

English Version	Authors: Marie Laperdrix, H�el�ene Guichard-Spica
Number of pages : 6	Archives nationales, France

Records management and Digital archiving as a “self-portrait”: A way to sketch the “*Digital National Archives*”

Overall context of the National Archives

The “Pierrefitte project” is a huge undertaking, as it is linked to the restructuring plan of the institution (a new organizational structure), a new building, and a new information system. Very quickly, the need to archive the documents produced by the institution itself became evident, if only for the removal of the teams from Paris and Fontainebleau to Pierrefitte-sur-Seine. This logistical obligation was combined with the need for all the scientific services in charge of collections to take on board the question of digital archiving, in so doing responding to one of the strategic themes in the National Archives’ new scientific, cultural, and educational programme¹. Even beyond adapting the institution to the new realities of collection, records management and digital archiving have been defined as two new means of participating in the restructuring plan of the institution. On the one hand, digital archiving has been a reality for the National Archives since 1983. The National Archives electronic archiving system CONSTANCE² will be celebrating its 30th birthday next year. Since it began, it has made it possible to collect over 15 TB of native digital data from more than 15 Ministries for which the National Archives is responsible. This old collection dating back to the 1950s is constantly being augmented via ever-more-active archival collection by the archivists within the Ministries. This practice is fully integrated into the institution’s process for collecting contemporary archives. On the other hand is a new mission created in 2009 for managing the National Archives’ own archives (Records Management unit for standard and intermediate archives) and to prepare the removal of the agents to the new Pierrefitte-sur-Seine building on the outskirts of Paris.

The merger of the Digital archiving and Records management disciplines right from the start, originally arising out of rationalization of resources and an obligation of service towards central administration, turned into a fully-fledged project for the evolution of the institution. By 2014–2015, the outcome of this is set to be the first milestones of the institution’s new electronic archiving system. After “Paper Pierrefitte”, the second phase of the restructuring plan is being embodied in a “Digital Pierrefitte” that is just as ambitious and complex to put in place. In this sense, the two functions of Records management and digital archiving are playing a part in constructing the new identity of the National Archives of France.

1. Step 1: Archiving Identity: building internal processes and methodology. The experiment of “Archives Playland”

In a French context, the records management and digital archiving disciplines are really quite distinct, all the more so at the National Archives, which only handles the final archiving for the central administration, since the State’s intermediate archiving is dealt with by the archivists in the Ministries, attached to the Archives of France’s Interministerial Service under the Ministry of Culture and thus, separate from the National Archives. The digital archives skills are now serving the institution’s records manager, with a view to building up the memory of the institution. Without great vigilance on our part, the mixing of each of their skills risks verging on schizophrenia!

1.1. Objectives

Although procedures did already exist within the institution in relation to electronic archiving for the ministries, everything still remained to be done. When the National Archives’ new building was about to see the light of day and the institution was starting a profound renewal, a Record Manager (Records management and archiving) was created in Summer 2009. The aim was to manage archiving from all the departments, with the prospect of the forthcoming removal, and implementing new document management processes.

The objective was to enable each person to take control of their own paper and electronic production, while

1 See National Archives website: www.archivesnationales.culture.gouv.fr/chan/chan/psce.html

2 CONSTANCE stands for *CON*servation et *ST*ockage des *Archives Nouvelles Cr e es par l’ *lectronique in French.

improving their technical and archival management of the electronic collections. And how better to do this than to apply to oneself the recommendations required for performing good archiving: to properly archive electronic records, they must have been properly managed right from the moment they were created. The renewal of the departments and working methods had just been started. In order to succeed with proper archiving, it was imperative to have put in place lasting records management procedures, while associating them, right from their preparation, with the prospect of digital archiving.

1.2. Records management at the National Archives: organization and tools

The task was both simple and immense, for there was no formalization of the internal working procedures for creating and organizing office computing work files.

The first step had been to assess document production and to put in place classification plans and retention schedules.

