DEVELOPMENT OF A DIGITAL PRESERVATION POLICY AS A BUSINESS OPORTUNITY FOR CONSULTING AND ARCHIVISTS

Jordi Serra Serra
Faculty of Library and Information Science
University of Barcelona

Today, thanks to research advances, digital preservation becomes a process increasingly affordable for any archive. Most of digital preservation challenges have already operational solutions, and it is expected that digital preservation market, especially in the cloud, grow significantly in coming years. Thus, and with cost decreasing, fewer archives will need to develop their own infrastructure for digital preservation, because they will found in the cloud the services that meet their needs (De la Vega, 2013, and Beagrie et al., 2014).

However here comes a new challenge. That's how a traditional archive without previous experience in electronic records management, and without any digital funds yet received, can discover which are his needs and risks, what services are best for it, and how to hire, use, control and maintain those services. If this archive wants achieve digital preservation successfully, the first thing it will need is to develop a comprehensive digital preservation policy, based on reliable models and standards, ensuring like this the commitment of its own institution in digital preservation project over time (Li & Banach, 2011, Sierman et al., 2013, and also Kulovits et al., 2013).

The University of Costa Rica (UCR)[1] is the main public university in Costa Rica. UCR has its origins in the University of Santo Tomás, founded in 1843, and was established in its present form in 1940 as the body responsible of Higher Level education in Costa Rica, through a constitutional decree giving autonomy to the University to transform society through education, research and social action. The UCR has six academic areas: Liberal Arts, Basic Sciences, Social Sciences, Agriculture and Food, Health and Engineering, that branch into 13 faculties. Each faculty is composed of different schools: in total there are 47 schools, 40 research centers and institutes, and 5 regional campuses. The UCR has a population of approximately 40,000 students, and has 5,000 teachers and professors, and 3,500 administrative staff members. Each year approximately 8,000 new students are admitted to UCR, and about 5,000 students graduate.[2]

According to regulations approved by the *Consejo Universitario* on 3 September 2008, the University of Costa Rica has an archival system leaded by a main archive, the *Archivo Universitario Rafael Obregón Loría* (AUROL), and composed by the central archives, the historical archives, office archives, specialized archives of all University departments, and by the appraisal commission, the *Comisión Universitaria de Selección y Eliminación de Documentos* (CUSED). The whole system is coordinated by the AUROL, which depends on the Rectory and exercises technical authority over all the archives of the University.

In 2012, the AUROL, together with the *Sección de Archivística de la Escuela de Historia de la Facultad de Ciencias Sociales de la Universidad de Costa Rica*, made contact with professors of the University of Barcelona, to give learning and awareness activities on electronic records management and digital preservation.[3] As a result of this actions, and at the request of the University of Costa Rica, on 2013 started a joint project, with the objective to create the Digital Archive of the University of Costa Rica. What we expose in this paper are the results of the first phase of this project.[4]

1. The need

Any organization with operating informations systems give rise immediately to the need for a digital archive (Beagrie et al., 2014). Although retention period being short, this conservation has always security requirements, and often exceeds the obsolescence span of most of file formats usual in the market. To people in charge of the custody of the records and data generated by this systems, and also for the organisation's archive, this general need splits in six concrete needs:

- To know the risks for records and the severity of each one (risks quantifying and measuring).
- To know which are the functionalities required to mitigate this risks (functional requirements).
- To know which market solutions meet this functional requirements, both in terms of software and services.
- Have an organizational committment, at the highest staff level, so it has the will to limit the risks for an extended and structural period of time.
- Have the minimal needed resources (financial, technical and human) to enable the use of the solutions that meet the functional requirements.
- Have the minimal control capability (internal procedural control) to use the solutions over records and data produced by operating information systems (what often requires a records management policy and recordkeeping system already implemented).

Most of this conditions were fulfilling in the University of Costa Rica, where the AUROL had been implementing a complete records management system (based on SISDOC tool), where the *Centro de Informática* (the computer center) was ready to make available the needed technological resources, and where the higher body of the University, the Rectory, had commissioned AUROL to start the actions facing the risks of digital preservation over time.

2. The project

The project to create the Digital Archive of the University of Costa Rica was started on June 2013 with four operational objectives:

- Evaluate and give methodological answer to digital preservation needs of the University of Costa Rica.
- Develop the institutional digital preservation policy for the University of Costa Rica.
- Identify the functional requirements to implement and launch a mid-long term digital archive.

- Define an submission agreement model (referring to OAIS terminology) for the ingest and preservation of records and data from UCR information systems.

This were the objectives of the **first phase** of the project which, within a short period of time, should provide the basis for the creation of the UCR Digital Archive, raising awareness, developing a digital preservation policy and defining the functional requirements of the technological solution. In a **second phase**, it should be established the main controller agent of the system, the *Comisión del Archivo Digital*, and created inside the AUROL the Digital Archive Unit, training all the assigned staff. At the same time, the *Centro de Informática* should complete the definition of the technological scenario through a feasibility and sizing analisys, and then selecting the technological solution. Finally, in a **third phase**, was planned carrying out the technological developments and launching the Digital Archive, activating at the same time with records producers the regular development of submission agreements.

Keep in mind that the first task that Rectory commissioned to the AUROL was to present a feasible and reasonable project to address a problem that, so far, had not yet been raised specifically. For this reason, the project scope was narrowed in two ways. First delaying its implementation in three consecutive phases, which guaranteed a rational periodization of investment and contemplated an intermediate-term for evaluating the actions and making possible adjustments in project plan. And second, modeling a system implementation based on records series related to information systems, which allowed to make profit from existing recordkeeping system inertia, and launch the Archive step by step, in line with a well-ordered growth of resources and technological investment.

