How to fit in?
Integrating a web archiving program in your organization

IIPC-Sponsored Workshop
Bibliothèque nationale de France, Paris
26th-30th November 2012

Workshop report and evaluation

Peter Stirling (Bibliothèque nationale de France)
Jaanus Kõuts (National Library of Estonia)

April 2013

Contents

Introduction

Aims and objectives

The aims of the workshop, as stated in the proposal to the IIPC, were as follows:

> A growing number of institutions are getting involved in web archiving. Tools are becoming mature, national and collaborative collections are blooming and growing, notably thanks to the efforts of the IIPC community. However, implementing a sustainable web archiving program and team in an institution is not only a matter of mastering web archiving technologies: web archiving is also an organizational challenge. In longstanding heritage and research institutions in particular, web archiving is still often considered unfamiliar, sometimes irrelevant. Web archiving project managers and curators therefore constantly need to advocate for their program and to demonstrate the continuities between collecting print and harvesting websites. They also need to find the best strategy to implement and distribute their program at the right place and at the right time within their organisations. At a time where budget constraint is generally the rule for public institutions, they therefore need to make the most of institutional resources outside of the web archiving team, notably collection curators, engineers and decision makers.

The workshop was therefore intended to deal with these issues, based mainly on the experience at the Bibliothèque nationale de France. Web archiving activity at the BnF dates back around ten years, and over this period an organizational structure has been put into place involving digital curators, IT staff and content curators, and placing the activity within the management structure and missions of the Library. Different speakers from the BnF would present different aspects of this organization. To provide a contrasting point of view the British Library was invited to give a presentation, and the aim was to facilitate discussion among the participants throughout the workshop, to compare the situations and solutions in different institutions and identify best practice. (The full program of the workshop is given in appendix 1.)

Participants

Regarding the intended audience for the workshop, it was aimed principally at web archiving project managers, along with curators or engineers who also take part in organizational decision making and team coordination. It was primarily designed for institutions who are starting out in web archiving and who want to find ways to implement web archiving in their internal organization, and for those who already have a first experience but now wish to move to a further step of implementation and therefore need to explore new organization schemes. As the organizational solutions presented were largely to be based on the experience of the BnF it was expected that the workshop would be of special interest to staff of National Libraries, but could be useful to colleagues working in other types of institution as well.

In the end there were 14 participants, representing 11 institutions (see appendix 2). With one exception the institutions were national libraries, and shared many issues. However the discussions showed that each institution has its own context and concerns and that no single solution will apply to all of them. Despite this, many of the elements discussed applied in multiple contexts, and in some cases issues being dealt with by one institution may become relevant to another institution in the future, as changes in the legal or organisational context require the web archiving activity to evolve.

Participants were asked at the beginning of the week what their aims and expectations were for the workshop, and were also asked in the feedback questionnaire if the workshop had met their expectations. The idea that appears the most was to learn about the overall workflow but also specific areas, not only at the BnF but at all the institutions participating, to be able to take lessons back that could be applied in each institution. In some cases this was related to changes in the context (new staff, cutbacks, changes in legislation…) but in others it was a case of seeing if new ideas could be applied to existing workflows.

Speakers and presentations

An important aspect of the workflow being presented was that it involves many different departments of the BnF, and the programme of presentations was intended to show different aspects of the organization of web archiving at the Bibliothèque nationale de France. In addition to the nine staff who work full-time on web archiving, and who are divided between the Legal Deposit and IT departments, there were also presentations from librarians from collections departments who select sites for focused crawls, and from the person responsible for updating the collection policy charter from the library, and finally from preservation experts who work on the BnF digital repository SPAR. The aim was to allow participants to discuss directly with people who perform different roles, to gain an immediate idea of the overall web archiving activity.To open up the discussion and allow comparisons with other organizational workflows, the was also a presentation from Helen Hockx-Yu of the British Library, an institution similar to the BnF in many ways but which has a very different model for its web archiving activity. Finally, the last day was given over to presentations from two other organizations based in Paris with again very different approaches, Internet Memory Foundation and the Institut national de l'Audiovisuel.

