000 full time employees indirectly. In 2012 there were about 28,000 peer reviewed journals, and collectively publishing about 1.8-1.9 million articles/year; growing about 3%/year. For the Journals of Economics the market in 2011 was estimated by Outsell at €16 billion, including journal, books, technical information and standards, databases and tools.

The STM book market worth €2,8 billion annually; e-books make about 17% of the market and growing faster than other areas[42]. Germany and France are ahead in digitally embracing books. According to the Outsell report the global market for e-books is €469 million with 23% of growth (17% of the STM market) in 2011. For e-books there is a considerable business model innovation associated to the educational market, potentially disruptive and including freemium models, class support tools, printing testing, etc. there has been a recent growth for open educational resources (OERs).

Regarding funding of DP in scientific publishing, most of cultural organisations have reported in PARSE Insight study (D3.6) that DP should be funded either by publishers or research organisations.

2.2.2 Memory Institutions

Memory institutions comprise Museums, Libraries and Archives. In Libraries and Museums the degree of exposure to DP challenges differ according to the size of the institution, its specific institutional mandate, and the type of digital content managed, among other criteria. Notwithstanding this, due to the digital content deluge and large efforts of content digitalization, libraries and museums are part of an increasing demand for DP.

National Libraries (71%), SRIs (60%), Research and Academic Libraries (56%) are regarded as important players of DP and it is their choice to take the responsibility of DP. Moreover, in the case of memory institutions, National archives and national libraries lead the DP initiatives since 5 years ago driven by the force of having high volume of information and wide variety of content. According to the Planets Project[43] in 2009 67% of national archives and 55% of national libraries have DP policies implemented; 81% of national archives and 65% of national libraries have budget for DP and 27% of National Archives and 15% of National Libraries have a DP solution in place. At the same time, when asking for who should fund DP activities, these organisations believe that Governments, research funders and EU should be paying for the DP.

<u>Museums</u>

European museums play an important role in showing the richness and diversity of cultures however, establishing the number of museums is very difficult; still a common understanding is needed. What is considered a museum in Spain may not be seen as such in Hungary or Finland. This huge difference arises from the lack of harmonisation in the definition of the term “museum” across and within the Member States. This results in Germany having some 6.000 museums, UK a number between 1.200 and 2.500 according to different official sources and France between 1.300 and 8.000.

Following the collection and comparison analysis of the works of the European Group on Museum Statistics (EGMUS), according to its statistics the number of Museums could be estimated between 151,361 and some 24.000 institutions[44]. Segmenting by sizes we found that there are about 4% of large state owned museums, with the Louvre being clearly the EU leader; 96% of the museums are local regional owned museums, private ones or other small public museums.

[42] Ware Mark. The STM Report. November 2012. An overview of scientific and scholarly journal publishing. http://www.stm-assoc.org/2012_12_11_STM_Report_2012.pdf

[43] www.planets-project.eu

[44] http://www.egmus.eu/en/statistics/

Libraries

Today, Europeana, the portal to Europe's digital cultural heritage, provides access to over 23 Million objects from more than 2.200 institutions (including museums, archives, audio-visual archives and libraries) from 33 countries[45]. Finally, in a recent report[46] on the results of a survey on "users' perceptions of the benefits of ICT in public libraries", the desk research estimates that there are more than 65,000 public libraries across the whole of the EU.

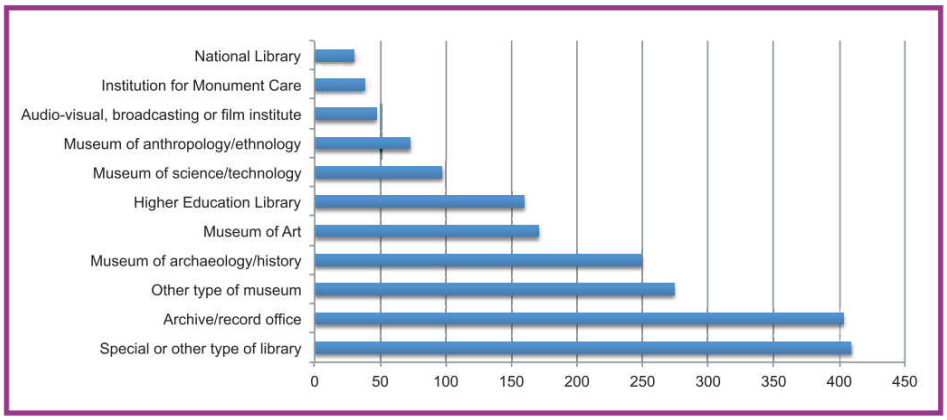


Figure 5: Distribution by organisation type from the first ENUMERATE survey

Statistics indicate that there are some 570 000 libraries worldwide, out of which 341 are classified as National Libraries[47]. Within Europe, the total number is around 81 000 and only some 50 institutions rank as National Libraries, including among them big organisations with more than 1000 employees, such as the British Library, the Deutsche Nationalbibliothek or the Bibliotheque Nationale de France, but also small organisations, according to the relative size of the country.

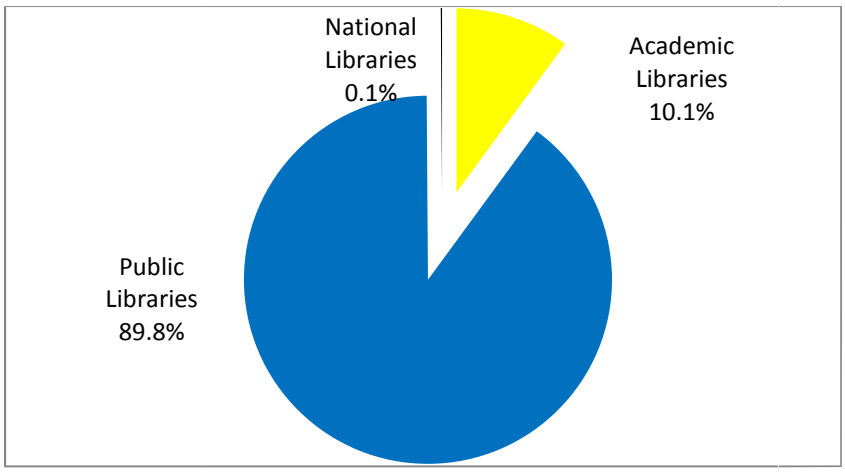


Figure 6: Types of libraries in Europe (source: ALBA, Association of Leading Visitors Attractions, 2013)

Similarly, as illustrated in USA there are some 120 000 Libraries, out of which some 1000 can be considered big organisations, including the Library of Congress and the libraries of leading

[45] http://conference.ifla.org/past/ifla78/181-niggemann-en.pdf
[46] http://www.minedu.fi/export/sites/default/OPM/Kirjastot/kansainvaelinen_ja_eu-yhteistyoe/Liitteet/Final_Report_-_Cross-European_Library_Impact.pdf
[47] Global Library Statistics (2003) – International Federation of Library Associations and Institutions

Universities and some State Libraries in the first rank. In 2001, investment in public libraries reached almost €8billions, while its operating expenses surpassed €7billion[48].

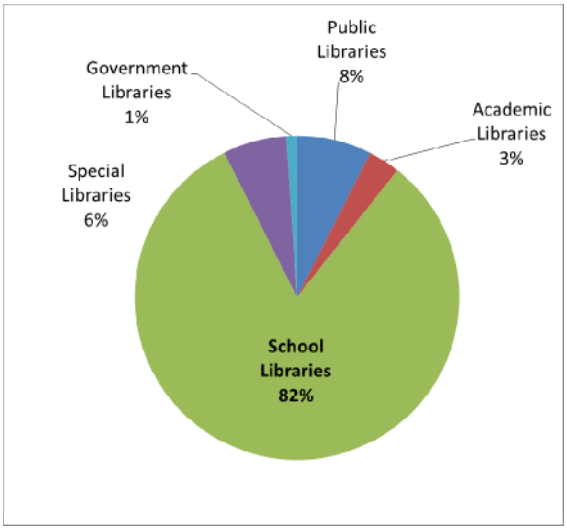


Figure 7: Types of libraries in the US according to ADL (ALA, 2013)[49]

The sales of book titles are directly related to the increasing importance of sales of bundles to library consortia. The consortia approach arose in order to provide efficiencies by centralising services (e.g. shared library management systems, catalogues, ILL, resources etc.) and centralising purchasing for increasing the negotiation power of libraries with publishers; due to the fact that prices, ownership concentration, and the number of journals have increased dramatically. In this case internet has become a barrier, as the electronic version should go cheaper. To maintain margins under the e-publishing trend, publishers have begun offering bundled packages known as the "Big Deal" contract, where the libraries enters into a long-term agreement; where libraries get access to a large digital library of journals for a discounted price while promising not to cut print subscriptions. [50] Consortia examples are: The Ringgold Consortia[51] Directory Online lists over 400 consortia representing over 26500 individual libraries worldwide; of which about 350 are responsible for licensing content. The International Coalition of Library Consortia comprises 200 libraries. In general the size and nature of consortia corresponds to the size segmentation, as it also varies depending on the sector (e.g. academic, medical, etc.) and status (regional vs. local or public vs. private).

Increasing usage of digital resources (e-books, digital audio, digital images, emails, video games, etc.) generates need for digitalization or digital born resources to cover their consumption; as well as new digital "access" platforms and business models (e-lending, e-commerce, etc.). Thus currently libraries need not only to go towards the development of workable models of e-lending with trade publishers and other stakeholders in the book industry, but also to cope with coming needs from e-learning and digital programs from schools and Universities; which already are moving to e-books. In addition social media and networks trends are also having impact in the integration of social platform into software and web platforms; such at it is the case of Mendeley (free reference manager) and ResearchGate (network dedicated to science and research. Connect, collaborate and discover scientific publications, jobs and conferences.).

[48] Library operating expenditures. A selected annotated bibliography. ALA Library fact sheet. http://www.ala.org/tools/libfactsheets/alalibraryfactsheet04

[49] Holton, B., Hardesty, L., and O'Shea, P. The Academic Libraries: 2006 First Look. http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2008337

[50] Edlin, Aaron; Rubinfeld, Daniel. The exclusión or efficient pricing? The Big Deal Bundling of academic journals.

[51] Homepage: http://www.ringgold.com/pages/cdo.html

Gartner[52] says that the worldwide consumer growth of digital content stored in PCs, smartphones, tablets, hard-disk drives (HDDs), network attached storage (NAS) and cloud repositories. There is an increasing trend in mobility, using tablets and mobile devices in libraries.

There is a decreasing trend of number of visitors (between -3% and -7%) in public libraries in Europe[53].

Archives

Modern archive services need to provide continuity, stability and a solid base for essential information and indispensable documents and archives. A coordinated system of support for local government archives that includes the National Archives and Records Administration, state archives, national and regional archives associations, and local governments and their associations. As a rough estimation, the number of archives in the EU is at least 16,000 units[54], while the order of magnitude worldwide is between 40 to 50 thousand organisations. Out of this total, some 1000 could be considered as national level archives.

There is strong anecdotal evidence for champion leading funding strategies; however champions are an essential element of local government archives programs.

2.2.3 Industry

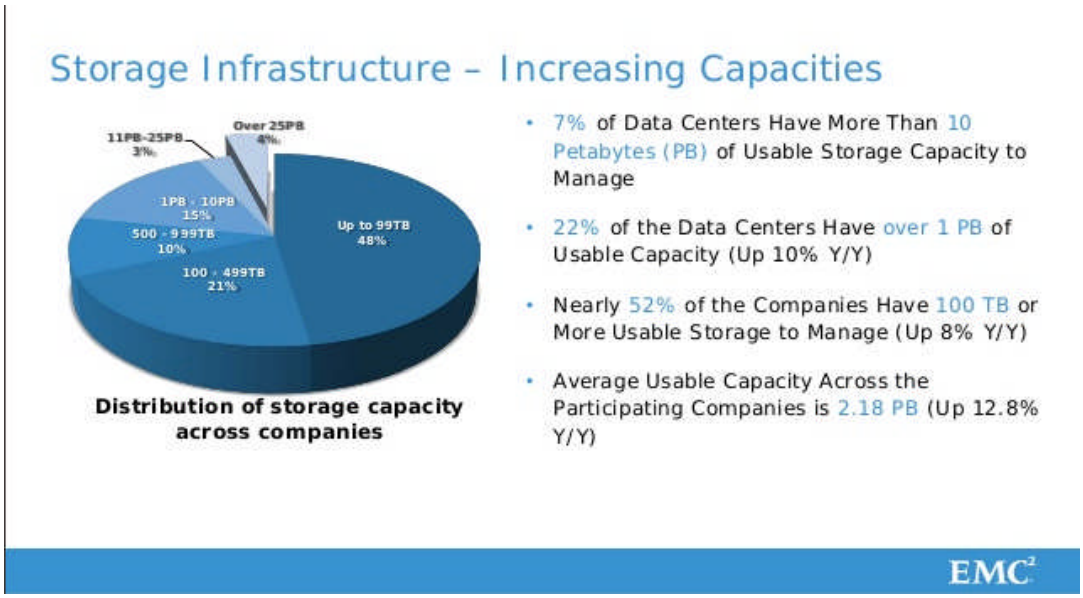


Figure 8: Capacity of storage infrastructure (Managing Storage: Trends, Challenges, and Options (2013-2014), EMC, 2013)

[52] http://www.gartner.com/newsroom/id/2060215

[53] COMPENDIUM. Cultural trends and policies in Europe. http://www.culturalpolicies.net/web/monitoring-laws-policies.php?aid=98&cid=116&lid=en

[54] Enseñat, Luis. The establishment and maintenance of an internet portal for documents and archives in Europe (APENET) Archives Portal of Europe on the Internet. 2008. http://www.arhiv.gov.si/fileadmin/arhiv.gov.si/pageuploads/EU/EAG_Portal.pdf

Related Sectors:

Enterprise organizations have adopted several practices for preserving needed information:

Information Lifecycle Management (ILM): This practice has been used by Records and Information Management (RIM).

Information life cycle management is an approach to data and storage management that recognizes that the value of information changes over time and that it must be managed accordingly. ILM seeks to classify data according to its business value and establish policies to migrate and store data on the appropriate storage tier and, ultimately, remove it altogether. ILM has evolved to include upfront initiatives like master data management and compliance.

Companies use ILM for: applying governance and policies to effective information management, there is enterprise data growth; growth in unstructured data; limitations in relational data base management system performance; information access and security concerns; lack of effective methods for classifying data; and difficulty in assessing productivity of systems, applications and databases. ILM allows increased control over data, regulatory compliance and reduced costs (by eliminating redundancies in data storage).

IT departments expect to nearly double the percentage of employees who have access to their enterprise content management system(s) in the next two years, and a significant percentage expect to increase their investments in capture, records management, and case management solutions (IDC 2013)

Market:

- According to Gartner ILM will grow between 5–7% in 2014
- Worldwide revenue for the content management market was €4 billion in 2013, up 8.7% over 2012[55].

Information Governance and Risk management (IG&R): Gartner defines **information governance** as the specification of decision rights and an accountability framework to ensure ***appropriate behaviour in the valuation, creation, storage, use, archiving and deletion of information***. It includes the processes, roles and policies, standards and metrics that ensure the effective and efficient use of information in enabling an organization to achieve its goals.

Companies use IG for:

- Information discovery, audit, risk assessment, and policy planning;
- Business process and IT process planning and design;
- Information retention implementation and deployment solutions, including testing, training, and organization change management;
- Legal hold and eDiscovery implementation and deployment solutions, including testing, training, and organization change management;
- Monitoring, audit, and reporting for adherence to corporate information retention, legal hold, and eDiscovery programs, as well as operational risks and security compliance programs for PCI-DSS, cybersecurity and data privacy, antifraud and antimoney laundering, and Basel II compliance
- Corporate processes involving information migration, management, and integration with existing corporate assets.

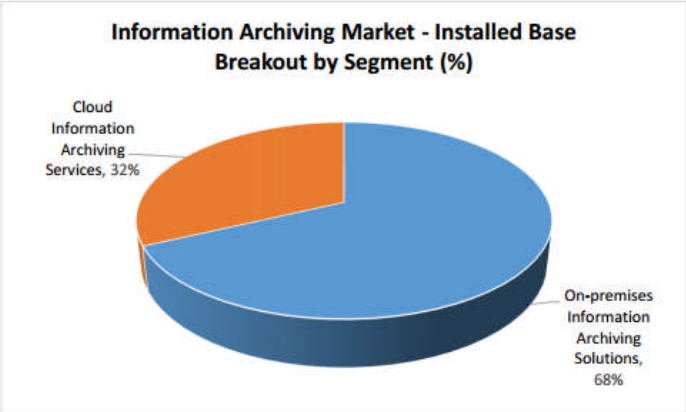
[55] COMPETITIVE ANALYSIS. Worldwide Content Management 2013 Vendor Shares. Melissa Webster June 2014, IDC #249014

- Corporate alliance partnerships (e.g. mergers and acquisitions), where comingled files, data, and messaging are shared, may benefit from a third-party service provider offering rapid information availability of information, with associated permissions, and continuity of content.

Market: mainly driven by growing pressure to improve governance strategies to meet the risks and regulatory/compliance challenges.

- The global market is expected to grow from € 3,5 billion in 2013 to €2,53 billion in 2018. Consulting services are expected to grow from € 1,6 billion in 2013 to €2,5 billion in 2018
- IDC has calculated that Governance, Risk and Compliance Infrastructure spending was €50,4 billion in 2013.

Archiving enables organizations to efficiently store, effectively manage, and easily discover and retrieve unstructured information as needed for business. The industry's most widely-deployed on-premise enterprise archiving solution



Information Archiving Market – Breakout by Segment, 2013

Figure 9: Information archiving market

Companies use archiving for :

- Compliance with strict regulatory requirements and compliance procedures. Growth in the government, education, and healthcare sectors. Also a growing number of businesses in less heavily regulated industries are beginning to deploy information archiving solutions for several different reasons.
- Provide support during litigation
- While not required by law to store and preserve electronic content, many businesses find great value in information archiving solutions
- Improve storage management and facilitate better backup and restore capabilities.

Market: driven in many cases for the need to re-evaluate their corporate storage policies due to content growth and source; Email compliance. Also proliferation of solutions such as, Microsoft SharePoint, Social media, and enterprise IM is also driving the need more archiving solutions.

The worldwide information archiving market, including both on-premises and cloud (i.e. hosted) solutions, is expected to be over €2,52billion in revenues by year-end 2013, and will grow to €5, 6 billion in 2017.

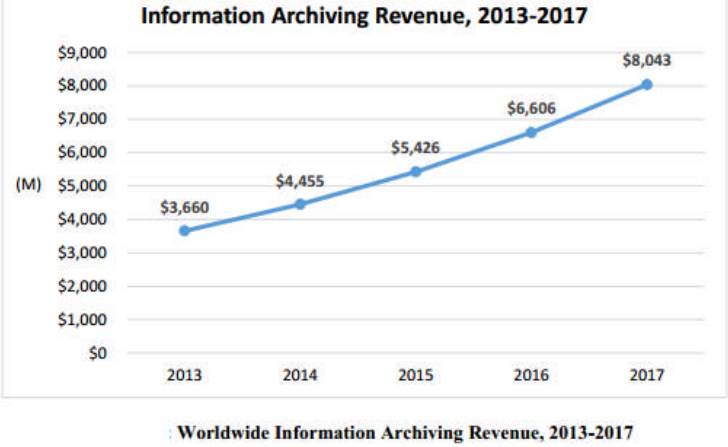


Figure 10: Information archiving revenue trends

Migration (services linked to DP): Data migration is the process of transferring data between storage types, formats, or computer system. Data migration phases (design, extraction, cleansing, load, verification). Migration is done due to: Storage migration, Database migration, Application migration or Business process migration.

