

**DIGITALISERINGSPLAN VOOR HET CULTURELE- EN WETENSCHAPPELIJKE  
PATRIMONIUM VAN DE FEDERALE WETENSCHAPPELIJKE INSTELLINGEN  
EN VAN HET KONINKLIJK BELGISCH FILMARCHIEF**

**Fase 1**

EINDVERSLAG

**DIGITALE COPIES VAN DOCUMENTEN**

**DI/00/03**



**Promotoren**

Lucie Verachten (ARA-AGR)

Laurent Meese (KBIN-IRSNB)

Dirk Martin (SOMA-CEGES)

**Auteurs**

Lucie Verachten (ARA-AGR)

Cathy Emery (KBIN-IRSNB)

Laurent Meese (KBIN-IRSNB)

Dirk Martin (SOMA-CEGES)



Uitgegeven in 2012 door het Federaal Wetenschapsbeleid

Louizalaan 231

B-1050 Brussel

België

Tel: +32 (0)2 238 34 11 - Fax: +32 (0)2 230 59 12

<http://www.belspo.be>

Contactpersoon: *Anna Calderone*

Secretariaat: +32 (0)2 238 34 65

Noch het Federaal Wetenschapsbeleid, noch eenieder die handelt in de naam van het Federaal Wetenschapsbeleid is verantwoordelijk voor het gebruik dat van de volgende informatie zou worden gemaakt. De auteurs zijn verantwoordelijk voor de inhoud.

Niets uit deze uitgave mag worden verveelvoudigd, opgeslagen in een gegevensbestand of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën of enige andere manier zonder de aanduiding van de referentie.

Lucie Verachten, Cathy Emery, Laurent Meese en Dirk Martin, *Digitale copies van documenten*, Eindverslag, Federaal Wetenschapsbeleid (Digitaliseringsplan voor het culturele- en wetenschappelijke patrimonium van de Federale wetenschappelijke instellingen en van het Koninklijk Belgisch Filmarchief – Fase 1), Brussel, 2012, 24 blz.

## INHOUDSTAFEL

RESUME

SAMENVATTING

SUMMARY

|  |    |
|--|----|
| 1. INLEIDING.....  | 1  |
| 1.1 Context.....   | 1  |
| 1.2 Internationale samenhang.....                                  | 1  |
| 1.3 Doelstellingen .....   | 2  |
| 2. UITVOERING .....  | 4  |
| 2.1 Methodologie, technische keuzes, beroep op onderaanneming..... | 4  |
| 2.2 Partnerschap ontwikkeld met andere FWI's .....                 | 8  |
| 2.3 Voorgekomen problemen en aandachtspunten .....                 | 9  |
| 2.4 Middelen (personeel en budget) .....                           | 11 |
| 2.4.1 Ten laste van Belspo.....                                    | 11 |
| 2.4.2 Ten laste van de FWI's .....                                 | 13 |
| 3. REALISATIES .....   | 15 |
| 4. VERSPREIDING EN VALORISATIE .....                               | 19 |
| 5. BALANS EN PERSPECTIEVEN .....                                   | 22 |
| 6. REFERENTIES .....   | 24 |

## RESUME

La numérisation est un moyen convivial par excellence pour mettre les collections précieuses et uniques à disposition de groupes-cibles très hétérogènes, qu'il s'agisse du grand public (composé de groupes cibles nouveaux et préexistants) ou de chercheurs scientifiques. Il importe également que ces documents fragiles et de grande valeur puissent être rangés en toute sécurité dans les dépôts, sans risques de perte, vol ou dommages dus à la manipulation.

L'objectif visé est 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 bibliothèque scientifique de l'IRSNB possède l'une des plus grandes collections en sciences naturelles de l'Europe de l'Ouest. Dans cette précieuse collection, les ouvrages les plus anciens datent du 16<sup>e</sup> siècle. Une des collections particulières est celle de la bibliothèque Dautzenberg qui comprend 7956 ouvrages spécialisés et livres rares, ainsi qu'un peu plus de 200 publications de Philippe Dautzenberg lui-même, dans lesquelles il a décrit 1895 nouveaux taxons. Cette dernière catégorie a dû être numérisée en priorité pour des raisons de conservation et de préservation. Une deuxième catégorie d'ouvrages à numériser concernait les publications scientifiques épuisées de l'institution. Une troisième catégorie d'ouvrages sélectionnés était composée de livres précédemment traités dans le cadre d'un projet de désacidification de masse à grande échelle.

Dans le cadre du projet DI/00/03, le CEGES a opté pour les dossiers d'enquête judiciaire d'après-guerre de la Juridiction militaire belge. Cette juridiction a été non seulement responsable après la guerre de la poursuite des collaborateurs, mais aussi de la poursuite des criminels de guerre et des gros bonnets du régime d'occupation militaire allemand en Belgique. Les dossiers doivent être vus dans le contexte des procès qui ont été intentés partout en Europe après celui du tribunal militaire international de Nuremberg contre les responsables allemands au niveau local.

Vu l'étendue du patrimoine des Archives de l'État, une numérisation intégrale serait inconcevable. Il faut donc opérer des choix en fonction de la fréquence de la demande de certains documents, de la valeur et de la fragilité des documents, de leur importance historique, des opportunités de collaboration avec d'autres organismes etc.

La valorisation figure en tête des critères qui sont utilisés pour fixer les projets de numérisation des archives. En effet, la numérisation vise en tout premier lieu la mise à disposition des documents et ceux qui sont le plus souvent consultés reçoivent la priorité. Vu le poids des généalogistes dans le total des visiteurs des Archives de l'État, la numérisation des sources généalogiques se trouve en tête des priorités.

Cependant, bien d'autres documents ont également été numérisés, parfois en collaboration avec des universités ou dans le cadre de projets de recherche scientifiques. La numérisation est un moyen important pour pouvoir explorer des sources peu connues.

Les réalisations de la numérisation au sens strict du terme sont disponibles au public via le moteur de recherche "search.arch.be", où l'utilisateur trouvera un accès à divers types de métadonnées (descriptions générales des fonds d'archives conservés, descriptions des unités archivistiques (inventaires classiques), des notices historiques détaillées relatives aux producteurs d'archives, des analyses détaillées des documents jusqu'au niveau de l'acte) ainsi qu'à des documents d'archives numérisés. Le moteur de recherche constitue le point de départ de toute recherche dans les archives, tant pour les internautes que pour les visiteurs de la salle de lecture.

La numérisation est un processus très conséquent pour les services d'archives. Les procédures de travail de tous les collaborateurs doivent être adaptées pour que les résultats soient directement exploitables dans l'environnement numérique. Dans le cadre du projet de numérisation, une importance toute particulière a été portée au développement d'un "back office" efficace où ont été prévus tous les aspects de la numérisation tels la préparation matérielle et intellectuelle, le contrôle de qualité ou le suivi de la disponibilité à long terme.

Avec les moyens du projet DI/00/03, en mobilisant des ressources propres et grâce à des financements externes (Loterie Nationale) et à la collaboration avec des tiers, il a été mis en place une infrastructure solide pour la création, la gestion, la conservation à long terme et la valorisation des contenus numériques. Les Archives de l'État ont investi dans du matériel (scanners, réseau, salle de lecture numérique,...), dans la formation du personnel (tant le personnel de l'institution que les collaborateurs du projet DI/00/03) et surtout dans des processus de travail sûrs et validés, avec l'utilisation de normes et de standards internationaux. Les nouvelles évolutions sont suivies, analysées et, si possible, appliquées immédiatement.

En acquérant des installations techniques (scanners, stockage) et lors de l'acquisition et de la programmation de logiciels, on a veillé à l'application de standards ouverts et aux possibilités d'une extension future. Les infrastructures des Archives de l'État sont conçues de façon à ce qu'elles puissent être développées et de telle sorte que tous les projets de numérisation à venir puissent être intégrés dans le même environnement. C'est pourquoi l'institution a opté pour l'uniformité des projets et pour leur intégration dans un ensemble plus large. Cette approche sert à rendre cet environnement gérable et elle permet aussi d'offrir aux utilisateurs un "point d'accès unique".

Mentionnons enfin comme point important la nécessité de la pérennité des résultats des projets de numérisation.

## SAMENVATTING

Digitalisering is het middel bij uitstek om waardevolle en unieke collecties op een gebruiksvriendelijke wijze ter beschikking te stellen van allerhande doelgroepen. Er wordt daarbij aandacht besteed zowel aan het grote publiek (bestaande en nieuwe doelgroepen) als aan wetenschappelijke onderzoekers. Belangrijk is eveneens dat de waardevolle en kwetsbare documenten veilig in het magazijn kunnen opgeborgen worden zonder gevaar voor verlies, diefstal en beschadiging door gebruik.

Het beoogde doel is de digitalisatie van een selectie van documenten binnen 3 FWI's: het Algemeen Rijksarchief en Rijksarchief in de Provinciën (ARA), het Koninklijk Belgisch Instituut voor Natuurwetenschappen (KBIN) en het Studie- en Documentatiecentrum Oorlog en Hedendaagse maatschappij (SOMA).

De wetenschappelijke bibliotheek van het KBIN beschikt over een van de grootste natuurwetenschappelijke collecties van West-Europa. In de waardevolle collectie dateren de oudste werken uit de 16de eeuw. Een van de bijzondere collecties is de Dautzenberg - Bibliotheek, deze bevat 7.956 gespecialiseerde werken en zeldzame boeken en ruim 200 publicaties van Philippe Dautzenberg zelf, waarin hij 1895 nieuwe taxa beschreef. Die laatste categorie diende om redenen van conservering en preservatie prioritair gedigitaliseerd te worden. Een tweede categorie van te digitaliseren werk betrof de uitgeputte wetenschappelijke publicaties van de instelling. Een derde categorie van geselecteerde werken bestond uit boeken die eerder in een grootschalig massaoontzuringproject behandeld werden.

In het kader van project DI/00/03 opteerde het SOMA/CEGES voor naoorlogse gerechtelijke onderzoeks dossiers van het Belgische krijgsgerecht. Dat was na de oorlog niet alleen verantwoordelijk voor de vervolging van collaborateurs, maar ook voor die van oorlogsmisdadigers en kopstukken van het Duitse militaire bezettingsbestuur in België. De dossiers moeten gezien worden in de context van de processen die na het internationale militaire tribunaal te Nürnberg overal in Europa gevoerd werden tegen de Duitse verantwoordelijken op lokaal vlak.

Rekening houdend met de grote omvang van de bestanden is het digitaliseren van het volledige patrimonium van het Rijksarchief echter geen optie. Er worden keuzes gemaakt op basis van de vraag naar bepaalde documenten, de kwetsbaarheid en de waarde van de documenten, het historisch belang van de documenten, de opportuniteiten voor samenwerking,...

Bij de identificatie van de prioritair digitaliseringprojecten staat valorisatie van archieven bovenaan. Digitalisering is immers in de eerste plaats gericht op beschikbaarstelling. De archieven die het vaakst geraadpleegd worden, staan hoog op de prioriteitenlijst. Gelet op het grote aandeel van de genealogen in het bezoekerspubliek van het Rijksarchief, staat "digitalisering van genealogische bronnen" dan ook bovenaan de prioriteitenlijst.

Maar ook tal van andere archieven werden al dan niet in samenwerking met universiteiten of wetenschappelijke onderzoeksprojecten gedigitaliseerd. Digitalisering is immers eveneens een belangrijk middel om ongekende bronnen beter te leren kennen.

De resultaten van de digitalisering in de strikte betekenis van het woord zijn voor het publiek raadpleegbaar via de zoekrobot "search.arch.be". Het betreft zowel de toegang tot de diverse types van metadata, gaande van de algemene beschrijvingen van de bewaarde archiefbestanden, over beschrijvingen van de archiefbestanddelen (klassieke inventarissen) en uitgebreide historische notities met betrekking tot allerhande archiefvormers tot gedetailleerde analyses van archiefdocumenten tot op het niveau van de akte alsook tot de gedigitaliseerde archiefdocumenten. De zoekrobot is voor de gebruiker zowel via internet als in de leeszaal het startpunt van het archiefonderzoek.

