

Digitalisering van de federale wetenschappelijke instellingen

Fase 1

Rapport d'état de la situation au 01/12/2009

La première phase du Plan de digitalisation des établissements scientifiques fédéraux et de la CRB a été lancée en novembre 2005. Le budget total réservé pour cette phase est de 16.59 millions EUR et est ventilé comme suit :

2005 : 2.20 millions EUR
2006 : 4.30 millions EUR
2007 : 4.30 millions EUR
2008 : 2.25 millions EUR
2009 : 3.54 millions EUR

Initialement, elle devait s'achever fin 2008 mais suite aux retards pris dans l'avancement des travaux, elle a été prolongée une première fois jusqu'à fin 2009 puis une seconde fois jusqu'à fin 2010 afin d'assurer le pont avec la deuxième phase.

1. Travaux de digitalisation menés par les ESF en dehors des projets de la phase 1

En dehors des activités exécutées dans le cadre des 9 projets de coopération prioritaires décrits au point 2, les travaux suivants de digitalisation ont été menés par les ESF et la CRB :

a. BRB

Le projet Belgica.kbr.be : réalisation de la bibliothèque virtuelle de la Bibliothèque royale

Belgica.kbr.be :

- offre un accès gratuit et permanent à différentes catégories de documents patrimoniaux :
 - o manuscrits, imprimés, cartes, partitions de musique, enregistrements sonores,
 - o collections de monnaies, de médailles, de dessins et d'estampes ;
- est une bibliothèque de référence orientée vers les Belgicana, c'est-à-dire le patrimoine en lien direct avec la Belgique et son histoire ;
- un moteur de recherche indexant plusieurs milliers de pages de journaux, des expositions virtuelles.

La numérisation est effectuée selon deux modes opératoires :

- la numérisation en interne,
- la sous-traitance à une firme spécialisée.

La numérisation en interne est mise en œuvre au cas par cas en fonction d'un certain nombre de priorités :

- les images destinées à étoffer et éclaircir les descriptions catalographiques (ex. Estampes),
- les sous-collections qui constituent les noyaux du patrimoine de la Bibliothèque royale (ex. la Librairie des ducs de Bourgogne),
- les Belgicana, c'est-à-dire les ouvrages spécifiques à la Belgique et les publications belges,
- les documents en rapport avec divers projets de recherche ou d'expositions menés ponctuellement à la Bibliothèque royale,
- les documents fragiles, dont la consultation doit être limitée au maximum,
- les pièces uniques et exceptionnelles.

En pratique, la numérisation en interne ne fait l'objet d'aucun financement structurel, ni ordinaire ni extraordinaire. Les opérations techniques sont réalisées au cas par cas, au service photographique et de manière décentralisée, au sein des sections où les documents sont conservés.

La sous-traitance à une firme spécialisée :

Ce mode opératoire est systématiquement envisagé pour deux types de collections :

- les journaux, en raison du grand volume de pages et de la nécessité d'un OCR (reconnaissance automatique de caractères),
- les documents sonores, parce que leur numérisation nécessite un équipement technique particulier.

En pratique, la sous-traitance n'est envisagée que dans le cadre de projets ponctuels de numérisation de masse dotés d'un financement extraordinaire.

Tous les documents accessibles dans Belgica.kbr.be sont d'une qualité d'image satisfaisante à l'écran mais empêchant toute impression en qualité d'édition. Tous ces documents sont par ailleurs assortis de métadonnées descriptives et techniques, respectant les prescriptions internationales adoptées en la matière.

Dans un premier temps, l'ensemble des documents numérisés par la Bibliothèque royale et mis en ligne dans Belgica.kbr.be appartiennent au domaine public. Il s'agit d'œuvres littéraires ou artistiques, dont le dernier auteur survivant est décédé depuis 70 ans au moins et, du point de vue des droits voisins (interprètes, producteurs de phonogrammes et organismes de radiodiffusion), d'œuvres inédites ou communiquées au public depuis plus de 50 ans.

Parallèlement, afin de ne pas écarter le XXe siècle du projet, la Bibliothèque royale étendra la mise en ligne de son patrimoine numérisé à des œuvres littéraires et artistiques soumises au droit d'auteur pour lesquelles elle aura préalablement conclu un accord en la matière avec les ayants droit. Elle donnera également accès à certaines œuvres dites orphelines, c'est-à-dire des œuvres soumises au droit d'auteur mais dont les ayants droit n'ont pu être identifiés et pour lesquelles aucune autorisation de numérisation et de mise en ligne n'a pu être demandée.

La Bibliothèque royale a mis tout en œuvre pour ouvrir au public son prototype Belgica.kbr.be à la mi-décembre 2008 (<http://belgica.kbr.be>).

Voir aussi dans Projet DI/02, sous-projet A, Autres projets de numérisation de journaux

b. MRAH

Pas d'autres projets de digitalisation que ceux financés par le Plan de digitalisation.

c. IRSNB

Projets de digitalisation récents:

- projet MARS ("Multimedia Archaeological Research System"), digitalisation des collections archéologiques (anthropologie et préhistoire)
- projet GNOSIS ("Generalized Natural sciences Online Spatial information System"), accès aux données de collection via une cartographie (répartition, biogéographique) de ces données (www.gnosis.be)

Etant donné que la gestion des collections scientifiques est une des missions de l'IRSNB, la digitalisation de ces données fait partie des tâches quotidiennes d'un bon nombre de membres du personnel de l'IRSNB et dont le financement provient de ressources propres ou de divers projets de recherche. Vous en trouverez ci-dessous quelques exemples (la liste n'étant pas exhaustive) :

- Base de données Physis (catalogue des habitats)
- BarCoVer (Barcoding Congolese Vertebrates): " DNA barcoding of selected vertebrates from the Congo Basin"
- ANT'PHIPODA ("Biodiversity Reference Centre for Antarctic Amphipod Crustaceans") - <http://www.sciencesnaturelles.be/amphi/>
- TNT ("The Nanderthal tools")
- African Rodentia (<http://projects.biodiversity.be/africanrodentia>)
- Collections de minéralogie: plus de 47.000 spécimens en collection dont la majorité est maintenant digitalisée
- CBD National Focal Point (collection iconographique Congo belge, période coloniale) digitalisation en cours (cartes d'index, diapos, cartes postales, aquarelles, ...)
- Bases de données écologiques (biologie de la conservation)
- Darwin: plateforme développée en interne permettant aux différents départements scientifiques de l'IRSNB de numériser les catalogues des collections d'histoire naturelle (consultation web via <http://www.naturalsciences.be/darwin/>)
- Numérisation continue en cours du catalogue de notre bibliothèque (650.000 livres et périodiques). Consultation possible via <http://www.libis.be>

d. MRAC

Buiten de collecties die het voorwerp uitmaken van een of meerdere van de negen digitaliseringsprojecten, werden reeds verschillende collecties en datasets van het KMMA met eigen middelen gedigitaliseerd. De resultaten van deze projecten zijn toegankelijk via de website van het KMMA, onder de sectie www.metafro.be. Meest opmerkelijk zijn de elektronische inventaris van het Henry Morton Stanley-archief, de "Tervuren Wood Xylarium" en de verschillende zoölogische databanken. Via online databanken wil het KMMA haar cultuur- en natuurhistorisch patrimonium beter toegankelijk maken voor zowel wetenschappers als voor het brede publiek.

Exemples :

- « *Le Film Colonial. Congo, Rwanda, Burundi. Projet de sauvegarde et de rapatriement* » : projet multipartenarial (MRAC/Section Histoire de la période coloniale, CRB, KADOC-KUL, ULB, FUSL, Université de Kinshasa, Université du Burundi, Université Nationale du Rwanda, Institut des Musées Nationaux du Rwanda). Projet en 3 phases : digitalisation des collections de films d'archives du MRAC-CRB-KADOC ; rapatriement des films numérisés en Afrique centrale ; valorisation de ces archives filmées par le renforcement du réseau d'échanges interuniversitaires belgo-africain autour de l'histoire et de la mémoire coloniales.
- *Microfotografie van commerciële houtsoorten (Microphotographie des espèces commerciales de bois)* : Digitalisation de microphotographies d'espèces commerciales de bois et saisie en base de données destinée à des projets de formation sur la connaissance du bois.

e. MRBAB

Il s'agit des projets suivants :

- la numérisation du fonds "Octave Maus" des Archives de l'Art contemporain en Belgique (AACB);
- la campagne de prises de vue scientifiques (infrarouges, etc.) de la collection d'oeuvres de Rubens conservées aux MRBAB, dans le cadre du projet de recherche "Rubens" et de l'exposition associée à ce projet;
- la numérisation de la correspondance de René Magritte, ainsi que de photographies et de divers documents le concernant (environ 800 documents au total), aux AACB.

f. IRM

- Projet GNOSIS (financé dans le cadre du "Multiannual Information Society Support Program", 2001-2008). IRM-MRAC-IRSNB. Projet terminé.

L'IRM a réalisé l'inventaire des données météorologiques du Congo, du Ruanda et du Burundi disponibles dans les archives de l'IRM pour la période entre la fin du 19e siècle et l'indépendance de ces pays. Un logiciel de type GIS a été mis en place pour l'exploitation de l'inventaire.

- Depuis 2000, projet général à l'IRM de centralisation des données et de suppression progressive de l'encodage manuel.

Installation à l'IRM en 2000 d'une base de données centralisée ORACLE et transfert progressif de l'ensemble des bases de données de l'IRM sur ce support.

Ces dernières années, remplacement de documents manuscrits ou de données encodées manuellement par un encodage informatique et centralisation de cette information sous forme digitale dans la base de données Oracle. Exemples actuels de produits/données digitalisés disponibles dans Oracle :

- o tous les textes de prévisions météorologiques (bulletins météorologiques);
- o tous les textes d'avertissements en cas de prévisions de situations météorologiques dangereuses. Ces informations sont maintenant envoyées par e-mail ou fax directement à partir des informations encodées sous forme digitale dans la base de données Oracle;

- depuis 2007, les observations climatologiques effectuées par les observateurs du réseau climatologique belge sont transmises sous forme électronique à l'IRM (via une application téléphonique) par les observateurs. Les données sont immédiatement transférées dans la base de données Oracle (jusqu'alors, depuis la fin du 19^e siècle, ces données étaient transmises à l'IRM une fois par mois sous forme manuscrite) ;
- depuis mi-2007, les observations climatologiques effectuées par les stations d'observations de Belgocontrol sont transmises à l'IRM sous forme électronique (en lieu et place de documents manuscrits). Ce même changement devrait être effectué pour les stations d'observations de la Force aérienne en 2010. L'IRM est en charge de la centralisation de toutes les données d'observations climatologiques en Belgique.

g. AGR

- In het kader van het project DIGIT maar eveneens met eigen middelen, externe financiering (Lotto) en dank zij samenwerking met derden heeft het Rijksarchief een solide infrastructuur opgezet voor de creatie, het beheer, de lange termijn bewaring en de valorisatie van digitale content. Er werd geïnvesteerd in materieel (scanners, netwerk, digitale leeszaal, ...), in opleiding van personeel (eigen personeel en personeel in kader van DIGIT) en vooral in solide en gevalideerde werkprocessen waarin het gebruik van internationale standaarden en normen centraal staat.
- In het kader van de digitalisering heeft het Rijksarchief twee doelstellingen. De eerste doelstelling is het on-line beschikbaar stellen via een zoekmotor van alle metadata die met betrekking tot de archieven beschikbaar zijn en die de onderzoeker de weg moeten wijzen naar de archiefdocumenten. Concreet betekent dit alle inventarissen, gidsen, nadere analyses, ... In het kader van een retroconversieproject wordt alle beschikbare informatie uit de gepubliceerde inventarissen, gidsen, ... gestructureerd en in digitale genormaliseerde vorm geplaatst. Nieuwe inventarissen worden onmiddellijk in deze genormaliseerde vorm gecreëerd. Momenteel zijn 99 % van de archiefbestanden beheerd door de verschillende vestigingen van het Rijksarchief terug te vinden via de website. Voor bijna 40 % van de bestanden is de volledige inventaris die een beschrijving tot op stukniveau geeft, doorzoekbaar. Het aanbod groeit dagelijks aan.
- Een tweede doelstelling is de digitalisering van de documenten zelf.
- En dehors du projet exécuté dans le cadre du projet de coordination « Copie numériques de documents » du Plan de digitalisation, des petits projets de numérisation sont exécutés par les Archives. Tous les dépôts en province ont eu la possibilité de commencer la numérisation de documents souvent demandés en salle de lecture et/ou en mauvais état et/ou pour répondre à des projets de recherche scientifique. Cette numérisation se fait en interne et le résultat est immédiatement intégré dans la base de données disponible en ligne. Quelques exemples : les matricules de l'université de Louvain, répertoire du notariat, chartes, Ceci nous permet aussi de collaborer avec les universités pour les travaux de séminaire des étudiants.
Presque 40 projets de numérisation sont en cours ou sont déjà terminés. Plusieurs milliers de documents sont à intégrer dans la « salle de lecture » dès que le stockage est disponible.
- Un projet important est la numérisation de la collection des matrices des sceaux. Le projet de numérisation permettra de mieux valoriser notre collection, qui est la deuxième la plus importante au monde.
- Vu sa fragilité, une collection de plaques de verre concernant les mines au Limbourg a été numérisée et sera mise en ligne dans un système wiki qui permettra au lecteur d'ajouter des données.
- "Digitalisering" is zeer belangrijk in het kader van samenwerking met andere instellingen (zowel in binnen- als buitenland). Het Rijksarchief sloot overeenkomsten af voor digitalisering van archiefdocumenten bewaard in het Rijksarchief met bij voorbeeld het Joods Museum voor Deportatie en Verzet, SOMA, Holocaust museum (Washington), ...
Het Rijksarchief is actief op het internationale vlak voor de valorisatie van het digitale erfgoed. In het kader van Europeana is het Rijksarchief actief lid van Apenet. Apenet is een project van een aantal

nationale archiefdiensten die als thematische agregator de correcte invoer van metadata vanuit archiefdiensten in Europeana voorbereiden.

- "Digitalisering" vormt vaak een belangrijk onderdeel in wetenschappelijke projecten gevoerd door het Rijksarchief. Bij voorbeeld: onderzoeksproject in samenwerking met Koninklijke Commissie voor Geschiedenis voor de ontsluiting van de briefwisseling met Wenen tijdens de Oostenrijkse periode. De documenten bewaard in het nationaal archief te Wenen die belangrijk zijn voor de Belgische geschiedenis worden gedigitaliseerd. Zij zullen bij de afronding van het project opgenomen worden in de digitale leeszaal.
- Een ander voorbeeld is het project rond armenbrieven waarbij documenten die in verschillende archiefdiensten bewaard worden gezamenlijk ontsloten worden en in digitale vorm ter beschikking gesteld worden.
- Een specifieke categorie zijn de projecten rond de digitalisering van kaarten en plannen. Naast een beperkt project in samenwerking met de FOD Financiën werd een ambitieus plan opgestart voor de valorisatie van historische kaarten in samenwerking met het Nationaal Geografisch Instituut, de Koninklijke Bibliotheek en het Afrikamuseum. Ook internationaal wordt een Europees project voorbereid in samenwerking met archiefdiensten in Nederland, Duitsland en Luxemburg rond digitale toegang tot *grenskaarten*.
- Het Rijksarchief is partner in het Histstat-project in samenwerking met de universiteiten van Brussel, Gent en Luik waarin statistische bronnen gedigitaliseerd en gevaloriseerd worden.
- Het Rijksarchief ontving een digitale kopie van een belangrijke collectie documenten die betrekking hebben op de vervolgingen tijdens de Tweede Wereldoorlog en die door International Tracing Service (ITS) in Bad Arolsen in verschillende archiefdiensten (ook in België) werden gedigitaliseerd. In totaal betreft het ca. 80 miljoen documenten.
- Les AGR ont investé dans la modernisation des salles de lecture afin de les adapter à la consultation de données numérisées. Alle leeszalen van het Rijksarchief zijn uitgerust met wireless netwerk. Lezers kunnen gebruik maken van de toestellen ter beschikking gesteld door het Rijksarchief of toegang krijgen tot de digitale documenten met hun eigen laptop.
- Momenteel is slechts een beperkt deel van de gedigitaliseerde content die in de voorbije jaren gecreëerd werd beschikbaar. De reden hiervoor is een gebrek aan storagecapaciteit. Niettegenstaande reeds verscheidene jaren een dossier ingediend werd bij de Regie der Gebouwen voor de inrichting van een serverlokaal met aangepaste electriciteits-voorziening en airconditioning voor de storagecapaciteit werd de uitwerking ervan steeds verdaagd. Gezien de Regie der Gebouwen eigenaar is van het gebouw in de Ruisbroekstraat 2-10 dient zij in te staan voor de structurele werken en kan het Rijksarchief deze werken niet zelf laten uitvoeren. De opleveringstermijn van de werken werd door de Regie der Gebouwen meermaals met enkele maanden uitgesteld zodat het zoeken naar een oplossing door eventueel tijdelijk uitbesteden van de storage aan een derde partij niet mogelijk was. De werken zijn gestart begin maart 2009 en de oplevering voorzien voor begin december 2009. Een belangrijk probleem is de complexiteit van de werken waarbij verschillende diensten van de Regie der Gebouwen betrokken zijn: infrastructuurwerken, HVAC, electriciteit en netwerkinfrastructuur. Verschillende diensten die bovendien van elkaar afhankelijk zijn voor de uitvoering van de werken.

h. IRPA

Eigen mensen werken aan de invoer van gegevens in de databanken, restauratiedossiers worden in een apart databanksysteem ingevoerd, begin tot opbouw van een groot archiverings- en backupsysteem op eigen middelen (in de hoop dat er later budget bijkomt hiervoor), onderzoek naar digitalisering radiografieën in gevorderd stadium, ...

i. CRB

A l'occasion de divers projets, la Cinémathèque Royale a numérisé des parties limitées de sa collection de films au cours des huit dernières années. En tout environ 500 heures de film ont été numérisées. Certains projets l'ont été pour compte de partenaires externes, d'autres font partie de projets d'édition DVD ou de projets de séance. Bon nombre de matériel a été digitalisé en vue de l'ouverture du nouveau musée du cinéma. Il s'agit aussi bien de films de fiction que de non-fiction, la plupart des films numérisés étant d'origine belge.

