



**FOR AN INTERNATIONAL
CONCERTED POLICY
OF LABOUR HISTORY
ARCHIVES DIGITISATION**

IALHI · PARIS · 14th february 2009

FOR AN INTERNATIONAL CONCERTED POLICY OF LABOUR HISTORY ARCHIVES DIGITISATION.

This Workshop, organized by the IALHI, is devoted to the various digitisation policies of its members. Its objectives are to coordinate at a European level the various digitisation policies, and to look for the creation of a digital library. Another aim is to define the future IALHI policy.

The morning's contributions will be about digitisation policies of traditional paper archives.

The afternoon will be more directly devoted to born digital archives.

Contributions are expected to open a debate.

It will be possible to speak in English or French.

The workshop will be really interesting only if there are many IALHI participants and if the debate leads us to effective new IALHI orientations and projects. Lieu : Centre d'histoire sociale du XXe siècle

9, rue Malher, 75 004 Paris

Date : 14th February 2009

Morning (9.30-12.30)

9.30 : General Introduction

Digitisation programs of labour history archives

9.45-11.30 :

1. Marien van der Heijden (International Institute of Social History, IISG) :

Netherlands : Digitisation policies in Netherlands : about IISG

2. Urs Kälin (Schweizerisches Sozialarchiv)

Switzerland : Networks and digitisation policy in Switzerland

3. Rüdiger Zimmermann (Friedrich Ebert)

Deutschland : Deutschland and Friedrich Ebert Stiftung's policy

4. David Bidussa (Feltrinelli)

Italy : Italian digital resources

5. Françoise Blum (CHS) and Franck Veyron (BDIC)

France : Digitisation and labour history in France

Pause

6. Maarten Savels (AMSAB-Instituut voor Sociale Geschiedenis)

Belgium : Digitisation in Belgium

7. Richard Temple (Senate House Library)

United Kingdom : United Kingdom policies

8. Lars Ilhshammar (ARAB-Arbetarrörelsens Arkiv och Bibliotek)

Sweden : Swedish policies

9. Andrew Lee (New York University Library)

United States : Some American examples

11.30-12.15 : Discussion : Examples of programs managed by other participants, and discussion about the role the IALHI could play in the general policy of digitisation of labour history archives (by gathering information about different local or national programs, giving them publicity on labourhistory.net, creating links to each database of digitised documents, etc.)

12.15-12.30 : Conclusions

12.30 - BUFFET

Afternoon (2 pm-6 pm)

2.pm • Panel

Problems and perspectives of born digital archives

- 1 - Aline Girard (BNF) : *About Gallica*
- 2 - Rudolf Schmitz (Friedrich Ebert) : *Web archiving*
- 3 - Gildas Illien et Lucie Fléjou (BNF) : *Web archiving : national and international perspectives*

3.pm • Discussion (in English and French)

Different themes could be discussed.

- *Creation of a Labour History Web portal*
- *Advice about technical guides and protocols or software (free ones like Heritrix for instance)*
- *Constitution of a digital library*
- *Possibilities of financial plans (European Community etc)*
- *Aspects of scientific policy (Research boost, knowledge diffusion, etc) links with scientific community (creation of a mailing-list about labour history archives with IALHI members but also with social historians)*

Pour une politique internationale concertée de numérisation des archives du monde ouvrier et social

Cette journée, organisée par l'IALHI, sera consacrée aux politiques de numérisation des différentes institutions européennes spécialisées dans le domaine de l'histoire ouvrière et sociale (Labour History). L'enjeu en est de coordonner les diverses politiques de numérisation au niveau européen, et d'envisager la création d'une bibliothèque numérique spécialisée. Il est aussi de contribuer à mieux définir les futures politiques de l'IALHI.

Les interventions seront faites en français ou en anglais.

Cette journée de travail ne sera réussie qu'à condition de la participation la plus large possible des membres de l'IALHI. Venez nombreux !

Lieu : Centre d'histoire sociale du XXe siècle
9, rue Malher, 75 004 Paris

Date : 14 février 2009

Matinée (9 h 30-12 h 30)

9 h 30 : *Introduction générale*

Table ronde :

Politiques de numérisation des archives en histoire ouvrière et sociale (labour history)

9 h 45 -11 h 30 :

1. *Pays-Bas : Politiques de numérisations aux Pays-Bas : le cas de l'Institut international d'histoire sociale d'Amsterdam* par Marien van der Heijden (IISG)
2. *Suisse : Mise en réseau et politiques de numérisation en Suisse* par Urs Kälin (Schweizerisches Sozialarchiv)
3. *Allemagne : L'Allemagne et le rôle de la Fondation Friedrich Ebert Stiftung* par Rüdiger Zimmermann (Friedrich Ebert Stiftung)
4. *Italie : Les ressources numériques en Italie* par David Bidussa (Feltrinelli)

5. *France : Numérisations et histoire ouvrière et sociale en France* par Françoise Blum (CHS) et Franck Veyron (BDIC)

Pause

6. *Belgique : Numérisations en Belgique* par Maarten Savels (AMSAB)

7. *Royaume-Uni : Politiques de numérisation au Royaume-Uni*, par Richard Temple (Senate House Library)

8. *Suède : Politiques suédoises* par Lars Ilshammar (Arab)

9. *Etats-Unis : Quelques exemples américains* par Andrew Lee (NY University)

11 h 30-12 h 15 : *Discussion* : Autres exemples de projets ou réalisations et débat sur le rôle que pourrait jouer l'IALHI dans le domaine des numérisations des archives de l'histoire ouvrière et sociale (coordonner l'information, la diffuser sur labourhistory.net, créer des liens vers les bases existantes etc.)

12 h 15-12 h 30 : *Conclusions de la matinée*

Buffet

Après-midi (14 heures-18 heures)

14 heures • *Table-ronde*

Problèmes et perspectives dans le domaine des Born digital archives

1 - Gallica, par Aline Girard (BNF)

2 - *Le L'archivage du web à la Friedrich Ebert* par Rudolf Schmitz

3 - *Archivage du web* par Gildas Illien et Lucie Fléjou (BNF)

15 heures • *Débat* (langues anglais et français)

Divers axes de discussion seront proposés.

- *Création d'un véritable Portail en labour history* :

- *Diffusion de Protocoles*

- *Proposition d'Outils à diffuser* (Logiciels etc. par exemple Heritrix...) *et conseils techniques à donner*

- *Aspects de politique scientifique (Impulsion à la recherche, diffusion des savoirs, réflexions sur les archives) et liaisons avec la communauté scientifique (créer une liste de discussion autour des archives avec les membres de l'IALHI mais aussi des historiens sociaux)*

Introduction générale

Je voudrais d'abord remercier tous ceux qui sont ici, dont certains ont fait le voyage de loin. D'autres sont parisiens, les membres du réseau français Codhos sont venus nombreux, ainsi que des représentants de la Bibliothèque nationale de France, et j'espère que les contacts qui pourront se nouer aujourd'hui vont fructifier à l'avenir.

Cette journée se place sous l'égide de l'IALHI, qui est le seul réseau mondial d'archives, bibliothèques et centres de documentation, dans le domaine de la Labour History. Nous réunissons toutes les institutions qui « comptent » dans ce domaine, tant par leur rayonnement international que par leur taille, leur réactivité et leur dynamisme. J'en profite pour remercier tout particulièrement l'Institut international d'histoire sociale d'Amsterdam, qui est représenté ici par notre ami de longue date, Marien et par son encore nouveau directeur, Erik-Jan Zürcher. J'ai fait avec plaisir la connaissance d' Erik-Jan Zürcher lors du passionnant workshop organisé par l'Institut en hommage à notre ami et ancien directeur de l'IISG, Jaap Kloosterman. Nous reparlerons aujourd'hui de bien des questions évoquées lors de ce workshop. Je tiens en tout cas à dire ou redire aussi à quel point l'Institut a pu être essentiel à l'IALHI, grâce à une étonnante science (ou prescience si l'on considère les actuels bouleversements), des enjeux internationaux. Jaap Kloosterman fut notre président discret et efficace. Il savait parfaitement que, aussi vrai qu'il ne peut y avoir de socialisme dans un seul pays, il ne peut y avoir de politique archivistique dans un seul pays. À côté de ces institutions « locomotives », nous avons toute une série de membres, moins visibles peut-être, qui effectue un travail remarquable pour la préservation du patrimoine des mouvements ouvriers et sociaux, avec les moyens du bord. Nous avons aussi, plus récemment, de nouveaux membres de pays « extra-occidentaux », où l'argent est rare, mais dont le cœur à l'ouvrage compense les difficultés financières : des pays d'Afrique ou d'Amérique latine. L'IALHI, plus que tout autre réseau, et ne serait-ce qu'en mémoire du vieil internationalisme ouvrier, se doit d'être global, dans un monde où les processus de domination s'inscrivent dans des logiques aussi bien horizontales, celles des rapports Nord-Sud, que verticales, celles de classe, de genre et d'ethnies.

À l'origine de cette journée, il y a la conférence de Johannesburg. Cette conférence fut une très belle conférence, et j'en remercie nos amis sud-africains. Dans un pays encore bouleversé par les lois de l'Apartheid (et nous étions réunis dans un ancien compound devenu non sans mal worker museum), tous les présents ont pu, grâce à nos hôtes, mieux comprendre les problèmes mémoriaux spécifiques à l'Afrique du Sud, et le rôle tout à fait particulier que peut y jouer l'archive. En Afrique du Sud, en un temps étonnamment proche, les promoteurs de l'Apartheid

ont voulu non seulement réorganiser l'espace selon les lois de la suprématie blanche, en expropriant et éloignant du cœur des villes leurs habitants de couleur mais ils ont voulu aussi éradiquer toutes traces de cette présence, pour fabriquer, en quelque sorte, une légitimité, la leur, en faisant table rase de toute histoire et de tout passé. Dans ces conditions, les descendants de ces expropriés visent à juste titre comme une nécessité, la restauration de cette mémoire que le pouvoir totalitaire de l'Apartheid a voulu définitivement éradiquer. Et c'est évidemment les archives, comme les vestiges archéologiques, qui sont les preuves, les pièces à conviction dans le long procès de l'Apartheid. Il va donc sans dire que les workers museums qui fleurissent en Afrique du Sud, et qui ont quelque chose à voir dans leur conception avec nos éco-musées, sont les nombreux jalons de cette histoire vivante, en train de se faire, qui est tous les jours un combat.

Nous n'étions malheureusement pas très nombreux (17 représentants de l'IALHI), ce qui était, évidemment, lié à des problèmes de coûts du voyage, et les présents ont déploré cette faible participation, au vu notamment de l'extrême intérêt de cette conférence. Nous avons donc pensé qu'il serait souhaitable d'organiser une journée, à l'occasion du bureau de l'IALHI, qui se tient toujours en février à Paris, qui permettrait à certains membres de l'IALHI de venir, alors qu'ils n'avaient pu être là en Afrique du Sud, journée qui serait une sorte de conférence médiane et allégée entre deux conférences, entre Johannesburg et Barcelone, qui est, comme vous le savez, notre prochaine destination. Journée également dont les résultats pourraient être exploités dans les futures conférences, et les prémices développées.

Nous avons d'autre part beaucoup discuté entre nous, à Johannesburg. Il semble que nous ayons tous conscience que l'IALHI est arrivé à un tournant de son histoire. Nous vivons actuellement une véritable révolution, la révolution numérique, qu'aucune institution, aussi puissante soit-elle, ne maîtrise totalement. Devant cette révolution numérique, de très nombreuses erreurs ont déjà été commises. Nous tâtonnons tous. Il y a néanmoins quelques certitudes : le temps du bricolage est d'une certaine manière finie. Cette révolution intervient dans un monde globalisé et est le fruit d'un monde globalisé. Elle offre des possibilités immenses mais elle a un coût. Elle rend indispensable les coopérations et les alliances à tous niveaux, local, national, régional et mondial. Dans ces conditions, le rôle des réseaux est plus que jamais important. Mais le rôle de ces réseaux doit être aussi repensé, à la lumière de cette révolution. Sinon, ils sont condamnés à mourir.

Il n'est pas question d'abandonner ce qui faisait l'efficace de ces réseaux, c'est-à-dire les échanges d'information et de savoir, les aides réciproques et la mutualisation, et bien sûr, je tiens à le rappeler, les longues amitiés qui sont nées de préoccupations, voir de valeurs communes. Mais il faut changer d'échelle. Il est sans doute nécessaire aujourd'hui de formaliser davantage ces échanges, de les inscrire aussi dans des analyses rigoureuses de la situation dans laquelle nous

vivons. Il faudrait que chacun se demande ce qu'il peut faire seul, et ce qu'il ne peut pas faire seul, il faudrait que chacun explicite clairement ses besoins, considère avec attention les aides financières de proximité qu'il peut obtenir et celles qu'il ne peut pas obtenir. L'argent existe mais il est vrai aussi que l'histoire ouvrière n'a pas toujours le vent en poupe. Il est vrai aussi qu'il faut déjà des savoirs experts pour formuler clairement les demandes, savoirs experts dont toutes les institutions ne disposent pas en interne. Beaucoup d'entre nous « s'arrachent les cheveux », pour parler vulgairement, pour savoir ce qu'ils vont faire pour archiver leurs courriers électroniques, ou leurs sites web. Faut-il refaire ce que fait Internet Archives ou ce que font les bibliothèques nationales ? Beaucoup d'entre nous « s'arrachent les cheveux » pour savoir ce qu'ils doivent d'abord numérisés, puisqu'on ne peut évidemment tout faire à la fois : livres, brochures, périodiques, images, archives etc. et s'ils numérisent, ils ne sont pas toujours sûrs de le faire selon des normes correctes, d'autant plus que ces mêmes normes sont en constante évolution. Maîtrisent-ils les standards OAI, l'EAD, la norme Isad (G) ? Or l'adoption des normes et des standards est beaucoup plus importante que lorsque régnait encore le seul papier, car pour exister dans le monde numérique, il faut être visible, et pour être visible, il faut partager et pour partager se plier à certaines règles. Comment se situer aussi, quand on est par exemple une petite institution militante par rapport aux grands catalogues nationaux en construction : la bibliothèque numérique de Google, Gallica pour la France, Europeana pour l'Europe etc. Pas question de refaire un travail déjà effectué mais par contre comment rendre visibles certains fonds noyés dans la masse ? Comment créer de la valeur ajoutée aux numérisations massives ? Comment aussi stocker ces masses numériques, avec quelle garantie de pérennité ? On se dirige de plus en plus vers des solutions type portail. Les portails en tout cas se multiplient, à en donner le vertige, ainsi que les réservoirs numériques, et l'on va vers des solutions de stockage partagé. D'un autre côté, alors même que les savoirs techniques deviennent de plus en plus nécessaires et sophistiqués, l'ancienne division archiviste/historien/chercheur a paradoxalement, me semble-t-il, beaucoup moins lieu d'être. Chaque savoir spécifique est bienvenu à l'heure du numérique, et se doit d'être versé dans le pot commun.

J'ai été par ailleurs fort étonnée en lisant les textes des membres du Codhos qui ont bien voulu me les envoyer (et je les en remercie) du nombre d'initiatives en matière de numérisation. Bien que je fréquente les membres du Codhos régulièrement et que nous ne soyons pas avares dans l'échange d'informations, je ne connaissais (et je ne suis sûrement pas la seule dans ce cas-là) que la partie émergée de cet iceberg numérique, ce qui pose bien sûr le problème de la visibilité ou lisibilité de ces multiples travaux et programmes. J'ai aussi appris beaucoup avec les textes de nos collègues de l'IALHI. Je connaissais certaines choses, grâce à nos conférences antérieures mais sûrement pas tout. Ne serait-ce que grâce à ce que

nous allons entendre et apprendre aujourd'hui les uns des autres, cette journée était nécessaire.

Avant de passer à l'organisation de cette journée, je voudrais vous donner l'exemple du partenariat que le réseau Codhos a mis en place avec la bibliothèque nationale de France. Le réseau Codhos bien sûr est membre de l'IALHI. Nous avons commencé très modestement, par un travail de recensement exhaustif des congrès du monde ouvrier, travail de recensement qui n'était possible que grâce au réseau, les divers documents étant dispersés. Nous avons ensuite passé une convention avec la BNF qui s'est chargée de numériser les documents du corpus qu'elle ne possédait pas et d'en donner une copie numérique à l'institution propriétaire. Nous avons dans la foulée construit une base des congrès avec évidemment des liens sur les documents numérisés dans Gallica. C'était une première expérience que nous sommes en passe de poursuivre, que nous espérons vivement poursuivre, le Codhos devenant le partenaire de la BNF en général et de Gallica en particulier sur ce que je vais appeler pour faire court la « question sociale ». Nous discutons actuellement du choix des premiers corpus thématique comme la Commune de Paris, ou la presse socialiste. Nous en reparlerons sûrement au cours de cette journée, puisque nos interlocuteurs de la Bibliothèque nationale ont eu la gentillesse de venir aujourd'hui, en la personne d'Aline Girard, responsable de Gallica et de son collègue Julien Gueslin. Une autre coopération s'est aussi mise en place avec la BNF sur l'archivage du web. Gildas Ilien y reviendra sûrement cet après-midi, comme il reviendra sûrement sur le fait que la Bibliothèque nationale de France, dépositaire légale du web.fr, travaille au sein d'un consortium international. En matière de web aussi, les partenariats et réseaux sont la règle indispensable. Et Gildas pourra confronter sa propre expérience avec celle de la Fondation Friedrich Ebert, dont parlera Rudolf Schmitz.

Ce matin, nous allons avoir toute une série d'interventions, portant sur les numérisations d'archives, archives étant pris dans un sens extensif. Nous avons pensé au départ faire un rapport général à partir duquel s'engagerait la discussion mais tous les textes n'étant pas parvenus à temps, cela s'est révélé impossible. Chacun interviendra donc, mais pas plus de 10 minutes, pour que nous puissions engager la discussion dans l'heure qui nous restera avant le buffet. Les interventions se feront selon la volonté des participants en français et en anglais. Quand les textes nous étaient parvenus à temps, nous en avons fait une traduction que vous pouvez trouver dans le petit volume qui vous a été remis. Marien avait proposé un très pertinent schéma d'intervention que j'ai répercuté aux différents intervenants : Que numérisons-nous ou qu'avons-nous le projet de numériser, dans quel projet plus général s'insèrent ces numérisations, à quels problèmes nous heurtons-nous, quels partenariats avons-nous ? Nous allons sûrement voir que beaucoup de problèmes sont les mêmes d'une institution à l'autre.

Pour conclure, je voudrais insister sur le fait que cette journée est essentielle pour l'IALHI, et que beaucoup de choses vont se jouer dans les discussions que nous allons avoir en fin de matinée et surtout en fin d'après-midi. Il s'agit de construire le nouvel objet du travailler ensemble dans le réseau, nouvel objet qui, quel que soit le différentiel des participations à sa construction devra être utile à tous. Certains ici, je le sais, ont les idées très claires et très précises sur ce qu'il faudrait faire, idées claires liées à une excellente compréhension de la donne internationale. Mais au-delà des idées, il faudra aussi, incontestablement, se donner de nouveaux moyens.

General introduction

Firstly I would like to thank you for being here this morning, especially those of you who travelled from abroad. Many of us are from Paris, and CODHOS's members are numerous. Some colleagues from the French national Library are also present. Some of you are meeting each other for the first time : I hope that today's discussions will be fruitful in the future.

This workshop is organised by the IALHI, which happens to be the only world network of archives, libraries and documentation centres specialised in the field of labour history. All the important institutions — because of their holdings, their work and their activities — belong to our association. I would like to thank our two colleagues from the International Institute of Social History for being with us : one of them, Marien, is a long-time friend and actor of the IALHI ; the other one is the new Director of the Institute, Erik-Jan Zürcher. I had the chance to meet him for the first time last December, when the Institute organised a very interesting international workshop to pay tribute to the work of our friend and previous director of the Institute, Jaap Kloosterman. Many of the questions that have been raised in December will be discussed again today. But I would like to underline how much the role played by Amsterdam Institute has been important for the IALHI, thanks to its very striking science of international politics. Jaap Kloosterman was at a time the general secretary, discrete and efficient, of our association. He knew that "socialism in one country" was a wrong theory ; but he also knew that archival policies have to be international.

Beside the big leading institutions, numerous other archive centres belong to the IALHI. Sometimes they are not so well-known ; still they are doing a very good job for the preservation of labour and social heritage, whatever are the means they can rely on. New members from non-western countries (Africa and South America) recently joined us ; money is lacking, but their activism is stronger. As a tribute to the old international labour movement, the IALHI, more than any other network,

has to be a global one, in a world where the domination processes are not only horizontal — North-South relations — but also vertical (class, gender or ethnic relations).

The idea of today's workshop came out last September during our conference in Johannesburg. This conference was a real success, and I would like to thank our South-African colleagues and friends. In a country that still feels strongly the effects of the Apartheid laws — our meetings took place in a previous workers' compound, that has been kept and transformed, not easily, in a workers' museum —, the IALHI members who attended the conference could understand, thanks to its organisers, the specific characteristics of heritage and memory questions in South Africa, and the role archives can play in that country. Apartheid promoters wanted to reorganise the space according to the laws of white supremacy. They expropriated the coloured inhabitants from their houses in inner cities areas, and obliged them to live in townships in the towns' outskirts. They managed to eradicate all traces of their presence, in order to get rid of the history and the past, and to build their own legitimacy. Considering that context, one can understand why the children of the people who have been pushed out of their houses are particularly interested, today, by the heritage and the memory the Apartheid regime wanted to erase. "Archaeological" remains of the pre-Apartheid era and archives can be used as pieces of evidence, exhibits, in the long trial against the previous political regime. There are numerous newly workers' museums in South Africa — museums that are not far from our western European heritage museums, or, in France, "écomusées". All contribute to the living and still developing history of contemporary South Africa, which is the result of a day after day struggle.

Unfortunately, mainly because of the travelling costs, only 17 members of the IALHI were represented in Johannesburg. So we thought that it could be a good idea to organise a day meeting in Paris at the time of the next coordination committee of the association. It's usually not such a problem, for IALHI members, to come to France, so members could have the opportunity to meet each other without having to wait for next September and the coming conference in Barcelona. We also thought of today's workshop as a meeting which conclusions could be used, developed and put in practice from next September.

We spoke a lot in Johannesburg, and we all shared to same statement : the IALHI is coming to a turning point of its history. We have to deal with a revolution — the digital revolution — that none of us, even the biggest institutions, can face alone. Some mistakes have already been made. We are all struggling to find, step by step, the right answers. Nevertheless we now have some precise ideas about what need to be made, and how : the time of do-it-yourself solutions is over. The digital revolution happens in an era of globalisation, and it is the result of globalisation. It offers a wide range of new possibilities, but its costs a lot. It

makes cooperation and alliances necessary, at all level (local, national, regional or worldwide). More than ever, networks have a key role to play. But their aims need to be redefined in the light of this revolution. Otherwise they will have no other future than death.

There is no reason to give up what used to be the core of the efficiency of our networks : the exchanges of information and knowledge, the mutual help, and of course — and I want to insist on that point —, our long friendships based on the share of common professional worries or even common values. But now we have to organise ourselves and work on a different scale. Our cooperation needs to be arranged more precisely, and must be the result of a close analysis of the present situation. We all have to think of what we can do in our institution, and what cannot be done without cooperation. We all have to define clearly our needs, and consider which financial help we can get by ourselves, or which are the credits we will not be able to obtain without presenting collective programs. The money exists somewhere, but it's true that labour history is not precisely fashionable these days. It's a fact, too, that we have to develop expertise in order to articulate our needs and present well-argued demands.

Many of us still panic about the saving and preservation of e-mails or websites. But do we have to do again by ourselves what is already made by Internet Archive or by some national libraries ? Many of us don't know what to decide when they have to choose which types of their documents (books, pamphlets, journals, photographs, archives, etc.) must be digitised more urgently ; they also wonder about changing protocols and standards. Do they know about OAIS (Open Archival Information System), EAD (Encoded Archival Description) or the ISAD(G) norm (General International Standard Archival Description) ? The respect of such standards is more important now than the one of norms used to be at the time of paper archives : in the digital world, you need to be visible to exist, and to exist you need to share your data, i. e. to respect the rules that allow their exchange.

If you belong to a small institution, with a political background, for example, other questions can be raised. How can you keep your originality and your identity when you have to deal or share documents with huge national or international catalogues like Google digital library, Gallica in France or Europeana in Europe ? There is no reason, of course, to do again a job that has already be done, but what can be made to facilitate the location of some special collections swamped by a mass of other digital documents ? How can we give an added value to massive digitisation programs ? Besides, how can we take into consideration the long-time preservation of digital holdings, with safety guarantees ? The "portal" programs are more and more numerous, like digital repositories projects, and it seems that the future will be based on shared storage solutions. Technical knowledge is getting more and more complicated and necessary but even if it can sound paradoxical, I think that the old division between archivists, historians and

researchers does not have so many reasons to exist. Each expertise should be shared and contribute to a common knowledge.

When I read the texts some CODHOS' members sent me to prepare our workshop - and I must thank them for kindly responding to my questions -, I've been surprised by the number and variety of their digitisation projects. We meet rather frequently, and share a lot of information, but I must say that I just knew the small emerged part of that digital iceberg - a fact that shows how difficult it can be to give visibility to our projects, and make them understandable. I also learned a lot by reading the contributions of our IALHI colleagues. In that case too, I already knew some things, thanks to our previous annual conferences, but certainly not the whole range of their digitisation activities. Considering what we are going to hear and learn today, this workshop was necessary.

Before telling you a few words about the practical organisation of the day, I would like to give some information about the partnership the CODHOS has established with the French national library.

We began our collective work on a small scale by an exhaustive inventory of the French labour movement congresses — an inventory that could only be the result of a cooperative work, since the French labour movement archives and printed materials are spread among different institutions. When the inventory was finished, we signed an agreement with the national library, which undertook the digitisation of the documents, using their collections or documents lent by CODHOS's members — in that second case, at the end of the process digital copies were given to the lending institutions. This was a first experience that now is certainly going to be continued — we greatly hope so —, as the CODHOS could become the major BnF partner, especially for the development of Gallica 2, in the field of "social questions", to say it quickly. At the moment, we have discussions about which subject corpus could be chosen first, printed materials and posters about the Paris Commune or socialist periodicals. We will certainly have time today to discuss that partnership, since colleagues from the National library, Aline Girard, in charge of the Gallica project, and her colleague Julien Gueslin, kindly accepted to join us.

We also began another partnership with the National library for web archiving. Gildas Illien will certainly speak of it this afternoon, as he will explain the cooperation of the BnF, which manages the legal deposit of French websites, with an international consortium. He will have the chance to confront the French experience to the German one, or at least the Friedrich Ebert Foundation's one, which will be presented by Rudolf Schmitz.