Digitalisering is een zeer ingrijpend proces voor de archiefdiensten. De werkprocessen van alle medewerkers worden aangepast opdat de resultaten onmiddellijk bruikbaar zouden zijn in de digitale omgeving. In het kader van het digitaliseringproject werd bijzonder veel aandacht besteed aan de uitbouw van een efficiënte *back-office* waarin alle aspecten van digitalisering materiële en intellectuele voorbereiding, kwaliteitscontrole, opvolging van langetermijnbeschikbaarheid,... vastgelegd zijn.

In het kader van het project DI/00/03 maar eveneens met eigen middelen en dank zij externe financiering (Nationale Loterij) en samenwerking met derden heeft men een solide infrastructuur opgezet voor de creatie, het beheer, de lange termijnbewaring en de valorisatie van digitale content. Bij het Rijksarchief werd er geïnvesteerd in materieel (scanners, netwerk, digitale leeszaal,...), in opleiding van personeel (eigen personeel en personeel in kader van DI/00/03) en vooral in beschreven en gevalideerde werkprocessen waarin het gebruik van internationale standaarden en normen centraal staat. Nieuwe ontwikkelingen worden opgevolgd, geanalyseerd en daar waar mogelijk onmiddellijk toegepast.

Bij de aanschaf van technische apparatuur (scanners, storage) maar ook en vooral bij de aanschaf of programmering van softwaretoepassingen wordt aandacht besteed aan de toepassing van "open standaarden" en de mogelijkheden voor verdere groei. De infrastructuur van het Rijksarchief is zo opgevat dat ze uitbreidbaar is en dat alle toekomstige digitaliseringprojecten geïntegreerd kunnen worden in de omgeving. Er werd dan ook geopteerd voor uniformiteit van de projecten en integratie in een groot geheel. Deze aanpak is belangrijk voor de beheersbaarheid van de omgeving maar heeft ook zijn nut bewezen als "unieke toegang" voor de gebruikers.

Een belangrijk aandachtspunt tot slot is de duurzaamheid van de resultaten van de digitaliseringprojecten.

## SUMMARY

Digitisation is admirably suited to put the precious and unique collections at the disposal of a wide variety of target groups, belonging to the large public (existing and new target groups) as well as to the scientific community. It is also important that the valuable and fragile documents are safely stored without the risk of loss, theft or damages due to their manipulation.

The aim is to digitise a selection of documents in 3 FSI's: the National Archives and the State Archives in the Provinces (ARA), the Royal Belgian Institute of Natural Sciences (RBINS) and the Centre for Historical Research on War and Contemporary Society (SOMA / CEGES).

The KBIN's (Royal Belgian Institute of Natural Sciences) scientific library contains one of the largest scientific collections in Western Europe. The earliest works in this valuable collection date back to the 16th century. One of the special collections is the Dautzenberg Library, which contains 7.956 specialised works and rare books and over 200 of Philippe Dautzenberg's own publications, in which he described 1.895 new taxa. For reasons of conservation and preservation the latter category was prioritised for digitisation. The second digitisation category encompasses the institute's exhausted scientific publications. The third category of selected works consists of books previously treated in a large-scale mass deacidification project.

SOMA / CEGES opted, in the context of project DI/00/03, for the post-war judicial investigation files of the Belgian military courts. After the war these courts tried not only collaborators, but war criminals and leaders of the German occupation administration in Belgium. The files should be viewed in the European context of local proceedings against the German leaders after the international military trials in Nuremberg.

Given the size of the patrimony of the State Archives a full digitisation would be inconceivable, so choices have to be made on the basis of the frequency of the demand for the documents, their value, fragility or historic importance and the opportunities for collaboration with other institutions.

Valorisation of the archives is one of the top criteria in defining the projects. Indeed, digitisation is in the first place about availability. The archival documents that are often asked for figure high on the priority list. As a large part of the State Archives' visitors are genealogists, the digitisation of genealogical sources is top priority.

However, many other archival documents have also been digitised, sometimes in collaboration with universities or in the context of scientific research projects. Digitisation is in fact an important means to discover as yet unknown sources.

The results of the digitisation projects in the narrow sense of the word can be consulted by the public by means of the search engine "search.arch.be", which gives access to different types of metadata (general descriptions of the records conserved, descriptions of archival units (traditional inventories), detailed historical notes concerning the archives' producers and analyses of archival documents up to the level of the acts) as well as digitised documents. The search engine is the starting point for all archival research, in the State Archives' reading room as well as on its website.

For the archival services, digitisation is a far-reaching process. The working processes of all personnel are being adapted to ensure that the results are immediately usable in a digital environment. In the context of the digitisation process, special attention has been given to the development of an efficient back office where all aspects (material and intellectual preparation, quality control, follow-up of long term availability,...) have been foreseen.

With the DI/00/03 project resources, but also with own resources and external financing (National Lottery) and thanks to the cooperation with third parties, a solid infrastructure for the creation, the



management has been set up, the long term conservation and the valorisation of digital content. The State Archives have invested in apparatus (scanners, networks, digital reading room,...), personnel training (own personnel as well as DI/00/03-personnel) and especially in solid and advanced working processes where international standards are being used.

When acquiring technical equipment (scanners, storage) and when purchasing or programming software applications, open standards and the possibilities for further growth are at the centre of attention. New developments are being followed, analysed, and applied where possible.

The State Archives' infrastructure has been conceived in order to be extendible and to be able to integrate all future digitisation projects. This is why uniformity and integration into a whole are being opted for. This approach is important to assure the digital environment can be managed and, at the same time, it can offer the users a unique entry point.

Finally, attention should also be given to the durability of the results of the digitisation projects.

## 1. INLEIDING

### 1.1 Context

- Digitalisering is een belangrijk middel om de missie van de wetenschappelijke instellingen en meer bepaald van het Algemeen Rijksarchief en Rijksarchief in de Provinciën uit te voeren. De wetenschappelijke bibliotheek van het KBIN beschikt over een van de grootste natuurwetenschappelijke collecties van West-Europa. In de waardevolle collectie dateren de oudste werken uit de 16de eeuw. Het Rijksarchief verwerft en bewaart (na selectie) archieven van minstens 30 jaar oud van hoven en rechtbanken, openbare besturen en notarissen, alsook van de private sector en particulieren (bedrijven, politici, verenigingen en genootschappen, notarissen, families, enz. die een belangrijke rol hebben gespeeld in het maatschappelijk leven). Het ziet erop toe dat overheidsarchieven volgens de archivistische normen worden overgedragen. Het beschikbaar stellen van deze archiefbescheiden aan het publiek, met bescherming van de privacy van bepaalde gegevens, behoort tot de belangrijkste hoofdpodochten van de instelling. Het Rijksarchief stelt in zijn 18 studiezalen infrastructuur ter beschikking van een ruim en gevarieerd publiek. De rechtstreekse dienstverlening via het internet (de digitale leeszaal) is één van de prioriteiten voor de instellingen.
- Eén van de hoofdtaken van het wetenschappelijk personeel bestaat in het toegankelijk maken van de immense hoeveelheid nauwelijks ontsloten archief die zich in de rijksarchieven bevindt. Dat gebeurt via de realisatie van een wetenschappelijk instrumentarium (weg- en zoekwijzers, archievenoverzichten en gidsen, inventarissen, institutionele studies) dat de onderzoeker moet toelaten de gewenste informatie vrij trefzeker en binnen redelijke tijd op te sporen. Ook deze worden in digitale vorm ter beschikking gesteld.
- "Digitalisering" is echter niet neutraal. Het is geen zuiver technische operatie maar het is een proces dat uitgevoerd dient te worden rekening houdend met de specifieke vereisten met betrekking tot de bewaring, de manipulatie, de ontsluiting, de valorisatie,... van archiefdocumenten. Het onderzoek naar de mogelijkheden, beperkingen, standaarden,... in verband met digitalisering sluit aan bij de missie van het Rijksarchief als kenniscentrum inzake archiefbeheer.
- "Digitalisering" is ook zeer belangrijk in het kader van samenwerking met andere instellingen (zowel in binnen- als buitenland). Het Rijksarchief sloot in de voorbije jaren overeenkomsten af voor digitalisering van archiefdocumenten bewaard in het Rijksarchief met bij voorbeeld het Joods Museum voor Deportatie en Verzet, SOMA, Holocaust museum (Washington), provincies, gemeenten, enz.

### 1.2 Internationale samenhang

- Digitalisering is een belangrijke opdracht voor culturele en wetenschappelijke instellingen. Vanuit "Europa" worden de culturele en wetenschappelijke instellingen aangespoord hun patrimonium te digitaliseren en via internet ter beschikking te stellen. In het rapport van het *Comité des sages. Reflection group on bringing Europe's cultural heritage online* wordt aan de instellingen een duidelijke opdracht gegeven.
  - o *"For centuries, libraries, archives and museums from across Europe have been the custodians of our rich and diverse cultural heritage. They have preserved and provided access to the testimonies of knowledge, beauty and imagination. The new information technologies have created unbelievable opportunities to make this common heritage more accessible for all. Culture is following the digital path and "memory institutions" are adapting the way in which they communicate with their public. Digitisation breathes new life into material from the past, and turns it into a formidable asset for the individual user and an important building*

*block of the digital economy. We are of the opinion that the public sector has the primary responsibility for making our cultural heritage accessible and preserving it for future generations. Digitising our cultural heritage is a gigantic task that requires large investments.*

- *We think that the benefits are worth the effort. These benefits are in the first place related to the wider access to and democratisation of culture and knowledge, as well as the benefits for the educational system - both schools and universities. Other major benefits lie in the economic sphere and concern the development of new technologies and services for digitisation, for digital preservation and for interacting in innovative ways with the cultural material. The digitised material can in itself be a driver of innovation and be at the basis of new services in sectors such as tourism and learning.*  
[\(\[http://ec.europa.eu/culture/documents/report\\\_comite\\\_des\\\_sages.pdf\]\(http://ec.europa.eu/culture/documents/report\_comite\_des\_sages.pdf\)\)](http://ec.europa.eu/culture/documents/report_comite_des_sages.pdf)
- 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/Apex. Apex is een Europees gesubsidieerd project van nationale archiefdiensten die als thematische agregator de correcte invoer van metadata vanuit archiefdiensten in Europeana voorbereiden.  
[\(\[www.archivesportaleurope.eu\]\(http://www.archivesportaleurope.eu\)\)](http://www.archivesportaleurope.eu).
- Digitalisering vormt vaak een belangrijk onderdeel in wetenschappelijke projecten gevoerd door het Rijksarchief in samenwerking met buitenlandse archiefdiensten en/of onderzoeksinstituten. In het Rijksarchief wordt bij voorbeeld een onderzoeksproject uitgevoerd 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. Omgekeerd worden ook documenten gedigitaliseerd die bewaard worden in het Rijksarchief maar die gevaloriseerd worden in het kader van onderzoek aan een buitenlandse onderzoeksinstituut of universiteit: bij voorbeeld samenwerking met het Holocaust Museum in Washington voor digitalisering van archiefdocumenten uit het archiefbestand "vreemdelingenpolitie".
- De bibliotheek van het KBIN ageert samen met de Plantentuin van Meise en het KMMA als content provider in The Biodiversity Heritage Library for Europe, een project binnen Europeana dat tot doel stelt de bibliotheken van natuurwetenschappelijke bibliotheken te ontsluiten. BHL-Europe is een 3-jarig Europees project dat startte op 1 mei 2009 en duurde tot 1 mei 2012. De gedigitaliseerde werken van het KBIN zullen op die manier raadpleegbaar zijn via het portaal van Europeana.

### 1.3 Doelstellingen

De globale doelstelling was de digitalisering van waardevolle, historische en uitgeputte werken en aansluitend het ter beschikking stellen van deze digitale kopieën.