Migrations are performed due to technical changes and business conditions which make companies to evaluate their server en storage infrastructure Drivers for this evaluation are:

- Leverage their IT investments for performance and efficiency
- Drive down hardware and software spending and management costs.

Main industry sectors for potential demand in the industrial sector are: Potential adopters of DP services of the VCoE in the business arena comprise companies from highly regulated industries with:

High economic value and/or growth of the sector
Sophisticated IT infrastructure
Sizeable international markets activities

www.project-consult.de/files/Iron%20Mountain%20Guide%202013%20European%20Retention%20Periods.pdf

DOCUMENT DESCRIPTION	COUNTRY														
MEDICAL/SAFETY RECORDS	AUT	BEL	DEU	DNK	ESP	FIN	FRA	HUN	IRL	ITA	NLD	POL	ROU	SWE	UK
Medical (occupational health & safety company doctor) files; medical documents in cases of a medical treatment contract	30 y	15-40 y	10 y	10 y	P+5 y R	S S	X	30 y	D+7 y 40 y*	P	15 y	20-50 y	D	-	S R
Floor plans and directions	-	-	-	-	R+5 y R	R	5 y	R	P	10 y*	R	2 y	50 y	-	S
Work-related medical examinations related to hazardous substances	-	30-40 y	40 y	40 y	P+5 y R	-	X	30 y	40 y - P	D	40 y	50 y	40 y	10 y	40 y
List of employees who have worked under dangerous conditions or whose health has otherwise been under threat	-	-	-	-	P+5 y R	-	P*	50 y*	S	D+ 10 y*	40 y	40 y	40 y	40 y	5-40 y
Register of employees who work with 3rd and 4th category biological agents	-	10-30 y	D D	10 y	10-40 y	40 y	10 y	-	10-40 y	D+ 10 y*	10 y	10 y	40 y	-	40 y
Lists/register of employees who have been exposed to asbestos dust	40 y	40 y	40 y	40 y	40 y	80 y	P*	50 y	40 y - P	40 y	40 y	40 y	40 y	40 y	40 y
Administration concerning measurements of radioactive substances	-	5 y	-	-	20 y	-	P*	5 y	S	D+ 10 y*	5 y	3-5 y	10 y	-	S
Records of radiation	-	30 y	30 y 100 y	-	30 y	5 y - R	P*	5 y	5 y - P	10 y*	5 y	30 y	10 y	10 y - P	S
Medical records of employees who have possibly been exposed to ionizing radiation	-	-	30 y 100 y	30 y	30 y	30 y	P*	50 y	S	30 y	30 y	30 y	40 y	30 y	5-40 y
Registration of work and rest periods (in appropriate format)	R	R	2 y	1 y R	D+4 y R	2 y R	D*	R	3-7 y* R	R	1 y 2 y	3 y D	R	3 y R	2-3 y R
Necessary data for emergency medical care, individual reintegration plans, individual treatment agreements, degree of incapacity for work, required workplace adaptations	R	-	R	10 y R	D+4 y R	R	D*	R	D+7 y - P R	R	R	20-50 y	R	R	D+6 y

LEGEND Black text: minimum retention period. Blue text: maximum retention period. *: Recommended retention period. X: Retention prohibited. D: Duration of the contract or permit. P: For the entire duration of the company and/or permanently. R: As long as required for the relevant purpose. S: Situation specific or too detailed for the scope of the summary and/or guide.

www.project-consult.de/files/Iron%20Mountain%20Guide%202013%20European%20Retention%20Periods.pdf

DOCUMENT DESCRIPTION	COUNTRY														
TAX AND ACCOUNTING RECORDS	AUT	BEL	DEU	DNK	ESP	FIN	FRA	HUN	IRL	ITA	NLD	POL	ROU	SWE	UK
General obligation of taxpayers to provide (upon request of the tax inspector) all information that may be relevant to their tax position, including all books, records and other data carriers	7 y	7 y	6-10 y	10 y*	10 y	6-10 y	6 y	7 y	6 y	4-10 y	7 y	5 y	5-10 y	7 y	1-6 y
A company is obliged to keep records of all delivery of goods or services, all intra-European Community acquisitions, all imports and exports, and all other information relevant for VAT purposes General obligation to keep at least the following records: (i) VAT invoices sent and received; (ii) documentation relating to supplies and acquisitions within the EU; (iii) documentation relating to goods imported from, and exported to, outside the EU	7 y	7 y	10 y	5 y	10 y	6 y	3 y*	7 y	6 y	10 y	7 y	5 y	10 y	7 y	6 y
Obligation to keep records and other data carriers relating to the ownership of real estate and rights relating thereto	22 y	-	-	10 y	10-13 y	13 y	6 y	-	D+6 y	10 y	9 y	-	10 y	-	-
Taxpayers are obliged to have available all information that deals with intra-group price setting, so that the Tax Authorities can check transfer prices and other conditions agreed upon in intra-group transactions	7 y	7 y	-	5 y	8-18 y	6-10 y	3 y*	7 y	6 y	4-5 y	7 y	5 y	5-10 y	7 y	-
Obligation to have an administration showing the payment of dividends, and the obligation to issue dividend notes	7 y	7-10 y	-	-	10 y	6-10 y	3 y*	-	S	10 y	-	-	10 y	-	-
Keep an administration as per the requirements set out in the General Customs Act, including books, records and other data carriers. General obligation of taxpayers to provide (upon request of the customs authority) all information which may be relevant to them, including making available all books, records and other data carriers.	3 y	-	-	-	3 y	S	3 y	7 y	3 y	4-5 y	7 y	5 y	10 y	5 y	3 y
Regional and municipal taxes	7 y*	S	S	-	10 y	S	1-3 y*	7 y	-	4-5 y	S	5 y	10 y	-	-

LEGEND Black text: minimum retention period. Blue text: maximum retention period. *: Recommended retention period. X: Retention prohibited. D: Duration of the contract or permit. P: For the entire duration of the company and/or permanently. R: As long as required for the relevant purpose. S: Situation specific or too detailed for the scope of the summary and/or guide.

Figure 11: Statutory retention times for certain types of official records

Cultural and Creative Sector

Sub-sectors	Cultural Industries/Media & Entertainment: Film & video, TV & Radio, Video Games, Music, Books & Press (Publishing) Creative Industries: Design (fashion, graphic, interior, product designs) Heritage: Museums, Libraries, Archives & Archaeological sites Other core arts: visual and performing arts
Compliance	20-50Years for music, prototypes and designs, +100 for long tail (e.g. film, cultural heritage) IPRs (copy rights and trademarks). Activities based on massive reproduction
Challenges	Transition to the digital content era Extract, combine and manage external and internal data Manage a complex cooperation / collaboration environment with new entrants and social media How to provide value-added services to potential customers (based on content) Maintain quality as a competitive advantage Create dynamic and interactive experiences based on existent content
Revenues	Cultural and creative sector turnover €541 billion 2003 (EU15) 3.5% of the GNP of the EU (2010) – cordis Broadcasting: €317billion 2014 to €405 billion in 2018 (+28%) Spending in ICT 12% Global broadcasting and cable TV industry expected to reach almost €332 billion in 2015 (MarketLine), 27% market growth in 5 years. TV advertising accounts for almost 48% of the overall market. Supports tourism: European arrivals reached 500 million. Tourism directly contributes, on average, 4.2% of GDP and 5.4% of employment (4.4% and 5.7% for EU members) in 2010.
Other	5.8 million of people worked in this industry 3.8 % of EU workforce (cordis) Creative industries are dominated by SMEs, with micro-SMEs and free-lancers representing 85% of all actors. SMEs co-exist with a few “global

	players", especially in publishing. The most important reason for publishers for DP: it will stimulate the advancement of science. Even most (96%) of small publishers DP is important. 84% of large and 55% of small publishers have a DP policy but 70% of them (large + small) do not have DP in place [56] Open to innovative ICT adoptions Shortages in IT budgets; do more with less

Energy and Utilities

Sub-sectors	
Compliance	Permanent retention – Energy database and pesticides database 3-20Y Copies of waste management 30Y Documents containing audits on radioactivity and results measurement 10Y data regarding chemicals or environmentally dangerous substances 10Y Metering database
Challenges	Regulation for utilities / energy industry continues to growth to improve industry security and reduce risks Market transparency and exchange of information regulation Customer and regulators are looking for ways to reduce energy costs Outage Prevention, Readiness, and Response is a major priority IoT (Internet of Things) business value is yet to be seen Capture the opportunity that represents HEM (Home Energy Management) before other competitors Sustainability is major concern for cities and citizens Every record has both content and metadata for indexing, searching, and formal auditability. For one company, there are typically 1,000 users and 6 million documents of various types and sizes to be managed on a daily basis regarding nuclear operations Increasing need of interconnecting countries and grids increase energy security and efficiency
Revenues	According to the World Nuclear Association, 435 nuclear reactors exist in 30 countries and generate 14% of the globe's electricity Globally, demand for electricity is set to continue to grow faster than for any other final form of energy. Demand expands by over 70% between 2010 and 2035, or 2.2% per year on average. In terms of electricity use, industry remains the largest end-use sector through 2035 The electricity sector's annual turnover of €420 billion represents more than 3% of European GDP (Electricity without Borders a plan to make the internal market work – BRUEGEL BLUEPRINT SERIES, 2013.)
Demand & Others	To maintain safety, security, and compliance at power plants, Management must have well-documented and highly visible information across the asset life cycle. Photovoltaic and wind energy capacity increasing

[56] PARSE.Insight D3.4 P5/83

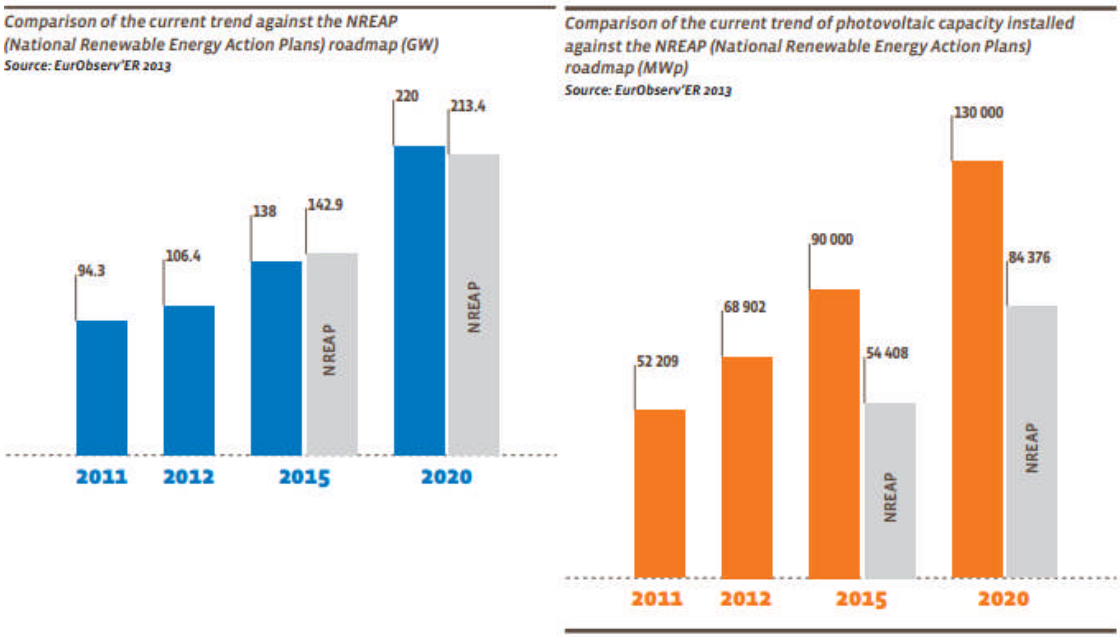


Figure 12: Trends in renewable energy capacity

Healthcare

Sub-sectors	Hospitals, medical technology and devices
Compliance	Highly regulated Patient lifetime Hospital safety records (i.e. incidents) 7-10 years X ray 30 years Ultrasound records (e.g. vascular, obstetric 20 years or 8 after death Post mortem Registers 30 years[57]
Challenges	Dealing with cost pressures Cope with more regulations Providing services under staff shortages Maintaining safety and quality of service New Business Models Drive at least 50% of Healthcare IT Growth Compliance will cost more than expected by players of the sector
Revenues	Public expenditure on healthcare in the EU 14% in 2030 UK £20 billion is exactly the extra money that the NHS will need every year by 2020 to meet patient demand
Other	IT spending: in western Europe were above 6% for 2014 Shortages in IT budgets; do more with less

[57] http://www.hse.ie/eng/services/list/3/hospitals/ulh/staff/resources/pppgs/rm/recret2013.pdf

Manufacturing

Sub-sectors	Automotive, Aerospace, Discrete manufacture (no automotive), chemical and process manufacturing, Food
Compliance	Highly regulated +50 years for design Automotive: +15 years for vehicles sold Aerospace: +50 Y Discrete manufacturing: 15-50 Y Chemical and process manufacturing: safety 20-50 years Food: 1-30Y (safety)
Challenges	Create globally integrated value chains (+ integration and discovery) Enablement of connected operations Move beyond automation, transform labour force into knowledge worker Decoupling of manufacturing and management functions in geographically dispersed manufacturing companies Move beyond automation, transform labour force into knowledge worker Increasing need of big data and analytics
Revenues	€271 billion -2001 EU manufacturing 7.4 % growth (2010) EU chemicals 10.2 % growth rate (2010) EU chemicals trade €41.7 billion (2011) Food & drink: €1,117 billion (2012, +6.8% than 2010) . 287,000 companies in Europe. Aerospace: €26 billion worldwide (2012) Automotive: €780 billion, with value added of over €140 billion (2011)
Other	ICT spending €541 billion Gross value added of ICT in Europe almost 120 billion (2010) Food: 4.25 million people employed in EU (stable to 2012) Chemical: 1,19 million people employed in EU (2011) Automotive: 2 million people directly employed (2011) Shortages in IT budgets; do more with less

2.3 FORMULATION OF THE USAGE BUSINESS CASE

2.3.1 Patterns in usage of digital (preserved or to be preserved) content offered by VCoE potential

This section focusses on cultural heritage/memory institutions. For such institutions the pattern is:

Demand profile: Libraries, Archives and Museums mainly that have implemented DP initiatives are of large size (e.g. National Libraries), with digital collections between 10 to50 TB and are driven by legal mandate, usually at national level; followed by the need of access and reusability; specific cases also mentioned: value for user, reputation, market penetration, need to comply to business change (from print to digital) etc.

The organizations that have no DP initiatives, 50% (mainly university libraries) have between 1-10 TB.

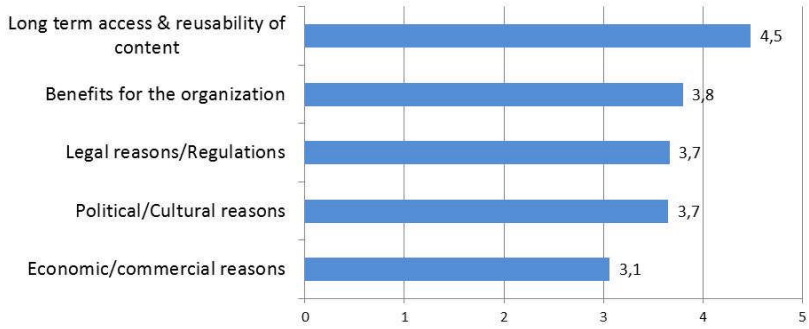


Figure 13: Motivations for digital preservation

Budget: budget constraints drives DP development, and looks for new initiatives (sharing and cloud). Organisations that have developed DP initiatives have done it mostly in-house (1/3 with budget). From this 1/3 of the universe, 97% of organisations have specific budgets, which amount over 50000€ for creating/acquiring/archiving digital content are often involved in digitisation and archiving activities and thus possess a specific budget for performing them. Nevertheless, the majority of organisations not performing DP have a budget below the 50000€ which primarily consists of university libraries located in Southern and Western Europe.

DP Staff: Scarcity of human resources is an issue that characterizes the preservation organizations; it also might reflect the scarcity of people in the market with the necessary skills. Around half of the survey respondents reported that their organisation has less than five (5) people involved in DP activities, 35% of them have between five (5) and ten (10) people, and only 16% of organisations have more than ten (10) people involved in DP.

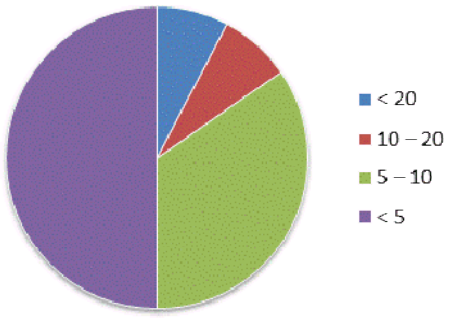


Figure 14: Number of DP staff per organisation

Benefits perceived: The functional nature of DP relies on the value of providing preservation managers with sufficient information to take appropriate actions to maintain the digital content's over the long-term and ensuring that the digital content can be rendered, interpreted and accessed in the future by optimizing lifecycle management and appraisal activities to overcome technological changes.

Main benefits of DP are:

- Increased use of content as a result of better availability and fundability
- Research results integrity
- Improved organisation and staff reputation (visibility, citations, recognition of being at the forefront of DP, etc.)

How DP adds value?: The way DP adds value to digital content management is affected by the extent to which operational and strategic processes are embedded in the organisation. The survey identifies that, overall and in organisations active in DP, the two most common ways that DP provides high value to digital content are through identification, interpretation and retrieval of digital objects (63%), followed by optimization of digital content lifecycle management (53%).

DP as a governance practice: Achieving economically-sustainable DP requires more than an investment in technical infrastructure for data storage, management, etc. Long term DP raises several organizational challenges, as some business processes across the whole organization are affected by DP. This issue, coupled with the increasing complexity of DP according to the particular organizational requirements, is generating a need of implication of the high level governance for a long-term proactive and sustainable DP management plan.

Buy-in: Decision makers are involved in DP as they are involved in managerial activities as supervisors (42%), executives (33%) or advisors (24%). However it is interesting that 43% of the respondents reported that decision makers are involved as financial funders, which also supports organisation's accountability and management buy-in.

Budgetary and funding resources: Internal institutional budgets are a financial source for 70% of participating libraries, thus DP is partially funded through general operating budgets. In this regard, significant differences are observed by regions.

Cooperation: cooperation is key to in moving towards economically-sustainable DP. DP is not only a costly process but also key in addressing the current challenge of establishing responsibilities, workflows, governance practices and infrastructures to systematically preserve exponentially growing digital content on an on-going basis.