The members of the project team were the management and technical team of the AUROL, an external collaborator expert in digital preservation acting as consultant (from PRESERVA research group of the University of Barcelona), the computer center of the UCR (*Centro de Informática*), the project development unit of the UCR (*Unidad de Gestión de Proyectos*), and finally the archival section of the UCR History School (*Sección de Archivística de la Escuela de Historia de la Universidad de Costa Rica*).

Working methodology was set using several models. Components and issues to be aware in the development of a digital preservation policy were obtained from PLATTER model (DPE *Repository Planning Checklist and Guidance*, 2008)[5], and estructured on two levels: responsibilities and commitments statement, and preservation strategy for processes, information objects an file formats used in ingest, preservation and dissemination processes. The functional requirements of the digital archive were validated with ISO 16363 model (*Audit and certification of trustworthy digital repositories*, 2011). And for the previous risks analysis, PLATTER was completed with the model used by the UK National Archives (*Risk Assessment Handbook*, 2011). Unfortunately the *SCAPE Catalogue of Preservation Policy Elements*, that will be probably an extremely useful tool in coming years, was not yet published and couldn't be used (Sierman et al., 2014).[6] For the whole project planning and definition of deliverables we don't use any specific model, but planning guidelines usual of records management consultancy.

With the aim to strengthen this relationship between digital preservation policy and records management system existing at the University of Costa Rica, a second phase was added to the project, to provide the criteria and instruments for melting traditional retention rules (appraisal schedules) with OAIS submission agreements, aiming to integrate the conservation of all records and data in a single preservation process map, regardless of its format and structure. To achieve this objective, we used a simple methodology of process analysis, inspired by Lean & Six Sigma and APS (Puig-Pey, Guiu & Agramunt, 2008) and developed in the framework of the *MGDIE Master* of the University of Barcelona.[7] This methodology identifies the information

elements of a process that has to be preserved by its evidential value, and helps setting its optimal preservation format (including not only file format but also structure and metadata).

The first step of the project had been given on 2012, with the completion of a training and awareness set of activities. For a week, and through different types of formal and less formal sessions (workshops, open discussions, role-playing), the critical challenges of digital preservation, as the solutions developed by the preservation community, had been raised. Teaching methodology used was that used by University of Barcelona for training in digital preservation at undergraduate or graduate level,[8] which emphasizes on making understandable and acquire the main concepts of digital preservation, to next train professionals to implement and maintain a digital preservation policy and control preservation services. But computer center staff couldn't attend the awareness sessions conducted in 2012, so during the working sessions of the consulting project on 2013 a second set of sessions was included into the work that was being done to develop the digital preservation policy. This allowed, among other benefits, to show synergies at infrastructure level beetwen the Digital Archive project and another project already underway at the University of Costa Rica, the creation of a private cloud computing infraestructure, called *Nube Académica Computacional* (Villalón & Loría, 2013).[9]

The implementation of this first phase of the project was divided into four steps:

2.1. Preliminary needs analysis

First, a survey were distributed to stakeholders to assess, on the one hand, risks for the digital continuity of UCR records, and moreover the will of the University in relationship with such continuity, expressed both in institutional mandates as in the vocation of main stakeholders (especially AUROL, *Centro de Informática* and *Unidad de Gestión de Proyectos*). Survey results were subsequently validated through interviews, with the aim of fixing not only the needs (risk assessment), but also the scope and characteristics of the archive to create (wills and mandates assessment). The result of this phase was gathered in the document **1A Risk and Context Assessment Report**.

2.2. Functional and technical definition

Second, and according to the results of the first step, a high-level definition of the Digital Archive was produced. This definition had two dimensions. First, functional dimension, in which the work processes of the Digital Archive were described, according to the OAIS model, and from them were derived the functional requirements. On the other hand the architectural dimension (components and integrations), where different scenarios were prepared, based on the records-producer systems map, and considering available open source solutions for each system component. These scenarios were assessed together with the Centro de Informática, selecting at the end the definitive scenario. All this work was collected in the document **1B Functional and technical model for Digital Archive.**

In this step, the main debate ranged around using the already available ECM tool of the AUROL, called SISDOC, as the only ingest way to Digital Archive system (forcing the other systems to interact with SISDOC), or develop integrations for each records-producer system. Finally we decided a hybrid scenario, with three possible ingest ways: business systems (automated ingest through integration with Digital Archive's Web service), legacy systems (bulk ingest with *ad hoc* original data processing, previous agreement with producer about processing costs) and office documents in local network (human-driven ingest using end user interface).

This scenario included a list of components to acquire or develop, and a proposal of available open source or market solutions for each component. This components included a Web service for SIP ingest and validation, an interface to manually create SIP from existing local network files, services for virus detection, format recognition, format normalization, XML containers signature and timestamp, a users management system (able to manage internal and external identities of several designated communities), and an interface for searching, viewing and requesting certified copies of documents.

Requirements definition was organized in three sections: functional requirements, technical requirements and integrations, and interoperability and standards. Functional requirements were defined according to the OAIS ingest, preservation and dissemination processes, while technical requirements and integrations were defined according to the technological scenario approved as definitive.

2.3. Responsibilities and processes definition

Third was developed the digital preservation policy. From an original draft, and following the conceptual elements of PLATTER model, the working group reviewed iteratively the draft until the consensus on a final document was reached. On July 18th 2013, the final version of de digital preservation policy was officially presented to Rectory, which undertook to approval and implementation. We explain the results in section 3 of this paper.

2.4. Pilot test of submission agreement development

In order to consolidate the full understanding by all agents of dynamics set by digital preservation policy, and show it clearer with practical demostration, the last step of the project was to develop a first submision agreement using the model contained in the policy, that in the context of this project was called ***Protocolo de ingreso y custodia***. The process chosen was "travel expenses management" (national or international traveling of the University of Costa Rica staff). The first action was the analysis of the process flow, identifying the records and data to be collected and preserved. The project team worked using the model of process analysis mentioned above. On the records and data substrate identified by the analysis was developed the pilot agreement,[10] at the level of records series contained inside the software used by UCR for human resources management (*Sistema de expediente único de personal*). Thus the pilot agreement set the commitments and deadlines of ingest process, the significant properties to preserve and the format for collecting and packaging the data. From this agreement were derived the modifications to do in the HCM software, that *Unidad de Gestión de Proyectos*, responsible for this system maintenance and development, added immediately to its development agenda.