From Autumn 2009, visits were made to the departments, to analyse in each of them the practices in the discipline and the document production. Needed in preparation for the removal of part of the agents into the new building in Pierrefitte-sur-Seine, the archiving of paper files was used as a basis for formalizing department by department the classification plans and the first retention schedules, which allowed paper and electronic documents to be handled without differentiation.

From their creation, the digital documents needed to be organized in a reliable way by standardized naming. This naming convention had to be aimed at easy assessment of the production, by identifying the status of the documents produced:

- whether it is a version of work in progress, a rough, or a final version,
- the identification of the producer,
- the nature of the documents, whether or not, they represent a commitment for the institution.

Therefore a very simple naming convention was proposed: a standardized identifier for the department (the producer), a year, a name relevant to the contents of the file, and a version number. On working versions, more individual identifiers and specific annotations can be added to this. A number of versioning rules were also adopted.

The second step was to analyse the use of messaging systems and the types of documents contained in the messages. This is a quick, easy way of producing or conveying electronic administrative documents whose value may be binding, probative, or historic for the institution. A legal analysis was carried out in parallel, in order to identify all the problems associated with collecting, conserving, and above all distributing the information contained in messaging services.

To inform people about new practices and put them in place, in addition to information meetings, a number of tools and training sessions were produced:

- a handbook for paper archiving and a handbook for digital archiving³
- a handbook for managing electronic messaging systems — the first part of a future handbook that will go as far as archiving messaging systems.
- training sessions on file naming, on what electronic archives are, on managing electronic messaging systems were carried out from 2010–2012, as well as input to new employee induction.

2012 is a key year for the National Archives: the one where the new organization of the departments was put in place and the agents were moved. From the work that has been carried out so far, functional general reference model for archiving were distributed early this year: the overlapping between the two organizations had to be able to be achieved without the departments referring to different work documents and without leaving too great a scope for interpretation. Therefore each work function has been detailed, in action and sub-actions. Then, for each action, a document analysis has been suggested, in accordance with the documents produced in all the departments. when this was necessary, durations of administrative usefulness and final fate have been suggested for each file or document. This functional management table has passed the test of the organizational change.

All of these provisions, relayed via the creation of classification plans within the tree structures of the new servers, make it possible to put the whole institution, whatever the geographical site, whatever discipline is being exercised, into the same records management process.

It is thanks to this strategy of starting with a fine document analysis that it is now possible to an electronic

3 You could download these handbooks on the website of the National Archives in the “Publications” section> “Publications professionnelles”.

document and records management system (EDRMS), suitable for the internal needs, that will be interfaced with the existing tools: the new Archives information system (SIA), the project for an institutional intranet, and the new electronic archiving platform at the design stage.

1.3 To help the National Archives to ensure the lasting conservation of the State's electronic output: the institution's "sandbox".

This work, carried out primarily for the needs of internal management, has been the means to put new management procedures and legal and technical analyses to the test, and above all, to renew deliberation around the archival description of unstructured electronic collections.

To carrying out this project, two major constraints had to be taken into account. Since 1978, the National Archives has had an electronic archiving system (CONSTANCE), but not automated, as we understand the term nowadays, and without any specific dedicated budget resources. Thinking of records management and digital archiving as a whole was an efficient means of experimenting with the whole of the processes in the archiving chain. The Archives' own archives thus became a true "sandbox" for a whole series of tests, for designing new processes, and deliberations for bringing about an evolution in the way the archivist's role is viewed in a digital world, where technology seems to be the first, easy answer to collecting, conserving, and communicating the data.

It is in this way that the internal work programme on the institution's electronic messaging systems has made it possible to propose concrete answers relating to collection, faced with the crucial issue of conserving the messaging services of the ministerial offices of the previous Presidential term.

2. **Step 2: Historical Identity of the National Archives: the archiving of the internal management tool of readers and collections, CARAN⁴**

2.1 Digital archives: a vector for a new identity?

In order to increase performance and accessibility for users, a new archival information system (Système d'information archivistique, SIA) was launched right at the start of July 2012 on the Paris sites: this allows internal management of the spaces, unavailabilities, and readers' orders, as well as offering a virtual reading room, once the great majority of the catalogues are put on line. This better-performing system replaces the former CARAN application (standing for "*Centre d'Accueil et de Recherche des Archives Nationales*"), itself launched when the new building dedicated to receiving readers opened in 1988.