Section 1: summary of discussions and lessons learned

Collection policy and legal deposit

Legal context and overall collection policy impact on web archiving collection policy and practice, and the legal framework and collection policy vary widely from one country or one institution to another. While the starting point for the workshop was the specific framework at the BnF, the discussion showed that whatever the local situation it is important to define the organizational approach to web archiving within this context, as this will structure the organizational and operational decisions to be taken.

The BnF works within a strong French legal deposit law, which provides a definition of the "French Internet", allows the BnF and INA to collect sites without asking permission from producers, and sets limitations on access[1]. This legal framework provides the overall structure not only for web archiving but for much of the library activity: legal deposit is a central mission of the library, of which Internet legal deposit is a part. The BnF applies a mixed model of broad and selective crawls to respect legal deposit obligations: this in itself has an organizational impact, since as described below URLs for selective crawls are chosen by content librarians, while the planning of crawls and the management of the broad crawl is handled by the web archiving staff in the Digital Legal Deposit team.

Different situations in different institutions can similarly be seen to affect the organizational choices to be made. The legal framework affects the collection policy: in a legal deposit structure the collections should be exhaustive, or failing that, representative of the online production, whereas without legal deposit the collections may be structured around ideas of quality or value of the material collected. This will structure many of the decisions to be made.

Overall, not only the legal context regarding legal deposit but also the general policy of the library will have a huge impact on the organizational choices put in place. Recognition of web archiving as a legal mission, as at the BnF, can help in positioning web archiving activity in the institution and securing the necessary staff and resources. But in all cases web archiving programs will have to work within the policy of the institution, and is necessary to demonstrate that they are not only important but also structured, economically viable and sustainable. The different aspects of this that were raised during the workshop are outlined in the following sections.

---

[1] For more details on the legal framework and the definition of the scope of French internet legal deposit, see "The state of e-legal deposit in France: looking back at five years of putting new legislation into practice and envisioning the future", ILLIEN G., SANZ P., SEPETJAN S., STIRLING P. In: IFLA journal, 2012, vol. 38, no° 1. [http://www.ifla.org/files/hq/publications/ifla-journal/ifla-journal-38-1_2012.pdf]

<u>Place of web archiving in the organization</u>

Web archiving combines skills from different disciplines; in particular, it requires a high level of IT involvement while also involving more traditional "librarian" skills such as collection policy, discussed above. This leads to the question of where to site web archiving activity within the organization, particularly in the case of large structures such as national libraries (highly represented among the participants at the workshop), which may have complex and somewhat rigid hierarchical structures. For organizations dedicated to web archiving, such as Internet Memory Foundation or Internet Archive, this problem will not apply in the same way.

A first question can thus be whether web archiving should be based in collections departments (i.e. run by librarians) or IT departments (i.e. run by IT staff). This can depend on how web archiving is approached at each institution: as legal deposit, as a collection, as preservation, as IT… The experiences of the workshop participants showed that, even with a group largely composed of national libraries or similar structures, there is a wide variety of approaches.

At the BnF, full-time web archiving team is split between two departments, Legal Deposit and Information Technologies. The policy decisions are made by the Legal Deposit department, within the structure of the policy of the library, but this process functions within the framework of a contract with IT that defines the resources in terms of staff, development, storage and other needs that will be allocated to the web archiving program. This contractualization is part of the overall structuring of web archiving activity described above, and finds its place within a three-tier policy structure comprising general library policy, overall focused crawls policy and policy for thematic and project crawls (described in more detail in session 2 of the workshop). Another aspect is the role of content curators in the different subject departments, as described below; this adds another level of organizational complexity as these departments are in the Collections Section, while Legal Deposit and IT are both in the Systems and Networks Section.