#### **ARA**

- 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 opgesteld.
- Een tweede doelstelling is de digitalisering van de documenten zelf en de terbeschikkingstelling van deze gedigitaliseerde documenten in de zoekmotor gelinkt aan de metadata. Aangezien het niet haalbaar is om alle documenten te digitaliseren werden prioriteiten vastgelegd. Er wordt een onderscheid gemaakt tussen enerzijds "landelijke prioriteiten" met andere woorden digitaliseringprojecten van een bepaald type bron bewaard in de verschillende rijksarchieven (bij

voorbeeld parochieregisters, registers van burgerlijke stand, kaarten en plannen). Daarnaast worden specifieke (vaak kleinere) projecten voor digitalisering van specifieke archiefbestanden bewaard in één rijksarchief (vb. matrikels van de universiteit Leuven, processen-verbaal van de Ministerraad, zegelafgietsel,...). Conform de archivalistische principes worden documenten steeds in hun context gedigitaliseerd. Meestal worden dan ook volledige archiefbestanden of betekenisvolle onderdelen van archiefbestanden gedigitaliseerd. Losse stukken uit een bestand worden zelden of nooit gedigitaliseerd.

- De derde doelstelling is de uitbouw van de backoffice voor alle aspecten van digitalisering en de creatie van een omgeving waarin nieuwe projecten kunnen gevaloriseerd en beheerd worden.

### **KBIN**

- Een van de bijzondere collecties is de Dautzenberg - Bibliotheek, deze bevat 7.956 gespecialiseerde werken en zeldzame boeken en ruim 200 publicaties van Philippe Dautzenberg zelf, waarin hij 1895 nieuwe taxa beschreef. Die laatste categorie diende om redenen van conservering en preservatie prioritair gedigitaliseerd te worden.
- Een tweede categorie van te digitaliseren werk betrof de uitgeputte wetenschappelijke publicaties van de instelling.
- Een derde categorie van geselecteerde werken bestond uit boeken die eerder in een grootschalig massa-ontzuringproject behandeld werden.

### **SOMA**

- Het gaat meer bepaald om de dossiers tegen Militärbefehlshaber Alexander von Falkenhausen en Militärverwaltungschef Eggert Reeder enerzijds, en de Beauftragte des Chefs der Sicherheitspolizei und Sicherheitsdienst (Sipo-SD) Constantin Canaris anderzijds<sup>1</sup>. De dossiers waarvan sprake zijn in feite de voor de advocaten van de beschuldigen gemaakte fotokopieën van de originele onderzoeksdossiers. Het SOMA/CEGES verwierf ze in de jaren 1970. De keuze viel op deze dossiers omdat ze een geprivilegieerde inblik verschaffen in de functionering en de doelstellingen van het Duitse bezettingsbestuur. Bovendien betreft het fotokopieën van de eerste generatie, die zich bij de start van het project in een zodanige staat bevonden dat raadpleging niet meer vanzelfsprekend was.
- Oorspronkelijk was het dus de bedoeling om aan de slag te gaan met deze gekopieerde, door het SOMA/CEGES verworven dossiers. Al snel echter bleek dat de door het militair gerecht gevormde dossiers ook nog bewaard worden, meer bepaald door de Federale Overheidsdienst Justitie, in het archief van het voormalige militaire auditoraat in het Brusselse justitiepaleis. Na een bezoek aan dat archief bleek dat de "originele" dossiers duidelijk uitgebreider zijn. Niet alleen waren regelmatig minder belangrijke stukken bij het fotokopiëren overgeslagen en had men zich vaak tevreden gesteld met de Franse vertaling van Duitse documenten, het onderzoeksdossier tegen Canaris telde volgens de eigentijdse toegang ook nog eens een 500-tal nummers meer dan het dossier bewaard op het SOMA/CEGES. Bovendien bevinden zich in het justitiepaleis ook de procesdossiers, een aantal administratieve stukken, en de onderzoeksdossiers tegen twee hoofden van de Oberfeldkommandantur van Luik.
- Heel wat elementen pleitten ervoor om de dossiers bewaard in het justitiepaleis te digitaliseren: de volledigheid, de aanwezigheid van originele stukken, de duidelijker samenhang tussen en de grotere kwaliteit van de documenten. Daarom werd uiteindelijk de voorkeur gegeven aan de dossiers van het militair auditoraat. Nochtans had die keuze enkele minder positieve consequenties: niet alleen vereiste de omvang (naar schatting nagenoeg eens zoveel als de oorspronkelijk gekozen dossiers) een duidelijk grotere tijdsinvestering, bovendien kon worden verwacht dat de scanning er niet eenvoudiger op zou worden gezien de grotere verschillen in formaat en kwaliteit van de documenten en de aanwezigheid van onder andere nietjes om verwante documenten samen te houden en touw om liassen te vormen. De samenwerking met het militaire auditoraat, wat eveneens tot problemen had kunnen leiden, verliep opvallend vlot.

---

<sup>1</sup> Procesbundels van het Militair gerecht zaak van Falkenhausen (SOMA/CEGES AA 278); Procesbundels van het Militair gerecht zaak Constantin Canaris (SOMA/CEGES AA 279).

## 2. UITVOERING

### 2.1 Methodologie, technische keuzes, beroep op onderaanneming

Onder de coördinatie van ARA werd de organisatie, methodiek en strategie voor het digitaliseren van documenten vastgelegd.

Archiefdocumenten - historische documenten is een zeer ruim begrip dat zeer verschillende soorten documenten omvat: handgeschreven, getypte of gedrukte documenten. De documenten kunnen losbladig zijn of ingebonden in registers. De herkomst en de leeftijd van de documenten kunnen al even uiteenlopend zijn

Niettegenstaande de grote verscheidenheid van de documenten is er toch een gelijklopende aanpak voor de verschillende projecten. Het betreft immers steeds digitalisering van in hoofdzaak tekstdocumenten waarop geen OCR toegepast wordt maar die in de vorm van een beeld ter beschikking gesteld worden. De documenten hebben een onderlinge samenhang en de gedigitaliseerde beelden moeten in hun context gepresenteerd worden. Om de documenten opzoekbaar te maken worden beschrijvingen en metadata toegevoegd.

#### Werkpakketten

Het project werd beheerd door middel van 15 werkpakketten. Deze werkpakketten omvatten het volledige traject van de digitalisering vanaf de voorbereiding tot de terbeschikkingstelling van de gedigitaliseerde documenten.

#### **WP 1 VASTLEGGEN VAN DE METHODOLOGIE**

- WP 1.1       STANDAARDEN
- WP 1.2       METADATA STANDAARDEN
- WP 1.3       BEST PRACTICES
- WP 1.4       ONDERZOEK VAN TECHNISCHE MOGELIJKHEDEN

Rekening houdend met het belang en de omvang van het project is het belangrijk van bij de aanvang de methodologie vast te leggen. Het project kan zich daarbij baseren op een aantal internationale standaarden waarvan sommige reeds binnen de wetenschappelijke instellingen gebruikt worden. Het betreffen zowel technische standaarden en beschrijvingsstandaarden als meer algemene standaarden voor het beheer en de behandeling van archiefdocumenten.

Een speciaal luik is voorzien voor de metadatastandaarden die gebruikt zullen worden. Een goede keuze van metadata is een belangrijk aspect om op lange termijn de leesbaarheid, bruikbaarheid, interpreteerbaarheid van de digitale documenten te kunnen waarborgen.

Het project onderhoudt contacten met andere gelijkaardige projecten en voorziet de mogelijkheid om goede praktijken van andere projecten toe te passen en te integreren.

Een afzonderlijk luik betreft het onderzoek van de technische mogelijkheden: welke technieken zijn beschikbaar en welke technieken zijn bruikbaar om de analoge documenten om te zetten in digitale bestanden.

Dit werkpakket is zeer belangrijk in de beginfase maar de opvolging van nieuwe ontwikkelingen en eventueel de bijstelling van de methodologie is een constante zorg gedurende het volledige project.

#### **WP 2 OPLEIDING PERSONEEL - OPERATOREN**

Digitalisering zal een belangrijke impact hebben op de werking van de instellingen. Alle personeelsleden en de klanten (in de leeszaal en via de digitale weg) zullen geconfronteerd worden met de resultaten van de digitalisering. Het is niet ondenkbaar dat bepaalde werkprocessen in de toekomst anders zullen georganiseerd worden.

Naast een algemene opleiding voor alle personeelsleden, worden specifieke opleidingen voorzien voor bepaalde personeelsleden die ingeschakeld worden in het digitaliseringproces. Personen die de

scanners zullen bedienen krijgen een opleiding digitaliseringstechnieken, personen die ingeschakeld worden in de voorbereiding van de documenten worden ingewijd in het verzamelen van metadata, ... Het werkpakket situeert zich na de vastlegging van de methodologie maar er kan voorzien worden dat op regelmatige tijdstippen de personeelsleden bijkomende opleidingen zullen volgen.

### **WP 3 SELECTIE EN VASTLEGGEN VAN PRIORITEITEN**

Een eerste selectie en een eerste prioriteitenlijst werd reeds opgesteld bij de voorbereiding van het project. Deze selectie en prioriteitenlijst zullen moeten bijgesteld worden in functie van de beschikbare capaciteit. In functie van de benodigde tijd voor de materiële voorbereiding, inhoudelijke voorbereiding en vervolgens de eigenlijke digitalisering zal op regelmatige basis de prioriteitenlijst bijgesteld worden.

### **WP 4 MATERIËLE VOORBEREIDING VAN DE DOCUMENTEN**

Alvorens de documenten gedigitaliseerd kunnen worden, moeten ze materieel voorbereid worden. Voor sommige archiefbestanden betekent dit dat ze gereinigd dienen te worden, andere documenten moeten eenduidig genummerd worden en adequaat verpakt te worden,... Tot slot blijft er de logistiek: transport vanuit de magazijnen naar het digitaliseringslabo en terugplaatsing.

### **WP 5 INHOUDELIJKE VOORBEREIDING VAN DE DOCUMENTEN**

Documenten kunnen pas gedigitaliseerd worden wanneer er een minimale inhoudelijke beschrijving beschikbaar is. Digitale bestanden zijn immers maar "vindbaar" aan de hand van metadata (in dit geval inhoudelijke metadata).

Dit werkpakket kan zeer verschillend zijn van deelproject tot deelproject. In sommige deelprojecten wordt in dit werkpakket een volledige inventarisering voorzien, terwijl voor andere deelprojecten een snelle controle van de inventaris of een beperkte beschrijving voldoende kan zijn.

### **WP 6 AANBESTEDING**

Indien beslist wordt om bepaalde digitaliseringwerkzaamheden uit te besteden of bij de aankoop van investeringsgoederen of diensten, dient het project de wet op de overheidsopdrachten toe te passen. Rekening houdend met de omvang van de administratieve last, wordt hiervoor een afzonderlijk werkpakket voorzien.

### **WP 7 VOORBEREIDING TECHNISCHE APPARATUUR DIGITALISERING IN EIGEN BEHEER EN ONDERHOUD**

Indien de instelling beschikt over de nodige technische apparatuur kan beslist worden om de digitalisering in eigen beheer uit te voeren. Werkpakket 7 voorziet in de tijd die nodig is om de toestellen operationeel te maken en te houden.

### **WP 8 DIGITALISERING**

Werkpakket 8 is de eigenlijke digitalisering: met andere woorden de opname van de documenten (scanning).

### **WP 9 INVOERING – CAPTURE VAN METADATA**

Tijdens de digitalisering worden een aantal metadata automatisch ingevoerd, andere metadata dienen ingevoerd te worden.

### **WP 10 MATERIËLE CONTROLE**

Na de digitalisering wordt een materiële controle uitgevoerd. Is de digitalisering volledig? Zijn alle documenten gedigitaliseerd. Zijn alle beelden leesbaar? Is het formaat correct? De materiële controle is noodzakelijk om de kwaliteit van het digitaliseringsproject te garanderen en dient uitgevoerd te worden zowel wanneer de digitalisering uitbesteed wordt als bij digitalisering in eigen beheer.

### **WP 11 INHOUDELIJKE CONTROLE**

Naast de materiële controle dient ook aandacht besteed te worden aan de inhoudelijke controle. Zijn de metadata correct? Komen de beschrijvende metadata overeen met de inhoud van de gedigitaliseerde documenten?

### **WP 12 OPSLAG EN LANGE TERMIJN ZORG VAN DE ORIGINELEN**

Na de digitalisering worden de beelden overgedragen aan de wetenschappelijke instelling die instaat voor de bruikbaarheid, authenticiteit,... van de digitale bestanden op lange termijn te garanderen. In dit werkpakket wordt de nadruk gelegd op de "masterbeelden" die op een zo optimaal mogelijke wijze opgeslagen worden en op regelmatige tijd gecontroleerd worden.