In order to foster cooperation between organisations pertaining DP or aiming at doing so, a coordination mechanism is needed to address synergies, initiatives and consistency of common DP activities across the European DP supply and demand side. 53% of respondents envisage this coordination mechanism as a consortium of peer organisations while 28% of them visualize under the figure of a national agency.

Gap: although having DP initiatives, still a gap in the integration of DP in Governance.

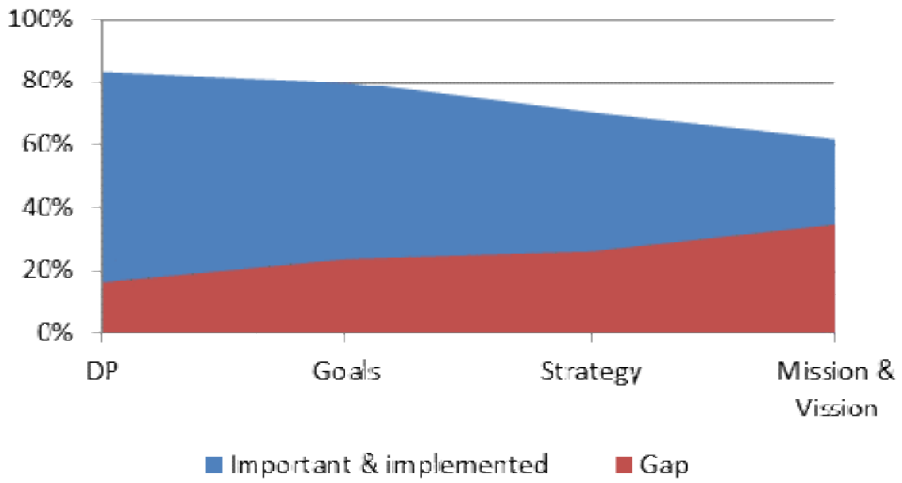


Figure 15: Gap in digital preservation aspirations

2.3.2 *Particular aspects of functioning of different market segments (cultural heritage and industry)*

Particular aspects of Cultural Heritage:

- Review the target market, map it and network: Differentiate by type and size of institutions (museum, academic library, SRI, etc.) and also investigate the potential amount of digital content, content growth, types and variety of digital formats to foresee the scope of preservation
- Understand how organizations defined digital preservation and how they managed these functions ahead: review mandates, if they participate in collaborative efforts and what do they think that is part of digital preservation
- DP is usually driven by:
    - A dedicated department/area with FET staffing; usually 1 in 3 Cultural organizations have FET for DP. If this is the case, many times there is one person who leads the activities, and the rest of the team also share other activities.
    - If it does not exist DP area or FET, DP activities are driven by
        - library, archive or other department that stewarded the collections
        - IT department
    - If the organisations are very small or DP initiatives are very incipient, then DP activities are led by managerial staff.

Thus decision makers are the audience and messages could be segmented as follows:

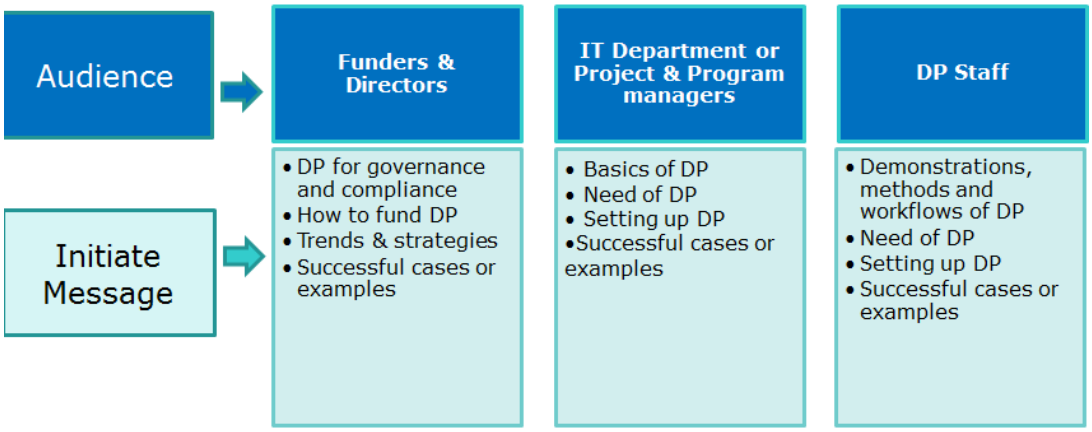


Figure 16: Segmentation of messages about DP

Directors and managers profiles to target:

- Library Operations Associate / Director / Manager
- Conservator
- Archives Librarian
- Librarian / Systems Librarian / Metadata Librarian / Preservation / Cataloguing Librarian
- Curator
- Archives Specialist
- Associate Archivist
- Policy & Program Analyst
- Head / manager of Special Collections

- Conservator and Preservation

More technical profiles to target:

- Technology Manager
- Manager of Digital Services
- Digital Content Initiative Coordinator / Manager
- Digital Imaging / Digital Preservation, Digital Collection / Digital Services Program Coordinator
- Head of Digital Collections
- Content Database Specialist
- Digitization & Metadata Coordinator
- Digital Preservation Librarian
- Government Records Archivist
- Electronic Records Archivist
- Senior/junior Records Analyst
- Metadata Librarian
- Technical Services Assistant
- Academic Technologist

Particular aspects in industry[58]:

- Review the target industry, map it and network: Differentiate by type and size of company and also "investigate" the potential amount of digital content, content growth, types and variety of digital formats to foresee the scope needed services in data management and compliance. DP services foresee long term preservation; which is not a key issue, generally, in industry.
- Understand how organizations see risk management on records or data: review mandates, and suppliers. Usually more than 80% of organisations (including SMEs) defend to have data/information risk management and data/information lifecycle practices, using retention rules, policies, and training or management practices.
- Industrial organisations claim that information lifecycle and retention/legal hold practices usually work, but not the deletion or destruction of information.
- Industrial organisations train people at least twice a year.
- Data management, Lifecycle management and compliance are usually driven by a dedicated for compliance: i.e. risk department or IT department; such as (by importance):
    - Compliance or regulatory department/affairs
    - Legal department
    - Risk management areas
    - Administrative services (i.e. health and safety or administrative compliance) or internal audits
    - IT department
    - Executive and high managerial team
    - Mid-level management

[58] Based on ARMA report. 2014 Governance benchmarking survey.

- From DP perspective, the IT areas managing ILM, IG and data management need to be approached, in this regard, profiles will comprise: directors, managers or responsible from the following areas will be related: Privacy, information security, legal holds and discovery, Business continuity, crisis management and recovery, data archiving or storage management, analytics including taxonomies and metadata, file plans and governance, application development.

2.3.3 Peculiarities of criteria, acquisition mechanisms and supplier selection

Cultural Heritage institutions need to comply with public procurement rules. Public procurement amounts to € 1000 billion per year of world trade and is a significant part of GDP (10-25% worldwide); in the EU, the public purchase of goods and services is 16% of GDP[59].

In this regard key aspects are:

- Make sure to comply with international standard procurement principles of EU; i.e. meet regulatory formalities to be qualified and registered to sell to the government.
- Know the business case: provide the long view and ensure compliance with their parameters: savings in the long run, improve efficiency, support mandate in better way, etc.
- Show that the tools comply with security or support localized emergencies
- In public sector usually a supplier from public sector or an organisation has advantages. Also universities provide support for training in Government. References from universities for training are good.
- Currently Government evaluates and support entrepreneurial organisations as suppliers
- For selling tools it is good to partner with current IT suppliers (e.g. hardware)
- Contacts, References –branding- and networking is important

Industry[60]: key aspects are:

- Know the business case: provide the long view and ensure compliance with their parameters: savings in short and long run, improve efficiency, support compliance in better/more secure way, etc.
- Show that the tools comply technological architecture
- Partner with other IT suppliers
- Current deployments with tangible results are important, especially in a similar or related segment. For example entertainment and culture industries are closely related to cultural heritage; publishing is related to broadcasting, etc.
- Contacts, References –branding- and networking is important

[59] http://ec.europa.eu/trade/policy/accessing-markets/public-procurement/

[60] Based in ARMA report. 2014 Governance benchmarking survey.

2.4 DETECTION OF UNSATISFIED NEEDS/COMMERCIAL OPPORTUNITIES.

In General for **Cultural Heritage** organisations needs are:

- DP becomes part of project proposals, thus DP could be attended to full extent (PARSE.Insight)
- Cohesive policies across departments to head towards real progress
- Increase the awareness of the fragility of digital content, and create consensus of what, when and how data have to be preserved among research institutes, thus collaboration could really happen.
- Available financial resources
- Better "cost data", not only to improve the complex cost models but to funders, and specially small and medium organisations, to provide systematic approaches and consistent data for DP operative procedures, and based on operative experience. This process includes the need of understanding the associated special requirements and variability of cost data[61].
- Dedicated staff or team for DP activities and adequate level of technical staff; thus it is required:
    - Training for local government officials, especially archivists in the Archives case. The most needed training is about advocacy, funding options, grant writing, and cost-benefit analysis. One potential strategy is to develop a certification program for local government archives, modelled after or integrated with the certified local government program used to manage the distribution of historic preservation funds[62]
    - Funding for training need to be increased; although in general about half of the organisations tend to have budget for professional development; larger organisations are the ones that tend to have funding for training however it tends to be perceived as limited, meanwhile small organisations have no funding for training[63].

For SRI:

- Publishers lack of a coherent near-term strategy for a return on investment in mobile (big deal the solution).
- There is a need of partnership and collaborations among publishers and data producers.
- SRI expect publishers to add value to core journal content, including active content, visualization an analytics and moving towards workflow tools and systems; which also favours aggregation.
- SRI need that publishers think terms of services rather than products, as well as greater understanding of user needs and behaviours.
- There is a need of semantic enrichment and linked data to make content smarter and improve discoverability.
- Define better, with publishers, the role of data as part of the research outputs and publishing of journals, including the citing, access integration of underlying data
- Research institutions need to develop its DP policies in an integrated way, currently they are developing DP policies separately as policy makers and funders.

[61] PrestoPrime. WP6. D6.3.1. analyses from P29/35

[62] Huth, Geof; Sustainable Funding for Local Government Archives. July 2007

[63] Digital Preservation Outreach and Education (DPOE). Trainning needs assessment survey.

- Publishers need tools, techniques and content, to embed these services in the workflows; similarly this could apply to the e-learning suppliers. At the same time, publishers need to handle compliance and exclusive rights of the content, in order to commercially offer or publish the content.
- There is a need to train researchers and to generate common procedures and standards for sharing data. Distrust towards digital archives generate that researchers keep their research data in personal computers at work (80%), portable storage carrier (66%) and organisational services (59%). The percentage of sharing is small, and usually does not take place through established digital archives (not even in the same discipline); major problems are need to have control over data, legal issues (41%), misuse of data (41%) and incompatible data (33%).

For Archives and Libraries:

- Training programme focusing on accessing e-content should be developed and offered to public library employees.
- Mainly 4 types of training and support (consulting) are needed by libraries: technical, strategic, basic DP knowledge and project management & planning.
- There is a need for hands-on experience and the how to IT. This includes case studies, successful experience, and how to transmit to the funders and managerial staff the importance of digital preservation (i.e. how to make the case)

For Industry:

- Tools and deployments that support storage growth management, as well as designing needs for storage, virtual storage, cloud services and disaster recovery solutions. The growth in data/information volume and value drives for better planning, governance and infrastructure for Big Data (included related analytics).
- Tools and services for more efficient content access, analytics and content management (especially disposal of files).
- Tools and services for strategic and big picture decisions as well as information management governance; still a lot of fragmentation in deployments
- Tools and services to support increasing need of security and risk based procedures, policies and governance. There are increasing concerns about risk emanate from different areas in the companies, including risk related with successful deployments of IT in the future, interoperability and standards.
- More cloud and storage professionals
- For industry there is a need for getting future-proofed products, with increasing willingness to take risk on forward looking market. This need is coupled with the need of convening and managing partnerships, alliances and more informal consortia; based on rich dialog and compelling solutions as few companies can develop a full suite of capabilities in-house[64].

[64] McKinsey on Smart Grids. Summer 2010.

Challenges Faced by IT and Storage Managers

Year-over-year Comparison

2013-14	2012-13	IT/Storage Managers and Professionals	
79%	77%	Managing Storage Growth	
43%	45%	Designing, Deploying and Managing Backup, Recovery, and Archive Solutions	
39%	36%	Making Informed Strategic/Big-picture Decisions	(+8%)
38%	39%	Designing, Deploying and Managing Disaster Recovery Solutions	
37%	31%	Designing, Deploying and Managing Storage in a Virtualized Server Environment	(+19%)
29%	27%	Lack of Skilled Storage Professionals	(+7%)
18%	16%	Designing, Deploying and Managing Storage in Cloud Computing Environments	(+13%)
15%	15%	Lack of skilled Cloud Technology Professionals	
11%	10%	Convincing Higher Management to Adopt Cloud	(+10%)
10%	7%	Infrastructure for Big Data Analytics	(+43%)
8%	4%	Managing External Cloud Service Providers	(+100%)

EMC²

Figure 17: Challenges faced by IT and storage managers

2.5 FORESEEABLE EVOLUTION, GENERAL FORECASTS ON DEMAND DEVELOPMENT AND MARKET SIZING

- Increasing demand on all sectors driven by PEST conditions: everything grows
- Internal demand of libraries
- Demand for digital content
- DP workforce in industry

	SRI	Cultural Heritage	Culture & Entertainment	Gas & Utilities Petroleum
Market Size	Research Universities : At least 24.743 SRI: at least 31,100	Libraries in EU: 65,000 -81.000. About 50 institutions rank as National Libraries Archives in EU: 10.000 to 15.000,	Books & publishers: +64000 Public service broadcasting media (PSM) in Europe with 73 large companieshttp://www3.ebu.ch/members in 56 countries	Gas & Petroleum companies: 82 (international Association of Oil & Gas producers (OGP)
Potential number of organisations that could be expected to join a membership organisation	Y1: 20 Y2: + 30 Y3: +50 Y4: +100	Y1: 10 Y2: +10 Y3: +10 Y4: +10	Y2: +10 Y3: +15	Y2: +1 Y3: +2
Tools and services	Tools, Consulting and training	Consulting and training	Tools, Consulting and training	Tools, Consulting and training

2.6 PRIORITISATION OF MARKET SEGMENTS TO BE ADDRESSED.

The recommendations arising from the above data and analyses are as follows.

- Develop new and strong clients in related markets:
  - "e-Science" (builds over brand and reputation on the heritage and e-science sector): this approach comes as the right segment to consolidate strong pilots with large projects due to:
    - The collaborative, computationally intensive and has the ability to work with massive volumes and data from different sources and diverse subject domains. New "computational laboratories" have been and are going to be enabled, performing new science by working on existing data
    - e-Infrastructures are at 25% of 2020 target. Thus the area will receive high budget to complete the roadmap in Europe.
- Expand on industry (builds on brand and reputation from e-science): expand on Industry and Entertainment, reinforcing the heritage sector.
- Expand cross-industry (builds on brand and reputation from e-science): industry segments related to the e-Science fields: thus probably Energy and Utilities, Oil and Gas and Healthcare are good industry segments where to start; as probable are the next step in the value chain using e-Science data.

3 VCOE POTENTIAL COMPETITORS

The offerings which the VCoE aims to provide are in the areas of:

- services
- tools
- training
- consultancy

In the sections below we present a first identification and analysis of potential competitors in each of these areas in turn.

3.1 SERVICES AND TOOLS

Identification and classification of the main and/or most dynamic DP related services suppliers: Direct/indirect competitors. Profile and identification. Main products/services

Non-commercial: non-commercial competitors comprise DP solutions, tools and services developed under holistic digital repository oriented software solutions that provide digital collection management and DP tools in open source. This offering is represented by MI and SRI, which having implemented DP solutions for their own needs for DP solutions have started to provide such solutions to other organisations, usually under non-profit conditions. Somehow this offering represents direct competition, as well as opportunities for the VCoE.

Offering Profile	Characterisation & Examples
DP infrastructure, software portfolios, frameworks and systems	Infrastructure, frameworks, systems and tools that provide support services, tools and storage/cloud-storage that enable digital preservation, data access, transformation, and data sharing. Examples are: • DuraSpace.org provides a portfolio of technology for the durability of digital content. Portfolio comprises: DSpace (http://DSpace.org/) is a turnkey institutional repository application, Fedora (http://Fedora-commons.org/) is a framework for building digital repositories, and DuraCloud (http://DuraCloud.org), an open source platform and managed service that provides on-demand storage and services for digital content in the cloud (DuraCloud is a pay as you go service as re-sells storage). • ePrints - University of Southampton is free and open-source software package for building open access repositories; works as a document management system mainly for IRs and scientific journals.
DP enabled Digital Library suites	This includes applications that manage automation of the full cycle of library processes and resources, including embedded DP tools which enable DP services. Fully developed examples are: Invenio – CERN Document Server Software Consortium (Invenio SW org). It was co-developed by CERN, DESY, EPFL, FNAL, SLAC, and originally developed to manage the CERN document service; it has a GLP license, and it has been deployed in about 30 SRI. Invenio is a free software suite that covers all aspects of digital library management from document ingestion through classification, indexing, and curation to dissemination. It is able to manage repositories from moderate to large sizes (several millions

Offering Profile	Characterisation & Examples
	of records). Invenio currently enables to run the new CERN repository Zenodo (OpenAireplus) which facilitates sharing of research outputs from a wide variety of formats across all fields of science. • National Archives of Australia Digital Preservation Software Platform (DPSP) is free and open source collection of software applications which support digital preservation of digital records. DPSP comprises 4 plugin modules: Xena that converts digital files to standards based, open formats; Digital Preservation Recorder (DPR) to handle bulk preservation of digital files via an automated workflow; the Checksum to monitor the contents of a digital archive for data loss or corruption and the Manifest Maker - Manifest Maker generate a tab-separated list of digital files in a specified location. Greenstone Digital Library Software- New Zealand Digital Library: it allows building, distributing and preserving digital library collections, with effective full-text searching, metadata-based browsing and making it available over the Internet or on CD-ROM. It has established a cooperative effort with UNESCO and deployed in more than 80 cultural heritage organisations and archives, such as in the New York Botanical Garden or iArchives of the British Library Newspapers or the Californian and Dublin Telegraph.
DP supporting tools and micro-services	Tools that support and enable services for DP. Usually comprise software developments that functions across the spectrum of DP processes and in/for different preservation environments: pre-ingestion, transformation or registries of formats, format validation, fidelity management, metadata creation and remediation, toolboxes for creating repositories, validation of integrity of files; etc. Increasing availability of tools also has brought to new flexible approaches that combine specialized DP solutions, depending on the organisation, its budget or challenges. There are at least 191 different tools available (Digital Preservation Services: State of the Art Analysis, presents and analyse those tools); amongst the most popular ones are DROID (Digital Record Object Identification) - Digital Preservation Department of the UK National Archives, and JHOVE (JSTOR/Harvard Object Validation Environment) - JSTOR and the Harvard University Library[65]. Thus combination of different tools can cover DP lifecycle; as affordable alternatives you can use Archivematica for ingesting and processing and DuraCloud for access, storage and maintenance of data[66].

