

Integrating physical preservation and digitization: a currently effective process in the French National Library

Philippe Vallas

Deputy Director of the Conservation Department

Bibliothèque nationale de France

Philippe.vallas@bnf.fr



Copyright © 2013 by **Philippe Vallas**. This work is made available under the terms of the Creative Commons Attribution 3.0 Unported License:

<http://creativecommons.org/licenses/by/3.0/>

Abstract:

In a few years, mainly from its priority mass-digitization programmes, the French National Library (BnF) built very huge digital collections, which are growing still very quickly. Its digital library Gallica has already more than 2 million documents. Therefore the costs of digital preservation are increasing very fast: BnF digital repository, SPAR, has more than 2 400 To, from increasingly various types of documents. In the same time, budgets for physical preservation are decreasing very quickly because of the economic crisis, and the staff is increasingly dedicated to digitization programmes.

In this difficult context, the Conservation Department (DSC), which is coordinating both digitization and physical preservation, is trying to reduce the global cost and the fall of activity in physical preservation by using each of the two types of conservation in aid of the other type. As digitization is obviously the priority, with a far larger budget, the strategy is to adjust and organize digitization programmes in order to compensate as much as possible for the decreasing of physical conservation activities:

- *Despite mass digitization requirements are generally considered to be difficult to reconcile with good conservation conditions for the original documents, BnF progressively advances (even in contracts run by private companies) from a digitization aiming at highlighting the collections to a true “**preservation digitization**”, for which damaged, brittle or very frequently communicated documents can be selected and are more protected against digitization damaging effects: We are more and more demanding in our contracts : We require non-damaging scanning equipment, high-quality pictures and file formats, we forbid 180° opening scanning for most documents, we check carefully the conditions of documents before and after digitization ; the contracts prescribe penalties when documents are damaged; a common training is systematically given to all staff from BnF and from the service providers involved in the programmes. Skilled restorers have been recruited for training and watching of the physical conditions of documents. We set up an audit process for an efficient and cheap control of digitization quality; the files from digitized documents get a high quality level of supervision in our digital repository.*

- ***Toward a physical preservation taking digitization into account:*** *We created integrated circuits combining physical and digitization treatments, the ones realized according to the others. For example, digitization is often immediately completed by a box to improve the preservation conditions of the original document; brittle documents get a physical preparation to make easier or even possible digitization. Two specialized workshops have been set up to prepare newspapers and printed books. More and more often, we choose to digitize documents which would have been rather restored a few years ago, and nowadays we don't restore any more all the documents which have been damaged by digitization. From now on, the head of the BnF emergency plan will coordinate the policies of risk management both for physical and for digital collections. Lastly, the public will not have access any more to the original documents (excepted for exceptional grounds) after digitization.*

The presentation will give an overview of these measures or projects, without hiding their drawbacks and limits.

1- Digitization as a protected priority

In a few years, mainly from its priority mass-digitization programs continued without any break since 2007, the French National Library (BnF) built very huge digital collections, which are growing still very quickly (of more than 300 000 items in 2012). Its digital library Gallica has already more than 2 million documents, and more and more often “heavy” in Mb or Gb. From now on, the official purpose of BnF is to put online a great part, if not all, of its huge collections (more than 30 million items), to highlight and spread them abroad more than to preserve them, and with an increasingly high level of technical requirement (minimal resolution of 400 dpi, pictures almost systematically made in colours). The first strategic objective of our “performance contract” of the five last years was to become “*a digital library of reference*”, at the national level.

Therefore the costs of digital preservation are increasing very fast: BnF digital repository, SPAR, has more than 2 400 Tb, the kinds of documents are increasingly various, and we are obliged to create more and more programs to keep them.