### **WP 13 INTEGRATIE IN RAADPLEEGSYSTEEM (INCLUSIEF NODIGE CONVERSIES)**

Naast de masterbeelden, kan een conversie uitgevoerd worden om de beelden in een formaat om te zetten dat meer geschikt is om ter beschikking te stellen. Bij de beschikbaarstelling gelden immers andere criteria dan bij de lange termijnbewaring: snelheid en gebruiksgemak staan daar voorop.

### **WP 14 BESCHIKBAARSTELLING**

Het digitaliseringsproject heeft als doel de digitale beelden ter beschikking te stellen van de onderzoekswereld. De wijze van beschikbaar stellen kan verschillen van deelproject tot deelproject.

### **WP 15 PRODUCTPROMOTIE: ARTIKELS, PUBLICATIES, PERS, COLLOQUIA**

Tot slot is het belangrijk dat aandacht besteed wordt aan de bekendmaking van het project. Er wordt daarbij aandacht besteed aan verschillende gebruikersgroepen: de onderzoekswereld, de amateuronderzoekers die reeds vertrouwd zijn met het erfgoed van de wetenschappelijke instellingen en het grote publiek.

## **Binnen de verschillende instellingen**

### **ARA**

Digitalisering betekent een ingrijpende verandering in de werking van de archieven. Zowel voor het personeel als voor het publiek betekent dit een totaal nieuwe aanpak. Gezien deze impact en voornamelijk ook gezien de kostprijs van digitalisering, is een systematische aanpak noodzakelijk. Het Rijksarchief heeft ervoor geopteerd om een vaste basis te creëren waarbinnen de huidige en toekomstige digitaliseringprojecten kunnen gevoerd worden. De nadruk ligt op de integratie van de verschillende digitaliseringprojecten in één groot geheel en op de lange termijnbruikbaarheid van de resultaten. Het Rijksarchief heeft niet gekozen voor kleinere afzonderlijke projecten met een resultaat op korte termijn maar voor de uitbouw van een digitale leeszaal met een goed gestructureerde back office.

Een aantal deelaspecten van een digitaliseringproject kunnen uitgevoerd worden door eigen personeel of kunnen uitbesteed worden. Het is echter gebleken dat heel wat aspecten en voornamelijk de inhoudelijke voorbereiding van de metadata dient uitgevoerd te worden door eigen, goed opgeleid en gekwalificeerd personeel. De digitalisering op zich, die trouwens slechts ca. 20% à 30% van de tijd van een digitaliseringproject bedraagt, kan uitbesteed worden. De opdrachten aan externe digitaliseringbedrijven werden toegekend op basis van een raamovereenkomst (na openbare aanbesteding). Er werd echter ook geïnvesteerd in de uitbouw van een digitaliseringlabo met aangepaste scanners voor het digitaliseren van diverse types van archiefdocumenten.

Voor het deelaspect "retroconversie" van bestaande toegangen dient het grootste deel van de opdracht eveneens uitgevoerd te worden door gekwalificeerd personeel dat toegang heeft tot de originele documenten en die indien nodig controles kan uitvoeren. Het automatische aspect beperkt zich tot de scanning en de OCR (indien geen elektronische versie beschikbaar is).

De competenties die noodzakelijk zijn in het kader van een digitaliseringproject zijn zeer divers. Er is geen enkel bedrijf dat zowel de diensten kan leveren voor de digitalisering van alle types documenten, de programmering van de raadpleegomgeving en de storage.

Er werd geopteerd om afzonderlijke aanbestedingen uit te schrijven voor de diverse onderdelen om op die manier voor elk onderdeel dat bedrijf te kunnen selecteren dat voldoende ervaring in de materie kan voorleggen en de beste prijs-kwaliteit kan aanbieden:

- aanbesteding voor **digitalisering**  
Deze aanbesteding werd uitgeschreven in de vorm van een raamovereenkomst.  
Er werden 15 verschillende loten gedefinieerd afhankelijk van het type document (registers, losbladige documenten, kaarten en plannen <A0, >A0, negatieven,...).  
Voor elk van de types documenten werden de opdrachten toegekend op basis van een shortlist.  
Aangezien de opdrachten vaak de maximale capaciteit van een bedrijf overschreden, dienden de opdrachten gesplitst te worden over meerdere bedrijven.
- aanbesteding voor **storage**
- aanbesteding voor **servers**
- aanbesteding voor **programmering van de zoekrobot**
- aanbesteding voor **aankoop van A2 bookscanners**
- aanbesteding voor **aankoop A0 scanner**
- aanbesteding voor **aankoop microfilmscanner**
- aanbesteding voor **aankoop van automatische doorvoerscanner**
- aanbesteding voor **hosting**

In het Rijksarchief is "digitalisering" geen extern gebeuren maar een proces dat integraal deel uitmaakt van de dagelijkse werking van de instelling. Alle personeelsleden zijn actief betrokken bij het digitaliseringsproces. De personeelsleden van de afdeling "digitalisering" maar ook de andere personeelsleden ontvingen een aangepaste opleiding "digitalisering" opdat ze vertrouwd zouden zijn met de diverse aspecten van digitalisering en voldoende technische kennis zouden hebben m.b.t. metadata, technische competentie voor digitalisering, verwerking van gegevens,...

Er wordt ook een permanente vorming met regelmatige herhalingscursussen georganiseerd. Normering, standaardisering, procedures, afspraken m.b.t. de verschillende aspecten van digitalisering zijn duidelijk vastgelegd en zijn in overeenstemming met internationale normen. Digitalisering is een uniform proces in alle Rijksarchieven.

### **KBIN**

Wat de methodologie betreft werden de internationale standaarden gevolgd. Voor de metadata werd geopteerd om MARC als bibliotheekstandaard te gebruiken. Dit hield in dat in een eerste fase alle te digitalisere volumes dienden geretrocatalogiseerd te worden. Op die manier konden ze in het LibisNet opgeladen worden en kon de barcodering voor de unieke identificatie van de documenten en kon de bibliografische data als metadata ingekapseld worden.

Deze retrocatalogisering ging in het kader van de preserverings- en conserveringsstrategie gepaard met een schade inventariserend onderzoek van de in te scannen werken.

De eigenlijke scanning gebeurde op de Copibookscanners van projectpartner ARA. Om de online consultatie in de toekomst mogelijk te maken werd gebruik gemaakt van de software BookRestorer. Dit programma laat toe om gescande TIFF-bestanden te corrigeren en te bewerken tot een consulterbaar PDF-formaat. De digitale gegevens werden op de lokale server bewaard en de backups werden op tape gezet. De Dautzenberg collectie werd eveneens ter beschikking gesteld van de Darwin applicatie, een online tool voor collectiebeheer (cfr. project DI/00/05).



## **SOMA**

Met dit project wilde het SOMA/CEGES historisch belangwekkend archief digitaliseren, maar niet alleen dat. Het was uiteraard ook de bedoeling om het gescande materiaal ter beschikking te stellen van het onderzoek. Aangezien dat nogal moeilijk is als de beelden in een groot ongeordend bestand zitten, namen we meteen als uitgangspunt dat het archiefmateriaal voldoende diep ontsloten diende te zijn. Er bestaan oude toegangen, opgesteld door de archiefvormende krijgsauditoraten zelf. Het gaat om toegangen in de vorm van tabellarische plaatsingslijsten die haast elk stuk summier beschrijven. Per stuk bevatten de oude toegangen (in principe) een volgnummer dat ook op de documenten zelf is aangebracht, een datering en een beknopte, niet altijd veelzeggende omschrijving. Een inhoudelijke beschrijving komt slechts zelden voor.

Voor het digitaliseringsproject volstonden deze oude toegangen met andere woorden niet; ze zijn te beknopt en te weinig precies. Om die reden werd beslist de gerechtelijke onderzoeksdossiers tegen von Falkenhausen en Reeder, resp. Canaris opnieuw te beschrijven in een nadere toegang<sup>2</sup>. Met het oog op de digitalisering, het toekennen van de nodige metadata en vanwege het belang van de documenten, gebeurde dat op stukniveau. In de beschrijving werden de volgende (traditionele) elementen opgenomen: een volgnummer, de redactionele vorm, de inhoudsbeschrijving en de datering. Gezien elke beschrijving in de regel slechts één stuk vertegenwoordigt, werd de uiterlijke vorm achterwege gelaten. Van andere mogelijk nuttige elementen, zoals de taal, fysische details of trefwoorden, werd wegens tijdgebrek afgezien. Wat wel werd opgenomen, is de titel die op het document staat, voor zover die afwijkt van de gebruikte redactionele vorm. In totaal werden een kleine 7.000 stukbeschrijvingen opgesteld. De minder belangrijke administratieve stukken en de meer gestructureerde en in omvang beperktere procesdossiers werden beschreven op het niveau van het archiefbestanddeel.

Een (her)ordering was niet aangewezen. De stukken gevormd tijdens het gerechtelijk onderzoek tegen von Falkenhausen en Reeder zitten gewoon in de volgorde waarin ze aan het dossier zijn toegevoegd, dus zonder thematische of andere systematische indeling. Het archief gevormd tijdens het gerechtelijk onderzoek tegen Canaris is wel min of meer thematisch ingedeeld. De krijgsauditeurs gingen hier blijkbaar iets systematischer tewerk.

Na de ontsluiting was het de beurt aan de scanning. Oorspronkelijk was het de bedoeling om dat onderdeel van het project uit te besteden. Grootste nadeel daaraan is dat men de controle op het digitaliseringsproces minstens voor een deel uit handen geeft. Er moeten vooraf dan ook duidelijke keuzes worden gemaakt over onder meer het te gebruiken opslagformaat of de te hanteren resolutie, terwijl aan het bedrijf dat de opdracht uitvoert duidelijke voorwaarden moeten worden opgelegd op het vlak van de omgang met het archiefmateriaal, de wijze van digitaliseren, de kwaliteit ervan, het croppen (knippen) van de beelden etc. Dat de scanning uiteindelijk binnenshuis bleek te kunnen worden uitgevoerd, is de efficiëntie ongetwijfeld ten goede gekomen. Het vergemakkelijkte de communicatie en foutieve scans en andere problemen konden veel sneller worden aangepakt.

### **2.2 Partnerschap ontwikkeld met andere FWI's**

In het kader van het digitaliseringproject hebben zich drie types van samenwerking ontwikkeld.

De eerste samenwerking heeft betrekking op het gezamenlijk gebruik van materiaal voor digitalisering. Het Rijksarchief heeft een volledig uitgebouwd digitaliseringlabo voor de digitalisering van diverse types van documenten. Voor kleinere of specifieke opdrachten en voor zover ruimte beschikbaar is, kunnen de andere FWI's gebruik maken van de infrastructuur. Tijdens de eerste fase van het project heeft bij voorbeeld het KBIN gebruik gemaakt van de bookscanners voor de digitalisering van een boekencollectie.

---

<sup>2</sup> Strikt genomen gaat het niet om inventarissen, wel om nadere toegangen op archiefbestanddelen (dossiers).

Een tweede vorm van samenwerking is "digitalisering van patrimonium" en digitale beschikbaarstelling in een andere FWI's ter aanvulling van de verzameling of ter ondersteuning van onderzoeksopdrachten. Tijdens de eerste fase van het project heeft bij voorbeeld het SOMA een aantal archiefbestanden van het Rijksarchief uitgeleend voor digitalisering en integratie in hun digitale collectie.

De derde vorm van samenwerking is de gezamenlijke terbeschikkingstelling van documenten in een portal. In de marge van het digitaliseringproject werd het Cartesiusproject opgestart als een samenwerking tussen het Nationaal Geografisch Instituut, de Koninklijke Bibliotheek, het Museum van Midden-Afrika en het Rijksarchief voor de valorisatie van historisch kaartmateriaal. Historisch kaartmateriaal dat in de diverse instellingen bewaard wordt, wordt gedigitaliseerd, gegeorefereerd en vervolgens door middel van een gezamenlijke website ter beschikking gesteld van de onderzoekers. Dit project was ook zeer belangrijk voor de uitwisseling van kennis met betrekking tot digitalisering en ontsluiting van kaartmateriaal.