This application was the first in France to have been developed for archive services. It was used as an example and as a support for other developments, in France and Italy. It opened the way to a new form of communication with the public and new reader tracking, offering services that no manual management had hitherto been able to provide. The name and the application itself are very present in the institutional and collective image.

So archiving the CARAN application was a fundamental element within the context of the experiments launched by Records Management and Digital Archives:

- the electronic data have great historical importance and not all of them are going to be migrated into the new application;
- right from when this database was set up, documentation was seen as an essential element. Thus it is very well documented;
- it was developed thanks to the National Archives' teams of archivists, some of whom are still working at the National Archives today;
- it has formatted the working processes in the reading room, and is the core discipline of many agents.

Thus the archiving of this database had to be not confined to a simple technical transfer of the data from one medium to another with simple archiving of paper records in parallel. It had to be achieved with an awareness of the emotional loss that is produced by moving to a new system and the grieving for working practices that some people have been operating for a great many years.

4 We would like to thank all the team at the National Archives who made it possible, especially Martine Sin Blima Barru and Yann Potin.

5 The system has been deployed since October 2011 on the National Archives site in Fontainebleau, which does not use the CARAN database.

2.2. Collection methodology

In the first instance, it was necessary to assess the volumes of data to be archived and to make the connections with the mapping of the computer applications. During this time, the paper documentation was identified and located, as being part of the same submission procedure.

CARAN was a tool whose physical side was based on the actions and practices of the staff and readers. It is them who are its soul. Archiving only the computer content would rob this database of all life. This is why it seemed necessary to associate actions and words with the tool:

The first step was to produce a filmed report, just before switching over to the new system, making it possible to follow a reader and the internal circuit of the archiving boxes, outlining the discipline practices and the functions of this application.

This initial work had a dual benefit: once the application went out of service, on 18th June 2012, its operation remained readable and understandable by everyone, without needing to “set up” the database again on a computer. In addition, for the Reading Room agents, this report seemed like a recognition of the quality of their work, and gave a sense of security in the face of the disappearance of the memory associated with their practices.

The second step, started in July 2012 and continuing throughout the Autumn of this same year, is the implementation of a collection of oral testimonies from key people in the CARAN project and the managers of this database. This makes it possible to follow the processes that led up to the creation of this tool, the choices that were made, the functional developments, the administrative and institutional context, personal experiences, and anecdotes.

This form of “total archiving” has also been a strong means of cohesion for all the participants in this project which has truly been building a new identity for the institution.

2.3. Towards a rethinking of archiving?

This collection, emblematic for the institution within the context of its restructuring plan, has also made it possible to reflect on the working processes for collecting digital archives: what is a digital archive? The representation information? Its documentary description?

In point of fact, the question of the identity and of the participation in managing the changes to the institution led the two departments to think up a complete “global” collection: archiving the CARAN database, its documentation, the processes (producing a video of the process, interviews filmed with the people who set up the database).

Since the early 1980s, the CONSTANCE system has been archiving numerous databases, along with what could be related to the representation information of this data, within a single submission. The electronic file never stands alone; it is always accompanied by all the elements of representation information (structural information and semantic information), whether these are on paper or electronic. Thus the CONSTANCE system, in order to guarantee the strong, indestructible link uniting file and representation information within the SIP (Submission Information Package), considers that an electronic archive submission is made up as follows:

1. An **Information Object** with

- article 1 corresponding to the file(s) to be archived (the data object according to *OAIS*)
- articles 2 to *n* corresponding to the representation information⁶

2. A **finding aid** that documents the production context.

Now up till now this principle laid down in the late 1970s had never been challenged. In view of the progress in electronic production and the new means available to us today (digital stills and video cameras, etc.) the National Archives’ Digital Archives department and the Records Management unit have been forced to rethink this description. The opportunity of archiving CARAN, with its strong identity dimension, was a golden opportunity to do this.