Other institutions have different approaches. At the British Library, for example, web archiving was initially split between IT and collections sections but later regrouped in the same section: first in collections, but later IT. In smaller structures, there may only be one or two individuals working full-time on web archiving, with support from other departments such as IT. Another approach is to use a third party service provider for some aspects, in particular the harvesting itself: the BnF broad crawls were performed by Internet Archive between 2004 and 2008 and the Biblioteca Nacional de Espana currently works with IA, while the Deutsche Nationalbibliothek works with OIA, a German company.

Different structures and different levels of experience in web archiving mean that there is no single solution that will work in all institutions. In all cases, the most important role of staff dedicated to web archiving will be to coordinate activity, putting the policy of the institution into application with other actors as necessary. The central role of communication within the institution is discussed below, after specific questions relating to the project approach and the role of content curators.

Project approach vs. integration in existing departments

Within the overall question of the place of web archiving in an organization is the more specific question of whether it is better to have a dedicated project team responsible for all aspects of web archiving activity, or whether this activity should rather be integrated in established departments representing the different disciplines involved.

Both models were represented by different participants at the workshop, and some, including the BnF, have seen their approach evolve over time. In the BnF, web archiving started as a project within the then Digital Library Department, with a multidisciplinary project team involving both librarians and IT staff. This situation evolved progressively; in 2008 the planning and coordination functions were transferred to the Legal Deposit Department, with the development and IT aspects based in the Information Systems Department, while selection for focused crawls were performed in the collections departments. This division, still in place today, has led to the structuring and definition of roles described above.

Each approach can be seen as having advantages: the project approach may give more freedom to act and to quickly develop and test out solutions; it can also be put into place with a small number of staff and limited resources. However integration in existing structures can provide more recognition within the institution, more security for funding and the possibility to use existing infrastructure and procedures (for development, storage, surveillance…) which may lead to economies of scale. The British Library, where web archiving has also moved from being a "work program" to a business as usual unit, this evolution allowed a consolidation of web archiving as a permanent activity, as a program budget is subject to renewal, whereas operational budgets may be cut but will always be maintained.

As with other points discussed, there is no one solution that will work in all situations, but based on the experiences of participants at the workshop it seems that the project approach may be particularly useful at the beginning of a web archiving program in an organization as it is well suited to an experimental approach. Once there is a move towards "business as usual" for web archiving activity, it is important that web archiving finds a recognized and stable position in the organizational structure, whether this is as a dedicated unit or split between several departments.

Role of content curators

One aspect of the BnF workflow that was highly appreciated by other participants was the organization of a network of content curators who contribute to web archiving activity. These librarians come from the different thematic departments of the Collections section of the library, and select sites to be collected in their areas of expertise for the focused crawls.

This input from content librarians is very important in the web archiving program at the BnF, not only from the collection point of view but also for the integration of web archiving activity in the institution. In French they are known as "*correspondants*", a term which has

been chosen to be larger than "selector" as they are also responsible for quality control of the sites collected, and for the promotion of these collections in reading rooms. Web archiving is thus seen as a continuation of the duties of the subject librarian: the sites selected are in some cases a continuation of material in the physical collections, or are new forms unrelated to printed material, but in all cases are seen as complementary to other types of material and part of the collections of the library.

Each department has a "coordinator" responsible for overseeing the work of the correspondents in the department, and there are quarterly meetings between the Digital Legal Deposit team and the coordinators where questions of policy and practical details are discussed. Within the Digital Legal Deposit team each digital curator works with specific departments as the initial contact for any questions from the coordinator or other curators. There are also regular training sessions. This contact is important in maintaining the network.

Other participants at the workshop also work with content curators, but all agreed that, as at the BnF, it can be difficult to get people to dedicate time to an activity that may be unfamiliar to them, of which they may not see the value, and which is often not a specific part of their job description. The overall position of web archiving in the institution has a role to play here, as if it is perceived as "business as usual" it will be easier to demonstrate the importance for the library. Much of this therefore relies on communication within the institution, as discussed in the next section.