Up to now, digitization programs are almost totally spared by money cuts, for most of them aren't funded on BnF budget, but by a special national tax managed by the Book National Center, which is a part of the Culture Ministry. The income of this tax is currently large and regular. In the Conservation Department/DSC, the staff of the Digitization Service, which coordinates BnF digitization activities in administrative and technical fields, is the only one the size of which is still increasing

2- A difficult situation for physical preservation

On the contrary, physical conservation is severely hit by the economic context, which is currently very difficult. For a few years, because of the economic crisis and the national debt which force the French governments to reduce public expenditure, BnF budget is progressively cut down. This situation has deteriorated a lot in the current year, and probably this trend will continue for several years. In that context, the sizeable budgets dedicated to restoration, binding, enclosures (boxes, sleeves), etc, which are crucial for physical preservation but the effects of which are not immediately perceptible by public, are decreasing more quickly than others. Therefore physical preservation activities are widely reducing. For example, the number of new documents yearly bound has fallen from about 46 000 to 23 000 between 2008 and 2012.

These financial difficulties have also led to cuts in staffs dedicated to physical preservation, mostly in the most traditional workshops (restoration, microfilming), where the average age is often the oldest : people are not replaced when they retire. This reduction is all the more difficult to manage since preservation staffs are increasingly dedicated to digitization programs, particularly in collections departments (choice and sorting of documents to digitize are partly made from physical criteria).

This situation is made even worse by extensive cuts in other budgets which are crucial for physical and digital preservation: building maintenance, air conditioning, cleaning, security, computers, which guarantee a safe environment and an efficient management.

In this difficult context, the Conservation Department (DSC), which is coordinating both digitization and physical preservation, has decided as a priority goal - in close collaboration with departments from the Collections Direction (DCO) which are responsible for the choices of documents and treatments - to reduce the global cost and the fall of activity in physical preservation by using each of the two types of activities in aid of the other type. As digitization is obviously the priority, with a far larger budget, the strategy is to adjust and organize digitization programs in order to compensate as much as possible for the decreasing of physical conservation activities; but, in the other hand, the use of physical preservation treatments is more and more often inevitable to make easier, and even possible digitization of damaged or brittle items.

3 – Toward a true “preservation digitization”

Digitization of library collections does in no way reduce to technical aspects like types of scanners, image or optical character recognition formats. One must never forget that in many cases the collections to be digitized are made of original documents, of various supports and physical types, which are generally of cultural heritage or documentary value. That value is often established by law (Legal Deposit).

This is the case at the French National Library. Documents selected for the numerous digitization programs are considered to keep their documentary and heritage value after they have been put online. With a very few exceptions (items bought only to be digitized), BnF collections have to be kept for long-term conservation. If there are several copies of an early printed book, all these copies are considered to be of heritage value.

Therefore the BnF has clearly stated, from the beginning of digitization programs, that the conservation of collections was a priority and cannot be compromised by digitization: documents the pictures of which are impossible to take without significant damages must be systematically excluded from digitization programs. Similarly, conservation concerns are always more important than production rates, even in large scale programs. These rules are clearly written in contracts concluded by BnF.

These precautions are all the more necessary since digitization is, and will probably stay, a delicate trial for heritage documents, particularly for old, brittle, in volumes and large-sized documents: in spite of the technical progress, all the types or models of scanners currently available may be dangerous for collections if they are not properly used (by lack of expertise, lack of respect for documents, or excessive scanning speed), or if they are unsuitable to the kinds of document which have to be digitized. A scanner model fitted to all types of document doesn't exist yet, and to be successful a program must be prepared in advance by dialog and consultation between all parties involved (curators, restorers and photographers particularly) to fit the handling and photography methods, and the range of equipment (scanners, book cradles) to physical features and conditions of collections.

Learning from the three mass digitization contracts successively run from 2007, the Conservation Department, in charge of the operational management of the contracts, could notice that if the programs weren't prepared in a careful and suited way (which means beforehand discussed between all the stakeholders), mass digitization requirements were difficult to reconcile with a safe preservation of items. Some problems encountered during the first BnF digitization contracts showed us how crucial it is to precisely adapt this treatment to collections: Because too much documents came back in bad condition after digitization, even if the percentage remained very low, some curators responsible for the collections got indignant and relations became strained into the different BnF teams involved in the programs (Conservation Department /DSC, , and Collections Direction/DCO, which is responsible for the selection of documents and documentary orientation), and a lot of curators didn't trust the digitization service provider any more, and became reluctant towards digitization itself. They restricted their selection to documents in perfect conditions, which greatly reduced the documentary relevance of the programs, and they systematically accused the private service providers for any damage noticed when the books were returned.