[65] Ruusalepp, Raivo; Dobreva, Milena. Digital Preservation Services: State of the Art Analysis. http://www.dc-net.org/getFile.php?id=467

[66] Schumacher, Jaime; Thomas, Lynne; VandeCreek, Drew. From Theory to Action:"Good Enough" Digital Preservation Solutions for under-resourced Cultural Heritage Institutions. Institute of Museum and Library services. POWRR project. August 2014.

Offering Profile	Characterisation & Examples
DP services	Comprises on-demand DP hosted services offered through the outsourcing model; mainly it has evolved in the e-journal sphere and extended to the e-book content. These services are mainly offered by non-profit organisations funded via combination of annual memberships from libraries and publishers, and endowments from various funders. Examples of these organisations are: LOCKSS, and private/public LOCKSS networks, Portico, Jstor, CLOCKSS, e-Depot (KB), The Keeper Registry (EDINA). Other DP services organisations participate in Keepers Registry (e.g. Portico, LOCKSS, e-Depot)
DP standardisation and Reference Model valuation efforts	It comprises the efforts of articulating the requirement for long-term or permanent preservation of digital content in a responsible manner. OAIS addresses this issue, but as a reference model without a checklist for conducting long-term preservation processes. Standardisation and validation efforts aim to create consensus practices for "trusted" environments (trusted repositories and organisations). The reference model (ISO 14721) currently also provides a common conceptual framework describing the environment, functional components, and information objects within a system responsible for the long-term preservation of digital materials Also many metadata standards which are addressed by different communities depending on their needs, as metadata is the backbone of digital curation; for example the PREMIS: Data Dictionary for Metadata Preservation which is restricted to the digital preservation activities of: maintaining viability, renderability, understandability, authenticity and identity[67].

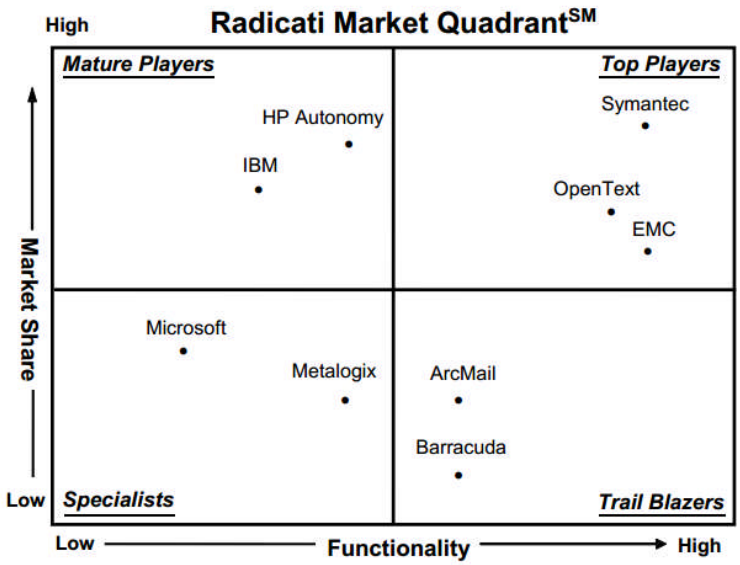
Commercial competitors: these competitors comprise those that directly target the DP as a niche market, and those indirect competitors that have embedded preservation modules or tools into content, information life-cycle or storage management solutions to enable preservation of content for at least a decade.

Commercial Direct Competitors	
DP and Library automation enabling DP products	They focus on Digital Preservation, Records retention, Archival solutions or Library automation solutions. Among competitors are: **Exlibris**: provides a modular library automation solutions that can be integrated as a suite for the discovery, management, and distribution of all materials—print, electronic, and digital. The offering includes: • Primo Discovery and Delivery • Alma library management services • bX Usage-Based Services (analytics and data mining) • Rosetta preservation of digital assets • Aleph, Voyager integrated library systems

[67] http://www.dcc.ac.uk/resources/briefing-papers/standards-watch-papers/what-are-metadata-standards#sthash.1nHhgsIO.dpuf

	• SFX scholarly linking • MetaLib metasearching • Verde e-resource management • DigiTool digital asset management **Preservica / Tesella:** Tesella Group has repositioned Tesella to target internationally the analytics, software and consulting services market, and **created Preservica**, a separate wholly owned subsidiary company. Preservica solutions range from cloud based Software as a Service "Preservica", to custom on premise installations of Tessella Safety Deposit Box (SDB), and positions itself as the "world-leader in digital preservation", ensuring digital assets to "be secure, accessible and readable for years to come – freeing you to focus on what you do best – safeguarding and sharing your valuable digital content, collections and electronic records." **Propylon**: Also companies as Propylon position in niche markets targeting financial or legal companies offering digital preservation. Propylon is a U.S.A based company that offers software to financial firms, legislatures, governments, legal teams and compliance teams worldwide for the creation, management, publishing, archival and research of content. Propylon software offers to "intelligently manage content using unique point-in-time technology, preserving a fully accessible audit trail with reports that demonstrate changes over time."
Commercial Indirect Competitors	
Archiving Solutions	Usually is a market taken by large and storage oriented vendors. According to Gartner, the "Archiving as a service (aka cloud archiving)" surpassed on-premises archiving. Gartner sees 60% to 70% that new or replacement email archiving implementations will be cloud-based. Archiving vendors claim to offer the following benefits through their solutions: • Reduce costs, improve operational efficiency, and enable compliance • *Help to* efficiently store, manage, and discover the organization's information • Reduce data footprint with a centralized, integrated data • Provide granular visibility into data stored enterprise-wide, ensuring storage utilization and compliance • Align the cost of archiving large volumes of data with the value associated with that data • Identify, manage, search, and preserve data across all formats and repositories in a defensible manner within a consolidated archive. Main players in Archiving market are: shown in the following quadrants: By Radical:

<table>
<tr>
<td></td>
<td>
MARKET QUADRANT – ON-PREMISE INFORMATION ARCHIVING SOLUTIONS



On-Premise Information Archiving Solutions Market Quadrant, 2013

By Gartner:

</td>
</tr>
<tr>
<td>Information Lifecycle Management (ILM)
Data Lifecycle Management (DLM)</td>
<td>ILM focus on managing data and storage as it recognizes that the value of information changes over time and that it must be managed accordingly. ILM seeks to classify data according to its business value and establish policies to migrate and store data on the appropriate storage tier and, ultimately, remove it altogether. ILM has evolved to include upfront initiatives like master data management and compliance (Gartner).
Data lifecycle management is the process of managing business</td>
</tr>
</table>

information throughout its lifecycle, from requirements through retirement (IBM).

ILM characterizes for managing data and information over its lifetime and through different application systems, databases and storage media.

The emergence of big data, with growing amounts of data, is increasing the need for ILM or DLM

Main vendors in this IT market are positioned in the following Gartner quadrant:

Main vendors in this IT market also have most of the revenues from the content management segment:

Worldwide Content Management Software Revenue by Vendor, 2011-2013 ($M)

	2011	2012	2013	2013 Share (%)	2012–2013 Growth (%)
IBM	765.6	864.4	924.8	15.9	7.0
OpenText	525.0	500.3	509.9	8.8	1.9
Microsoft	354.3	389.7	428.2	7.4	9.9
EMC	422.8	414.6	415.5	7.2	0.2
Oracle	279.9	281.0	274.5	4.7	-2.3
HP	276.9	248.1	257.5	4.4	3.8
Hyland Software	185.6	198.4	230.5	4.0	16.2
Adobe	90.6	154.8	209.9	3.6	35.6
Dropbox[1]	20.8	70.6	157.5	2.7	123.1
Lexmark[2]	128.6	143.6	157.1	2.7	9.4
Subtotal	3,050.1	3,265.5	3,565.4	61.4	9.2
Other	1,981.7	2,077.8	2,243.2	38.6	8.0
Total	5,031.8	5,343.3	5,808.6	100.0	8.7

Information Governance / Information	This area is defined similarly by different bodies: • It comprises multi-disciplinary structures, policies, procedures, processes and controls implemented to manage information at an

Governance Risk / Governance, Risk and Compliance Infrastructure	enterprise level, supporting an organization's immediate and future regulatory, legal, risk, environmental and operational requirements (Wikipedia) • It is a software platform organizations can use to identify information-related compliance across the enterprise, drive improvements, and develop metrics for measuring information governance (IG) program maturity (ARMA). • Provides a cross-disciplinary IT solutions view of the enabling technologies and services that allow companies to address the following objectives: information integrity and confidentiality, process integrity and application availability, information retention and disposition, and enterprise risk management (IDC) For IG, the visibility, actionable intelligence, and automation of structured and unstructured data is critical to manage the explosion of content, thus in this area has become important to leverage the metadata framework technology (IDC[68]) Main players in this area are: RDS, EMC, SAP, IBM, Software AG, Metric Stream, SAS enterprise GRC, Acaveo, Ostia, Reltio
Services	
On-demand hosted services	They comprise provision of secure and distributed storage or cloud services environments, some of them with preservation capabilities. For example, • Amazon Simple Storage Service (S3) • Amazon Glacier: online storage/archiving solution that starts; cheaper than S3 • EVault Long-Term Storage Service (EVault LTS2) with elastic capacity and "pay-as-you-go" pricing to scale your storage usage up or down as needed (by Segate) • Others: Microsoft Azure, CloudSigma, GreenQloud Also we can include vendors that offer DP software with storage services, such as: DuraSpace/DuraCloud, Preservica or Archivematica
DP related supportive services	They comprise services such as consulting, documentation, training, integration, cataloguing, metadata and planning of storage services, which play a key role in DP deployment. These players mainly sell consulting project management, data planning (e.g. data disposal) and customisation services based on different DP architectures (e.g. DuraCloud etc.). Among players in this area we have: PwC's Advisory Risk Consulting, Deloitte, IBM, Accenture, KPMG, McKinsey, Booz¬C As for training a number of Universities such as the University of London Computer Centre has developed a training program The Digital Preservation Training Programme (DPTP), which targets people working in institutional information management involved with issues of digital preservation. It combines technological and organizational perspectives

[68] https://varonis-assets.s3.amazonaws.com/pdfs/IDC_1014.pdf

Positioning: What messages are competitors sending vs. markets addressed

Non-Commercial suppliers:

DuraSpace:

DuraSpace (http://duraspace.org) organization is an independent not-for-profit providing ***leadership and innovation for open technologies that promote durable, persistent access and discovery of digital data***. Its values are expressed in its organizational by-line, "***Committed to our digital future."***

DuraSpace going towards sharing information

DuraSpace works collaboratively with organizations that use VIVO to advance the design, development and sustainability of the project.

VIVO Project

VIVO is an open source! semantic web! application for integrating and sharing information about researchers and their activities and outputs at a single institution, while supporting discovery of related work and expertise across a distributed network of linked data profiles. VIVO is fundamentally interdisciplinary in nature and enables collaboration across traditional boundaries of geography, organizational structure, and type.

Vivo: connect share discover. VIVO for integrating and sharing information about researchers and institutions to support collaboration and discovery

e-Prints Services:

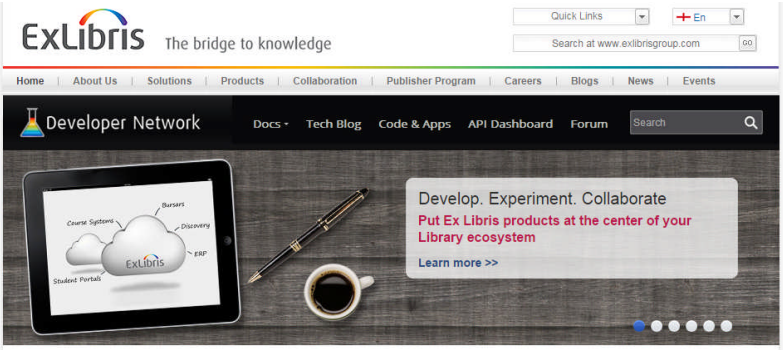
ePrints will not compete from software deployment point of view, but from services perspective. "The EPrints team is committed **to working closely with clients to develop tailor-made repositories that fulfil their exact requirements**, and we are proud to be supporting EPrints installations throughout the world."

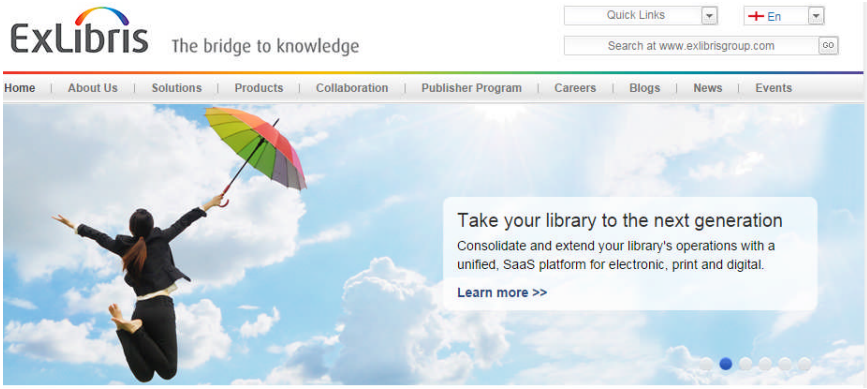
Greenstone

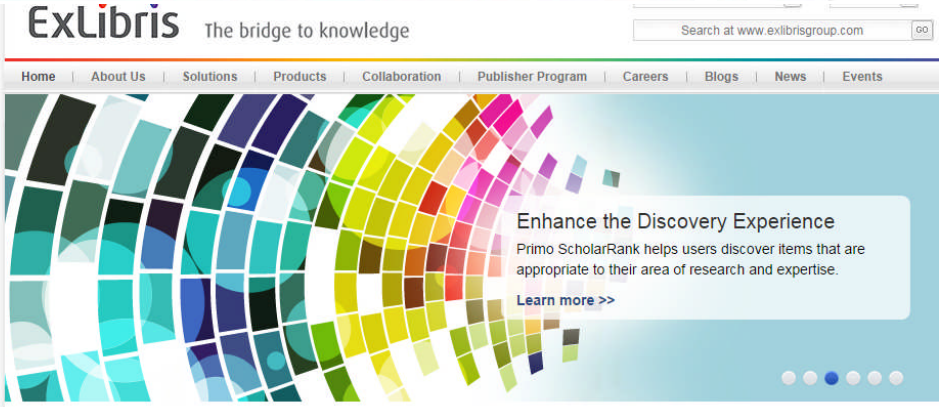
"The aim of the Greenstone software is to empower users, particularly in universities, libraries, and other public service institutions, to build their own digital libraries"

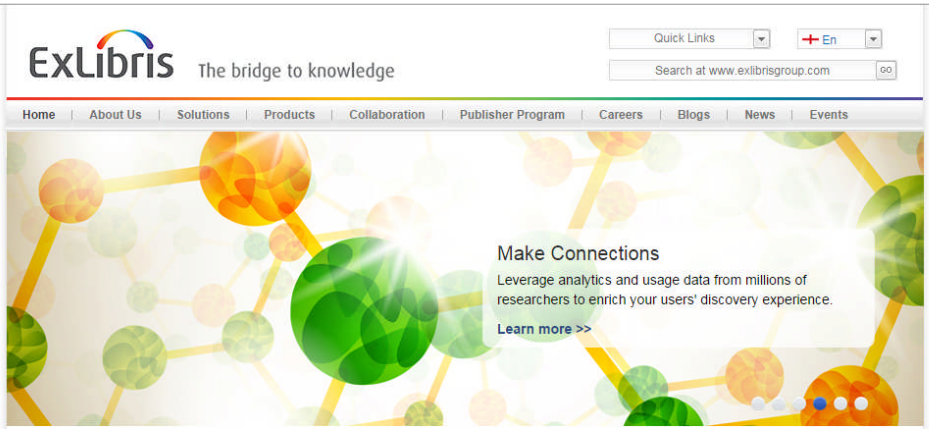
Commercial suppliers:

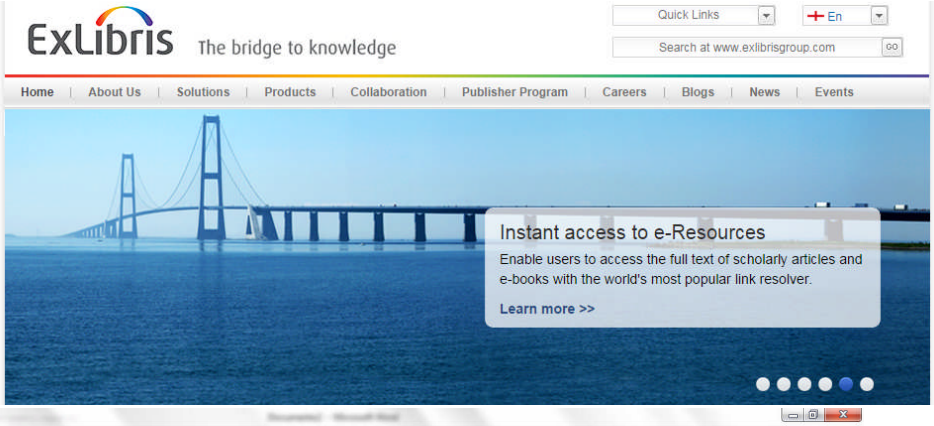
Exlibris: **"**under the Motto "Bridge to knowledge", they offer Discovery, Management and Preservation of information for Libraries. They build their messages based on: easiness and flexibility, search capabilities, advance and be on the edge, analytic and re-use of data, access to the information anytime, as well as end-to-end services for preservation.