One interesting point that was raised was that content curators at the BnF have more responsibility for choosing technical parameters for the sites selected (depth, frequency of harvest and budget in terms of files collected); in many other institutions curators select URLs but it is the web archiving staff that decides what technical settings are necessary. Furthermore, as noted above content curators are also asked to contribute to quality assurance and relations with users regarding web archive collections. While this requires more training for the content curators, it was suggested that this extra investment in web archiving activity creates a greater sense of engagement in the work which serves to motivate and maintain the network of curators.

Communication and organization: management, content curators, IT staff

As suggested above, for web archiving to find its place in an institution a number of different actors, from management to different departments with different skills and priorities. It is therefore vital to communicate both with management and other actors within the institution, and particularly IT and librarians who may select sites. Frequently, the main role of the web archiving team, however it is constituted, will be to provide this communication and coordination to promote and make possible web archiving activity. The role of "digital curator" therefore requires a mixed profile with both librarian and IT skills and knowledge, but particularly the ability to talk in both "librarian language" when speaking with curators and management and "IT language" when speaking to IT staff. As web archiving activity evolves, the kind of communication needed, and therefore the kind of person involved, may change.

Communication with management is vital when starting a web archiving program, but remains so once it becomes "business as usual". At the beginning of a web archive program, it is necessary to send the message that preserving Internet content is important: often this can involve concrete examples (publications that no longer exist in print form), metaphors ("ephemera" collections such as store catalogues) and comparisons with other institutions or other countries. Over time the messages to management change, becoming concerned first about how best to perform web archiving in a way that is effective, and then once the workflow is in place, the important point becomes maintaining the collections, ensuring that the full cycle is sustainable. As previously stated, the experience at the BnF is that integration of web archiving both within library policy and within the existing organizational structure helps in creating a durable workflow.

At the BnF, the Digital Legal Deposit team has frequently the role of communication and coordination among the different groups that work on web archiving activity. This is important in everyday activity so that each group can see the overall picture, and the importance of their role in it. One example is the integration of the web archive collections with SPAR, the centralized digital repository of the BnF. The web archiving team (both Legal Deposit and IT staff) works with the multidisciplinary team of IT and preservation experts to define the specific preservation needs of web archive material, but also in a wider sense to ensure that the SPAR team understands the role and importance of web archiving. At the same time, SPAR can seem very technical and of interest only to the dedicated team working on it: as one participant remarked, "how much do content curators know or care about any of this?". But it is the the collections that the curators help to constitute that are held in SPAR, and it is important that the librarians understand how it works and what they can gain from it. The British Library also works with a centralized digital library system which handles the WARC files that hold the web archives, and is therefore faced with similar questions.

Another important aspect of communication was shown by some of the smaller web archiving teams represented at the workshop. Where only one or two people are involved in web archiving, there is a high risk that if staff leave everything will have to start from scratch. Integration in the organization and documentation of what has been put in place are important to prevent this; in the end, the ideal is to have a workflow that is "business as usual" and which is sufficiently robust to handle changes in personnel.

Technical aspects of web archiving and quality assurance

Although the workshop was based around organizational questions, many technical points were raised during the discussions, regarding the effectiveness of crawlers, or aspects of preservation and access. While such questions were not the focus of the workshop, it provided an occasion for colleagues to compare problems and seek solutions. It was also noted that problems with technical aspects create a need for communication within the institution, and therefore relate to the organization in place.