In order to spare the original documents, and to save time and money, our library widely used of course the numerous microforms produced from its collections in the last decades: a specific contract for microform digitization was established, and a lot of the other contracts for the digitization of original documents included microform treatment too. But the results have not always been of high quality, in particular in the case of old microforms, and, of course, when the original document was in colors. And this seems to be all the worse because the level of technical requirements has increased very quickly in the last years. In an increasing number of cases original documents are used for digitization if in good conditions, even if they had been microfilmed before. In other cases, fortunately less frequent, documents had to be digitized a second time to match the current standards.

Therefore the Conservation Department, in cooperation with collections departments, is working to more and more systematically organize digitisation activities in order to reduce as much as possible damages caused to collections, and on the contrary to use these activities as an help for safe conservation of collections.

Learning from its almost ten-year experience in digitization programs entrusted to private firms, the BnF began in the last years to take measures and establish rules to allow both its staff and the service providers (which realize the most part of its digitization programs) to reach these goals.

Firstly, we must give in these programs an important and increasing part to damaged, brittle and/or very often consulted items, as it was the case in the past decades, when reproduction was made on microfilms and microfiches, but that change can occur only very progressively, because it needs beforehand that we make collections curators sure that the documents they are responsible for will be treated with care and respect, even in mass programs run by private companies.

Up to now, in fact, the most brittle (notably in acid paper) and valuable documents are mostly digitized in BnF workshops (if possible on-site, to limit transfers as much as possible), by its own staff which are already sensitive to preservation requirements and which can provide restoration treatments if necessary. . In fact, BnF collections departments trust BnF digitisation workshops, for they have already tested them, they are very close of them, and therefore they send them a lot of damaged items – and selected precisely because they are damaged.

To increase our digitization capacities we also signed digitization contracts dedicated to special collections – manuscripts, precious printed books from the Rare Books Department for example – and in these cases the digitization service providers had to set up their workshops on site, closely watched by BnF curators.

But internal workshops and contracts for special collections represent only a minor part of BnF digitisation activities; now the main goal to reach is to integrate damaged and brittle documents in mass digitization programs, of which they were up to recently systematically eliminated.

The current contracts give obligation to digitization companies to treat very carefully all the items they have to reproduce. The new rules are in use in the current contracts, which prescribe **penalties** if documents are damaged, giving a detailed range of penalty prices corresponding to types of damages. Before sending off an item for digitization, a survey of its physical conditions is carried out and recorded jointly by BnF staff (Conservation and Collections Departments) and staff from the service provider. A computerized file is created, which is compared with the corresponding book by a BnF restorer each time a damage is recorded upon return of the item after digitization.

If the damage is recognized as a result of digitization, a detailed invoice is sent to the digitization service provider. At the end of the first mass digitization program for printed books (2010), the amount of penalties allowed to recruit a private restorer for almost one year, who repaired several hundreds of damaged items. It is the same thing for the current mass contract.

However, we also understood that the preventive policy, by consciousness-raising and training, was an essential complement of the penalties: At the beginning of the current mass digitization program (2011), in order to minimize damages, and thus to permit that digitization treatments are conducted quickly and without conflicts, some restoration and binding experts from the BnF Conservation Department set up a **special training**. Its goal is to give the tools to decide if and how a document can be digitized, in accordance to its physical condition and features (in particular opening angle), and to the equipment (scanners) used by the service provider. The training is also crucial to teach how to accurately handle documents, especially during scanning itself. This training is systematically given to all BnF teams and teams from the digitization service provider, who select, handle and transfer documents at the different steps of the digitization circuit, and who take part in the physical conditions surveys before digitization.

This short and practical training permitted to give **shared knowledge of respect for the document** to all people involved in collections handling at the different steps of the digitization circuit, and contributed to improve relations, not only between the BnF and its service providers, but also internally between teams from the Conservation Department and teams from the Collections Departments, the latter feeling reassured for the safety of their documents. Nevertheless, the training must be frequently given again, because of the extensive all together in the teams, and in the private companies teams the digitization operators must be called to order very often, for they tend to forget the rules of appropriate handling when they try to reach the high production rates required in mass-digitization programs.