Different archive digitisation programs — archive being taken in its wider definition — will be presented successively this morning. We thought of writing a general summary report which could have been the basis of our discussions, but we didn't receive all the texts soon enough. So people will speak one after another. I

would like you to take no more than 10 minutes for your communication, so that we will have enough time — around one hour — for a debate before lunch. You can speak in French or in English. When we received written texts, we managed to do little translations that you can find in your little leaflet. Excuse us for not being professional translators. Marien proposed a scheme for the communications, which I forwarded to everybody. These 4 points were : 1. Which of your collections are you digitising ? Which ones are you planning to digitise ? Which projects or programs do you have in connection with digitisation (storage, access, etc.) ? 2. With whom do you cooperate, at national or international level ? Which national or international funding possibilities do you have ? What are the advantages or disadvantages of these funding schemes ? 3. What are your main 'problems' and limitations (technical, conceptual, financial, organizational) ? 4. Which opportunities do you see within IALHI ? Which possibilities for future cooperation ?

I guess some remarks and problems will be the same from one institution to the other.

As a conclusion, I would like to insist on the fact that today's workshop is very important for the future of the IALHI. One aim of our discussions this morning and this afternoon will be to define altogether new objectives for our collective and mutual work — and whatever will be the future involvement of each institution, these objectives must be useful for everybody. Some of us, I know, already have clear and well-defined ideas, based on an accurate understanding of the international situation. We need ideas, but we will also have to find means to make them concrete.

Digitisation policies in Netherlands : about IISG

Marien van der Heijden

(International Institute of Social History, Amsterdam)

What are we digitizing / are we going to digitize

Archives: digitization on project basis, dependent on funding. In 2009 the archives of Dora Russell (20 m.) and Ferdinand Domela Nieuwenhuis (5 m.) will be digitized within the « Metamorfoze program » (national programme for the preservation of paper heritage).

Books, brochures, serials: digitization on project basis, dependent on funding. In 2009, the International Review of Social History will be digitized (1936-..., over 25.000 p.)

Images: photos, posters, prints etc. are digitized on a high quality (100%, 300 DPI) within the normal workflow. The master files are stored, low resolution images are generated for access through the catalogue. Certain collections are digitized within projects, dependent on funding (e.g. 5.300 posters from Dutch alternative youth centers).

Film, video, sound : in 2009 a pilot project will be carried out to digitize 100 hours of sound material, 50 hours video and 10 hours film. The results will be used to draw up a project proposal to take on the digitization of the complete collection (15.000 hours of sound material, 8,000 hours of video, 2.000 hours film).

The 'Centrale project': in 2009 a project will start to digitize the collections bought before 1940 by 'De Centrale', the workers' insurance company that funded the IISH in its early years. These collections include 180 m. of archives (Michail Bakunin, Karl Marx, Max Nettlau, Leo Trotski, Lucien Descaves, Louise Michel, Eduard Bernstein, Karl Kautsky, etc.) and important collections of books and serials. 2009 will be the 'pilot year'. Within this project, organizational and technical procedures will be developed for digitization, permanent storage of and access to digitized material. These include workflows within the organization, quality standards and control, permanent identifiers, data exchange protocols. The project is funded by the SNS Reaal insurance company, successor to 'De Centrale'.

Projects/programs in connection with digitization

Repository: the IISH has developed in-house facilities for the permanent storage of large volumes of digital files (digital born or digitized). These are based on the OAIS-standard for Trusted Digital Repositories. The basic ingest and storage

functions of the 'TDR' are operational, administrative functions and access are under development.

Access: the search platform (search.iisg.nl) that already provides access to the catalog of books, brochures, serials and image collections, archive inventories and other metadata will be developed further to provide access to digitized collections.

Web 2.0 platform: in 2009-2010 a platform will be built that provides new forms of presentation of and interaction with digitized collections. As to presentation, visualization of collections on timelines and maps will be made possible. As to interaction, digitized material will be uploaded to popular 'content pools' or 'social sites' such as Flickr for photos and Youtube for video, where the public can reuse and enrich the items. The platform will provide functionalities to 'embed' these items in web presentations made by users, to add other items, texts and other documents to these presentations. Also, the platform can be used by 'communities' of users to add their knowledge and - for instance - enrich the metadata and add background information. The material digitized in the 'Centrale project' will be used for pilot projects in this context.

HITIME: a four-year project will start in 2009 to develop a 'software toolkit' for historical text mining. The software should automatically recognize 'entities' such as names of people and organizations, geographical locations and dates, and relations between these entities, and present them as an associative network, linked to a timeline. This should enable researchers to obtain answers to questions that now require too large amounts of human effort. The IISH's large collection of primary sources, secondary sources and metadata will be the 'case study' for this research project, mainly carried out by the Tilburg centre for Creative Computing, Tilburg University. The results will be published as open source, and should be applicable to multilingual sources.

With whom do we cooperate (national/international)

On a national level, the IISH is mainly active in the exchange of information and participation in joint forums with a host of relevant institutes and organizations, such as the Royal Library, National Archives, Netherlands Coalition for Digital Preservation, etc. A very practical form of cooperation is developing with the IIAV (information center and archive of the women's movement), that will move in with the IISH by 2011. The existing digital infrastructures will have to be 'merged' to a large extent.

Internationally, the IISH is in contact with OCLC/RLG Special Programs and with a number of EU-bodies in the field of digitization and 'digital heritage'. The IISH participates in OCLC's Worldcat and Archivegrid. With the EU, a first pilot to participate in the Europeana 'portal' is planned for late 2009.

Which national/international funding possibilities do we have

The IISH is reasonably successful in getting a regular flow of projects financed by various national programs, such as *Metamorfoze* (digitization of 'paper heritage') and *'Geheugen van Nederland'* (images). An attempt will be made to tap into the national program for the digitization of audiovisual material too. The 'Web 2.0 platform' project is funded by the national program for 'concerted digitization'. The HITIME project is funded by the 'CATCH' program of the Netherlands Organisation for Scientific Research.

In the private sector, our most important funder is the SNS Reaal fund. Other funds are the Mondriaanstichting, the VSB-fonds and the Prins Bernard Cultuurfonds, each with its own policies, regulations and priorities. In this sector, the economic crisis seems to have its effects already: one fund recently announced that in 2009 it will spend *half* of what it spent in 2008...

Internationally, the EU funds related to 'digital libraries' seem to be the most important possibility by far. It is only a possibility for large projects, and involves a lot of bureaucratic formalities. Until now IISH has not submitted a proposal.

What are advantages and disadvantages of these funding schemes

Obviously, all these funds can generate money to do things we cannot afford otherwise. The programs provide a platform where knowledge is exchanged and can be reused. Solutions are developed for problems that are much too large for individual organizations.

On the other hand, with 'regular money for regular tasks' going down while costs are going up, this increasing dependency on project money has an impact on the organization and its traditional tasks. Basic tasks of cataloguing and preservation have come under a pressure that cannot be relieved by project funding. Basic digitization - as important as it is - is not covered by many funding programs.

The national programs in the Netherlands have mainly been developed by the big National institutions (Royal Library, National Archives etc.). They have 'created' the funds, so who could blame them for spending most of it too? But the policies and priorities developed do not always fit the ideas and needs of other libraries and archives. There are - for instance - important differences between state archives working with government bodies and a 'special' archive working with informal organizations and private persons.

To make things worse, there seems to be a fair amount of monopolism and nationalism in the heritage debate nowadays. If you believe the policy makers and the Big Players, our national heritage is saved when the collections of the big national institutions have been digitized. Collections from 'small', 'private' or 'specialist' institutes can sometimes be of 'additional' value, so it can sometimes be digitized with project funds as long as the results are 'donated' to the national repositories and portals (the long-term consequences as to what these national repositories can do with this material are very unclear). And, even worse, heritage

material is used as a tool to create a 'national identity', or a 'European identity', whatever that may be. Many of our most important archive creators would be horrified by the idea...

But to get back to the advantages: *the funds exist and there is a lot of money involved*. We have no choice but to try to understand how they work, to invest in network building, and to develop proposals that fit our needs in their terms.

What are our main 'problems' and limitations

The first has to do with the size of the institute in relation to the challenges involved. We do not have the money, staff or expertise to solve all problems by ourselves. We have to cooperate with other organizations and to make use of solutions developed elsewhere. Even then, we need to have enough expertise in house to make these choices. And because of the speed of developments, this expertise has to be kept up to date. This requires an investment on our side.

The second has to do with the scale of projects. There is a strong drive to work in projects, with a time frame of one to four years. Much of our work, however, has to be relevant for a longer period of time. The projects and their results need to fit in a bigger strategy, the results have to be valid for a longer period and in a wider context. We need to find a balance between 'projects' and 'programs'. This requires an organizational re-orientation.

Which opportunities do we see within IALHI, which possibilities for cooperation

Information on digitization projects: For the exchange of information on digitization and the handling of born-digital collections, we believe IALHI is not the best platform for most members. Local, regional or national contacts will probably be more effective. On those levels, contacts and cooperation already exist, and national procedures and funding opportunities are better known. That is not to say that the exchange of information between IALHI members would be useless. Especially information about ongoing and future digitization projects would be extremely useful to announce and publish - for instance on labourhistory.net. We have the Labnet and Labour History News services available, and could also publish a permanent list on labourhistory.net.

Storage facilities: Several times, IALHI members have expressed the opinion that the storage of digital files (digitized or digital born) will become a problem. It is expensive, once you start digitizing on a large scale or start to collect digital archives on a large scale. Using external storage facilities, commercial or with local/national colleague institutes, can also be expensive, or come with many conditions and limitations. Perhaps a joint storage facility, developed in IALHI context, could be an alternative.

Labour History Index/Platform: For several years, we have tried to get a "Labour History Index" or "Labour History Platform" project going. A first pilot version was released in 2005. IISH has not been able to expand this index, to rebuild it on a new software platform and to keep it up to date all on its own and at its own costs. The general idea of a Labour History Index or Platform seems to be as valid as ever - and it might just be possible to redefine it in a way that would make a funding application funding possible. IALHI is an interesting multi-national and multi-lingual community, with institutes holding important 'special collections' in a relatively field. Combining these collections' metadata and digitized versions with the 'expert knowledge network' that IALHI is could be an interesting proposal to EU's funders. We will have to rethink the original idea, and to think of it less as an individual website - perhaps it should be a special portal in the Europeana Digital Library (to name just one possibility).

To get a project like this well defined, and to draw up a feasible proposal to the EU funds, we will need to form a project group. This project group should consist of representatives of institutes willing and able to invest time. A concrete goal could be a project proposal to the EU by September 2010.

PAYS-BAS

Politiques de numérisations aux Pays-Bas : le cas de l'Institut international d'histoire sociale d'Amsterdam

Marien van der Heijden

(International Institute of Social History, Amsterdam)

Que numérisons-nous ou qu'allons nous numériser ?

Archives: numérisation en début de projet, dépendante des financements. En 2009 les archives de Dora Russell (20 m) et de Ferdinand Domela Nieuwenhuis (5 m) seront numérisées dans le cadre du « Programme Metamorfoze » (Programme national pour la conservation du patrimoine papier).

Livres, brochures, périodiques: numérisation en début de projet, dépendant des financements. En 2009, l'International Review of Social History sera numérisée (1936- à peu près 25 000p)

Images: photos, affiches, imprimés sont numérisés en haute définition (100, 300DPI) de façon automatisée. Les fichiers originaux sont stockés, des images basse résolution sont générées auxquelles on a accès via les catalogues. Certaines

collections sont numérisées dans le cadre de projets liés à des financements spécifiques (par ex. 5 300 affiches des centres de jeunesse alternatifs des Pays-Bas)

Films, videos, sons: en 2009, on a lancé un projet pilote pour numériser 100 heures de matériel sonore, 50 heures de video et 10 heures de films. Les résultats seront utilisés pour lancer une proposition de projet de numérisation sur l'ensemble de la collection (15 000 heures de matériel sonore, 8 000 heures de video, 2 000 heures de films).

Le "Projet Centrale": en 2009, débutera un projet de numérisation des collections achetées avant 1940 par 'De Centrale', la compagnie d'assurances ouvrière qui finança l'IISG dans ses premières années. Ces collections incluent 180 m d'archives (Bakounine, Marx, Nettlau, Trotski, Lucien Descaves, Louis Michel, Eduard Bernstein, Karl Kautsky, etc.) et d'importantes collections de livres et périodiques. 2009 sera « l'année pilote ». Dans ce projet, des procédures d'ordre technique et organisationnel seront développées pour la numérisation, le stockage et l'accès au matériel numérisé. Cela inclut des processus d'automatisation, des standards de qualité et de contrôle, des identifiants, et des protocoles d'échange de données. Le projet est financé par la compagnie d'assurances SNS Reaal, qui a succédé à 'De Centrale'.

Projets/programmes liés à la numérisation

Réservoir numérique: L'IIHS a développé des équipements maison pour le stockage permanent de grands volumes de fichiers numériques (« born digital » ou numérisés). Ces équipements sont fondés sur les standards OAIS pour les réservoirs numériques (Trust digital repositories). Les fonctions de base de captation et de stockage du 'TDR' sont opérationnelles, les fonctions administratives et d'accès sont en cours de développement.

Accès: La plateforme de recherche (search.iisg.nl) qui offre déjà l'accès au catalogue des livres, brochures, périodiques et collections d'images, aux inventaires d'archives et autres métadonnées sera développé pour offrir un accès aux collections numérisées.

La Plateforme Web 2.0: en 2009-2010, une plateforme sera construite qui offrira de nouvelles formes de présentation et d'interaction avec les collections numérisées. En matière de présentation, il sera possible de visualiser les collections à partir de calendriers et cartes.

En matière d'interaction, le matériel digital sera téléchargé vers des 'content pools' ou des 'sites sociaux' comme Flickr pour les photos et Youtube pour les videos, à partir desquels le public pourra réutiliser et enrichir les divers éléments. La plateforme offrira des fonctionnalités qui permettront de consigner ces informations en une présentation web fabriquée par les utilisateurs, d'ajouter d'autres éléments, textes et autres documents à cette présentation. Ainsi, la

plateforme pourra être utilisée par des 'communautés' d'utilisateurs qui ajouteront ainsi leurs connaissances et - par exemple- enrichiront les métadonnées et ajouteront des informations contextuelles. Le matériel numérisé dans le cadre du 'Projet Centrale' sera utilisé dans des projets pilote de cet ordre.

HITIME : un projet de 4 ans commencera en 2009 pour développer un 'logiciel boîte à outils' pour l'exploitation historique des textes. Le logiciel pourrait reconnaître automatiquement des 'entités' tels que des noms de personnes ou d'organisations, des lieux et des dates, ainsi que les relations entre ces entités, et les présenter en un réseau d'associations, lié à un calendrier. Cela pourrait aider les utilisateurs à trouver des réponses à des questions qui actuellement demandent une trop grande somme d'efforts humains. Les importantes collections de l'IISH en matière de sources primaires, secondaires et de métadonnées sera le 'cas d'école' pour ce projet de recherche, porté principalement par le 'Tilburg Centre for creative computing' de l'Université de Tilburg. Les résultats seront publiés en Open source, et devraient être applicables à des sources multilingues.

Avec qui coopérons-nous (au niveau national/international) ?

Au niveau national, l'IISH est principalement actif dans les échanges d'informations et la participation à des forums conjoints animés par les institutions et organisations concernées, telles que la 'Royal Library', les archives nationales, le consortium hollandais pour la conservation numérique (Netherlands Coalition for Digital Preservation), etc. Une forme de coopération très concrète est développée avec l'IIAV (Centre d'information et d'archives sur les mouvements de femmes), qui s'intégrera à l'IISH en 2011. Les infrastructures numériques existantes devront en grande partie être fusionnées.

Au niveau international, l'IISG est en contact avec les programmes d'OCLC/RLG et avec de nombreuses institutions européennes dans le domaine de la numérisation et du « patrimoine numérique ». Avec les communautés européennes, une participation pilote à Europeana est prévue pour fin 2009.

Quelles possibilités nationales ou internationales de financements avons-nous ?

L'IISH réussit de façon globalement appréciable à mener régulièrement des projets via différents programmes nationaux, tels que « Metamorfoze » (Numérisation du patrimoine imprimé) et « Geheugen van Nederland » (images). Nous allons tenter également d'accéder au programme national sur la numérisation du matériel audiovisuel. Le projet 'plateforme Web 2.0' est financé par le programme national pour 'une numérisation coordonnée' (Concerted digitization). Le projet HITIME est financé par le programme 'CATCH' de l'organisation néerlandaise pour la recherche scientifique.

Dans le secteur privé, notre principal financeur est SNS Reaal. Nous sommes également financés par le « Mondriaanstichting » le « VSB-Fonds » et le « Prins

Bernard Cultuurfonds », chacun ayant sa propre politique, ses protocoles et ses priorités. Dans ce secteur, la crise économique semble avoir déjà des effets : un des fonds a d'ores et déjà annoncé qu'il ne financerait en 2009 que la moitié de ce qu'il avait financé en 2008...

Au niveau international, les financements européens relatifs aux bibliothèques numériques semblent être de loin les plus importants. Cela n'offre de possibilités que pour les très grands projets, et cela implique de nombreuses formalités bureaucratiques. Jusqu'à présent, l'IISH n'a pas soumis de projet.

Quels sont les avantages et désavantages de ces modes de financements ?

De façon générale, tous ces fonds permettent d'avoir de l'argent pour faire des choses que nous ne pourrions nous permettre autrement. Les programmes offrent des plateformes où les connaissances peuvent être échangées et réutilisées. Des solutions peuvent être développées pour des problèmes beaucoup trop importants pour chaque organisation.

D'un autre côté, avec la baisse des financements réguliers pour des tâches régulières, alors que les coûts augmentent, cette dépendance croissante vis-à-vis des financements sur projets a un impact sur l'organisation et ses fonctions traditionnelles. Les fonctions de base de catalogage et de conservation obéissent à des contraintes qui ne peuvent être prises en charge par des financements sur projets. Les numérisations de base - aussi importantes soient-elles - ne sont pas couvertes par les financements sur projets.

Les programmes nationaux aux Pays-Bas ont été principalement développés par les grandes institutions nationales (Bibliothèque royale, Archives nationales etc.). Elles ont 'créé' les fonds et qui pourrait les blâmer d'en dépenser la plus grande partie ? Mais les politiques et les priorités développées ne correspondent pas toujours aux idées et aux besoins des autres bibliothèques et centres d'archives. Il y a, par exemple, de grandes différences entre les archives d'état travaillant avec des organisations gouvernementales et un centre d'archives spécialisé travaillant avec des organisations et des personnes privées.

Pour rendre les choses encore pire, il semble y avoir une croissance aussi bien de l'esprit de monopole que de celui du nationalisme dans le débat sur l'héritage aujourd'hui. Si vous en croyez les politiciens et les maîtres du jeu, notre héritage national sera sauf quand les collections des grandes institutions nationales auront été numérisées.

Les collections des petites institutions, privées ou spécialisées, peuvent parfois représenter un plus, et ainsi peuvent parfois être numérisées à l'aide de financements sur projets tant que les résultats sont 'donnés' aux réservoirs numériques nationaux et aux portails (les conséquences à long terme et ce que ces réservoirs peuvent faire de ce matériel sont peu claires). Et, encore pire, cet héritage matériel est utilisé pour créer une 'identité nationale', ou une 'identité

européenne', quoique ce puisse être. Beaucoup de nos plus importants producteurs d'archives auraient été horrifiés par cette idée.

Mais pour en revenir aux avantages : les sources de financement existent et il y a énormément d'argent engagé. Nous n'avons pas d'autre choix que d'essayer de comprendre comment ils fonctionnent, pour nous investir dans la construction des réseaux, et développer des propositions qui incluent nos besoins dans leurs logiques.

Quels sont nos problèmes principaux et quelles sont nos limites ?

Le premier de ces problèmes a à voir avec la taille de l'institut par rapport aux enjeux. Nous n'avons sûrement pas l'argent, le personnel ou les expertises pour résoudre tous les problèmes tout seul. Nous devons coopérer avec d'autres organisations et utiliser des solutions développées ailleurs. Et même alors, nous devons avoir suffisamment d'expertise à demeure pour faire ces choix. A cause de la rapidité des évolutions, cette expertise doit être toujours mise à jour. Cela nécessite un investissement de notre part.

Le second problème a à voir avec l'échelle de ces projets. Il y a un investissement fort à travailler dans ces projets, pour un temps qui peut aller de un à quatre ans. La plus grande partie de notre travail, cependant, porte sur un plus long terme. Les projets et leurs résultats doivent être intégrés dans une stratégie plus large, les résultats doivent être valides pour une plus longue période et dans un contexte plus large. Nous avons besoin de trouver un équilibre entre projets et programmes. Cela nécessite une réorientation organisationnelle.

Quelles opportunités nous offre l'IALHI, et quelles possibilités de coopération ?

Information sur les projets de numérisation : En matière d'échange d'information sur les numérisations et de collections « born digital », nous pensons que l'IALHI n'est pas la meilleure plateforme pour la plupart d'entre nous. Des contacts locaux, régionaux ou nationaux seront probablement plus efficaces. A ces niveaux, contacts et coopération existent déjà et les procédures nationales ainsi que les possibilités financières sont mieux connues. Cela ne veut pas dire que les échanges d'informations entre les membres de l'IALHI soient inutiles. En particulier, il pourrait être extrêmement utile de diffuser l'information sur les projets de numérisation en cours et à venir - par exemple sur labourhistory.net. Nous pouvons utiliser Labnet et Labour History News, et nous pourrions également publier une liste permanente sur labourhistory.net.

Facilités de stockage : A plusieurs reprises, les membres de l'IALHI ont exprimé l'opinion que le stockage de fichiers numériques (numérisés ou « born digital ») allait devenir un réel problème. C'est cher, une fois que l'on commence à numériser à grande échelle ou que l'on commence à collecter des archives numériques à grande échelle. Externaliser le stockage auprès d'institutions commerciales ou de collègues

nationaux ou locaux, peut aussi se révéler cher, ou assorti de nombreuses conditions et limites. Peut-être qu'un entrepôt collectif, développé dans le cadre de l'IALHI, pourrait être une alternative.

Labour History Index/Plateforme : durant plusieurs années, nous avons essayé de développer un projet de « Labour history Index » ou « Labour History Plateforme ». Il en a résulté une version pilote en 2005. L'IISH n'a pas été en mesure de développer cet Index, de le reconstruire sur une nouvelle plateforme logicielle, de le maintenir à jour à lui seul et avec ses moyens. L'idée générale du Labour History Index ou Plateforme semble plus que jamais valable et il pourrait juste être possible de le redéfinir de manière à obtenir des financements. L'IALHI est une intéressante communauté multilingue et multinationale, avec des institutions détenant d'importantes collections spécialisées dans le même champ. Associer les index de ces collections et leurs versions numérisées en même temps que le réseau de connaissance et d'expertise qu'est l'IALHI pourrait être une proposition intéressante à faire pour obtenir des financements auprès des communautés européennes. Nous devons repenser l'idée initiale et moins y réfléchir comme à un site web isolé. Cela pourrait peut-être être un portail spécifique d'Europeana (pour ne parler que d'une des possibilités).

Pour bien définir un tel projet, et pour en faire une proposition recevable pour les financements européens, nous devons former un groupe du projet. Ce groupe pourrait être constitué des représentants des institutions qui le désireraient et seraient susceptibles d'y investir du temps. Un but concret pourrait être de proposer le projet en septembre 2010 aux Communautés européennes.

Networks and digitisation policy in Switzerland

Urs Kälin (29-01-2009)

(Schweizerisches Sozialarchiv ·
Stadelhoferstrasse 12, CH-8001 Zürich)

Digitization of cultural goods such as books and archival documents in Switzerland is still in its infancy, comprehensive strategies or appropriate coordinating bodies are missing. An important milestone in the debate was the report "*Politique de la mémoire nationale*", published in order of the Federal Office of Culture in 2005. In a position paper, Swiss authorities summed up the state of affairs as follows:

- a) Archives and libraries are planning and executing digitization projects in an uncoordinated manner.
- b) Standards of digitization, description, safety of links and networking are not guaranteed.
- c) Problems of access and long-term archiving are not processed.
- d) The requirements of future research in the digitization, accessibility, production of links and networking are not taken into account.
- e) The major problems in the documentation of contemporary events are not processed.

The reasons for this very unsatisfactory situation are manifold:

- There is no comprehensive memo policy at the national level.
- There is no national digitization program.
- Contrary to microfilming, which is promoted in the perspective of cultural heritage protection, there are no public funds for digitization, especially no funding from the National Swiss Science Foundation.
- Priority has been attributed to the archiving of born digital documents and images. In this area a lot of money has been invested and there are several concrete projects ongoing.

Generally, it can be stated that libraries are much more advanced in digitization projects than archives.

CURRENT LIBRARY PROJECTS

- *SEALS* [<http://retro.seals.ch/digbib/en/home>]

Swiss Electronic Academic Library Service, part of the project E-Archiving of the Consortium of Swiss Academic Libraries; retro digitization of printed scientific journals.

Journals: 40, Volumes: 952, Articles: 56'729, Pages: 642'563

- *DigiBern* [<http://www.digibern.ch/en/angebot.html>]

DigiBern is a web offer by the University Library of Berne and includes digitized texts and maps pertaining to the history and culture of Berne city and the canton of Berne. DigiBern has been online since 2002. So far, the most widely spread and used printed texts have been digitized. The documents are fully searchable for keywords. They can be traced via the index on the DigiBern web site. Moreover, they are indexed and web linked in the online library catalog IDS Basel/Bern.

- *e-Helvetica*

[http://www.bak.admin.ch/slb/slb_professionnel/01693/index.html?lang=en]

The objective of e-Helvetica is to collect, list and conserve documents which arrive at the Swiss National Library in digital format, either online Helvetica or data on mediums such as CDs, DVDs or CD-ROMs etc. With the increasing importance of electronic publications in our society, the need to develop conservatory solutions in this area has become more pressing. As this area is rather new, e-Helvetica has to cooperate with other Swiss and foreign institutions to develop and implement conservation measures for digital information.

- *E-Collection ETH* [<http://e-collection.ethbib.ethz.ch/>]

The ETH E-Collection is the Institutional Repository of the ETH Zurich. This service allows members of the ETH community to publish relevant scientific information with open access.

- *Digital library RERO DOC* [<http://doc.rero.ch/?ln=en>]

RERO DOC is the digital library of RERO, the Library Network of Western Switzerland. It has been created to promote, preserve and provide public access to the scholarly literature produced by the institutions participating to the project (electronic theses, dissertations and e-prints) as well as to the digitalized collections of RERO libraries. In the framework of the Open Access initiative it provides a platform for researchers in member universities to self-archive their publications.

CURRENT ARCHIVAL PROJECTS

- *DoDiS* [<http://www.dodis.ch/e/datenbank.asp>]

DoDiS is the Internet database of Swiss Diplomatic Documents (DDS). It contains comprehensive information on documents, personalities, organizations, geographical terms and bibliographical references concerning the foreign relations of Switzerland. The DoDiS Internet database was conceived as a way of publishing the documents for the volumes of the second period (from vol. 16, 1945 and following years) in digital form. Unlike the printed publication, the database provides no scientific annotations and annexes. Parallel use of the printed publication therefore offers considerable benefits. DoDiS displays all documents of the second series as image (by PDF format). In addition to this first group,

DoDiS includes more documents which haven't been published in the printed edition. A third group in DoDiS are those documents without image (including literature). For more detailed explanations see Contents of the database. The Thematic list supports your search of subjects or countries.

- *University of Zurich: Matrikeedition* [<http://www.dodis.ch/e/datenbank.asp>]

On the preparations for the 150th anniversary of the University of Zurich a edition project emerged: The digital Matrikeedition has all students enrolled since 1833, and provides biographical information on these people. There are currently around 27,000 entries.