3. The results

The results of the project, considered the first step for the creation of the UCR Digital Archive, were collected in several documents (see figure 1), the main of which was the **1C Institutional Digital Preservation Policy**. This policy was composed by the following elements.

First, the **policy** itself, understood mainly as an allocation of responsibilities. Policy was preceded by an initial policy statement, containing the mission, objectives, principles and revision and upgrade procedure. Next, the core of the policy included the composition and responsibilities of the Digital Archive, the control and management body (the *Comisión Institucional del Archivo Digital*), the scope of records and data to preserve,[11] the duties and committments of staff, the due integration of every UCR information system with the Digital Archive, the committment of archive's financial sustainibility, and finally the referral to succession and disaster prevention plans.

Second, the technical side of the policy, called **preservation strategy**, that was composed by the following specific plans:

- **Data plan**: detailed description of the information to preserve, and the structure and metadata of OAIS information packages (SIP, AIP and DIP).

- **Ingest plan**: description of ingest process, with all possible functional scenarios for ingest ways, depending on the internal or external origin of records and data. This plan fixed also the obligation, for each source of records or data, to develop and approve a mandatory submission agreement, and described its structure, development and approval process. The plan made also mandatory ensuring archival requirements compliance in every technological project undertaken by UCR, through active participation of Digital Archive staff in the development team, getting ready from the system start the submission agreement.

- **Preservation plan**: description of preservation planning and migration processes (as migration was adopted as preferent strategy). This plan fixed three complementary ways to detect obsolescence risks: subscription to external alert services, periodic scanning using format registries updates, and random testing of a significant sample of preserved formats. The plan included also an specification of records preservation conditions into the Digital Archive, according to significant properties set in the submission agreements and formats recognized as preferred in the policy appendix.

- **Dissemination plan**: identification of designated communities,[12] description of dissemination channels according to designated communities knowledge base, and procedures for searching and copy delivering of preserved records, with particular emphasis in authentification of copies. The designated communities were considered entities with variable features, so the plan included a procedure for monitoring and updating its knowledge base.

- **Technology plan**: description of general security and integrity requirements, service availability and continuity, scalability and infrastructure upgrade.

- **Continuity plan**: planned contingency measures in the areas of continued financial capability, continued staff knowledge base, institutional continuity (succession plan) and continuity in case of disaster.

Three appendices were added to the preservation strategy:

- The **schedule of periodic duties** for all agents involved in Digital Archive processes.
- The list of **mandatory activity logs** (audit trail) of the Digital Archive.
- The list of **preferred file formats** for ingest, preservation, copy delivering (by designated community), packages authentification and electronic signature.

The third element contained in the policy was the **submission agreement model** (*modelo de protocolo de ingreso y custodia*). This is probably the most particular contribution of this project, thus we will discuss it specifically.

The main problem identified in digital archives experiences analyzed to support the development of this project was how to ensure that records and data produced by the organization are effectively and regularly submitted to Digital Archive, in a context more suited to Information Governance than to digital preservation pure and hard. We look at the digital archive as one more element of one set of guarantees covering the full records lifecycle, that starts with a capture and use control by means of a records management system implemented and active. However, to many times digital preservation projects seem to be like islands detached from usual organization's work processes (and therefore from records and information management), so often information vital for the organization remains in the limbo of the border between the permanent archive, the records management systems and the "silo" systems.

If we consider that output control of a records management system is formalized by retention schedules, and input control of a digital archive system is formalized by submission agreements,[13] it seemed evident that these two instruments became the linkage point between both systems, and being connected, or even merged, both instruments would allow us to increase substantially the level of control over the preservation of records produced by UCR information systems. Considering the digital archive as a natural extension of records management system would do the archive less vulnerable to institutional continuity and sustainability risks, and in the opposite direction it would add value to the records management system, increasing its service offer with an element of great impact, as long term preservation is.

But here rises a second problem. Records management works well with aggregations, but digital preservation requires, in addition, to work at file and bitstream level, thus finally information is detached from source systems and gets an independent life (and often independent structure) for archival purposes. It was necessary, therefore, to reach an adequate level of detail to identify exactly what records and data must be kept from information systems. Thus we made use here of the process oriented records management systems analysis methodology to which we have previously referred, which allows us to build significant relationships between records series and information systems through the development of evidence matrix (records or data), convertible to document types for records management purposes, and to SIP's for preservation purposes (Serra 2006, 2010, 2011, 2013 and 2014).[14]

The development of a UCR submission agreement begins, therefore, with the identification of records series contained in one or several information systems (guided by appraisal priorities scheduled by AUROL), and continue with the analysis of this series, the building of the evidence matrix and its normalization. Then the ingest parameters are defined: ingest channel, supported ingest formats, mandatory metadata, data checks to be performed, and eventual file formats normalization parameters. For each document type (or set of types) associated with an specific format or formats family, the significant properties to be preserved are defined, in the scope of appearance, behaviour and authenticity. Next are fixed the deadlines for records and data preservation, not only the whole deadline, but also specific deadlines for each kind of significant property (appearance, behaviour and authenticity).[15] The following step is to fix the rights that

will be transferred with the ingest into the Digital Archive, in the scopes of dissemination and access, use and copy, and changes that could be made for preservation purposes. Finally, the last section contains a list of commitments between the two parties that sign the agreement: the producer or custodian of documents (*instancia transferente o cedente*), and the Digital Archive. Formalizing the agreement is especially important when there are costs of development, adaptation and/or data processing rising from agreement implementation.