2. Etat d'avancement des 9 projets en réseau prioritaires

Projet DI/01 : Catalogues informatisés des bibliothèques des ESF et de la CRB

Ce projet met en commun tous les ESF (à l'exception de l'Institut d'Aéronomie spatiale) ainsi que le Jardin botanique de Meise et la Cinémathèque royale de Belgique autour de la constitution d'un catalogue centralisé numérique des ouvrages des bibliothèques des ESF. C'est la Bibliothèque royale de Belgique qui en assure la coordination.

La Bibliothèque royale de Belgique

Conversie van steekkaartencatalogi

Le travail préparatoire de conversion a été assuré pour la Bibliothèque royale par 10 personnes durant 6 mois au début 2007. Il a débouché sur la rédaction d'un cahier des charges destiné à désigner l'opérateur privé chargé de la numérisation proprement dite du premier lot d'environ 350.000 références bibliographiques. L'Institut royal des Sciences naturelles de Belgique, l'Institut royal du Patrimoine artistique se sont associés à ce cahier des charges en vue de sous traiter la numérisation (les autres ESF impliqués ont préféré mener la numérisation en interne avec du personnel affecté à cet effet). C'est la société AEL DATA qui fut sélectionné.

De gegevens van het eerste bestek zijn geïmporteerd in de elektronische catalogus. In het totaal zijn 215.240 beschrijvingen opgenomen, waarvan 12.716 voor het fonds Van Hulthem. Omwille van het belang van deze collectie, werden de geleverde gegevens aan een grondige controle onderworpen. Vooral de schrijfwijze van auteursnamen werd geverifieerd. De geleverde gegevens werden vergeleken met gegevens in andere catalogi, en in databanken, inzonderheid Hand Press Book Data.

De gegevens van het tweede bestek zijn geleverd, en worden gecontroleerd. In totaal gaat het om 330.458 beschrijvingen. Hiervan zijn er 59.880 gecontroleerd en klaar om in de catalogus geïmporteerd te worden. Dit zal gebeuren in de loop van de maand december.

Op 27 februari 2009 werd een lot, geraamd op 64.650 steekkaarten gegund aan AEL Data, Chennai, India. De steekkaarten zijn afkomstig van een aantal speciale afdelingen; Handschriften, Muziek, Kostbare Werken, Officiële Documenten. Het aantal gemaakte beschrijvingen bedraagt 56.998.

Op 20 maart 2009 werd de digitalisering van een lot van 1.300.000 steekkaarten gegund aan de firma Lumenscan, Brussel. Dit lot heeft betrekking op de catalogus van de Afdeling Handschriften, en bevat bibliografische informatie betreffende Handschriften in de Koninklijke Bibliotheek en in andere bibliotheken. De gegevens van de steekkaarten worden niet omgezet naar gestructureerde informatie en worden niet opgenomen in de elektronische catalogus; de gegevens zullen aangeboden worden als beelden. Voor getypte steekkaarten worden OCR technieken toegepast, zodat opzoeken mogelijk zijn op de volledige tekst van de steekkaart. Tevens wordt een zoekindex gemaakt op de begin- en eindsteekkaart van elke steekkaartenbak. Op die manier zal de gebruiker een aantal steekkaarten in sequentie kunnen bekijken.

Omdat het aantal steekkaarten voor deze catalogus uiteindelijk slechts 1.095.000 bedroeg, in plaats van de geraamde 1.300.000, werd begonnen met het digitaliseren van een onderwerpscatalogus met publicaties uit de 19^{de} eeuw. Hierdoor worden ook deze publicaties via het Internet op onderwerp ontsloten. In het totaal zullen voor deze catalogus 447.500 steekkaarten worden gedigitaliseerd.

De beschikbaarstelling van beide catalogi, via het Internet, is gepland voor januari 2010.

Kwaliteitsverbetering

1. Verbetering van dubbele records

De retrospectieve catalogus is gescand op identieke plaatskenmerken. Het gaat om ongeveer 25.000 beschrijvingen met dubbels. De records werden gedrukt, en nagelezen.

Naargelang het geval worden de gepaste maatregelen genomen:

- van identieke records met hetzelfde plaatskenmerk wordt er slechts 1 overgehouden;
- bij verschillende records met hetzelfde plaatskenmerk wordt nagegaan in het magazijn welke het juiste boek is.
- er worden ook correcties gedaan in de overblijvende beschrijvingen wat betreft het plaatskenmerk, schrijffouten, structuur van de beschrijving, ...

Dit werk is in september 2009 beëindigd.

2. Correcties in de retrospectieve catalogus

De hoofdwoorden in de retrospectieve catalogus moeten geüniformiseerd worden. Dit gebeurt op basis van alfabetische lijsten van de auteursindex. Dit werk verloopt in twee fasen:

- het uniformiseren van auteursnamen in de retrospectieve catalogus,
- in een aantal gevallen nagaan of de auteur aanwezig is in de algemene catalogus en een verbinding leggen met het authority record van de auteur.

De inhoud van het editie- en verwantschapannotatie is nagekeken. Dit veld bevat ondermeer verwijzingen naar overdrukken uit tijdschriften. Er werd nagegaan of de titel van het tijdschrift steeds op dezelfde manier is aangegeven.

Het aantal records dat in 2009 werd gecontroleerd bedraagt 200.039.

3. Correcties in de algemene catalogus

Correcties in de algemene catalogus hebben vooral betrekking op het verbeteren van authority records. Het aantal authority records, verbeterd in het kader van het retroconversieproject bedraagt, voor het jaar 2009 is nu 11.648.

Besluit

In 2009 werd, zowel voor de conversie van de steekkaartencatalogi als voor de kwaliteitsverbetering van de catalogi een aanzienlijke vooruitgang geboekt.

Musées royaux d'Art et d'Histoire

Les 4 bibliothèques des MRAH, auxquelles s'ajoute celle du Musée des instruments de Musique, hébergent plus de 300.000 ouvrages. L'automatisation des catalogues avait déjà débuté en 2003 et l'actuel projet permet de poursuivre le travail entamé. La priorité, tant aux MRAH qu'au MIM, est de se concentrer sur les ouvrages les plus récents; la rétroconversion d'ouvrages plus âgés suivra ultérieurement.

De beoogde doelstelling - ¼ van het bestand beschikbaar stellen via een geïnformatiseerde catalogus - is behaald. Referenties werden allemaal manueel, vertrekkende van het boek, in de catalogus ingevoerd.

Realisaties

- Algemene coördinatie van het project binnen de KMG (WP0)
- In samenspraak met het bibliotheekpersoneel selectie van de catalogi voor retroconversie (WP1)
- In samenwerking met bibliotheekpersoneel voorbereiding van de te converteren gegevens (WP2)
- In samenwerking met bibliotheekpersoneel retrocatalogering (WP6)
- In samenwerking met bibliotheekpersoneel controle geconverteerde gegevens (WP7)
- Coördinatie opname in Centrale Catalogus en eigen opac (WP8)

- Opvolging methodologie en preservering bibliografische records (WP9)

Resultaten

Stand van zaken ingevoerde titels:

- ca. 81.000 monografieën
- ca. 2.500 tijdschriften
- ca. 200 partituren
- ca. 2.000 AVM

Werkzaamheden verwant aan het project

Deelname aan Centrale Catalogus van de Federale Bibliotheken (financiering FOD P&O)

Institut Royal des Sciences naturelles de Belgique

“Retroconversieproject” betreft de digitalisering van de catalogus van bibliotheek van de Belgische Geologische Dienst (BGD). Het gaat hierbij om een te digitaliseren lot van 60 768 steekkaarten.

Stand van zaken november 2009

Tot op heden werden volgende werkzaamheden verricht:

- Selectie en controle van de te converteren catalogoog. Referenties van de vermelde fiches werden nagekeken op hun aanwezigheid in de catalogus van het LIBISNET enerzijds en in de interne, reeds geautomatiseerde catalogus van de bibliotheek van KBIN –IRSCNB anderzijds.
- Scanning en conversie: uitbesteding aan AEL DATA. Het volledige lot van 60 768 kaarten werd ter plaatse ingescand in het eerste kwartaal van 2008.
- Levering geconverteerde data: laatste levering eerste trimester 2009. Alle materiaal werd ondertussen in MARC21 formaat geleverd.
- Verificatie data: circa 50 % voltooid, finalisering kwaliteitscontrole voorzien voor eind maart 2010.
- Migratie naar Aleph (huidig bibliotheekbeheerssysteem KBIN) incl. conversie naar de gemeenschappelijke catalogi van BIB.BELGIUM.BE en LIBISNET : in voorbereiding, finalisering voorzien laatste trimester 2010.

Musée royal de l’Afrique centrale

Brève description du projet

La bibliothèque centrale du MRAC, participe au projet DI/00/01 au même titre que de nombreuses autres bibliothèques des Etablissements Scientifiques Fédéraux : Bibliothèque Royale (BR), Musées Royaux d’Art et d’Histoire (MRAH), Institut Royal des Sciences Naturelles de Belgique (IRSNB), Cinémathèque Royale de Belgique (CRB), Musées Royaux des Beaux-Arts de Belgique (MRBAB), ORB (Observatoire Royal de Belgique), CEGES (Centre d’Études et de Documentation Guerres et Sociétés Contemporaines), Archives Générales du Royaume (AGR), IRPA (Institut Royal du Patrimoine Artistique). Un des objectifs du projet est, notamment, la participation des ces institutions au catalogue informatisé des bibliothèques des établissements scientifiques.

Les bibliothèques du MRAC disposent de richesses qui sont souvent uniques et donc essentielles pour la recherche scientifique. Contrairement aux collections des bibliothèques des autres ESF, les collections des bibliothèques du MRAC sont physiquement décentralisées dans les différentes sections du Musée et sont donc souvent, localisées dans différents bâtiments. Dans les sections, les collections de bibliothèques n’étaient pas encore répertoriées de manière optimale.

Pour le MRAC, les objectifs de ce projet sont :

- de créer un catalogue centralisé commun des bibliothèques du MRAC afin de localiser au travers d’une interface web (et donc de rendre accessible au public et aux chercheurs internes) l’ensemble des collections de livres, périodiques et tirés à part disponibles au MRAC

- de gérer plus efficacement les collections de bibliothèques du MRAC, malgré la décentralisation physique des collections
- de réaliser la catalographie dans un format standard international de catalographie (MARC21) permettant:
 - o de collaborer à divers réseaux de bibliothèques qu'ils soient nationaux ou internationaux
 - o d'effectuer des recherches efficaces et pointues au travers d'une interface web (via le site du Musée ou le site de tout autre réseau de bibliothèques) et d'augmenter de la sorte la visibilité du MRAC et de ses richesses.

Résultats

La première phase du projet a permis à la bibliothèque du MRAC et de l'Institut Royal des Sciences Naturelles de se doter d'un outil de gestion documentaire en participant au consortium Libisnet de la KULeuven (logiciel Aleph500).

Sur les 130.000 ouvrages que contient la bibliothèque du MRAC, plus de 21.000 ont été inscrits dans le système de gestion documentaire Libis depuis le début du projet de digitalisation.

Le projet DI/00/01 a permis au MRAC d'avancer considérablement surtout pour ce qui concerne les collections de livres en sciences humaines. La catalographie des livres au format MARC21 a été réalisée "livres en main" pour les sections suivantes : anatomie du bois, publications du MRAC, ethnographie, anthropologie sociale et culturelle, linguistique, histoire coloniale, musicologie.

Le projet est loin d'être terminé car il s'agit encore d'intégrer les collections de livres des départements de géologie et de zoologie, de la section préhistoire et archéologie ainsi que de la section d'histoire du temps présent. Ensuite, il s'agira de catalographier les collections de périodiques et de tirés à part pour toutes les sections du MRAC.

Actuellement, des bases de données réalisées dans le cadre du projet METAFRO contiennent pratiquement 70.000 enregistrements d'ouvrages (Library Catalogue des différentes sections scientifiques). Ceux-ci peuvent faire l'objet d'un versement dans Libis. Dans un premier temps ils seront versés dans une bibliothèque temporaire où chaque enregistrement sera validé « livre en main » afin de bénéficier d'un catalogue exempt d'anomalie.

Cinémathèque royale de Belgique

Le travail des deux personnes engagées à mi-temps à charge du budget dans le cadre du projet consiste à ré-encoder dans la base de données (catalogue numérique) du Centre de documentation (base de données accessible en ligne sur le site de la Cinémathèque et sur le portail bib.belgium.be) des métadonnées-métier encodées initialement sur des fiches papier, résultat du traitement documentaire des collections « non-film » (livres, publications périodiques, catalogues de festival, annuaires,...) de 1938 à 1992. Ces métadonnées ou références se rapportent à des titres de films, à des noms de personnes, à des éditions de festival de cinéma et à des entrées sujets, dont une fiche théorique existe déjà ou est encore à créer dans la base de données. Elles consistent en un renvoi d'un article dans une revue spécialisée en cinéma ou de pages ou chapitres d'ouvrages sur le cinéma vers un titre de film, une personne donnée, une édition de festivals ou une entrée sujet.

Outre les métadonnées, le catalogue numérique présente bien entendu l'inventaire des collections « non film ». Cet inventaire est à 100 % pour les livres (50.000 volumes), les publications périodiques (4.928 titres de publication), les fardes de presse (122.000), les photos (45.000 entrées) et à environ 30% pour les affiches.

Etant donné que ces métadonnées ne sont pas standardisées et sont parfois manuscrites, ce travail de récataloguage doit se faire manuellement et ne peut être automatisé, les informations devant être « décodées » voire vérifiées dans les collections mêmes.

Le personnel scientifique de l'institution participe à ce travail de rétrocatalogage dans le cadre de ses missions quotidiennes d'inventorisation et de dépouillement des collections non- film.

A titre indicatif, au 01/01/2008, donc avant l'engagement du personnel à charge du budget (1 équivalent temps plein engagé au 01/06/2008), la base de données contenait déjà 689.054 références pour 146.000 entrées, toutes ces références ayant été encodées par du personnel de l'institution.

En date du 23/11/2009, la base de données comptait 905.346 références pour 200.000 entrées, dont plus de 27.000 références provenant du rétrocatalogage effectué par le personnel engagé à charge du budget de l'état sur le projet.

Ce travail de rétrocatalogage est un travail de longue haleine car il porte sur le traitement documentaire effectué par l'institution sur près de 50 ans. Actuellement, seul environ 1/30ème du travail de rétrocatalogage a été réalisé.

Musées royaux des Beaux-Arts de Belgique

Van 1 maart 2007 tot 30 juni 2008 werden een 3600-tal werken in de catalogus online geïntegreerd door een bibliothecaris niveau B, aangeworven in het kader van het project. Deze persoon werd tot op heden niet vervangen. Ondertussen werd wel verder geretrocatalogiseerd door het vast bibliotheekpersoneel. Van 2005 tot 2009 werden ca. 300 publicaties door hen behandeld. De methodiek blijft dezelfde: de beschrijvingen worden gemaakt vertrekkend van de boeken zelf.

Observatoire royal de Belgique – Institut royal météorologique

Les catalogues des fichiers bibliographiques des bibliothèques de l'ORB et de l'IRM étaient déjà automatisés en partie (livres et séries depuis 1951 et périodiques depuis 1996). L'objectif du projet est donc de compléter ce travail pour les ouvrages antérieurs à 1951 et les périodiques antérieurs à 1996, ainsi que de donner un code-barre à toutes les unités bibliographiques. Il faut en outre contrôler et améliorer les descriptions des documents datant de 1951 à 1995. Le budget alloué dans le cadre du projet a permis l'engagement, de février à décembre 2008, d'une personne qui a encodé les fiches des livres et séries acquis entre 1900 et 1938 (10.125 fiches catalogiques).

Jardin botanique national de Belgique

Comme prévu, le budget mis à disposition du Jardin botanique national de Belgique, suite à la défection de l'Institut d'Aéronomie spatiale, a été utilisé pour engager un collaborateur de niveau B durant 4 mois pour le tri et l'encodage des ouvrages confiés en dépôt au jardin botanique par le Musée royal de l'Afrique centrale. Un millier de monographies et une centaine de périodiques ont été introduits dans le catalogue Vubis. Le travail est poursuivi par du personnel statutaire ou par des contractuels engagés sur fonds propres.

Centre d'Etudes et de documentation Guerre et Société contemporaine

Le rétrocatalogage d'un lot de livres reçu en dotation a été réalisé en interne par une personne engagée à cet effet sous la supervision d'un membre du personnel scientifique du CEGES.

Met het nog resterende beperkte saldo konden twee jobstudenten worden aangeworven. Ze stonden in voor de selectie (opsporen ontbrekende nummers), het ophalen en invoeren van de voor het SOMA relevante titels uit de collectie dubbels van de tijdschriften van het vroegere ministerie van Landbouw, die nu beheerd worden door de bibliotheek van de FOD-Economie. De studenten zorgden voor de invoer van deze tijdschriften in het systeem Pallas en de plaatsing in de magazijnen, wat een kleine interne reorganisatie

noodzakelijk maakte. Verder werden enkele collecties giften ingevoerd, die vaak toelaten onvolledige reeksen aan te vullen. Het gaat om vaak omvangrijke collecties zoals Tenir en Entre Nous uit het archief van het Onafhankelijkheidsfront, een gift van Jacques Wynants (onder meer over Verviers). Uit het fonds Pourvoyeur werd een honderdtal boeken ingevoerd. Trefwoorden worden toegekend door een medewerker van de bibliotheek van het SOMA.

Archives générales du Royaume

Le budget de ce projet est utilisé afin d'étendre les licences permettant aux catalogues des bibliothèques des archives de l'État dans les provinces de s'intégrer au système de gestion centralisé des catalogues. Le rétrocatalogage pour les dépôts des Archives de l'État dans les provinces a commencé et est consultable sur le site des archives.

In het kader van de gemeenschappelijke catalogus van de bibliotheken, leverde het Rijksarchief de gegevens aan met betrekking tot het boekenbezit van de centrale bibliotheek Brussel. Intussen werd eveneens gestart met de invoering van het boekenbezit van de Rijksarchieven in de provinciën. De nodige toestellen, licenties, opleidingen, ... werden voorzien opdat het eigen personeel de invoering van de catalogus zou kunnen verzorgen. Ook deze gegevens worden opgenomen in de centrale catalogus van de wetenschappelijke instellingen.