- *Memoriav: Memobase* [<http://en.memoriav.ch/memobase/default.aspx>]

MEMOBASE is a core product of Memoriav. The database's objective is to gather together newly indexed, restored or digitised collections from Memoriav projects. This information on Switzerland's cultural heritage is then available to researchers, educators and any other interested parties - free of charge. Currently, MEMOBASE offers access only to metadata in the areas of sound/radio, film and video/TV in general. One of Memoriav's top priorities is to provide step-by-step access to all content. Due to copyright reasons, however, this will only be possible in streaming mode for specific locations and workstations.

OTHER CURRENT PROJECTS

- *Poster collection*

[*Swiss National Library: <http://ccsa.admin.ch/cgi-bin/gw/chameleon>*]

The poster collection encompasses a wide variety of posters from the period at the end of the 19th century to today. 35'000 posters from various stocks: Swiss National Library, Bibliothèque publique et universitaire de Genève, the Bibliothèque publique et universitaire de Neuchâtel, Médiathèque Valais and Swiss Transport Museum.

- *e-codices Virtual Manuscript Library of Switzerland*

[<http://www.e-codices.unifr.ch/en>]

The purpose of "e-codices" is to provide access to the medieval codices in Switzerland by creating a virtual library. At the moment, the virtual library contains 363 manuscripts from 16 different libraries. The virtual library will be continuously updated and extended.

Mise en réseau et politiques de numérisation en Suisse

Urs Kälin (29-01-2009)

(Schweizerisches Sozialarchiv ·
Stadelhoferstrasse 12, CH-8001 Zürich)

La numérisation de biens culturels tels que livres et documents d'archives en Suisse est encore en enfance, les stratégies d'ensemble et les organismes de coordination appropriés manquent. Une pierre importante dans le débat a été le rapport « Politique de la mémoire nationale » publié sous l'égide de l'Office fédéral de la culture en 2005. Dans un texte programmatique, les autorités suisses ont résumé ainsi la situation :

- Centres d'archives et bibliothèques programment et mènent des projets de numérisation sans coordination.
- Les standards de numérisation, description, sécurité des liens et des réseaux ne sont pas garantis.
- Les problèmes d'accès aux données et d'archivage à long terme ne sont pas traités.
- Les exigences d'une recherche future en matière de numérisation, accessibilité, production de liens et mise en réseau ne sont pas prises en compte.
- Les problèmes majeurs en matière de documentation des événements contemporains ne sont pas pris en compte.

Les raisons de cette situation très insatisfaisante sont multiples :

- Il n'y a pas de politique générale de mémoire au niveau national.
- Il n'y a pas de programme national de numérisation
- A l'inverse du microfilmage, qui est promu dans une perspective de protection de l'héritage culturel, il n'y a pas de fonds publics pour la numérisation, et en particulier il n'y a pas de financement émanant de la Fondation nationale scientifique suisse.
- La priorité a été accordée à l'archivage des « Born digital archives » et des images. Dans ce domaine, beaucoup d'argent a été investi et il y a plusieurs projets en cours.

De manière générale, on peut dire que les bibliothèques sont plus avancées en matière de numérisation que les archives.

PROJETS DES BIBLIOTHEQUES EN COURS

- **SEALS** [<http://retro.seals.ch/digbib/en/home>]

Swiss Electronic Academic Library Service, partie du projet E-Archiving du Consortium des bibliothèques académiques suisses; rétro-numérisation des périodiques scientifiques imprimés suisses

Titres: 40, Volumes: 952, Articles: 56'729, Pages: 642'563

- **DigiBern** [<http://www.digibern.ch/en/angebot.html>]

Digibern est une offre web de la bibliothèque universitaire de Berne et inclut des textes numérisés et des cartes se rattachant à l'histoire et à la culture de la cité et du canton de Berne. Digibern est en ligne depuis 2002. Ainsi, les textes imprimés les plus répandus et utilisés ont été numérisés. Les documents sont accessibles par mots clefs. Ils peuvent être retrouvés via l'index du site web DigiBern. De plus ils sont indexés et liés au catalogue en ligne de IDS Basel/Bern.

- **e-Helvetica**

[http://www.bak.admin.ch/slb/slb_professionnel/01693/index.html?lang=en]

L'objectif de e-Helvetica est de collecter, lister et conserver les documents qui arrivent à la Bibliothèque nationale Suisse sous format numérique, soit le web helvétique soit les données de CD, DVD ou CDROM. Avec l'importance croissante des publications électroniques dans notre société, la nécessité de développer des solutions de conservation dans ce domaine s'est fait plus précis. Comme ce domaine est encore neuf, e-Helvetica doit coopérer avec d'autres institutions suisses et étrangères pour développer et mettre en œuvre des mesures de conservation pour l'information numérique.

- **E-Collection ETH** [<http://e-collection.ethbib.ethz.ch/>]

The ETH E-Collection est le réservoir institutionnel de ETH Zurich. Ce service permet aux membres de la communauté ETH de publier de l'information scientifique pertinente en open source.

- **Bibliothèque numérique RERO DOC** [<http://doc.rero.ch/?ln=en>]

RERO DOC est la bibliothèque numérique de RERO, le réseau des bibliothèques de Suisse de l'Ouest. Il a été créé pour promouvoir, préserver et donner accès à la littérature scolaire produite par des institutions participant au projet (thèses électroniques, dissertations et publications électroniques) ainsi qu'aux collections numériques des bibliothèques RERO. Dans le cadre des initiatives en Open access, cela offre une plateforme aux chercheurs des universités membres pour auto-archiver leurs publications.

PROJETS ARCHIVISTIQUES EN COURS

- *DoDiS* [<http://www.dodis.ch/e/datenbank.asp>]

DoDis est la base de données de Swiss Diplomatic Documents (DDS). Elle contient une information détaillée sur les documents, les personnes, les organisations, les termes géographiques et les références bibliographiques concernant les relations internationales de la Suisse. La base de données en ligne DoDis a été conçue comme une façon de publier les documents pour les volumes de la seconde période (du vol. 16, 1945 et années suivantes) au format numérique. Comme la publication imprimée, la base de données n'offre pas d'annotations scientifiques ni d'annexes. L'usage parallèle de la publication imprimée offre néanmoins de considérables bénéfices. DoDis permet de visualiser tous les documents de la seconde série en mode image (Format PDF). En plus de ce premier groupe, DoDis inclut plus de documents qu'il n'en avait été publié dans l'édition imprimée. Un troisième groupe dans DoDis sont des documents sans image (incluant la littérature). Pour plus d'explications, voir Contents of the database. Une liste thématique aide à la recherche par sujets ou pays

- *Université de Zurich : Matrikeedition* [<http://www.dodis.ch/e/datenbank.asp>]

Au moment des préparatifs du 150^{ème} anniversaire de l'Université de Zurich, un projet d'édition émergea : le Matrikeedition numérique contient tous les étudiants enrôlés depuis 1833, et fournit des informations biographiques sur ces personnes. Il y a actuellement 27 000 entrées.

- *Memoriav: Memobase* [<http://en.memoriav.ch/memobase/default.aspx>]

MEMOBASE est un produit central de memoriav. L'objectif de la base de données est de rassembler les collections nouvellement indexées, restaurées ou numérisées des projets Memoriav. Cette information sur l'héritage culturel suisse est alors disponible pour les chercheurs, enseignants ou autres intéressés, gratuitement. Actuellement, MEMOBASE offre accès aux métadonnées dans les domaines du son/radio, du film et de la vidéo/TV en général. Une des priorités de Memoriav est d'offrir petit à petit accès à l'ensemble. Pour des raisons de copyright, cependant, cela sera seulement possible en streaming pour des emplacements spécifiques et des lieux de travail.

AUTRES PROJETS ACTUELS

- *Collection d'affiches* [Swiss National Library: <http://ccsa.admin.ch/cgi-bin/gw/chameleon>]

La collection d'affiches englobe une grande variété d'affiches de la fin XIX^e à nos jours. 35 000 affiches de provenance diverse : bibliothèque nationale suisse, bibliothèque publique et universitaire de Genève, bibliothèque publique et universitaire de Neuchâtel, Médiathèque Valais et Musée suisse du transport.

• *e-codices Bibliothèque virtuelle Suisse de manuscrits*

Le propos de e-codices est de donner accès aux Codex médiévaux suisses en créant une bibliothèque virtuelle. En ce moment, la bibliothèque virtuelle comprend 363 manuscrits de 16 bibliothèques différentes. La bibliothèque virtuelle sera mise à jour et étendue.

The indexing of digital documents in Germany

Rüdiger Zimmermann
(Friedrich Ebert Stiftung)

I would like to speak about the national strategies in Germany to digitize its cultural heritage and the ideas of archiving digital resources on the national level. And I would also like to outline the strategies of the library and the archive of the Friedrich Ebert Foundation in this context.

Germany is a federalist state. Thus, the responsibility for many aspects of cultural life lies with the German federal states. On the national level the Deutsche Forschungsgemeinschaft DFG (German Research Foundation) works as a steering committee. For more than ten years the DFG pursues an ambitious strategy in the context of digitizing the cultural heritage. Funds of considerable size are allocated for this program, though the initiative for specific projects must be taken by single libraries and archives.

By now the DFG has compiled a kind of "bible" for digitizing activities (code of practice¹). This paper clearly defines the standards that need to be met to successfully apply for the funds. Nevertheless, to date only 1% of the printed national heritage is available on the Internet.

Of particular importance is the public private partnership of the Bayerische Staatsbibliothek (Bavarian State Library) with Google. This cooperation aims at digitizing all documents in Germanys largest historical library that are not copyrightable any more. This will comprise the heritage of the German labour movement as well, provided that such documents made it into this "feudal" library. The project is discussed quite controversial in the library community, but the benefits for research become more and more obvious.

The Bundesarchiv (Federal Archives) now also cooperates with Google with the intention to provide images free from copyright on the Internet.

The Internet and its archiving

In 2006, the Bundestag (German Parliament) passed the bill on the Deutsche Nationalbibliothek DNB (German National Library)². For almost three years now the DNB in Frankfurt has been responsible for the archiving of German Internet pages. Everyone who hosts Internet pages in Germany is legally obligated to

¹http://www.dfg.de/forschungsfoerderung/wissenschaftliche_infrastruktur/lis/aktuelles/download/praxisregeln_kulturelle_ueberlieferung_0208.pdf

² <http://bundesrecht.juris.de/dnbg/BJNR133800006.html>

cooperate and deliver these pages to the DNB. The DNB in Frankfurt received extensive financial support to increase its members of staff. Meanwhile the bill has been supplemented by several regulatory statutes. The subject of archiving the Internet is rather complicated, and only the future will show how it works in practice.

So far it seems rather unlikely to reach a complete web archive. The National Library aims to find partners for this complex challenge and already cooperates with IBM in terms of long-term preservation projects.

The strategies of the Friedrich Ebert Foundation

The FES Library digitizes sources on the history of the German labour movement but is still far from having digitized only 1% of all its library holdings. We started with the most important journals of the German labour movement. At several IALHI conferences we already presented corresponding projects³. The library managed to raise funds from the German Research Foundation as well as from other sponsors.

We digitized printed sources like the "Sozialistische Monatshefte", the "Neue Zeit", "Gewerkschaftliche Monatshefte" or social democratic press releases. The Historische Forschungszentrum (Historical Research Centre) of the FES supported us as well. Currently we work on an application concerning the digitization of the printed heritage of the German trade unions before 1933, i.e. we would like to digitally republish trade union newspapers.

The library collects and indexes current digital publications by political parties and trade unions from Germany, Western Europe and North America by the same library standards that also apply for printed publications. In this context a main focus of the library's work lies in the collection of digital press releases from political parties and trade unions that are of particular interest for our special subject collection. We also presented this project on the occasion of IALHI conferences⁴.

At present the archive digitizes small holdings of the early German labour movement without external support. The outstanding project among the diverse digitizing activities is the archiving of the Internet presence of the German Social Democratic Party. This comprises all political levels of the party. Likewise, this pilot scheme received the support of the German Research Foundation. Rudolf Schmitz as responsible person presented this trend-setting project at many national and international conferences. He gained considerable recognition for his

³ <http://library.fes.de/library/fr-digbib.html>

⁴ <http://library.fes.de/fulltext/bibliothek/01103.htm>

solutions. Amongst others he gave presentations at IALHI's Milan and Zurich conferences⁵.

In the past the archive collected mass files of German trade unions and the German social democracy. Today many documents are only available as digital versions. In close cooperation with several German trade unions the archive experimentally took over digital documents and meta data from document management systems. All considerations and concepts for this document transfer and for the inevitable editing (file conversion, editing the meta data structure) and further use of this digital transmissions base upon the well-known OAIS reference model (ISO 14721:2003).

To guarantee integrity and authenticity of digital archival holdings for both the depositor and the subsequent user appropriate measures of protection have been developed and tested successfully. Besides this, the archive contacted several trade unions (ver.di, IG Metall) to give advice in terms of archival requirements for document management systems to enable a subsequent transfer and utilisation of these digital documents in the Archives of Social Democracy at all.

At present the archive prepares the take over of digital records from the former trade union Öffentliche Dienste, Transport und Verkehr ÖTV (Public Services and Transport Workers' Union).

In terms of long-term preservation for digital journals the library tries to find sustainable and yet affordable solutions in cooperation with the "Kooperativer Bibliotheksverbund Berlin-Brandenburg", KOBV (cooperative library network Berlin-Brandenburg). A major German data processing centre is affiliated to the KOBV. The German Research Foundation only awards funds to libraries if projects involve a sustainable concept in cooperation with a major data processing centre.

The Library of the Friedrich Ebert Foundation does not own the exclusive right for the digitization of printed labour movement sources. Nevertheless, we have to visualize these sources. At the same time we are obliged to find solutions for the integrative presentation of digitized documents and archived Internet sources within the FES. For this purpose we developed the "Portal deutscher Arbeiterbewegung" (portal of German labour movement) which enables us to present the particular sources that are visible world-wide under one surface. Jacques Paparo reported on this during the IALHI conference in Rome⁶. This portal follows the ideas of the web2.0. It furthermore confirms our hegemony within our main subject collection.

In my opinion a portal solution would be an adequate project for international cooperation. For the future of the IALHI it will be essential to develop such a

⁵ <http://library.fes.de/pdf-files/bibliothek/02625.pdf>

⁶ <http://library.fes.de/pdf-files/bibliothek/04805.pdf>

cooperative project. This scheme might be organised in a rather simple way - like that of our Swiss friends. But without additional personal resources such a portal cannot be realised either. In my opinion this will become the biggest stumbling block.

ALLEMAGNE

L'indexation des documents électroniques en Allemagne

Rüdiger Zimmermann
(Friedrich Ebert Stiftung)

Je voudrais vous parler des programmes de numérisation du patrimoine culturel mis en place en Allemagne, et des réflexions en cours, toujours à l'échelle nationale, sur l'archivage des documents numériques. Je présenterai aussi les projets de la bibliothèque et des archives de la Fondation Friedrich Ebert qui s'inscrivent dans ce contexte.

L'Allemagne est un Etat fédéral. De nombreux aspects de la politique culturelle sont donc du ressort des Etats fédérés, des Landers. A l'échelle nationale, la Fondation allemande pour la recherche (DFG, Deutsche Forschungsgemeinschaft) fait office d'organisme central pour la promotion de la recherche scientifique. Elle mène depuis plus de 10 ans un ambitieux projet de numérisation. Des ressources financières importantes ont été allouées à ce programme, mais il est du ressort des bibliothèques ou des centres d'archives de prendre eux-mêmes l'initiative de projets concrets.

La DFG a désormais établi une sorte de "bible" de recommandation des bonnes pratiques pour les travaux de numérisation. Celle-ci est disponible en ligne, et définit précisément les normes requises pour espérer obtenir une aide financière de la DFG. A ce jour, seul 1 % du patrimoine imprimé allemand est numérisé et accessible sur Internet.

Il faut souligner, car il est important, le partenariat établi par la Bibliothèque d'Etat de Bavière (Bayerische Staatsbibliothek) avec Google, dont l'objet est de numériser tous les documents de la plus grande bibliothèque d'histoire en Allemagne non encore soumis au droit d'auteur. Un tel projet inclura le patrimoine imprimé du mouvement ouvrier allemand, à condition toutefois que ce type de documents ait fait leur chemin jusqu'à cette bibliothèque « féodale ». Ce partenariat n'est pas sans entraîner quelques polémiques dans le monde des bibliothèques, mais ses bénéfices pour la recherche historique apparaissent de plus en plus évidents.

Les archives fédérales (Bundesarchiv) travaillent désormais elles aussi avec Google, et envisagent de rendre accessibles en ligne des images libres de droit.

L'archivage de l'internet

En 2006, le Parlement (Bundestag) a adopté une loi fixant les missions de la bibliothèque nationale (DNB, Deutsche Nationalbibliothek). Depuis trois ans, la DNB, installée à Francfort, a ainsi la responsabilité de l'archivage des pages Internet allemandes. Quiconque en Allemagne héberge des pages Internet est désormais obligé de coopérer avec la bibliothèque nationale, et de lui donner des copies de ces pages. La DNB a bénéficié de crédits importants pour pouvoir embaucher du personnel. Depuis trois ans, la loi initiale a été complétée par plusieurs décrets d'application. La question de l'archivage de l'Internet est assez complexe, et seul le futur nous dira comment le mettre vraiment en pratique.

Pour l'instant, il semble assez improbable de parvenir à un archivage complet du Web. La bibliothèque nationale est à la recherche de partenaires pour atteindre cet objectif, et travaille par exemple avec la société IBM pour la mise en place de stratégies de conservation des données à long-terme.

La politique et les projets de la Fondation Friedrich Ebert

La bibliothèque de la Fondation numérise des documents concernant l'histoire du mouvement ouvrier allemand, mais nous sommes encore loin de l'avoir fait pour ne serait-ce qu'1 % de l'ensemble de nos fonds. Nous avons commencé ce travail avec nos collections de journaux les plus importants. Ces initiatives ont déjà été présentées à l'occasion de conférences de l'IALHI, et une bibliothèque numérique regroupant ces collections numérisées est consultable en ligne⁷. Pour leur financement, la bibliothèque a pu obtenir des crédits de la DFG, ainsi que d'autres sponsors.

Nous avons numérisé des sources imprimées comme par exemple le *Sozialistische Monatshefte*, le *Neue Zeit*, le *Gewerkschaftliche Monatshefte*, ainsi que d'autres publications social-démocrates. Le Centre de recherche historique (Historische Forschungszentrum) de la Fondation nous a aussi aidé. Nous travaillons maintenant au développement d'une application liée à la numérisation du patrimoine imprimé des syndicats allemands antérieur à 1933 : nous aimerions pouvoir publier de nouveau, sous une forme électronique, les journaux syndicaux de cette époque.

Notre bibliothèque collectionne, catalogue et indexe des publications électroniques actuelles des partis politiques et des syndicats allemands, européens (Europe de l'Ouest) et nord-américains, en respectant les mêmes normes bibliothéconomiques que pour les publications imprimées. Nous travaillons ainsi beaucoup au développement de notre archive des communiqués de presse diffusés électroniquement par des partis politiques ou des organisations syndicales,

⁷ <http://library.fes.de/library/fr-digbib.html>

documents particulièrement intéressants pour nos collections spécialisées. Nous avons déjà eu l'occasion de présenter cette initiative lors de conférences de l'IALHI (des détails sur les objectifs de ce travail, ainsi que des précisions techniques sur les conditions de sa mise en pratique, sont accessibles en ligne sur le site de notre bibliothèque⁸).

Le département Archives de notre Fondation numérise aussi désormais, sans aide extérieure, quelques petits fonds concernant les débuts du mouvement ouvrier allemand. Mais le projet le plus important est consacré à l'archivage de l'Internet du parti social-démocrate, à tous les niveaux de l'organisation. Ce projet pilote a lui aussi reçu le soutien de la DFG. Rudolf Schmitz, qui en a la charge, a déjà présenté cette initiative, susceptible de servir de modèle à d'autres projets du même genre, lors de nombreuses conférences allemandes ou internationales (lors des conférences IALHI à Milan et à Zurich, par exemple⁹), et les solutions qu'ils proposent ont été largement saluées.

Dans le passé, nos Archives ont collecté de très nombreux documents concernant les syndicats allemands et la social-démocratie. Aujourd'hui, beaucoup d'entre eux n'existent plus que sous une forme numérique. En partenariat avec différents syndicats, le département Archives a commencé la récupération expérimentale de ces documents électroniques, et de leurs méta-données, collectées directement à partir des systèmes de gestion de contenu¹⁰ utilisés par leurs producteurs. Les aspects pratiques de ces transferts de documents, ainsi que leur mise à disposition finale en tant qu'archive (conversion des fichiers, "publication" des structures des données), sont fondés sur le désormais bien connu modèle OAIS, Open Archival Information System (norme ISO 14721:2003)¹¹.

Afin de garantir l'intégrité et l'authenticité de ces nouvelles collections d'archives électroniques — nécessaires tant pour les donateurs de documents que pour leurs futurs utilisateurs —, des mesures de protection ont été définies et mises en place avec succès. Des contacts ont aussi été pris avec plusieurs syndicats (dont ver.di¹² et IG Metall), afin de les conseiller sur les exigences de l'archivage, de manière à ce que leurs systèmes de gestion de contenu permettent ensuite le transfert et l'utilisation de leurs productions électroniques en tant qu'archives de la social-démocratie.

⁸ <http://library.fes.de/fulltext/bibliothek/01103.htm>

⁹ Le texte de sa communication à Milan (en anglais) est disponible à l'adresse : <http://library.fes.de/pdf-files/bibliothek/02625.pdf>

¹⁰ Un outil de CMS (Content Management System, en français Gestion de contenu) est un site web disposant de fonctionnalités de publication et offrant en particulier une interface d'administration permettant à un administrateur de site de créer ou organiser les différentes rubriques.

¹¹ L'OAIS est un modèle conceptuel destiné à la gestion, à l'archivage et à la préservation à long terme de documents numériques.

¹² Acronyme de Vereinte Dienstleistungsgewerkschaft (Fédération unifiée des services).

Nos Archives préparent actuellement la récupération d'archives électroniques de l'ancien syndicat des salariés des services publics et des transports ÖTV (Öffentliche Dienste, Transport und Verkehr)¹³.

Pour ce qui est de la conservation à long-terme des journaux électroniques, notre bibliothèque s'efforce aujourd'hui de trouver des solutions pérennes mais financièrement abordables, en partenariat avec l'Association coopérative des bibliothèques Berlin-Brandenburg (Kooperativer Bibliotheksverbund Berlin-Brandenburg, KOBV). L'un des principaux centres allemands de traitement des données est affilié à cette association. La Fondation allemande pour la recherche n'accorde des crédits spécifiques aux bibliothèques que si les projets qu'elles présentent prennent en compte les questions de conservation durable, en collaboration avec un centre de traitement des données.

La bibliothèque de la fondation Friedrich Ebert ne dispose pas d'un droit exclusif pour la numérisation des archives imprimées du mouvement ouvrier. C'est pourtant notre rôle de rendre aussi visibles les documents numérisés ailleurs que chez nous. Dans le même temps, nous sommes obligés de trouver des solutions pour une présentation intégrée des documents numérisés ou des archives électroniques dont nous disposons. Ainsi avons-nous développé un « portail de l'histoire mouvement ouvrier allemand » (Portal deutscher Arbeiterbewegung¹⁴) qui nous permet de présenter à partir d'un seul site les différentes sources sur ce sujet disponibles en ligne partout dans le monde. Jacques Paparo a présenté ce portail lors de la Conférence IALHI tenue à Rome¹⁵. Il fonctionne selon les préceptes du web 2.0, et nous permet de confirmer notre hégémonie dans le domaine qui constitue le cœur de nos collections.

Selon moi, une solution de type « portail » serait la mieux à même de permettre une coopération internationale. Il sera essentiel, pour l'avenir de l'IALHI, de développer ce genre de projet coopératif. Un tel travail pourrait être organisé de manière assez simple — sur le modèle de l'initiative de nos collègues suisses. Mais sans ressources supplémentaires en personnel, un tel portail ne peut pas être mis en place. Voilà à mon avis quel sera le principal obstacle sur notre chemin.

¹³ Fondé en 1949, ce syndicat a été en 2001 l'un des syndicats fondateurs du ver.di.

¹⁴ <http://library.fes.de/portalarb>

¹⁵ Le texte de sa communication est disponible en ligne à l'adresse <http://library.fes.de/pdf-files/bibliothek/04805.pdf>

Digital experience in Italian libraries

David Bidussa

(david.bidussa@fondazionefeltrinelli.it)

General situation and general criteria

Generally digitisation in Italy increased in the last ten years (The first national and official document about digitization was approved in February 2001 at the end of the 3rd library national conference (Padua, 14-16 February 2001), and the main aim was preservation.

Digitisation in Italy is, first of all, preservation. Second preservation means digitization of single texts or documents. In digitisation program and projects don't mean digital library, but digitisation of single texts.

It's enough to analyse the digitisation program of the BDI (*i.e.* : BIBLIOTECA DIGITALE ITALIANA, Italian Digital Library) the digitisation dept. under supervision of the Ministero dei beni culturali e ambientali (all the data are legible at the following webpage :

« <http://www.librari.beniculturali.it/genera.jsp?s=31&lingua=it> » see document enclosed)

It's possible to distinguish three different plans :

- Reproduction of collection. Digitisation is as the microform 30 or 40 years ago. (Most of the projects are the reproduction of the paper collection) ;
- Organization and creation in an unified collection of homogenous documentation (*i.e.* the digitisation of cartographic collection at the National Library in Venice, or digitisation of the collection of photos and pictures of the 1WW Bibliotheca di Storia Moderna e contemporanea, in Rome ; liturgical manuscripts, Civic Museum in Bologna) ;
- Creation of sw particularly devoted to project MANUS. MANUS is a database including the unified and national catalogue on manuscripts, incunabula and XVIth century printed editions in Italy and includes MANUS is a database including description a digitised images of manuscripts in library both public and private, both civic and religious.

Practises for Digital library

Bibliotheca nazionale Braidense (www.braidense.it) National library in Milan, may be represent the most developed digital experience in Italy today. Digital Library is an active dept. in library.(<http://www.braidense.it/risorse/index.php#digitale>)

Braidense acts in the general criteria - both i) and ii) - according to a particular chronological plan : the XIX Century and the beginning of the 20 th in Lombardy. Digitisation includes monographs (above all novels and particularly popular literature) and periodical collections.

Monographies. Two digitization project are working out

1. *Progetto Di. Re.* (Digital Resources projects)

Monographs are readable in XHTML Os by download *pdf format.

A complete list both of monographies is on the web page

<http://www.braidense.it/risorse/dire.php>

2. *Scaffale Aperto* (Open Library) A short list of works, manuscripts devoted to didactic purpose.