### **2.3 Voorgekomen problemen en aandachtspunten**

Digitalisering wordt soms aangezien als een puur technische aangelegenheid. Gedigitaliseerde objecten moeten echter voorzien worden van of verrijkt met metadata. Dit is een zeer specifieke taak die enkel door gekwalificeerd wetenschappelijk personeel kan uitgevoerd worden. De wijze waarop de metadata en de gedigitaliseerde objecten ter beschikking gesteld worden, dient aan te sluiten bij de specifieke vereisten van de archivaliek. Ook een aantal op het eerste gezicht eerder technische of administratieve handelingen in het kader van de digitalisering zijn specifiek voor elke instelling en vereisen een specifieke opleiding en kennis van het personeel. Het onderdeel "scannen" in het totale digitaliseringproject neemt nauwelijks 20 tot 30 % van de tijd in beslag (zie hoger). Digitalisering in de brede betekenis van het woord maakt integraal deel uit van alle werkprocessen. Bij de uitbesteding van diverse deeltaken in het kader van het digitaliseringproject is gebleken dat slechts een beperkt aantal bedrijven over de nodige kennis en ervaring beschikken om probleemloos de vereisten van een lastenboek uit te voeren. Het voorbeeld van de "nieuwe zoekrobot" heeft aangetoond dat heel wat inbreng van de instelling noodzakelijk is om tot het gewenste resultaat te komen.

Over het algemeen is de samenwerking met de diverse bedrijven die in het kader van het digitaliseringproject materiaal en diensten geleverd hebben aan het Rijksarchief zeer goed verlopen. Voor wat betreft de digitalisering in de strikte betekenis van het woord is gebleken dat kwaliteitscontrole een absolute must is en ook op een systematische wijze dient te gebeuren gedurende de volledige duur van het volledige project. Deze kwaliteitscontrole heeft betrekking op de volledigheid van de digitalisering, de kwaliteit van de beelden en de overeenstemming met de technische vereisten (bestandsformaat, resolutie, kleurruimte,...). Gelet op de waarde van de archiefdocumenten en op de hoeveelheid documenten is een goede opvolging van de logistieke processen een belangrijk aandachtspunt. Het is dan ook zeer nuttig gebleken dat bij de aanvang de vereisten duidelijk geformuleerd waren en als referentie konden gebruikt worden bij de controle en aanvaarding van de opdrachten.

Zoals ook reeds in het hoofdstuk "Methodologie" werd aangehaald, is het gezien de zeer uiteenlopende competenties bijna onmogelijk om één bedrijf te vinden dat alle types van documenten kan digitaliseren, de programmering van de raadpleegomgeving verzorgen, de levering van storage, servers,... kan garanderen. Het Rijksarchief heeft ervoor geopteerd om de coördinatie tussen de verschillende aspecten van de digitalisering zelf te verzorgen en niet te kiezen voor een bedrijf dat het gehele project "aanneemt" en vervolgens samenwerkt met diverse onderaannemers. Het voordeel van deze aanpak is dat het Rijksarchief een volledig zicht heeft op de gevolgde procedures en keuzes en dat een "open" omgeving kon gecreëerd worden die ook in het kader van toekomstige projecten kan gebruikt worden. Anderzijds betekende dit ook dat de investering in tijd en middelen door het Rijksarchief bijzonder hoog lag maar tegelijkertijd dat heel wat ervaring en kennis in het Rijksarchief opgebouwd werd.

Voor de uitbouw van het digitaliseringlabo en het serverlokaal diende de gebouweninfrastructuur van het Rijksarchief te worden aangepast. Deze werken dienen uitgevoerd te worden in samenwerking met de Regie der gebouwen. Het vastleggen van de prioriteiten door de Regie der gebouwen volgt niet altijd de planning van het Rijksarchief en vertraagde in bepaalde gevallen de voortgang van het digitaliseringproject.

Het Rijksarchief is een gedecentraliseerde instelling met diensten verspreid over het gehele land. Invoering van nieuwe procedures in het kader van digitalisering houdt telkens in dat er aandacht besteed moet worden aan communicatie en opleiding voor meerdere personen in de drie landstalen. De nieuwe procedures worden omwille van diverse redenen (personeelstekort, andere prioriteiten,...) niet in alle rijksarchieven volgens hetzelfde tempo opgevolgd. Het Rijksarchief heeft echter resoluut gekozen voor een uniforme oplossing die van toepassing is in alle rijksarchieven, van Brugge tot Arlon. Deze oplossing is de meest kostenefficiënte en is eveneens voor de gebruikers de meest transparante oplossing en maakt samenwerking met andere instellingen mogelijk. Binnen de instelling, heeft het Rijksarchief heel wat ervaring kunnen opdoen in samenwerking tussen diensten en procedures opgesteld voor opvolging van projecten en afhandeling van de taken door de diverse diensten.

Meertaligheid is een belangrijk aandachtspunt in de verwerking van de metadata. In de verschillende diensten wordt respectievelijk Nederlands, Frans, Nederlands/Frans en Frans/Engels als officiële voertaal gebruikt. Daarnaast wordt voor een internationaal onderzoekspubliek ook informatie in het Engels aangeboden. De metadata zijn echter nog niet in alle talen beschikbaar: sommige zijn ééntalig, sommige tweetalig, sommige drietalig (met verschillende combinaties). De verwerking en de presentatie van de gegevens in een zoekrobot op een wijze dat de gebruikers maximaal de gezochte informatie terug vinden en bij voorkeur in de eigen taal of indien dit niet beschikbaar is in een andere taal, is een hele uitdaging gebleken. In de internationale normen is deze problematiek van meertaligheid bovendien niet volledig ingebed.

Een ander aandachtspunt is de specificiteit van archiefbeschrijvingen. In tegenstelling tot metadata in de bibliotheekwereld of in andere culturele en wetenschappelijke omgevingen worden archiefbeschrijvingen hiërarchisch gestructureerd en worden de beschrijvingen steeds binnen een context gemaakt. Vaak beantwoorden meerdere (soms meerdere honderden) digitale objecten aan één beschrijving. Het is gebleken dat deze hiërarchische aanpak die wezenlijk is binnen de archivaliek uitwisseling met andere culturele instellingen (cf. portal Cartesius maar ook naar Europeana) vaak bemoeilijkt waar veelal uitgegaan wordt van een "platte beschrijving" (al dan niet *Dublin core*) met een 1 op 1 relatie tussen metadata en digitaal object.

De belangrijkste doelstelling van de digitalisering is een verbetering van de toegankelijkheid van het patrimonium van de wetenschappelijke instellingen. Gezien de diversiteit van het bezoekerspubliek en het potentiële bezoekerspubliek dat via het internet kan bereikt worden, is het belangrijk om regelmatig de prioriteiten - zowel voor wat betreft de gedigitaliseerde objecten als de tools voor de raadpleging - te evalueren en bij te stellen. De vragen of de eisen van sommige gebruikers zijn soms zeer hoog en niet altijd realistisch. Omwille van budgettaire redenen, personeelomkering en praktische redenen dienen keuzes gemaakt te worden. "Alles gratis onmiddellijk à la google beschikbaar" is wellicht niet realistisch. Een goede communicatie en open discussie met de diverse gebruikersgroepen is een must om frustraties te vermijden en al te grote verwachtingen te voorkomen.

Digitaliseringprojecten betekenen een belangrijke investering voor de wetenschappelijke instellingen. Een belangrijk aandachtspunt is dat er op lange termijn budgetten voorzien worden voor het instandhouden en toegankelijk houden van de digitale objecten. Er dient permanent personeel ter beschikking te zijn voor de controle en opvolging van de lange termijnbewaring, de opvolging van nieuwe technologieën, de eventuele migraties van bestanden, aanpassing van softwaretoepassingen,... en een regelmatige vernieuwing van het materiaal dient ingepland te worden.

Het onderdeel digitalisering viel uiteen in verschillende stappen die voor SOMA werden eerst de nadere toegangen opgeladen in Pallas, het geautomatiseerde, geïntegreerde ontsluitingssysteem waarvan het SOMA/CEGES gebruikmaakt. Vervolgens werd elk stuk afzonderlijk in een omslagje gestopt. Daarop werd het nummer aangebracht van het record in Pallas met de corresponderende stukbeschrijving. Vooraleer met het eigenlijke scannen gestart kon worden, moesten een aantal technische keuzes gemaakt worden. Een eerste vraag luidde of het nodig is de stukken in kleur te scannen. Omdat er in de dossiers maar weinig kleur voorkomt, werd beslist te scannen in grijswaarden (met een kleurdiepte van 8 bits). Als scanresolutie kozen we voor 300 dpi, de vaakst gebruikte resolutie voor het digitaliseren van documenten op een papieren drager. Een volgende vraag was of het zinvol is om op de beelden optische tekenherkenning (OCR) toe te passen. De slechte kwaliteit van de getypte en de vele handgeschreven stukken deden ons daarvan afzien. Qua bestandsformaat tot slot kozen we voor JPEG in plaats van het gebruikelijkere TIFF (Tagged Image File Format). Het SOMA/CEGES gebruikt dat formaat al vele jaren voor de gedigitaliseerde foto's. Voordelen van JPEG zijn onder andere de geringere opslagruimte en de brede softwareondersteuning voor het formaat. Pas toen konden de personeelsleden van het SOMA/CEGES-digitaliseringsatelier met hun bookscanner aan de slag. Zij zorgden ook voor het eventuele, zorgvuldige verwijderen van nietjes, koorden en andere hechtmaterialen. De digitale beelden van elk gedigitaliseerd stuk sloegen ze op in een map die genoemd werd naar het nummer op de omslag.

De kwaliteitscontrole op de digitalisering, de volgende stap, gebeurde niet steekproefsgewijs. Omdat de dossiers niet behoren tot de SOMA/CEGES-collectie en er dus na verloop niet zomaar zou kunnen worden herscand, leek het ons noodzakelijk elk beeld, elke pagina afzonderlijk te controleren. Een bijzonder tijdrovende bezigheid die echter zijn nut bleek te hebben. Een klein, maar niet onbelangrijk percentage (vergeten pagina's, onduidelijke scans en onvolledige beelden) moest opnieuw gescand worden. Naderhand volgde een nieuwe kwaliteitscontrole en werden de foute beelden geschrapt en vervangen door de opnieuw gescande beelden. Toen alles in orde was, konden de meer dan 36.500 beelden (35 gigabyte) opgeladen worden in de databank Pallas. Zoals uit een steekproef naar voren kwam, blijkt dat ook een dergelijk geautomatiseerd proces niet altijd foutloos verloopt.

Een laatste belangrijk element bij digitalisering is de (langdurige) opslag van de beelden. Als ze slechts op één plaats worden bewaard, is het risico niet denkbeeldig dat ze als gevolg van een computercrash bijvoorbeeld, verloren zouden gaan. Daarom werd een bewaarstrategie ontwikkeld. Op dit ogenblik worden de scans op vier plaatsen bewaard: op twee verschillende servers, waarvan één reservekopie, en op twee externe harde schijven. Al deze kopieën worden bewaard in een en hetzelfde gebouw. Om de veiligheid nog te verhogen, komt er nog dit jaar een kopie bij, die op een andere locatie zal worden bewaard. Tot slot zullen de nodige voorzorgsmaatregelen worden genomen opdat de beelden ook over 20 of 30 jaar nog leesbaar zijn.

