

Expertise in digitalisering

→ Met vallen en opstaan

Al enkele jaren zien we de belangstelling voor het verleden in beeld en geluid steeds toenemen. Men heeft interesse voor foto's, postkaarten, films en andere opnamen van de tijd van toen. Veel van deze herinneringen zijn echter slechts voor een kleine kring toegankelijk omdat de opnames enkel in familiale kring bekend zijn, achter slot en grendel bewaard worden, nooit worden getoond. Bovendien verdwijnen deze herinneringen al te vaak omdat de kwaliteit van het materiaal achteruit gaat, of omdat ze worden weggegooid of vernield.

Is het niet jammer dat zoveel verborgen blijft van zowel het dagelijkse leven als die bijzondere, te koesteren, of zelfs te verfoeien dagen? Meer en meer wil de erfgoedsector inspelen op die vraag. En nu zijn er ook meer hulpmiddelen om dit te verwezenlijken. 'WESTHOEK verbeeldt' wil voor de regio net die meerwaarde bieden: een beeld op het dagelijkse leven van een stukje West-Vlaanderen, en dit door de boer op te gaan en beelden uit de privésfeer, maar met een duidelijk gemeenschappelijk belang, breder te ontsluiten door ze te digitaliseren en aan te bieden via de website.

De afgelopen jaren heeft ook de Universiteitsbibliotheek Gent veel werk gemaakt van het opbouwen van expertise in het digitaliseren, ontsluiten en archiveren van delen uit haar omvangrijke collectie. Die bedraagt ca. 48 lopende km. Een vierde is erfgoed, waaronder 380.000 boeken van voor 1900, een collectie ephemera (de Vliegende Bladen) van één miljoen stuks, 15.000 affiches, 5.500 handschriften en noem maar op. Een grote schat aan informatie die soms nauwelijks bekend is, laat staan geraadpleegd wordt. Logisch dus dat de Universiteitsbibliotheek zich ging verdiepen in de digitalisering van sommige werken.

STEEKKAARTENCATALOGUS UNIVERSITEITBIBLIOTHEEK GENT

De eerste stap in het opbouwen van de kennis was het scannen van de steekkaartencatalogus in 2004. Bij het ontstaan van de Universiteitsbibliotheek in 1817 werd de collectie enkel ontsloten via handgeschreven registers. Per thema was er een register waarin de collectie alfabetisch beschreven werd. Pas op het einde van de 19^e eeuw, onder impuls van de hoofdbibliotheecaris Ferdinand Vander Haeghen, werd gestart met een manuele steekkaartencatalogus. Een eeuw later, vanaf 1984, deed de geautomatiseerde catalogus haar intrede. Aangezien de steekkaartencatalogus dan al een omvang van bijna 2.650.000 fiches had, bleek het een onmogelijke taak al deze

fiches retroactief in te voeren. Dus veroorzaakte de invoer van een elektronische catalogus een splitsing in de ontsluiting naar de gebruiker toe. Een uitzondering niet te na gesproken, moest men de steekkaartencatalogus raadplegen om een ouder werk terug te vinden. Dit betekende een fysieke verplaatsing naar de cataloguszaal, de juiste lade zoeken, de lade doorbladeren en (hopelijk, want kennis van het Pruisische classificatiesysteem was een vereiste) de publicatie vinden. Gedurende enkele jaren werd gezocht naar oplossingen om een betere toegankelijkheid mogelijk te maken. De piste van het laten intikken in bijvoorbeeld de Filippijnen werd al snel verlaten. Want hoe moet de gebruiker de oude collectie vinden als de catalogus weg is? En komen de catalogusfiches wel (onbeschadigd) terug? In 2004 werd besloten de steekkaartencatalogus in huis te laten scannen door een gespecialiseerde firma en ze vervolgens beter te ontsluiten met behulp van OCR. Gelijkaardige projecten werden met succes uitgevoerd in de universiteitsbibliotheken van Princeton en Heidelberg, de Koninklijke Bibliotheek te Kopenhagen, de Österreichische Nationalbibliothek te Wenen en de Zentralbibliothek in Zürich. Gemiddeld kon men 30.000 steekkaarten per dag scannen. Dagelijks werden slechts enkele laden uit circulatie genomen en verder ondervond de gebruiker geen hinder van het project. De fiches werden één per één gescand en opgeslagen in twee beeldbestanden, gekleurd en geoptimaliseerd in grijswaarden. Waarom twee beelden? Het ene wordt gebruikt voor de optische herkenning (OCR), omdat men in grijswaarden veel nauwkeuriger de tekst kan herkennen. Het kleurbeeld wordt aan de gebruiker getoond. Die kleur was belangrijk omdat een groene band het bezit van de vroegere seminariebibliotheken van die van de Universiteitsbibliotheek onderscheidt. Op vier en een halve maand werd de klus geklaard. 2.647.069 steekkaarten, verdeeld over ca. 2.675 laden, werden omgezet naar een digitaal beeld, het begin van een digitaal bestaan. De volgende stap was de steekkaarten doorzoekbaar te maken voor de gebruiker. In eerste instantie werd een bladeraar ontwikkeld zodat men virtueel door de catalogus kan bladeren, net zoals je in de fysieke catalogus doet. Ondertussen werden de gescande fiches aan de optische beeldherkenning onderworpen. De gedigitaliseerde tekst is nodig om de steekkaarten doorzoekbaar te maken. Het is deze tekst die door de gezamenlijke zoekinterface doorheen de geautomatiseerde catalogus Aleph en de gedigitaliseerde steekkaartencatalogus, Cat-Fich, wordt doorzocht. Op deze manier hoeft de gebruiker niet meer van zijn computerstoel te komen om het bezit van de