A complete list of monographs and manuscripts is on the web page

<http://www.braidense.it/risorse/testi.php>

Periodicals. Actually 867 periodical collections digitized, XIXth XXth century to 'Fifties. Most collections are local and regional publications

www.braidense.it/risorse/emeroteca.php

Braidense works as a library hub and take digitisation project from different libraries. Particularly 3 different catalogs :

EPOCA Emeroteca politica e culturale antifascista (Antifascist Cultural a and political Newspaper and periodical collections) [in cooperation with Istituto Nazionale per la storia del movimento di liberazione in Italia www.parrimilano.italia-liberazione.it/ita/istituti.php?rete=1

EVA Open virtual Newspaper and periodical collection) GEA Daily weekly and monthly collections without index

GEA Giornali e altro (daily, weekly and bi-weekly periodical collections without index.

Single institutions as Fondazione Basso Issoco or Fondazione Gramsci digitised periodical collection or single works, but they have not a programme neither as preservation nor as digital library.

BIBLIOTECA DIGITALE ITALIANA
RIEPILOGO PROGETTI ANNI 2001 - 2004

ESERCIZIO FINANZIARIO	PROGETTO	ENTE PROPONENTE
2001	Digitizzazione cataloghi	ICCU ROMA
2001	Fondo Sommariva	Biblioteca Nazionale Braidense MILANO
2001	Digital library	Biblioteca Nazionale Braidense MILANO
2001	Microfilm giornali	Biblioteca Nazionale Braidense MILANO
2001	Digitalizzazione fondi librari	Biblioteca Reale di TORINO
2001	Digitalizzazione fondo fotografie	Biblioteca Vallicelliana di ROMA
2001	Catalogazione patrimonio librario	Regione Piemonte Biblioteca del Seminario Vescovile di BIELLA
2001	Catalogazione patrimonio librario	Regione Piemonte Biblioteca Provinciale dei Cappuccini TORINO
2002	Digitalizzazione fondi musicali (ADMV)	ICCU ROMA in collaborazione con B. Naz. Centr. TORINO, B. Estense di MODENA B. Naz. Centr. ROMA e Bibl. Naz.le Marciana di VENEZIA
2002	Digitalizzazione fondi librari	Biblioteca Reale di TORINO
2002	Adeguamento hw e sw	ICCU ROMA
2002	Digitalizzazione inventari biblioteche medievali	Società Internazionale per lo studio del Medioevo Latino FIRENZE
2002	Biblioteca Digitale Galileiana	Istituto e Museo di Storia della Scienza FIRENZE
2002	Rinascimento virtuale	Biblioteca Medicea Laurenziana FIRENZE
2002	Contributo acquisto scheda informatica MANUS	Biblioteca Medicea Laurenziana FIRENZE
2002	Digitalizzazione partiture autografe di Giovanni Simone Mayr	Biblioteca Civica Angelo Mai BERGAMO
2002	Digitalizzazione fondi storico culturali della biblioteca	Conservatorio Statale di Musica "Luigi Cherubini" FIRENZE
2002	Digitalizzazione manoscritti liturgici	Civico museo Bibliografico Musicale BOLOGNA
2002	Musica sacra in Augusta - digitalizzazione mss. musicali	Regione Umbria B. Augusta PERUGIA
2002	adeguamento hw e sw	Biblioteca Marucelliana di FIRENZE
2002	Mare Magnum	Biblioteca Marucelliana di FIRENZE
2002	Aggiornamento standard	CIBIT ROMA
2002	Adeguamento hw r sw	Biblioteca Estense di MODENA
2002	Conversione digitale cataloghi	Regione Umbria B. Augusta PERUGIA
2002	Catalogazione in SBN fondi librari	Regione Piemonte Centro Culturale Diocesano di SUSA
2002	Catalogazione in SBN - A di ca. 1700 cinquecentine	Regione Piemonte Seminario Vescovile di ASTI
2003	Rinascimento virtuale	Biblioteca Medicea Laurenziana FIRENZE
2003	Contributo acquisto scheda informatica MANUS	Biblioteca Medicea Laurenziana FIRENZE
2003	Aggiornamento standard	CIBIT ROMA
2003	Biblioteca Digitale Galileiana	Istituto e Museo di Storia della Scienza FIRENZE
2003	Digitalizzazione inventari biblioteche medievali	Società Internazionale per lo studio del Medioevo Latino FIRENZE
2003	Digitalizzazione periodici eruditi	ICCU ROMA
2003	Digitalizzazione raccolte Fondo Calogerà	Istituto e Museo di Storia della Scienza FIRENZE
2003	Digitalizzazione fondi storico culturali della biblioteca	Conservatorio Statale di Musica "Luigi Cherubini" FIRENZE
2003	Digitalizzazione manoscritti musicali	Conservatorio di Musica di San Pietro a Maiella NAPOLI
2003	Digitalizzazione manoscritti fondo antico presso Sacro Convento di Assisi	Società Internazionale di studi Francescani di ASSISI
2003	Digit. Periodici e quotidiani locali altoatesini	B. Prov.le Italiana Claudia Augusta BOLZANO

Grassetti/BDI/riepilogoprogetti2001

2004	Rinascimento virtuale	Biblioteca Medicea Laurenziana FIRENZE
2004	Digitalizzazione manoscritti fondo antico presso Sacro Convento di Assisi	Società Internazionale di studi Francescani di ASSISI
2004	Partecipazione "Progetto Crawler dell'International Internet Preservation Consortium"	Biblioteca Nazionale centrale FIRENZE
2004	Digitalizzazione Fondo Pannunzio	Biblioteca Nazionale centrale FIRENZE
2004	Biblioteca Digitale della prima Accademia dei Lincei e Biblioteca Digitale della Scuola Galileiana	Istituto e Museo di Storia della Scienza FIRENZE
2004	Testi storiografici relativi alle regioni, città, monasteri e conventi italiani dall'antichità all'età moderna	SISMEL
2004	La fabbrica dell'italiano. Cinque secoli di strumenti	Accademia della Crusca FIRENZE
2004	Gli storici italiani dell'Umanesimo e del Rinascimento	Istituto Nazionale di Studi del Rinascimento FIRENZE
2004	Digitalizzazione della Collana degli Scrittori italiani	CIBIT ROMA
2004	Digitalizzazione del Fondo manoscritto appartenuto a Clementino Sala	Accademia Filarmonica Romana ROMA
2004	Digitalizzazione del Fondo della Fondazione Maria e Goffredo Bellonci	Fondazione Maria e Goffredo Bellonci
2005	Digitalizzazione Fondo fotografico Sommariva	Biblioteca Nazionale Braidense di MILANO
2005	Digitalizzazione immagini della grande guerra	Biblioteca Storia Moderna e Contemporanea di ROMA
2005	Sapere Scientifico nei secoli XVIII e XIX secolo	Istituto e Museo di storia della scienza
2005	Sapere Scientifico nei secoli XVIII e XIX secolo	Accademia delle scienze di Torino
2005	risorse HW teca per progetto incunaboli	ICCU
2005	digitalizzazione bandi ed editti	Bibl. Casanatense di Roma
2005	comitato di redazione portale internet culturale	Dipartimento di Italianistica Università la Sapienza ROMA
2005	digitalizzazione incunaboli in volgare	Dipartimento di Italianistica Università la Sapienza ROMA
2005	Digitalizzazione dei fondi cartografici	Bibl. Nazionale Marciana di VENEZIA
2005	Migrazione in SBN e creazione MAG per il materiale cartografico della Società Geografica Italiana	ICCU
2005	Digitalizzazione fondo Donizetti Biblioteca Angelo Mai	ICCU
2005	digitalizzazione manoscritti della biblioteca Medicea Laurentiana di Firenze	Biblioteca Medicea Laurentiana di Firenze
2005	coordinamento scientifico progetti manoscritti della biblioteca Medicea Laurentiana di Firenze	Biblioteca Medicea Laurentiana di Firenze
2005	evento di comunicazione progetto manoscritti	Biblioteca Medicea Laurentiana di Firenze
2005	risorse HW e SW teca progetto manoscritti	Biblioteca Medicea Laurentiana di Firenze

Expériences de numérisation dans les bibliothèques italiennes

David Bidussa

(david.bidussa@fondazionefeltrinelli.it)

Situation et critères généraux

De manière générale, les numérisations en Italie ont augmenté ces dix dernières années (le premier document officiel national sur la numérisation a été approuvé en février 2001 à la fin de la 3e conférence de la bibliothèque nationale (Padoue, 14-16 février 2001) et l'objectif principal était la conservation.

La numérisation en Italie est d'abord conservation. Ensuite la conservation signifie numérisation de textes seuls ou de documents. Les programmes et projets de numérisation signifient numérisation de textes et non bibliothèque numérique.

C'est assez pour analyser le programme de numérisation de la BDI (Bibliothèque numérique italienne), le département de numérisation sous l'égide du ministère des biens culturels (toutes les données sont disponibles à l'adresse suivante :

<http://www.librari.beniculturali.it/genera.jsp?s=31&lingua=it>)

Il est possible de distinguer trois plans différents :

- Reproduction des collections. (La numérisation correspond aux microformes 30 ou 40 ans plus tôt)
- Organisation et création dans une collection unifiée d'une documentation homogène (Numérisation de la collection cartographique de la bibliothèque de Venise et de la collection de photos et de peintures de 1WW Biblioteca di Storia Moderna e contemporanea, à Rome ; manuscrits liturgiques, au Musée civique de Bologne)
- Création de sources spécifiquement dévolues au projet Manus. Manus est une base de données incluant les catalogues unifiés et nationaux des manuscrits, incunables et éditions du XVI^e siècle en Italie.

Expériences de bibliothèques virtuelles

La Bibliothèque nationale Braidense (www.braidense.it) , la bibliothèque nationale de Milan représentent peut-être les expériences numériques les plus développées en Italie aujourd'hui. Le département numérique est un département très actif de la bibliothèque

« <http://www.braidense.it/risorse/index.php%23digitale> »

Les numérisations de la Braidense correspondent aux critères 1) et 2) selon un plan chronologique, XIX^e et début du XX^e en Lombardie. La numérisation inclut les

monographies (surtout romans et littérature populaire) et des collections de périodiques.

Monographies. Deux projets de numérisation sont en cours :

1) Progetto Di. Re.

Les Monographies sont en XHTML ou téléchargeables au format PDF.

<http://www.braidense.it/risorse/dire.php>

2) Scaffale Aperto (« Open Library ») Une courte liste de travaux pédagogiques.

Liste complète voir :

<http://www.braidense.it/risorse/testi.php>

Périodiques. Actuellement 867 collections de périodiques sont numérisées : XIX^e et XX^e jusqu'aux années 50. La plupart sont des collections régionales et locales

www.braidense.it/risorse/emeroteca.php

La Braidense travaille comme un *hub* et intègre les projets numériques de différentes bibliothèques

Particulièrement, trois catalogues

EPOCA "Emeroteca" politique et culturelle antifasciste (en coopération avec l'Institut National pour l'histoire des mouvements de libération en Italie)

www.parr-milano.italia-liberazione.it/ita/istituti.php?rete=1

EVA Newspapers and periodical collections)

GEA Giornali e altro (Périodiques quotidiens, hebdo et bi-hebdo sans Index)

Des institutions comme la Fondation Basso Issoco ou la Fondation Gramsci numérisent certains ouvrages ou périodiques mais n'ont pas de véritables programmes en ce domaine.

Numérisations réalisées en histoire ouvrière et sociale

Françoise Blum et Franck Veyron
(CHS du XX^e siècle et BDIC)

Pour commencer, nous allons rappeler quelques données générales, qui ne concernent pas spécifiquement l'histoire ouvrière et sociale, mais plus généralement les programmes nationaux (eux-mêmes articulés à des programmes internationaux) en matière de numérisation d'archives, et d'accès au patrimoine numérisé. <http://www.archivesdefrance.culture.gouv.fr/gerer/numerisation/>)

Les archives de France et le ministère de la culture mettent en œuvre un plan national de numérisation, qui concerne évidemment au premier chef les archives publiques (nationales ou départementales) mais qui peut profiter aussi à d'autres institutions publiques ou privées qui ont la possibilité de répondre aux appels du plan numérisation national. Un portail donne ainsi accès au patrimoine numérisé de 540 institutions, qui ont bénéficié (ou non) des soutiens publics. Tout fonds numérique peut être, sous certaines conditions, intégré à ce portail. Les archives de France participent d'autre part au projet européen Michael (Multilingual Inventory of cultural heritage in Europe) qui offre un portail de ressources numériques européennes. Parmi les fonds des institutions du Codhos ici représentées, les affiches de Génériques par exemple sont visibles via ce portail.

À l'échelle nationale, toujours, il existe un programme nommé Adonis dont un axe comprend la création à moyen terme d'un Centre de Sources et de Ressources Numériques des mondes contemporains (XIX^e-XX^e siècles) dont la partie liée à l'histoire du mouvement ouvrier et de la pensée critique, est une des dimensions.

Pour ce qui est des « born digital archives », les archives de France sont actuellement engagées dans le projet « Pilae », dont le but ultime est de mettre en œuvre une plate-forme pilote, testée par les archives de France et susceptible d'être réutilisée. Le projet est actuellement en phase test.

Enfin, il faudrait signaler une troisième initiative nationale, émanant de la BNF, extrêmement intéressante, bien qu'encre, semble-t-il, dans les limbes. C'est le programme SPAR (Système de Préservation et d'Archivage Réparti), qui vise à offrir des magasins, y compris à des partenaires extérieurs. La Bibliothèque nationale de France se constituerait ainsi en tiers-archiviste, suivant des modalités contractuelles encore à définir.

Nous ne parlerons pas ici des numérisations massives des documents imprimés, qui ont constitué la base Gallica, articulée à Europeana, puisqu'il en sera question cet après-midi. D'autre part, les numérisations de la BNF concernent d'abord les imprimés, et non les archives, même si on trouve dans Gallica de très nombreux documents iconographiques.

Il s'agit des grandes lignes d'une politique nationale qui se cherche parfois encore, ce qui, je crois, n'est pas propre à la France. Il s'agit aussi d'initiatives qui peuvent être soumises à des aléas liés aux diverses décisions du pouvoir (et celui-ci, en France cherche actuellement à faire le plus d'économies possibles privilégiant des objectifs parfois définis selon des modalités politiques qui ne sont pas toujours les nôtres) Mais évidemment, à côté de ces grands chantiers, il existe de multiples initiatives, des centaines et des milliers de documents sont numérisés ou en cours de numérisation sur le territoire.

Nous ne donnerons que quelques exemples, sans viser, bien entendu, à l'exhaustivité, exemples pris dans le champ de la Labour History et donc émanant des institutions membres du Codhos. Je remercie tous ceux qui ont bien voulu m'envoyer un état des lieux de leurs numérisations.

Un gros effort a été fait, semble-t-il, de manière générale sur les images, affiches et photographies. La BDIC a une base images accessible en ligne comprenant 17 000 photos, 6 000 dessins et peintures, 6 000 affiches comprenant notamment le fonds Gabriel Ersler sur la guerre d'Espagne, un très beau fonds d'archives polonaises, les fonds Kagan et Pottier sur Mai 68. Génériques a numérisé 1 850 affiches de l'immigration et en numérisera 500 de plus en 2009. La Fondation Jean Jaurès a mis en ligne 4 000 clichés provenant de l'Université syndicaliste et/ou représentant les principaux acteurs du socialisme français. L'Institut d'histoire sociale de la CGT a une photothèque de 600 000 images, provenant des services iconographiques des journaux *Antoinette* et *La Vie Ouvrière* mais aussi des archives de la Confédération et de celles de certaines fédérations comme la Métallurgie. L'IRHSES, qui est l'Institut de documentation du SNES (Syndicat national de l'enseignement secondaire) a numérisé 5 000 portraits de dirigeants nationaux, mais qui ne sont pour l'instant accessibles que sur place. On commence donc à avoir des fonds iconographiques numérisés en ligne qui couvrent de larges domaines de la vie ouvrière, syndicale ou socialiste.

À mi-chemin entre l'image et l'archive, il faut donner sa place à la très intéressante expérience faite par le CAARME, (Centre d'animations, d'archives et de recherches sur les mouvements étudiants) en partenariat avec d'autres institutions telles que le CHS et la BDIC : la numérisation des tracts et (je cite) des « expressions éphémères » de mai-juin 1968. Le CAARME est par ailleurs à l'initiative de la création du CME (Centre des mémoires étudiantes). Ces tracts sont numérisés en mode image mais indexés de façon assez sophistiquée.

Toujours à mi-chemin entre l'image et l'archive, Le centre Georges Chevrier et la MSH de Dijon sont engagés dans un projet de numérisation des plans de mine :

Ce projet, à mettre en œuvre dans le cadre d'un partenariat entre l'Académie François Bourdon (située au Creusot), le Muséum d'Histoire Naturelle d'Autun et la Maison des Sciences de l'Homme de Dijon porterait sur deux exploitations du bassin de Bourgogne-Nivernais : la houillère du Creusot, les schistes bitumineux d'Autun

En ce qui concerne la presse, les initiatives sont trop nombreuses pour être citées : Nous avons commencé à discuter avec les représentants de la BNF des numérisations de l'ensemble de la presse socialiste et ouvrière, ensemble qui a été inventorié par le Codhos, dans un guide récent.

Les différents documents liés aux congrès ouvriers et socialistes sont des corpus particulièrement prisés dans les projets de numérisation. Je vous ai déjà mentionné la première expérience menée par le Codhos avec la BNF sur le corpus 1870-1940, je n'y reviendrai donc pas. Pour les périodes postérieures, la Fondation Jean Jaurès a numérisé et mis en ligne les sténotypies des débats des organismes centraux du Parti socialiste et l'IRHSES a numérisé les congrès du SNES depuis 1953, ceux de la FEN des années 70s, 80s. La Fraternelle et l'AMSAB ont monté un projet pour numériser les congrès, dispersés dans plusieurs centres de l'Alliance universelle des ouvriers diamantaires.

Enfin en matière d'archives stricto sensu une place d'honneur doit être réservée aux initiatives d'une part du Centre Georges Chevrier et de la MSH de Dijon, et d'autre part à la BDIC. Les premiers sont les créateurs de la base INCOMKA d'archives du communisme, poursuivie actuellement avec les numérisations des documents de l'Internationale concernant la France. Les seconds ont déjà numérisé 63 000 pages d'archives, dont beaucoup - mais pas seulement - autour de la guerre de 14. Je ne peux résister à l'envie de citer l'exemple des journaux des tranchées, opération menée en collaboration avec la BNF.

Au vu de toutes ces expériences et projets passionnants, un constat cependant s'impose : chacun y va de manière dispersée, sans véritable coordination d'initiatives. Les moyens mis en œuvre d'une opération à l'autre sont fort divers, rudimentaires ou plus sophistiqués, avec la BDIC par exemple qui met en place l'OAI en collaboration avec la BNF. Quant aux formats des numérisations et aux appareils scientifiques et techniques mis en place pour en assurer la visibilité, ils sont aussi de tous ordres : Numérisation en mode image pour les archives le plus souvent, en mode texte pour les documents imprimés, indexation au nom des documents.

Tout cela est malgré tout fort encourageant. Il me semble néanmoins que la leçon à en tirer est bien la nécessité d'une coordination sur la base de la réunion de tous les partenaires concernés, voir d'une planification des efforts dans le domaine de l'histoire ouvrière et sociale, et la très nécessaire construction d'une plateforme

OAI et d'un portail. Peut-être cette journée va-t-elle faire avancer les choses, je le souhaite en tout cas. Le Codhos comme l'IALHI est confronté à la nécessité de changer d'échelle et de se donner les outils techniques qui apporteront un plus aux numérisations de masse, tout en ne laissant pas aux bords de la route les plus petites institutions, celles qui justement possèdent parfois les fonds les plus originaux et les plus rares.

FRANCE

Digitisation in labour history

Françoise Blum et Franck Veyron
(CHS du XX^e siècle et BDIC)

Firstly, we would like to give you some general information about national digitisation programs — programs that, of course, do not particularly deal with the history of the labour movement.

The French department of Archives (Archives de France) and the Ministry of culture (Ministère de la culture et de la communication) have the charge to plan a national program of digitisation¹⁶. Of course, they are mainly interested in public archives — the ones of the National archives or the local "Archives départementales" — but other public or even private institutions can also present their initiatives and join the program. They developed an Internet portal (called Patrimoine numérique) through which one can get access to the digitised heritage of 540 different institutions, who did receive (or not) a public financial support. In fact, under certain conditions, any digital collection can be add to that portal. At the international level, the Archives de France contributes to the European project Michael (Multilingual Inventory of cultural heritage in Europe), a multilingual service through which you can "explore the European digital cultural heritage", i.e. find "digital collections from museums, archives, libraries and other cultural institutions from across Europe"¹⁷. The poster collections of the association Génériques, from the CODHOS, can be seen, for example, on their website.

¹⁶ <http://www.archivesdefrance.culture.gouv.fr/gerer/numerisation>

¹⁷ <http://www.michael-culture.org/en/home>

At the French national scale, another digitisation program called ADONIS is also developed under the authority of the CNRS (the French national office for scientific research)¹⁸. Its general aim is to offer access to digital data and documents in the field of Social and Human Sciences. As part of its program, it offers to create, in the next years, a Centre of digital sources and resources for contemporary studies (Centre de sources et de ressources numériques des mondes contemporains, XIX-XXe siècles), on which there will certainly be some resources about the history of the labour movement.

If we consider the specific preservation of born-digital archives, the Archives de France are at present developing a digital preservation system to allow ingestion, secure storage and dissemination of digital records at the French National Archives. After having been tested, this system, called PILAE, could be used by other institutions.

And finally, some words must also be said about a program developed by the French national library (Bibliothèque nationale de France, BnF), which seems very interesting for the future, though only at its beginnings. Called SPAR (Système de préservation et d'archivage réparti -Distributed archiving and preservation system), this digital preservation repository could be open to other institutions, under certain conditions that still need to be defined.

We will not speak now of the different mass digitisation of printed documents, which were the basis of the first Gallica digital library, now Gallica 2, linked to the European project Europeana : our colleagues from the National library will tell you everything about them this afternoon. I just would like to underline the fact that the digitisation programs launched by the National library are mostly about printed materials, books and periodicals, more than real archives, even if you can find many photographs in Gallica.

So here are the main characteristics of the French national digitisation policy, which in a sense is still not definitively defined — and I think the situation is more or less the same in other countries. This general policy and their particular applications can also suffer the effects of the changing views of the French political bodies — our Government is now trying to save money as much as possible, and its purposes considering digitisation are not always based on the same preoccupations than ours.

Of course, besides the main national programs, many French institutions undertook and still undertake smaller digitisation projects. Hundreds of thousands of documents are already digitised.

This initiatives are too numerous to be all presented today. We will just give you some brief examples of what has been made in the field of labour history. I thank again all CODHOS members who sent me information about their activities.

¹⁸ <http://www.tge-adonis.fr>

Firstly, it seems that a big effort has been made for documents with an iconographic element images — posters and photographs. The BDIC, for example, built a database of digital images that already holds 17 000 photographs, 6 000 drawings and paintings and 6 000 posters, among them the Gabriel Ersler collection of photographs about the Spanish civil war, a collection of Polish posters from the after war period up to the 1980s, or the Elie Kagan or Jean Pottier's collections of photographs¹⁹. *Génériques*, dedicated to the history and memory of immigration, already digitised more than 1 800 posters and will keep on adding 500 more posters to its digital collection next year²⁰. The Jean Jaurès Foundation offers online on its website 4 000 photographs from their holdings, representing the main figures of the history of socialism in France. The Institute for social history from the CGT trade-union (HIS-CGT) began the digitisation of its photo-library of 600 000 images coming from the photograph collections of the newspapers *La Vie ouvrière* et *Antoinette*, the archives of the Confédération or even some federal archives like the metal workers' ones²¹. The IRHSES — the documentation and research centre created by one of our secondary education professors' trade union, the SNES (Syndicat national de l'enseignement secondaire) — digitised 5 000 portraits of national leaders of their trade-unions²². As you can see, digital collections of photographs already cover a large part of the French labour movement and socialist history.

Halfway between images and archives, one must underline the very interesting experience undertaken by our colleagues from the CAARME Mission (Centre d'animations, d'archives et de recherches sur les mouvements étudiants), whose work is dedicated to the archives and history of French student movements. Thanks to cooperation with other institutions like the Centre d'histoire sociale du XX^e siècle and the BDIC, they recently digitised the leaflets and other "ephemeral expressions" — that was their expression — of the May-June 68 student movement to finally create an online Digital Journal of the student Commune (*Journal électronique de la Commune étudiante*)²³. The actual 520 documents have been digitised in an image format, but their precise description allows different text researches.

The Georges-Chevrier centre in Dijon also began with the local MSH (Maison des sciences de l'homme ; Social Sciences and Humanities Research Institute) the digitisation of mine plans. That project will also be the result of a partnership with other local institutions.

¹⁹ <http://www.bdic.fr>

²⁰ <http://www.generiques.org>

²¹ <http://www.ihs.cgt.fr>

²² <http://www.irhses.snes.edu>

²³ <http://www.cme-u.fr/commune68>

In we now look at what has been made for the digital preservation of newspapers and periodicals, examples are too numerous to be all quoted. Let just say that we recently discussed with our colleagues from the National library about the possible digitisation of the French socialist and labour publications published between the Paris Commune and the First World War — publications that have already all been listed in an inventory guide the CODHOS published in 2007.

The different papers — accounts and reports, published speeches, etc. — that document the workers and socialists' congresses seem to particular interest our French colleagues. We already presented you several times, during previous IALHI conferences, the collaboration between the CODHOS and the National library for the documents published between the Paris Commune and 1940. For the post-war period, some of us went further by themselves. For example the Foundation Jean-Jaurès digitised and made available online the full reports of the debates of congresses, national conventions and director boards held by the Socialist party between 1971 and 1997. The same way, the IRHSES did digitise the reports of SNES congresses held since 1953, as FEN's ones — FEN is a teachers' trade unions — for the 1970s and 1980s. The archives of La Fraternelle (in Saint-Claude, Jura) participates, altogether with the AMSAB and the International Social Institute in Amsterdam, to an international program for the digitisation of the congresses of the Universal Alliance of diamond workers.

If we now speak of archives according to their literal definition, the projects developed in Dijon by the Georges Chevrier center with the local Social Sciences and Humanities Research Institute, and by the BDIC, must be underline. The firsts contribute, by digitising the Communist International archives concerning France, to the international Communist International (Comintern) Archives Project (INCOMKA)²⁴. And the BDIC already added to its digital library 63 000 pages of archives, mostly concerning, a choice one can understand considering the origins of this library, the First World War period : a partnership with the National library permitted the digitisation of a large collection of rare "journaux de tranchées" — periodicals written, designed and published by the soldiers in the trenches during the First World War)²⁵.