The submission agreement model was completed with another document: **Guidelines for evidence identification**. This document contained the basic guidelines and the data model to build evidence matrix, define normalization parameters and transform sets of evidences into documents types and SIP.

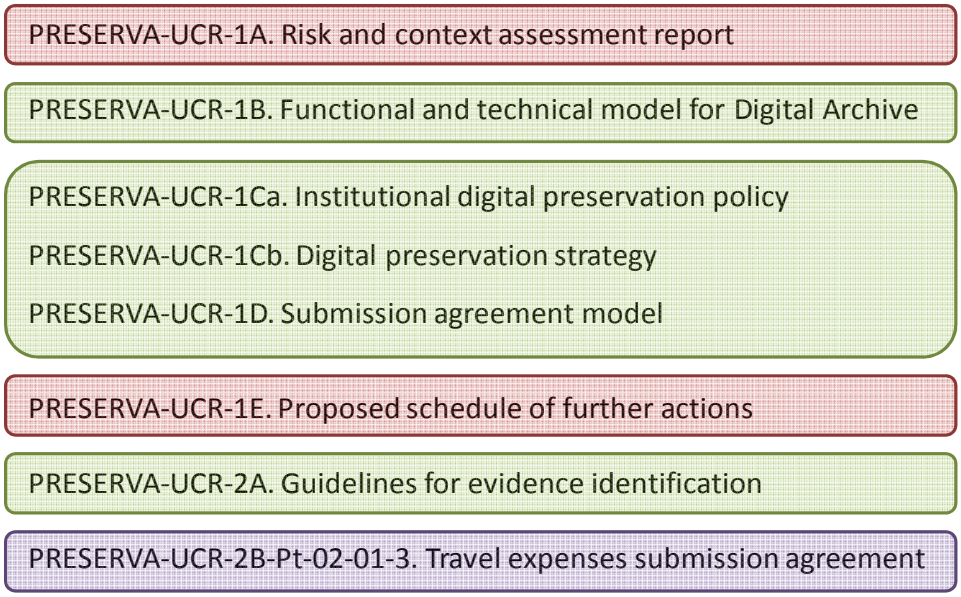

Figure 1: Project deliverables

In order to promote the continuity of the project and take advantage of the first phase inertia, a **Proposed schedule of further actions** (1E) was included as a roadmap in the deliverables of the project, trying to establish a timetable for, once approved the policy, continue with underway implementation of the Digital Archive. The actions proposed for the second and third phases were the following:

1. Start-up the governing body (designation of members, announcements and workplan)

2. Start-up of the organizational structure (creation of the Digital Archive unit within the AUROL, recruitment and training of staff)

3. Start-up of the technological scenario (sizing and feasibility analysis, architecture definition, development and integrations)

The current status of the project is the following. The process for policy approval by UCR Rectory is underway, and this approval will start the sequence of actions that need specific investment. The three leading agents of the project are in parallel implementing the remaining actions. The AUROL has started the creation of the Digital Archive unit, and the training of assigned staff. Meanwhile, the *Centro de Informática* is working in the synergies between the

Digital Archive project and the OAI institutional repository project of the UCR,[16] aiming to share software, as with the project to replace the UCR data center with a new powerful infrastructure, the *Nube académica computacional* (NAC), to which we referred above. Also, since UCR is strongly committed to open source software, and much of digital preservation software is open source, the alignment between development teams of *Centro de Informática* and the aim of Digital Archive project is complete. Finally, the *Unidad de Gestión de Proyectos*, which establishes the link between functional needs of UCR bodies and technological developments, as vital partner for success in submission agreement's development and compliance, has begun to include the digital preservation policy requirements in the development of all technological projects.

4. The business model

As it was a project started from scratch, in an institution without precedents in digital preservation and with a complete involvement of all stakeholders, this project allowed to test a set of worklines that, according to current trends in digital preservation, become potential business lines. We will try to enumerate this lines.

The level of complexity and dependencies between different functional components of a digital archive makes today essential a close cooperation between archives, developpers and institutions, so it is hard to imagine digital preservation as a project individually affordable by a single archive (Gracio et al., 2013). This complexity rises specially by higher technological nature of digital preservation. Solutions development, as services delivery, require advanced technological skills not included in standard archivist's curriculum. The evolution of research and development in digital preservation, with a cross-cutting approach including libraries and archives, and emphasizing in technological issues, has tended to untie digital preservation from traditional archival processing, giving to some extent its own entity.

We do not value this transformation negatively, since it is the result of a progressive interdisciplinarity highly positive for the development of archival discipline, which has allowed an spectacular progress if we look at which was the scene one decade ago. But at the same time, for institutional archives that gets used to run the custody process with its own resources, and which consider start receiving and keeping digital funds, this specialization implies that digital preservation is difficult to assume only with own resources, and the need to resort to highly specialized external providers arises strongly. Thereby digital preservation is becoming one of archival processes more clearly candidate for outsourcing, both in terms of infrastructure and software (SaaS) as service delivery, specially when it includes data processing (eg migration). In Latin or Mediterranean areas, where the archive is still few used to outsource and share solutions, this issue means a considerable cultural shock.

At the same time, this issue affects archivist's profile, specially of institutional archivist, a professional who often covers a vast area of expertise ranging from implementation of records management systems to custody and dissemination of historical funds.[17] This multi-purpose archivist warns that must deal adequately with digital funds, but at the same time it is not able to reach the level of specialization that digital preservation solutions require, nor has the needed knowledge to find and hire such solutions (*Staffing for effective digital preservation*, 2013). Every delay make worse the situation because, with digital preservation progress, the distance between the reality of the archive (the starting point) and the model to implement (the point of arrival) increases. This situation offers interesting business opportunities for digital preservation professionals, especially in the field of consulting, having as customers the archivists of every kind of organizations.