UGentbibliotheken te doorzoeken². Hoewel dit project heel eenduidige tekstfiches behandelde, was er toch heel wat uit te leren zoals het belang van een goede voorbereiding, een vlotte workflow en een werkbare ontsluiting. In de voorbereiding wordt het doel van de digitalisering bepaald en hier is het online ontsluiten van de steekkaarten het prioritaire doel. Uit die doelstelling volgen de beslissingen die in de verdere *workflow* worden genomen, zoals de wijze van scannen (bitdiepte, resolutie, kleur), de eikpunten van de kwaliteitscontrole en de manier van ontsluiten.



Steekkaarten op onlinecatalogus van bibliotheek UG

RECOLLECTING LANDSCAPES

In een volgend project, *Recollecting Landscapes*³, werd de vakgroep Architectuur en Stedenbouw – in een onderzoek naar de transformatie van het landschap en de verstedelijking – door het team van de Universiteitsbibliotheek ondersteund bij de digitalisering van een collectie foto's en het aanbieden van een gebruiksvriendelijke user interface. Ondersteunen van onderzoek is immers een kerntaak van de Universiteitsbibliotheek en bovendien kadert dit in de Architectuurfocus, een vijf jaar durende samenwerking met de UGentvakgroep Architectuur en Stedenbouw⁴.

Begin vorige eeuw trok de ULB-professor Jean Massart Vlaanderen rond om landschappen te fotograferen. Als bioloog probeerde hij vooral de plantengroei vast te leggen en doordat hij ook een uitstekend fotograaf was, resulteerde dit in een schitterende fotoreeks. In 1980 gaf de Nationale Plantentuin een tweede fotograaf, Georges Charlier, de opdracht om eenzelfde plekjes van Massart opnieuw te fotograferen. Dat kon doordat Massart zijn coördinaten en kijkhoek genoteerd had. In 2003-2004 deed Jan Kempenaers nog eens hetzelfde, nu in

opdracht van het Vlaams Architectuurinstituut en geïnspireerd door de vakgroep Architectuur en Stedenbouw van UGent. Het onderzoek resulteerde in een tentoonstelling in het SMAK met daaraan verbonden een boek en een website. Zowel de bewaarbibliotheek als de digitale bibliotheek werden heel nauw bij het onderzoek betrokken. De ene voor de betere bewaring van de foto's, waarvan één reeks op stevig karton met nog steviger lijm gekleefd was, en voor het beschikbaar stellen van het materiaal in de eerste onderzoeksfase. De andere vond in de tweede fase in deze collectie een echt testbed. De kernvragen die in dit project centraal stonden, betreffen de manier van scannen (de scans waren niet enkel bestemd voor ontsluiting maar ook voor archivering), de problematiek van auteursrechten naar ontsluiting in de tentoonstelling, het boek en de website toe, en de vraag hoe de informatie en de beelden overzichtelijk te presenteren in een digitale omgeving. Het resultaat is te zien in een overzichtelijke website, met een speciaal voor de gelegenheid ontworpen 'timescape' waarbij de beelden van de drie periodes in elkaar overvloeien.



Recollecting Landscapes

TOPOGRAFISCHE COLLECTIE

Deze twee projecten deden de nood aan een volwaardige beeldbibliotheek nog duidelijker voelen. Een beeldbibliotheek die de waardevolle documenten van de Universiteitsbibliotheek op eenzelfde manier kan aanbieden. Dus werd het tijd voor het grotere werk: het digitaliseren en ontsluiten van de topografi-