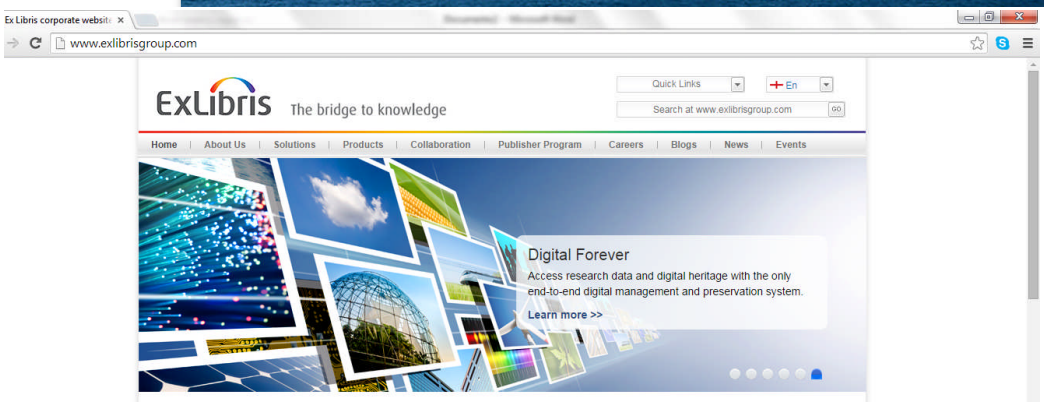


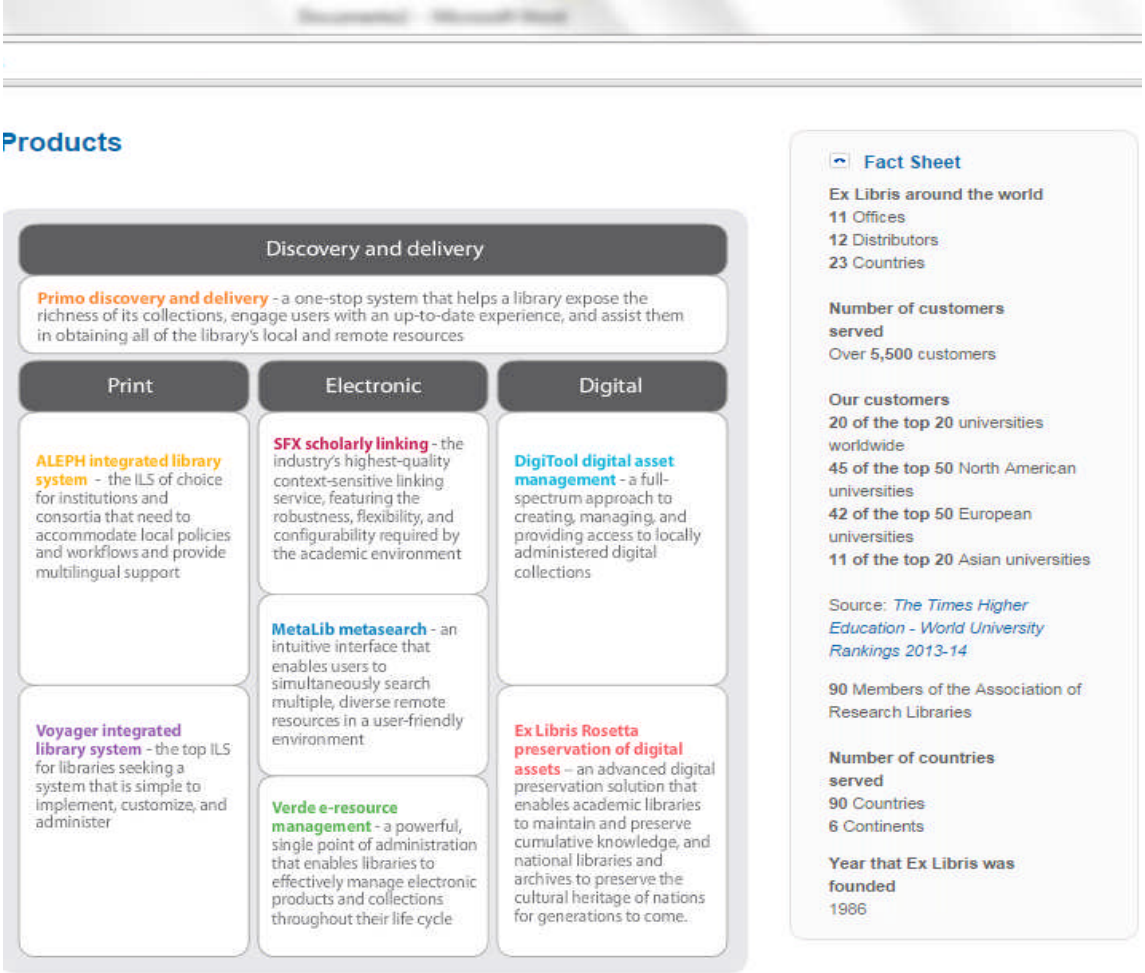







ExLibris
The bridge to knowledge
Quick Links
En
Search at www.exlibrisgroup.com
Home | About Us | Solutions | Products | Collaboration | Publisher Program | Careers | Blogs | News | Events
Instant access to e-Resources
Enable users to access the full text of scholarly articles and e-books with the world's most popular link resolver.
Learn more >>


www.exlibrisgroup.com
ExLibris
The bridge to knowledge
Quick Links
En
Search at www.exlibrisgroup.com
Home | About Us | Solutions | Products | Collaboration | Publisher Program | Careers | Blogs | News | Events
Digital Forever
Access research data and digital heritage with the only end-to-end digital management and preservation system.
Learn more >>



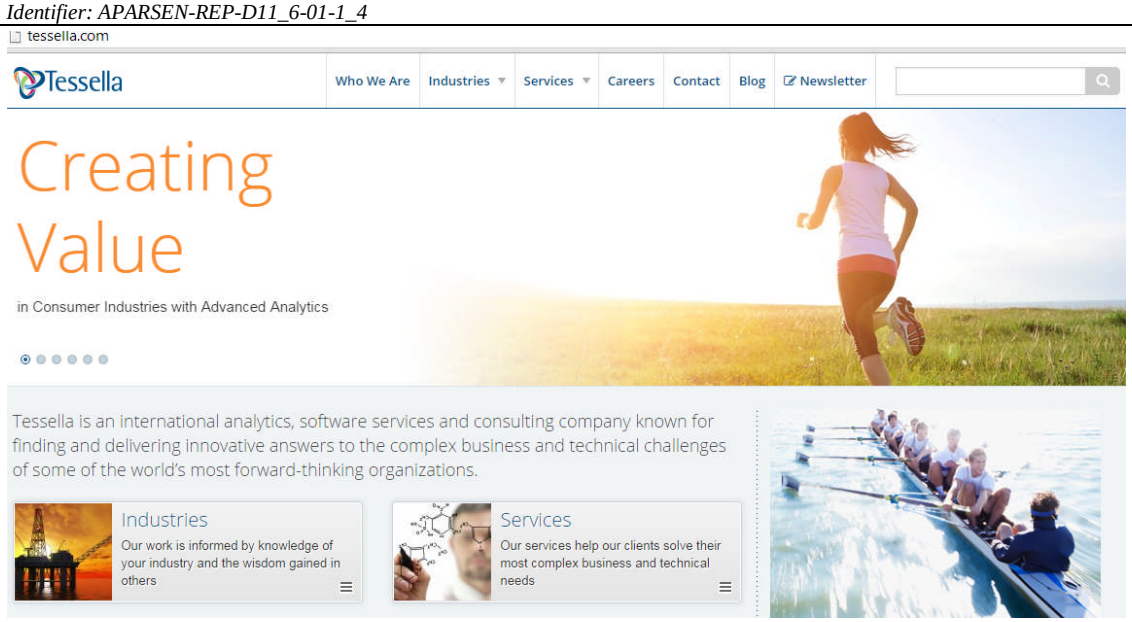
Messages from ILM and IG are equivalent to functional concept and benefits of DP

IBM provides information lifecycle governance solutions for: *Defensible Disposal*, *eDiscovery*, *Legacy Data Cleanup*, *Records and Retention*, *and Value-based Archiving*, helping customers manage enterprise information according to its business value while driving down cost and risk.

Tessella

Positions on the advanced analytics: "Tessella is an international analytics, software services and consulting company known for finding and delivering innovative answers to the complex business and technical challenges of some of the world's most forward-thinking organizations".

But maintains the link to preservation and an integration approach.



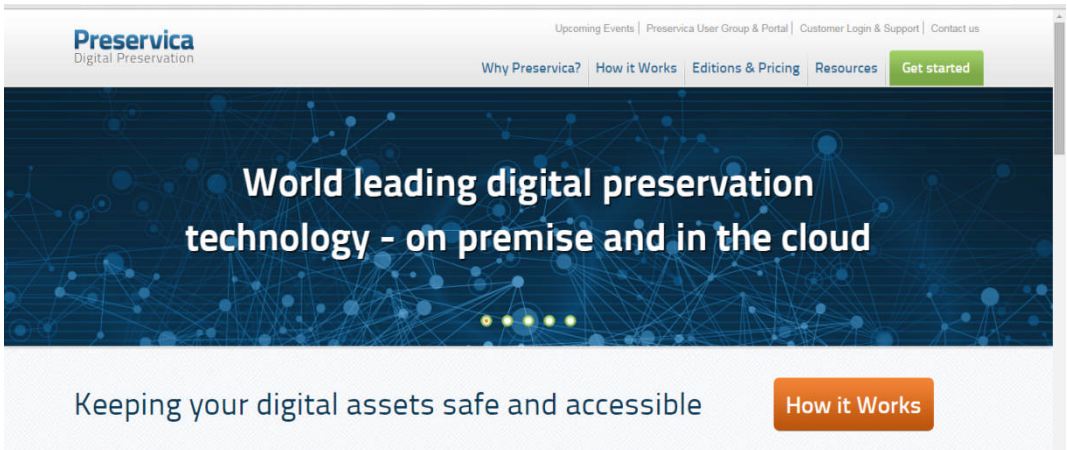
Preservica: Tesella wholly owned subsidiary

Positions on all sectors; for types of organization and sizes; to be used as affordable and complex solutions, and not as a costly solution. Preservica uses business language "solution for decades". Umbrella message underpins on Safeguarding and Safekeeping as well as availability of information

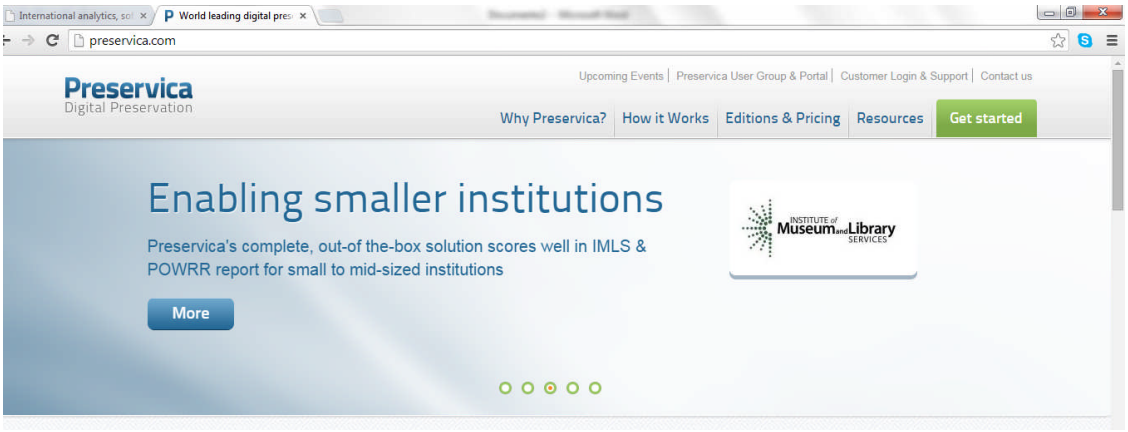
Builds on through storage solutions: from on premises to cloud based services.

Builds on Tesella Experience and standardization:

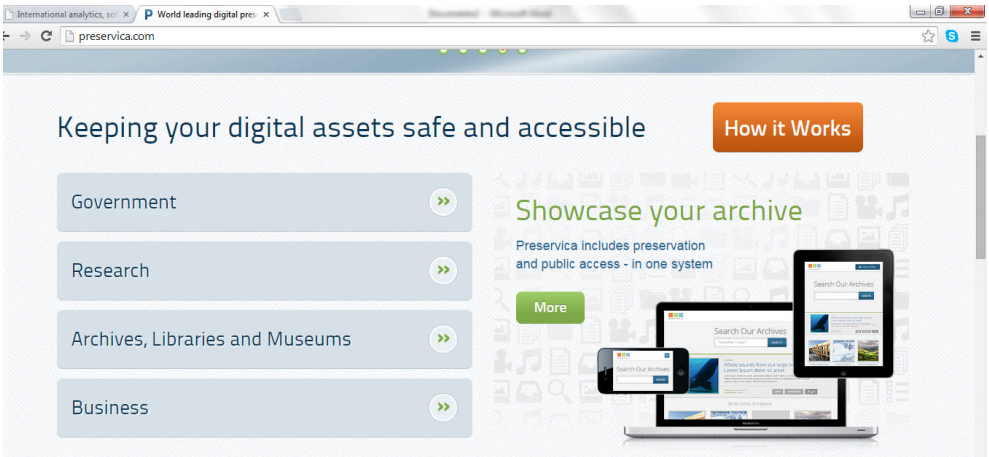
"Preservica have been instrumental in developing and contributing to many important standards including PRONOM, DROID and OAIS"

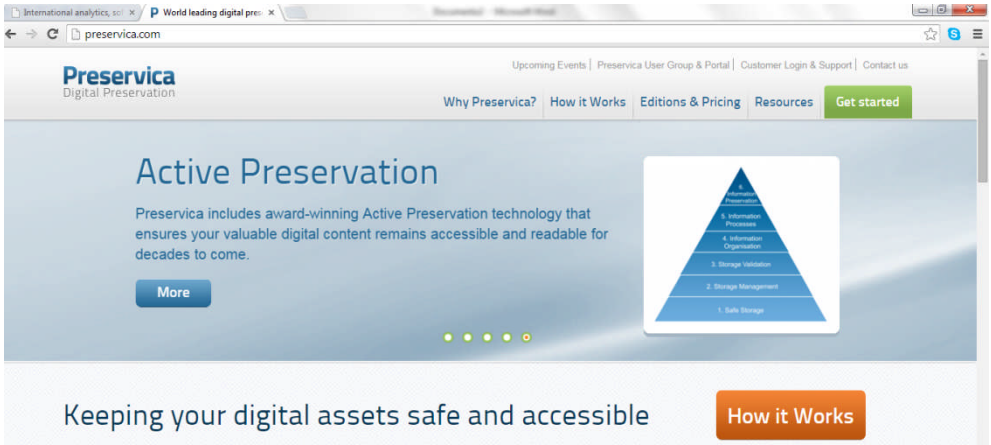


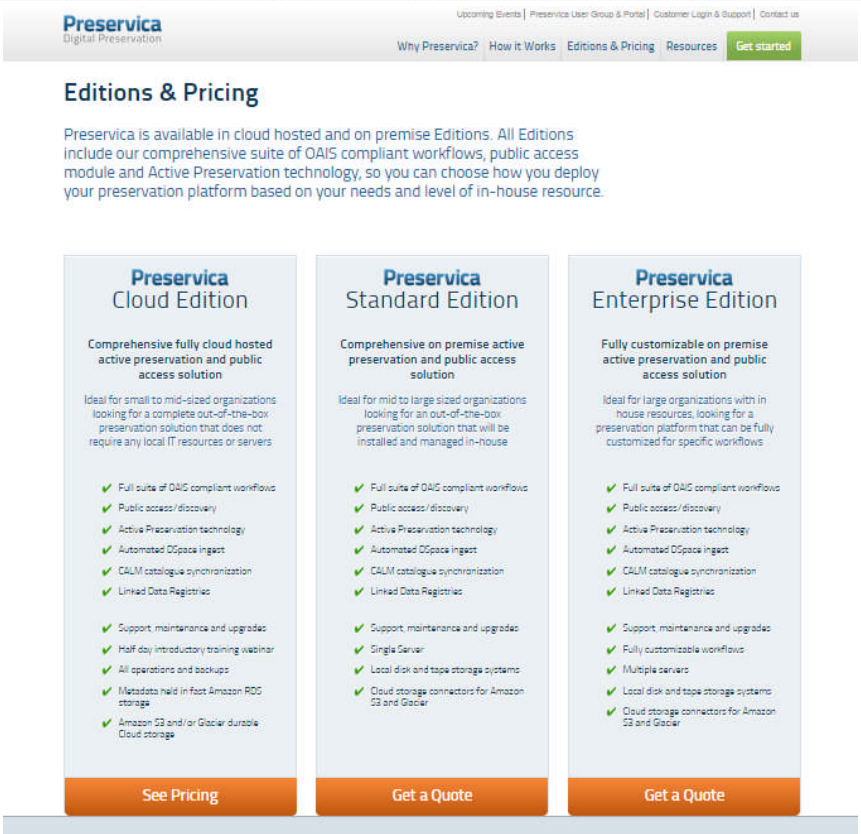
preservica.com
Preservica
Digital Preservation
Upcoming Events | Preservica User Group & Portal | Customer Login & Support | Contact us
Why Preservica? How it Works Editions & Pricing Resources Get started
Protecting State records
The Kentucky Department for Libraries and Archives is one of the largest state archives collections in the U.S. and has been a pioneer in archiving electronic records since the early 80s
More

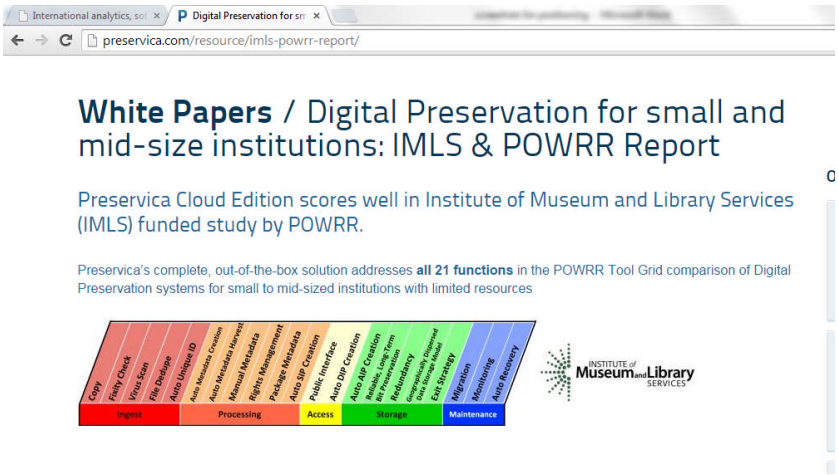

preservica.com
Preservica
Digital Preservation
Upcoming Events | Preservica User Group & Portal | Customer Login & Support | Contact us
Why Preservica? How it Works Editions & Pricing Resources Get started
Enabling smaller institutions
Preservica's complete, out-of the-box solution scores well in IMLS & POWRR report for small to mid-sized institutions
More
Institute of Museum and Library Services

preservica.com
Preservica
Digital Preservation
Upcoming Events | Preservica User Group & Portal | Customer Login & Support | Contact us
Why Preservica? How it Works Editions & Pricing Resources Get started
Bring your archive to life
Preservica now includes preservation and public access - in one system
More


preservica.com
Keeping your digital assets safe and accessible
How it Works
Government
Research
Archives, Libraries and Museums
Business
Showcase your archive
Preservica includes preservation and public access - in one system
More


preservica.com
Preservica
Digital Preservation
Why Preservica? How it Works Editions & Pricing Resources Get started
Active Preservation
Preservica includes award-winning Active Preservation technology that ensures your valuable digital content remains accessible and readable for decades to come.
More
Keeping your digital assets safe and accessible
How it Works


Preservica
Digital Preservation
Why Preservica? How it Works Editions & Pricing Resources Get started
Editions & Pricing
Preservica is available in cloud hosted and on premise Editions. All Editions include our comprehensive suite of OAIS compliant workflows, public access module and Active Preservation technology, so you can choose how you deploy your preservation platform based on your needs and level of in-house resource.
Preservica Cloud Edition
Comprehensive fully cloud hosted active preservation and public access solution
Ideal for small to mid-sized organizations looking for a complete out-of-the-box preservation solution that does not require any local IT resources or servers
Full suite of OAIS compliant workflows
Public access/discovery
Active Preservation technology
Automated DSpace ingest
CALM catalogue synchronization
Linked Data Registries
Support, maintenance and upgrades
Half day introductory training webinar
All operations and backups
Metadata held in fast Amazon RDS storage
Amazon S3 and/or Glacier durable Cloud storage
See Pricing
Preservica Standard Edition
Comprehensive on premise active preservation and public access solution
Ideal for mid to large sized organizations looking for an out-of-the-box preservation solution that will be installed and managed in-house
Full suite of OAIS compliant workflows
Public access/discovery
Active Preservation technology
Automated DSpace ingest
CALM catalogue synchronization
Linked Data Registries
Support, maintenance and upgrades
Single Server
Local disk and tape storage systems
Cloud storage connectors for Amazon S3 and Glacier
Get a Quote
Preservica Enterprise Edition
Fully customizable on premise active preservation and public access solution
Ideal for large organizations with in house resources, looking for a preservation platform that can be fully customized for specific workflows
Full suite of OAIS compliant workflows
Public access/discovery
Active Preservation technology
Automated DSpace ingest
CALM catalogue synchronization
Linked Data Registries
Support, maintenance and upgrades
Fully customizable workflows
Multiple servers
Local disk and tape storage systems
Cloud storage connectors for Amazon S3 and Glacier
Get a Quote



Overall evolution

Non-commercial DP suppliers have evolved in recent years towards a more productized offering:

- In the case of the DP enabled library suites (e.g. Invenio) and DP services (outsourced; e.g. eDepot), we can see that they have gone from a standalone in-house customized solutions, to work collaboratively and with other institutions as co-developers and extend the services to similar institutions or organisations. For DP services, also the integration with libraries is generating stronger partnerships which are generating a change in their services. However, still not efficient and generally lacks of underlying business model.
- In the case of frameworks such as DuraSpace, vertical integration (DSpace + Fedora + Storage +Vivo) is producing a robust framework that covers the whole infrastructure spectrum, tools and moving towards added value services for data/information re-use and sharing.
- In the case of DP tools, generation of increasing number of supporting DP tools have created a large enough (more than 190 tools) selection that could be integrated in the OASIS workflow. This trend is allowing a new wave of micro-services for DP, as well as could increase the potential of embedding DP tools in other solutions (i.e. Financial risk solutions).
- In the case of commercial suppliers, we can see that they have been strengthening their positioning in the market, either growing in the market segment (e.g. ExLibris in the library segment) or cross-sectors (e.g. Tesella). Evolution of Tesella towards Preservica also shows the potential of the industrial market as well as the potential of Cloud-based services.
- Regarding the large IT players, we can see that they are evolving towards effective information access and analysis for value added services. Increasing cloud based services is also freed demand organisations from vendor lock-in (specially for storage services), thus large players have increased their product portfolio towards analytics and governance. ILM and Governance are increasing in importance, due to the exponential growth of data volume and value.