Physical conditions surveys on documents before and after digitization, and training activities are all carried out by skilled restorers, which greatly contribute to reassuring collections managers and to creating links between digitisation and physical preservation.

So BnF teams pay more and more attention to the brittleness of documents, and especially to the degree of opening, which can vary a lot from a book to another. We are more and more demanding on the digitization conditions and equipment: now **partial opening of volumes** (110°/120°, and even 80° opening) is increasingly required, and most BnF books cannot be digitized on a 180° opening; the service providers have to buy special book cradles. However, it wouldn't be useful to be excessively demanding, because in the current state-of-the-art, systematic partial opening digitization is much more slow and expensive than 180° opening digitization, and therefore it doesn't really allow a mass output.

- To realize a real "conservation digitization", we have to express technical demands of course, but above all we have to check that they are effectively respected; because in "conservation digitization" the digital files have to be definitive as much as possible- it is that we have to reach the highest quality level for consultation. Therefore BnF staff set up a preventive organization of **quality control** for the digitization output: some people from the Digitization Service of the Conservation Department, who have got special training, realize regularly audits of the workflows in contracts run by private companies. These people are working with a very detailed check-list. Of course this organization is completed by automatic computerized, and then visual controls of the digital files produced when they

are sent to BnF. These last controls are made partly by a specialized team of the DSC, partly by curators from the collections departments (for special and invaluable documents).

- Lastly, to reduce as much as possible the risks of loss, a specific program, with the highest level of watch, has been created in the BnF digital repository, named SPAR, to keep the files of digitized items. We have planned in the future to create several other programs in SPAR in order to adjust the level of watch to the degree of brittleness and durability of the original documents.

4- Toward a physical preservation associated with digitization

Thus, our new strategy is to use as much as possible digitisation programs in favour of physical preservation; the logical counterpoint of this is that means and skills of physical preservation are increasingly often employed to make digitisation easier and less harmful, and to reduce the percentage of documents which cannot be reproduced because of their bad conditions. Physical treatments are also frequently used in association with digitisation, to strengthen its protective effects.

Therefore we progressively set up integrated circuits of both digital and physical treatments, interdependent and carried out the ones according to the others:

- Before to be sent back in the storage rooms, documents digitized by internal workshops or by private companies are increasingly often –if not systematically – put in boxes or sleeves immediately after they have been reproduced. These enclosures have a protective effect of course, and they are also useful to warn, by a specific stamp or mark put onto, of the digital substitute which is available for the public.

- More and more often too, documents which have to be restored before an exhibition are digitised at this moment, particularly if they need to be discarded before restoration. Digitization is made systematically, as a security measure, in case of external loan for exhibition.

- On the contrary, because of the fall of budgets and staff, **digitisation is increasingly used as a substitution treatment**, which permits to spare time-consuming and expensive physical treatments; documents aren't restored after digitisation any more, excepted for the most invaluable ones, or for the ones which have obviously been damaged by private service providers; But nowadays it can even happen that items damaged in the external programs are only put in boxes or sleeves.

As a logical consequence of this evolution, the consultation of the original documents, even if they are still in good conditions, will probably not be authorized any more as soon as the digital substitute will be available online. Exceptional consultations would remain possible of course, but this last measure, if it is really decided, will suppress handle by staff and public, which is the main factor of documents deterioration.

But one of the most significant steps made in the two or three last years is to use physical repairing treatments to make the digitisation workflow more fluid both in internal and external programs:

- In order to reduce even more the number of damages in the current mass-digitization program for printed books, and also the number of documents which can't be selected because of their brittleness or their bad conditions, the Conservation Department has set up a small **special workshop**, run by two half-time staff (a skilled restorer and a library assistant qualified in binding). Properly speaking they don't restore but do quick repairs for some types of books to prepare them for scanning: strengthening the back of early non-bound books or pamphlets to prevent them from breaking, replacing staples by some stitches on brochures, repairing the largest tears, the most brittle head capes and section joints, etc. Last year more than 2 000 documents were repaired or prepared in this workshop, which is about 5 % of the total number of documents digitized over the year within this program, and a sizeable addition of the total number of items which underwent small repairs in all the BnF workshops.