One topic which came up several times during the workshop, although not central to questions of organization, was quality assurance. The different ways of doing Quality Assurance seem to interest everyone, with questions such as who should do it, with what

tools, and how does one define "quality". Different approaches are applied in different institutions. At the BnF, the digital curators monitor the crawls to identify problems, and analyze the indicators after the crawls, while the visual QA of the sites collected is largely performed by the content curators, although the Digital Legal Deposit team will also check a sample of sites. The British Library, which performs QA on all sites before they are made available to the public, presented its new QA module which automatically identifies sites which may pose problems, based on a range of criteria (time taken to crawl, number of URLs compared to previous harvest, unusual URLs…). These sites are then the subject of manual QA by curators. An important step is the definition of indicators for web archiving by ISO, which should allow more standardized practice between institutions.

Other points regarding harvesting and access demonstrated that technical aspects are at the heart of everyday web archiving business, and that different institutions share similar questions:

- The evolution of Heritrix
- The harvesting of complex content
- How to harvest or collect subscription-based material, e-books…
- How to provide access in a meaningful way for users

With the context of the workshop, these technical aspects of course underlie many of the other questions discussed: the relation between curators and IT, the roles of different departments, communication with management to have resources, etc. The priorities in terms of technical development can in some cases be defined or imposed by the organizational or legal context: thus the BnF has worked extensively on its harvesting workflow to handle large amounts of data under its legal deposit obligations, while the British Library, working under a permission-based system, has concentrated more resources on access solutions.

All the participants agreed that improving harvesting of certain kinds of content is vital; while outside the scope of the workshop this shared experience may be of use to the IIPC in defining its priorities. Different solutions for providing access are also of interest to many institutions, whether it be full-text indexing, data and link mining or new forms of visualization. As suggested at the evaluation session, these and other issues could form the basis of subsequent IIPC-sponsored workshops.

Section 2: evaluation of the workshop

The evaluation of the workshop took place in two stages. There was an open discussion during Session 8 about the findings, lessons learned, tips to share and organization models so the organizing team and the participants could recap and draw conclusions from the workshop. After the workshop, an anonymous online feedback form was made available from 10th to 21st December 2013 and was used by 11 out of 14 participants.
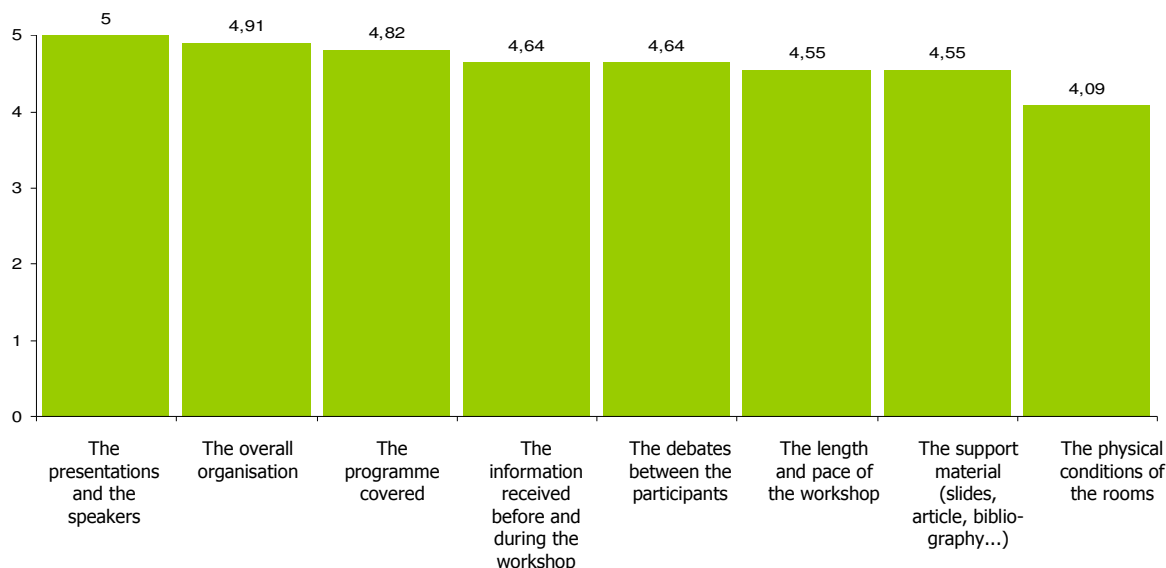
Summary of responses