We will explore first the opportunities for **consulting**, especially those related with initial change management. In this sense, a first business line is **training**, which can cover from initial awareness of preservation needs to staff training or specialization in digital preservation tasks and tools. The curriculum of the archivist specialised in digital preservation, and the different levels of such training, have already been detailed in projects like DigCCurr (Kulovits et al., 2013).[18] Considering that archival institutions often manage also networks of archives, and that implementation of a digital archive system can be deployed in such networks waterfall, this training can also take the nature of training of trainers, as recently made Library of Congress through a joint program with the NDSA Outreach Working Group of the National Digital Stewardship Alliance, the *Digital Preservation Outreach and Education program* (DPOE).[19]

A second business line is the development of organizational and technical guidelines for the implementation of a digital archive, that's the **development of digital preservation policies and strategies**.[20] Several analysis of published preservation policies, as the compilation made by SCAPE project[21] or the surveys of Kristen Snawder and later Madeline Sheldon (Sheldon, 2013), revealed that, despite the large number of digital funds managed by archives and libraries, still few institutions have a formalized preservation policy, and in many cases this policy don't cover all the aspects of the preservation processes (Li & Banach, 2011). The development of a preservation policy is also a complex process in which institutional mandates or local models about how to approach preservation do not allow to consider the replication of a generic model. Thus, the progressive availability of standardized methodological references as the SCAPE Catalog (Sierman et al, 2014), even though they mean a major breakthrough in line to promote interoperability between policies, do not replace the need for a detailed analysis and development tailored to each institution.

These policies can be fixed in an institutional document, or can also be deployed progressively by sets of rules, preferably in the form of submission agreements or migration plans. The development of these submission agreements, prior **analysis and normalization of the evidence** generated by work processes and information systems, is a third business line, directly stemmed with records management systems design and process reengineering. Finally, after identifying digital archive needs (functional requirements), a fourth business line is to **provide advice for locating and hiring services**, for solutions configuration and for integrations development, once within the scope of technology consulting.[22]

We should add another business line to the possibilities we exposed above, that's **auditing**. This audit may be prior to implementation of a digital archive system, in order to assess risks and needs, and then define the scope of the project (Vermaaten et al., 2012). But may be also performed later, once implemented the system, related to the required periodic updating of preservation policies (Sierman et al., 2013), as well as to the **certification** of repositories with ISO 16363 standard. This is a market that will grow with the forthcoming publication of ISO 16919 standard for bodies providing audit and certification of digital repositories. This business line offers numerous synergies with records management, which is also working in certification processes with the development of the ISO 30300 family of standards.

A second set of business lines are those related to **services delivery**. In this regard, as well as integration and technological development services (on premise) and trusted cloud services (SaaS, and specifically AaaS or *Archive as a Service*), we would remark services of technology obsolescence monitoring and alert, such as SCOUT,[23] or services for normalization, migration or display of obsolete file formats, as the emulation framework bwFLA, which significantly uses to define itself the term *Emulation as a Service* (EaaS).[24] This area has grown significantly in recent years, both in the field of cooperative solutions in an open software environment (eg MetaArchive with LOCKSS or DuraCloud with Archivematica), as in those of comprehensive proprietary solutions, moving ever deeper into the addition of digital preservation functionalities,

with spearheads as Preservica.[25] It will be probably from the hand of *Enterprise Information Archiving* (EAI) that digital preservation functionalities will spread decisively across propietary software market, taking advantage of needed growth of these solutions as requirement for eDiscovery, and probably the preservation of accessibility is going to be a mandatory maturity requirement of EAI solutions.[26]

The role of institutional archivist becomes transformed with preservation outsourcing. Even when archive system keeps on premise, much of the management is performed by a technology unit (specialized or not in preservation), and direct activity of the archivist over the fund tends to decrease. This must be read not as a threat but as an opportunity, because, while in today custody the archivist becomes increasingly customer and less actor, at the same time it is empowered in the role of funds ingest negotiation and control, especially for extended use that can be given to appraisal schedules, as link between producers, preservation needs and records management. Archivists are consumers of a set of tools and services at their disposal in order to meet the need for long-term preservation. Given that these services make digital preservation an increasingly automated process (see for example the approach of Graf & Gordea, 2013, or the will to make preservation policies interoperables in Sierman et al., 2013), archival activity is increasingly focused in early stages of lifecycle, those traditionally related to records management.

In this context, is vital to develop mechanisms to link records producer systems and records management system with digital archive system, a subject that seems not yet enough worked.[27] For example, Preservica has developed a connector to import content from DSpace, and uses CMIS for content dissemination via browser, but don't advertise yet connectors with common ECM systems. This is probably one of the most important challenges for immediate future of digital preservation, closing a gap long time identified (Korb & Strodl, 2010). Beyond ensure proper management of records in its active phase, a records management system should identify what evidence (documents or data) must be preserved, and also assess whether its original form makes possible or sustainable such preservation, proposing actions for normalization according to digital archive guidelines available in the organization. Keep in mind that digital preservation itself does not improve record's quality,[28] thereby create a high-quality digital archive needs a previous and detailed work in capture phase. This work is not only normalization, it is previous to records creation, identifying with process reengineering what evidence should be created, and in what formats and/or structures, as it make easier the subsequent ingest into the digital archive. Thus, a digital preservation policy becomes a natural extension of the records management policy of an organization, sharing the objectives of quality, reliability and efficiency of records and data, according to the absolutely integrated perspective that *Information Governance* promotes.

5. Bibliografía

A guide to distributed digital preservation (2010). Katherine Skinner & Matt Schultz (ed.). Atlanta: Educopia Institute.<http://www.metaarchive.org/sites/metaarchive.org/files/GDDP_Educopia.pdf>. [25/08/2014].