3.2 TRAINING

Results collected in D42.1 have come to the following conclusions, based on the topics of training currently available, 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.

Moreover, 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.

Furthermore, as also noted in APARSEN Deliverable D43.1, which addressed more formal curricula about digital preservation, there is a lack of common terminology.

Perhaps more importantly there is a lack of consistency of the various approaches.

The target markets seem to be largely the preservers of rendered objects.

The VCoE training is based on the Common Vision, a glossary of digital preservation and an overall consistent and coherent plan. A description of the training modules is provided in a spreadsheet[69]. The modules are each described in the following terms:

Module/Sub-Module title	Based on the common vision
Vision for module	Explanation of the purpose of the module
Sources	Sources of information, especially from APARSEN deliverables
Length (Mins)	15-20 Minutes for the introductory "pills" 20-30 Minutes for the more detailed sub-modules into which each module is divided
Person responsible to prepare module	
Stated Goal	The aim of the module, in terms of how it fits into an overall picture
Participant	Target audience
Benefits for participant	The specific benefits which the participant should obtain through taking the module
Key features for participant	Specific topics which the module should contain, for example links to other modules
Material for manual and further info	For example from existing APARSEN deliverables
Outline of exercises	Description of exercises for students with detailed examples

[69] Available at https://www.dropbox.com/s/40530bzub5ybds5/Training%20plan%20outline_v3.xlsx

Guided by the plans for the modules, itself guided by the common vision, allow the VCoE the means by which to provide a most advanced and consistent training offering.

3.3 CONSULTANCY

There are a number of areas of consultancy in digital preservation. It is difficult to provide any generalisations about them.

Consultancy provided by the VCoE will have the advantage of being based on the members, which have a great number of experienced practitioners, supplemented by the evidence which has been collected by the APARSEN project.

4 THE BUSINESS CHALLENGE FOR THE VCOE

4.1 OPPORTUNITIES AND THREATS

The analysis of the context for the creation of the VCoE is framed by a general positive trend: If society becomes truly digital as it is expected, its heritage and, more in general, its memory, will be based in digital forms. In DP demand terms, this means that this growth in creation of digital content should trigger a growing demand for DP technologies, infrastructures, tools, and solutions."

But the landscape for DP related practices is not simple one. There is a wide array of Threats and Opportunities arising from the current **political, economic, social and technological trends configuring the competitive context for deploying and supporting sustainable DP practices**:

Opportunities for the VCOE

- **Consolidation of the Information Society**: everything and everyone will involve the digital in future. There will be new patterns of knowledge acquisition coupled with rising demand from industry and diverse business sectors, ranging from energy production equipment to pharmaceuticals and from financial services to nationwide healthcare systems. **Demand for DP solutions will accelerate as the Information Society continues to spread, increasing the need to preserve and enable access to increasingly large volumes of digital content.**
- **A data deluge**: explosion of digital content and formats such as social media content, multimedia
- Memory institutions such as libraries, are becoming **digital organisations**, and recognise themselves as such in their own Mission statements. This change in the nature of core businesses is likely to become a top driver for the adoption of DP practices, especially among those organisations charged will legal mandates to preserve new digital realm (Legal deposit)
- **Lack of evidence**: There is an increasing deployment of evidence-based policy, along with a mounting demand on the public sector to provide "more with less". Evidence-based policymaking practices demand measurability of results, impact assessments and return on investment. Such evidence may come too late and, unfortunately be the result of a 'digital disaster'.
- The stage of **development of DP as a business practice is still far from being mature**, even in the most advanced economies of the world:
    - DP deployment is still concentrated in the activities of a few pioneering institutions, to a great extent implemented through short term- tactical measures (projects, pilots,…), funded by fragile sources, and to a large extent still fragmented. In other words, the typical picture of an incipient –far from matureness activity. Out of more than half Million Memory institutions, just a minority of "**early adopters"** have implemented DP practices.
    - In global terms, demand for both DP technological solutions and services have so far not generated enough traction to attract enough leading ICT industry players. Memory institutions have been forced to address the development of their own solutions

Threats for the VCoE:

- **Misconceptions around "digital lasts forever":** lack of awareness around ownership and responsibility for preserving digital content. There is a perception that simply converting from an analogue format will ensure long term preservation, even though digitally encoded information, is itself, not simple to preserve
- **Compliance with institutional mandates for the leading providers of knowledge are proving not to be adequate**: the monopolies in Public Sector Information supply are falling one after the other and being replaced by other sources such as general purpose search engines or specialised providers of information.
- **Being part of public sector is not a 100% guarantee for long term support:** Increasing requirements for demonstrating Return-On-Investment , increasing political demands for proper economic/employment/wealth creation. Austerity in public sector expenditure, cross

sector/generalised budget cuts/ freezes/ "Fiscal Cliffs", deaths/abrupt termination of entire lines of public sector driven activities and institutions.

- **Inadequacy of regulatory frameworks:** By its nature digital preservation is a long-term economic activity. Development of this kind of activity requires a favourable, robust and stable legal and institutional framework as the basic conditions to survive. In the digital realm, this requirement is far from being satisfied. Regulatory frameworks are constantly changing and will keep on changing: they form a structural component for future DP management and sustainability (e.g. what to preserve?).

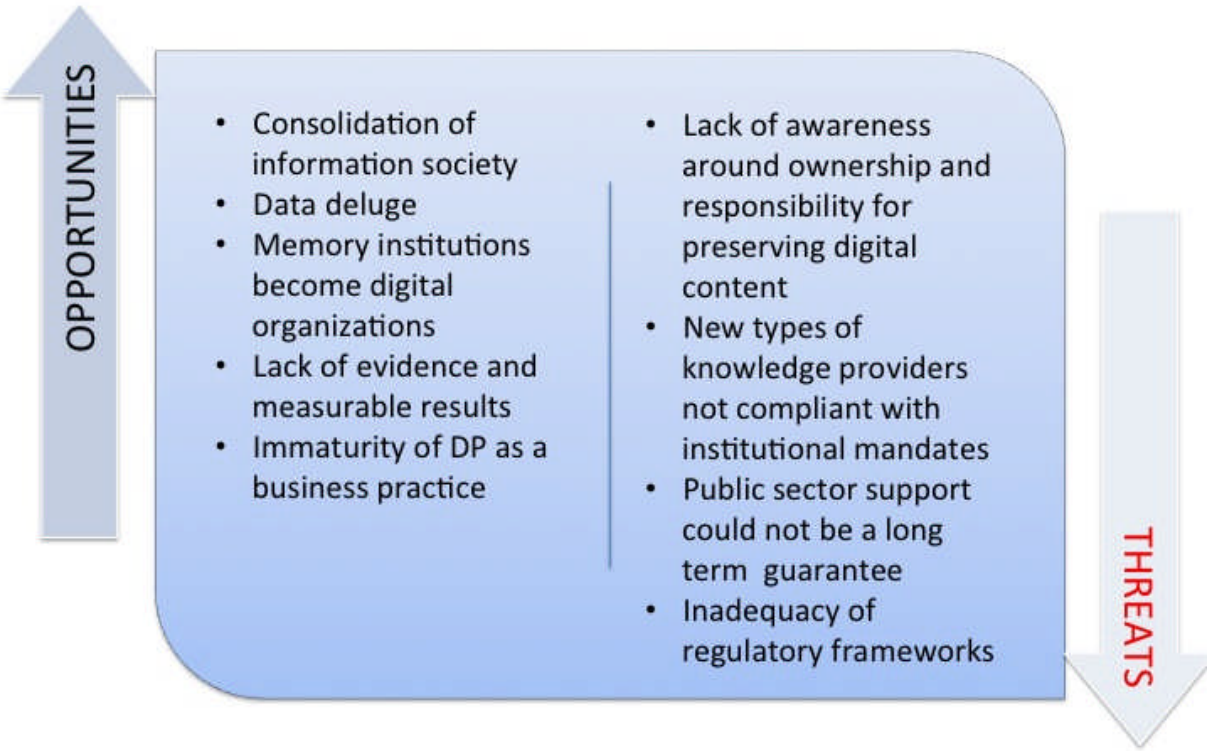


Figure 18: VCoE opportunities and threats

4.2 OPPORTUNITIES GENERATED FROM DEMAND SIDE OF MARKET

For Services and positioning:

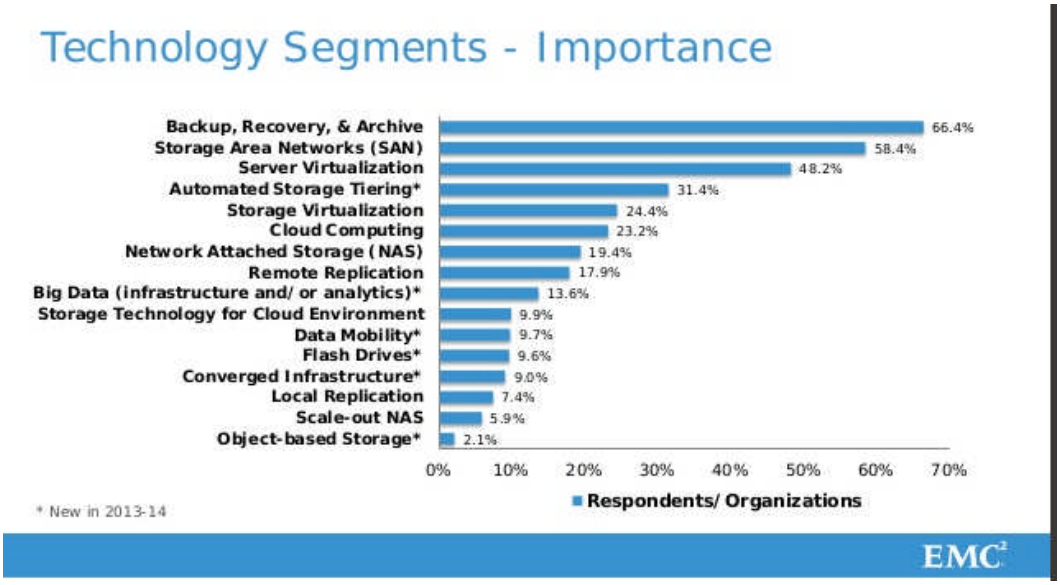
- Increasing adoption of DP in the organisation`s Governance and policies areas.
- Increasing need of efficient and cost effective DP approaches
- Opportunities for integrating vendors for providing and piloting price oriented scalable solutions. VCoE has a very good position to influence the shape of the "DP industry" structure and ecosystem.

Governance and Compliance:

- Increasing need to have an overall view for governance
- Opportunity to support conveying the messages of the DP value to executive team

Favourable market conditions:

- For SRIs still 60% of needed developments pending and with budget allocation. Also there is an increasing need of research data, data computing, intensive research; thus DP is needed as well as the need of new computational services to meet the explosion of data produced by e-Science infrastructures.
- For industry IT spending in the ILM and IG areas steady and slowly increasing.
- E-publishing and e-books markets and trends increasing
- Number of researchers expected to grow by 1M, thus the publications also
- Libraries willing to share and deploy DP piloted systems, tools and developments
- Deep commitment of the involved team. DP is considered as important amongst the cultural heritage organisations; although budget constraints.
- Increasing need to create value added services over current DP infrastructure developments.
- Increasing needs of DP tools to support storage and governance areas in the industry side.



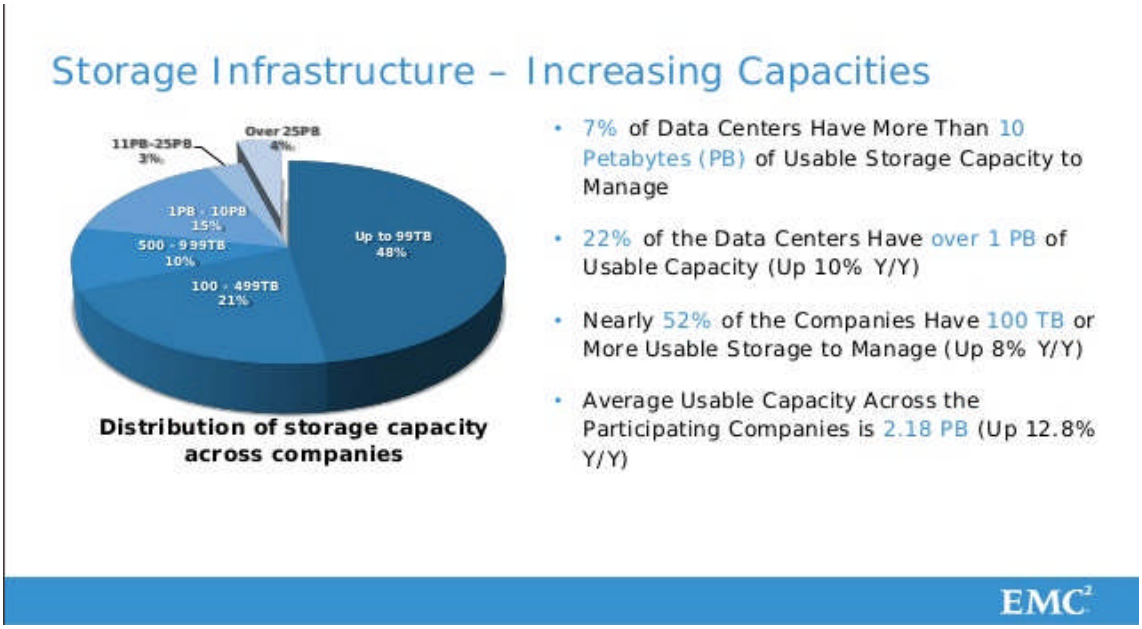


Figure 19: Managing Storage: Trends, Challenges, and Options (2013-2014), EMC, 2013

4.3 OPPORTUNITIES GENERATED IN SUPPLY SIDE FAILURES

Maturity:

- Increasing development of tools and new deployments of already "piloted" DP solutions or tool; but DP suppliers not yet integrated vertically or horizontally to compete better. Thus there is a huge opportunity to act as a DP technology broker to create DP ecosystems.

Leadership

- Strong expertise in DP technology, standards, frameworks, interoperability

Gaps in positioning and Governance

- Take advantage of the opportunities in the IG for Cultural Heritage Sector
- Different perceptions of the value of DP, opportunities to support in conveying a message
- Lack of shared priorities and good metrics
- Lack of mature central process

Budget Constrains

- Cultural Heritage organisations are understaffed, thus opportunities for outsourcing and consulting

4.4 WHY THE VCOE COULD OUTPERFORM CURRENT SUPPLIERS

Positioning in gap areas in which VCoE has real strengths; Sandwich strategy top and bottom, building towards the ecosystem centre. On the top: VCoE can play in the Governance area, building on the experience from standards, reference frameworks and providing the overall vision. VCoE need to build partnerships with suppliers and position as a neutral "DP broker" that provides the best advice, solution and overall vision. VCoE can provide the knowledge to support vendors to integrate horizontally and vertically in DP to grow as an ecosystem that could grow to play strongly in the industry arena against larger IT vendors.

VCoE needs to build on its contacts, to support not only the demand but also the DP vendors to: partner with other DP vendors to create new solutions and partner with the demand to provide the best suitable solution. At the same time VCoE through consulting services can act as a catalyser for DP, supporting DP inclusion in organisations policies and governance.

4.5 WHY DEMAND (AT LEAST AT PART OF IT) WILL ACTUALLY CHOOSE THE VCOE

- Thought for digital environments (not adapted neither brought from other environments) and digital growth
- Made by scientist for science and industry related with science
- Branding and experience (state of the art)
- Neutral
- Broker of DP tools
- Pioneer
- Flexibility
- Provide certification from APARSEN?
- Lower overheads, good when working with government and more competitive
- A start up from a project with long standing experience and branding in the preservation area.
- Identify as entrepreneur, government sees entrepreneurship as an essential element in rebuilding economy
- VCoE can show demonstration of value of their potential purchase (STM report): →many use cases and long standing implantation

5 THE VCOE BUSINESS DESIGN

5.1 BUSINESS MODELLING

To develop and present the VCoE business model we adopted the widely used "Business model Canvas" described within the book Business Model Generation[70] which provides an applied and tested method for development of business models, especially those related to innovative products and services.

The Business Model Canvas is based on four areas: customers, offer, infrastructure and financial viability; which contain nine building blocks: customer segments, value propositions, channels, customer relationships, revenue streams, key resources, key activities, key partnerships, and cost structures.