The overall reaction from participants was very positive, both regarding program and content of the workshop and the practical organization. Regarding the content in particular, several participants noted that they had had doubts before attending about whether a workshop based principally on the experience of one institution could be valuable for other institutions with different legal and organizational contexts, but that in fact the information and discussions could be useful in different contexts. Participants also noted that though the program was dense, the sessions were organized in a logical manner that meant the information was easier to follow and absorb. One element that was particularly appreciated was being able to meet the actual people who perform different parts of the workflow, from content curators to IT.

Perhaps most importantly, the main negative point raised by several participants was the limited time for discussions between participants and presentations of the different situations in each institution. The organizers had planned to allow time at the end of each session for discussions but in practice the presentations took up most of the time. In future workshops care should be taken to ensure that sufficient time is set aside for this. One interesting idea from the feedback form was to ask each institution to present their web archiving activities; again this would be something to consider for future workshops.

Finally, during the wrap-up session participants were asked specifically to think of how future IIPC workshops could be organized. All were in agreement that such events were valuable to allow exchanges with colleagues, share experience, get new ideas and find renewed motivation. Rerunning a similar workshop with a wide-ranging program such as that presented by the BnF, but perhaps organized by a different institution, could be useful, for other institutions or colleagues who had not been able to attend; several participants said that such a workshop worked almost as a training session in web archiving. However it was also suggested that certain subjects could be treated more in-depth, either with dedicated workshops or as focused sessions within a general program, perhaps with parallel sessions for curators and technical staff. Suggested subjects included improving Heritrix and harvesting of complex content, improving access and research uses of web archives, collecting subscription-based content such as ebooks, and quality assurance. It was suggested that such dedicated sessions could lead to more practical recommendations and actions which could be taken up by IIPC working groups.

Examples of feedback from participants

Figure 1. Average ratings on the organisation of the workshop
Scale from 1 (very poor) to 5 (very good)

The following examples of different answers reflect the impact of the workshop and possible change in the way of participants' thinking and acting on the web archiving field at their own organizations.

What are your overall impressions of the workshop?

- The overall impression of the workshop is excellent. This goes for organization, content presented, materials handed out and the BnF web archiving team effort.
- The number of participants was well suited and favoured conversation and exchange.
- I now feel like I completed the training program that I never had for the work I am doing with web archiving.

What were your expectations before attending and did the workshop meet them?

- The agenda sent beforehand was very detailed and so I had a good idea of what to expect before I came.
- I was doubtful that a program based primarily on one institution's experiences could be that helpful since there are distinct differences in workflow, tools, and particularly the controlling legal regime, but this proved not to be a problem.
- My expectations involved achieving a better knowledge of the challenges behind web archiving activities, putting our initiative in perspective amid other international projects, dealing with organizational issues, networking and sharing. Not only did I meet them all,

but attained goals beyond my wildest prospects (a better understanding of how to integrate web archiving back home, arguments in favour of legitimizing the use of web archives for cataloguing and beyond, awareness of content policy issues, etc.).

- I was not expecting the sessions to be as tiring and intense as it turned out to be.
- The workshop exceeded my expectations in that the presentations were very clear and the notes that I took still make sense.

What were the main strengths of the workshop?

- Presentations by people who actually do the work that was presented (first hand information).
- Coverage of all different aspects of web archiving (seeing the whole picture, not just individual pieces of the puzzle).
- Not only presentations, but also practical demonstrations of certain tools and activities. Also the possibility to ask very specific and practical questions and get the answers.
- The agenda was great, well organized, flowed well and included variety of speakers and rooms.
- It was great that we were such a small group, it allowed for more informality so you were comfortable asking questions and were able to get to know fellow participants.
- Getting a deep overview of BNF from which I can extract what would be applicable to my organization, and also being introduced to other models.
- The self-assurance of the BnF and their motivating spirit; the careful chronogramatic planning of sessions, the combination of viewpoints: web archiving is not seen through the eyes of engineers as through the eyes of subject librarians. Every perspective counts, and the 360º combination you managed to offer was unique.