Audit and certification of trustworthy digital repositories (CCSDS 652.0-M-1) (2011). Recommended Practice, Issue 1. Washington, DC: Consultative Committee for Space Data Systems. También referenciada como *ISO 16363:2012. Space data and information transfer systems - Audit and certification of trustworthy digital repositories*. <http://public.ccsds.org/publications/archive/652x0m1.pdf>.

[18/06/2013].
BEAGRIE, Neil; CHARLESWORTH, Andrew &MILLER, Paul (2014). *The National Archives Guidance on cloud storage and digital preservation: how cloud storage can address the needs of public archives in the UK*. 1st edition. The National Archives.
DE LA VEGA, Ricard. "Preservació digital al núvol" (2013). *Item: revista de biblioteconomia i documentació*, vol. 10, núm. 57.
<http://www.recercat.net/bitstream/handle/2072/219218/preservacio-al-nuvol-rdelavega-item-57.pdf>.
[16/07/2014].
Digital preservation maturity model (2014). White paper. Preservica.
<http://preservica.com/resource/praesent-ante-stiam-white-paper/>.
[25/08/2014].
Digital Preservation Policies: Guidance for archives(2011). The National Archives.
<http://www.nationalarchives.gov.uk/documents/information-management/digital-preservation-policies-guidance-draft-v4.2.pdf>.
[29/06/2013].
Digital preservation policy(2013). National Archives of Australia. <http://www.naa.gov.au/about-us/organisation/accountability/operations-and-preservation/digital-preservation-policy.aspx>.
[18/06/2013].
Digital preservation policy(2013). State Library of Victoria. <http://www.slv.vic.gov.au/about-us/plans-policies/digital-preservation-policy>. [18/06/2013].
Digital preservation policy (2013). The National Archives.
<http://www.nationalarchives.gov.uk/information-management/projects-and-work/digital-preservation-policy.htm>.
[18/06/2013].
Digital preservation policy(2013). 4th ed. National Library of Australia.
<http://www.nla.gov.au/policy-and-planning/digital-preservation-policy>.
[18/06/2013].
Digital preservation program(2013). J. Willard Marriott Library, University of Utah.
<http://www.lib.utah.edu/collections/digital/digital-preservation.php>.
[18/06/2013].
DOLLAR, Charles M. & Ashley, Lori J. (2013). *Assessing digital preservation capability using a maturity model process improvement approach*. White paper.
<https://docs.google.com/file/d/0BwbqtwrvKHokR3g0RVR4bmNIWjg/edit?pli=1>.
[25/08/2014].
DOWNS, Robert R.&CHEN, Robert S. (2012). “Testing the International Standard for Audit and Certification of Trustworthy Digital Repositories”. In: *Digital Preservation 2012*. Arlington, Virginia.
DPE Repository Planning Checklist and Guidance DPED 3.2. (2008). Digital Preservation Europe.
<http://www.digitalpreservationeurope.eu/publications/reports/Repository_Planning_Checklist_and_Guidance.pdf>.
[18/06/2013].

Establishing a Digital Preservation Policy (2013). JISC Digital Media.
<http://www.jiscdigitalmedia.ac.uk/guide/establishing-a-digital-preservation-policy>.
[18/06/2013].
GBAJE, E. S. &MOHAMMED, Z. (2013). “Digital preservation policy in National Information Centres in Nigeria”. *Electronic Library*, vol. 4, núm. 31.
GRACIO, José Carlos Abbud; FADEL, Bárbara&VALENTIM, Marta Lígia Pomim (2013). “Preservação digital nas instituições de ensino superior: aspectos organizacionais, legais e técnicos”. *Perspectivas em Ciência da Informação*, vol. 18, núm. 3.<http://dx.doi.org/10.1590/S1413-99362013000300008>.
[30/06/2014].

GRAF, Roman &GORDEA, Sergiu (2013). "A risk analysis of file formats for preservation planning". In: *Proceedings of the 10th International Conference on Preservation of Digital Objects, IPRES 2013*.
<http://purl.pt/24107/1/iPres2013_PDF/A%20Risk%20Analysis%20of%20File%20Formats%20for%20Preservation%20Planning.pdf>.
[30/06/2014].
HITCHCOCK, Steve &TARRANT, David (2011). "Characterising and preserving digital repositories: file format profiles". *Ariadne*, núm. 66.
HOFMAN, Hans (2014). “The revision of ISO 15489: an overview”. In: *40th Annual ICA/SIO meeting "Digital preservation"*. Noordwijk, The Netherlands.
ICPSR Digital Preservation Policy Framework (2013). Inter-university Consortium for Political and Social Research.
<http://www.icpsr.umich.edu/icpsrweb/content/datamanagement/preservation/policies/dpp-framework.html>.
[18/06/2013].
ISO/FDIS 16919. Space data and information transfer systems - Requirements for bodies providing audit and certification of candidate trustworthy digital repositories.
KNIGHT, Gareth &PENNOCK, Maureen (2009). "Data without meaning: establishing the significant properties of digital research". *The International Journal of Digital Curation*, vol. 4, núm. 1.
KORB, Joachim &STRODL, Stephan (2010). "Digital preservation for enterprise content: a gap-analysis between ECM and OAIS." In: *Proceedings of the 7th International Conference on Preservation of Digital Objects, Österreichische Computer Gesellschaft, IPRES 2010*.
KULOVITS, Hannes; BECKER, Christoph &RAUBER, Andreas (2013). "Roles and responsibilities in digital preservation decision making". In: Luciana Duranti & Elisabeth Shaffer (eds.). *Proceedings of The memory of the World in the digital age: digitization and preservation*. Vancouver.
LI, Yuan &BANACH, Meghan (2011). "Institutional repositories and digital preservation: assessing current practices at research libraries". *D-Lib Magazine*, vol. 17, núm. 5/6.
<http://www.dlib.org/dlib/may11/yuanli/05yuanli.print.html>.
[25/08/2014].
PHILLIPS, Megan; BAILEY, Jefferson; GOETHALS, Andrea & OWENS, Trevor (2013). *The NDSA Levels of digital preservation: an explanation and uses*. National Digital Stewardship Alliance.
<http://www.digitalpreservation.gov/ndsa/working_groups/documents/NDSA_Levels_Archiving_2013.pdf>.
[25/08/2014].
PROENCA, Diogo (2013). *The role of risk analysis to support cost models for digital preservation*. In: Archiving 2013: Final program and proceedings.
PUIG-PEY, Antoni; GUIU, Pere &AGRAMUNT, Helena (2008). *Circuits administratius: Disseny i millora*. València: Associació d´Arxivers i Gestors de Documents Valencians - Tiravol.