Aangezien het personeel slechts een beperkt deel van de werktijd kan besteden aan de invoering van de boeken in de catalogus zal deze invoer nog meerdere jaren in beslag nemen. Sommige rijksarchieven hebben een zeer rijke en soms unieke collectie boeken. Deze collectie is slechts valoriseerbaar naarmate deze via een webcatalogus ontsloten is.

In het kader van een uitbreiding van het project Digit01 wordt voorgesteld drie personeelsleden in dienst te nemen die de invoer van de gegevens in een centrale catalogus in de verschillende rijksarchieven ter plaatse uitvoert. Sommige rijksarchieven beschikken reeds over een lijst van het boekenbezit al dan niet ingevoerd in een catalografische databank. In het kader van het project wordt eveneens een budget voorzien voor de omzetting van deze gegevens. De kwaliteitscontrole wordt uitgevoerd door de in het kader van DIGIT01 aangeworven personeelsleden.

Institut royal du Patrimoine artistique

Les 60 à 80.000 fiches bibliographiques manuscrites que compte la bibliothèque de l'IRPA sont rétroconverties pour la plupart via OCRisation – via sous-traitance dans le cadre d'un appel conjoint avec l'IRSNB et la BRB. Certaines sont réalisées directement par le personnel de l'IRPA. Les fiches qui ne peuvent être converties sont disposées dans un catalogue virtuel spécifique qui migrera, selon la disponibilité future des budgets, dans un système ADLIB.

Alle fiches zijn gescand en de MarcXML bestanden zijn geleverd. Deze zullen in het nog op te leveren databanksysteem worden geïmporteerd na bijkomende controle in 2010.

De invoer van gegevens in de huidige bibliotheek databank wordt voortgezet met eigen middelen.

La première phase du « Plan de numérisation du patrimoine culturel et scientifique des Etablissements scientifiques fédéraux », constitue, à ce jour, le volet le plus considérable de la numérisation des journaux à la KBR, mais aussi le plus important projet de numérisation de la presse en Belgique.

Associant initialement les dossiers de coupures de presse de la Cinémathèque royale de Belgique (CRB-KBF), le « Projet DI/00/02 – numérisation de journaux et de coupures de presse a du être redéfini en deux sous-projets, la KBR et le CEGES-SOMA assumant le « sous-projet A : Numérisation de la presse belge » et la CRB le « sous-projet B : Numérisation des dossiers de presse de la CRB ». En effet, tant les techniques à mettre en œuvre pour la numérisation (semi-automatisées pour la KBR et le CEGES-SOMA, manuelles pour la CRB-KBF), que les droits d'utilisation, les publics cibles et les applicatifs à mettre en place étaient fondamentalement différents. Ce n'est que fin octobre 2006 qu'un arrêté royal libéra les budgets 2005 et 2006, ce qui officialisa enfin le lancement effectif du projet de numérisation de la presse belge.

Sous projet A : Numérisation de la presse belge pour la BRB et le CEGES

Mise en œuvre

C'est en 2007 que la mise en œuvre effective du projet DI/00/02 a effectivement été entreprise, avec :

- La réalisation d'une seconde phase d'études des collections de journaux, ayant pour objectif de déterminer le nombre total de volumes reliés concernés par la numérisation, d'évaluer le nombre global de pages concernées, d'analyser l'évolution des formats de pages et les épaisseurs de volumes, d'identifier les lacunes ou problèmes matériels divers et de définir préalablement toutes les mesures utiles à prévoir pour le traitement de ces volumes (qualité des papiers et des reliures, pages manquantes ou déchirées, langues d'édition, usage de la couleur, existence de supports de substitution, etc.) (janvier - mai 2007)¹;
- La rédaction du cahier des charges pour marché de services (février - juin 2007);
- La procédure par appel d'offres général, qui a abouti à l'attribution du marché à Belgacom, associé à la société luxembourgeoise Infotechnique (juin - octobre 2007);
- La mise en œuvre de la phase de numérisation externalisée (entamée en décembre 2007).

Objectifs

Les buts du projet DI/00/02 (sous-projet A) sont définis comme suit dans le cahier des charges: « La numérisation des journaux permettra de présenter l'oeuvre au grand public sur Internet et/ou intranet. La valeur ajoutée de cette démarche est la possibilité de faire des recherches pointues dans la structure et le texte intégral de l'oeuvre. L'accès numérique à l'oeuvre permettra de préserver les originaux fragilisés. Les données numériques (images et métadonnées) demandées respectent les demandes de la préservation numérique à long terme ».

Ces objectifs doivent être réalisés par :

- La numérisation, l'OCR et la production des métadonnées standard (selon les recommandations européennes) - METS/ALTO - de ca. 2.700.000 pages de quotidiens belges généralistes (KBR) ;
- La numérisation, l'OCR et la production des métadonnées standard de ca. 110.000 pages de journaux censurés et de journaux clandestins des deux guerres mondiales (CEGES-SOMA) ;
- L'exploitation de ces données numérisées, par la mise en place d'une plate-forme informatique de stockage et d'accès en ligne du type full-text, utilisant exclusivement des logiciels du domaine public (le format de consultation étant PDF)² ;

¹ Effectuée en 2005-2006, la première phase d'études portait sur l'ensemble de la collection des journaux belges conservés à la KBR et la sélection des titres devant être numérisés par priorité.

² Linux, Lucene, MySQL, Apache, Java, Nutch.

- La préservation à long terme de ce matériel numérisé.

Budget

Pour ce faire, le projet dispose en premier lieu d'un substantiel apport budgétaire de la Politique scientifique fédérale, d'un montant total de 2.272.000 euros. Ce budget extraordinaire couvre notamment :

- La sous-traitance pour la phase de numérisation/segmentation/OCR (KBR) et d'OCR/segmentation (CEGES/SOMA)³;
- Des frais de personnel (coordination générale du projet ; développements, coordination et mise en oeuvre de « Belgica », bibliothèque numérique de la KBR ; élaboration d'une base de données des journaux censurés et clandestins des deux guerres mondiales, assistance administrative) ;
- L'acquisition de matériel pour la numérisation (achat d'un bookscanner A2 par le CEGES/SOMA pour le scannage en couleur de la presse clandestine), le Quality Control (achat d'un progiciel sous Windows), le stockage, l'accès et la préservation à long terme des données numérisées (achat de disques durs pour stockage par le CEGES/SOMA, achat de serveurs + installation + maintenance par la KBR).

De leur côté, les deux institutions ont effectué sur fonds propres d'importants investissements :

- En personnel (élaboration et gestion scientifiques, techniques, administratives et pratiques du projet, workflow des volumes de journaux à numériser, Quality Control, développement et fonctionnement d'une interface de recherche, intégration des journaux numérisés dans « Belgica », volet éditorial, information, communication et publications, opérations de numérisation en « interne » des journaux clandestins des deux guerres) ;
- En matériel (achat d'un Bookeye Grayscale System A2 par le CEGES/SOMA pour le scannage en noir et blanc de la presse clandestine, mise en place d'une plate-forme de stockage et de préservation à long terme par la KBR).

Aspects qualitatifs et quantitatifs

Tous les projets de numérisation des journaux, en Belgique comme à l'étranger, ont mis l'accent sur l'importance primordiale de la sélection des matériaux prioritaires, tant du point de vue de la sauvegarde des collections que sous l'angle de la mise à la disposition du public⁴.

Le projet DI/00/02 vise au total la numérisation, l'OCRisation, la segmentation, la consultation et la préservation d'environ 2.800.000 pages de journaux belges, soit :

- Ca. 2.700.000 pages de quotidiens belges généralistes de 1831 à 1950 (initialement, 30 titres).
- Ca. 110.000 pages de journaux belges censurés des deux guerres mondiales (c'est-à-dire des journaux autorisés à paraître par les autorités d'occupation, soit 42 titres en 1914-1918 et 34 titres en 1940-1944 = 76 titres) et de journaux belges clandestins (soit 66 titres en 1914-1918 et 724 titres en 1940-1944 = 790 titres).

Ce qui représente un « grand total » de 896 titres, ce chiffre mêlant des journaux ayant été publiés durant plus d'un siècle avec des journaux n'ayant eu qu'une parution éphémère, voire unique.

En ce qui concerne les titres de la presse belge dite « généraliste »

La sélection a été effectuée au sein de la KBR (2005-2006), en concertation « informelle » avec des institutions (CEGES/SOMA) et des personnalités qualifiées du monde scientifique, universitaire et académique belge.

Tout en se conformant au plus près à l'objectif fondamental de sauvegarde des collections dégradées et menacées, cette sélection a été effectuée afin d'obtenir un éventail aussi représentatif et pertinent que

³ Soit un montant global de 627.000 € TVAC.

⁴ Cf. M. D'HOORE, D. LUYTEN et Th. DELPANCQ (eds.), *Le journal dans tous ses états – Het rijk van de krant*, op.cit.

possible. Ceci implique tout à la fois la prise en compte des équilibres chronologique (bonne couverture d'une période de 120 ans), géographique, linguistique, culturel et idéologique.

Dans cette optique, il a donc été tenu compte d'éléments essentiels, tels que l'existence d'importants journaux de province qui suppléent des décennies durant l'absence d'une presse d'audience « nationale », la variété d'un spectre idéologique très mouvant tout au long de ces 120 ans, le développement et l'affirmation de la presse en langue néerlandaise à la fin du XIXe siècle, ou encore l'émergence progressive d'une presse pluraliste, neutre ou « d'information ». Même si leur existence a été parfois brève, des organes émanant de courants d'opinion plus ou moins « marginaux » ont également été incorporés à la sélection.

Soulignons en outre que plusieurs journaux importants font ou ont fait l'objet d'une numérisation dans le cadre d'un autre programme. C'est le cas de *Het Belang van Limburg* et du *Soir*, qui ne figurent donc pas dans la sélection. C'est aussi – partiellement – le cas du *Grenz Echo* (de 1946 à 2006).

Les éléments suivants méritent également d'être soulignés :

- Dates extrêmes : le choix a été fait de numériser chaque journal à partir de sa date de fondation ou à défaut, du plus ancien numéro de sa collection conservée à la KBR. Le *terminus ad quem* arbitraire a été fixé au 31 décembre 1950.
- Journaux vivants et journaux morts : 17 des 30 quotidiens sélectionnés (57%) ont cessé de paraître, parfois depuis plusieurs décennies (1936 ou 1940). 13 autres quotidiens (43%) paraissent toujours à ce jour, relevant de 7 éditeurs différents.

En ce qui concerne les titres de la presse censurée et la presse clandestine des deux guerres mondiales

La sélection a ici visé au contraire à l'exhaustivité, par le biais d'une identification systématique des collections conservées non seulement au CEGES/SOMA, mais également dans les autres institutions (e.a. KBR, AGR, Musée royal de l'Armée).

La phase de numérisation externalisée : situation au 31.12.2009

Entreprise en décembre 2007, la réalisation de la phase de numérisation externalisée s'est heurtée à deux types de difficultés sérieuses éprouvées par le prestataire désigné au terme de la procédure par appel d'offres général : des problèmes d'ordre technique, d'une part (OCRisation et cropping), et une inadéquation préoccupante de son rythme de production, compte tenu des quantités à traiter et des délais impartis, d'autre part.

Cette situation a mené à une situation de blocage formel du projet, en septembre 2008, suivie d'une longue phase de renégociation du projet, qui a trouvé une issue favorable en janvier 2009 (résolution des problèmes techniques, transfert du projet du site de production de La Walck (France) à Luxembourg, et redéfinition du calendrier des collectes de volumes physiques et des livraisons des fichiers numériques.

Grâce aux décisions prises de concert à l'issue de cette renégociation, le retard considérable subi dans la réalisation du projet a pu partiellement être rattrapé.

A la date du 31.12.2009, la totalité des 2.700.000 pages de journaux KBR et la totalité des 110.000 pages de journaux CEGES/SOMA se trouvent effectivement traitées selon les prescriptions du cahier des charges. Il convient de souligner qu'à titre compensatoire, le traitement d'un surplus a pu être pris en charge par le prestataire dans le cadre du bon de commande passé en exécution de l'attribution du marché. En outre, la « marge de sécurité » dont la KBR a tenu compte lors de la seconde phase d'étude des collections de journaux a permis la numérisation d'une dizaine de quotidiens belges supplémentaires (période 1831 à 1950), sélectionnés selon les critères analogues à ceux ayant présidé à l'établissement de la liste des titres prioritaires en 2007.

Validé le 7 juillet 2009 pour mise en application effective à partir du 24 juillet 2009, le calendrier des livraisons de données numériques s'est traduit quant à lui par la livraison de 60% des pages numérisées (ca.

1.700.000 pages), avant contrôle de qualité interne (chiffre au 01.12.2009). En effet, tous les fichiers produits (TIFF, PDF, JPEG, METS et ALTO) font l'objet d'un contrôle de qualité à leur livraison, avant d'être validés. Ces opérations de validation par échantillonnage s'effectuent au sein de la KBR, au moyen d'un progiciel sous Windows développé par le prestataire de services.

La préservation à long terme des données numérisées

Parallèlement à la phase de numérisation sous-traitée, la KBR a progressivement développé une plate-forme d'accès, de stockage et de préservation à long terme des données numérisées.

Le stockage réseau (Network Storage), qui assure l'importation des livrables informatiques et la mise en oeuvre du contrôle de qualité, s'articule sur 3 machines SUN (d'une capacité nette de 33 téraoctets chacune) tournant sous le système d'exploitation Solaris 10. Une procédure de stockage sur bandes magnétiques est intégralement automatisée par l'intermédiaire d'un serveur de gestion de stockage relié au stockage réseau. Par sécurité, un second exemplaire de ces bandes devra être réalisé pour être conservé à l'extérieur de la KBR.

Dans le cadre de la réflexion développée sur la préservation à long terme des données numériques, différentes possibilités d'intégration de cette plate-forme locale à une plate-forme commune élaborée avec un ou plusieurs partenaires ont par ailleurs été envisagées. Cette réflexion porte également sur le recours éventuel au microfilm « post-numérisation ».

Au 31.12.2009, la plate-forme de préservation à long terme est pleinement opérationnelle.

La mise en ligne des journaux numérisés

Conçu et développé en interne en 2008, amélioré et appliqué en 2009 dans le cadre du lancement de « Belgica » (bibliothèque numérique de la KBR), une interface de recherche prévoit la double consultation/recherche dans « Belgica » des journaux numérisés, OCRisés et segmentés, selon un mode « calendrier » (année-mois-jour) et un mode de recherche « full text » (http://belgica.kbr.be/fr/search/search_fr.html).

Au terme des opérations de contrôle de qualité et de stockage, l'ensemble des journaux numérisés dans le cadre du projet DI/00/02 devront être complètement intégrés à « Belgica » selon des modalités identiques. Des contacts étroits avec le CEGES/SOMA doivent par ailleurs aboutir à déterminer les modalités pratiques de consultation et de recherche parmi les journaux censurés et les journaux clandestins des deux guerres mondiales.

Il importe de souligner l'importance et les contraintes du volet éditorial de la mise en ligne des journaux numérisés, qui implique notamment :

- La rédaction d'une « manchette » pour chaque journal intégrant « Belgica » (FR/NL) ;
- La rédaction d'une « étiquette » catalographique (FR /NL) ;
- La rédaction d'une notice détaillée (FR/NL) pour chaque titre concerné (avec orientation bibliographique);
- La relecture, la correction et l'harmonisation de ces textes (+ mise en ligne).

Jusqu'à ce jour, ce travail a lui aussi été intégralement pris en charge en interne.

La question des droits d'auteur

« Idéalement », la KBR tout comme le CEGES/SOMA souhaitent la mise en ligne, à terme, de tout ou de la plus grande partie des journaux numérisés. Dans cette perspective, de nombreuses questions se posent en matière de droits d'auteur, notamment quant :

- A la définition aussi claire et précise que possible de ce qui peut être considéré, titre par titre, comme tombé dans le domaine public, d'une part, et ce qui doit être considéré comme protégé d'autre part ;

- Au statut des journaux « généralistes » ayant cessé de paraître, des journaux censurés et des journaux clandestins des deux guerres mondiales, notamment du point de vue de l'identité effective des éditeurs ou ayants droit des éditeurs de tout ou partie d'entre eux ;
- A la formule la plus adéquate de solution globale et durable à cette problématique (de la non-mise en ligne pure et simple de ce qui est protégé à la mise en ligne de l'ensemble avec l'accord formel et dans le cadre d'une convention - contre rémunération - avec les parties intéressées) ;
- Aux autres aspects qui pourraient idéalement être intégrés dans une éventuelle convention (participation des éditeurs à l'effort de numérisation, question du dépôt électronique volontaire) ;
- Aux conséquences éventuelles pour les usagers de la conclusion d'une convention (consultation payante, restrictions à la consultation, etc.)⁵.

De l'avis de toutes les parties intéressées, le débat sur la problématique des droits d'auteur relatifs aux œuvres numérisées et mises en ligne par les institutions scientifiques et culturelles est « extrêmement complexe » et appelle des « solutions spécifiques » développées dans le cadre de « négociations approfondies » associant les dites parties (institutions concernées, éditeurs de presse et sociétés de gestion des droits d'auteur).

Plus concrètement, la question fondamentale est d'assurer la sécurité juridique des institutions qui procèdent à des opérations de numérisation à des fins patrimoniales et scientifiques, et à la mise en ligne de ces patrimoines numérisés.

Sous cet angle, deux logiques distinctes, également légitimes mais volontiers contradictoires, doivent être conciliées :

- La réalisation de programmes de numérisation effectués par la KBR dans le cadre de ses missions légales prioritaires (préservation des patrimoines menacés de destruction irrémédiable et garantir à chacun le libre accès à l'information scientifique et culturelle)⁶;
- L'obtention, par les éditeurs et/ou par les auteurs, d'une « juste rémunération » à titre de reconnaissance explicite de leurs droits patrimoniaux sur les œuvres protégées légalement.

Nous devons en outre relever que :

- Du point de vue de la KBR, les implications de ce débat portent non seulement sur la nature des collections numérisées par priorité et leurs modalités d'accès, mais aussi sur l'objectif de recherche et de fidélisation de nouveaux publics (en Belgique comme à l'étranger) ;
- Du point de vue des éditeurs et auteurs, ce débat particulièrement sensible a pris récemment une acuité nouvelle, compte tenu tout à la fois des initiatives développées par des opérateurs privés actifs au niveau international (e.a. Google) et du contexte économique actuel, qui s'est traduit par une baisse sensible des revenus publicitaires de la presse écrite.