## 2.4 Middelen (personeel en budget)

### 2.4.1 Ten laste van Belspo

- Personeel

| ARA  |            |
|--|------------|
| Personeel (Categorie, Specialiteit)                | Aantal M/M |
| Wetenschappelijk personeel                         | 110        |
| Administratief en technisch personeel (niveau C/D) | 738        |
| <b>TOTAAL</b>                                      | <b>848</b> |

Het aantal personeelsleden ten laste van DI/00/03 heeft zeer sterk geschommeld over de volledige periode van het project. Bij het beëindigen van het DI/00/03 -project in maart 2012 zijn nog 2 archivariissen en 12 contractuele personeelsleden in dienst.

| KBIN   |            |
|--|------------|
| Personeel (Categorie, Specialiteit)                          | Aantal M/M |
| CA1 Projectmedewerker:<br>metadata, catalografie en scanning | 42         |
| <b>TOTAAL</b>  | <b>42</b>  |

- Budget

| Globaal        |                  |
|----------------|------------------|
| (EUR)          | TOTAAL           |
| Personeel      | 2.548.772        |
| Werking        | 426.852          |
| Overheads      | 144.085          |
| Uitrusting     | 1.396.402        |
| Onderaanneming | 561.939          |
| <b>TOTAAL</b>  | <b>5.078.050</b> |

| ARA            |                  |
|----------------|------------------|
| (EUR)          | TOTAAL           |
| Personeel      | 2.426.458        |
| Werking        | 398.788          |
| Overheads      | 141.263          |
| Uitrusting     | 1.396.402        |
| Onderaanneming | 561.939          |
| <b>TOTAAL</b>  | <b>4.924.850</b> |

| KBIN           |                |
|----------------|----------------|
| (EUR)          | TOTAAL         |
| Personeel      | 104.314        |
| Werking        | 28.064         |
| Overheads      | 2.822          |
| Uitrusting     | 0              |
| Onderaanneming | 0              |
| <b>TOTAAL</b>  | <b>135.200</b> |

| SOMA           |               |
|----------------|---------------|
| (EUR)          | TOTAAL        |
| Personeel      | 18.000        |
| Werking        | 0             |
| Overheads      | 0             |
| Uitrusting     | 0             |
| Onderaanneming | 0             |
| <b>TOTAAL</b>  | <b>18.000</b> |

## 2.4.2 Ten laste van de FWI's

## Personeel

| ARA  |            |
|--|------------|
| Personeel (Categorie, Specialiteit)                | Aantal M/M |
| Wetenschappelijk personeel                         | 66         |
| Administratief en technisch personeel (niveau C/D) | 114        |
| <b>TOTAAL</b>                                      | <b>180</b> |

- Alle wetenschappelijke personeelsleden leveren metadata aan en worden ingeschakeld in de kwaliteitscontrole van de metadata;
- Alle personeelsleden van niveau C/D tewerkgesteld in de Rijksarchieven in de provinciën worden voor een beperkt percentage van hun arbeidstijd ingeschakeld in de digitalisering van archiefdocumenten (voornamelijk kleinere digitaliseringprojecten);
- Personeel van de ICT-dienst ondersteunt de werking van de afdeling "DIGIT";
- In bovenstaande tabel wordt enkel het personeel van de afdeling DIGIT in rekening gebracht.

| KBIN                                |             |
|-------------------------------------|-------------|
| Personeel (Categorie, Specialiteit) | Aantal M/M  |
| Promotor                            | 22,8        |
| Coördinator                         | 3,2         |
| <b>TOTAAL</b>                       | <b>26,0</b> |

| SOMA                                |              |
|-------------------------------------|--------------|
| Personeel (Categorie, Specialiteit) | Aantal M/M   |
| A                                   | 135,0        |
| C                                   | 389,0        |
| D                                   | 51,5         |
| <b>TOTAAL</b>                       | <b>575,5</b> |

## - Budget

| Globaal        |                  |
|----------------|------------------|
| (EUR)          | TOTAAL           |
| Personeel      | 496.737          |
| Werking        | 6.376            |
| Overheads      | 0                |
| Uitrusting     | 801.000          |
| Onderaanneming | 0                |
| <b>TOTAAL</b>  | <b>1.304.113</b> |

| ARA            |                  |
|----------------|------------------|
| (EUR)          | TOTAAL           |
| Personeel      | 302.917          |
| Werking        | 0                |
| Overheads      | 0                |
| Uitrusting     | 801.000          |
| Onderaanneming | 0                |
| <b>TOTAAL</b>  | <b>1.103.917</b> |

| KBIN           |                |
|----------------|----------------|
| (EUR)          | TOTAAL         |
| Personeel      | 126.945        |
| Werking        | 6.376          |
| Overheads      | 0              |
| Uitrusting     | 0              |
| Onderaanneming | 0              |
| <b>TOTAAL</b>  | <b>133.322</b> |

| SOMA           |               |
|----------------|---------------|
| (EUR)          | TOTAAL        |
| Personeel      | 66.875        |
| Werking        | 0             |
| Overheads      | 0             |
| Uitrusting     | 0             |
| Onderaanneming | 0             |
| <b>TOTAAL</b>  | <b>66.875</b> |

### 3. REALISATIES

#### ARA

Het hoofdaandeel van de digitalisering in de strikte betekenis van het woord, betreft "documenten": losbladige documenten, registers, kaarten en plannen, fiches,... Er werden eveneens een aantal kleinere projecten gerealiseerd rond specifieke niet-papieren dragers: glasnegatieven en fotonegatieven alsook rond audiovisuele archieven (geluidsbanden).

Voor de presentatie van archiefdocumenten is het cruciaal dat de gedigitaliseerde documenten in hun context (meer bepaald per archiefbestand) gepresenteerd worden ongeacht het type document, de drager,... Bij de digitalisering daarentegen dient een onderscheid gemaakt te worden tussen de diverse dragers aangezien deze met een ander type scanner dienen gedigitaliseerd te worden.

Het resultaat van alle projecten wordt via één centrale toegang aangeboden. Het voordeel van deze aanpak is dat deze "digitale leeszaal" naadloos aansluit bij de bestaande dienstverlening in de leeszalen. Naarmate digitale informatie beschikbaar is, wordt deze onmiddellijk zichtbaar voor de lezers. Zij kunnen dan de digitale informatie raadplegen in plaats van de originele documenten op te vragen in de leeszaal. Deze werkwijze is noodzakelijk aangezien het niet de bedoeling kan zijn "alle" archiefdocumenten (ca. 65 km) te digitaliseren, maar het is belangrijk dat het voor de lezers duidelijk is welke documenten ze kunnen raadplegen en in welke vorm.

#### ***Realisatie 1: uitbouw van een volledig en uniform metadatasysteem voor de beschrijving van archiefdocumenten***

Het metadatasysteem is volledig gebaseerd op de internationale standaarden voor archiefdiensten:

- ISO 15111 - International Standard Identifier for Libraries and Related Organizations
- ISAD(G) - General International Standard Archival Description
  - \* EAD Encoded Archival Description
- ISAAR(CPF) International Standard Archival Authority Record For Corporate Bodies, Persons and Families
  - \* EAC Encoded Archival Context
- ISDF International Standard for Describing Functions
- ISDIAH International Standard for Describing Institutions
- ISO standaarden voor beschrijving van talen, landen, dateringen, ...
- Alle metadata in XML formaat (EAD en EAC)
- Technische metadata METS (Metadata Encoding and Transmission Standard)
- Unimarc (bibliografische beschrijvingen)

Deze ervaring en realisatie zijn niet enkel belangrijk voor de eigen toepassingen maar zijn bijzonder nuttig gebleken in de samenwerking en uitwisseling van informatie met andere archiefdiensten (nationaal en internationaal: vb. Apex/Apenet).

Het Rijksarchief heeft als enige Europese nationale archiefdienst een zoekrobot "archiefvormers" en wordt als voorbeeld gebruikt worden in het Europees project Apex.

In het kader van het Cartesiusproject werd gestart met de geolokalisatie van archiefdocumenten en voornamelijk van kaarten en plannen. Het is de bedoeling een webapplicatie te bouwen waarbij het mogelijk wordt uitgaande van een huidige kaart historisch kaartmateriaal op te zoeken. In gezamenlijk overleg met de andere partners in het Cartesiusproject werden de standaarden voor geolokalisatie vastgesteld. Er werd ook gewerkt aan een uitwisselstandaard van de metadata voor archieven naar de ISO-standaarden gebruikt in geografische informatiesystemen conform de Inspire-richtlijn (ISO 19115, 19119 en 19169).



In samenwerking met het APENET/Apex project worden ook de uitwisselstandaarden vastgelegd met Europeana (achtereenvolgens de ESE (Europeana Semantic Element) en EDM (Europeana Data Model)).

### **Realisatie 2: metadata repository**

Metadata zijn cruciaal voor het terugvinden van de informatie. Er werd voor geopteerd om alle metadata te integreren in één centrale repository. Deze repository is de basis voor de presentatie van de gegevens online en voor het beheer van de informatie. De metadata worden in XML opgeslagen conform de geldende standaarden (cf. realisatie 1)

De metadata werden deels manueel ingevoerd en deels geretroconverteerd vanuit bestaande toegangen, inventarissen,...

Cijfergegevens van de informatie beschikbaar via de digitale leeszaal:

- 22.252 archiefbestanden
- 7.807.720 bestanddelen
- 20.888 toegangen
- 32.768 archiefvormers
- 5.358 inventarissen in EAD-formaat (*full text* + hiërarchisch doorzoekbaar)

### **Realisatie 3: uitbouw digitaliseringlabo**

Het Rijksarchief beschikt over een volledig ingericht digitaliseringlabo voor digitalisering van archiefdocumenten. Het digitaliseringlabo is uitgerust met 3 bookscanners voor digitalisering van documenten tot A2, 2 digibooks voor digitalisering van kaarten, plannen en registers tot A0, 1 microfilmscanner en 1 automatische doorvoerscanner. Daarnaast zijn alle leeszalen van het Rijksarchief uitgerust met een bookscanner voor digitalisering van documenten tot A2, die gebruikt wordt voor digitale reproductieaanvragen en kleine digitaliseringsprojecten.

### **Realisatie 4: vastlegging standaarden voor digitalisering, werkprocessen, kwaliteitscontrole, indicatoren,...**

Het volledige werkproces "digitalisering" werd vastgelegd en beschreven. In elke fase wordt bijzondere aandacht besteed aan kwaliteitscontrole. Gezien de grote hoeveelheid digitale informatie is het van essentieel belang dat alle unieke identificatoren correct zijn en dat de link tussen de beelden en de metadata gegarandeerd is. Speciale aandacht wordt besteed aan de volledigheid van de digitalisering en de kwaliteit van de beelden.

Er wordt zoveel als mogelijk gebruik gemaakt van open standaarden om de interoperabiliteit en de lange termijnbewaring te garanderen.

### **Realisatie 5: infrastructuur voor bewaring en raadpleging**

Het Rijksarchief beschikt over infrastructuur voor de bewaring en de raadpleging van de metadata en de gedigitaliseerde documenten.

Een aangepaste en gebruiksvriendelijke raadpleegomgeving zowel voor de metadata als voor de gedigitaliseerde documenten is cruciaal voor het succes van de digitale leeszaal. Hierbij moeten de technische mogelijkheden in overeenstemming gebracht worden met de vaktechnische vereisten en beperkingen en met de wensen van de klanten (zeer uiteenlopende groep!).

Het resultaat zijn de zoekroboten die via <http://search.arch.be> raadpleegbaar zijn. Deze werden in verschillende fasen online beschikbaar gesteld. Telkens werd een bevraging georganiseerd zodat terdege rekening kan gehouden worden met de wensen en opmerkingen van de klanten.

### **Realisatie 6: beschikbare gedigitaliseerde documenten**

In de leeszalen van het Rijksarchief waren op 1 januari 2012: **133.246.189 bladzijden** in digitale vorm beschikbaar.

Hierbij enkele van de belangrijkste projecten:

- Parochieregisters
- Burgerlijke stand
- Notulen ministerraad, secretarissen-generaal
- Vreemdelingenpolitie
- Matrikels universiteit Leuven
- ITS (Bad Arolsen)
- Zegelafgietsels
- Glasnegatieven en filmnegatieven mijnen Limburg
- Kaarten en plannen

Iets meer dan **40 miljoen bladzijden** zijn reeds gedigitaliseerd maar dienen nog aan een kwaliteitscontrole onderworpen te worden of gekoppeld te worden aan de respectieve metadata. Zodra de zoekrobot en de opslagcapaciteit beschikbaar is, kan deze digitale informatie eveneens ter beschikking gesteld te worden.

Via internet is momenteel slechts een fractie van deze digitale beelden raadpleegbaar. Zodra de aanbesteding "hosting" afgerond is, wordt het aanbod uitgebreid.