Seeing the variety of all these experiences, we can only come to the following conclusion : all institutions seem to follow their own private agenda, and they do not always try to coordinate their projects with other organisations. From one institution to another, the technical and financial means an institution can rely on are very different. Whereas some projects are totally self-supported and developed by benevolent activists (we could give the example of the work made by another CODHOS member, the newly named RADAR — an association for the

²⁴ <http://www.loc.gov/rr/european/comintern/comintern-project.html>

²⁵ http://www.bdic.fr/journaux_tranchees_intro.html

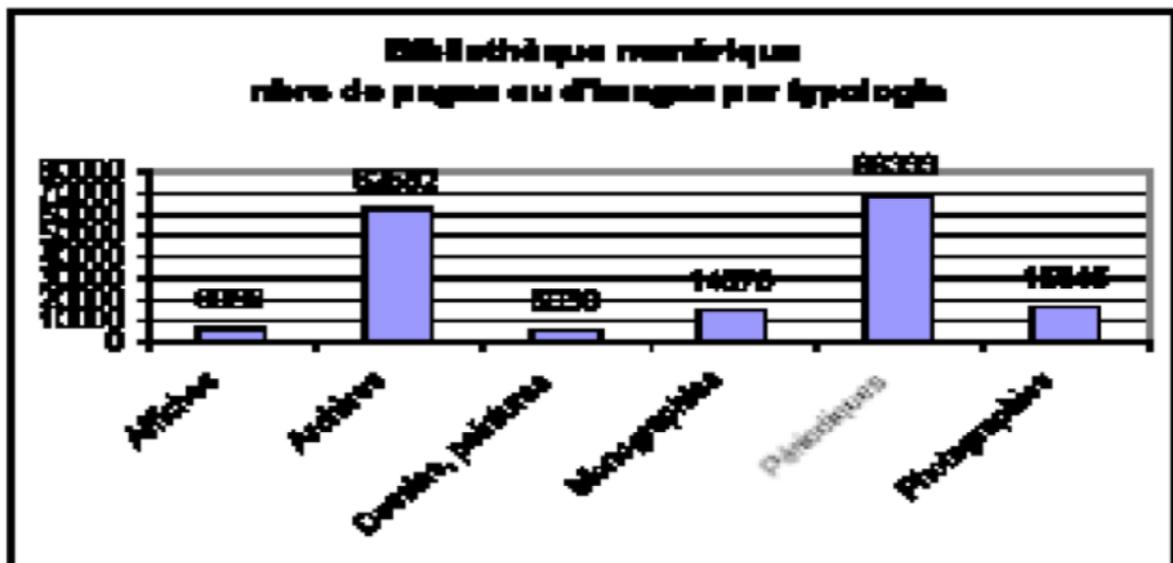
preservation of the memory of the Fourth International French section —, to digitised newspapers, leaflets and other printed documents), other institutions can implement more sophisticated models, like the BDIC who installed an OAI repository which will facilitate its cooperation with different partners, and particularly the National library.

Standards used for the digitisation process and protocols then developed to facilitate the visibility and use of digital documents also change a lot from one institution to another, from different image files formats to text files ones (especially for originally printed materials), and the attention accorded to the cataloguing and indexing of digital documents vary also a lot.

The work made makes us optimistic. Nevertheless better exchanges and a better coordination between institutions seems to be more and more necessary. They could lead, for example, to a to the development — by the CODHOS — of a real portal for the archive of the French labour movement, and of an OAI platform. I hope today's debates and discussions will help us to go further that way. Like the IALHI, the CODHOS now has to rethink its activities on a new scale, and needs to find a way to appropriate the technical tools that allow massive digitisation programs. But whatever our future will be, we will not forget the particular worries and needs of our smallest members — institutions that sometimes hold the rarest resource and archives.

BDIC

La BDIC a commencé la numérisation de ses collections en 2003. Au 31/12/2008, **175 000 pages ou images ont été numérisées** : 63 000 pages/archives, 17 000 photographies, 68 000 pages/périodiques, 15 000 pages/monographies, 6 000 dessins et peintures, 6 000 affiches



COLLECTIONS NUMERISEES

GUERRE 1914-1918

- Journaux de tranchées : journaux rédigés et réalisés par les soldats pendant la guerre (en partenariat avec la BnF)
- Académie de Lille (archives) : enquête réalisée en 1920 sur la guerre dans l'Académie de Lille (Aisne, Ardennes, Nord, Pas-de-Calais, Somme)
- Affaire du Bonnet Rouge (archives) : révision du procès du « Bonnet Rouge » (publication anarchiste de Miguel Almereyda, accusé de trahison en 1917)
- Fusillés de Souain (archives) : campagne de réhabilitation de la Ligue des Droits de l'Homme en faveur des Fusillés de Souain (1917)
- Fonds Mantoux (archives) : Conférences diplomatiques et militaires entre Alliés, Conférences du Conseil Supérieur de la Guerre
- Conférences de la paix : procès-verbaux et résolutions 1919-1920
- Dessins et peintures : Vallotton, Steinlen... et de nombreux combattants partis sur le front avec un carnet de croquis ou d'aquarelle
- Gaston Lavy : récit manuscrit et illustré de la vie quotidienne au front par un soldat
- Historiques de régiments de l'infanterie française
- Collection Valois (photographies réalisées de 1914 à 1920 à la demande de la Section Photographique des Armées) : Verdun, le département du Nord, la

ville de Reims, l'Alsace et la Lorraine, les photographies prises par Charles Gicquet de Preissac.

PACIFISME (Fonds Jules Prudhommeaux) : activité pacifiste des associations, organisations... de 1887 à 1938 (brochures, correspondances, tracts...)

GUERRE D'ESPAGNE : photographies de Gabriel Ersler, engagé au Service Santé de la 13e Brigade Internationale ; affiches.

UNEF : correspondance de l'Union Nationale des Etudiants de France de 1940 à 1945 (en partenariat avec Germe)

AFFICHES POLONAISES : 1945 à 2004

IMMIGRATION

Textes préparatoires à la Convention de Genève

Périodiques de l'immigration espagnole

Photographies de Jean Pottier : bidonville de Nanterre (1956-1975), travailleurs immigrés, manifestations contre le racisme...

MAI 1968 : photographies d'Elie Kagan et Jean Pottier

ARGENTINE (archives d'associations, brochures et périodiques)

Droits de l'homme en Argentine

Répression en Argentine

Madres de la Plaza de Mayo

Politique et gouvernement en Argentine : brochures et périodiques sur le péronisme et les partis et mouvements en exil

TYSAE : travailleurs et syndicalistes en exil

COBA : Comité pour le Boycott de la coupe du Monde en Argentine

Movimiento Peronista Montoneros : documents divers et périodiques

REVUE MATERIAUX : accessible sur le portail Persée

BASE ARCHIVES ET IMAGES

Accessible sur le site internet de la BDIC, **elle permet l'accès à tous les documents numérisés**. Pour des raisons de droits d'auteur, certaines images sont accessibles seulement dans nos locaux (ex : affiches polonaises). Elle donne accès également au catalogue du musée en cours d'informatisation : affiches, dessins, peintures, objets et photographies.

TECHNIQUES

La plupart des documents sont numérisés par un prestataire extérieur. Depuis 2005, un atelier interne permet de numériser les documents les plus fragiles (objets, dessins et peintures...).

La numérisation se fait en mode image pour les fonds iconographiques et les archives, en mode texte (permettant la recherche intégrale plein texte) pour les documents imprimés.

PROJETS

Thématiques : Poursuite des projets : guerre 1914-1918, périodiques de l'immigration.

Nouveaux projets : Affiches de la Commune et de la guerre 1870-1871, archives de l'immigration russe, collections audiovisuelles (sauvegardes urgentes des fonds sonores de la BDIC).

Valorisation : **Signalement** des fonds numérisés dans le Sudoc et dans notre catalogue. Un lien URL permet l'accès direct au document numérisé. - Mise en place de l'**OAI** : les journaux de tranchées numérisés par la BDIC et la BNF seront par l'OAI accessibles à la fois sur Gallica et sur le site BDIC. - Achèvement du **parcours thématique** sur l'Espagne républicaine et de l'**exposition virtuelle** sur l'Ozerlag.

Pérennisation : La BDIC a acheté en 2008 un serveur dédié à la conservation des données numériques (9 To). Les données y sont maintenant stockées. A terme, nous souhaitons confier à la BNF la pérennisation de nos données numériques (projet SPAR).

PARTENARIATS

La numérisation permet et rend obligatoire le partenariat. La BDIC développe chaque année de nouveaux partenariats, qui permettent un meilleur accès aux documents pour les chercheurs. Ce travail en coopération permet d'éviter de numériser plusieurs fois le même document et donne accès aux chercheurs à l'ensemble des fonds, même dispersés dans plusieurs bibliothèques. Pour être efficace dans ces partenariats, il nous faut utiliser l'OAI.

BNF. Le partenariat avec la BnF se développe : 14-18, le pacifisme, la presse de l'immigration, la Commune. L'OAI permettra de mieux valoriser ce partenariat.

Mission CAARME. La mission Caarme continue de numériser des fonds de la BDIC sur les mouvements étudiants.

Ministère de la Défense. En 2007, la BDIC a entrepris la numérisation des historiques de régiments 14-18. Un partenariat se met en place à compter de 2008 avec le Ministère de la Défense qui souhaite numériser tous ses historiques pour cette période.

Génériques. La BDIC a prêté à Génériques des compléments de collections pour son projet de numérisation des périodiques portugais et maghrébins de

l'immigration. Les numérisations seront accessibles sur les sites de l'un et de l'autre.

FINANCEMENTS

Une part du budget de fonctionnement de la bibliothèque (sous tutelle du Ministère de l'Enseignement et de la Recherche) est dévolue annuellement à la numérisation.

Pour les projets menés en partenariat avec la BNF, une subvention annuelle est versée à la BDIC, dans le cadre de la convention de pôle associé en histoire contemporaine.

EXPERIENCES FRANÇAISES • 2

Génériques et la numérisation du patrimoine de l'immigration

Génériques - 34 rue de Citeaux - 75012 PARIS
Téléphone : 01 49 28 57 75 - Télécopie: 01 49 28 09 30

Génériques a commencé la publication de sa base de description d'affiches en 1998, et les premières images y furent attachées deux ans plus tard. La base présente actuellement plus de 1850 affiches de l'immigration.

En 2008, nous avons amplifié le volume mis à disposition grâce à l'éligibilité des projets sur le patrimoine de l'immigration pour un financement par le plan de numérisation du ministère de la Culture et de la Communication : 500 nouvelles affiches, et près de 15 000 pages de « journaux de l'immigration » seront mis en ligne via notre outil Internet en cours de refonte.

Notre programme 2009, s'il est financé entièrement, prévoit de continuer sur les mêmes volumes que l'année précédente : 500 nouvelles affiches, 15 000 pages de journaux, mais s'ouvrent aussi sur de nouvelles typologies : quelque 400 partitions manuscrites de chansons francarabes, 500 chansons (sous forme d'extraits sur notre site, en audition complète dans nos locaux) viendront compléter les sources mises à disposition en ligne sur notre site, après sa refonte.

Nous nous concentrons sur le patrimoine privé de l'immigration, prolongeant ainsi le travail d'archives de *Génériques* engagé depuis plus de 20 ans. Mais nous voulons aussi permettre aux lecteurs de porter un autre regard que celui de l'encadrement étatique sur des acteurs de la vie publique française souvent minorés, en ayant en vue la production des immigrés eux-mêmes.

Nous élaborons des politiques de publication de corpus cohérent et de masse critique, tenant à la fois compte de l'offre déjà présente en ligne et des programmes en cours, notamment le plan de numérisation partagée de la presse entre la BNF et la BDIC qui traitent des populations allemande et italienne (BNF) ou espagnole (BDIC), des collections et des droits que nous arrivons à recueillir.

La « presse de l'immigration » que nous publierons est donc issue des populations maghrébine et portugaise, ou de l'immigration sans spécification de communautés. Le coeur de nos journaux va des années 1960 aux années 1990, englobant le moment charnière du début des années 1980. Nous souhaitons ainsi former un *hub* sur la presse de ces populations, en attendant la création d'un portail thématique sur l'ensemble des ressources pour l'histoire de l'immigration.

Les affiches proposées sont issues de notre collecte auprès des acteurs de l'immigration et sur les murs des villes de nos partenaires, et débutent, sauf de rares exceptions, au cours des années 1960.

Mais nous souhaitons également élargir la variété des types de sources accessibles en publiant des partitions manuscrites de chansons d'un Tunisien installé à Paris, à la fois auteur, compositeur, interprète et musicien central de la scène parisienne de l'immigration de 1946 aux années 1980. Complété par des biographies de leurs auteurs et interprètes, ainsi que des enregistrements des chansons de l'immigration issues d'avant 1958, cet ensemble formera un corpus permettant l'étude et la ré-interprétation contextualisée d'une production culturelle mal- et méconnue, au moment de l'âge d'or du « cabaret franco-maghrébin » à Paris et de la professionnalisation des musiciens de l'immigration maghrébine au cours des années 1940 et 1950. Nous pensons ainsi rendre leur statut d'oeuvre aux musiques migrantes, leur offrant une occasion de légitimation publique ainsi qu'une possibilité de ré-interprétation contemporaine.

La publication de ces collections d'origines privées permet de mettre en avant auprès d'un large public les activités culturelles, sociales ou politiques de ces composantes de la société française, par un regard des acteurs portés sur eux-mêmes, avec un aperçu du foisonnement et de la variété de ces activités.

Nous voulons ainsi permettre l'émergence de nouvelles figures d'acteurs des populations immigrées depuis la Seconde Guerre mondiale, aussi bien dans l'historiographie que dans les représentations sociétales, en permettant à la figure de l'immigré comme créateur culturel de prendre corps et racine dans plusieurs de ses modes d'expressions.

La Fraternelle

Alain Mélo, archiviste et historien,
La fraternelle, Saint-Claude, Jura, France

En 2006, suite à une première rencontre organisée par l'AMSAB entre des délégués patronaux et syndicaux diamantaires belges, puis à une seconde rencontre, en 2007, entre la branche diamantaire du syndicat ouvrier belge FGTB, nous avons imaginé de numériser les congrès de l'Alliance Universelle des Ouvriers Diamantaires (fondée en 1905) afin de mettre en ligne ces informations de base, alors que les archives papier se trouvent dispersées entre l'AMSAB d'Anvers (Belgique), l'IISG d'Amsterdam (Pays-Bas, fonds de l'ANDB) et les Archives de la Maison du peuple de Saint-Claude (La fraternelle, Jura).

Il convenait d'abord de repérer les fonds, puis de faire le point d'une part sur les congrès effectifs et sur les rapports rédigés de ces congrès. C'est chose faite. Il reste à trouver le financement de cette numérisation. Pour la part Fraternelle, une partie devrait être couverte par une subvention régionale.

Mission CAARME

Quelques éléments : projets de numérisation en cours de la Mission CAARME/
Cité des mémoires étudiantes (nom après préfiguration)

• *Que numérisez-vous actuellement/Quels sont vos projets de numérisation ?*

Des corpus homogènes et cohérents de journaux étudiants (des années 1950 à nos jours) relatifs à la santé et à la mutualité (étudiantes) - conservés dans notre structure et dans différentes institutions-, mais aussi de la presse étudiante régionale (notamment Rhône Alpes et Champagne-Ardenne) ainsi que des tracts et expressions éphémères du mouvement de mai-juillet 1968.

Les journaux champardennais sont issus des fonds classés de la Mission CAARME/Cité des mémoires étudiantes.

L'ensemble de ces documents numérisés sont mis en ligne au fur et à mesure sur le site de mutualisation du réseau du Conservatoire des mémoires étudiantes (www.cme-u.fr).

• *Quels autres projets/programmes sont menés parallèlement (ou en liaison) avec les numérisations (Archivage, accessibilité, valorisations etc.) ?*

Collecte et traitement documentaire et multimédia de témoignages oraux sur le thème de la santé et de la mutualité étudiantes : anciens militants mutualistes, anciens étudiants ayant séjourné en sanatorium ou salariés de la Fondation Santé des Étudiants de France (FSEF).

• *Sur quelles coopérations vous appuyez-vous : nationales/internationales en ce qui concerne les numérisations ?*

En dehors de la numérisation de nos propres fonds/collections, deux types de coopération nous permettent de mener à bien les différents projets de numérisation :

- La coopération directe : partenariat privilégié avec une autre institution : nous avons une convention avec la BDIC qui permet notamment la numérisation de la presse nationale étudiante issue des fonds qu'elle conserve. Nous avons également une convention avec les Archives du communisme pour la numérisation et la mise en ligne du journal *Clarté*.

- Un travail en réseau plus large avec d'autres partenaires dans le cadre du réseau du CME, dont certains sont également membres du réseau du CODHOS.

Regroupant sous forme associative institutions patrimoniales et professionnels de la conservation (archivistes, bibliothécaires, documentalistes, ...), anciens responsables du mouvement étudiant et chercheurs, le CME vise à mettre en réseau les différents acteurs voulant s'impliquer dans la sauvegarde et la valorisation des archives et mémoires étudiantes.

Outre le GERME, l'AAUNEF et la Mission CAARME, le CME compte déjà parmi ses membres actifs : les Archives du communisme, les Archives du rectorat de Paris, les Archives municipales de Grenoble, Lyon et Saint-Étienne, les Archives de l'Université de Haute-Alsace, Atelier @lternet, la BDIC (Bibliothèque de documentation internationale contemporaine, Nanterre) et le Centre d'histoire sociale du XXe siècle (Paris 1/CNRS).

La Mission CAARME, à Reims, y assure un rôle de tête de réseau afin d'y renforcer la mutualisation et la logique de projet partagé.

C'est dans ce cadre que les Archives municipales de Lyon, la BDIC (Bibliothèque de Documentation internationale contemporaine) de Nanterre, le CHS (Centre d'histoire sociale) du XXe siècle (Université Paris 1) et la Mission CAARME (Centre d'animation d'archives et de recherches sur les mouvements étudiants) ont uni leur force pour numériser, mettre en ligne et indexer les diverses expressions éphémères de mai-juillet 1968 (tracts, journaux, motions rapports de commissions et comptes-rendus de réunion...) : projet intitulé « Le journal électronique de la Commune étudiante ».

Et c'est également dans ce cadre que nous menons le projet « Mémoires vives étudiantes : santé et mutualité » (numérisation de journaux étudiants, recueil et traitement de témoignages oraux, mise en ligne) avec la Fondation Santé des

Étudiants de France (FSEF), l'association Guy Renard, la BDIC et les Archives municipales de Grenoble.

- *Avec quels financements fonctionnez-vous, nationaux et/ou internationaux ?*
« Mémoires vives étudiantes en Champagne-Ardenne » : principalement mené par la Mission CAARME, avec le soutien de la Ville de Reims et de l'Université de Reims-Champagne-Ardenne.

Le projet « Mémoires vives étudiantes : santé et mutualité » est financé à hauteur de 50 % par la Mission recherche et technologie du Ministère de la Culture (appel à projets de numérisation 2009)

- *Quels sont vos principaux problèmes et limites (techniques, financiers, conceptuels, organisationnels etc.) ?*

Alors que, pour le « Journal électronique de la Commune étudiante », le traitement documentaire permet une description très fine des documents, pour la presse étudiante c'est une simple reconstitution numérique des collections, sans réelle indexation ou autre traitement documentaire, qui est proposée.

Tout ce qui est numérisé est mis en ligne sur le site de mutualisation du réseau du CME mais les collections numérisées de chaque institution n'apparaissent pas sur leur propre site internet.

- *Quelles perspectives supplémentaires dans ce domaine vous semblent offrir les réseaux (Codhos/IALHI) ?*

La possibilité de mettre en place un entrepôt OAI avec moissonnage est à la fois une perspective et un enjeu majeurs.

La poursuite et l'amplification du travail sur le mouvement étudiant de mai-juin 1968 au niveau international est également une perspective intéressante.

EXPERIENCES FRANÇAISES • 5

IRHSES

CONDITIONS ET RESULTATS DE LA NUMERISATION DU FONDS D'ARCHIVES

Le contexte et la décision

L'IRHSES (Institut de Recherches Historiques sur le Syndicalisme dans les Enseignements de Second degré) est une association créée en 1985 par le SNES (Syndicat National des Enseignements de Second degré) dans le double but de conserver les archives du syndicalisme enseignant de second degré et de participer

à la formation syndicale. Le syndicat disposait de son propre fonds documentaire (congrès, circulaires, correspondance, comptes-rendus d'instances...), de fonds privés de militants, de documents provenant d'autres organisations (syndicats, fédération, associations pédagogiques...) et d'un important fonds photographique. La conservation de ces documents dont la masse augmentait rapidement, fut engagée dès la création de l'institut et, au milieu des années 90, un réel travail de classement confié à des militants aidés d'une salariée à mi-temps fut entrepris selon la méthode du classement linéaire en série W. Un déménagement et l'incapacité des nouveaux locaux à accueillir les quelque 450 ml de boîtes d'archives conduisirent à la décision de numériser tout ou partie de notre fonds. Un plan sur cinq ans fut mis en place en 2005 qui donnait la priorité à la documentation interne du SNES, dans un premier temps les congrès nationaux, les circulaires internes et les réunions des instances syndicales ; une seconde vague fut consacrée au fonds photographique et enfin la troisième étape, actuellement en cours concerne l'action syndicale, les relations avec le Ministère de l'Éducation Nationale (réunion des Conseils) et les différents partenaires (syndicats, associations de parents...), les documents liés au paritarisme (commissions) et enfin les productions des sections régionales (académiques et départementales). Déjà, suite à des déménagements successifs, une partie des archives des sections départementales et académiques du SNES ont disparu et il y a encore des difficultés à obtenir de ces niveaux une prise en compte réelle de l'archivage pour les recherches futures.

La mise en place du travail de numérisation

Le travail préparatoire fut très important, il l'est encore pour trier ce qui est numérisable et ce qui ne l'est pas, tant au niveau de l'intérêt du document que de sa qualité. Il fallut ensuite rendre lisible notre classement initial par l'organisation de dossiers et de sous-dossiers identifiables. Ensuite il fallait choisir une entreprise avec qui travailler, se garantir de son savoir-faire pas vraiment au plan technique, encore que la question ne soit pas secondaire, mais surtout au plan de la compréhension de notre demande. Il nous fallut à la fois organiser le travail du numérisateur et anticiper sur celui du chercheur, assurer un suivi des opérations. Ici le logiciel de recherche est très important, il doit pouvoir s'adapter à la demande et être suffisamment souple d'utilisation. Avec l'outil utilisé (Visual doc) nous avons pu mettre en place une nomenclature précise de classement et de recherche qui permet aux chercheurs plusieurs entrées, par exemple par année, événement, lieu, acteur... et également pour chaque document une rubrique commentaire sur laquelle nous pouvons intervenir pour préciser le contenu. Pour tous les documents, le logiciel permet un recadrage et bien sûr la reproduction totale ou partielle. Ce logiciel ayant été mis au point par l'entreprise à notre demande, il peut encore évoluer pour s'adapter toujours mieux au travail de recherche.

Dans une petite structure comme la nôtre, la question du coût n'est pas mince c'est pour cette raison que nous avons mis en place un plan pluri annuel.

Enfin se pose la question du devenir des documents originaux, notre choix fut celui du reversement pour partie aux Archives du Monde du Travail de Roubaix et pour ce qui est des documents concernant la région parisienne aux Archives Départementales de la Seine Saint- Denis.

Le travail réalisé.

Nous avons aujourd'hui numérisé la totalité des congrès nationaux du SNES depuis 1953, la totalité des circulaires nationales sur la même période et une bonne partie des congrès de la FEN des années 70 et 80. En ce qui concerne les archives photographiques, une partie du fonds provenant de l'Université Syndicaliste est actuellement numérisée soit plus de 5 000 images ainsi que les portraits en situation des militants nationaux des années 70 et 80. L'ensemble est actuellement stocké sur le serveur du SNES (plusieurs sauvegardes sur DVD ont été réalisées) et disponible à la consultation dans nos locaux. Une sélection des ces documents concernant mai juin 1968 est en outre en ligne sur le site www.irhses.snes.edu et le site du CODHOS consacré à mai 1968.

La question de l'accessibilité. Quelle diffusion et comment ?

La qualité du travail produit est ici essentielle, les documents doivent être lisibles et facilement reproductibles.

Pour le moment nos archives numérisées sont disponibles dans les mêmes conditions que les archives papier, c'est-à-dire dans nos locaux. C'est à la fois un outil de documentation interne et un outil pour les chercheurs. Nous nous posons la question d'une mise à disposition plus large par le biais d'une mise en réseau interne (documentation pour les militants du SNES) et externe dans le cadre de réseaux partagés (dans le cadre du CODHOS par exemple) ou d'une mise en ligne sur un site propre.

La difficile question de la conservation des archives numériques

Le stockage et l'archivage des documents numérisés, hormis qu'il nécessite un suivi technique, pose toute une série de questions auxquelles il est difficile de répondre aujourd'hui telles que la conservation, la nature des supports et la durée de vie de ces derniers.

De plus un champ nouveau s'ouvre actuellement à l'archiviste, la conservation des documents produits par les outils informatiques. Nous sommes actuellement le plus souvent dans le domaine de l'éphémère, du fait entre autre de la vitesse de communication nouvelle, le temps du mail n'est pas celui de la lettre. Il faut d'abord faire acquérir un réflexe de conservation (mails, page web, communication interne et externe), puis de tri (le nécessaire nettoyage des mémoires informatiques saturées ne devant pas se résumer à l'écrasement des fichiers) et enfin de classement sous la forme de dossiers sur des modèles très proches du classement

papier. Ce domaine nécessite des compétences techniques nouvelles et une formation adaptée des archivistes et documentalistes.

Conclusion

La numérisation est à beaucoup d'égards un moyen de conservation et surtout de diffusion des archives très intéressant et aujourd'hui incontournable. C'est une réponse moderne et efficace à la conservation de la mémoire dans une période où la masse des informations augmente très rapidement. Se posent toutefois deux questions bien réelles, celle du coût d'une part et d'autre part celle de la fiabilité dans la durée des supports des produits numérisés et des moyens de les lire.

EXPERIENCES FRANÇAISES • 6

Fondation Jean-Jaurès

Programme de numérisation de documents d'archives

BILAN 2001-2009

La Fondation Jean-Jaurès a fait le choix en 2001 de mettre en place un *programme de numérisation* d'envergure concernant les documents d'archives du mouvement socialiste français dont elle assure la conservation et la mise en valeur. Ce projet a toujours suivi un double objectif : changer de support des documents de plus en plus fragiles (transfert sur CD-ROM et DVD-ROM), et les mettre à la disposition du public, soit dans la salle de consultation du centre d'archives, soit en ligne sur le site Internet de la Fondation, soit par la présentation de reproductions de qualité dans des publications ou lors d'expositions.

Les opérations de numérisation sont confiées depuis l'origine à un prestataire spécialisé dans le conseil en NTIC et en management de l'information, la société OUROUK (www.ourouk.com Paris), choix constamment renouvelé compte tenu de la satisfaction procurée par leur travail. Chaque année, depuis 2001, le Ministère de la culture et de la communication nous accorde une subvention pour nos projets.

Numérisation des débats du PS et mise en ligne d'une base de données

En 2001, la numérisation des sténotypies des débats des organismes centraux du Parti socialiste (congrès, conventions nationales, comités directeurs, conseils nationaux) a débuté, et plusieurs phases sont en cours ou achevées :

2001-2003 : numérisation des documents pour la période 1971-1981 (environ 6 000 pages). Il s'agissait concrètement de numériser les documents originaux (fichiers PDF), et de les transformer en fichiers texte (type Word ou .RTF).