In collaboration with the National Archives IT Department, the extraction of the CARAN database was

6 Open Archival Information System (OAIS) standard (ISO 14721/OAIS), p. 21: “Representation Information: information that translates a Data Object into more explicit concepts. For example, the ASCII code definition describes how a sequence of bits (a Data Object) is converted into characters”.

scheduled for 18th June 2012, the roll-out date for the new archival information system on the National Archives Paris site. It turned out that the traditional representation information for the CONSTANCE system was sufficient (data dictionary, structure dictionary, conceptual data model). It was relatively easy to reconstruct all this information, since the current administrators of the CARAN database had kept the majority of the original database documents (Special Technical Specifications, Specifications, operational specifications, etc.)

However, the identity question in collecting the CARAN database caused us to wonder about the documentation of this database in respect of the finding aid. The end of the CARAN database meant in fact not just shutting down a management application that was almost 25 years old, but also ending a certain number of work processes that were modified by the deployment of the new inter-site application and had an impact on a larger number of processes for processing and communicating paper and electronic archives. The “traditional” finding aid would then turn out to be cut down in order to document the discipline processes and the number of agents involved, and to take account of the geography of the Paris site, which had led to working processes and database functions being adapted.

The next step in our deliberations is the incorporation of these elements within the context of the description of a traditional finding aid: how to produce links to the documentation that itself now forms National Archives’ archives? The simplest solution found is to incorporate it into the research instrument’s additional sources. But this choice is showing its limitations. Certainly, the information collected for the CARAN documentation does play a part in contextualizing the application and in the description of its environment – but can they for all that be reduced to this simple function?

So what is the place of this documentation that contextualizes the application? Must it be found in the **Representation network**? The film of the CARAN process and the interviews carried out in parallel with archiving the database do seem to constitute a Representation network, since they are additional references to other elements of representation information⁷. Over and above the traditional screen captures, the film makes it possible to have a dynamic, contextualized representation of the database. So this forms part of the rest of the documentation intrinsic to the application, just like the User Manual, for example.

But despite everything, this solution is far from satisfactory, since it is easy to appreciate all the different possible angles for approaching these electronic documents, which are at one and the same time representation information and native digital documents. In the context of these deliberations on research instruments and their connection with electronic elements, the paper research instrument appears too limited. In the paper world, the existence of additional sources made cross-referencing of collections possible, but made crossed exploitation of their contents difficult. In the digital world, this approach is simplified, and indeed even necessary, when there is a need to make the representation information itself last, which becomes, as in the case of CARAN, an archive document and hence a true Data Object in its own right. The contextualization approach is thus in a network with different angles. The considerations for definitive electronic archives are quite similar to the records management approach to discipline documents. Each department may have different search metadata for the same document within an Information System, and this, depending on its missions. In the same way, can we not think of an electronic document as having a destiny other than being a Data Object or Representation Information? In this sense, the finding aid appears inadequate, indeed simplistic. Now, armed with this observation, it only remains for us, the archivists, to find a discipline solution to this problem: how to integrate these multiple angles of approach to the final electronic document? The deployment of the National Archives’ new archival information system, and above all its documentary phase in the second half of 2012, will be bringing a number of solutions. It only remains to be seen if this will be enough...

3. Step 3: Digital Identity: Building “Pierrefitte 2.0”: The VITAM Project⁸

The new step in this collaboration between records management and digital archiving is embodied in the third process in renovating the institution: “Digital Pierrefitte” or “Pierrefitte 2.0”. The expertise of the National Archives combined with its “sandbox” of discipline tests and its internal digital archiving campaigns

⁷ OAIS standard, p.66.