Reference Model for an Open Archival Information System (CCSDS 650.0-M-2) (2012). Recommended Practice, Issue 2. Washington, DC: Consultative Committee for Space Data Systems.También referenciada como*ISO 14721:2012 Space data and information transfer systems - Open archival information system (OAIS) - Reference model*.
<http://public.ccsds.org/publications/archive/650x0m2.pdf>.
[18/06/2013].
Risk Assessment Handbook(2011). The National Archives.
<http://www.nationalarchives.gov.uk/documents/information-management/Risk-Assessment-Handbook.pdf>.
[29/06/2013].
SERRA, Jordi (2004). "La firma electrónica y el archivo digital". In: Primeres Jornades de Signatura Electrònica. Agència Catalana de Certificació (CATCert).
<http://hdl.handle.net/10760/5659>.

[25/08/2014].
SERRA, Jordi (2006). "La gestió per expedients i la preservació de la documentació electrònica: el sistema BAULA". *Lligall*, núm. 25.
<http://hdl.handle.net/10760/9360>.
[25/08/2014].
SERRA, Jordi (2010). "La inclusión de requisitos archivísticos en proyectos de tramitación electrónica: el mapa de series y sus formas de implantación". *Revista d'arxius*, núm 8.
<http://hdl.handle.net/10760/15326>.
[25/08/2014].
SERRA, Jordi (2011). "B.1 Diseño del sistema de gestión documental: la construcción del mapa de series documentales y la definición de las políticas del sistema". Material docente del *Máster de Gestión Documental e Información en las Empresas*. Universidad de Barcelona.
SERRA, Jordi (2013). "Una interpretación metodológica de la norma ISO 15489 para la implantación de un sistema de gestión de documentos". In: *Jornadas Ibéricas de Arquivos Municipais; Políticas, Sistemas e Instrumentos*. Lisboa.
SERRA, Jordi (2014). *Análisis y mejora de procesos orientados a la gestión documental electrónica*. Material docente. Pamplona: Asociación de Archiveros de Navarra.
<http://hdl.handle.net/10760/22880>.
[25/08/2014].
SHELDON, Madeline (2013). *Analysis of current digital preservation policies: archives, libraries and museums*. Library of Congress.
<http://www.digitalpreservation.gov/documents/Analysis%20of%20Current%20Digital%20Preservation%20Policies.pdf>.
[25/08/2014].
SIERMAN, Barbara; JONES, Catherine; BECHHOFER, Sean; ELSTROM, Gry (2013). "Preservation policy levels in SCAPE". In: Proceedings of the 10th International Conference on Preservation of Digital Objects, IPRES 2013.
<http://purl.pt/24107/1/iPres2013_PDF/Preservation%20Policy%20Levels%20in%20SCAPE.pdf>.
[30/06/2014].
SIERMAN, Barbara; JONES, Catherine & ELSTROM, Gry (2014). Catalogue of preservation policy elements. SCAPE Project.
<http://www.scape-project.eu/wp-content/uploads/2014/02/SCAPE_D13.2_KB_V1.0.pdf>.
[25/08/2014].
Staffing for effective digital preservation (2013). National Digital Stewardship Alliance.
<http://digitalpreservation.gov/ndsa/documents/NDSA-Staffing-Survey-Report-Final122013.pdf>.
[25/08/2014].
active preservation of digital information (2014). White paper. Preservica.
<http://preservica.com/download/4050>.
[25/08/2014].

VERMAATEN, Sally; LAVOIE, Brian &CAPLAN, Priscilla (2012). "Identifying threats to successful digital preservation: the SPOT model for risk assessment". *D-Lib Magazine*, vol. 18, núm. 9/10.
<http://www.dlib.org/dlib/september12/vermaaten/09vermaaten.print.html>.
[25/08/2014].
VILLALÓN, Ricardo &LORÍA, Luis (2013). "Proyecto de nube académica computacional para las Universidades públicas de Costa Rica". In: *Encuentro Temático TICAL, RedCLARA*.
<http://www.redclara.net/news/DV/Tical/NAC_UCR_version2.6.pdf>.
[25/08/2014].

Notes

[1] See http://www.ucr.ac.cr/.