2006-2007 : Numérisation et indexation des sténotypies des congrès et conventions nationales du Parti socialiste, de 1981 à 1997 : documents

originaux sur support papier et reliés, mais également sous forme de fichiers informatiques à partir de 1994 pour certains (fichiers Mac). 17 000 pages environ

2008-2010 : En 2008, et 2009 sera poursuivie la numérisation et l'indexation de l'ensemble des débats de la SFIO et du Nouveau parti socialiste de 1939 à 1971 (50 000 pages de débats environ, congrès, conseils nationaux, journées d'études, conférences, comités directeurs), entre 1939 et 1971. Ce projet d'envergure, étalé sur trois années, a commencé en 2008 avec le traitement de 30 % des documents, comprenant les congrès de 1939 à 1955. Il est prévu en 2009 de poursuivre la numérisation et l'indexation des congrès et des conseils nationaux de 1955 à 1971, soit 40 % de l'opération environ. Les volumes seront massicotés, numérisés, et reliés à nouveau. Chaque manifestation numérisée sera sauvegardée et disponible en fichier PDF sur DVD. Chaque document sera ensuite transformé en fichier Word, relu et corrigé, et intégré dans la base de données en ligne.

La base de données concernant les manifestations de 1971 à 1997 est en ligne depuis l'automne 2008 sur www.jean-jaures.org

Elle sera complétée progressivement en 2009 et 2010 pour les documents de la période 1939-1971.

Ont été ajoutés à chaque manifestation les numéros des journaux internes du Parti (téléchargeables, fichiers PDF) se rapportant à la manifestation (par exemple, pour un congrès, textes des contributions et motions)

Presse socialiste / Base de donnée en ligne L'Unité

Depuis 2003 la presse socialiste, très consultée par les chercheurs, fait l'objet d'un projet de numérisation ambitieux. Ces journaux, souvent manipulés par les chercheurs alors qu'ils sont de plus en plus fragiles, doivent changer de support afin de leur assurer une meilleure conservation.

En 2003, *Le Poing et la Rose*, journal interne (135 numéros, 1972-1992). et *Le Poing et la rose supplément Responsables*, (557 numéros, 1979-1993) ont ainsi été numérisés, le résultat est sous format PDF et consultable sur CD-ROM par les chercheurs. Certains numéros ont été mis en ligne sur notre site Internet pour compléter la base de données des débats des organismes centraux.

Depuis 2004, et jusqu'en 2008 pour l'achèvement définitif, l'opération la plus importante de numérisation de presse s'est concentrée sur l'hebdomadaire *L'Unité* (675 numéros, 1972-1986) : ce projet qui a été planifié sur plusieurs années, compte tenu de l'ampleur et des difficultés techniques de la tâche, est actuellement achevé. Chaque numéro du journal devient un fichier image, et chaque article est découpé électroniquement grâce au logiciel Newsclip® ; parallèlement au fichier image, un fichier texte permet d'isoler le titre, la date et le corps de l'article, ce qui rend celui-ci accessible pour une recherche en texte intégral. Chaque article est ensuite indexé dans une base de données, selon un thésaurus

préalablement défini. Les unes du journal, les photographies, les dessins et les caricatures sont aussi extraits, facilitant l'exploitation iconographique par Internet ou plus globalement pour de la recherche documentaire.

Fin 2006, la base de données a été mise en ligne sur le site Internet de la Fondation Jean-Jaurès. Durant l'année 2007, elle a fait l'objet d'un contrôle qualité rigoureux de la part du prestataire qui l'a conçue. Fin 2008, une journée d'études à la Fondation Jean-Jaurès a permis d'en faire la présentation officielle.

Les affiches

Depuis 2004, numérisation des affiches du Parti socialiste en plusieurs lots : deux produits numériques seront remis à l'issue des opérations techniques : un DVD-ROM de conservation (fichiers TIFF) et un DVD-ROM, de consultation (format JPG). Au total, 700 images à ce jour.

En 2009, numérisation de la collection de diapositives de Joseph Daniel, 700 clichés environ, consultant en communication et membre de la commission propagande / communication du PS de 1971 à 1981 (affiches, projets d'affiches, affichage commercial, etc.)

A terme, projet de constitution d'une base de données en lignes des affiches du mouvement socialiste (PS, MJS, fédérations départementales socialistes).

Les photographies

Collection de 4 000 clichés numérisés, disponibles en ligne sur www.jean-jaures.org ; DVD de conservation fichiers TIFF ; droits d'auteur partagés.

Numérisation en 2006 de photos diverses extraites des archives du Parti socialiste et des fonds privés de Claude Estier et de Robert Pontillon : 1 500 photographies, de formats divers, sur supports divers (quelques diapositives notamment) ont été numérisées au printemps 2006 avec une résolution de 600 points par pouce et 256 niveaux de gris. Le résultat est fourni sur DVD en format TIFF. Des DVD de consultation sont disponibles, format JPG.

Dijon • ressources numériques

Centre Georges-Chevrier
(UMR 5605) - MSH Dijon

Le centre Georges-Chevrier et la MSH de Dijon ont acquis depuis une dizaine d'années une expérience solide dans la numérisation des archives du communisme : base INCOMKA et numérisation des archives de direction du Parti communiste français. Depuis, ils se sont engagés dans un programme de numérisation de certaines ressources liées à l'histoire du monde ouvrier et de la pensée critique. Ce programme, qui s'inscrit dans le cadre de la constitution des Très Grands Equipements (TGE) Adonis, vise à moyen terme à construire un Centre de Sources et de Ressources Numériques des mondes contemporains (XIX^e-XX^e siècles) dont la partie liée à l'histoire du mouvement ouvrier et de la pensée critique, n'est qu'une des dimensions.

Autour de l'histoire du mouvement ouvrier et de la pensée critique, plusieurs opérations de numérisation sont en cours ou prévues :

Numérisation des brochures de la Bibliothèque Marxiste de Paris (BMP)

Les brochures de la Bibliothèque Marxiste de Paris forment un fonds important pour l'histoire sociale du XIX^e et du XX^e siècles. Ces documents, qui ont été souvent édités par des partis politiques ou des mouvements syndicaux n'ont pas toujours fait l'objet du dépôt légal et ne figurent pas dans le fonds de la Bibliothèque Nationale de France ou des bibliothèques universitaires. Au total, ce sont près de 2 500 brochures qui seront prochainement accessibles.

Conservées pendant de nombreuses années à la Bibliothèque Marxiste de Paris, ces brochures ont dû être déménagées lors de la fermeture du local. Les brochures fragilisées par leur support vieillissant (les pages cassent), et donc difficilement communicables, ont été numérisées au service de numérisation de la MSH de Dijon avec un numériseur de livres précieux et plans. Les fichiers images obtenues ont fait l'objet d'un traitement informatique et sont consultables actuellement dans un format pdf interrogeable en texte intégral.

Numérisation des plans de mine du bassin de Bourgogne-Nivernais

La conservation des archives minières et en particulier des plans d'exploitation constitue un enjeu de taille dans la mesure où elle reste souvent le seul moyen de connaître précisément la localisation des travaux miniers passés, en particulier ceux qui ont eu d'importants développements au cours de la période contemporaine. Le projet de numérisation des plans de mine à mettre en œuvre dans le cadre d'un partenariat entre l'Académie François Bourdon (située au Creusot), le Muséum

d'Histoire Naturelle d'Autun et la Maison des Sciences de l'Homme de Dijon porterait sur deux exploitations du bassin de Bourgogne-Nivernais qui correspondent à des horizons géologiques légèrement différents :

- la houillère du Creusot ;
- les schistes bitumineux d'Autun.

Ces deux exploitations ont été retenues pour plusieurs raisons. Il s'agit d'exploitations de moyenne dimension, dont on n'a jamais extrait plus de 350 000 tonnes par an. Par conséquent, même si, dans les deux cas, les travaux miniers se sont étalés sur plus d'un siècle, le travail de numérisation à entreprendre ne débouche pas sur une ambition démesurée comme aurait pu constituer un travail identique avec les mines de Blanzay.

Les archives concernant ces deux exploitations sont de nature différente, mais présentent des séries assez complètes. Dans le cas du Creusot, il s'agit des archives minières des établissements Schneider et Cie, conservées au Creusot, dans les locaux de l'Académie François Bourdon. Dans le cas des schistes bitumineux, il s'agit des archives « concessions » conservées par la DRIRE à Dijon et actuellement déposées au muséum d'Autun.

Grande Edition Marx et Engels (GEME)

Dans le cadre de la Grande Edition Marx et Engels réalisée sous l'égide des Editions La Dispute, qui prévoit la traduction intégrale des œuvres de Marx et Engels sous un format électronique, une convention a été signée entre la Dispute, l'association GEME, le centre de recherche de l'histoire des systèmes de pensée moderne, la fondation Gabriel Péri, et le Centre Georges Chevrier. Dans ce cadre, une numérisation des traductions anciennes de Marx et Engels a été réalisée, qui sera à terme accessible en ligne, et qui permettra d'analyser le corpus marxien accessible en français et ses usages depuis le XIX^e siècle.

La « question française » dans les archives de l'Internationale communiste (1919-1943)

Fort de son expérience de la numérisation des archives du communisme, la MSH de Dijon entreprend avec le soutien de la Direction des archives de France la poursuite de la dématérialisation des archives de l'Internationale communiste permettant l'étude de ce phénomène politique en France. Une première campagne vise à numériser l'ensemble des documents provenant du fonds 517 inventaire 1 actuellement disponible en France (provenant aussi bien de la collection des microfilms autrefois déposée à la Bibliothèque marxiste de Paris et aujourd'hui consultable aux Archives départementales de la Seine-Saint-Denis que de la collection de microfilms complémentaire achetée par l'université de Bourgogne). Une deuxième campagne visera à numériser les autres documents concernant la France provenant des autres fonds de l'Internationale communiste à l'exception des documents numérisés dans le cadre du projet Incomka.

International workshop for an international concerted policy of labour history archives digitization : digitization and digital archives in Belgium and Amsab

Maarten Savels

(archivist at Amsab-Institute for Social History, Ghent)

Introduction

This presentation aims to give you a concise overview of projects and programs on digitization and digital archives in Belgium and in Amsab-ISG. It consists of three chapters:

- Digitization
- Digital born archives and digital storage
- Noteworthy projects in Belgium

Digitization in Amsab-ISG

This first part focuses on the digitization strategies employed in Amsab. Our digitization policy is primarily based upon three questions, where the answer to each question has an important influence on the others: why, what and how.

First of all: why do we digitize? We can identify three important reasons.

We digitize in order to preserve our archival material. Preservation is a strategy employed for materials that are quite popular, but that start to suffer from frequent use. They start to discolour and fray around the edges, but have lost none of their informational value yet. However we know that, if they'll be given to researchers in the future, they might start to suffer more seriously. The digitization is then used to create a snapshot of a source that is still in good condition. The materials themselves are then no longer shown to the public, except for special purposes, such as an exposition. Researchers will only be given access to the digital copy.

In more critical cases, this strategy can be modified. This happens when we digitize to replace the original. This substitution strategy is only used in very special cases, when the materials have suffered so much damage, that the informational value might be lost forever. The aim of this strategy is to make the best possible copy, even if it means that the original can get severely damaged or even destroyed in the process. Naturally, this is a worst case scenario and as such

rarely used. In this strategy, it is paramount that the scans are done at the best possible quality, since the digital copy becomes the new original.

Lastly, we can digitize for dissemination purposes. This strategy is mainly used to accommodate the needs of the users. One benefit is the possibility of putting the digital copies on a website, so that they can be viewed worldwide. Another important, but often-overlooked benefit is the ability to enlarge texts, which is very practical for hard-to-read manuscripts or for researchers suffering from a visual impairment.

The second question is: what do we digitize? We cannot scan every single scrap of paper in our collection, as that is too costly and simply unnecessary. The choice of materials is determined by three factors between which we must find a compromise. These factors are popularity, importance and physical condition. Material that is popular, important and has suffered severe damage has without a doubt top priority. Rarely consulted, trivial and mint-condition papers have lowest priority.

The third question is: how do we digitize? We can either do it internally on our own -albeit primitive- scanners or contact a specialised firm. For large quantities or for very vulnerable originals, an external firm is addressed. The same applies for very technical scanning (such as Betamax tapes) or scanning where quality is the top priority. For smaller series (say up to fifty volumes) where the quality of the scan is less important (like for preservation or dissemination scans), we use our own equipment and software. One of the reasons for this is that external scanning tends to cost more per page when the total amount of pages is rather low.

Digital born archives and digital storage

Over the past few years we have begun acquiring digital born archives. Luckily for us, up to now, we weren't confronted with huge masses and the ones we did get, still have to remain shut for at least a few more years. In other words, now is the time to start thinking about how we should handle them. I believe they will present a big challenge for a small institution such as Amsab. The enormous bulk, the technical nature, the complexity of the archives and the difficulty of adding metadata pose important questions that need to be answered. Of course, a lot has already been written on the subject, but a lot of these reports don't really satisfy. They tend to be either very technical, or they tend to be too generic and require a lot of further research from the archival institution, which is something that is usually hard to do for the smaller ones. For the moment, we are still relying on the methods that we use for paper documents, when it comes down to selection and registration. We can't integrate metadata with the digital objects, so we just make use of the possibilities that are offered by our catalogue software. The same goes for digital storage. The past years I have attended a few meetings where enthusiastic archivists show models and prototypes of their digital repositories to

the equally enthusiastic crowd. However, when the costs are mentioned, the spirits tend to get dampened a bit. Sums of 250.000 and up to 3 million euros were mentioned. And that's just the start-up costs. I believe this amount of money is too much for a lot of institutions. However, I don't think this proves that a small institution is incapable of setting up a digital repository, it just proves that it will be very hard if you do it on your own. The point I'm trying to make is that, for me, as a member of a smaller institution, cooperation is the best guarantee for success. In the meantime, while we are trying to find possible candidates for potential cooperation, we are using a sort of simplified solution that keeps costs low and keeps the door open for a lot of possibilities. To put it in a more concrete way: we're using a simple server, running on windows, with a very rigorously managed folder structure to which very few people have access. Back-ups are taken on a regular basis and after every major change (e.g. the ingest of new records). Whenever possible, we migrate the old bitstreams to TIFF or PDF.

With cooperation, however, I don't just mean working together as Labour History Institutions, but also working with records creators. I think most institutions represented here today share the same problem. Most of us are not structurally bound to the records creators. In other words, we don't have a legal framework that gives us some kind of grip on the way records are created and stored before they are delivered to us. A city archive, for example, is normally part of the city's administration, so it's safe to say that the head of the department should have at least something to say about the way the city's administration creates and manages its records before they are handed over to the archive. We often don't have this luxury, so we must find other ways of steering the records creator. Luckily, badly managed archives are as much a problem for us as for the records creator. So we offer help. We try to come to some sort of mutually beneficial agreement that gets put down in a cooperation contract. Up till today we have made contracts with six records creators. These contracts can have quite a broad perspective and don't necessarily focus only on assisting with records creation. Furthermore, there is no fixed procedure in the assistance. This would prove to be very difficult, since every institution has its own problems and questions. The sort of help we offer ranges from simple advice and information to a comprehensive analysis of the institutional structure, the workflow and the hard- and software in order to give very thorough advice on adapting some of the institution's policies. This sort of assistance works quite well for promoting the use of standards and open source software as well as helping to create better folder structure and preventing the enormous bulk of redundant copies. The downside of this approach is that it often is a sort of one-time project, so for the future we are also considering to create a more permanent helpdesk-function through which the archive creators can ask questions.

Noteworthy projects in Belgium

In Belgium there are a few interesting projects currently running both in the fields of digitization and digital storage.

One of the most interesting projects is the denBell-project. The city council of Antwerp decided a few years ago to centralise their entire administration and to introduce a more flexible working environment. Two key points in this project are the centralisation of all the files and documents in one single cluster of servers and the redesign of the work- and document flows. The city archives of Antwerp - who already had a good reputation when it comes to digital archiving- were actively consulted during the preparation and design phases and got abundant funding for creating a top-quality digital repository. Macz -as the repository is called- is currently in its test-phase, but should be up and running before the end of this year. Before and during the preparation phase of this project, the Antwerp city archives created a very valuable batch of documents concerning every aspect of document creation and digital storage.

The second interesting project is the BOM-VI-project. The acronym stands for storage and dissemination of multimedia data in Flanders. This project is a testbed for digitization, storage and dissemination of audiovisual data. It aims to create an example of how an archival institution should handle its audiovisual data and it focuses on every single aspect of this task, ranging from selection over metadata to the creation of a working prototype of a digital repository. The project normally ends this year, but there are plans for making the repository permanent and offering trustworthy storage to participating institutions.

The third project is the Presse-Papier-project. This project is a large-scale digitization project, carried out by the national library in order to better protect and disseminate the huge collection of newspapers in their collection, but also in other collections throughout Belgium. The aim is to digitize thirty newspapers from the nineteenth century up to 1950 and a collection of clandestine papers from both world wars. In total this should amount to about three million pages. A side project is the Abraham-database. This database was constructed to collect detailed information about every possible newspaper printed in Belgium. This information includes the availability and quality of every newspaper, so that adequate steps can be taken to protect -i.e. digitize- threatened collections.

Conclusions

That wraps it up for what I have to say. I just want to make a short summary of what I believe are the most important points for me.

I believe that the entire issue of digital born archives and digital storage is a big challenge, but that we should not be afraid of addressing it. The longer we wait, the bigger the problem will be. The key to success -especially for smaller institutions- is cooperation and communication. Furthermore, this cooperation

should be as broad as possible. Networks like IALHI are very important, but we should also look towards records creators, national archives and -if possible- soft- and hardware firms. Only through a collaborative effort will we be able to fully and adequately overcome this challenge.

BELGIQUE

Numérisation et archives numériques en Belgique et à l'AMSAB

Maarten Savels
(archivist at Amsab-Institute for
Social History, Ghent)

Introduction

Cette présentation vise à donner une vue synthétique des projets et programmes en matière de numérisation et d'archives numériques en Belgique et à l'AMSAB. Cette présentation comporte trois chapitres :

- numérisations ;
- « Born digital archives » et stockage numérique ;
- grands projets en Belgique.

Numérisations à l'AMSAB

Cette première partie porte sur les stratégies de numérisation utilisées à l'AMSAB. Notre politique de numérisation est essentiellement fondée sur trois questions dont la réponse à chacune influence les autres : pourquoi, quoi et comment.

Avant tout : Que devons-nous numériser ? Nous pouvons dégager trois importantes raisons.

Nous numérisons pour sauvegarder notre matériel archivistique. La sauvegarde est une stratégie employée pour des matériels qui sont plutôt populaires, mais qui commencent à souffrir de la fréquente demande. Ils commencent à se décolorer, s'émiettent sur les bords, mais n'ont aucunement perdu leur valeur informationnelle. Cependant nous savons que, si nous continuons à les donner à des chercheurs, ils vont souffrir beaucoup plus sérieusement. La numérisation est alors utilisée pour créer une photo d'une source qui est encore bonne. Les matériaux eux-mêmes ne

sont plus communiqués au public, exception faite d'occasions exceptionnelles telles des expositions. Les chercheurs n'ont plus accès qu'à la copie numérique.

Dans des cas plus critiques, la stratégie peut être modifiée. Cela arrive quand nous numérisons pour remplacer l'original. Cette stratégie de substitution est utilisée seulement dans des cas très particuliers, quand le matériel a souffert de tels dégâts que la valeur informationnelle peut être perdue à jamais. Le but de cette stratégie est de faire la meilleure copie possible, même si cela signifie que l'original peut être sévèrement endommagé ou même détruit dans le processus. Naturellement, c'est le pire scénario rarement utilisé. Dans cette stratégie, il est capital que les scans soient de la meilleure qualité possible, puisque la copie numérique devient le nouvel original.

Dernièrement, nous pouvons être amenés à numériser dans une logique de diffusion. Cette stratégie est principalement utilisée pour répondre aux besoins des usagers. Un avantage est la possibilité de mettre les copies numériques sur un site web, de manière à ce qu'elles puissent être vues partout dans le monde. Un autre avantage important, mais souvent mal perçu, est la possibilité d'agrandir les textes, ce qui est très pratique pour les manuscrits difficiles à lire ou pour les chercheurs souffrant de déficience visuelle.

La seconde question est : que devons-nous numériser ? Nous ne pouvons scanner chaque bout de papier de nos collections, cela serait trop coûteux et tout simplement inutile. Le choix des matériaux est déterminé par trois facteurs entre lesquels nous devons trouver un compromis. Ces facteurs sont la popularité, l'importance et la condition physique. Le matériel qui, souvent demandé, important et en mauvaise condition a sans aucun doute la priorité. Les matériaux rarement consultés, communs et en bonne condition ont évidemment une priorité moindre.

La troisième question est : comment numérisons-nous ? Nous pouvons soit le faire en interne avec nos propres scanners primitifs ou prendre contact avec une entreprise spécialisée. Pour de grandes quantités ou des originaux très fragiles, une entreprise est recommandée. C'est la même chose dans le cas de scans très techniques ou les scans où la qualité est prioritaire. Pour de plus petites séries (disons jusqu'à 50 volumes) où la qualité du scan est moins importante (comme pour la sauvegarde ou la diffusion), nous utilisons notre propre équipement et logiciel. L'une des raisons en est que le scan extériorisé coûte plus cher à la page quand le nombre total des pages n'est pas élevé.

"Born digital archives" et stockage numérique

Il y a quelques années, nous avons commencé à acquérir des « born digital archives ». Heureusement pour nous, jusqu'à maintenant, nous n'avons pas été confronté à d'énormes masses et celles que nous avons acquies doivent rester fermées au moins quelques années encore. En d'autres termes, il est maintenant temps de commencer à penser comment nous pourrions les traiter. Je pense que

cela représentera un grand défi pour une petite institution comme l'AMSAB. L'énormité du volume, la nature technique, la complexité de ces archives et la difficulté d'ajouter des métadonnées posent d'importantes questions que nous devons résoudre. Bien sûr, beaucoup a déjà été écrit sur le sujet, mais nombre de ces rapports restent insatisfaisants. Ils ont tendance à être soit très techniques soit trop généraux et demandent des recherches supplémentaires aux centres d'archives, ce qui est souvent difficile à faire pour les plus petits d'entre eux. Pour le moment, nous sommes encore en train d'utiliser les méthodes que nous avons pour le papier quand cela n'est plus valable pour la sélection et l'enregistrement. Nous ne pouvons intégrer les métadonnées aux objets numériques, aussi nous utilisons seulement les possibilités offertes par notre logiciel de catalogage. Les années passées, j'ai assisté à de nombreuses réunions où des archivistes enthousiastes montraient des modèles et des prototypes de leurs réservoirs numériques à une foule également enthousiaste. Cependant, quand il fut question des coûts, les esprits furent quelque peu rafraîchis. Il fut fait mention de sommes allant de 250 000 et plus jusqu'à 3 millions d'euros. Et ce ne sont que les coûts de démarrage. Je pense que ces sommes sont beaucoup trop importantes pour bien des institutions. Cependant, je ne pense pas que cela prouve qu'une petite institution soit incapable de mettre en place un réservoir numérique. Cela prouve seulement que cela sera très difficile à faire avec ses propres forces. Ce que je veux dire par là est que, pour moi, en tant que membre d'une petite institution, la coopération est la meilleure garantie du succès.

En même temps, tout en essayant de trouver de possibles candidats pour cette coopération, nous utilisons une sorte de solution simplifiée à bas coûts et qui laisse la porte ouverte à de multiples possibilités. Pour dire cela plus concrètement : nous utilisons un simple serveur, sous Windows, avec une structure de répertoire très rigoureusement organisée à laquelle très peu ont accès. Les sauvegardes sont faites régulièrement et après chaque changement important (comme de nouveaux enregistrements). Aussi souvent que possible, nous faisons migrer les anciens fichiers en Tiff ou PDF.

En parlant de coopération, cependant, je ne veux pas seulement parler de travailler ensemble comme institutions en Labour history, mais aussi de travailler avec des « créateurs d'archives » (Producteurs d'archives). Je pense que la plupart des institutions représentées ici aujourd'hui ont le même problème. La plupart d'entre nous ne sont pas structurellement rattachés à des producteurs d'archives. En d'autres termes, nous n'avons pas le cadre légal qui nous donnerait quelque maîtrise sur la manière dont sont créées et conservées les archives avant qu'elles nous parviennent.

Les archives municipales, par exemple, sont normalement une partie de l'administration de la ville, aussi peut-on dire sans risque que les chefs de départements pourraient avoir quelque chose à dire sur la manière dont

l'administration municipale crée et gère ses archives avant qu'elles soient versées aux archives. Nous ne pouvons généralement nous payer ce luxe, et devons donc trouver d'autres manières d'orienter les producteurs d'archives. Heureusement, les archives mal gérées sont autant un problème pour nous que pour les producteurs. Aussi offrons-nous notre aide. Nous essayons de trouver des agréments mutuels qui puissent être matérialisés sous forme de contrats. Jusqu'à ce jour nous avons passé des contrats avec 6 producteurs d'archives. Ces contrats peuvent s'inscrire dans une perspective assez large et ne se focalisent pas forcément sur l'assistance à la création d'archives. De plus, il n'y a pas de procédure figée à propos de cette assistance. Cela pourrait s'avérer très difficile, puisque chaque institution a ses propres problèmes. L'aide que nous offrons va du simple conseil jusqu'à l'analyse détaillée de la structure de l'institution, des flux d'information et du matériel logiciel de manière à donner des conseils conformes aux politiques de l'institution. Ce type de travail d'assistance est bien pour promouvoir l'utilisation de standards et de logiciels open source aussi bien que pour aider à créer de meilleures structures de répertoires et prévenir l'énorme volume des copies redondantes. Le plus mauvais côté de cette approche est qu'il s'agit là d'une première étape et que pour l'avenir nous sommes en train de mettre en place un dispositif d'aide à partir duquel les producteurs d'archives pourront poser leurs questions.

Grands projets en Belgique

En Belgique, il y a quelques intéressants projets en cours dans le domaine de la numérisation et du stockage numérique.

L'un des plus intéressants est le denBell-project. Le conseil municipal d'Antwerp a décidé il y a quelques années de centraliser toute son administration et d'introduire un environnement de travail plus flexible. Deux points clés dans ce projet sont la centralisation de tous les fichiers et documents dans un seul noeud de serveurs et la restructuration du travail - et du flux de documents. Les archives municipales d'Antwerp - qui avaient déjà une bonne réputation avant le passage au numérique- furent activement consultée pendant la phase de préparation et de structuration et devinrent un fonds important pour la création d'un réservoir numérique de grande qualité.

Macz - comme est appelé le réservoir - est actuellement dans sa phase test, mais pourrait être opérationnel avant la fin de l'année. Avant et pendant la phase de préparation, les archives de la municipalité d'Antwerp créèrent une riche documentation à propos de création de documents et de stockage numérique.

Le second projet intéressant est le BOM-V1-project. L'acronyme désigne un projet de conservation et diffusion de données multimédia en Flandres. Ce projet est testé pour la numérisation, le stockage et la diffusion de données audiovisuelles. Il a pour but de donner un exemple de la manière dont une institution pourrait s'occuper de ses données audio-visuelles et porte sur chaque

élément de cette opération, allant de la sélection des méta-données jusqu'à la création d'un prototype de réservoir numérique. Le projet se termine normalement cette année, mais on prévoit de pérenniser le réservoir et d'offrir ainsi un stockage fiable aux institutions participantes.