Dans le cadre du projet DI/00/02, les données à prendre spécialement en compte sont les suivantes :

- Certains titres de journaux sont détenus par les éditeurs de presse. C'est le cas de tous les titres encore actifs aujourd'hui (13 quotidiens) mais cela peut être le cas également pour des titres disparus, dans des circonstances parfois complexes à retracer (cessation de parution brutale, faillite, reprise par un autre éditeur) ;

⁵ Il importe néanmoins de contextualiser soigneusement cette problématique, en rappelant qu'à ce jour, elle ne concerne ni les journaux antérieurs à 1831, ni d'innombrables titres de la période 1831-1950, ni les journaux de la période de 1951 (1971 dans le seul cas du *Soir*) à nos jours, ceux-ci ayant délibérément été tenus hors du champ des collections prioritaires.

⁶ Lesquelles missions légales reposent sur l'AR du 20 avril 1965 relatif au statut organique des établissements scientifiques de l'État (Moniteur Belge du 15 mai 1965), l'AR du 30 octobre 1996 désignant les établissements scientifiques et culturels fédéraux (Moniteur Belge du 7 décembre 1996), l'AR du 31 décembre 1965 portant exécution de la loi du 8 avril 1965 instituant le dépôt légal à la Bibliothèque royale de Belgique (Moniteur Belge du 19 janvier 1966) et l'AR du 14 février 2008 modifiant l'AR du 31 décembre 1965 portant exécution de la loi du 8 avril 1965 instituant le dépôt légal à la Bibliothèque royale de Belgique (Moniteur Belge du 21 mars 2008).

- Les articles, dessins et photos sont couverts par le droit d'auteur. Une exploitation nouvelle doit donc requérir l'autorisation préalable des auteurs, sauf en ce qui concerne les œuvres tombées dans le domaine public, 70 ans après le décès de l'auteur ou 70 ans après la date de la publication dans le cas des œuvres anonymes, pseudonymes ou collectives. Les sociétés de gestion collective (e.a. SABAM, SACD et SOFAM), représentant chacune une fraction importante d'auteurs, sont à même de garantir une relative sécurité juridique. Mais tous les auteurs ne sont pas représentés par une société de gestion collective, et il est impossible à priori de connaître et de contacter chaque auteur individuellement, étant donné les quantités considérées et les fortes variations observées selon les époques et les titres en matière de pratiques d'identification des œuvres.

En tout état de cause, une application *stricto sensu* des normes actuelles sur le droit d'auteur limiterait la libre mise en ligne des journaux numérisés aux années 1870⁷. L'exploitation en ligne de ce patrimoine numérisé doit donc faire l'objet d'un accord tant avec les éditeurs qu'avec les représentants des auteurs. Cependant, l'application effective des dispositifs législatifs sur le droit d'auteur se heurte aux interprétations parfois éloignées des éditeurs de presse, d'une part, et des sociétés d'auteurs, d'autre part. La définition d'une position commune aux uns et aux autres constitue donc un préalable indispensable à la conclusion d'un accord durable et sûr entre toutes les parties.

D'emblée, la KBR et le CEGES/SOMA ont tenu à exprimer leur pleine et entière prise en compte des dispositifs législatifs sur le droit d'auteur, en accordant une priorité aux journaux ayant cessé de paraître et en ne dépassant en aucun cas la date du 31 décembre 1950. Toutefois, s'il est indéniable que « *Les droits de propriété intellectuelle sont un outil essentiel pour promouvoir la créativité, les biens culturels de l'Europe doivent être numérisés, mis à disposition et conservés dans le respect absolu des droits d'auteur et droits voisins* »⁸, la réalisation aussi complète que possible des objectifs de préservation du patrimoine rend inévitable et indispensable la numérisation de documents protégés.

En conséquence de quoi, le choix a été fait de ne mettre en ligne, à ce jour, aucun journal qui puisse soulever quelque problème que ce soit en matière de droits d'auteur. Seule la consultation sur place (dans l'enceinte de la KBR) est autorisée sans possibilité d'impression ou de sauvegarde sur un quelconque support que ce soit.

Au cours de l'année 2009, la réflexion entreprise sur la question des droits d'auteur s'est notamment concrétisée par :

- La préparation de la publication des actes de la Journée d'études KBR-CEGES/SOMA du 18 novembre 2008 (« 'La presse est libre'... L'exploitation en ligne des journaux belges des 19^{ème} et 20^{ème} siècles » - « 'De pers is vrij...' Digitalisering en online ontsluiting van Belgische kranten uit de negentiende en twintigste eeuw ») ;
- L'organisation de plusieurs réunions de travail avec les sociétés de gestion des droits d'auteurs (janvier - mars 2009) ;
- Le développement de contacts réguliers et étroits avec l'Office de la propriété intellectuelle (SPF Economie, P.M.E., Classes moyennes et Énergie) ;
- L'adhésion au groupe de travail « Copyright » de la *Conference of European National Librarians (CENL, septembre 2009)* ;
- La participation active de la KBR et du CEGES/SOMA au groupe de travail constitué à l'initiative du président du SPP Politique scientifique fédérale (première réunion : 16.10.2009) ;
- La participation active au workshop « Problématique de la propriété intellectuelle » de la conférence «Europeana en de digitale ontsluiting van cultureel erfgoed» organisée par Kunsten en Erfgoed (16.12.2009).

⁷ Selon l'hypothèse de travail basée sur les droits exigibles par les héritiers ou ayants droit d'un journaliste qui aurait signé son premier article à l'âge de 20 ans et serait décédé à l'âge de 90 ans.

⁸ Recommandation de la Commission européenne 2006/585/CE/ du 24 août 2006 sur la numérisation et l'accessibilité en ligne du matériel culturel et la conservation numérique.

Autres projets de numérisation de journaux

L'expertise scientifique, technique et administrative développée à la KBR dans le cadre de l'élaboration et de la réalisation du projet DI/00/02, ainsi que la richesse exceptionnelle de ses collections de journaux, ont notamment contribué à la mise en œuvre, dans le cadre de partenariats bi- ou multi-latéraux, d'une série d'autres projets de numérisation d'organes de presse.

Outre le projet KBR-Archives de la Ville de Bruxelles portant sur la numérisation du quotidien bruxellois *Le Soir* (1887-1970), abouti en novembre 2007, et le projet KBR-CRNS Lyon 2 consacré à la numérisation de la *Gazette de Leyde* (1750-1789), entrepris en 2007 pour finalisation en 2011, nous pouvons épingler :

- La numérisation des journaux du Waasland (KBR-Erfgoedcel Waasland) (projet réalisé en 2009) ;
- La numérisation des journaux de la ville de Dixmude (KBR-Stadsarchief Diksmuide) (projet arrêté en novembre 2009, pour réalisation en 2010) ;
- La numérisation des journaux de la ville de Tongres (KBR-Erfgoedcel Tongeren, en cours d'élaboration depuis 2009) ;
- La numérisation de périodiques belges artistiques et littéraires des 19^e et 20^e siècles (KBR-Université libre de Bruxelles, projet arrêté en décembre 2009, pour réalisation en 2010) ;
- La numérisation de la *Gazet van Antwerpen*, 1891-1911 (KBR-Concentra Uitgeversmaatschappij, projet élaboré en 2009 pour réalisation en 2010).

Par ailleurs, plusieurs journaux belges particulièrement rares et précieux ont été numérisés en interne par la KBR, en 2009, e.a. le premier quotidien belge de langue néerlandaise, *Vlaemsch België-De Vlaemsche Belgen* (1844-1845), la *Deutsche Brüsseler Zeitung*, bi-hebdomadaire dirigé par Karl Marx (1847-1848) et l'hebdomadaire satirique *Uylenspiegel* (1863).

Communication et valorisation

Tout au long de l'année 2009, la réalisation du projet DI/00/02 a fait l'objet de nombreuses initiatives de valorisation et de communication, parmi lesquelles nous citerons notamment :

- La conférence de presse de lancement de « Belgica », bibliothèque numérique de la KBR (17.02.2009) ;
- Diverses conférences relatives à « Belgica », notamment pour l'Association belge de Documentation (ABD-BVD) et des établissements d'enseignement supérieur ;
- Le *Bulletin-Berichtenblad* du CEGES-SOMA (septembre 2009) ;
- La *Feuille de contact - Contactblad* de la KBR (septembre 2009) ;
- La journée « Portes ouvertes » de la KBR (20 septembre 2009)
- La journée de travail KBR-Koninklijke Bibliotheek Nederland (25.09. 2009) ;
- L' « Europeanaconferentie », organisée par Kunsten en Erfgoed (16.12.2009).

Articulations internationales

Au plan international, le projet de numérisation des journaux belges demeure l'un des apports prochains les plus substantiels de la Belgique aux politiques de numérisation développées au niveau européen, spécialement « Europeana », bibliothèque numérique européenne lancée à la Bibliothèque royale de Belgique le 20 novembre 2008 (<http://www.europeana.eu/portal/>).

Rappelons enfin que depuis l'été 2008, la KBR participe également au portail du Réseau francophone des bibliothèques nationales numériques (www.rfbnn.org).

En ce qui concerne plus spécifiquement le CEGES

De (interne) scanning van de sluikepers werd in 2009 voortgezet zodat alle bekende titels zullen zijn ingescand. Het ging om belangrijke collecties zoals CARCOB, AMSAB-ISG. Parallel werden waar mogelijk, nieuwe titels en nummers opgespoord. Dat was met name het geval voor de collectie van het Onafhankelijkheidsfront die in 2008 aan het SOMA werd overgedragen. Deze nummers werden in de collectie van het SOMA geïntegreerd en vervingen soms microfilms, fotokopieën of onvolledige exemplaren. Deze originelen werden verpakt in zuurvrije dozen. In de collectie van CARCOB werden nog onbekende titels en nummers gevonden. De microfilms van het SOMA werden verder ingescand. Deze gegevens

werden toegevoegd aan de database die bij het begin van het project werd aangemaakt. De scanning van de sluiipers is daarmee afgerond.

De gecensureerde pers werd gescand door de onderaannemer. De meeste kranten werden teruggevonden in de collecties van de Koninklijke Bibliotheek. Het SOMA stond in voor de voorbereiding van het transport naar de onderaannemer, in nauwe samenwerking met de collega's van de KBR. Titels/nummers uit andere bibliotheken werden verzameld en voorbereid voor transport naar de onderaannemer. De scanning van de gecensureerde pers zal eveneens afgerond kunnen worden. Het aantal te scannen bladzijden was hoger dan aanvankelijk geschat, omdat een aantal tot nu ongekende titels kon worden teruggevonden. De scanning kan nog gefinancierd worden binnen het voorziene budget. Deze titels werden toegevoegd aan de database. Twee titels werden op verzoek van de ontlenende bibliotheek door medewerkers van het SOMA gescand. Enkele titels zijn alleen beschikbaar in PDF. Deze en andere bijzondere formaten worden eerstdaags verwerkt en geïntegreerd in het hele pakket sluiipers.

De intern gescande gegevens werden verder verwerkt door de onderaannemer en geïntegreerd in de globale workflow. Dit impliceert dat veel arbeidstijd moest worden besteed aan de kwaliteitscontrole, in nauwe samenwerking met de collega's van de KBR, die ook instonden voor het kopiëren van de gegevens. Voor enkele problemen werd een definitieve, voor alle partijen bevredigende oplossing gevonden. Dat was o.m. het geval voor het eerste lot van de sluiipers, dat in zwart-wit en niet in kleur werd verwerkt, zoals oorspronkelijk afgesproken.

De KBR werkt aan de integratie van de gedigitaliseerde krantencollectie in Belgica, zodat de kranten online consulteerbaar zijn. Het SOMA werd bij deze taak betrokken.

Een medewerker van het SOMA volgde mee het dossier op van de auteursrechten waarvoor door de KBR naar een oplossing wordt gezocht. Er werden contacten gelegd met de dienst voor intellectueel eigendomsrecht van de FOD Economie en er werd een vergadering belegd met de vertegenwoordigers van de auteurs en de uitgevers.

Wat de promotie van het project betreft kan verwezen worden naar een artikel in het Berichtenblad van het SOMA en naar het bezoek van medewerkers van de Koninklijke Bibliotheek Den Haag aan de Koninklijke Bibliotheek en het SOMA in oktober.

Scanning, OCR, inclusief kwaliteitscontrole zullen zijn afgerond in 2009. Ook werd de basis gelegd voor de on line ontsluiting. Eén van de uitdagingen zal erin bestaan een oplossing te vinden voor het probleem van het copyright. Dit dossier wordt niet alleen van nabij opgevolgd – voor het SOMA heeft de problematiek een zekere specificiteit – er werden ook initiatieven genomen en aansluiting gevonden bij de werkgroep die in de schoot van het federaal wetenschapsbeleid werd opgericht.

Sous projet B : Numérisation des dossiers de presse de la Cinémathèque royale de Belgique

La Cinémathèque royale a pour objet de constituer et conserver une collection de films possédant un intérêt esthétique, technique et historique permanent, et d'assurer, dans un but d'intérêt esthétique et scientifique, la consultation de ces films. La collection compte à l'heure actuelle plus de 110.000 copies de films correspondant à plus de 55.000 titres.

Mais elle est aussi chargée de réunir la plus large documentation possible ayant trait à l'art cinématographique. Cette documentation est non seulement un instrument de travail pour le personnel scientifique de l'institution mais aussi pour les chercheurs, les historiens, les journalistes, les étudiants et autres passionnés de cinéma.

Le Centre de documentation de la Cinémathèque compte, dans le domaine du cinéma, parmi les plus importants d'Europe. Y sont conservés des ouvrages, des périodiques de cinéma, des annuaires, des catalogues de festival de films du monde entier, des photos, des affiches, des coupures de presse et des fonds d'archive.

En 2003, les conclusions de l'étude menée par le Bureau van Dijk portant sur la numérisation du patrimoine culturel et scientifique des ESF et de la Cinémathèque montraient que l'incertitude demeurait grande en matière de numérisation des films et notamment en ce qui concernait les standards de numérisation et la préservation à long terme des films numérisés.

De ce fait, la Cinémathèque a, dans le cadre de la première phase du plan de numérisation, mis la priorité sur la numérisation de sa collection de dossiers de presse. Celle-ci se compose de plus de 120.000 dossiers dans lesquels sont conservés une estimation de 3.000.000 de pages de coupures de presse provenant de quotidiens généralistes et de magazines belges et étrangers depuis les années 1910 jusqu'à aujourd'hui, ainsi que d'autres documents papier, tels les dossiers de presse officiels sur les films. Ces documents ont trait exclusivement au cinéma sous toutes ses formes et offrent un éclairage historique, esthétique, critique et scientifique sur tous les aspects du cinéma, ses courants esthétiques, ses tendances, l'évolution de la critique professionnelle, de l'exploitation des films, de la réception par les spectateurs, etc.

Du fait de la nature même du support (papier journal) et de la consultation quotidienne qui en est faite, cette collection subissait une dégradation endogène et exogène qui mettait en péril sa conservation sur le long terme, raison de l'aspect prioritaire de sa numérisation. La mise en valeur de cette collection et une consultation plus aisée de celle-ci offrent évidemment une valeur ajoutée.

Plusieurs institutions internationales de conservation du patrimoine cinématographique ont déjà initié une politique de digitalisation de parties de leurs collections « non film », à des fins de conservation, de valorisation ou de diffusion. La Bibliothèque du Film à Paris, Le British Film Institute à Londres ou le Nederlands Filmmuseum à Amsterdam (qui vient de terminer un travail identique de numérisation des coupures de presse portant sur 100.000 coupures) en font partie. La digitalisation de la collection de coupures de presse de la Cinémathèque s'inscrit parfaitement dans l'élan des réalisations menées hors de nos frontières.

La particularité de cette collection rendait une automatisation du processus de numérisation impossible. De ce fait, tout en faisant partie du même projet DI/00/02 portant sur la numérisation de journaux, il a été décidé d'en faire deux sous-projets, l'un portant sur une numérisation automatisée (Bibliothèque royale et CEGES), l'autre manuelle (Cinémathèque royale).

Via un appel d'offre européen, la Cinémathèque a cherché une firme extérieure chargée de la numérisation de la collection et de la fourniture d'une infrastructure informatique permettant le stockage et l'archivage des fichiers digitaux, la mise en consultation de ceux-ci pour le public et la poursuite de la numérisation en interne des nouveaux documents entrants.

De très longues tractations ont eu lieu avec le candidat retenu, les négociations portant notamment sur les solutions techniques et opérationnelles proposées par la firme extérieure (software et hardware, processus de numérisation,...) et sur les aspects financiers, l'enveloppe budgétaire disponible ne permettant pas de prendre en compte l'ensemble des solutions proposées par le soumissionnaire en réponse au Cahier des charges.

Le contrat a finalement été signé fin juillet 2008. Celui-ci, à la demande du soumissionnaire, prévoyait une phase de set up de 6 mois et le lancement de la numérisation de masse en janvier 2009 pour une livraison définitive fixée au 31/12/2009.

En date du 15 janvier 2009, le soumissionnaire a du reconnaître ne pas pouvoir garantir de mener à bien le projet et a décidé de bloquer également la commande de l'infrastructure informatique.

Après des mois de tergiversations, le soumissionnaire a finalement signé en mars un contrat de cession de la numérisation à une société turque désireuse de s'implanter sur le marché belge et chargée de la numérisation en Belgique, et de l'indexation et de l'océrisation des documents digitaux en Turquie.

La société turque a à son tour demandé quelques mois pour mettre sur pied son unité de numérisation en Belgique, recruter des opérateurs de scanning, tester, commander et paramétrer les scanners et déterminer le processus et le planning des activités.

En fin de compte, la numérisation n'a véritablement commencé que mi-mai 2009 et la livraison de l'infrastructure, prévue pour début avril 2009, n'aura finalement été opérationnelle que début septembre 2009.