Het merendeel van de archiefdocumenten is beschikbaar als "beeld". Voor een beperkt aantal archiefdocumenten is een full-text search mogelijkheid voorzien als (notulen van de ministerraad 1918-1979)

#### ***Realisatie 7: modaliteiten voor samenwerking met derden (archiefdiensten, portals, ...)***

Diverse modaliteiten voor samenwerking met derden werden onderzocht en vastgelegd in gestandaardiseerde samenwerkingsovereenkomsten: bv. samenwerking voor digitalisering (met GSU), voor online terbeschikkingstelling (vb. met stads- en gemeentearchieven), voor uitwisseling van metadata (vb. met Apex, Archiefbank Vlaanderen)

#### ***Realisatie 8: web 2.0, crowdsourcing***

De nieuwe informatietechnologie biedt heel wat mogelijkheden voor samenwerking. Vooral *social tagging* is in het kader van de ontsluiting van archiefdocumenten zeer nuttig. Het Rijksarchief heeft een netwerk van "vrijwilligers" die via digitale weg gedigitaliseerde archiefdocumenten nader ontsluiten. (Project Demogen – e-gov Award Innovatie 2009)

Het resultaat van de *social tagging* wordt via internet ter beschikking gesteld in de zoekrobot "zoeken naar personen". Deze zoekrobot telde op 1 januari 2012 meer dan 15 miljoen namen. De databank groeit gestaag aan.

## **KBIN**

De digitalisering van de publicaties van Philippe Dautzenberg werd gefinaliseerd eind 2008 en ingebracht in het nieuwe beheerssysteem Aleph. De aldus verkregen bibliografische gegevens in MARC-formaat werden vervolgens als metadata gebruikt bij de upload van de digitale content naar BHL-Europe en Europeana. De export is aan de gang, met als einddatum april 2012.

In totaal betreft het voor de Dautzenberg Bibliotheek 7.809 TIFF-bestanden of pagina's. De digitale bestanden zijn in PDF-formaat beschikbaar voor gebruik in de bibliotheek en de afdeling malacologie.

Aansluitend werden in de periode 2008-2009 de uitgeputte publicaties van het KBIN gescand. Op deze manier werden deze uitgaves opnieuw beschikbaar voor het publiek. Het betreft circa 20.000 TIFF-bestanden of pagina's.

Van de waardevolle collecties en ontzuurde werken werd eveneens een ruime selectie gescand.

In totaal werden er door de KBIN-bibliotheek ruim 65.000 TIFF'S (PDF's ) aangemaakt die in de loop van 2012 via Europeana beschikbaar zullen zijn voor het publiek.

## SOMA

De collectie dossiers die uiteindelijk gedigitaliseerd werd, bestaat uit documenten gevormd in het kader van:

- het gerechtelijk onderzoek van het auditoraat-generaal tegen Karl Constantin Canaris (hoofd van de Sipo-SD in België en Noord-Frankrijk) (SOMA/CEGES AA 2146),
- het proces tegen Karl Constantin Canaris voor de krijgsraad te Brussel (SOMA/CEGES AA 2146),
- het gerechtelijk onderzoek van het krijgsauditoraat te Luik tegen Alexander von Falkenhausen (Militärbefehlshaber van België en Noord-Frankrijk) en Eggert Hans Reeder (Militärverwaltungschef van België en Noord-Frankrijk) (SOMA/CEGES AA 2143),
- het gerechtelijk onderzoek van het krijgsauditoraat te Luik tegen George Franz Bertram (Oberfeldkommandant Luik) (SOMA/CEGES AA 2144),
- het gerechtelijk onderzoek van het krijgsauditoraat te Luik tegen Bernhardt von Claer (Oberfeldkommandant Luik) (SOMA/CEGES AA 2145),
- het proces tegen Alexander von Falkenhausen, Eggert Hans Reeder, George Franz Bertram en Bernhardt von Claer voor de krijgsraad van Brabant (SOMA/CEGES AA 2143),
- het gerechtelijk onderzoek van het krijgsauditoraat te Brussel tegen Karl Theodor Moskopf (sinds de zomer van 1944 leider van het Ersatzkommando Wallonien der Waffen-SS) (SOMA/CEGES AA 2118).

Inhoudelijk hebben de stukken voornamelijk betrekking op de werking van Militärverwaltung en Sipo-SD, de verhoudingen tussen beide instanties, de represaillemaatregelen na verzetsacties en het executeren van gijzelaars, de verplichte tewerkstelling, de deportatie van Joden en het kamp van Breendonk. Behalve naoorlogse verhoren bevatten de dossiers ook tal van (Duitse) documenten die tijdens (of zelfs voor) de oorlog werden opgesteld<sup>3</sup>.

---

<sup>3</sup> Verwant archiefmateriaal: Strafvordering, minuten zittingsverslagen, pleidooi en persdossier meester Veldekens in de zaak von Claer, 1945 (SOMA/CEGES AA 273); Archives de l'avocat Jules Wolf concernant résistance, répression, défense des droits de l'homme et Ligue belge des Droits de l'Homme: Procès von Falkenhausen et consorts (SOMA/CEGES AA 1836/11-21); Dossiers de l'avocat Bruno Wuyts, défenseur du général Bertram (SOMA/CEGES AA 1937).

## 4. VERSPREIDING EN VALORISATIE

### ARA

Centraal staat de website search.arch.be waar dagelijks vernieuwd wordt. De belangrijkste zoekingen zijn de archieven, de archiefvormers en de personen vermeld in archiefdocument. Daarnaast zijn verscheidene thematische "ingangen" voorzien op basis van type archiefvormers, type archieven,... Deze website geeft toegang tot zowel de metadata als de gedigitaliseerde objecten.

Op regelmatige tijdstippen wordt het gebruik van de website geëvalueerd. Het eerste deel van de website "zoeken naar personen" werd gelanceerd half juni 2011. Het gedeelte archieven, archiefvormers kwam in april 2012 online. Het gemiddeld aantal bezoekers is over de volledige periode gestegen. Momenteel telt de website meer dan 1.200 bezoekers per dag. Gemiddeld gebruiken zij de website gedurende 20 minuten.

The screenshot shows the website's search interface. At the top left is the logo of the National Archives of Belgium, consisting of a stylized 'a' inside a square frame. To its right is the text 'Het Rijksarchief in België' and 'Zoeken in het Rijksarchief in België'. Below the logo is a button that says 'Vul onze enquête in'. A search bar is located on the left side with the placeholder text 'Zoekterm...' and a 'Zoeken' button. The navigation menu includes 'Thuispagina', 'Archieven', 'Archiefvormers', 'Personen', and 'Thema's'. The main content area has a heading 'Zoeken in het Rijksarchief in België' and contains several paragraphs of text:

De 'zoekrobot' zijn dé toegang tot de archieven die bewaard worden in het Rijksarchief. De vernieuwing van deze zoekrobot is één van de belangrijkste projecten voor het Rijksarchief. Niet alleen de technische infrastructuur maar ook de zoekmogelijkheden en de presentatie worden aangepast. De verschillende onderdelen van de zoekrobot zullen één na één beschikbaar gesteld worden. Het eerste onderdeel betrof zoeken naar personen'. Vanaf 26 maart is ook de bètaversie van 'zoeken naar archieven' en 'zoeken naar archiefvormers' beschikbaar.

De zoekmogelijkheden zijn vernieuwd en bij de presentatie van de gegevens werd rekening gehouden met de suggesties die we vanuit diverse hoeken de voorbije jaren ontvangen hebben. Belangrijke vernieuwing is dat full-text zal gezocht kunnen worden in de algemene bestandsbeschrijving. De historie van het archiefbestand, de archiefvormer, e.d. ... worden opgenomen in de zoekrobot.

Het onderdeel 'archiefvormers' wordt stelselmatig aangevuld met informatie uit de vele gidsen die in de voorbije jaren gepubliceerd werden door het Rijksarchief.

Vooraf het scherm voor de raadpleging van de gedigitaliseerde documenten werd grondig herzien. Het scherm is groter, het wordt mogelijk om onmiddellijk naar een bepaalde bladzijde in een register door te klikken, nieuwe hulpmiddelen om de beeldkwaliteit aan te passen werden toegevoegd, ...

Hoewel de meeste gedigitaliseerde archieven nog enkel beschikbaar zijn in de leeszalen van het Rijksarchief, betekent de vernieuwing van de zoekrobot toch een zeer belangrijke stap richting online brengen van de digitale beelden. We zijn daartoe momenteel de voorbereidingen aan het treffen.

We nodigen alle gebruikers van de zoekrobot ook uit, deel te nemen aan de enquête. Net zoals dat het geval was voor de 'zoekrobot personen' zullen uw evaluatie en suggesties zeer nuttig zijn om dit werkinstrument verder te verbeteren.

Bij de lancering van de nieuwe zoekrobot werd aan de gebruikers gevraagd hun waardering over de nieuwe zoekrobot uit te spreken. De enquête heeft betrekking op de zoekmogelijkheden, de snelheid van de toepassing en de presentatie. Voor elk van deze elementen wordt een waardering gevraagd op een schaal van 1 tot 10 en wordt de mogelijkheid geboden bijkomende informatie in te voeren. De enquête kon anoniem ingevuld worden of men kon ervoor opteren om naam en e-mailadres te vermelden. Een eerste analyse na 50 dagen toonde aan dat de gebruikers zeer positief zijn over de nieuwe mogelijkheden. Meer dan 75% gaf een score van 7 of meer aan de zoekcriteria. Meer dan 82% van de respondenten gaf 7 of meer op het aspect snelheid. En de presentatie werd door 75% van de gebruikers gequoteerd met een cijfer van 7 of meer, meer dan de helft van de respondenten gaf zelfs een score van 8 of hoger. De enquête leverde ook heel wat interessante suggesties op die gebruikt konden worden voor kleine verbeteringen in de zoekrobot en voor extra toelichtingen. De volledige analyse is terug te vinden op:

- [http://extranet.arch.be/arch/newsletter19/evaluatie\\_nieuwe\\_zoekrobot.pdf](http://extranet.arch.be/arch/newsletter19/evaluatie_nieuwe_zoekrobot.pdf) (NI)
- [http://extranet.arch.be/arch/newsletter19/evaluation\\_moteur\\_de\\_recherche.pdf](http://extranet.arch.be/arch/newsletter19/evaluation_moteur_de_recherche.pdf) (Fr).

Op regelmatige tijdstippen wordt in de elektronische nieuwsbrief van het Rijksarchief de voortgang van de digitaliseringprojecten vermeld.

In de periode januari - december 2010 werd een wekelijkse elektronische nieuwsbrief "digitale parochieregisters" verspreid aan meer dan 4.000 abonnees.

Bij belangrijke mijlpalen in het project werd een persconferentie georganiseerd. Op 23 oktober 2009 werd het digitaliseringlabo van het Algemeen Rijksarchief plechtig geopend. Op 19 oktober 2010 werd in de Wetstraat 16 in samenwerking met de Kanselarij van de Eerste Minister de website gelanceerd met de gedigitaliseerde notulen van de Ministerraad. Op 11 januari 2011 werd het resultaat van de samenwerking met het Stadsarchief Brussel voor de digitalisering van de parochieregisters van Brussel voorgesteld.

In de periode januari - februari 2010 werd een tentoonstelling "Rijksarchief gaat digitaal" in de inkomhal van het Algemeen Rijksarchief te Brussel.

Medewerkers van de afdeling digitalisering hebben deelgenomen aan meerdere studiedagen georganiseerd door genealogische verenigingen (Roeselare 2011, Mol 2012, Tournai 2010,...).

Het digitaliseringproject werd voorgesteld en toegelicht op internationale conferenties (Genève 2010, 8th European Conference on Digital Archiving; Luxemburg, Belgisch-Duits-Nederlands-Luxemburgs Archievenverleg; Tholen en Middelburg, Zeeuws archievenoverleg, enz.).

Op 3 december 2009 mocht het Rijksarchief de *e-gov award* voor meest innovatieve project ontvangst nemen voor het Demogenproject, een *social tagging*-project in samenwerking met vrijwilligers voor de ontsluiting van gedigitaliseerde archiefdocumenten.