Le troisième projet est le Projet Presse-Papier. Ce projet est un projet de numérisation à grande échelle, porté par la bibliothèque nationale afin de mieux protéger et diffuser les grandes collections de périodiques de ses collections, mais d'autres collections à travers la Belgique. Le but en est de numériser 30 titres de périodiques du XIX^e à 1950 et une collection de journaux clandestins des deux guerres mondiales. Au total, cela se monterait à environ 3 millions de pages. Une partie du projet est la base de données Abraham. Cette base de données fut construite pour collecter des informations détaillées à travers tous les journaux possibles imprimés en Belgique. Cette information inclut l'état et la qualité de chaque périodique, cela peut ainsi permettre d'adapter les mesures de protection à prendre pour les collections menacées.

Conclusions

Voilà ce que j'avais à vous dire. Je voudrais juste faire un bref résumé de ce que je crois être les points les plus importants.

Je crois que toute la question des « born digital archives » et du stockage numérique est un grand défi, mais nous ne devons pas avoir peur de le relever. Plus nous attendrons, plus grand sera le problème. La clé du succès, en particulier pour les plus petites institutions - est la coopération et la communication. De plus cette coopération devrait être aussi large que possible. Des réseaux comme l'IALHI sont très importants, mais nous devons aussi voir du côté des producteurs d'archives et - si possible - des fabricants de logiciels. Seul un effort en collaboration sera capable de venir à bout de façon adéquate de ce défi.

IALHI international workshop

Summary report for Amsab-ISH

Maarten Savels

(archivist at Amsab-Institute of Social History, Ghent)

What are we digitizing / will we digitize?

We are currently digitizing two important series from our collection: the meeting reports of the bureau and the general council of the BWP (Belgian Labourer's Party) for the period 1892 - 1940 and the meeting reports of the bureau of the ABVV (Belgian Socialist trade union) for the period 1945 - 1984. Furthermore, the important socialist newspaper *Vooruit* is being digitized by the Belgian Royal Library, some audiovisual materials from the SOM (Socialist broadcasting corporation) are being digitized as part of the BOM-VI-project (storage and dissemination of multimedia in Flanders). Last but not least, we are preparing the digitization of a collection of smaller newspapers from the 19th and beginning of the 20th century that have suffered a lot and are in urgent need of rescue.

Which projects/programs do we have in connection with digitization?

We have a small digital repository running, but this is not a definite version. For a trusted digital repository that complies to the internationally accepted standards, we are lacking funds and staff. Our repository was conceived to be a compromise between simplicity, flexibility and low cost. The aim is to enable ourselves to easily cooperate with other initiatives that are slowly taking shape in Belgium (e.g. the aforementioned BOM-VI-project). Unlike the Netherlands or the United Kingdom, digital repositories in Belgium are still in their infancy. We believe a collaborative effort is necessary, but it will take a few more years before we can pick the fruit of this collaboration. Our repository is more or less a transitional effort to help bridge that period.

Concerning access and dissemination, we have built a new website, enabling us to present the content that has already been digitized. As with our repository, this is a system designed for simplicity and flexibility and, again, is a transitional effort.

With whom do we cooperate (national / international)?

We currently do not cooperate on an international level on the subject of digitization. On a national level, however, we are actively and passively involved in a number of projects. The two most important projects are:

The BOM-VI-project. This project is a testbed for digitization, storage and dissemination of audiovisual data. It aims to create an example of how an archival institution should handle its audiovisual material and it focuses on every single aspect of this task, ranging from selection over metadata to the creation of a working prototype of a digital repository. We are mainly involved as content and metadata providers, but we are also providing information about the needs of archival institutions and researchers. The project normally ends this year, but there are plans for making the repository permanent and offering trustworthy storage to participating institutions.

The Presse-Papier-project. This project is a large-scale digitization effort, carried out by the national library in order to better protect and disseminate the huge collection of newspapers that are available from their collection, but also from other collections throughout Belgium. The aim is to digitise thirty newspapers from the nineteenth century up to 1950 and a collection of clandestine papers from both world wars. In total this should amount to about three million pages. Here too, we are involved as content and metadata providers. A side project is the Abraham-database. This database was constructed to collect detailed information about every possible newspaper printed in Belgium. This information includes the availability and quality of every newspaper, so that adequate steps can be taken to protect -i.e. digitize- threatened collections. Here we are part of the policy group that evaluates the outcome of the project.

Which national / international funding possibilities do we have?

We get indirect funding through the projects we are participating in. This indirect funding comes as a certain amount of archival objects that can be delivered and that will be digitized for us, so that we don't have to pay for it. In both of the aforementioned projects, we get about € 20'000 worth of material digitized.

Furthermore, each year we receive a certain amount of gifts from private persons and institutions. These gifts can take two forms:

- The donor asks that a certain collection of documents are digitized and proposes a working budget. This works in much the same way as indirect funding we get through projects. The only difference is that we take care of the digitization ourselves.
- The donor gives a certain amount of money, but asks that it is spent according to his or her wishes.

In most cases the grants come from persons or institutions that are related to the collections we hold. E.g. persons that used to work for the party or the trade union or institutions that have already donated their archives.

What are advantages and disadvantages of these funding schemes?

Besides the obvious advantage of being able to digitize a lot more than our normal budget allows, we are able to strengthen ties with the people and institutions we get funding from. E.g. we hope that, through our cooperation in the BOM-VI-project, we will be given the possibility of participating in the further development and use of their digital repository. That way we can remove the burden of financing staff, hardware and software needed for creating a full scale digital repository and allowing us to focus our efforts more on assisting the records creators or providing metadata and context.

The main disadvantage is that this kind of funding is not structural: they are one-time donations, that cannot be used for hiring staff, long-term maintenance or setting up a research project. Furthermore, we are often bound by the wishes of the donor: we cannot digitize whatever we want, we digitize whatever the donor approves of. Although in most cases we can make propositions about which material we would like to be digitized, the donor has the final word.

What are our main problems and limitations?

A lot has already been written on the technical and conceptual problems related to digitization and a lot of solutions have already been proposed, so we think that this is not the main problem. The main problem is implementing these solutions and for that we need money. As mentioned before, the main problem with funding is the fact that it is not structural, however, for storing and providing access to digitised material over a longer period of time, we need structural funding. As long as this is not available, setting up a long-term project or building a repository will be difficult.

Another problem is staff. We employ archivists, librarians, documentalists and some IT-staff as well, but the problem with digitization and digital storage is the fact that it lies on the border between these fields of expertise. Currently, not a lot of people have expertise in all these fields and that complicates matters. The solution naturally lies with education, but retraining some employees can be very expensive and hiring new staff could mean having to fire old staff, so both solutions lead to difficult decisions.

Which opportunities do we see with IALHI, which possibilities for cooperation?

First of all, IALHI is an existing network, which is a great starting point. Workshops, such as this, are good means for exchanging ideas and the website could be a good medium for spreading best practices or research outcomes.

Second, we share some of the same problems, especially when it comes down to records creators. E.g. we all know the problems that are related to working with papers from ministers such as the confidentiality or the rapid changes after elections. Research into some of these problems could be bundled into a collective effort.

BELGIQUE

IALHI international workshop Summary report for Amsab-ISH

Maarten Savels

(archiviste à l'Amsab-Institut d'histoire sociale, Gand)

Que numérisons-nous ? Que voulons-nous numériser ?

Nous numérisons actuellement deux importantes séries de nos collections : les rapports des réunions du bureau et du conseil général du parti ouvrier belge (BWP) pour la période 1892-1940 et les rapports des réunions du bureau de l'ABVV (syndicat socialiste belge) pour la période 1945-1984. De plus, l'important périodique socialiste *Vooruit* est en cours de numérisation par la Bibliothèque royale de Belgique, quelques matériaux audiovisuels de la SOM (Société socialiste de diffusion) sont numérisés dans le cadre du projet BOM-V1 (stockage et diffusion des multimédia en Flandres). Last but not least, nous préparons la numérisation d'une collection de petits périodiques du XIX^e et du début XX^e qui ont beaucoup souffert et ont un besoin urgent de restauration.

Quels programmes et projets avons-nous en liaison avec la numérisation ?

Nous avons un petit réservoir numérique, mais ce n'est pas la version finale. Pour un réservoir numérique fiable qui réponde aux standards internationaux, nous manquons de fonds et de personnel. Notre réservoir a été conçu comme compromis entre simplicité, flexibilité et bas coûts. Le but est d'être capables de coopérer avec d'autres initiatives qui prennent lentement forme en Belgique (comme le projet ci-dessus mentionné BOM-V1). Contrairement aux Pays-Bas et au Royaume-Uni, les

réservoirs numériques en Belgique sont encore en enfance. Nous croyons qu'un effort de collaboration est nécessaire, mais cela prendra quelques années de plus avant que nous puissions recueillir le fruit de cette collaboration. Notre réservoir est plus ou moins un effort de transition pour passer cette période.

En ce qui concerne l'accès et la diffusion, nous avons construit un nouveau site web, qui nous rend capable de présenter ce que nous avons déjà numérisé. Comme pour le réservoir numérique, c'est un système choisi pour sa simplicité et sa souplesse et, là aussi c'est un effort de transition.

Avec qui coopérons-nous (national/international) ?

Nous n'avons pas actuellement de coopérations internationales à propos de la numérisation. Au niveau national, nous sommes impliqués passivement et activement dans nombre de projets. Les deux projets les plus importants sont :

The BOM-VI-project. Ce projet est un projet pilote pour la numérisation, le stockage et la diffusion de données audiovisuelles. Il a pour but de créer un modèle suivant lequel un centre d'archives puisse traiter son matériel audiovisuel et il porte sur chaque aspect de cette tâche, allant des métadonnées jusqu'à la création d'un prototype de réservoir numérique. Nous sommes impliqués au niveau du contenu et de la production des métadonnées, mais nous fournissons aussi de l'information au sujet des besoins des centres d'archives et des chercheurs. Le projet s'achève normalement cette année, mais il y a des plans pour rendre cet entrepôt permanent et offrir un stockage fiable aux institutions participantes.

Le troisième projet est le Projet Presse-Papier. Ce projet est un projet de numérisation à grande échelle, porté par la bibliothèque nationale afin de mieux protéger et diffuser les grandes collections de périodiques de ses collections, et d'autres collections à travers la Belgique. Le but en est de numériser trente titres de périodiques du XIX^e à 1950 et une collection de journaux clandestins des deux guerres mondiales. Au total, cela se monterait à environ 3 millions de pages. Une partie du projet est la base de données Abraham. Cette base de données fut construite pour collecter des informations détaillées à travers tous les journaux possibles imprimés en Belgique. Cette information inclut l'état et la qualité de chaque périodique, cela peut ainsi permettre d'adapter les mesures de protection à prendre pour les collections menacées.

Quels financements nationaux ou internationaux avons-nous ?

Nous obtenons des financements indirects via les projets auxquels nous participons. Ce financement indirect correspond à un certain nombre d'objets archivistiques qui sont numérisés pour nous, et pour lesquels nous n'avons donc pas à payer. L'ensemble des projets mentionnés représente 20 000 € de numérisation.

De plus, chaque année, nous recevons un certain nombre de dons venant de personnes privées et d'institutions. Ces dons peuvent prendre deux formes.

Le donneur demande qu'une collection particulière soit numérisée et propose un budget de travail. Ce travail revient au même que ce que nous obtenons par les projets ci-dessus mentionnés. La seule différence est que nous devons prendre nous-mêmes en charge les numérisations.

Le donneur donne un certain montant, mais demande à ce que ce soit utilisé selon ses vœux.

Dans la plupart des cas, les donations viennent de personnes ou d'institutions qui ont un rapport avec nos collections, telles des personnes qui sont habituées à travailler pour le Parti ou les syndicats ou les institutions qui ont déjà donné leurs archives.

Quels avantages ou inconvénients y a-t-il à ces financements ?

Au-delà de l'avantage général qu'il y a à être capable de numériser beaucoup plus que ne nous le permettrait notre budget habituel, nous pouvons fortifier les liens avec les personnes ou les institutions desquels nous obtenons des financements. Comme nous espérons qu'à travers notre coopération au projet BOM-V1, nous aurons la possibilité de participer au développement et à l'utilisation du réservoir numérique. Ainsi nous pouvons déplacer le poids du financement, des logiciels nécessaires pour créer un réservoir numérique et focaliser nos efforts en aidant davantage les producteurs d'archives ou en fournissant des métadonnées et du contenu.

Le principal inconvénient est que cette sorte de financement n'est pas structurel : les dons ont lieu une fois et ne peuvent être utilisés pour embaucher du personnel, pour la maintenance ou pour promouvoir un projet de recherche. De plus, nous sommes souvent liés par les désirs du donneur : nous ne pouvons numériser ce que nous voulons, nous numérisons ce que veut le donateur. Bien que dans la plupart des cas, nous puissions faire des propositions, le donateur a le dernier mot.

Quels sont nos principaux problèmes et nos limites ?

Beaucoup a déjà été écrit sur les problèmes conceptuels et techniques relatifs aux numérisations et beaucoup de solutions ont déjà été proposées, aussi voyons-nous que ce n'est pas le problème principal. Le problème principal est de mettre en œuvre ces solutions et pour cela nous avons besoin d'argent. Comme nous l'avons mentionné, le principal problème avec le financement est le fait que ce ne soit pas structurel, et pourtant pour entreposer et fournir un accès au matériel numérique sur une plus longue période, nous avons besoin de financements structurels. Aussi longtemps que cela ne sera pas, monter un projet à long terme ou construire un réservoir sera difficile.

Un autre problème est celui du personnel. Nous employons des archivistes, des bibliothécaires, des documentalistes et quelques IT, mais le problème avec la numérisation et le stockage numérique est le fait que cela se situe à la limite de tous ces champs d'expertise.

Actuellement, peu de gens sont experts dans ce domaine et cela complique les choses.

La solution réside bien entendu dans l'apprentissage, mais former quelques employés peut coûter très cher et embaucher un nouveau personnel signifierait renvoyer l'ancien, ainsi les deux solutions engagent à de difficiles décisions.

Quelles opportunités voyez-vous avec l'IALHI, quelles possibilités de coopération ?

D'abord, l'IALHI existe, ce qui est un grand point de départ. Les ateliers comme celui-là sont de bons moyens d'échanger des idées et le site web pourrait être un bon moyen de diffuser de bons usages et des résultats de recherche.

Ensuite, nous partageons les mêmes problèmes, particulièrement en ce qui concerne la disparition de producteurs d'archives. Par exemple nous savons tous les problèmes que pose le fait de travailler avec des papiers émanant de ministères, tel que la confidentialité ou les changements rapides après les élections. Les recherches sur ces problèmes pourraient être mises en commun.

Digitization issues

Richard Temple
(Senate House Library)

Senate House Library has already digitized University student records
<http://www.shl.lon.ac.uk/specialcollections/archives/studentrecords.shtml>

We hoped to create pdf copies with a background text but these files were too large. So we created OCR versions of the originals and then turned them into pdf copies (without the background text). These online pdf copies are searchable although the format varies slightly from the originals

1. One problem is storage for digital images. When we were envisaging a more ambitious student records projects, I enquired about the costs of digital storage. This would have cost us at least £ 5 000 per year

2. We have not co-operated with any other institutions as yet although there are discussions about a project

3. There is Joint Information Systems Committee (JISC) money available in the UK at the moment but this depends on joint projects between UK and US institutions.

4. Labour history digitisation projects in the UK :

- TUC Library : (Christine Coates knows much more than me about these !)
- Modern Records Centre has digitized teaching modules and images from its exhibitions

5. Problems :

- digitization is often seen as replacement for cataloguing rather than as a complement to it. Resources for cataloguing are often very limited but it is important to add digitized copies to archive descriptions. Digital copies also need final, specific reference number, which implies full cataloguing
- metadata. If images are added to an ISAD(G) catalogue, is this enough ? Or should we think of Dublin Core metadata which would allow cross-searching of images by library users ? Other metadata standards are OAI-PHM and the METS standard for digital libraries
- Data Protection Act 1998 (and other privacy legislation). This circumscribes way in which the personal data, may be mounted online. Our student lists are at least 80 years old. Putting old lists of (e.g.) trade union members online would be an excellent way of drawing family historians into labour history.

- Copyright. UK legislation is extremely restrictive especially as it relates to images (which may not be copied unless they are out of copyright or have the copyright holder's permission to do so). One way of dealing with this is to put a disclaimer up with images to inform copyright holders should they contact the institution. It is important not to charge for access to digital images if there is any doubt about copyright

6. Co-operation :

- Shared digital space ?
- International themes (e.g.) Trotskyism
- EU money?

ROYAUME-UNI

Expériences de numérisation

Richard Temple
(Senate House Library)

La Senate House Library a numérisé des dossiers universitaires

<http://www.shl.lon.ac.uk/specialcollections/archives/studentrecords.shtml>

Nous avons essayé de créer des copies en pdf à partir d'un texte de base mais ces dossiers étaient trop grands. Nous avons donc créé des versions de ROC des originaux et les avons ensuite transformés en copies pdf (sans le texte de base). Ces copies pdf en ligne sont accessibles bien que le format soit légèrement différent de celui des originaux

1. Un problème est celui du stockage des images numériques. Quand nous envisagions un projet plus ambitieux pour l'enregistrement de ces dossiers universitaires, je me suis renseigné au sujet des prix du stockage numérique. Cela nous aurait coûté au moins £ 5 000 par an.

2. Nous n'avons coopéré avec aucune autre institution jusqu'à présent bien qu'il y ait des discussions autour d'un projet possible.

3. Il y a le Comité de Systèmes informatiques Collectif (JISC). Il y a de l'argent disponible actuellement au Royaume-Uni mais tout dépend des projets collectifs entre les institutions britanniques et américaines.

4. Numérisation de l'histoire du parti travailliste en projet :

- Bibliothèque des TUC : (Christine Coates en sait beaucoup plus que moi à ce sujet) ;
- le Centre de Dossiers Moderne a numérisé des modules d'enseignement et des images de ses expositions.

5. Problèmes :

- la numérisation est souvent considérée comme une solution de remplacement pour un catalogue plutôt que comme un complément. Les ressources pour le catalogage sont souvent très limitées mais il est important d'ajouter des copies numérisées pour archiver des descriptions. Les copies numériques ont aussi besoin du numéro de référence final, spécifique, qui implique le catalogage complet ;
- métadonnées. Si les images sont ajoutées à un catalogue ISAD (G), est-ce assez ? Ou devrions-nous penser aux métadonnées du Dublin Core qui permettraient la recherche des images par les utilisateurs ? ;
- copyright. La législation britannique est extrêmement limitative d'autant plus qu'elle s'étend aux images (qui ne peuvent être copiées à moins qu'elles ne soient tombées dans le domaine public ou aient l'autorisation du détenteur de copyright). Une façon de s'en occuper est de relever un problème avec les images pour informer les détenteurs de copyright qu'ils devraient contacter l'institution.

6. Coopération :

- espace numérique partagé ?
- les thèmes internationaux (eg) Trotskyism
- argent d'UE ?

Labour Movement Archives and Library

Lars Ilshammar

(ARAB)

I would like to talk here today very briefly about our experience of digitization and receipt of electronic files at the Labor Movement Archives and Library in Stockholm.

We are a large private archive institution, perhaps the largest in Sweden, with 24 full time employees and a budget of approximately SEK 20 million a year. We are also an old institution, founded already in 1902, rich in experience and knowledge. Yet we are not in principle accepting against electronic archive material at all - neither digitized nor born digital material. The reason is as simple as regrettable - today and for time being we cannot guarantee the long-term preservation, and therefore we don't want to be responsible for digital material.

At present, we require that the organizations that provide their archival material to us, including major popular movements and large trade unions, first print the content on archival paper - that also applies to web sites - which is a bit contradictory, and ultimately of course an unsustainable strategy. We therefore regard it as a kind of sustained warfare, a defensive action in anticipation of better and cheaper methods for long-term digital preservation.

So far, the organizations have been surprisingly understanding. That there none the less exists some floppy disks, hard drives and old tape reels in our archives shelves may be well blamed on the archivists, who generally are kind and rather would turn a blind eye to the principles than consider useful archival material go to waste. With that said, we still cannot guarantee the preservation of digital content, and some really old disks from the 'Annu Dazumal' are perhaps best used as curling stones today.

In the case of the digitization of our existing records, the instinct has been to accelerate slowly, learn from others - especially the large national memory institutions as the National Archives - and not let us be pushed into some mass digitization that could turn out to be technically questionable and economically costly. What actually has been digitized to date is almost exclusively archival material of a non-traditional kind and then in the form of specific projects financed with external resources:

In cooperation with the Swedish Film Institute and the National archives of audio and video - now a part of the Swedish national library - we have digitized most of our film stock, about 400 films, from the 1920s onwards. Original films will in future be kept in climate-secured warehouses of the Film Institute's documentary film archives in Grängesberg in Dalcarlia in central Sweden, while we can make the digitized versions available for visitors in our premises in Stockholm. However, because of copyright reasons the films cannot yet be published online.

Together with the National archives of audio and video, and with temporary funding from the Swedish Arts Council, we are currently digitizing our extensive portfolio of audiotapes. The project is three years, has employed three persons full time, a project manager and two sound technicians. Here, I may in particular mention the wire tapes in former Prime Minister Tage Erlanders archives, a unique political material which in practice would be inaccessible without digitization. When the project is concluded, the digitized material will be available to visitors both in our premises and at the National archives of audio and video. Copyright issues prevent us, unfortunately, in the current situation from publishing audio files on the Internet.

The sound project has so far shown that it's not the actual scanning, which requires large resources, but the pre- and post-work. Therefore it is particularly important to consider the purpose of digitization. Is it primarily about conservation, or about making the material accessible, and if so, physically in our premises or virtually online? If the material is to be published on the net that requires, in addition to the discussion of copyright, a soundly based policy for how the individual documents or files is to be presented to a larger audience.

An interesting and complicating factor about the Internet is that it's not only

turning to academic researchers, but to a wider public who may require extensive explanatory texts, which impose specific requirements on the competence of the staff to describe the documents or files. Moreover, it should be made clear whether a material is presented in its entirety, or only in the form of a selection. In the case of limited selection descriptions of all materials not available in digital form is also required, to prevent this matter from disappearing into the analog darkness.

With funding from the Olof Palme Memorial Fund, we have built a dedicated web site with material from former Prime Minister Olof Palme's archives: <<http://www.olofpalme.org>>. In the summer of 2007, we scanned about 25 000 pages, mostly drafts and manuscripts for speeches, in cooperation with the company Info Disk Media in Riga, Latvia. In the autumn of 2008, we scanned a further 20 000 pages. Parallel to the scanning, the construction of the web site with photos and audio recordings, topic pages, help texts and a timeline, took place.

The project has already provided useful experience of web publishing. One such experience is that the material not published on the web tends to become "darkened" and therefore relatively more inaccessible than before. Scientists, like other people, are indeed often following the law of least resistance. The web also appears to raise the number of internationalized questions, which is positive, but also put other demands on our staff. A clear trend is that the number of international requests for information increases as a consequence of web publication.

Through IALHI we are also involved in the project Labor History Index <<http://search.labourhistory.net>> aimed at developing a common web-based search interface for the member institutions' collections. Since you are all familiar with Labor History Index, I will not waste more time on that project here and now.

Finally, banners, posters and parts of our large photo collection have been scanned through cooperation with the Foundation for Artifact Preservation or photographed with digital camera. These materials can therefore be depicted digitally and in some cases also be downloaded. The Foundation for Artifact Preservation, which is now threatened by drastic downsizing, has digitized more than 60 000 photos, including skiopticon pictures, glass negatives and slides, also

posters of the smaller format, and as a test still film, a hitherto neglected historical source - as it has been difficult for researchers to access.

These were a few digital highlights from a broadly still analog archives reality. The big question therefore remains: Should we push into a digitization of the major archives stocks, and who is in that case willing to bear the economic burden? In the current situation of decreased funding, that looks a somewhat bleak calculation. If our clients don't want to inject resources to guarantee that current operations are keeping pace with price and wage increases, how can we then hope to get more funds for digitization? To make things the even more complicated: In the general awareness, digitization is the tool to save resources. It's not supposed to create a cost increase. We can call that the digital paradox.

One thing is absolutely certain - requirements and expectations from the outside world that the archives should be digital and available on the web one keystroke away will continue to grow. We are therefore forced to somehow learn to live with the digital paradox for the foreseeable future.

Suède

Bibliothèque et Archives du mouvement ouvrier

Lars Ilshammar

(ARAB)

Je voudrais parler ici très brièvement de notre expérience de numérisation et de collecte de fichiers électroniques à la Bibliothèque et Centre d'archives du Mouvement ouvrier de Stockholm.

Nous sommes une importante institution d'archives privées à Stockholm, peut-être la plus importante de Suède avec 24 salariés à plein temps et un budget d'à peu près 20 Millions de SEK par an²⁶. Nous sommes aussi une institution ancienne, fondée en 1902, riche d'expérience et de savoir. Cependant nous n'acceptons en principe aucun matériel électronique - ni numérisé ni « Born digital ». La raison en est aussi simple que regrettable -aujourd'hui et pour l'instant, nous ne pouvons garantir la conservation à long terme des documents électroniques.

Pour le moment, nous demandons que les organisations qui nous confient leur matériel électronique, dont les principaux mouvements populaires et les grands syndicats, impriment d'abord le contenu sur papier - cela s'applique aussi aux sites web - ce qui est un peu contradictoire ;. Nous la considérons donc comme une sorte de stratégie provisoire, une action défensive en attendant des méthodes meilleures et moins coûteuses pour la conservation électronique à long terme.

Jusque là, les organisations ont été étonnamment compréhensives. Certes, grâce à nos archivistes, qui, toujours pleins de bonnes intentions, préfèrent transiger avec les principes de conservation plutôt que de voir de potentiels ressources archivistiques finir à la poubelle, nous conservons tout de même, sur les étagères de nos magasins, quelques disquettes, quelques disques durs et des bandes sonores. Mais nous ne pouvons pas encore garantir la préservation des contenus numériques, et quelques disques vraiment anciens de 'Annu Dazumal' pourraient peut-être plus utilement servir aujourd'hui de palets pour le curling.

Pour ce qui est de la numérisation de nos propres enregistrements, notre instinct nous a conduit à accélérer lentement, apprendre des autres - en particulier des

²⁶ Soit environ 1.750.000€

grandes institutions nationales de mémoire telles que les Archives nationales. - et de ne pas nous laisser entraîner dans des numérisations massives qui pourraient se révéler techniquement douteuses et économiquement coûteuses. Ce qui a été en fait numérisé à ce jour est presque exclusivement du matériel archivistique non traditionnel et d'autre part sous forme de projets spécifiques financés avec des ressources extérieurs.

En coopération avec l'Institut du Film suédois et les archives nationales audio-visuelles - qui font maintenant partie de la Bibliothèque nationale de Suède - nous avons numérisé la plus grande partie de notre fonds de films, environ 400 à partir des années 20. Les films originaux seront à l'avenir conservés dans des entrepôts à température adaptée des Archives du film documentaire de l'Institut du film à Grangesberg dans la Dalcarlia de la Suède centrale, tandis que nous pouvons rendre les versions numérisées accessibles aux visiteurs dans nos locaux à Stockholm. Cependant, pour des raisons de copyright, les films ne peuvent pas encore être diffusés en ligne.