Des problèmes techniques et opérationnels, dont la responsabilité incombe au sous-traitant du soumissionnaire, ont empêché la Cinémathèque de commencer les opérations de contrôle qualité des documents numérisés avant fin juillet 2009, lesquelles opérations ont été interrompues fin août du fait de mauvaises interprétations du contrat, contestées par le soumissionnaire, et qui oblige in fine la société turque à reprocesser l'ensemble des documents déjà numérisés (plus de 1.700.000 pages en date de fin août), principalement pour des questions de pages non pertinentes et d'erreurs d'indexation.

Dès le mois de février 2009, le retard accumulé et l'incertitude quant à la bonne fin du projet ont en outre contraint la Cinémathèque à détacher à partir de mars 2009 une personne supplémentaire de son Centre de documentation sur le projet, pour inventoriser, indexer et préparer la numérisation en local (à la Cinémathèque) des nouveaux documents entrants et de documents non encore inventorisés et indexés.

En ce début novembre 2009, les opérations de contrôle qualité n'ont pas encore pu reprendre et doivent donc à nouveau porter sur l'ensemble des documents numérisés et retravaillés, soit probablement un peu moins de 3.000.000 de pages à valider par la Cinémathèque. Il est prévu que ces opérations reprennent début décembre 2009.

Toujours en ce début novembre, la société a numérisé 2.300.000 pages (représentant 112.000 fardes de presse) et océrisé 2.100.000 pages. L'indexation réalisée sur 2.000.000 pages doit elle être retravaillée avant que l'ensemble des pages ne repassent par le stade du contrôle qualité réalisé par la Cinémathèque et validant les documents numérisés. Il est par ailleurs prévu que la société rescanne certains documents qui ont été mal scannés lors de la production de masse et repasse un OCR sur certains autres documents, le degré de qualité de l'OCR devant être contractuellement entre 95 et 98 %.

Le contrôle qualité porte sur un échantillon aléatoire de 4 % des documents numérisés, soit sur 120.000 documents environ (certains multi-pages). Ce travail nécessitera donc encore des mois de travail pour la Cinémathèque et ses opérateurs de contrôle qualité. Sans problème majeur, la bonne fin du projet est prévue avant fin 2010.

Par ailleurs, la Cinémathèque a commencé début septembre 2009 l'indexation et la numérisation en local, avec la nouvelle infrastructure acquise dans le cadre du projet, des nouveaux documents entrants (coupures de presse et dossiers de presse) et de parties de collection en attente de traitement documentaire. A ce stade, près de 5.000 documents ont été indexés et numérisés et sont versés en consultation dans les serveurs de stockage et d'archivage, et sont d'ores et déjà consultables via le logiciel de consultation ad hoc.

Les informaticiens de l'institution sont en charge de la maîtrise de la gestion des politiques de sauvegarde, des backups des fichiers numériques et des serveurs de stockage et d'archivage. Ils ont pour ce faire suivi une formation spécifique, laquelle devra être suivi de remises à niveau régulières.

Une réflexion doit encore avoir lieu sur la pérennité des données numériques et la pérennité de leur exploitation via toutes les formes adéquates.

Le respect des droits d'auteur, droits intellectuels et autres droits de tiers contraignent la Cinémathèque à ne proposer dans un premier temps la consultation des fichiers digitaux qu'en intranet dans sa salle de lecture sur des terminaux ad hoc et pour le personnel scientifique de l'institution. Néanmoins les deux parties (la Cinémathèque et le soumissionnaire) se sont engagées contractuellement à examiner les différentes possibilités d'exploitation licite, c'est-à-dire dans le respect des ayants droits, de ces archives numérisées, notamment via le web.

Pour rappel, l'inadéquation entre l'enveloppe budgétaire disponible et l'ensemble des solutions proposées par le soumissionnaire en réponse au Cahier des charges a obligé la Cinémathèque à opérer des coupes sombres dans l'offre technique du soumissionnaire. De ce fait, plusieurs logiciels, notamment un logiciel spécialisé en gestion des coupures de presse, n'ont pas pu être commandés à ce stade. De même, la migration de la base de données du Centre de documentation de la Cinémathèque, nécessaire pour avoir un langage informatique équivalent entre celle-ci et le logiciel DocCenter de consultation des versions digitales des coupures de presse, a dû être postposée. De ce fait, aucune corrélation entre les deux bases de données n'est possible, et les métadonnées encodées dans la DB du Centre de documentation doivent être une deuxième fois introduites dans DocCenter, sans possibilité de mise à jour. Enfin, le hardware finalement pris en compte dans le contrat définitif est minimal et non suffisant, par exemple en terme de capacité de stockage et d'archivage. Une grosse partie de ces manquements pourrait être prise en compte dans le cadre d'une extension du projet pour 2010 et l'attribution de budgets supplémentaires.

Projet DI/03 : copies numériques de documents

Ce projet a pour objectif de numériser une sélection de documents au sein de trois ESF : les Archives générales du Royaume et archives de l'État dans les Provinces (AGR), l'Institut royal des Sciences naturelles de Belgique (IRSNB) et le Centre d'Etudes et de Documentation Guerre et Société contemporaine (CEGES).

La sélection des documents, laissée à l'appréciation de chaque institution, avait pour point commun la haute valeur historique des documents, les risques encourus de dégradation ainsi que le potentiel de valorisation auprès des publics cibles.

Les deux premières années du projet ont essentiellement été consacrées à la mise au point du projet proprement dit : mise en place de l'infrastructure, des modes de gestion du projet, des synergies entre institutions, des processus de traitement des documents, de la formation du personnel, à la mise au point des métadonnées...

Archives générales du Royaume

Aux Archives, la priorité au sein des près de 200km de documents fut donnée aux sources dites « généalogiques » : registres paroissiaux, registres d'état civil – soit 1.5% du patrimoine total.

Les Archives ont opté pour une stratégie de numérisation diversifiée : certains documents ont été numérisés en interne (à l'aide d'une vingtaine de scanners A2 et d'un scanner A0 qui arriva fort tard en 2008 en raison de travaux d'infrastructures à finaliser par la Régie des Bâtiments pour accueillir l'outil).

Om op een vlotte manier documenten voor digitalisering aan derden te kunnen uitbesteden werd een openbare aanbesteding gelanceerd. Het resultaat is een raamovereenkomst waarin 15 verschillende loten gedefinieerd zijn op basis van de materiële kenmerken van het document.

D'autres documents ont été numérisés en partenariats avec les antennes provinciales des archives, d'autres encore sur base de collaborations externes (ainsi la Genealogical Society of Utah numérise gratuitement les documents envoyés en échange des métadonnées établies par les Archives) et une dernière partie a été sous traitée. Des softwares sur mesure ont été développés en interne aux Archives.

Institut Royal des Sciences naturelles de Belgique

A l'IRSNB, c'est principalement la collection unique du Malacologue Philippe Dautzenberg (1849-1935) qui a été retenue. L'IRSNB a réalisé lui-même la numérisation de ces collections à partir du 04 décembre 2006 jusqu'à présent en faisant usage de l'infrastructure mise à disposition par les Archives.

Op 25 november 2009 staat de teller van het aantal gescande pagina's op:

- 7979 tiff-bestanden Dautzenberg
- 20.787 tiff-bestanden andere KBIN publicaties

De omzetting van deze moederbestanden naar pdf-publicaties werd tijdens dit project ook reeds aangevat. De teller van het aantal pdf-publicaties staat tot op heden op:

- 222 pdf-bestanden Dautzenberg
- 133 pdf-bestanden andere KBIN publicaties

Deze publicaties zijn echter nog niet beschikbaar online voor het grote publiek.

Centre d'Etudes et de documentation Guerre et Société contemporaine

Au CEGES, ce sont les documents du procès Von Falkenhausen et Canaris qui sont convertis sous format digital eu égard à leur intérêt historique et scientifique.

De digitalisering van het procesdossier-von Falkenhausen kon in het voorjaar van 2009 worden afgerond. Intussen werd ook de systematische kwaliteitscontrole ervan voltooid. Voor dit dossier rest enkel nog het herscannen van een beperkt aantal pagina's en het via Pallas on line toegankelijk maken van de digitale beelden. De materiële voorbereiding van de digitalisering van het onderzoeksdossier tegen Canaris was eveneens begin 2009 afgewerkt. Dat betekende dat onmiddellijk van start kon worden gegaan met de digitalisering ervan. Onder meer dankzij het inzetten van twee jobstudenten is dit dossier bijna verwerkt. De kwaliteitscontrole, door personeel van het SOMA, is nog volop aan de gang. Verwacht mag worden dat die in januari 2010

Réalisations globales du projet

Er dient een onderscheid gemaakt te worden tussen de realisaties van het project en de realisaties van het project die momenteel nog niet gevaloriseerd kunnen worden gezien de reeds hoger genoemde problemen met de Regie der Gebouwen voor de inrichting van de DigiStore.

- Metadata / Toegangen (on-line opzoekbaar)

Beschrijving van

- o 20.956 archiefblokken
- o 26.603 archiefvormers
- o 5.583 inventarissen on-line

Analyses van documenten (in samenwerking met vrijwilligers)

- o 3.278.322 akten
- o 9.644.717 namen

- Alle parochieregisters die in het kader van het project Digi03 geselecteerd werden voor digitalisering werden beschreven (inhoudelijke metadata).
 - Tijdens de voorbereiding van de parochieregisters werd eveneens een overzicht opgemaakt van de ontbrekende registers. Een aantal hiervan zijn nog bewaard in de gemeente of de parochie. Er werd eveneens een vergelijking gemaakt tussen de originele registers en de dubbels en/of tafels. Aan de hand van dit onderzoek wordt contact opgenomen met de diverse *bezitters* van registers en worden overeenkomsten afgesloten om de verzameling te vervolledigen. Dit is een zeer belangrijke aanvulling voor de onderzoekers.
 - De registers van de Burgerlijke stand werden beschreven voor zover als de registers reeds in digitale vorm aangeleverd werden door de GSU na digitalisering van de microfilms. Momenteel werden reeds 126 leveringen met een gemiddelde van 120.000 à 150.000 gedigitaliseerde bladzijden beschreven (in totaal +- 17 miljoen bladzijden). Indien de aanvoer door de GSU constant blijft zal het project in 2012 afgerond zijn.
 - De registers van de Burgerlijke Stand voor de periode 1890-1910 die nog niet microverfilmd werden, werd eveneens een overeenkomst afgesloten met de GSU. Deze gedigitaliseerde documenten kunnen pas verwerkt worden zodra de nodige storagecapaciteit beschikbaar is.
- Gedigitaliseerde documenten

Via de digitale leeszaal zijn momenteel 1,9 miljoen gedigitaliseerde bladzijden archief raadpleegbaar. Zodra de DigiStore beschikbaar is, kan gestart worden met de verplaatsing van de bestanden naar de opslagcapaciteit. Gezien de hoeveelheid hebben we berekend dat alleen het kopiëren van de beschikbare bestanden die momenteel op diverse hard drives (en gedeeltelijk op andere externe media) bewaard zijn meer dan 8 maanden 24/24 in beslag zullen nemen.

We gaan er echter van uit dat we binnen de 6 maand na de ingebruikname van de DigiStore de kaap van de 20 miljoen bladzijden on-line gekoppeld aan de reeds beschikbare metadata kunnen halen.

Gezien het ontbreken van de nodige storagecapaciteit werd de digitalisering door externe bedrijven in het kader van de raamovereenkomst van september 2008 verstraagd. We beschikten immers niet over de nodige infrastructuur om bij ontvangst van de gedigitaliseerde documenten de nodige controles uit te voeren en de verdere stappen in het werkproces te vervolledigen. De laatste uitbestedingsprojecten werden recent gelanceerd zodat halfweg 2010 alle parochieregisters zoals vooropgesteld in het initiële project gedigitaliseerd zullen zijn. Voor wat betreft de registers van de Burgerlijke stand wordt het tempo gevolgd zoals vastgelegd in de overeenkomst met de GSU.

La digitalisation de documents concernés par le projet n'est pas une OCRisation (c'est-à-dire une conversion en texte digital exploitable) mais consiste à numériser un document sous forme d'image TIFF, d'y adjoindre des métadonnées situant le contexte du document et d'en permettre un accès Internet.

In de marge van het DIGIT-project werd een test gedaan met OCR op vroeg 20ste eeuwse documenten. In een eerste fase zullen 18.000 full-text doorzoekbaar zijn. In de loop van 2010 zullen twee nieuwe projecten respectievelijk van +- 30.000 bladzijden en 45.000 bladzijden on-line geplaatst worden.

Au final, il est clair que le projet ne couvre qu'une infime partie du patrimoine archivistique des partenaires à ce projet. Aux Archives tout particulièrement, seul 1.5% du patrimoine fut concerné. Bien qu'une digitalisation exhaustive au sein de cette institution soit difficile à envisager, le projet fut loin de permettre d'atteindre le seuil critique pour le grand public et les communautés spécifiques d'utilisateurs. Ceci dit, il s'est accompagné de modifications en profondeur des mentalités, l'infrastructure mise en place fut d'emblée conçue pour accueillir des données de nature diverse, le personnel a eu l'occasion de se former en interne et à intégrer la préservation digitale dans ses préoccupations quotidiennes.

Ce projet a pour objectif d'informatiser des collections au sein de deux ESF : les Musées royaux des Beaux-Arts de Belgique (MRBAB) et les Musées royaux d'Art et d'Histoire (MRAH).

Un partenariat est développé entre les deux institutions afin d'étudier la façon dont la méthodologie développée par les MRBAB dans le cadre de la réalisation de la base de données FABRITIUS, peut profiter aux MRAH pour l'informatisation de leurs propres collections.

Musées royaux des Beaux-Arts de Belgique

Introduction

Mené entre mars 2000 et mars 2004, le projet de recherche « Développement d'un format pour une base de données bilingue de la collection de sculptures du XIXe siècle, Musées royaux des Beaux-Arts de Belgique à Bruxelles », a débouché sur la création d'un catalogue informatisé accessible via l'Internet, et la mise en ligne au printemps 2003 de la collection de sculptures du XIXème siècle. Cette base de données est dénommée FABRITIUS, pour Fine Arts BRussels InTernet and Intranet USers.

Objectifs

Le projet **DI/04** « Bases de données d'objets de collections artistiques et historiques » comporte deux volets distincts :

- poursuite de l'informatisation des collections des MRBAB (DI/04/01) ;
- développement d'un partenariat avec les Musées royaux d'Art et d'Histoire, afin d'étudier la façon dont la méthodologie développée par les MRBAB dans le cadre de la réalisation de FABRITIUS peut profiter aux MRAH pour l'informatisation de leurs propres collections (DI/04/02).

Compte tenu de l'importance prise par la numérisation des collections pour leur gestion, et le développement de divers projets européens liés à cette problématique, un troisième volet, « projets connexes », a été ajouté.

Réalisations

Le processus d'informatisation des collections des MRBAB se compose de deux parties, complémentaires :

- gestion et développement du catalogue en ligne FABRITIUS, étendu à toutes les collections des MRBAB (www.opac-fabritius.be) ;
- prises de vues numériques très haute définition des œuvres, intégrées aux notices de ce catalogue en ligne.

1. Le catalogue informatisé

Le choix des collections prioritaires pour la poursuite de l'informatisation (2005-2008) repose sur plusieurs critères :

- accès en ligne aux chefs-d'œuvre des MRBAB et à toutes les œuvres exposées ;
- accès en ligne aux œuvres présentées dans les expositions temporaires (expositions Alechinsky, Rubens, Spilliaert, etc) ;
- mise en valeur des ensembles significatifs (toutes les œuvres de Magritte, de Khnopff, de Rubens, de Paul Delvaux, etc) ;
- mise en valeur des œuvres rarement visibles du public, telles que les œuvres sur papier, en particulier du département d'art ancien ;
- mise en valeur de nouvelles acquisitions.

Le tableau ci-dessous fournit un aperçu de l'état de l'informatisation des collections des MRBAB fin 2009, avec 7.932 notices bilingues FR/NL disponibles pour le public sur l'Internet et, dans une version plus complète et avec des images haute définition, sur l'intranet de l'institution :

	Collection MRBAB	Dans FABRITIUS	%
Peintures	5719	4035	71%
Sculptures (dont films et tapisseries moderne)	2245	1811	81%
Oeuvres sur papier (dont photos)	12556	2058	16%
Affiches	945	18	2%
Tapisserie (ancien)	12	10	83%

2. Photographies numériques

Le plan de numérisation des collections a permis de développer une méthodologie de prises de vues numériques très haute définition, avec corrections chromatiques et impression sur imprimante calibrée. Un appel d'offre a été lancé, conditionné par un cahier des charges très précis quant au type et à la qualité du matériel utilisé. Il a été conclu par l'établissement d'un contrat avec un photographe spécialisé, et l'organisation systématique de campagnes de prises de vues – en fonction des demandes, ou de l'actualité des MRBAB (réalisation d'un catalogue pour une exposition, projet de recherche, etc). Plus de 1.100 photographies numériques haute définition ont été réalisées dans ce cadre, et introduites dans le catalogue en ligne ; actuellement, plus de 70% des notices du catalogue en ligne sont accompagnées d'une illustration au moins. Il s'agit d'ektachromes ou de clichés ACL numérisés, ou d'images numériques haute définition (1.707 images de ce type, soit 31% de l'ensemble des images en ligne) ; un travail rétroactif est en cours afin de parvenir à 100% de notices illustrées.

3. Partenariat avec les MRAH

Le partenariat a permis aux MRAH de se doter d'un outil de gestion informatisée des collections, basé sur un modèle correspondant aux besoins spécifiques de l'institution (voir rapport MRAH).