What were the main points that could be improved?

- I believe it would have been useful to receive the support material (slides, etc.) in advance, as that would have helped us in order to take more accurate and tidier notes.
- For one thing, as a compelling basis for selling the cost of the workshop that title didn't help me - to get funding from my management I cast it as "Training on how to support curators, in particular to make them more effective in their roles. Acquire valuable understanding for comparison/contrast with our practice from successful BnF program."
- A little bit more time between the sessions for more discussions and exchange between the participants.
- It would have been good if one person from each institution presented a 20 minute synopsis of their web archiving activities.

What benefits do you think the workshop will bring to your work personally?

- I think I can now better understand what it takes to implement a really effective web harvesting process in a (national) library (in terms of resources, communication, and infrastructure). That results in a more realistic view of what can be achieved in smaller institutions.

- Much greater confidence in what I'm doing.
- There are more people outside there which I can contact if I have questions concerning web archiving.
- I got a good example of how web archiving can be organized in a national library. This example confirmed my own views of web archiving organization and I will be able to proceed more confidently with development and integrating web archiving activities in my library.
- Main benefits refer to:
  - promotion and presentation of web archiving to top management
  - development of good cooperation between different profiles of employees working in different departments (collections, IT, web archiving)
  - ways to work with and motivate content curators from different departments
  - preparation of strategic and other necessary documents for web archiving
  - planning costs and resources for web archiving.
- More confidence in the way to do things. Motivation to bring the program further. Better knowing of the context of web archiving in the world.
- New faith in the importance of web archiving, solid argumentation to support my stance, grounding, a bond with colleagues beyond theoretical issues, on a personal, human-aware basis. We are sending a message in a futuristic bottle, and we're all on the same shore.
- I was able to make many new contacts at the workshop and have already been communicating with them on web archiving matters. This has been extremely beneficial as I now have a wider pool of knowledge and experience to draw on.

What benefits do you think the workshop will bring to your institution?

- More archived websites.
- Better organized workflows.
- Improved techniques/tools.
- Reviews of the different web archiving tools such as NetarchiveSuite and BCWeb has helped formulate ideas for improving our own workflows.
- Hearing the experiences of librarians and subject specialists at the BnF gave me inspiration to embed some of aspects of selection and QA into our future workflow with Legal Deposit. This will make the job of our own selectors a lot easier and more rewarding. I will be able to sell the benefits of Web archiving to colleagues outside of our team.
- Building web archive collections should be part of Collection Policy of National Library.
- It will help us to better integrate web archiving into our library and give it the place it deserves - not just as some exotic experimental work but as an important part of library's collections development.

Appendix 1: Workshop program

How to Fit In?
Integrating a web archiving program in your organization

IIPC-Sponsored Workshop
Bibliothèque nationale de France, Paris
26th – 30th November 2012

Program

Monday 26th November 2012

9:30am - Session 1: Introduction and overview
- Presentation of workshop program, objectives and participants' goals and expectations
- Overview of the BnF's missions and structure, focusing on legal deposit and acquisitions
- Overview of the BnF web archiving team, workflow and tools from selection to preservation: BCWeb curator tool (BnF Collections on the Web), NetarchiveSuite, Heritrix, access interface, SPAR digital repository (Scalable Preservation and Archiving Repository)

(Speakers: Clément Oury, Peter Stirling, BnF Web Archiving Team)

2pm - Session 2: Integrating web archiving in acquisition practices
- Presentation of the BnF network of subject librarians responsible for selecting websites
- Use case (a): integrating web archiving in the BnF Collection Policy Charter
- Use case (b): building and maintaining a cooperation network for the French Presidential and General Elections project