[2] See facts and figures of UCR at http://www.ucr.ac.cr/acerca-u/ucr-en-cifras.html. See also http://www.shanghairanking.com/World-University-Rankings/University-of-Costa-Rica.html.
[3] Person in charge of collaboration, first as trainer and later as consultant, was the author of this paper, Jordi Serra Serra, Professor at the Faculty of Library and Information Sciences of the University of Barcelona and member of the research group PRESERVA, also of the University of Barcelona.
[4] I would like thank the excellent collaboration in the project of all UCR staff, and especially Raquel Umaña, Director of the AUROL and promoter of the project, Luís Loria, of the UCR computer center, Michel Angulo and Pedro Céspedes, of the projects management unit, and María Teresa Bermúdez, professor of the School of History of the University of Costa Rica.
[5] More information about the method at http://www.digitalpreservationeurope.eu/platter/.
[6] See http://wiki.opf-labs.org/display/SP/Catalogue+of+Preservation+Policy+Elements.
[7] *Master's Degree in Records and Information Management in Business* (MGDIE), of the University of Barcelona, where we work with a process-oriented methodology for records management systems definition, linked to digital preservation. More information at http://www.mgdie.net/en/.
[8]The *Master's Degree in Digital Content Management* of the University of Barcelona (http://www.ub.edu/biblio/master-de-gestio-de-continguts-digitals/presentacio-2.html), now in its tenth edition, was pioneer in Spain to include content and digital preservation practices. From teachers and students of Master's Degree comes the staff of the research group PRESERVA (http://bd.ub.edu/preservadigital/), which among other results produces a collection of tutorials of digital preservation tools, basically for teaching support purposes.
[9] For more information about NAC project see http://ci.ucr.ac.cr/node/319.
[10] Submission agreement number Pt-02-01-3: Travel expenses.
[11] A wide-ranging content because it included administrative, teaching and research records.
[12] At a first phase were identified six communities: teachers and researchers of UCR, management and services staff of UCR, students and alumni of UCR, Public Administration staff of Costa Rica, citizenship of Costa Rica, and people with accessibility issues by disability.
[13] ISO 14721:2012, section 2.3.2.
[14] All regulated process can be reduced to a matrix of variables, so that documents do not exist as such, but only as temporal data aggregations that are part of a unique data set (a unique record for each process), working directly at file level (although some systems simulate sequences of documents, as occurs in SAP with infotypes). For a development of this concept, see Puig-Pey, Guiu & Agramunt, 2008, and also Serra, 2006, 2010, 2011, 2013 and 2014. The draft of the new version of ISO 15489 is going clearly to this line, devoting the section 5 to process analysis as a source of records requirements identification (Hofman, 2014).
[15] The meaning of this measure lies in the fact that the complexity, and therefore cost, of retaining indefinitely certain records functionalities can be very high, and hard to justify if the value or use of the records decreases as time goes by. For example, preserving the authenticity of an electronic signature using the long-term signature format XAdES-A involves a cost of validation and periodic re-signature, which is justified only when the document has a potential legal use. When the document is kept only for its historical value, the signature ceases to be a significant property (in the sense of relevant to the original significance and use of the document), and therefore the cost of updating the signature becomes superfluous (Serra, 2004, and also Knight & Pennock, 2009).

[16] The institutional repository of the University of Costa Rica, called Kérwá, is an open digital archive (OAI-PMH) that stores, preserves and disseminates scientific and scholarly publications of the University of Costa Rica. Kérwá offers open access to books, technical and working papers, articles, theses, audio and video records, and research reports. This repository uses DSpace as a platform, began operating in mid-2010, and was declared in 2013 Institutional Repository. To access Kérwá go to: http://www.kerwa.ucr.ac.cr/.
[17] In Mediterranean and Latin world, the fact that the distinction between records manager and archivist profiles is not formalized has the consequence to overwhelm the institutional archivist with a range of skills often beyond the resources he can get from its institution.
[18] See http://ils.unc.edu/digccurr/index.html.
[19] See http://digitalpreservation.gov/education/.
[20] We can see an example of business opportunities raised by the development of strategic documents in Spain in the Act 15/1999, which regulates the protection of intimacy and personal data. The article 88 of the Act's regulation (Royal Decree 1720/2007) requires the organizations to have a corporate document called "Security document", where all technical and organizational measures in line with the current safety regulations are collected. As the Act's regulation was approved, the preparation of this document offered

immediately a significant business opportunity for independent security and data protection consulting companies operating in Spain (see *Real Decreto 1720/2007, de 21 de diciembre, por el que se aprueba el Reglamento de desarrollo de la Ley Orgánica 15/1999, de 13 de diciembre, de protección de datos de carácter personal*).

[21] See http://wiki.opf-labs.org/display/SP/Published+Preservation+Policies.

[22] The uncertainty in the allocation of responsibilities keeping digital funds increases specialization, and therefore diversification, of digital preservation solutions, depending on the kind of archive and documents to preserve (Web archives, Data archives, DAM, etc.). Given this increasing specialization, the challenge for organizations in the future will be to have a single integrated policy and a single control system for a set of different operating archive solutions (Hitchcock & Tarrant, 2011).

[23] See http://scout.scape.keep.pt/web/.

[24] See http://bw-fla.uni-freiburg.de/.

[25] See http://preservica.com/.

[26] The *Enterprise Information Archiving* (EAI) Gartner's Magic Quadrant 2013 emphasizes both the preference of organizations for cloud archiving for evidential and historical purposes, as the growing multiplicity of kinds of preserved data (email, file systems, Sharepoint contents, instant messaging and social networks, etc.). In this sense is also interesting the Preservica white paper that identifies the maturity levels that allow talking about digital preservation in Archiving (*Digital preservation maturity model*, 2014).

[27] An evidence of this is the evolution of the two reference standards for records management and digital preservation, ISO 15489 and ISO 14721, which have not been clearly stated as a continuous sequence bound by the disposal-ingest process. While in its first version (Blue Book) ISO 14721 standard made no mention of the ISO 15489 standard, in its second version (Magenta Book) includes a mention, particularly when describing the functional entity of system administration (section 4.1.1.5), and specifically the process of negotiating the submission agreement, stating that in the definition of the submission agreement may be applicable, among others, the ISO 15489-1 and ISO 15489-2 standards. Meanwhile, the draft of the new version of ISO 15489 standard strengthens the link with digital preservation processes, especially in sections *7.9 Maintaining usability* and *7.10 Migrating and converting records beetwen systems* (Hofman, 2014).

[28] We refer to strictly archival or diplomatic quality of records, although we could add quality also in terms of efficiency.