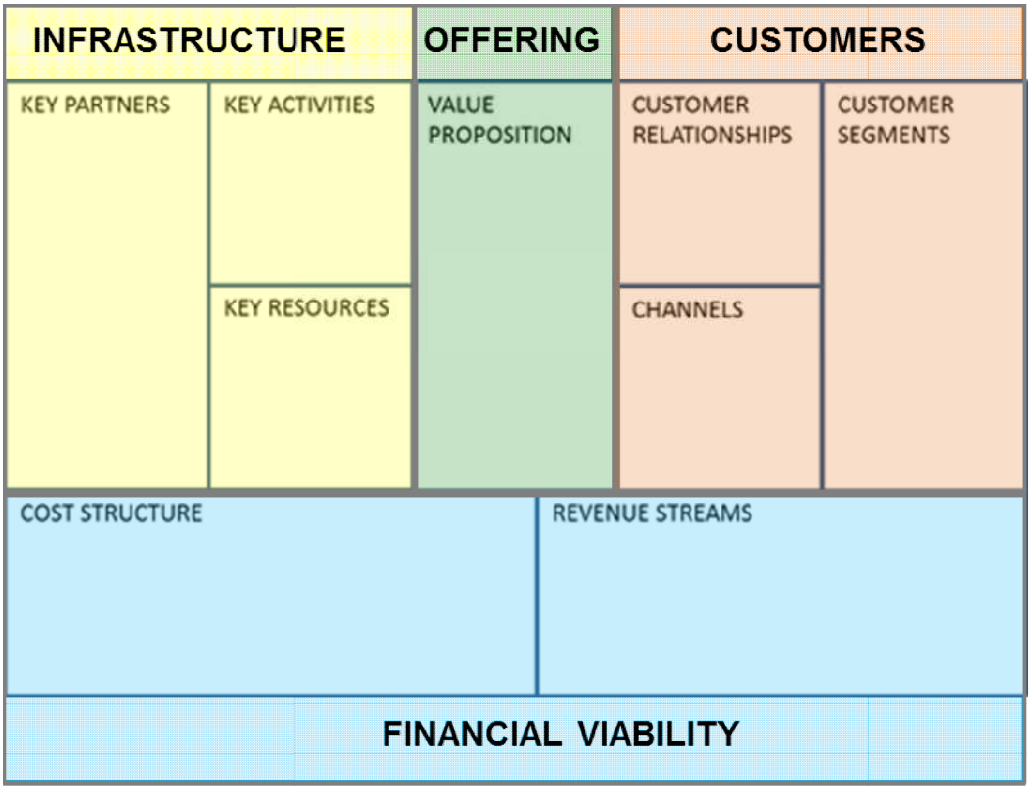


Figure 20: Business Model Canvas

5.1.1 Target markets, value propositions and revenue streams

The VCoE will primarily target organisations which have digitally encoded information which they wish to or they must preserve. The VCoE should be able to support related stakeholder organizations including research institutions, private companies, third-party archives, professional societies working with domain and preservation experts to ensure that personnel are fully equipped with the technical skills needed for selecting, curating, and preserving materials, as well as fund internal preservation and access activities as core infrastructure.

The VCoE will be able to provide training, consultancy and software services to support organisations which need to ensure that their digital resources remain understandable and usable. It can also help to create the business case to justify the resources needed.

[70] Oesterwalder, Alexander; Pigneur Ives. "Business Model Generation".Self published. 2009. ISBN: 978-2-8399-0580-0

The unique selling point is that the offering is coherent and consistent and should be applicable to any type of digital objects. While there are specific data areas such as documents or audiovisual which have community recognised preservation techniques, we believe that most if not all repositories will be called on to preserve many different types of information.

The actual mix of services offering will come from members themselves. Members obtain benefits including, but not limited to, being able to access additional, non-public material, discounts on access to the offerings and to present a unified voice to policy makers and funders.

The following table summarises the preliminary analysis[71] of key components of the business model for the VCoE, including the value proposition, targeted customer segments and potential revenue streams.

Value proposition	Customer segments	Revenue streams
• Pointers to appropriate skills and experts	ALL	Free pointers but could lead to charges e.g. need to become member i.e. Membership
• Help in creation of DP business cases • Forecasting future demand of preserved objects • Including cost/benefit modelling tools and examples? • Advice on how to add value • Case study formats for business cases and best practices?	Data holders Data creators	Usage fee
• Help in becoming Trustworthy Digital Repository • facilitate the maintenance tools for DP • Pointers to evaluation or rating tools • Audit and certification support	Data holders	Consultancy Training fees Usage fee
• Advice on which tools and services are best to try first (i.e. reduce effort in trails) for DP • Evaluation/rating of tools/services • Licensing advice, PI etc. • Standards • Brokerage e.g. advice on outsourcing and services	Data holders, DP Accreditation and certification bodies	Free pointers but could lead to charges e.g. need to become member, hence Membership Training fees Usage fee
• R&D for new tools and services • Incubator for new project for members • Surveys about challenges/issues in DP	• DP Research institutions, • DP Service providers	• Projects
• Supply infrastructure to be shared by members (association management, phone)	• small members	• Usage fees • Membership

[71] This work was carried out in a series of dedicated workshops, following the methodology proposed in the book "Business Model Generation" by Osterwalder and Pigneur (2010).

Value proposition	Customer segments	Revenue streams
• Lobbying services and awareness raising • Visibility for digital preservation—lobbying • Responding to/being included in consultations etc. • Targeted workshops to find needs of big industry—what do they need to access? • Conferences	• ALL	• [included in Membership]
Qualifications (basic/intermediate/advanced) pointers to • People • Organisations • Systems • Data	• Data holders, • DP Service providers, • Data creators	• Training fees • Consultancy
• Satisfy a need for training • Recommendations for DP courses/material • Data management training (approved? recommendation)	• Data holders, • Data creators	• Training fees • Brokerage fees
• Answers to questions (initial consultation free?) - What are your problems? • Web site • Email • Phone helpline? • Consultancy—need to understand problems	• ALL	• Free initially then • Usage fees, • Consultancy • Sponsorship for website?
• Help in creating data management plans • guidance to appropriate experts (who may charge) • web resources	• Data creators	• Consultancy
• Prestige through membership • Membership criteria dependent	• ALL	• Membership
• Provide visibility and marketing channels for members e.g. on web site	• ALL	• Sponsorship (even if members) for specific products
• Digital Preservation services e.g. from SCIDIP-ES (www.scidip-es.eu)	• ALL	• Free or Usage fee and/or Training
• Help on customisation of preservation tools and services • Guidance on systems integrators etc.	• Data holders	• Consultancy
• Preserving DP related resources e.g. training, tools etc.	• DP Research institutions	• Free or Usage fee

5.1.2 Customer relationships

The APA, potentially positioned as the core organisation within the VCoE, has the benefit of a membership of which, each individually, has international repute with connections into many disciplines and thousands of users. Complementing this is the international reputation in digital preservation of the APA director and the reputation of the APA as a united voice in Europe.

The marketing of the APA will leverage both these important points.

Digital preservation is intrinsically cross-disciplinary and restricted neither by national boundaries nor by timescale. Moreover it is clear that there is a demand that society benefits from its digital intellectual capital. Therefore the positioning of the APA is global, addressing the preservation of and adding value to all types of digitally encoded information.

The branding will be based on the "Common Vision" figure:

Figure 21: Common vision overview diagram

This is explained fully in the document dedicated to the Common Vision document D11.5.

The APA website is the prime source promotional outlet, supported by academic publications and presentations in a wide variety of conferences and workshops.

In addition the promotional material (glossy publications) produced by the various projects of which the APA and its members have been part, should remain usable and timely for several years.

5.1.3 Key activities

The **functional and operational activities** for making the VCoE business model work comprise the following:

To maintain a customer centric approach.

To develop needed service delivery practices and secure needed skills

To offer a level of operational excellence that meets the customers‘ expectations; and develop the SLA/service level agreement needed for providing VCOE products and services.

To maintain the market research and competitors/price watch to establish a market positioning

To develop IPR protection practices

To dynamically manage the service and suppliers mix to secure the best possible matching between the overall VCoE offering on the one hand and the changing structure and behaviour patterns in the demand side, on the other hand.

The specific activities include:

- advertise the offerings
- maintain the list of solution contributors
- maintain the software and database resources
- obtain business contracts
- put together proposals for solutions
- put the solutions into operation

5.1.4 Key resources

Key **resources comprise the necessary capabilities to create value for the VCOE customer segments**, and include human, intellectual and marketing resources (financial are taken in account in financial viability).

The key resource created by APARSEN is knowledge of:

- the capabilities which the APARSEN organisations are willing and able to make available for the VCoE
- how to fit the various capabilities together to solve problems
- how to guide the selections of tools and services to address issues about digital objects, including new types of digital objects.
- training resources

This knowledge underpins and identifies the resources needed to allow the VCoE to perform the key activities.

To support the key activities the following key resources are required:

Key Activity	Key Resources
Advertise the offerings	Web site, presentations, advertising
Maintain the list of solution contributors	Email list and personal contacts
Maintain the software and database resources	Effort to maintain and update the databases (1) software (2) standards and also software (e.g. from SCIDIP-ES[72])
Obtain business contracts	One or more people and a legal entity for contractual obligations

[72] See http://www.scidip-es.eu and also http://int-platform.digitalpreserve.info

Put together proposals for solutions	People with the skills to put together appropriate solutions by integrating the appropriate components Arrangements with solution contributors , possibly: • contractual arrangements with the purchaser and with individual solution contributors • contractual arrangements with purchaser. The purchaser has contractual arrangements with the individual solution contributors
Put the solutions into operation	People with skills to perform the integration e.g. • Organise training courses • Analyse existing systems in order to provide appropriate advice • Integrate preservation software/ services into existing systems
Maintain the organisation financially	Provide any necessary financial “buffering” to smooth out income from the offerings

6 ORGANISATIONAL OPTIONS: IMPLEMENTING THE BLUEPRINT

6.1 GENERAL PRINCIPLES

The foregoing surveys, analyses and discussions have described the characteristics of the landscape in which the VCoE must operate, particularly with respect to the demand for digital preservation, the existing offerings that are available, and aspects of business modelling that direct and constrain the VCoE. The final step is to consider what options exist for the organisation of the VCoE so that it may operate effectively in the landscape, adopting structures and modes of functioning that are adapted to the opportunities and constraints of the market and the environment in general.

In the early stages of the APARSEN project, lessons were taken from existing Centres of Excellence/Centres of Competence with a view to understanding the appropriateness of these models for the situation of APARSEN. However it is necessary to keep in mind that the aim is not to set up a particular type of organisation, but to try to ensure that benefits can flow from the combined expertise and experience of the APARSEN members, built up during the life of the project and beyond—this is real sustainability. The question then is what are the options for implementing the VCoE beyond the life of the project—that is, establishing a sustainable legacy as envisaged by the Description of Work.

We cannot unquestioningly adopt a pre-existing structure and implement it within the APARSEN context. All that can be said in general is that the VCoE will comprise members or participants and that it will 'bring the capabilities, knowledge and expertise together from diverse teams across geographical and organizational boundaries to create something exemplary and distinguishable within its domain' (see APARSEN Description of Work), with the key being the vision which is shared by members of APARSEN, including the know-how and the know-why of digital preservation. Furthermore the VCoE is expected to be long-lasting, which implies sustainability in the way it operates.

The concept of VCoE implies a recognised entity of some kind to which people and organisations turn in search of a special kind of expertise that they believe will be of use to them. The 'market' thus defined may consist of those who participate in the VCoE themselves, and those outside it who recognise its special qualities. The VCoE satisfies this demand by offering access to skills and capabilities—either of individual organisations or in concert—to create solutions.

There may be several options for a suitable organisational structure that allows this kind of sustainability. Indeed the VCoE might be 'virtual' in the sense of several distinct implementation routes coexisting. What is clear is that if it is to meet with success it must build on the diversity of the APARSEN partners unified by the common vision.

6.2 ROLES IN THE VCOE

Regardless of what kind of structure is adopted in the implementation of the VCoE, it is possible to identify three roles within it. These roles might overlap, in that one organisation could fulfil more than one role. The roles are:

- members of the VCoE (could also be labelled as also participants or contributors)
- systems integrators
- a core organisation

Of these only the first is absolutely essential: without any identifiable grouping of members or participants the VCoE would not exist at all. The other two roles are optional, though the second is likely to arise naturally from the outputs of APARSEN.

Members/participants/contributors are organisations inside the VCoE that have some common understanding and commitment to the unifying vision and may provide offerings (capabilities/expertise) and also benefit from each other's knowledge. Taking the existing APARSEN partners as a basis for the membership of the VCoE, different organisations have different expectations in terms of capacity and willingness to provide services externally, and under what conditions, but ultimately there is a pool of expertise that is in some way available for the benefit of each other and the

community at large. The VCoE can therefore be a potential outlet for services offered by an organisation to a range of customers. For leveraging the results and benefits, the members would also represent the VCoE in other networks. A prime example is the Research Data Alliance, discussed below, where by taking an active role (as many APARSEN partners do) the discussion will be taken to a global audience.

Systems integrators borrow a name from mainstream information technology for an organisation that brings together component subsystems into a whole and ensures that those subsystems function together. In the case of the VCoE, a systems integrator would operate in the field of digital preservation and, building on the knowledge base of APARSEN, have enough information to make a recommendation for a solution that is likely to work in particular circumstances and have access to components to implement the solution. The foundation on the common vision and a knowledge of the offerings of VCoE members is key to this role. The role could be fulfilled by one of the members with its own offerings, or a third party.

A **core organisation** is a single body at the centre of the VCoE, which could be seen as making it less virtual and more concrete. The VCoE could operate without the existence of a core organisation provided that coherence is maintained and the VCoE is sufficiently strong to maintain its identity. The core organisation could also provide services, and would be a candidate to play the role of systems integrator.

The core organisation would have to function as a business. This means that there are risks to its viability—at least initially, the revenue from the consultancy, training, tools and services might not balance costs; and more importantly the cash-flow, which is generally the greatest threat to a business, is likely to be uneven.

The table below summarises the contributions which the core organisation must make to the VCoE. These are all functions that provide an integrating and unifying face for the VCoE.

Key Activity	Key Resources	Core contribution
Advertise the offerings	Web site, presentations, advertising	Organise and host website
Maintain the list of solution contributors	Email list and personal contacts	IT support system for contacts and email
Maintain the software and database resources	Effort to maintain and update the databases (1) software (2) standards and also software (e.g. from SCIDIP-ES)	IT support Solicits and assesses inputs about software tools and standards. Also maintains the software
Obtain business	One or more people and a legal entity for contractual obligations	Legal entity for receiving money. Legal support for contracts.
Put together proposals for solutions	People with the skills to put together appropriate solutions by integrating the appropriate components Arrangements with solution contributors , possibly: • contractual arrangements with the purchaser and with individual solution contributors	Expertise of core staff plus possible advice from others within the organisation

	• contractual arrangements with purchaser. The purchaser has contractual arrangements with the individual solution contributors	
Put the solutions into operation	People with skills to perform the integration e.g. • Organise training courses • Analyse existing systems in order to provide appropriate advice • Integrate preservation software/ services into existing systems	Depends on the skills of the VCoE core staff and advice from potential solution contributors.
Maintain the organisation's cash flow	Provide any necessary financial "buffering" to smooth out cash flow of income from the offerings	Provide another source of funding to smooth. If the core organisation is a large then this will have the financial base which can maintain cash flow. A membership organisation could use membership fee income to smooth cash flow.

6.3 EXAMPLES OF VCOE IMPLEMENTATION OPTIONS

The following outline a number of options within the above roles. These are not intended to be exclusive—they can and probably will co-exist.

6.3.1 APA as core organisation

The APA Executive Board is considering the next five-year strategic plan. This will take into account the fact that the APA has achieved the aims of its first two strategic plans, while at the same time the Research Data Alliance is gaining a great deal of momentum globally and the position of the APA with respect to the RDA must be clarified. The plans for the future of the APA will be completed before the summer of 2015. An important input to this plan will be this VCoE blueprint to position APA as the core organisation of a VCoE.

The solution contributors would obviously be the APARSEN members—with capabilities detailed in the Annex. Other contributors would be expected to be added over time.

The previous deliverable D11.4 included a detailed financial plan containing forecasts of income and operating expenses. This is not reproduced here.

6.3.2 Giaretta Associates as core organisation

Giaretta Associates will also be a potential core organisation for a VCoE. Giaretta Associates works closely with APA and has a number of specific additional contributions to help provide solutions:

- ISO16363 audit and certification (see http://www.iso16363.org) – to help repositories identify the areas which need to be improved
- SCIDIP-ES project – the SCIDIP-ES software and documentation and services (see http://www.scidip-es.eu and http://int-platform.digitalpreserve.info). These are tailored to help to address the most common areas of improvement for repositories
- Arrangements with the European Grid Infrastructure (EGI) which will help to provide scalable preservation capabilities
- Extensive contacts outside Europe, especially USA, India and China

- Leading the Active Data Management Plans and the Preservation e-Infrastructure Interest Groups in the RDA

These contacts provide a global set of opportunities for a VCoE.

6.3.3 Italian Centre of Excellence as national instance

The possibility of extending the VCoE services by specializing them at national level has been investigated, and has led to the pilot implementation of the *Italian CoE portal* by CINI, one of the Italian partners of the APARSEN project. There are several motivations to this kind of action:

- *National regulations*. Some countries have specific regulations on digital preservation, and these must be taken into account in addition to European and International regulations and standards.
- *Market peculiarities*. These are mostly triggered by national regulations. Taking them into account, may help in identifying and targeting interesting groups of potential CoE customers.
- *Language*. In some countries language may still be a barrier, and there is a clear demand for a national language interface. This is especially true for training activities.

In Italy, that has been selected for a pilot activity, the situation is indeed peculiar, since specific and detailed national regulations on digital records and digital preservation have been in place for more than 20 years. For instance, the digital signature and digital timestamps have been given full legal validity in the mid-1990s. Moreover, for more than a decade, all corporate official documentation must be kept in digital form, notably financial statements, and, more recently, electronic invoicing has become mandatory, at least when the Public Administration is billed. Similarly, the protocol registry, the gateway to all recordkeeping activities in the Italian Public Administration, has moved to digital since 1998. All these regulations have led to a strong drive towards electronic records with full legal value and, consequently, to the need to conveniently preserve them, since paper originals do not exist any more.

As for the digital preservation process, recent regulations have provided quite detailed technical guidelines and introduced, among other innovations, *Preservation Providers* that, as a third party with certification granted by a national agency and based on ISO 16363, offer preservation as a service, and take full legal responsibility of the process. But, even if preservation will be often delegated to certified providers, there is still plenty of room left for action, since some relevant user communities would like anyway to maintain direct control of the structure and the quality of their preservation process, and would consequently need adequate assistance to do that, thus becoming an appropriate target for CoE services.

According to this scenario, the Italian CoE portal, that can be accessed at www.conservazionedigitale.org, is meant to be the main reference point for the Italian digital preservation community, and is aimed at understanding its specific needs, at formulating practical proposals and at promoting the CoE activities to the target audience, by setting the stage for appropriate disseminating actions.