4. Projets connexes

La cellule chargée aux MRBAB du projet DI/04 participe à différents projets connexes, qui ont en commun de développer des applications au départ de l'expertise acquise dans le domaine de la numérisation des collections :

- conception d'un système de gestion des déplacement des œuvres à l'intérieur de l'institution (salles, réserves, expositions temporaires) ou à l'extérieur (dépôts, prêts) ;
- développement de fonctionnalités pour le grand public, telles que les « visites guidées » virtuelles (chaque exposition temporaire fait l'objet d'une présentation dans FABRITIUS ; ces visites sont archivées et restent accessibles en ligne) ;
- conception et gestion d'un catalogue informatisé spécifique afin de valoriser, auprès du grand public et de la communauté des chercheurs, les résultats du projet de recherche « Rubens » ;
- participation au développement de la base de données en ligne du projet de recherche « Octave Maus » ;
- participation au développement du Musée Magritte en ligne : développement, en collaboration avec la société Ineo, de nouveaux outils de recherche multicritères sur plusieurs bases de données disponibles en ligne (catalogue FABRITIUS, catalogue de la bibliothèque, archives de l'art contemporain en Belgique et base de données multimédia Mediaview) ;
- partenariat avec la société Lumière Technology (Paris) pour la réalisation de mesures multispectrales d'œuvres des MRBAB, dans le cadre de recherches sur les applications de l'imagerie numérique appliquée à la conservation ;
- partenariat avec la VUB (laboratoire Etro) dans le cadre de l'application, à des fins de conservation, de schémas mathématiques à des images numériques haute définition de tableaux du XIXème siècle ;
- suivi de la problématique liée au droit d'auteur (images en ligne, publications) ;
- suivi des projets européens de numérisation du patrimoine (Europeana, Athena).

Bilan

Le premier objectif du plan de numérisation des MRBAB, qui était d'étendre l'accès en ligne à un ensemble représentatif de la richesse de ses collections, a été atteint de manière très satisfaisante. S'il reste beaucoup à faire pour que la totalité des collections des MRBAB soient accessibles en ligne, l'on constate que la quasi-totalité des collections de peintures et de sculptures le sont déjà.

Les œuvres sur papier (qui représentent près de 60% de l'ensemble des collections) nécessitent davantage de temps et de recherches (spécifiées typologiques, vocabulaire étendu, etc) ; d'une manière générale, la création de nouvelles notices dans la base de données en ligne s'accompagne de recherches complémentaires sur les œuvres elles-mêmes (iconographie, typologie, historique des expositions, état de conservation), menées par des historiens de l'art : il ne s'agit pas d'un simple encodage, mais bien d'un travail sur plusieurs sources (dossiers des œuvres, catalogues, rapports de conservation, etc).

Il faut souligner que, maintenant qu'une « masse critique » d'œuvres est accessible en ligne, FABRITIUS est devenu un véritable outil tant pour le personnel des MRBAB (scientifiques, gardes de collections, régisseur, bénévoles des Amis, ...) que pour les visiteurs extérieurs (environ 3.000 visiteurs différents par mois) ; le revers de la médaille est que l'équipe chargée de la numérisation des collections consacre davantage de temps à la gestion et au développement de la base de données, qu'à son seul accroissement (gestion des déplacements des œuvres, mise à jour de la liste des expositions, développement de nouvelles fonctionnalités, collaboration avec d'autres projets de recherche afin de les rendre accessibles en ligne, réponses aux demandes de renseignements, etc) : toutes ces tâches sont directement liées à la numérisation des collections, mais ralentissent le processus de création de nouvelles notices.

Concernant les images, les MRBAB ont maintenant développé une méthodologie de numérisation qui permet l'obtention de fichiers de haute résolution, pouvant servir tant à des publications qu'à des rapports de conservation, et sont également disponibles en basse définition sur l'Internet. Le nombre de photographies ainsi réalisées est également très satisfaisant, surtout si l'on considère la complexité d'organisation d'une journée de prises de vue impliquant différents départements, le déplacement des œuvres, et la disponibilité de plusieurs membres du personnel.

Le second objectif a également été atteint, dans la mesure où les Musées royaux d'Art et d'Histoire disposent également aujourd'hui d'une base de données pour la gestion de leur collection (voir rapports MRAH).

Musées royaux d'Art et d'Histoire

Het project InfoCol is de vertaling van het digitaliseringbeleid van de POD Wetenschapsbeleid in de KMKG inzake de informatisering en digitalisering van de collecties van de KMKG. Er is een pilootproject opgestart binnen de KMKG waarbij 7 van de 9 oorspronkelijk geselecteerde collecties (Nabije oosten, Iran, Egypte, Oceanië, Rijtuigen, Siret, Fotografie & film, Afrikaanse muziekinstrumenten en Europese blaasinstrumenten) uit de diverse departementen worden geanalyseerd en met de hulp van de firma Zetcom worden vertaald naar een gebruiksvriendelijke databank MuseumPlus. Dit project wordt nauw gevolgd door de stuurgroep InfoCol die bestaat uit vertegenwoordigers van alle wetenschappelijk departementen en informatici.

Sinds september 2008 werken conservators en hun medewerkers samen aan deze gemeenschappelijke databank waarin vandaag al **30.000 notities** werden ontsloten. Een platform kwam tot stand waarin wetenschappers en informatici bereid waren samen te werken aan één gemeenschappelijk project. Het aantal personen dat werd ingeschakeld op het niveau van het encoderen werd uitgebreid tot 16 gebruikers waaronder conservators, administratief personeel, restaurateurs en studenten.

Gecentraliseerde (thesaurus)lijsten werden samengesteld binnen de instelling voor materiaal/techniek, objectnaam en geografische referentie, met vertalingen en procedures om deze verder uit te breiden.

Beeldmateriaal werd met eigen middelen gedigitaliseerd, centraal en op een gestructureerde wijze bewaard op de multimediaserver.

Er werd heel veel gewerkt aan het signaleren en documenteren van bugs in Museumplus (wachtijden bij het linken van termen uit thesauri) en aan het uittesten van nieuwe functionaliteiten in de testversie.

Mappings (transfers van gegevens tussen twee data modellen) werden gemaakt om de oude thesauruslijst Materiaal/techniek te vervangen door de nieuwe.

De context van bepaalde metadata (gegevens die de karakteristieken van bepaalde gegevens beschrijven, data over data) zoals bij 'objectnaam', 'titel', 'beschrijving', 'afmetingen', 'materiaal en techniek', 'geografische referentie' werd geanalyseerd en naar de gebruikers toe vastgelegd in documenten die via de helpfunctie van Museumplus kunnen worden geraadpleegd.

Conventies in verband met dateringen (ISO-standaarden) werden opgesteld en via het intranet verspreid.

Er werd gewerkt aan de integratie van een aantal catalogi in een website in samenwerking met de firma Zetcom, zodanig dat een deel van de gedigitaliseerde data binnen de instelling voor iedereen toegankelijk wordt en later ook voor het publiek.

Er werd veel ervaring opgedaan betreffende organisatie, best practising, procedures, technologie, thesauri, metadata, datastructuren en exportformaten.

Het project infocol heeft ook een samenwerking tussen de museumdiensten tot stand gebracht die er voorheen niet was (conservators, informaticadienst, cel digitalisering, fototheek, archieven). Met de dienst fototheek werden er afspraken gemaakt waardoor administrators voortaan toegang krijgen tot deze server.

Door deelname aan Europese projecten zoals Europeana (de Europese digitale bibliotheek), Athena, Mimo..., zullen ook delen van onze collecties worden geïndexeerd in Europese databanken en zo gekoppeld worden aan digitale collecties van andere cultuurinstellingen. ATHENA staat voor 'Access to cultural heritage networks across Europe' en is één van de voorbereidende projecten voor Europeana. Het richt zich op het voorbereiden van de koppeling van digitale collecties aan Europeana. Bevindingen uit de praktijk worden hier verwerkt tot tools en richtlijnen rond onderwerpen als meertaligheid en semantiek, metadata en thesauri, rechtenproblematiek. De KMKG leiden samen met het Britse Collections Trust het werkpakket rond metadata en coördineert de federale contentpartners.

Het inzicht dat onderzoek naar (inter)nationaal toegepaste standaarden mbt datastructuren en metadata onontbeerlijk is voor het slagen van ons project, is nu reëel en zal verder worden uitgewerkt.

Er staan nog veel nieuwe collecties te wachten om te worden gedigitaliseerd. De kostprijs hiervoor is hoog (licenties, infrastructuur, stockage, gekwalificeerd personeel, beheerders, wetenschappers, invoerders, vertalers, ICT helpdesk..) en het budget ontbreekt.

Anderzijds zijn er ook nog veel metadata binnen Museumplus zoals auteur (type, functie, adres, biografie, reproductierechten), periode en type datering, inscriptie... die niet in gebruik zijn en waarvoor nog analyse nodig is. Veel modules werden nog niet grondig benaderd zoals de module bibliografie, adresbeheer, multimedia. Er moet nog veel kennisoverdracht gebeuren ondermeer over het opstellen van lijsten en rapporten, het uitzuiveren van typelijsten, het exporteren van gegevens, het aanpassen van thesaurusstructuren, de meertaligheid en de problematiek van synoniemen. Dit zijn zaken die dringend moeten worden aangepakt en waarvoor middelen moeten worden vrijgemaakt.

Ce projet consiste à cataloguer les collections scientifiques de l'Institut royal des Sciences naturelles de Belgique (IRSNB). Les données des collections doivent au total représenter entre 35 et 37 millions d'objets.

Les collections de l'IRSNB, et notamment celles concernant la zoologie, font parties des plus importantes tant au niveau européen que mondial. Le nombre de *spécimens "type"* en collection est estimé à 110.000 exemplaires. La priorité est donc mise sur la numérisation du catalogue de ces exemplaires uniques (holotypes et paratypes, mais aussi allotype, cotype, genotype, lectotype, locatype, neotype, paralectotype, plastotype, plesiotype, syntype et topotype). Les données de base ainsi que les références aux descriptions originales seront encodées dans DARWIN, logiciel commun de gestion des collections (base de données développée par le service ICT de l'IRSNB) par du personnel engagé grâce au projet mais aussi par du personnel sur fonds propres et encadrés par des scientifiques, gestionnaires de collections. Fin octobre 2009, 21541 enregistrements ont été encodés dans DARWIN, représentant l'information concernant 52.271 objets « type » en collection (soit 47,5% de l'objectif initial).

Le projet vise également la numérisation d'environ 500.000 échantillons de « *matériel belge* ». Au début du projet, seulement une infime partie des collections "belges" (vertébrés récents, invertébrés, entomologie, paléontologie, minéraux et données géologiques) se trouvait sous forme informatisée. Le but de cette première phase était donc d'encoder dans DARWIN (à l'instar du matériel type) une sélection de spécimens "belges" en collection pour lesquels les données de base sont connues. Une fois encodées, ces données (tout comme celles relatives au matériel type) sont immédiatement disponibles via l'interface WEB de consultation de DARWIN (déjà disponible depuis le 21/12/2007): <http://www.naturalsciences.be/darwin>.

Au fur et à mesure de l'avancement du projet, même si la priorité a toujours été mise d'abord sur l'encodage des données relatives aux collections dites « belges » mais afin de ne pas ralentir le rythme de travail des encodeurs, la numérisation s'est également étendue à d'autres collections de spécimens présents dans nos collections et dont les données étaient également prêtes à l'encodage. Fin octobre, 278.284 records étaient présents dans DARWIN (soit 55,6% de l'objectif initial), représentant l'information concernant 1.935.882 objets en collection.

Parallèlement un effort a également été apporté à l'encodage des données relatives aux collections du département d'anthropologie et de préhistoire de l'Institut, ces dernières nécessitant une approche différente (besoins spécifiques) pour leur gestion. Pour ce faire, toute l'information est encodée dans une plateforme dédiée (« MARS »), développée à partir de solutions « open source ». Actuellement et grâce notamment à l'apport du projet DIGIT05, pour les collections d'anthropologie et de préhistoire:

- tous les numéros d'inventaire général de ce département ont été encodés à 100% dans MARS
- les ossements humains de la collection de SPY, Goyet et Ishango sont entièrement encodés
- tous les inventaires préexistants sous format papier et électronique ont été numérisés par plateau dans MARS
- prise de vue à haute résolution des plateaux en fonction des demandes d'études et d'inventaire + intégration dans MARS (« conservatoires virtuels »).

S'y ajoutent encore le catalogage de 500.000 *données de baguage*. Les données concernant le travail de baguage en Belgique sont centralisées depuis 1926 à l'IRSNB. Actuellement, seule la base de reprises qui comprend 450.000 données est complètement informatisée. Les autres données (plus de 15.000.000 d'enregistrements) sont pour la plus grande partie uniquement disponibles sur des fiches manuscrites. Il s'agira donc d'encoder un ensemble de données présélectionnées parmi ces 15.000.000 de données (sélection retenue: grandes espèces d'oiseaux - bagues larges) dans le logiciel Papageno utilisé depuis plusieurs années à l'IRSNB. Les données qui y sont encodées sont conformes au format EURING. Au 31 août, 548.268 fiches de baguage ont été encodées dans Papageno, dépassant ainsi l'objectif initial de presque 10%! Depuis le 1 septembre, le travail d'encodage pour le service de baguage a donc été arrêté pour permettre de concentrer toutes les forces vives à la poursuite des deux premiers objectifs du projet.

Le problème de la numérisation à l'IRSNB est incompréhensible si on ne comprend pas le but de la collection et si on ne mesure pas la quantité et la variété des objets concernés. Les objectifs fixés dans le cadre de ce projet ne représentent qu'une infime partie des collections de l'IRSNB pouvant faire l'objet d'un travail de numérisation et de digitalisation comme en témoignent les résultats de l'étude du bureau Van Dijk menée en 2002-2003 recensant à ce moment-là pour l'IRSNB pas moins de 46 collections classées suivant la nature des objets qui les composent (livres, périodiques, archives, données scientifiques sous forme "analogique" ou 3D, photos, plans, ...).

Pour mener à bien ce genre de projet de numérisation, il faut du temps et du personnel, rien n'ayant encore réussi à remplacer la saisie humaine des données des inventaires. En effet, les inventaires ont un but de recherche. Les règles ont changé, les normes de nomenclature aussi, il ne suffit donc pas de scanner les fiches d'inventaire avec un logiciel de reconnaissance de caractères, il faut transcrire les données existantes dans les formats reconnus internationalement par la communauté scientifique. La numérisation s'accompagne donc d'une vérification. Il faut donc des opérateurs ("encodeurs") formés et un encadrement scientifique suffisant sans oublier, pour ce qui concerne les aspects informatiques, le personnel ICT nécessaire au développement, à la gestion et au support de ce genre d'outil. C'est pourquoi dans le cadre du projet DIGIT05, la quasi-totalité des ressources financières a été affectée à l'engagement de personnel. En moyenne depuis 2007, un pool de 10 encodeurs a pu ainsi être mis en place. Ils ont pu s'attaquer de manière effective à la numérisation des collections sous l'encadrement d'un scientifique de l'Institut, gestionnaire de collections.

Projet DI/06 : Banque de données d'observations climatologiques manuscrites

Ce projet réunit l'Institut royal météorologique de Belgique (IRM) et l'Institut d'Aéronomie spatiale de Belgique (IASB). Dans le cadre de la Phase 1 du projet « Digitalisation », le choix des données à digitaliser en priorité a été fait en fonction de l'intérêt scientifique de données disponibles uniquement sous forme manuscrite dans les archives des deux Instituts. Pour l'IRM, il s'agit des données relevées dans le réseau climatologique belge depuis la fin du 19^e siècle. Pour l'IASB, il s'agit des mesures en ozone effectuées en 1965-1966 à la base Roi Baudouin en Antarctique.

Sur le plan technique, il a été décidé d'encoder manuellement les données afin d'en faciliter leur traitement ultérieur. L'encodage des données des deux Instituts est effectué en interne à l'IRM par du personnel contractuel engagé et formé à cet effet. Le travail se passe à l'IRM car le volume des données IRM à encoder représente 99% du volume d'encodage du projet.

Institut royal météorologique de Belgique

Le plan stratégique de l'IRM insiste sur le renforcement de la sécurité des personnes et des biens par la connaissance du temps, du climat et de la géophysique. Un tel objectif dépend de manière cruciale de l'existence de données d'observations fiables et continues. Dans la banque de données de l'Institut, les plus anciennes observations effectuées de manière régulière dans l'ensemble du pays remontent à la fin du 19^e siècle, au moment où le réseau climatologique belge a été mis en place. Pour la période avant 1950, la plupart de ces données uniques ne sont accessibles que sur papier ou diagrammes. A l'occasion de la première phase du plan de digitalisation, le choix de l'IRM s'est porté sur l'encodage systématique des principales données climatologiques belges depuis la fin du 19^e siècle. La valorisation de ces données permettra à terme de caractériser l'évolution spatio-temporelle des climats régionaux dans notre pays à l'échelle du siècle, contribuant ainsi, d'une part, à l'objectif stratégique mentionné plus haut et, d'autre part, à l'effort international actuel de recherche dans le domaine du climat.

Les données encodées concernent plus précisément les données journalières de précipitations et de températures extrêmes (températures maximale et minimale) relevées dans le réseau climatologique belge

depuis le début des années 1880. L'encodage couvre la période 1881-1950. Combinées aux données climatologiques plus récentes déjà existantes dans la banque de données Oracle de l'IRM, les données climatologiques encodées pourront par la suite être mises à la disposition des chercheurs pour l'étude de l'évolution du climat belge, à l'échelle régionale, sur une période de plus de 100 ans.

Le volume des données à encoder représente un total de 115.000 bulletins mensuels, soit environ 4% du total des feuilles dans les archives de l'IRM pouvant être digitalisées.

Pour les données climatologiques de l'IRM, des travaux préliminaires ont été nécessaires avant la phase d'encodage proprement dite : (i) une interface d'encodage spécifique a été développée par les informaticiens de l'Institut et (ii) le fichier historique des stations du réseau climatologique depuis 1881 a été réalisé et ses données introduites dans la banque de données Oracle. L'encodage proprement dit n'a pu commencer qu'au premier trimestre 2008. Au 31 décembre 2009, 65% des données sont encodées, dont toutes les mesures du 19^e siècle. La totalité des données sera encodée fin 2010.

Du point de vue scientifique, l'IRM prévoit de collaborer avec les chercheurs belges (ou étrangers) intéressés dans le domaine de l'analyse de la qualité et de l'homogénéité des séries encodées et dans celui des climats régionaux belges et de leur évolution. En 2010, des contacts seront pris avec les universités en vue de mettre en place des projets de collaboration pour l'analyse et l'exploitation des données encodées. L'IRM essaiera aussi d'exploiter ces données dans le cadre des projets ECSN (European Climate Support Network) d'Eumetnet.

Dans un deuxième temps (à partir de 2011), après sélection des séries de données dont la qualité aura été jugée suffisante, des séries de données climatologiques simples, déduites des séries encodées, seront mises à la disposition du public sur le site web de l'IRM, pour caractériser les climats régionaux dans notre pays, ainsi que leur évolution.