⁸ To have an overview of the digital archiving collection of the National Archives, see : Article from the Ministry of Culture website, C/blog: <http://cblog.culture.fr/2012/06/27/archiving-electronique-archives-nationales> (article from 27th June 2012, consulted on 19th July 2012).

is today feeding the new project to update their 30-year-old electronic archiving system CONSTANCE. Actioned by the National Archives and the IT Department of the Ministry of Culture (Information Systems Sub-directorate), its first objective is to respond to the evolution in the central administration and to adopt a “mass-production” approach to collecting its digital archives. This project was moreover included in the Ministry of Culture’s IT outline plan in December 2011. Since then, the Ministry of Foreign Affairs has joined the project. The project’s philosophy is directly in line with the legacy of CONSTANCE: simplicity, neutrality, durability, integrity.

Christened VITAM⁹ (*Valuable Items Transferred to Archives for Memory*), the project entered its operational phase in July 2012 after a year of studies and proposals for structuring the project. The first iterations of the project are scheduled for 2014. These will cover the re-use of the data from the old system and the archiving of the Acts of the French Republic (from SOLON, a discipline-specific application from the Prime Minister’s office).

In addition, the project is benefiting from the pilot experiment by the Ministry of Culture for its own archives through the mapping of the ministry applications and the development of EDM tools (project SA2). The project Monitoring Committee brings together the Ministry of Culture, the Ministry of Foreign Affairs, the Ministry of Defence, and the Ministry of Finance (Archivists and IT departments).

The approach adopted is iterative, so as to be able to respond to the most urgent files without having to put in place a global system that would be like some endless building site. Several pre-studies have been launched from a technical standpoint (storage, workflow, interfacing), and likewise from a disciplines standpoint (submission processes and input channels).

The project is being accompanied internally by a real change management policy so that all the National Archives sites can speak the language of digital archiving. Up till now, the electronic archiving system only actually affected the Fontainebleau site, which was the original site for collecting contemporary archives. Now, within the context of the reorganization, the Pierrefitte-sur-Seine site must also be in a position to understand and receive these collections. Thus the National Archives, through the VITAM project, have put in place a change management strategy in order to update the identity of the institution internally. Taking the archivists’ new expertise on board in this way is fundamental to achieving this renewal. To this end, a not-inconsiderable part of the project team tasked with designing and putting in place this new system will be dedicated to this change management: training sessions, participation in the discipline work groups drawing up the document handling processes, interfacing with the National Archives’ other information systems. There are still many avenues to be explored, but the new face of the National Archives is already being sketched out within the central administration, offering all archive users a lasting solution for the conservation of the memory of the State, in a direct line from the work processes put in place since the late 1970s and which still today are enabling the National Archives to conserve over 15 TB of native digital data, the oldest of which date back to the early 1950s! *Ad vitam eternam...* a play on words that is admittedly easy, but which represents a constant challenge in the digital world.

Conclusion

The re-founding of the National Archives and the removal of its agents to the Pierrefitte-sur-Seine site has obliged the institution to rethink its policy in terms of records management associated with definitive electronic archiving. The merging of skill areas has made it possible to participate in the internal change management through skills upgrades for agents on the question of the digital world, the collection of electronic archives that is fundamental for the institution, the move on from the paper archives building project to the National Archives digital project. The institution’s internal archiving has proved to be very beneficial to the institution in many respects. It has allowed us to put our own archiving system and the choices made in the late 1980s back into perspective. These deliberations are today feeding the “Digital Pierrefitte” project, both in the field of collection and in the field of research and documentation of the electronic collections. It would not have been possible to put this project in place without a programmed and fairly exceptional collection of the institution’s digital archives, such is the extent to which this is time- and energy-consuming! It remains to find a solution for our research instruments and the communication of all these archives collected. The fusion of these three identities (Archiving, Historical, and Digital) within the institution lets us arrive at an observation: after the National Archives as the foundation of the past memory of the State (Archiving and Historical Identity), the institution is currently getting involved in the future by becoming, through its new VITAM project, the guarantor of the current and future memory of the State (Digital Identity), fulfilling to an ever-greater extent its fundamental role in guaranteeing a democratic,

9 VITAM stands for *Valeurs immatérielles transférées aux Archives pour Mémoire* in French.

republican State such as the French Revolution had envisaged in 1792 when the National Assembly Archives were founded, the precursor to the National Archives.