The portal provides a variety of contents:

- Introductory and detailed explanatory materials on the main issues in digital preservation, as for instance the organization and the management of ERMS and preservation repositories, the file formats and the metadata sets suited for preservation, the management of authenticity and provenance and the transmission of electronic resources.
- State of the art on digital preservation: what is currently done at national and international level, which proposals have been made by individual countries and by supranational organizations, which are the reference organizations in these countries and which research projects have contributed to the topic and are currently underway.
- Standards and guidelines issued at national and international level, thematically organized by issue, and complemented by introductory and explanatory materials in Italian, in order to allow the understanding of their relevance and their range of applicability. Information is also provided about availability of the standards together with a variety of links.

- Italian national regulations on digital preservation, and related issues (digital signature, protocol systems etc.), which are presented in an historical perspective and discussed in detail.
- Training materials on all the above topics. Most training modules are in Italian and provide a complete coverage of the main issues. Additional training modules in English are available as well. For each module a summary and a preliminary comment is given to guide the user in selecting the appropriate ones. Moreover, a few training itineraries are suggested, based on the needs of different classes of users.

As already stressed, a central feature of the portal is the national language interface, which has been designed to allow easy surfing and access to the contents. Most of the materials, especially on-line training materials, are in Italian, but, even for the ones in other languages, some kind introductory information in the national language is always supplied, to provide at least a preliminary understanding of the problems and to point out the relevance of the main issues. This promise to be a crucial element in attracting the national audience, since in Italy, unfortunately, there are still too many people, not sufficiently acquainted with the English language, especially in the humanities and cultural heritage milieu that is traditionally interested in the digital preservation issues.

6.3.4 The Research Data Alliance

The Research Data Alliance (www.rd-alliance.org) is an open and global organisation bringing together a large community of individuals and organisations to 'build the social and technical bridges that enable open sharing of data'. The RDA, like APARSEN, recognises fragmentation in the current situation which is a barrier to open data sharing. Working Groups and Interest Groups formed of experts from around the world from academia, industry and government operate to break down the barriers.

The scope of RDA is wider than digital preservation, of course, but digital preservation is certainly one of its themes. At the APARSEN General Assembly meeting in September 2014 it was agreed that the RDA is a very suitable vehicle for taking on the mantel of promoting digital preservation research and practices. Many APARSEN members are already active in RDA, and the reality is that the RDA is a perfect forum for this exchange. The focus on 'research data' is not a restriction as there are Interest Groups covering wider areas such as multimedia and research libraries. The RDA could therefore fulfil the role of maintaining the common vision of APARSEN, providing a formal forum for the members to cooperate, and allow dissemination to the wider community.

The following partners in APARSEN are RDA organisational members:

- CSC
- DANS
- STM
- STFC
- LIBER

The following RDA Working Groups are directly related to aspects of digital preservation.

RDA Working Group	APARSEN members participating
Data citation	secure sba-research.org; HELMHOLTZ ASSOCIATION; CSC; Science & Technology Facilities Council
Metadata standards	Science & Technology Facilities Council
Repository Audit and Certification DSA–WDS Partnership	HELMHOLTZ ASSOCIATION; UK DATA ARCHIVE

The following RDA Interest Groups are directly related to digital preservation.

RDA Interest Group	APARSEN members participating
Active data management plans	APA Alliance for Permanent Access; Science & Technology Facilities Council; CERN; UK DATA ARCHIVE Note: David Giaretta (APA) is co-chair
Data in context	Koninklijke Bibliotheek National Library of the Netherlands; Science & Technology Facilities Council; HELMHOLTZ ASSOCIATION; UK DATA ARCHIVE; CSC
Persistent identifiers	BRITISH LIBRARY; Fondazione RINASCIMENTO digitale; UK DATA ARCHIVE; DEUTSCHE NATIONAL BIBLIOTHEK; CERN; secure sba-research.org
Preservation e-infrastructure	secure sba-research.org; Koninklijke Bibliotheek National Library of the Netherlands; APA Alliance for Permanent Access; Science & Technology Facilities Council; UK DATA ARCHIVE; CERN; DEUTSCHE NATIONAL BIBLIOTHEK Note: David Giaretta (APA) and Jamie Shiers (CERN) are co-chairs

RDA/WDS Certification of Digital Repositories	secure sba-research.org; KB Koninklijke Bibliotheek National Library of the Netherlands; UK DATA ARCHIVE; Data Archiving and Networked Services DANS; CSC; CINES Centre Informatique National de l'Enseignement Supérieur
Research data provenance	secure sba-research.org; KB Koninklijke Bibliotheek National Library of the Netherlands; UK DATA ARCHIVE; Science & Technology Facilities Council

ANNEX 1. SERVICES PORTFOLIO CONFIGURATION

The members of APA and APARSEN have provided information about the potential contributions which could be provided through th e Virtual Centre of Excellence. The tables focus in turn on consultancy, software related, training and other.

I. Consultancy

Member Name and contact details	Organisation	0	Consultancy	0
0	0	Fee	Description	Partners in delivery
0	Airbus	0	0	0
David Giaretta & Simon Berriman E.mail: director@alliancepermanentaccess.org Tel: +44 1935 872660	APA (Alliance for Permanent Access)	500 per day (average)	Consultancy in digital preservation to research, library, government and commercial organisations, adding value, data management plans, best practice, preparation for audits etc	0
Kirnn Kaur	BL	0	No services offered	0
Jamie Shiers, Jamie.Shiers@cern.ch	CERN	N/A	Consultancy on "bit preservation" techniques and experience. This is based on CERN's experience at the 100PB scale and on the results of the (recent) coordination effort that has started across the main High Energy Physics (HEP) institutes.	HEPiX members: https://www.hepix.org/

Mariella Guercio, e-mail: m.guercio@mclink.it, tel. +39 3476058341; Silvio Salza, e-mail salza@dis.uniroma1.it, tel. +39 3388492106.	CINI	500 Euro per day (average)	Consultancy to private and public organisations on design and implementation of digital preservation repositories. Includes: analysis and assessment of existing repositories and current practices, feasibility studies, design of OAIS compliant repositories, preparation for audits and certification procedures. Special emphasis on workflow design, metadata, authenticity management, redundant and persistent storage solutions and interoperability issues.	0

DANS - Data Archiving & Networked Services Rene van Horik. Email: rene.van.horik@dans.knaw.nl. Tel: +31 70 344 64 84	DANS - Data Archiving & Networked Services	100 euro per hour	DANS promotes sustained access to digital research data. For this purpose, DANS encourages scientific researchers to archive and reuse data in a sustained manner, e.g. by means of the online archiving system EASY. DANS also provides access, through NARCIS, to thousands of scientific datasets, e-publications and other research information in the Netherlands. In addition, the institute provides training and advice and performs research into sustained access to digital information. Consultancy in research data management to research organisations (data curation, data storage, data documentation, etc.).	0
Sabine Schrimpf Email: s.schrimpf@dnb.de Tel: +49 69 1525 1761	DNB	0	Risk management during the ingest process based on DNB's ingest level policy DNB has also access to the know how in the nestor network	0

William Kilbride, Digital Preservation Coalition, Innovation Centre, York Science Park, Heslington, York, YO10 5DG, UK, info@dpconline.org	DPC	Price on application. DPC is a membership organisation and existing members have priority access to these services.	DPC offers occasional consultancy around the following topics which are drawn from its strategic plan: Workforce Development; Advocacy; Knowledge Exchange; Assurance and Practice; and Developing Partnerships. DPC also facilitates and organises events on behalf of third parties. DPC is a membership organisation.	DPC Members if appropriate to the request.
Yannis Tzitzikas, Information Systems Lab, FORTH-ICS, Crete, GREECE	FORTH	Variable	Semantic Technologies for Digital Preservation. Automated Reasoning for Interoperability	0
Holger Brocks, Jana Becker, Beate Deska E.mail: hbrocks@ftk.de, jbecker@ftk.de, bdeska@ftk.de Tel: +49 231 9750560	FTK	Various	Consultancy in Digital Preservation to SMEs, Memory Institutions, Public Administration, Authorities, Municipalties, Ministries, Intermediaries. Consultancy in Advanced Research & Development as well as Proof-of-Concept Prototypes, Feasibility Studies and DP Market Research/Segmentation.	0

open-access@oa.helmholtz.de	Helmholtz	500 per day (average)	Consultancy in open science policy; data publication journals; data and publications repository management	0
Hans-Ulrich Heidbrink & Andreas Hundsdörfer / email: hans-ulrich.heidbrink@incontec.de andreas.hundsdoerfer@incontec.de / Tel: +49-5251-931494	InContec	TBD	Consulting services for developing strategies, roadmaps, requests for proposals, market surveys and project structuring regarding infrastructures for long-term preservation of digital-born datasets, multimedia objects and electronic documents from libraries and archives, academic research and industrial engineering; Collaborative Development Environments for Mechatronics Systems	0
Andreas Rauber (rauber@ifs.tuwien.ac.at), Stefan Pröll (sproell@sba-research.org), Rudolf Mayer (rmayer@sba-research.org), Stephan Strodl (sstrodl@sba-research.org), Elisabeth Weigl (eweigl@sba-research.org)	SBA	0	Consultancy services for libraries, governments and commercial companies in the areas of digital preservation and preservation planning.	0

simon.lambert@stfc.ac.uk	STFC	Probably free for initial contact in spirit of sharing knowledge and experience in areas of mutual interest; more substantial and formalised work at rate consistent with other VCoE members	(1) General consultancy in research data management and curation for (scientific) data-holding organisations (2) Advice on repository audit/certification, for example in preparation for audit (3) Advice on digital preservation policies and linking policies to operations (from SCAPE)	APA

Jon Tilbury Email: Jonathan.Tilbury@tessella.com Tel: +44 1235 555511	Tessella http://www.digital-preservation.com http://preservica.com	£800 per day (average)	Consultancy for: - Identifying the drivers for archiving. - Categorizing the data storage needs. - Planning the data archiving workflows. - Devising a preservation strategy. Specifying the security requirements. - Helping ensure compliance with legal and regulatory requirements. - Investigating how users will search and access the contents of the archive. - Planning data disposal – what happens to information once the requirement to keep it expires. - Exploring how the archive integrates with operational systems. - Selecting the appropriate software solution. - Planning and deploying the selected solution. - Advising on digitization strategies and integration with non-digital archives.	0

email: comms@data-archive.ac.uk Tel: +44 1206 872001	UKDA	500 per day (average)	Consultancy in research data management, and digital curation to research, government and commercial organisations	0
Paolo Bouquet & Barbara Bazzanella email: bouquet@disi.unitn.it; barbara.bazzanella@unitn.it	UNITN	TBD	Consultancy in Persistent Identifiers and interoperability solutions and services to create a global interoperable data exchange framework, to different stakeholder communities such as publishers, libraries and digital libraries, researchers, research institutes, universities, funding agencies, e-infrastructure operators, data centers, author rights organizations, governmental organizations	UNITN & OKKAM srl

II. Software products and software support and upgrades

Organisation	0	Software Products	0	0	Software Support & Upgrades	0
0	Fee	Description	Partners in delivery	Fee	Description	Partners in delivery
APA (Alliance for Permanent Access)	Variable	Dedicated implementation of SCIDIP-ES services and customisation of service installation and tools	0	15% of software costs	Annual upgrades and bug fixes	0
CINI	The software is freely distributed. Consultancy fees for customisation 500 Euro per day (average)	SCIDIP-ES Authenticity Toolkit, is a modular framework, based on REST (REpresentational State Transfer) architectural style, designed and implemented by a joint APARSEN-SCIDIP-ES team. It can be conveniently exploited by any organisation implementing a LTDP repository, or improving an existing one, to introduce suitable procedures to manage the authenticity evidence of preserved digital resources. This approach allows to reduce the implementation effort, by exploiting efficient APIs (Application Programming Interfaces) that provide a complete set of functionalities to implement the APARSEN authenticity model. A consultancy service is	Engineering SpA	0	0	0

		provided for customisation and integration in a specific repository.				
DANS - Data Archiving & Networked Services	Variable, based on project plan	Software development concerning the management and dissemination of research data (e.g. Data analysis, Data visualisation, Data linking, Data sharing)	0	0	0	0
DNB	variable	International URN:NBN cluster as a common infrastructure for URN resolving with one resolving service for multiple URN namespaces	DANS, KB	0	0	0
FORTH	0	Several tools and systems ranging from semantic registries and semantic repositories, to tools that can automate the ingestion process or interoperability testing.	0	Depending on the level of support and the terms of the contract	Annual upgrades, improvements, extensions	0
InContec	Variable	Custom-tailored software implementations in the areas of long-term preservation and collaborative engineering environments	0	TBD	Technical support for custom-tailored software	0
STFC	Variable	Customisation of some of the toolkits coming out of SCIDIP-ES	APA	0	0	0
Tessella http://www.digital-preservation.com http://preservica.com	Variable £1,100.00	SDB: Fully customisable award winning Digital Preservation system. PRESERVICA: Digital Preservation as a Service on the Cloud (based on SDB technology with pre-fixed functionality)	0	PRESERVICA: Support & upgrades included in price SDB: Yearly support fee	PRESERVICA: Services available with an uptime service availability level of at least 98% during each Subscription Quarter (see details in contract). Regular	0

	1000 GB/Month Variable	PRESERVICA LE: Version of Preservica intended for installation within the client's own data centre connected to own storage system. Delivered as a single pre-built virtual appliance. All products support Active Preservation and all aspects of the Digital Archiving maturity model: http://www.digital-preservation.com/digital-archiving-maturity-model/			upgrade. SDB: Support service available with SLAs. Yearly upgrade.	
UKDA	Apply for details	HASSET (Humanities and Social Sciences Electronic Thesaurus). See: http://data-archive.ac.uk/find/hasset-thesaurus	0	0	0	0
UNITN	Free	Entity Name System (ENS), a system for managing the lifecycle of globally unique identifiers and their interoperability with alternative identifiers. http://api.okkam.org/ The not-for-profit public part of the ENS infrastructure will be made available for the entire community of web users in the form of a public TRUST (the OKKAM Trust), under the control of a high-level international and independent board of	OKKAM srl	TBD	Technical support and development of customized services on top of the ENS	OKKAM srl

		protectors. The VCoE of APARSEN could coordinate this independent authority in order to ensure the proper use and development of the system and guaranteeing its trustworthiness and long term sustainability.				

III. Training courses

Organisation	0	Training Courses	0
0	Fee	Description	Partners in delivery
APA (Alliance for Permanent Access)	Various	Most aspects of digital preservation, from fundamentals of OAIS and preservation of all types of digitally encoded information. Covers software implementations. Courses from single day to week long summer schools.	0
CINI	Variable	Courses at several different levels: management executives, technical and operational staff. Possibly taylored on specific needs, may cover most aspects of the management and preservation process of digital data. Modules from a single day to week long training courses. E-learning module can be provided as well.	0
DANS - Data Archiving & Networked Services	Various	Most aspects of research data management. The training can be directed towards the implementation of a data management policy for research organisations such as universities. Also practical hands-on training concerning data curation. Some training modules will become available online for free. Occasionally face-to-face meetings will be organised for which a fee will be charged. These training modules can be adjusted for specific usage on request.	Research Data Netherlands (RDNL) Alliance
DNB	0	Annual nestor school (available in German only; not offered by DNB directly but by a group of nestor partners)	0
DPC	Various	Range of training courses from elementary through to advanced and subject specialist courses. DPC also facilitates and organises events on behalf of third parties. The DPC publishes a variety of reports which support training including its highly respected Technology Watch series and the Digital Preservation Handbook.	DPC Members if appropriate to the request.
FORTH	Various	Courses on topics on Knowledge Management for Digital Preservation	0
FTK	Various	Construction and coordination of regional, national and international training programs and materials.	0
InContec	Various	For all InConTec consulting services and custom-tailored software	0

SBA	0	Focused training courses, seminars, tutorials and interactive workshops. Special topics: preservation planning, content analysis, software escrow, process preservation, data citation, data security	0
STFC	Depends on requirement for development of training materials	(Later, possibly growing out of the consultancy) Training for other scientific data organisations in data preservation issues	0
Tessella http://www.digital-preservation.com http://preservica.com	Variable	User and/or Admin training available on both PRESERVICA, PRESERVICA LE and SDB. Developer training available for 3rd parties	0
UKDA	c£250	Annual 'How to Run a Data Service' training courses. (e.g., http://www.data-archive.ac.uk/news-events/events.aspx?id=3543) Various training courses in research data management (for researchers, with elements of DP.)	0
UNITN	TBD	Training activities on services to "okkamize" entities in digital documents and resources (i.e. to assign ENS persistent identifiers to entities of different types) and integrate information across systems and platforms.	0

IV. Other services/ Products/ Tools

Organisation	0	Other Services/Products/Tools	0
0	Fee	Description	Partners in delivery
DANS - Data Archiving & Networked Services	(1) Depending on data volume to be stored in TDR. (2) 75 Euro per hour	(1) Trusted Digital Repository (TDR) for research data sets. Storage of research data up to 1 GB in Trusted Digital Archive is free, provided that the data is available as Open Access data (or within 2 years after deposit). For storage above 1 GB a fixed fee will be charged. The TDR will curate the data for an indefinite time period. URL of TDR = http://easy.dans.knaw.nl (2) Research data sets to be archived in the TDR should be properly documented and organised. Additional documentation activities to correct and adjust the documentation and organisation costs 75 euro per hour.	0
DNB	OSS, available from http://kopal.langzeitarchivierung.de/index_koLibRI.php.en	Didigo (Diagnose Digital Objects) service for format identification, validation, technical metadata. Didigo allows the detection of DRM protected publications.	0
DPC	Various	The DPC supports its members through peer review of practice and accreditation of training. The DPC would be willing to also facilitate these processes for organisations outwith our membership.	DPC Members if appropriate to the request.
FTK	Various	Continuing the roadmapping processes for training and education as well as for stakeholder communities in the field of DP infrastructures and continuing the validation process for the DP curriculum.	0
GLOBIT	Various	Operation of the Online Training Portal (OTP) for providing training and qualification in DP; Content management for DP learning resources; Production of educational content in DP; Scientific community support for DP; Scientific community management for DP; Community DP knowledge management; Preservation support for scientific communities; DP resource mediation; Roadmapping for Continuing Professional Education (CPE) in the field of DP;	0

LIBER	0	DART-Europe Thesis Portal	info@dart-europe.eu Networked Digital Library of Theses and Dissertations
SBA	0	Case studies, feasibility studies, focused research projects	0
Tessella http://www.digital-preservation.com http://preservica.com	0	0	0
UNITN	0	A wide set of services are offered to support applications to retrieve entity identifiers. These are available to developers as SOAP Web Service or HTTP REST-like Web services, and to final users through web applications.	OKKAM srl
Finnish Academy of Science & Letters	The cost of keeping publications available is difficult to estimate since it is connected to the production of journals and books and is a relatively small part of that	The Finnish Academy of Science and Letters publishes a number of journals and books (the books are organized in series). Our main journal (Ann. Acad. Sci. Fenn. Math.) uses OPEN ACCESS and all the articles published since 1993 are completely freely available in our homepage. The books are available in the bookstore with whom we have an agreement. We try to keep our books available and if the recent books are outsold we take a new print.	0