Institut d'Aéronomie spatiale de Belgique

Pour l'IASB, les données à encoder concernent les mesures journalières du contenu total en ozone effectuées en 1965-1966 à la base Roi Baudouin en Antarctique. Ces données originales permettront, après contrôle de qualité, de mieux caractériser le contenu en ozone à cet endroit avant la découverte du fameux trou d'ozone au dessus de l'Antarctique, au début des années 1970. Dans le cadre du présent projet, le volume des données de l'IASB à encoder est très faible par rapport au volume des données de l'IRM (environ 1%).

Pour les données d'ozone, l'encodage a été effectué dans des fichiers Word et Excel, après définition d'un lay-out standard et concerne plusieurs types de données : (i) une documentation explicitant les méthodes de mesures, (ii) les résumés des mesures journalières et (iii) l'ensemble des feuilles de mesures journalières complètes réalisées sur la période 1965-1966.

L'encodage de ces données a été terminé en août 2008 et elles ont été transmises à l'IASB pour un examen préliminaire de leur qualité. En novembre 2009, une page web a été mise en ligne par l'IASB pour présenter à un large public le résultat du travail d'encodage, ainsi que le contexte général des mesures d'ozone dans l'Antarctique. Cette initiative permet la mise à disposition des données aux scientifiques intéressés via un accès électronique simple. On se réfèrera au site <http://ozonehistory.aeronomie.be/>.

Brève description du déroulement global du projet

Dans ce projet de numérisation des supports photographiques originaux (plaques de verre, films négatifs et positifs) quatre ESF sont impliqués. L'Observatoire royal de Belgique (ORB) et l'Institut d'Aéronomie spatiale de Belgique (IASB) possèdent 25.000 plaques astrophotographiques, le Musée royal d'Afrique centrale (MRAC) 200.000 plaques de verre de photos aériennes d'Afrique centrale (partiellement en dépôt de l'Institut géographique National – IGN) et l'Institut royal du patrimoine artistique (IRPA) 900.000 clichés de photos d'objets d'art, dont une grande partie sont des plaques en verre. Un cinquième institut fédéral, l'Institut géographique National – IGN possède également des clichés de photos aériennes sur film et sur verre. Partenaire dans le projet précurseur D4A ; ce dernier institut est associé au projet DI/00/07 indirectement via un partenariat avec le MRAC ; toute activité de scanning à très haute résolution pour le compte de l'IGN se faisant 'sur le crédit' du MRAC.

Le scannage en haute résolution de ces collections apporte un matériel de haute valeur scientifique. L'élément commun réside dans les images dont le support est une plaque de verre ou un film photographique de grande dimension, qui sont scannées dans un scanner spécialement conçu pour atteindre et reproduire de très hautes précisions géométriques et radiométriques.

De par différents obstacles rencontrés durant la construction de son infrastructure d'accueil par la Régie des Bâtiments, l'installation du scanner à très haute résolution DAMIAN (Digital Access to Metric Images Archives Network) s'est trouvée retardée de quelques années. Construit déjà en 2003-2004 aux EU, il a finalement été installé fin 2007 à l'ORB, et est entré en phase test opérationnel en septembre 2008. Durant cette période quelques centaines d'images aériennes de l'IGN et du MRAC nécessitant une très haute résolution ont pu être digitalisées ; elles ont servi comme données d'essai durant toute la phase test du scanner. Depuis septembre 2009, le DAMIAN est opérationnel et une ligne de production automatique est en phase test (voir plus bas).

Les travaux de pré-scanning (« quick looks » ou format thumbnail) et de catalogage informatisé de la collection de photo aériennes du MRAC ont cependant bien avancé.

Résultats

De DAMIAN-scanner is uniek niet zozeer wegens de hoge resolutie (3630dpi) maar vooral wegens de zeer hoge geometrische en radiometrische precisie.

De DAMIAN-scanner is de snelste digitaliseermachine ter wereld (er bestaan actueel slechts 2 andere gelijkaardige toestellen: 1 in de VS, 1 in Duitsland) en tevens de enige uitgerust met een automatische platenwisselaar en een automatisch filmroltransport en een positioneer-nauwkeurigheid en herhaalbaarheid beter dan 50 nanometer.

Binnen DI/00/07 werd een synergie ontwikkeld, waarbij een beroep werd gedaan op de kennis aanwezig op het NGI, de US Naval Observatory (USNO, Washington DC) en AGFA-Gevaert (Mortsel), betreffende de technische vereisten voor de digitalisering en exploitatie van (lucht)foto's en astro-fotografische platen, alsook op de kennis opgebouwd in het voorafgaand D4A-pilootproject (I2/AE/103).

Observatoire royal de Belgique

L'ORB possède une collection de photos de 22.000 prises de vue directes du ciel et 3.200 prises de vue spectrales. La plupart sont sur plaques de verre, une minorité sur film.

La numérisation porte sur la constitution d'un catalogue digital des métadonnées concernant 21.530 prises de vue directes et 2.680 prises spectrales.

De catalogus bevat metadata betreffende 25.239 hemelbeelden en spectra (meestal meervoudig op een drager), gekoppeld aan 12.582 prescans en quick-looks.

De par différents obstacles rencontrés durant la construction de son infrastructure d'accueil par la Régie des Bâtiments, l'installation du scanner à très haute résolution DAMIAN s'est trouvée retardée de plusieurs années. Construit déjà en 2003-2004 aux EU, il a finalement été installé fin 2007 à l'ORB, et est devenu opérationnel pendant l'année 2009. Les anciennes images aériennes nécessitant une très haute résolution sont d'abord transférées sur des rouleaux de film polyester et digitalisées en mode automatique.

Werkzaamheden verwant aan het project

Internationale samenwerking tussen de KSB, het IMCCE van de Observatoire de Paris en de US Naval Observatory, Washington, DC voor de digitalisatie en de astrometrische gegevensextractie van fotografische platen van planetaire satellieten (Externe financiering door IMCCE en eigen middelen). Momenteel werden reeds een 250 tal glasplaten gedigitaliseerd.

Werkzaamheden gepland voor komend semester

Gezien de vertraging met de bouw van de DAMIAN scanner, zal de nadruk liggen op het maximaal uitvoeren van WP4 (Scannen objecten KMMA) en WP5 (Scannen objecten KIK).

Musée royal d'Afrique central

Au MRAC, environ 100.000 "quick looks" de photos ont été numérisés et enregistrés sur DVD et sur un NAS (network attached storage system) de 4Tb, acquis par le projet. Les travaux de scanning continuent jusqu'à ce jour. En parallèle, le catalogue informatisé, finalisé lors du projet précurseur D4A, a été mis en ligne et est consultable sur GNOSIS. (<http://www.gnosis.be>).

Durant la phase test du DAMIAN et depuis le mois de septembre, en coordination avec l'IGN qui possède les plaques de verre et films originaux des photos aériennes de la collection CSK (comité spécial du Katanga), à peu près 500 originaux ont été scannés à très haute résolution. Hormis des test effectués par l'IGN sur la qualité radiométrique et photogrammétrique des images scannées, ces premières scènes scannées ont permis d'évaluer la mise en œuvre d'une ligne de production la plus automatisée possible (faisant appel aux automatismes incorporés dans le DAMIAN). Ce scanning automatique nécessite cependant un prétraitement des originaux ayant pour but d'en faire une copie contact sur film graphique AGFA par un reproducteur SCANATRON (Agfa). Les rouleaux de film graphique peuvent contenir jusqu'à 100 originaux, et sont chargés dans le DAMIAN qui peut ainsi être opérationnel sans supervision durant plusieurs jours d'affilé. Le prétraitement se fait, pas l'intermédiaire des opérateurs du DAMIAN, chez AGFA à Mortsel. Pour des formats originaux hors gabarit du SCANATRON, un système de chargeurs pour 10 négatifs existe permettant le scanning automatique. Ce système nécessite cependant la présence d'un technicien-opérateur pour charger / décharger les chargeurs.

Les prises de vues originales (négatives) de la collection IGCB (institut géographique du Congo belge) ont été perdues, ou n'ont jamais été rapatriées en Belgique durant / après la période coloniale ; le MRAC possède néanmoins un jeu complet de premières copies contact sur diapositive. Nous avons lancé en collaboration avec l'IGN une série de tests pour évaluer si ce jeu de copies pourrait livrer les mêmes propriétés radiométriques et photogrammétriques que les originaux après scanning sur le DAMIAN. Les premiers résultats semblent positifs.

Bien que la ligne de production du DAMIAN soit opérationnelle, il reste néanmoins à développer une ligne de gestion de cette nouvelle archive digitale qui sera créée. Les moyens de stockage dans le DAMIAN sont limités (production d'une semaine maximale), ce qui implique que les instituts respectifs collaborant au

projet DI/07 doivent implémenter leurs propres systèmes d'archivage... ou que des moyens centralisés soient développés.

Problèmes rencontrés et solutions

De par différents obstacles rencontrés durant la construction de son infrastructure d'accueil par la Régie des Bâtiments, l'installation du scanner à très haute résolution DAMIAN s'est trouvée retardée de plusieurs années. Construit déjà en 2003-2004 aux EU, il a finalement été installé fin 2007 à l'ORB, et est devenu opérationnel que début 2010. A ce jour quelques centaines d'images aériennes (de l'IGN) nécessitant une très haute résolution ont pu être digitalisées ; elles ont servi comme données d'essai durant toute la phase test du scanner.

Les WP4 à WP7 n'ont jusqu'à ce jour pas pu être entamé, pour leur volet scannage à (très) haute résolution sur le DAMIAN.

200.000 quick looks nécessitent des téraoctets de stockage de préférence accessible on-line via des serveurs dédiés. Les scans de très haute résolution nécessitent une capacité de stockage offline de plusieurs ordres de magnitude plus important que les quick looks. Pour les deux collections, il s'en suit des problèmes d'archivage, stockage, sauvegarde et gestion de la collection digitale qui devront être résolus ; l'infrastructure et know-how manque actuellement au MRAC.

Activités prévues pour le prochain semestre

Continuer la digitalisation à très haute résolution des clichés originaux CSK et éventuellement des diapositives IGCB ; développement de l'outil de gestion informatisé de cette nouvelle collection digitale. Finalisation du catalogue on-line et connexion entre le catalogue, la collection de quick looks et celle des scans de très haute résolution (interconnexion de banques de données et de serveurs via un portail unique). Développement du know-how et de l'infrastructure pour gérer des collections digitales de centaines (milliers) de téraoctets. Surtout ce dernier point devrait être développé en synergie avec d'autres projets/instituts.

Gezien de vertraging met de bouw van de DAMIAN scanner, zal de nadruk liggen op het maximaal uitvoeren van WP4 (Scannen objecten KMMA).

Institut royal du Patrimoine artistique

A l'IRPA, 532 plaques de verre de 40 sur 40 cm ont été sur place contrôlées, nettoyées, scannées (photographie digitale haute résolution) et enregistrées dans un système d'archivage développé en interne : *artiPACS*. Une présentation de ce système est donnée fin octobre 2008 en Chypre pour la *14th International Conference on Virtual Systems and MultiMedia (VSMM)* (<http://www.vsmm08.org>). L'IRPA a aussi commencé avec la photographie numérique de la série des plaques en verre en format 30 x 40 cm (un peu plus de 1.000 clichés).

En parallèle, le catalogue se constitue en s'appuyant sur la base de données déjà existante de la *photothèque online*. A ce jour, 900.000 photos (donc pas les clichés de ces photos) ont déjà été scannées en résolution moyenne pour une reproduction 1/1, et environ 400.000 descriptions d'objets d'art liés à ces photos ont été introduites. Tout cela est ouvertement consultable pour le public.

Testen van verschillende formaat glasnegatieven. Controle en selectie van de prioritaire negatieven en positieven die moeten gescand worden, rekening houdend met de eventuele aantasting veroorzaakt door ouderdom en zuurbetrandende dozen die in het verleden werden gebruikt. Schoonmaken van de aangetaste negatieven ter voorbereiding van de scanning. Fotograferen van groot formaat negatieven (afwerking). Blijvend verder aanvullen van wetenschappelijke en administratieve data in de fototheek databank.

Problemen en oplossingen

Omdat de scanner pas min of meer operationeel werd in 2009 en er nog steeds een aantal technische problemen moeten worden opgelost in verband met het voorkomen van onvolkomenheden in de resulterende beeldbestanden (een aanpassing van bepaalde algoritmes is vereist) zitten we nog in een testfase wat betreft de specifieke reeks negatieven van het Lam Gods.

Werkzaamheden verwant aan het project

- Met de middelen van de Nationale Loterij is er industriële kwaliteitsapparatuur en software aangekocht voor de digitalisering van radiografieën zowel op rol als op platen. (Achat d'un scanner conçu uniquement pour le scannage des plaques et bandes radiographiques de formats différents. L'achat comprend également un écran en noir et blanc de haute qualité et un ordinateur adapté à capter, traiter et stocker provisoirement des images scannées. Le scanner servira à scanner toutes nos radiographies actuelles ainsi que nos archives : 11.000 radiographies en plaques de 30 x 40cm² et de 470 radiographiques de grand format comprenant plusieurs bandes (largeur de 35,5cm ; longueur jusqu'à 3,5m). Il est à noter que nous avons un besoin immédiat pour ce genre de scanner, et que nous avons gardé plusieurs travaux en attente d'une nouvelle machine.
- Tegelijkertijd is er intern ook door middel van hoge resolutie fotografie een aantal supergrote röntgenopnamen gefotografeerd met een Phase One camera (30 miljoen pixels).
- Het fotograferen van de glasnegatieven 30x40cm is voltooid.

Werkzaamheden gepland voor komend semester

- De verwerking van al de bijkomende digitale data op artiPACS (een eigen ontwikkeld digitaal archiveringssysteem met versiecontrole voor grote bestanden zoals hoge resolutie scans) ligt gedeeltelijk stil wegens gebrek aan informaticapersoneel, maar kan hopelijk terug opgestart worden als de voorziene IT'er aangeworven is.
- De voorbereiding voor de overzetting van de fototheek- en bibliotheeksoftware is voorbij, de bestelling is gedaan bij de leveranciers (ook nieuwe MS server met MSSQL) en de eerste testen zullen in de loop van het 1^{ste} trimester 2010 beginnen.
- De nadruk ligt op het maximaal uitvoeren van WP10 (invoeren objectgegevens KIK), het in eigen huis fotograferen op hoge resolutie van resterende grote glasplaten, het omzetten op rol van glasnegatieven die in kritieke staat verkeren (Agfa)
- Ondertussen is er begonnen met het scannen van een reeks van 1.000 negatieven over het Lam Gods (zie problemen en oplossingen).

Er wordt gewacht op de resultaten van het scannen van de bovenvermelde negatieven. Er zal ook worden uitgekeken naar de wijze waarop de samenwerking tussen de verschillende FWI's in dit project verder moet verlopen, en wat er kan nog kan gerealiseerd worden op het gemeenschappelijke budget (2005-2008) rekening houdend met de verloren tijd.

Institut d'Aéronomie spatiale de Belgique

Les recherches de collection n'ont pas abouti à la constitution d'une collection suffisante permettant de valoriser la digitalisation et donc la participation de l'IASB-BIRA a constitué en un support à l'ORB-KSB et celui-ci n'a pas été requis pendant la période couverte par le présent rapport.

Ce projet concerne le scannage d'une part importante du patrimoine photographique du Musée royal d'Afrique centrale (MRAC) et de certaines collections des Musées royaux d'Art et d'Histoire (MRAH) et du Centre d'Etude et de Documentation Guerre et Société contemporaine (CEGES).

Musée royal d'Afrique central

Brève description du projet

Dans le cadre du plan mis en place pour la numérisation des collections des ESF et de la CRB, le MRAC Tervuren finalise actuellement la première phase du projet de digitalisation DI/00/08 qui porte sur une importante partie de son patrimoine photographique, soit : plusieurs collections de taille considérable qui ont été constituées et/ou acquises par le Musée au fil de son histoire. Dans les années à venir, le MRAC ambitionne effectivement de s'appuyer sur les technologies de l'information et de la communication, et d'utiliser les possibilités offertes par Internet, pour renforcer l'accès à ses collections et proposer certains services sous de nouvelles formes. Il est donc indispensable que les collections soient numérisées : les images comme le texte doivent passer d'un support analogue à un support numérique.

Dans cet ensemble unique, on distingue les collections de photographies historiques – impressionnant défilé de personnages, de lieux et de paysages, d'événements marquants et de scènes de la vie quotidienne – de celle des photographies d'objets ethnographiques.

Les collections de photographies historiques comprennent l'ancienne documentation visuelle du « Musée du Congo belge », soit plus de 37.000 vues imprimées à partir de clichés sur verre dont les plus anciens datent des années 1890. Environ 7.000 photographies proviennent du fonds de l'Office Colonial, organe du Ministère des Colonies qui fut créé au moment de l'annexion du Congo par la Belgique en 1908, en vue de la promotion de la colonie par le biais de conférences et d'expositions ainsi que dans la presse. Près de 14.000 photos de terrain issues de la photothèque de la section d'ethnographie du MRAC, qui continue d'ailleurs de s'enrichir actuellement, sont également incluses dans le projet, ainsi qu'une importante collection d'albums photo de l'Union Minière du Haut Katanga. Le groupe a effectivement documenté, dès 1910 et le début de l'exploitation de la première mine du Congo, tous les aspects de son développement.

Enfin, la collection Inforcongo fait également partie de cette première sélection : à travers plus de 19.000 photos, elle aborde des thèmes très divers qui illustrent les transformations socio-économiques du Congo au 20^e siècle. C'est le Centre d'information et de documentation du Congo belge et du Ruanda-Urundi (C.I.D.), et plus tard Inforcongo) qui, en 1950, fut à la base du développement de cette collection.

L'ensemble de ces collections recouvre plus d'un siècle de l'histoire de l'Afrique centrale et constitue une documentation tout à fait unique. De par les thèmes abordés, on y retrouve nombre d'aspects témoignant des profonds changements qu'a connus ce sous-continent depuis la fin du 19^e siècle. La mise à disposition de ce patrimoine exceptionnel, sous forme digitale, sera à l'origine d'une nouvelle impulsion donnée à la recherche scientifique, et ravivera probablement le souvenir de nombreux Africains et Européens.