- For the **treatment of early newspapers**, physical preparation of collections before digitization is not only helpful but really essential: Most of the huge collection of newspapers printed during the Third French Republic (1870-1940), the “Golden Age” of the French press, is actually in very bad conditions (acid and brittle paper, folded leaves because of the lack of space in storage rooms) and therefore has to be flattened and often repaired to allow digitization. Nowadays, we don’t restore any more but make minimal repairs to allow an easier and high-quality digitization, while being respectful to the integrity of documents. The usual treatment is to flatten newspapers sheets by hot pressure with electric irons. When the sheets are too brittle to be easily flattened and handled, some repairs are made with Filmoplast®, a hemp paper of 10 g. gluing when heated, which has good long-term conservation properties but cannot be easily removed. When possible, these repairs are made only in the most brittle or damaged places of the documents (original folds, margins, other places used for handling), and with narrow stripes of Filmoplast®, but sometimes it is necessary to strengthen one side of the full sheet by lining with a hot press and wide rolls of Filmoplast®. We estimate that about 20% of the newspaper sheets need a slight repair (in addition to flattening) and 5% a full lining. When a newspaper has been bound, it is often necessary to remove the binding of the volumes (that is to remove the cover and the glue of the spine) to allow a 180° opening and a perfect digitization. All these treatments are time-consuming (from a few seconds to 4 minutes per sheet, and up to 2 hours to remove a cover), and the strengthening product is quite expensive: globally they increase the total cost of press digitization of about 30%, although they are not made by high-skilled restorers but by library assistants who have received special training. BnF has set up a 4 people workshop (equivalent to 3 full-time people), which treated about 112 000 newspaper sheets in 2012. But to build a mass-digitization program reaching the huge needs of our newspapers collections we would have to recruit more than twenty people.

A lot of these activities that I have just described are carried out by skilled restorers or other staff qualified in repairing documents. With the development of various and massive digitization programs, it becomes increasingly obvious that it is useful, and even necessary, to have restoration (or repair or preparation) workshops close to the digitization laboratories: In fact, thanks to the treatments they carry out on books, but also thanks to their advice and training activities, these workshops staffs take a decisive part to prepare these programs, to make digitization easier, better and faster, to simplify selection of items and reduce the number of those which can’t be digitized, to reassure the curators, to establish better relations between the different teams working on these programs, and of course to minimize damages caused by digitization and the handling required by it. One skilled restorer has been recruited at the beginning of the current mass digitization program, to train all people involved in the treatment process, to set up and manage the preparation workshop, to supervise the surveys of the physical conditions of the documents, to establish the damage estimates, to make sure that instructions are applied, to give advice to photographers and colleagues who select documents. We have planned that in the next years our two repairing workshops will get more staff. Actually, in addition to our two dedicated workshops that I have just described, almost all the other BnF restoration workshops are now dedicating a significant part of their time to documents going to, or coming back from digitization.

Thus, digitization doesn’t make obsolete physical treatments, and the entrance of BnF in the digital era provides paradoxically new opportunities and prospects for its restoration workshops.

5 – Limits and risks

If some of the measures described beforehand are not only necessary in the current crisis, but also desirable and energizing by themselves, it would be a nonsense, of course, to describe all the current trend as an ideal and intrinsically positive evolution. We must acknowledge that in some ways it is only a last resort, tactics (even if the most suitable and efficient as possible) to limit the damages which will inevitably occur from the decreasing of physical preservation activities, whatever BnF digitization policy could be. This trend involves some risks too:

- Up to now, the decisions described beforehand, already taken or still only at the planning stage, don't change very significantly the objective to which our programs are oriented, which is mostly highlighting and spreading abroad of collections. For the most part, items currently reproduced (old monographs, special collections) are not the ones which would be selected if our programs would be focused on preservation, because they don't belong to the most damaged nor the most consulted part of BnF collections. On the contrary, we don't run mass programs to digitize newspapers which yet are the most brittle and the most in demand part of our collections, because nowadays our sponsors are not interested by the press. Up to now, only 5% of the BnF brittle newspapers collections have been digitized.