De digitaliseringprojecten kwamen ook ruimschoots aan bod tijdens persconferenties en studiedagen rond specifieke thema's zoals "vreemdelingendossiers", "archief en democratie", "tentoonstelling Breekbaar verleden Limburg",... De digitale toegankelijk is immers vaak een belangrijke start voor wetenschappelijke onderzoeksprojecten.

In oktober 2010 werd het Rijksarchief gehoord door het zogenaamde "Comité des sages", een commissie die een advies diende te formuleren voor "La numérisation du patrimoine culturel Européen".

### **KBIN**

L. MEESE, *No innovation without inspiration*, in Into Info, 2011, p. 8-9.

L. MEESE, Voordracht en presentatie "*KBIN bibliotheek, verleden, heden en toekomst*", 28 februari 2008, Colloquium Libis-Net, Boerenbond, Leuven.

T. MALLAERTS & N. HANQUART, Voordracht en presentatie "*The Biodiversity Heritage Library (BHL-Europe): towards a global library of life*", 18 mei 2011, Dag van de bibliothecaris, Brussel, 17 mei 2011.

### **SOMA**

Zoals reeds aangehaald, was het de bedoeling om het gedigitaliseerde materiaal aan het publiek ter beschikking te stellen. Probleem is echter dat de dossiers tal van persoonsgegevens bevatten. Zeker wanneer die nog levende personen betreffen, kunnen die niet zomaar in de openbaarheid worden gegooid. Dat geldt a fortiori voor een aantal categorieën van gegevens die volgens de wetgeving op de bescherming van de persoonlijke levenssfeer extra zorg vergen. Het gaat om gerechtelijke gegevens, gezondheidsgegevens en andere gevoelige informatie (gegevens over raciale of etnische achtergrond,

politieke opvattingen, levensbeschouwelijke overtuigingen, lidmaatschap van een vakvereniging of het seksuele leven)<sup>4</sup>. Nu komt precies dat soort gegevens relatief vaak voor in de door ons gedigitaliseerde dossiers. Daarom vroegen we de Belgische privacy commissie (voluit: Commissie voor de Bescherming van de Persoonlijke Levenssfeer) om een advies.

Het advies stelde dat zowel de toegang als de stukken enkel in de leeszaal van het SOMA/CEGES kunnen worden bekeken of via een beveiligde toegang op het internet. De aanbeveling van de Commissie stelde onze instelling voor een probleem. Via de online databank Pallas is het momenteel immers niet mogelijk om de stukken enkel in de leeszaal ter beschikking te stellen. Ze zijn ofwel zichtbaar voor iedereen op het wereldwijde web, ofwel enkel voor het personeel van het SOMA/CEGES. Daarom werd gekozen voor een voorlopige oplossing: van de toegang werd een statische html-pagina gemaakt, die op de raadpleegcomputers van de leeszaal werd geplaatst. In die pagina vormt elke stukbeschrijving een hyperlink. Als men erop klikt, worden de bijhorende beelden geopend. We hopen de digitale beelden binnen afzienbare tijd wel via Pallas in de leeszaal te kunnen aanbieden, gezien de grotere gebruiksvriendelijkheid en de uitgebreidere zoekmogelijkheden van dat systeem. Een eventuele beperkte en tijdelijke toegang met paswoord via het internet blijft voorlopig nog toekomstmuziek. Het is wel al mogelijk om onderzoekers via e-mail een EAD-inventaris in PDF op te sturen.

---

<sup>4</sup> Zie <http://www.privacycommission.be/de/>, laatst geraadpleegd op 20/03/2012.

## 5. BALANS EN PERSPECTIEVEN

De hierboven vermelde resultaten tonen aan dat de objectieven gehaald zijn en dat we het project dus als "gelukt" mogen beschouwen. Dit betekent niet dat het geen uitdaging was.

Bij de aanvang bleek al zeer snel dat uitbesteding van "digitalisering tot valorisatie" geen optie was. Er werd dus van bij de aanvang geopteerd voor een opdeling van het grote project in deelprojecten waarbij voor elk onderdeel beslist werd of het in eigen beheer uitgevoerd werd dan wel uitbesteed werd. Deze werkwijze was zeer arbeidsintensief in het bijzonder voor het Rijksarchief. Dit heeft er ook voor gezorgd dat de eerste resultaten langer op zich hebben laten wachten dan misschien oorspronkelijk verhoopt. Anderzijds heeft dit er ook voor gezorgd dat het Rijksarchief op geen enkel onderdeel afhankelijk van één externe partner en dat het een gedegen kennis heeft opgedaan m.b.t. digitalisering en over een infrastructuur beschikt die schaalbaar is. Alle lopende en toekomstige digitaliseringprojecten kunnen probleemloos geïntegreerd worden in de bestaande omgeving voor "preservatie" en "valorisatie" (mits uitbreiding van de server- en storagecapaciteit).

Er is bijzonder veel aandacht besteed aan de toepassing van open standaarden om elke mogelijke afhankelijkheid van derden zoveel als mogelijk te vermijden. Deze investering in tijd rendeert.

De ervaringen m.b.t. uitbesteding van "digitalisering" zijn over het algemeen zeer positief. Het is echter belangrijk om voorafgaand aan elke aanbesteding duidelijke instructies te geven en de objectieven duidelijk te bepalen.

De retroconversie van bestaande metadata (in inventarissen, toegangen,...) en de structurering volgens de hedendaagse normen is zeer arbeidsintensief en vraagt gespecialiseerde kennis. Bij wijze van test werd een kleine hoeveelheid inventarissen uitbesteed voor retroconversie. De invoer van de gegevens (intypen of scannen en OCR) verliep correct. Het structureren van de gegevens volgens de EAD of EAC-standaard stelde echter grote problemen. Er werd voor geopteerd om dit te laten uitvoeren door personeelsleden die een aangepaste opleiding kregen in het Algemeen Rijksarchief.

Het dient herhaald dat kwaliteitscontrole een zeer arbeidsintensief proces is. Niettegenstaande de duidelijke afspraken en de vastlegging van de werkprocessen, dient veel aandacht besteed te worden aan nazorg en kwaliteitscontrole.

Vertragingen in de uitvoering van enkele relatief eenvoudige infrastructuurwerken in het gebouw van het Algemeen Rijksarchief, heeft maandelang een hypotheek gelegd op de voortgang van het globale project. Zowel de uitbouw en de inrichting van het digitaliseringlabo als de beschikbaarstelling van de resultaten van de digitalisering diende meermaals uitgesteld te worden.

Het belang van vakgerelateerde metadatastandaarden wordt vaak onderschat. Het is bijzonder moeilijk om de diverse partners ervan te overtuigen de metadata in een gestandaardiseerd formaat aan te leveren. Ook bij samenwerking met andere instellingen in het kader van "portals" is gebleken dat het niet eenvoudig is om tot een aanvaardbare oplossing voor alle partijen te komen (vb. samenwerking archiefdiensten en Europeana).

De dienstverlening in het Rijksarchief is gericht op een dienstverlening tijdens de werkuren. Online diensten dienen echter 24/24 7/7 beschikbaar te zijn. Er dient beroep gedaan te worden op externe partners om deze diensten te monitoren.

Het vinden van gekwalificeerd personeel voor digitalisering is bijzonder moeilijk zowel voor wat betreft de wetenschappelijke personeelsleden als voor wat betreft de technische personeelsleden. Er bestaat geen echte "vakopleiding" zodat interne opleiding en kennisoverdracht cruciaal is.

De digitale leeszaal is een dienst aan de klant. De verwachtingen van de klanten zijn zeer hoog gespannen. Zij hopen dat de digitale leeszaal een pasklaar antwoord biedt op alle mogelijke vragen. Het is belangrijk om enerzijds zoveel mogelijk in te spelen op hun verwachtingen maar anderzijds ook duidelijk aan te geven welke de technische, praktische, juridische en financiële grenzen zijn van de mogelijkheden.

Om de investeringen op lange termijn te kunnen valoriseren is het belangrijk dat de **continuïteit** kan gegarandeerd worden. In het kader van het project werd geïnvesteerd in de opleiding van personeel. Een aantal personeelsleden werden opgeleid in zeer specifieke taken die belangrijk zijn in het kader van de digitalisering door de wetenschappelijke instellingen. Om de toekomstige projecten efficiënt te kunnen voeren, is het belangrijk deze personeelsleden een perspectief te kunnen bieden.

Het personeel is eveneens nodig om de bestaande projecten blijvend te kunnen exploiteren en valoriseren. Het **onderhoud van de digitale objecten** is een dagelijkse zorg. Ook hier is reeds heel wat kennis en ervaring bij personeelsleden in het kader van het Digit-project. Het onderhoud heeft ook een financieel aspect: de levensduur van informaticamateriaal en toepassingen is beperkt in de tijd. Updates en vernieuwingen zijn noodzakelijk om de resultaten van de reeds uitgevoerde projecten in stand te houden.

De problematiek van de digitale duurzaamheid dient verder onderzocht te worden. Het antwoord op de vraag hoe de wetenschappelijke instellingen de verkregen digitale informatie en objecten op langere termijn dienen te bewaren en toegankelijk houden ("preservation, sustainability and access") vormt een belangrijk criterium in het al dan niet succesvol beschouwen van deze eerste digitaliseringsfase.



## 6. REFERENTIES

Een selectie van publicaties met betrekking tot "digitalisering":

- Artikel: *Anno 2010. Parochieregisters treden het digitale tijdperk in.* (Annelies Coenen en François Baptiste), in "Archief- en Bibliotheekwezen in België. Het congres van Brussel 1910-2010".
- Artikel: *Archiefdiensten: meer dan archiveren!* (Lucie Verachten), in "Archief- en Bibliotheekwezen in België. Het congres van Brussel 1910-2010".
- Artikel: *Digitalisering in het Rijksarchief*, (Science Connection, februari 2010, p.40-43.
- Artikel: *Zoeken naar personen. Evaluatie 50 dagen nieuwe zoekrobot "search.arch.be". 27 juni 2011-15 augustus 2011*
- François Baptiste, Annelies Coenen en Lucie Verachten. *Een bloemlezing uit de parochieregisters. Un florilège des registres paroissiaux.* Brussel, 2011.
- *Het bezoekersprofiel van het Rijksarchief: 2004-2008.* Miscellanea Archivistica 201. Brussel, 2011.
- Lucie Verachten en Karel Velle. DIPHOTA. *Digitizing Personal Photographs for Archiving: Ein Projekt zum Anlâgen von Archiven, zwischen Erwartungen, Chancen und Herausforderungen.* In: Gfornner, Sabine en Follner, Michaela. *Beruf(ung) Archivar: Festschrift für Lorenz Mikoletzky.* Wien, 2011, p. 287-296.
- *Zoekgedrag in digitale archieven. Case study. Notulen van de ministerraad 1918-1979. On-line toegang met full text search.* In: Karel Velle, Claude de Moreau de Gerbehaye en Els Herrebout (eds). *Liber Amicorum Alfred Minke Studia 127.* Brussel, 2011, p. 487-503.
- Lucie Verachten, Annelies Coenen en François Baptiste. *Parish registers. Resource for family history and much more.* Het artikel wordt gepubliceerd in het Genealogisch Jaarboek Lublin (Poland).
- *BHL-Europe - The Biodiversity Heritage Library for Europe*, <http://www.europeana.eu/portal/>.
- F. Boudrez, *Standaarden voor digitale archiefdocumenten*, Antwerpen, 2001-2005. [http://www.edavid.be/docs/eDAVID\\_standaarden.pdf](http://www.edavid.be/docs/eDAVID_standaarden.pdf)
- J.L. Van Goethem & T. Backeljau, 1997, *History and value of the malacological collections of the Royal Belgian Institute of Natural Sciences*, In: J.R Nudds. & C. W. Pettitt (Eds.), *Proceedings of the international Conference, Manchester 1995*, p. 258-262, fig. 1-6.
- L. MEESE, *Deacidifying Books with "CSC Book Saver" A Preservation Project of the Royal Belgian Institute of Natural Sciences (RBINS)*, in *International Preservation News*, 35, may 2005, 4-12.
- Verschillende publicaties in Francia, revue de l'IHA Paris, Tijdschrift van de BVNG, Mededelingenblad SOMA en website SOMA (2012).