En marge des collections de photographies historiques, le MRAC conserve une importante documentation visuelle et manuscrite sur les milliers d'objets ethnographiques qui peuplent ses salles et ses réserves. Il est prévu de digitaliser l'intégralité de ces informations également.

Aperçu de l'ensemble des collections traitées:

Code	Collection	Type de support	Section	Dimensions	Encodage	Quantité
AP	Collection de photographies anciennes	Photographies collées sur des fiches	AP1	17,5 x 20,5 cm	oui	35.814
			AP2	17,5 x 20,5 cm	oui	1.644
OC	Collection Office Colonial	Photographies collées sur des fiches	OC1	17,5 x 20,5 cm	oui	5.733
			OC2	17,5 x 20,5 cm	oui	1.264
EP	Collection de photographies de terrain à orientation ethnographique	Photographies collées sur des fiches	EP1	21,5 x 27,5 cm	non	2.000
			EP2	42,5 x 27 cm	partiel	13.982
EO	Collection d'objets ethnographiques	Fiches avec texte	EO1	21,5 x 27,5 cm of 21,1 x 29,6 cm	non	74.928
		Fiches avec texte et photos	EO2		non	29.054
UM	Collection Union Minière	Pages d'albums avec photographies		20 x 33,5 cm of 22,5 x 34 cm of 24 x 34 cm	non	7.248
INFOR	Collection Inforcongo	Photographies collées sur des fiches		27 x 42,5 of 27,5 x 43cm	oui	19.147
TOT						190.901

Résultats du projet à la fin de l'année 2009

La préparation de ces immenses fonds, plus de 190.000 pièces au total, a été intégralement réalisée par l'équipe du MRAC. A côté des phases d'inventorisation, nettoyage, récolement, comptage, répartition en lots et emballage, tous les aspects pratiques et surtout, techniques, ont été étudiés pour rédiger un premier cahier des charges complet et rigoureux, en vue de la sous-traitance de la numérisation proprement dite et d'une partie de l'encodage des données.

Het lastenboek werd gepubliceerd in juli 2008 en de volledige selectieprocedure werd afgerond in november. Een informatievergadering en de organisatie van een technische test maakten deel uit van deze procedure. Een gemotiveerd gunningsverslag werd opgesteld en zonder aanmerkingen goedgekeurd door de Inspecteur van Financiën. De opdracht werd toegekend aan een bedrijf uit Nederland, Pictura Imaginis, gespecialiseerd in de digitalisering van beeldarchieven. Zodoende kon de uitvoering van dit eerste lastenboek starten in december 2008.

Intussen loopt de scanning en de invoer van gegevens naar zijn einde toe. Een deel van de opdracht werd reeds correct door de onderaannemer opgeleverd. Deze beeldbestanden werden, na een interne kwaliteitscontrole, gekoppeld met de overeenkomstige records in de centrale collectiedatabank, en deze werden aangevuld met de gegevens die door de onderaannemer in XML werden aangeleverd. De uitvoering van dit eerste lastenboek wordt intern opgevolgd door een projectteam bestaande uit 9 personen die elk inkomende lot onderwerpen aan een strikte kwaliteitscontrole.

De uitvoering van de opdracht heeft een kleine vertraging opgelopen door de technische complexiteit van de gegevensinvoer, die grotendeels dient te vertrekken van de diverse handschriften die op de oude steekkaarten terug te vinden zijn. In het begin werd dit door de onderaannemer onderschat. De vooropgestelde kwalitatieve norm werd voor enkele loten aanvankelijk niet gehaald. Hierdoor moest een deel van het werk opnieuw worden uitgevoerd, wat tot een vertraging in de uitvoering heeft geleid. De

methodologische correcties toegepast door de onderaannemer werpen intussen hun vruchten af, en er wordt verwacht dat de volledige opdracht correct zal opgeleverd zijn eind januari 2010.

Intussen is een tweede lastenboek opgemaakt voor een selectie van fotografische collecties, die een meer delicatesse behandeling vereisen. Deze collecties bestaan uit originele glasnegatieven, naast negatieven uit kunststof. De voorbereiding van deze collecties gebeurt op een andere manier dan de collecties die deel uit maken van het eerste lastenboek, en de technische vereisten voor het scannen zijn ook verschillend. Om deze redenen werd beslist om hiervoor een apart, tweede lastenboek op te maken. Dit lastenboek wordt momenteel afgewerkt en zal, na vertaling, worden gepubliceerd begin 2010.

La mise à disposition de ce vaste patrimoine, au format digital, est prévue pour 2010. Cent ans après l'ouverture du musée, et à l'occasion du cinquantenaire de l'indépendance du Congo, un tout nouveau site web est actuellement en développement. L'accès à une partie des collections numérisées constituera un volet important de ce nouveau site. La valorisation de l'ensemble des archives visuelles du Musée par ce biais sera certainement à l'origine d'une nouvelle impulsion donnée à la recherche scientifique. Les collections photographiques du MRAC constituent un témoignage matériel incontournable pour toute étude historique ou anthropologique liée au culte, à la politique ou l'économie, ou à n'importe quel aspect de la vie quotidienne en Afrique, le site offrira dès lors à un public varié une source de documentation inestimable concernant une multitude d'aspects liés à l'histoire de l'Afrique centrale depuis la fin du 19^e siècle.

Musées royaux d'Art et d'Histoire

Historique

2005

Le projet de scannage des photos et autres documents a été initié en pôle avec le Musée d'Afrique centrale et le CEGES. Il était prévu que Hein Vanhee en assurant la coordination.

Deux grands axes sont définis par les MRAH :

- la photographie : les collections de photographies anciennes sont composées de photographies de techniques différentes. Si certaines œuvres sont de photographes renommés, d'autres sont de photographes anonymes. Leur intérêt réside dans leur aspect de témoignage de la vie quotidienne, de vues de sites urbains ou ruraux, ou du folklore belge. Une dernière partie des photos traitent de l'histoire de l'art et du patrimoine ;
- les documents graphiques sur papier : plans, dessins, affiches. Le fonds le plus important est celui des multiples dessins et plans de Paul Hankar. Ces documents sont parfois rehaussés d'aquarelles et donc sensibles à la lumière.

Le volume est estimé à 100.000 photos, 5.000 affiches, 3.000 dessins et plans.

Le budget souhaité : 2 personnes administratives 2,5 ans, soit 150 000 €

2006

19/01/2006 : réception de l'arrêté ministériel octroyant le budget pour 2005

07/09/2006 : réception des « Spécifications techniques » du projet ciblé, vu le budget alloué, le projet sera ciblé essentiellement sur le scannage des dessins du fonds Hankar.

Description globale du projet

Le fonds Hankar est constitué de 3.050 dessins, 460 pièces écrites (devis, rapports, échanges de lettres, comptes,...) et de 13 documents visuels (photographies anciennes de dessin ou d'édifices) sur papier fort ou sur calque. Ces documents sont réalisés au crayon et encre de chine, parfois rehaussé de lavis bruns, gris, verts.... Quelques dessins sont réalisés au fusain. Tous les dessins ont des dimensions différentes et varient du format de 690x1640mm au format A5.

La fragilité du support et le caractère éphémère des traits nécessitent une sauvegarde du document lui-même. De plus, la digitalisation permettra une consultation des documents sans avoir à les manipuler. Cette manipulation représente, en effet, la cause la plus importante de dégradation.

Ce fonds est en cours d'inventaire et de mise en conservation passive (rangement dans de nouvelles armoires par pochettes de carton neutre).

La première phase (2005-2006) porte sur la préparation des documents fragiles, une étude sur les firmes pouvant effectuer le scannage et la mise au point du matériel de transport des documents.

La deuxième phase (2007) devait porter sur l'établissement du cahier des charges, le choix de la firme et le début du scannage.

La troisième phase (2008) devait permettre le classement des originaux dans les réserves et l'intégration des digitalisations dans la banque de données de l'inventaire et dans le website du musée.

Comité de suivi aux MRAH

- Claudine Deltour-Levie, responsable du projet aux MRAH
- Werner Adriaenssens, responsable de la collection
- Bénédicte Rigoli, restauratrice spécialisée dans le papier

Travaux réalisés 2006-2008 : dessins Hankar

1. 2005-2006 : Préparation des documents

Une nouvelle réserve dédiée exclusivement aux documents « Papier » a été initiée : dégagement du local, équipement en armoires à plans. Plusieurs armoires ont été réservées à la collection des dessins Hankar.

En vue de la préparation à la digitalisation, les 3.000 documents ont été classés et rangés selon 4 facteurs :

- Documents déjà publiés et donc susceptibles d'être demandés
- Dessins rehaussés à l'aquarelle car très attractifs comme documents par rapport au public
- Dessins d'intérieur ne rentrant dans aucune des deux premières catégories
- Documents techniques.

Les documents sont classés dans des fardes avec leur numéro d'inventaire et regroupés par catégorie. Le scannage était prévu par ordre du classement.

2. 2007 : Préparation du cahier des charges pour appel d'offre auprès de sociétés pour effectuer le scannage

- Contacts et réunions avec les autres partenaires du projet
- Établissement des conditions spécifiques envoyées par Werner Adriaenssens à Hein Vanhee le 20 avril 2007

3. Bilan (fin 2008)

- Les 3.000 documents sont rangés, classés et prêts à être scannés ; l'inventaire existant a également été scanné
- Un descriptif pour l'établissement d'un cahier des charges a été établi

Ces points ont été effectués par le personnel du musée et sur les ressources propres pour les matériaux de rangement.

Nouvelle orientation : digitalisation des photographies historiques

Début 2009 : demande de changement d'affectation du budget

Demande accordée par l'avenant reçu le 03/08/2009

Motif du changement

Le projet de digitalisation des documents fragiles portait sur la numérisation du fonds Hankar. La préparation est déjà bien avancée : classement des documents, début de la sélection.

Mais un fait nouveau accapare toute l'attention de Werner Adriaenssens, responsable de la collection Hankar : la préparation des 6 salles d'exposition consacrée à l'Art Nouveau et à l'Art Déco. En effet, grâce au sponsoring de Inbev Baillet Latour (1.500.000 €), le projet de rénovation des salles est remis en route mais avec un délai d'échéance serré. Les trois premières salles doivent être achevées pour juin 2010. Dans ce contexte, Werner Adriaenssens a demandé de pouvoir être déchargé du projet de digitalisation.

Objectifs

Pour rester dans l'objectif fixé par le projet, le musée souhaite affecter la totalité du budget à l'engagement de personnel pour l'inventorisation et à la digitalisation des photographies historiques de la collection Photographie et Cinéma.

Il s'agit, notamment, de la collection des Pictorialistes achetées par le Gouvernement belge en 1896 afin de fonder ce qui aurait pu être le premier musée de la photographie. Ces photographies, régulièrement demandées en prêt, sont l'œuvre de photographes comme Misonne, Demachy, Alexandre ou encore Boissonnas. Si la collection des pictorialistes a fait l'objet d'une plaquette (« La collection des pictorialistes » par M.C. Claes, MRAH, 1998, 63 p.), le reste n'a pas encore été étudié. Il s'agit de photographies originales réalisées à partir de procédés aussi différents que l'albumine, les tirages au charbon ou, bien sûr, en procédé argentique.

Cette collection avait d'ailleurs été évoquée dès le début du projet, en même temps que les dessins Hankar, mais, étant donné l'enveloppe budgétaire disponible, le musée avait préféré ne retenir que les dessins Hankar.

La tâche de la personne engagée consistera à préparer le travail de digitalisation en phase avec le programme d'informatisation des collections en cours d'installation, Museumplus. Un premier temps sera consacré à l'analyse en tenant en compte des normes de Museumplus et du travail d'étude déjà réalisé par une spécialiste de la photographie afin de pouvoir entrer toutes les données dans le système commun. Un second temps sera consacré à la digitalisation des photos. Nous espérons pouvoir ainsi traiter environ 450 photographies anciennes.

Travaux réalisés (2008-2009)

1. Choix des photographies :

Définition des ensembles à traiter par la responsable de la collection, Claudine Deltour-Lévie en collaboration avec Florence Bertrand et Céline Quairiaux : le choix s'est porté sur l'ensemble des Pictorialistes, soit une trentaine de photographies car il s'agit du « noyau » de la collection et sur la collection des cartes de visites (soit près de 300 photos) représentatives des ateliers photographiques de la fin du XIX^e et du début du XX^e siècle.

Ces photographies avaient toutes fait l'objet d'un nettoyage effectué par Nathalie Minten.

2. Insertion dans le programme d'inventaire du musée : Museumplus

Analyse du programme de Museumplus avec l'équipe informatique pour l'intégration des l'inventaire des photographies historiques dans le nouveau programme.
Insertion des premières fiches en test.

3. Préparation des photographies par la restauratrice en photographie Céline Quairiaux :

Préparation de la suite des photographies à intégrer : choix sur les photos sur papier salé et à l'albumine, procédés fragiles à la lumière.

4. Début de la digitalisation des photographies :

L'ensemble des photos « cartes de visites » (350 n°d'inventaire), la collection des pictorialistes (32 photos), photographies « fragiles » (en cours) par Céline Quairiaux

5. Intégration dans la banque de données : en cours

Composition de l'équipe en charge du projet:

Pour la collection photographie

- Claudine Deltour-Levie, responsable de la collection (musée)
- Florence Bertrand, mi-temps, suivi scientifique et pratique (projet)
- Céline Quairiaux, restauratrice, un jour semaine (projet)

Encadrement informatique :

- Els Angenon (musée)
- Philippe Lambert (musée)
- Chris De Loof (musée)

Planning pour le développement du projet (2010-2014)

Le présent projet a permis d'initier un travail indispensable à la bonne conservation des œuvres tout en garantissant leur consultation. De plus, introduites dans une base de données, les photographies peuvent être accessibles aux recherches tant scientifiques que de vulgarisation. Le projet prévoit que les photographies une fois digitalisées, soient conditionnées d'une manière conforme aux normes de conservation passive. L'équipe en place estime qu'il faudra encore 5 ans pour digitaliser l'ensemble de la collection.

Centre d'Etude et de Documentation Guerre et Société contemporaine

In 2009 was het project afgerond. De gedigitaliseerde foto's zijn via Pallas consulteerbaar.

Projet DI/09 : Scannage du matériel audio-visuel

Deux sections de recherche du Musée royal de l'Afrique centrale se sont réunies autour d'un projet commun qui visait à digitaliser les collections d'archives sonores qu'elles possédaient. Ainsi, le service d'Ethnomusicologie, liée au Département Anthropologie culturelle, possède une importante audiothèque composée de musiques traditionnelles d'Afrique subsaharienne, enregistrées sur supports divers (supports sonores mécaniques, supports magnétiques, supports numériques) et dont les plus anciens datent du début du 20^{ème} siècle.

Parallèlement, la section Histoire de la période coloniale conserve plusieurs centaines d'archives sonores produites après la 2^{ème} Guerre mondiale par les services radiophoniques du Gouvernement colonial au Congo (reportages, interviews, chroniques, enregistrements musicaux, discours) et réalisées durant la période de la colonisation belge au Congo, au Rwanda et au Burundi.

L'objectif du projet était de digitaliser et de documenter trois mille heures d'enregistrements sonores dans deux bases de données, une spécifiquement adaptée à la discipline ethnomusicologique, l'autre à la discipline historique.

Ce projet s'inscrit dans la continuité du projet « Digitalisatie van het Ethnomusicologisch Klankarchief van het KMMA » (DEKKMMA) qui était déjà financé par le SPP Politique Scientifique Fédérale (période 01/10/2003-30/09/2007).

Sur base de ces acquis et avec l'assistance de deux techniciens, d'un collaborateur scientifique et d'un collaborateur administratif, les trois mille heures d'enregistrements sonores ont ainsi été digitalisées dans les délais impartis. Grâce aux gains de temps réalisés dans les étapes de digitalisation des archives sonores, des collections supplémentaires issues des collections de la section d'Histoire du temps présent, ont pu également être digitalisées.

En parallèle à la digitalisation proprement dite, un travail d'encodage des métadonnées liées a été effectué pour répondre à d'autres objectifs muséaux bien spécifiques comme la gestion de la documentation et la diffusion des connaissances vers un public de spécialistes et vers un plus large public.

Deux bases de données distinctes ont été produites dans cette optique. L'une est spécifique à l'Ethnomusicologie (DEKKMMA), réalisée dans le cadre du projet précédent et dont les résultats sont mis en ligne sur le site internet <http://music.africamuseum.be>; l'autre est spécifique à l'Histoire sous l'appellation *Base de données des Archives Sonores* et n'est momentanément qu'en utilisation intranet avant sa finalisation pour un accès internet.

Se retrouveront donc accessibles les informations suivantes : la description de la source sonore et sa documentation bibliographique en trois langues (français, néerlandais et anglais) ; l'insert de fragments sonores en MP3 ou un autre format ; la possibilité de commander en ligne certains fragments qui seront compilés sur CD. Cette dernière fonction doit encore être travaillée de manière à combiner de manière adéquate l'aspect commercial et éthique de cette entreprise. L'avantage est qu'il permet au personnel technique de pouvoir mieux s'adapter aux demandes extérieures en possédant, au départ, un stock d'enregistrements qui peuvent être copiés plus rapidement et facilement selon un système automatisé.

Le point de discussion, à ce stade, porte aussi sur l'importante question des droits d'auteurs, de reproduction et de diffusion de ces archives sonores et qui reste à régler au cas par cas, en fonction des divers modes de réalisation et d'acquisition des documents sonores, de leur producteur ou de leurs ayant droits. Les dernières étapes du projet, qui sont actuellement organisées, concernent la validation scientifique de la Base de données des Archives Sonores de la section Histoire de la période coloniale de même que sa valorisation par son accès internet.

La validation de la « Base de données des Archives Sonores » de la section Histoire de la période coloniale est en voie d'achèvement.

La valorisation des ses résultats (finalisation de la page de présentation du site, insert de fragments sonores en MP3 et la mise en accès internet) constituera la principale activité dans le courant de 2010. Dans le contexte des archives audio-visuelles (sonores exclues), il est aussi envisagé d'entreprendre la numérisation de certains films d'archives provenant de sociétés (Union minière du Haut-Katanga, Institut des Parcs nationaux du Congo Belge) et de particuliers et l'encodage de leurs métadonnées dans une base de données adaptées aux films d'archives.