Avec les archives nationales audio-visuelles et avec un financement provisoire du conseil suédois des arts, nous sommes en train de numériser notre considérable collection d'enregistrements. Le projet est prévu pour durer trois ans, et mobilise trois personnes à plein temps, un chef de projet et deux techniciens du son. Ici, je voudrais faire une mention particulière des enregistrements téléphoniques conservés parmi les archives de l'ancien Premier Ministre Tage Erlanders, un matériel politique exceptionnel qui serait pratiquement inaccessible sans la numérisation. Quand le projet sera achevé, le matériel numérique sera accessible à la fois dans nos locaux et aux archives nationales audio-visuelles. Les problèmes de copyright nous empêchent malheureusement de publier pour l'instant les fichiers audio sur Internet.

Ce projet autour d'enregistrements sonores a jusque là montré que ce n'est pas le processus de numérisation lui-même qui nécessite de grandes ressources mais le travail qui le précède et le suit. Pour cette raison, il est particulièrement important de prendre en compte les buts de la numérisation. Est-elle d'abord à usage de conservation, ou a-t-elle pour objectif de rendre le matériel accessible, et, si c'est le cas, selon quelles dispositions : sur place dans nos établissements ou en ligne sur Internet ? Si le matériel doit être publié sur le net, cela nécessite, en plus des questions de copyright à résoudre, une politique bien définie en vue de savoir comment faire pour présenter documents ou fichiers à un large public.

En effet, un facteur intéressant mais qui complique les choses avec Internet est que cet outil n'est pas uniquement destiné à un public de chercheurs, mais touche aussi un public plus large qui peut demander avoir besoin d'explications approfondies ; en amont, il faut donc que le personnel chargé de la description des documents ou des fichiers numériques dispose de compétences particulières. Il faut aussi indiquer clairement si les documents numérisés mis en ligne le sont de manière intégrale, ou s'il s'agit seulement d'une sélection. Dans ce second cas, il faut penser à une description des matériaux non sélectionnés, afin qu'ils ne disparaissent pas dans le néant analogique.

Avec un financement du Fonds Memorial Olof Palme, nous avons construit un site web dédié avec des matériaux venant des archives d'Olof Palme : <http://www.olofpalme.org>. Durant l'été 2007, nous avons scanné environ 25 000 pages, pour la plupart des brouillons et manuscrits de discours, en coopération avec la compagnie Info Disk Media de Riga, en Lettonie. Puis à l'automne 2008, nous avons scanné 20 000 pages de plus. Nous avons aussi construit un site web avec des photos et des enregistrements audio, des pages thématiques, des textes d'aide et une chronologie.

Au sein de l'IALHI, nous sommes aussi impliqués dans le projet d'Index d'histoire ouvrière <http://search.labourhistory.net> dont le but est de développer une interface de recherche commune pour les collections des membres de l'institution. Puisque l'Index vous est à tous familier, je ne perdrai pas plus de temps sur ce projet.

Enfin, des bannières, des affiches et une partie de notre vaste collection de photographies ont été numérisés en coopération avec la Fondation pour la préservation d' *artifacts* ou des documents photographiés avec des appareils numériques. Nous disposons donc de reproductions numériques de ces matériaux, qui dans certains cas peuvent être téléchargés. La Fondation pour la préservation des Artefacts, qui est maintenant menacé de réduction drastique de ses effectifs, a numérisé plus de 60 000 photos, dont des synoptiques skiopticon des illustrations pour lanternes magiques, des négatifs sur plaque de verre, des diapositives et des affiches de petit format, et même à titre d'expérience, quelques photogrammes - une source jusqu'à présent négligée par les historiens - parce que difficile d'accès pour les chercheurs.

Voilà donc quelques remarques sur une réalité archivistique encore largement analogique. La grande question reste donc : Devrions-nous pousser à la numérisation des principaux fonds d'archives, et qui dans ce cas pourra en supporter le coût ? Dans une situation générale de baisse des financements, cela semble un calcul un peu inapproprié un pari assez hasardeux ? Si nos clients ne veulent pas injecter

des ressources pour garantir que les opérations en cours tiennent le rythme avec l'augmentation des prix et des salaires, comment pouvons-nous dans ces conditions obtenir plus de financements pour les numérisations ? Pour rendre les choses encore plus compliquées, de manière générale, la numérisation est vue comme un outil de sauvegarde. Elle n'est pas supposée générer de nouvelles dépenses. Nous pouvons appeler cela le paradoxe numérique.

Une chose est absolument certaine - les besoins et les attentes venues du monde extérieur en matière de numérisation et d'accessibilité des archives sur le web, à portée de clic, vont continuer à croître. Pour les années à venir, Nous sommes donc obligés d'apprendre à vivre avec le paradoxe numérique .

Some American examples

Andrew Lee
(New York University Library)

I want to make some features of the cultural and political make up of the United States clear for my audience because these all have a tremendous impact on digitization in our country. In almost everything in the United States we have commercial and non-commercial providers and creators of digital content, as do most of you. One of the key things to remember is that we do not have a national library. The Library of Congress serves as the *de facto* national library, but ultimately it is responsible to Congress, and not to the nation. Digitization efforts in the country are divided up between commercial and non-commercial providers, with the private providers usually charging large fees, either for subscription or purchase, with restricted access - Google is the major exception to this generalization. The digital projects created by non-commercial entities are usually either free projects, demonstration sites, or have restricted access such as the Google Books project (which is also driven by advertising). The United States has a form of government in which power and control over education is devolved to the most local authorities. We have a system of mixed education between schools that are funded by private and public entities, and increasing a mixture of the two. Curriculum is decided locally, but standards are established at the state level. Not only do these standards differ between the states, what the facts are also are not consistent. John Brown is a hero in Iowa, a criminal in the south. Funding is also on the same process, so that private schools often have a large endowments and can be very rich while state supported institutions can be struggling and in desperate need of cash. Tuition fees vary radically between universities and colleges, and so resources for items beyond the basics such as digitization are heavily reliant on outside funding. What is digitized by institutions is also decided based on the holdings of other New York institutions of higher learning. We also compete for federal money as well, and for private foundation grants, by copyright and by a local institutions determination of what constitutes « fair use. » Fair Use is a rather slippery concept that is applied unevenly across the country in determining whether or not the use is a violation of the holder's copyright. Commercial vendors are rapidly digitizing out of copyright materials, frequently from microfilm, but they are also asserting proprietary rights over the digital form of the material, especially through the search engines, without which the material is inaccessible.

The largest public digitization project is the Library of Congress' « The Making of America » and its «Digitizing American Imprints» program which now has

25 000 titles. Numerous libraries and academic institutions have their own digital programs, and the U.S. members of the Google Book project and some associated members have created the Hathi Trust. There are some local collections and a few state wide endeavors, such as the California Digital Library, which is also a member of the Google Book program through the University of California Berkeley and subsequently also a member of the Hathi Trust. However, the Hathi trust is still in the early stages and is trying to find a way to make their work discoverable. There are some endeavors to make this material accessible, such as the Digital Library Federation's Aquifer (<http://www.dlfaquifer.org/>) and the union catalogue OAIster (<http://www.oaister.org/>).

I now switch my talk to New York University, which despite the name is a private institution, heavily dependent on tuition income. As a major private university we receive some aid from the state government, usually in the form of grants for specific programs which we apply for and compete against other New York institutions, aid from the federal government, again in competition with other institutions. Some digitization projects are in conjunction with other institutions. We are not a participant in the Google Book Project, a partial participant in the OAIster, and full participant in Aquifer. The hardest things for our and almost any digitization project are discovery: how do we make our digital content visible for users, and whether appropriate, available and accessible to the wider world. Even internally, not many know exactly what we have done digitally, and finding these from the library's web side is next to impossible. There has been a great deal of atomization of staff and projects, and while there are efforts to improve communications, the fact remains that we still operate in a system that is based on the older and now outdated as the sole paradigm of printed materials.

We have several digitizing projects which should be divided into preservation and access projects. Digitizing as a preservation process tries to retain as much as possible the information and characteristics of the original item, such as color, typographical errors, and includes extensive metadata. The goal here is persistence: to forever maintain in an electronic environment. An Access project on the other reverses the focus is less concerned about preserving the non-print characteristics of the object in favor of short term gains and turnaround, ease of access, file size, and simple searching with less attention paid to metadata. While these are gross simplifications, overall a preservation project requires substantially more time and expertise, as well as advanced systems for checking and preserving data than an access project.

I have listed some projects which I will not talk about but I am happy to try and answer any questions you might have.

- 1) The first is a federal project, the Digitizing American Imprints: The Library of Congress will digitally scan *The Heroic Life of Abraham Lincoln: The Great Emancipator* as the 25,000th book in its *Digitizing American Imprints*

program, which scans aging "brittle" books often too fragile to serve to researchers. The program is sponsored by a \$2 million grant from the Alfred P. Sloan Foundation. The Library, which has contracted with the Internet Archive for digitization services, is combining its efforts with other libraries as part of the open content movement. The movement, which includes over 100 libraries, universities and cultural institutions, aims to digitize and make freely available public-domain books in a wide variety of subject areas. Shortly after scanning is complete, the books are available online at www.archive.org. Books can be read online or downloaded for more intensive study. The Library of Congress is actively working with the Internet Archive on the development of a full-featured, open-source page turner. A beta version, called the Flip Book, is currently available on the Internet Archive site

- 2) Archiving the Political Web is a joint project with the California Digital Library, and others. NYU's Digital Library Team is seeking to develop Web archiving tools that will be used by libraries to capture, curate and preserve collections of Web-based government and political information. The collections will focus on local political activities and movements, such as the California gubernatorial recall election of 2003.
- 3) Afghanistan Digital Library will retrieve and restore works published in Afghanistan between 1870 and 1930; the long-term objective is to collect, catalogue, digitize and provide access to as many of this period's publications as possible. This is a project in cooperation with the U.S. State Department. It is a preservation project in contrast to the University of Arizona Library's Afghanistan Digital Libraries.
- 4) Archivists' Toolkit is another collaborative project to develop an open source application that will provide archives with a turn-key system to support major archival functions and activities. The objective is a system that will support all aspects and functions of archival administration and description (including components for digital and analog media asset tracking and management), and which will be easily deployable in a range of archival repositories from small historical societies to large multi-repository consortia such as the Online Archive of California.
- 5) The Database of Recorded American Music: In partnership with New World Records, a company specializing in documenting American music largely ignored by the commercial recording companies, NYU's Digital Library Team is building a database of American music. This system will provide subscribing institutions with a searchable database of American music recorded by New World Records, as well as access to complete digital versions of the recordings in New World Record's catalog.

- 6) Digital Archive of Public Television: In collaboration with the Educational Broadcasting Corporation (EBC), WGBH Educational Foundation, and the Public Broadcasting Service, New York University (NYU) is participating in a project sponsored under the National Digital Information Infrastructure Preservation Program to establish the first procedures, structures and national standards necessary to preserve public television programs produced in digital formats. EBC and WGBH are the two largest producers of public television content in the United States. Through PBS, their productions are made available to audiences from coast-to-coast. Together, these three entities produce and distribute the majority of public television in the United States. The four partners will focus on such influential series as "Nature," "American Masters," "NOVA" and "Frontline," which are increasingly being produced only in digital formats, including the new high-definition standard (HDTV). The project will also examine issues associated with the preservation of important corollary content, such as Web sites that accompany broadcasts.
- 7) The Hemispheric Institute Performance Video Archive: 515 hours of video, The Hemispheric Institute of Performance and Politics is a consortium of institutions, artists, and scholars dedicated to exploring the relationship between expressive behavior (broadly construed as performance) and social and political life in the Americas. The Digital Library Team, working in conjunction with the Hemispheric Institute and NYU's Humanities Computing Group, is developing an online archive of over 200 performance videos to be made publicly available over the Internet.
- 8) Second Avenue Online: Second Avenue Online is a multimedia Yiddish Theatre digital archive established to preserve the history of the Yiddish stage of yesterday, and to bring together a new audience for its present and future. The project is named after Maurice Schwartz's Yiddish Art Theater on Second Avenue and 12th Street, the longest running repertory theater in New York City's history. The archive contains histories of the Yiddish theater, oral histories, scripts in Yiddish and English, photographs, video clips and music.

Some others are the Abraham Lincoln Brigade photograph collections, the Levy Collection of Dime Novels, and Washington Square.

Quelques exemples américains

Andrew Lee

(New York University Library)

Je voudrais commencer par quelques remarques générales sur l'histoire politique et culturelle des Etats-Unis, car le passé a un impact considérable sur nos politiques de numérisation. Comme partout dans le monde, existent aux Etats-Unis différentes entreprises susceptibles de mener des chantiers de numérisation, et celles-ci peuvent être ou non à but commercial. Il faut se souvenir que nous n'avons pas de bibliothèque nationale. La bibliothèque du Congrès en tient lieu, *de facto*, mais en dernier recours elle ne dépend que du Congrès, et pas de la nation. Les expériences de numérisation menés aux Etats-Unis se partagent entre initiatives à visées lucratives et projets non-commerciaux ; dans le premier cas, les fournisseurs de données numériques imposent en général des tarifs d'accès élevés, qu'il s'agisse d'abonnements ou d'achats, et des conditions de consultation restrictives — les services de Google étant la principale exception à cette règle. Les projets de numérisation portés par des groupements non-commerciaux proposent eux, pour la plupart, une consultation gratuite, ou mettent en place un système d'accès réservé — ce qui est le cas du programme Google Search (qui bénéficie aussi de ressources publicitaires). Le gouvernement est organisé de telle manière aux Etats-Unis que les politiques éducatives sont du ressort des autorités les plus locales. Notre système scolaire est divisé entre écoles publiques et écoles privées, de plus en plus d'établissements étant cependant « mixtes ». Les programmes sont définis localement, mais en fonction de recommandations établies à l'échelle des Etats. Non seulement celles-ci varient d'un Etat à l'autre, mais, de plus, les faits eux-mêmes finissent pas de ne plus avoir de réalité. Ainsi John Brown est-il considéré comme un héros dans l'Iowa, et comme un criminel dans les Etats du Sud. Les financements sont eux aussi très variables, si bien que les écoles privées sont en général bien pourvues et parfois même très riches, tandis que les établissements financés par les Etats peuvent manquer de crédits et courir désespérément après l'argent. De la même manière, les frais de scolarité changent radicalement d'un lycée ou d'une université à l'autre ; les crédits disponibles pour des projets autres que les missions de base - les projets de numérisation, par exemple - reposent donc le plus souvent sur des subventions extérieures.

Dans le cas de mon université, les choix de numérisation dépendent aussi des collections disponibles dans les autres établissements d'enseignement supérieur de

la ville. Nous cherchons à obtenir des subsides fédéraux, ainsi que d'éventuelles bourses privées, et respectons les règles du copyright, ou tout au moins celles découlant d'une utilisation « raisonnable » des documents. Ce concept de "fair use », dont la définition est assez vague, s'applique partout dans le pays pour établir si l'utilisation d'un document viole ou non les droits dont bénéficie son détenteur. Dans le cas de projets à visées commerciales, les documents anciens échappant à la règle du droit d'auteur sont parfois numérisés très vite, à partir souvent de copies microfilmées, mais leur version électronique finale est soumise à de nouveaux droits de propriété, en particulier via la création de moteurs de recherche sans lesquels ces nouvelles ressources ne sont pas consultables.

Le plus grand projet de numérisation est celui de la Bibliothèque du Congrès, Digitizing American Imprint, mené dans le cadre du programme The Making of America, qui compte désormais 25 000 ouvrages numérisés. De nombreuses bibliothèques et différents établissements universitaires ont aussi engagé des programmes de numérisation, tandis que les participants américains au projet Google Book ont créé, en partenariat avec d'autres membres associés, le Hathi Trust, projet de bibliothèque numérique partagée²⁷. Divers établissements aux collections d'importance locale participent à ce projet, ainsi que quelques institutions oeuvrant à l'échelle d'États, comme par exemple la California Digital Library (Bibliothèque numérique californienne), affiliée au programme Google Book via l'Université de Californie à Berkeley, et donc automatiquement membre du Hathi Trust. Les activités de ce consortium sont encore embryonnaires, et ses membres essaient actuellement de donner plus de publicité à leurs travaux. Des efforts sont faits pour rendre les documents numérisés accessibles, comme par exemple le portail Aquifer de la Digital Library Federation, consacré à l'histoire sociale des États-Unis²⁸ ou le projet de moteur OAIster, dont l'ambition est de faciliter l'accès aux ressources électroniques à orientation universitaire²⁹.

Je vais maintenant vous parler plus précisément de ce qui se passe à la New York University, qui en dépit de son nom est un établissement privé, dont les ressources financières dépendent pour beaucoup des frais de scolarité. En tant qu'établissement important, nous recevons certes de l'argent en provenance de l'État de New York, mais en général sous forme de bourses accordées pour la réalisation de programmes spécifiques, obtenues sur candidature et après que nous ayons été mis en compétition avec d'autres institutions. Quelques unes de nos activités autour de la numérisation sont menées en partenariat avec d'autres établissements. Nous ne participons pas au projet Google Book, ne contribuons que partiellement au projet OAIster mais sommes en revanche investis totalement

²⁷ www.hathitrust.org

²⁸ www.dlfaquifer.org

²⁹ www.oaister.org

dans le projet Aquifer. Comme c'est le cas pour la plupart des programmes de numérisation, les questions pour nous les plus difficiles à résoudre concernent la mise en valeur de nos collections numériques : comment les rendre visibles pour leurs utilisateurs potentiels, et faciliter leur consultation de partout dans le monde ? A l'intérieur même de notre université, peu de personnes connaissent vraiment l'étendue de nos travaux de numérisation, et il est à peu près impossible d'accéder à nos ressources numériques à partir de notre site Internet. Nos numérisations ont souvent été menées par des équipes disparates, et sur des petits corpus. En dépit de nos efforts pour améliorer la communication entre nous, c'est un fait que nous continuons de travailler selon des règles anciennes — et désormais obsolètes —, établies autrefois pour la prise en charge des documents imprimés.

Nous avons actuellement plusieurs programmes liés à notre politique de numérisation : certains sont des projets de conservation, et d'autres visent à faciliter l'accès à nos collections numérisées. Dans le premier cas, il est important que le fichier numérique finalement obtenu rende compte au maximum des informations disponibles sur le document original, ainsi que de ses différentes caractéristiques (comme par exemple sa couleur ou ses éventuelles erreurs typographiques), et disposent aussi de méta-données détaillées. Sa conservation à long terme doit être prévue : le fichier numérique devra exister pour toujours dans un environnement électronique. Les questions relatives à l'accès aux documents numérisés renversent, elles, la perspective : la conservation des caractéristiques du document original ne constitue plus la préoccupation principale ; qu'il s'agisse des moyens d'accès aux documents, de la taille des fichiers ou des modes de consultation des documents, on recherche plutôt des solutions faciles à mettre en place et efficaces à court terme. Ce sont là, évidemment, des remarques très générales, mais on voit bien que les programmes de conservation à long terme sont ceux qui demandent les plus gros investissements en temps et en expertise.

Afin de ne pas abuser de mon temps de parole, je me suis contenté de faire la liste de quelques programmes de numérisation. Je ne les présenterai pas en détail, mais je tacherai plus tard de répondre, si vous le souhaitez, aux questions que vous pourriez me poser.

Premier projet : un projet fédéral intitulé **Digitising American Imprints**, consistant donc en la création d'une bibliothèque d'ouvrages numérisés. La bibliothèque du Congrès compte numériser très bientôt l'ouvrage *The heroic life of Abraham Lincoln : the Emancipator*, qui sera ainsi le 25 000^e livre de sa bibliothèque numérique, constituée principalement de copies d'ouvrages anciens et fragiles, désormais en trop mauvais état pour être manipulés par les lecteurs. Ce programme a été financé par une bourse de 2 millions de dollars offerte par la Fondation Alfred P. Sloan. La bibliothèque du Congrès a établi pour la numérisation un partenariat avec Internet Archive, et collabore aussi avec d'autres bibliothèques dans le cadre du mouvement pour un accès libre au savoir (« open

content movement »). Celui-ci regroupe plus de 100 bibliothèques, universités et autres établissements culturels ; il a pour ambition de numériser et de rendre disponibles gratuitement sur Internet des ouvrages tombés dans le domaine public et concernant un choix très large de sujets. Dès que numérisés, les livres sont mis en ligne, et consultables sur le site www.archive.org, à partir duquel il est possible aussi de les télécharger. En partenariat avec Internet Archive, la Bibliothèque du Congrès travaille aussi activement au développement d'un logiciel open source pour le feuilletage des livres en ligne, disposant de toutes les fonctionnalités nécessaires. Une première version de ce logiciel, appelée « The Flip Book », fonctionne déjà sur le site d'Internet Archive.

Il existe aussi un programme d'archivage du Web politique, dirigé conjointement par la California Digital Library et d'autres établissements. Dans ce domaine, l'équipe en charge de la bibliothèque numérique de la New York University travaille actuellement au développement d'outils d'archivage de sites web qui pourraient être utilisés ensuite par différentes bibliothèques pour récupérer, cataloguer et conserver des publications politiques éditées en ligne, gouvernementales ou autres. Un intérêt particulier devraient être accordé aux activités et aux mouvements politiques de dimension locale, comme par exemple les événements ayant conduits en 2003 à l'élection d'Arnold Schwarzenegger comme gouverneur de Californie.

Dans un autre registre, l'Université de New York travaille aussi à la mise en place d'une Bibliothèque numérique d'Afghanistan (**Afghanistan Digital Library**), qui permettra la collecte, la restauration, le catalogage, la numérisation et la mise en ligne d'un maximum d'ouvrages publiés dans ce pays entre 1870 et 1930³⁰. Mené en partenariat avec le Département d'Etat américain, ce projet a clairement des objectifs de conservation, à la différence du programme du même nom initié par la Bibliothèque de l'Université d'Arizona³¹.

Archivists' toolkit — la boîte à outil des archivistes — est un programme collaboratif dont l'ambition est de développer un logiciel libre, facile à installer et à utiliser, tout en disposant de toutes les fonctionnalités nécessaires à la gestion et au catalogage d'archives (y compris électroniques)³². Cet outil est mis gratuitement à la disposition de tous les archivistes, qu'ils travaillent dans des petites structures ou dans de gros consortiums comme par exemple Online Archive of California.

Autre initiative de la New York University : la création, par l'équipe chargée du développement de notre bibliothèque numérique, d'une base de donnée consacrée à la musique américaine enregistrée (**DRAM, Database of Recorded American Music**)³³ Elle a été construite en partenariat avec la compagnie de disques New

³⁰ afghanistandl.nyu.edu

³¹ www.afghanresource.org/joomla

³² www.archiviststoolkit.org

³³ www.dramonline.org

World Records, spécialisée dans les enregistrements de types de musique délaissés par les grandes compagnies commerciales. Le système mis en place offre aux établissements souscripteurs l'accès à des versions numérisées de l'ensemble des enregistrements du catalogue New World.

Après le son, les images. Mené en collaboration avec l'Educational Broadcasting Corporation (EBC), la Fondation pour l'éducation WGBH et la chaîne de télévision à but non lucratif PBS (Public Broadcasting Service), notre université participe aussi à un projet d'archivage des programmes de la télévision publique diffusés sous une forme numérique (**PTV Digital Archive, Digital archive of Public Television**)³⁴. Financé dans le cadre d'un plan national pour la conservation de ce nouveau type de productions (National Digital Information Infrastructure Preservation Program), il a pour objectif de définir les premières procédures, structures et normes nécessaires à la mise en place d'une véritable politique d'archivage des images numériques. L'EBC et la Fondation WGBH sont les deux principaux producteurs de programmes pour la télévision publique aux Etats-Unis, diffusés ensuite dans tout le pays par PBS. A eux trois, ces organismes représentent l'essentiel de la télévision à but non lucratif. Les quatre partenaires s'intéresseront particulièrement à des programmes aussi importants que « Nature », « American Masters », « Nova » ou « Frontline », qui de plus en plus ne sont désormais produits que sous une forme numérique, en fonction par exemple du nouveau standard de diffusion haute définition HDTV. La question de l'archivage des autres productions électroniques — sites Internet, par exemple — qui accompagnent parfois la diffusion des programmes est elle aussi prise en compte.

L'équipe de notre Bibliothèque numérique a participé aussi à la mise en place de la vidéothèque numérique du Hemispheric Institute (**Hemispheric Institute Digital Video Library**)³⁵, un consortium d'établissements, d'artistes et d'universitaires dont l'ambition commune est de réfléchir aux relations qui existent entre les différentes formes d'expressions physiques et corporelles (résumées par le terme "performance») et la vie politique et sociale des Amériques. Avec l'aide du centre d'informatique pour les sciences humaines (Humanities Computing Group) de l'université, plus de 200 enregistrements vidéo de performances ont ainsi été mis en ligne, regroupés dans une base de données librement consultable.

Enfin, je terminerai par l'évocation du projet **Second Avenue Online**, destiné à conserver, sous forme d'enregistrements numériques, la mémoire du théâtre yiddish d'autrefois, tout en lui permettant aussi de toucher peut-être de nouveaux publics. L'intitulé de ce projet renvoie au Yiddish Art Theatre, créé en 1926 par Maurice Schwartz sur la 2^e avenue au croisement de la 12^e rue. On trouve parmi ces archives des histoires du théâtre yiddish, des éléments d'histoire orale, des textes

³⁴ dev.ptvdigitalarchive.org

³⁵ www.hemisphericinstitute.org/eng/hidvl/index.html

de pièces rédigés en yiddish ou en anglais, des photographies, ainsi que des enregistrements vidéo ou audio.

Contents

Table des matières

Introduction générale	5
General introduction	9
Netherlands. Digitisation policies in Netherlands : about IISG	14
Pays-Bas. Politiques de numérisations aux Pays-Bas : le cas de l'Institut international d'histoire sociale d'Amsterdam	18
Switzerland. Networks and digitisation policy in Switzerland	24
Suisse. Mise en réseau et politiques de numérisation en Suisse	27
Deutschland. The indexing of digital documents in Germany	31
Allemagne. L'indexation des documents électroniques en Allemagne	34
Italy. Digital experience in Italian libraries	38
Italie. Expériences de numérisation dans les bibliothèques italiennes	42
France. Numérisations réalisées en histoire ouvrière et sociale	44
France. Digitisation in labour history	47
1. BDIC	52
2. Génériques	55
3. La Fraternelle	57
4. Mission CAARME	57
5. IRHSES	59
6. Fondation Jean-Jaurès	62
7. Centre Georges-Chevrier, Dijon	65
Belgium. International workshop for an international concerted policy of labour history Archives digitization: digitization and digital archives in Belgium and Amsab.....	67
Belgique. Numérisation et archives numériques en Belgique et à l'Amsab.....	71
Belgium. Ialhi international workshop. Summary report for Amsab-ISH	76
Belgique. Ialhi international workshop. Summary report for Amsab-ISH	79
United Kingdom. Digitization issues	83
Royaume-Uni. Expériences de numérisation	84
Sweden. Labour movement Archives and Library.....	86
Suède. Bibliothèque et Archives du mouvement ouvrier.....	90
United States. Some American examples.....	94
Etats-Unis. Quelques exemples américains.....	98