(Speakers: Clément Oury, Caroline Rives, Sophie Derrot)

Tuesday 27th November 2012

9:30am - Session 3: Tools, tutorials & training for web archiving
- Presentation of the day-to-day work of BnF subject librarians and the tools they use:
  - Selection of sites using BCWeb
  - Quality control and reader access using in-house Wayback Machine
- Internal training programme to develop web archiving skills and motivation within the Library
- Presenting web archives to the end users in reading rooms
- Open discussion on training strategies, evaluation and tools.

(Speakers: Annick Lorthios, Peter Stirling, Marguerite Sablonnière, Françoise Jacquet)

2pm - Session 4: Web archiving for decision-makers: Internal Advocacy & Organization
- Promotion of Web archiving towards top management
- Internal organization issues: departments, units, networks and boards
- External organization issues: cooperation with other communities and institutions (national and international)
- Choice of NAS and development community issues

(Speakers: Gildas Illien, Sara Aubry)

Evening (optional): dinner (paid for by IIPC)

<u>Wednesday 28th November 2012</u>

9:30am - Session 5: integrating Web archiving in IT operations
- Definition of responsibilities between legal deposit and IT departments
- Management of daily operations and infrastructure

(Speakers: Annick Lorthios, Bert Wendland, Sébastien Pivain)

2pm - Session 6: integrating Web archiving in preservation workflows
- Integrating web archives preservation in digital preservation standards: the standardisation of the WARC format
- Presentation of the SPAR digital Repository
- Integration and preservation of the Web archives in SPAR (data model, preservation policy, costs, risks)

(Speakers: Sébastien Peyrard, Louise Fauduet, Clément Oury, Sophie Derrot)

<u>Thursday 29th November 2012</u>

9:30am - Session 7: presentation by Helen Hockx-Yu, British Library
- Helen Hockx-Yu, Head of Web Archiving at the British Library, will share her experience with the group to put the BnF approach into perspective

2pm - Session 8: conclusions and perspectives
- Open discussion with participants: findings, lessons learned, tips to share, organization models
- Perspectives and future projects at the BnF
- Recap and conclusion of the workshop

(Speakers: BnF Web Archiving Team)

<u>Friday 30th November 2012</u>

Morning
- Presentation by Internet Memory Foundation

Afternoon
- Presentation by l'Institut National de l'Audiovisuel (INA, the National Audiovisual Institute)

Choice of other sessions and visits depending on participants' interests to be arranged as required.

Appendix 2: Participants and speakers

Participants

Name	Institution
Martin Cremer	National Library of Germany
Helen Hockx-Yu	British Library
Nicola Johnson	British Library
Janko Klasinc	National and University Library of Slovenia
Jaanus Kõuts	National Library of Estonia
Matjaž Kragelj	National and University Library of Slovenia
Anna Kugler	Bavarian State Library
Rosalie Lack	California Digital Library
Mireille Laforce	National Library and Archives of Québec
Icíar Muguerza Lopez	National Library of Spain
Magdy Nagi	Bibliotheca Alexandrina
Michael Neubert	Library of Congress
Mar Pérez Morillo	National Library of Spain
Victor-Jan Vos	National Library of the Netherlands

BnF staff

Name	Department
Sara Aubry	IT
Sophie Derrot	Legal Deposit
Louise Fauduet	Conservation
Nicolas Giraud	IT
Gildas Illien	Digital and Bibliographic Information
Françoise Jacquet	Literature and Art
Annick Lorthios	Legal Deposit
Clément Oury	Legal Deposit
Sébastien Peyrard	Digital and Bibliographic Information
Sébastien Pivain	IT
Géraldine Raison	Legal Deposit
Caroline Rives	Collections Coordination
Marguerite Sablonnière	Music
Peter Stirling	Legal Deposit
Bert Wendland	IT