- Because of the high number and the high priority of digitization programs, and because the extensive cuts in staffs, a lot of people previously dedicated to physical preservation are progressively appointed (even if for one part of their time) to digitisation tasks. This trend does worsen the decreasing of physical preservation activities: sometimes it happens that budgets are not completely spent because of lack of people to select and prepare items for preservation treatments. Thus, although digitisation takes a real part in collections preservation, it also seriously challenges physical preservation in access to staff resources.

- Despite their positive effects, digitization programs, and even mass programs, will not be able to stop deterioration of physical heritage collections caused by too extensive and too long-lasting reduction of physical preservation budgets. On the contrary, this deterioration will eventually make these programs more difficult and more expensive to run.

- The durability, or at least the complete development of our digital repository SPAR (which only gives us a guarantee of the long-term preservation of digital files) is not completely sure, because SPAR is partly funded on the BnF decreasing budget. For the first time in the current year, some cuts have been made in it.

- This situation may also create difficulties in staff psychology and management: A significant part of our crews, although non reluctant to changes and new technologies, don't agree to the absolute priority given to digitization. Some of them believe that a risk does exist that the culture of physical documents, and respect for them, will disappear in the future, and that afterwards nobody will be able to properly care and even handle them. Many restorers, in particular, are still very suspicious of digitization, because they are not interested by works of physical preparation or repairing created by that activity, that they regard as too simple to show their skills in a flattering light; they also are afraid that the rise of digitization might make more rare the difficult and meticulous restoration treatments they particularly appreciate and which justify to their mind the existence of their profession. In our hard times it is often difficult for BnF managerial staff to manage both teams completely and enthusiastically involved in digitization programs or research, and teams still very reluctant and feeling themselves sacrificed in the general trend.

That suspicion is maintained by the ambiguity of some official speeches, from which one can understand that the rise of digitisation and digital-native documents could compensate for, and even make acceptable, the decreasing of means dedicated to physical preservation.

- If the current cuts are continued in the next few years, we can fear that an excessive fall of resources in staff and budgets dedicated to physical preservation might quickly jeopardize the integration of the two activities which is now under way : thus digitisation would seem excessively privileged, and relationships might become strained again into the different staffs, because a lot of people wouldn't agree that physical preservation could be reduced to a subsidiary activity, to a mere complement of digitization.

6 – Conclusion

Despite these reserves and the remaining difficulties, the global assessment of that evolution is really positive, without any doubt: Although initially rejected by a large part of the staff as an exogenous, too extensive and too priority activity, mass digitisation has progressively found its place among BnF activities, and links have been built with physical preservation, thanks to BnF conservation Department which coordinates the two activities. People with very different skills and job cultures have learned to better know each other and to work together, in a more transverse and less hierarchical way than usually. A common culture is progressively built by work meetings and trainings. In an unexpected but finally logical way, digitization creates new activities and responsibilities for people belonging to the most traditional preservation specialities, who are the restorers, but also for a lot of library assistants who take part to surveys of physical conditions of documents and assess the possibility to digitize them. We now clearly understand that our restoration workshops, situated very close to digitization laboratories, are a great asset for BnF, and that therefore we haven't to fear that digitisation might be totally made outside in the future.

In our digital repository SPAR the physical features of original documents are progressively taken into account for long-term preservation of the digital files. In order to reduce the damages caused by public access, the use of the digital substitute will be promoted and will probably become mandatory.

To summarize, and globally speaking, digitization is playing and will play a more and more important part in collections preservation, and physical preservation treatments are increasingly employed to make digitization easier and more efficient to protect documents. A crucial step has not really done yet, which is to integrate in great number damaged and brittle items in mass digitisation programs, and to use widely these programs to treat this kind of documents (in particular newspapers and periodicals). But integration of physical preservation and digitization could be endangered by persistent and excessive cuts in physical preservation resources, which could also jeopardize the continuation of BnF cultural heritage collections digitization.