sche collectie. Deze verzameling van 40.000 afbeeldingen van gebouwen en landschappen in België uit de periode 1875 – 1925 kwam in het bezit van de Universiteitsbibliotheek via een schenking van het HIKO in 2000. Met dit project stonden ons verschillende doelen voor ogen: het beter conserveren van de collectie (afstoffen en in zuurvrije mappen opbergen), het digitaliseren, het realiseren van een efficiënte werkomgeving voor de behandeling van digitaal materiaal, het toevoegen van de juiste beschrijvende metadata (gegevens over de beelden om op te kunnen zoeken) en het aanbieden via een aantrekkelijke webinterface. Het scannen van de collectie gebeurde in huis door jobstudenten en tijdelijke werkrachten (onder supervisie) met een A2-flatbedscanner. Vrij snel kwam er een groot probleem aan de oppervlakte: de opslag. Elk beeld is immers 30 à 40 Mbyte groot en de hele collectie neemt al snel 1.2 Terabyte in beslag. De trefwoorden en plaatsnamen zijn Franstalig en werden ingevoerd in het bibliotheekstelsel Aleph. Ze werden vertaald naar het Nederlands en het Engels; de plaatsnamen werden geüniformiseerd op basis van het huidige postnummer. In de tweede fase, die nu nog loopt, worden de beelden een per een gedetailleerder beschreven. Dit is nog steeds noodzakelijk want goede beschrijvingen maken beelden vindbaar. Voor het beheer van de beelden werd een DAM (*Digital Asset Management*) aangekocht, een software die niet alleen toelaat beelden te beschrijven, toe te voegen, te verwijderen en te publiceren, maar ook zeer snel diep in te zoomen in de beelden zodat je elk detail kunt vergroten. Een ander belangrijk aspect van de beelddatabank is de presentatie van het geheel van gescande beelden en de bijhorende beschrijvingen op het web met verschillende zoekingen. De eerste is een grafische zoekactie. Op basis van de postnummers werd een kaartje van België gemaakt waarop je een gemeente kunt aanklikken om er de beelden van te zien. Verder kun je zoeken door een plaats (regio – gemeente – deelgemeente) en een type gebouw te combineren. Ten slotte kun je ook vrij zoeken op alle woorden die in de beschrijving zijn opgenomen. De originele foto's zijn nu veilig opgeborgen en hun digitale beeld is 24 uur op 24 en 7 dagen op 7 vrij toegankelijk: voor onderzoek en onderwijs, maar ook voor het grote publiek.

ANTIFONARIUM TSGROOTEN

Nu hadden we de smaak te pakken. Er was echter nog veel werk aan de winkel. De vorige projecten hadden ons geleerd dat er een grote nood was aan een meer gespecialiseerde aanpak. Niet alleen gingen we op zoek naar professionele scanapparatuur, met de hulp van een expert, maar ook naar gekwalificeerd personeel, mensen met een technische kennis en een fotografisch oog voor detail. Het heeft wat voeten in de aarde gehad, maar er werden budgetten gevonden om een scanatelier op te bouwen met een flexibele A0-scanner, een A2-book-

scanner, een A2-flatbedscanner en enkele eenvoudige A4- en doorvoerscanners, alsook kalibreerbare beeldschermen en printers. Ondertussen werkt 1,5 VTE in dat scanatelier. Zij behandelen, naast de eigen digitaliseringsprojecten, de vele aanvragen van studenten, onderzoekers en het grote publiek. Daarnaast werd er werk gemaakt van een gedegen *workflow*. Daarvoor zijn we te rade gegaan bij enkele voorbeelden zoals NARA⁵, Het Geheugen van Nederland⁶, TASI⁷ en andere. De *workflow* werd ondertussen met vallen en opstaan geïmplementeerd. Een grote nadruk ligt op het voortraject, omdat daar de beslissingen over de stappen van de te volgen *workflow* worden genomen, zoals het formaat waarin zal worden gescand en volgens welke standaarden, welke metadata van belang zijn, welke kwaliteitscriteria moeten worden gevolgd, hoe het digitale beeld en de bijhorende informatie zal worden opgeslagen, en ten slotte op welke manier het geheel zal worden ontsloten.

De ervaring die werd opgebouwd in deze projecten, resulteerde in de supersnelle digitalisering van het *Antifonarium*. Het handschrift, aangekocht door Vlaams minister Bert Anciaux voor de Vlaamse Kunstcollectie, werd in bruikleen gegeven aan de Vlaamse Erfgoedbibliotheek. Het werd in amper een week tijd gedigitaliseerd en via het internet voor het brede publiek ter beschikking gesteld. De Universiteit Gent wijdde een website aan dit nieuwe topstuk: www.antifonarium-tsgrooten.be. Het *Antifonarium* zal in de komende jaren in de Universiteitsbibliotheek bewaard worden en daarna verhuist het naar andere partners binnen de Vlaamse Erfgoedbibliotheek.



Antifonarium Tsgrooten online

VAN BOM-VL TOT WESTHOEK VERBEELDT

De focus bij de Universiteitsbibliotheek ligt nu op archivering en het uitbreiden van de kennis naar andere te digitaliseren materialen. Om die reden participeert ze ook aan het project BOM-VI. BOM-VI⁸ wil komen tot één groot digitaal depot waarin al het Vlaamse multimediale materiaal veilig gestockeerd zit. Het is een bewaarproject, waarin alle problemen rond digitale preservatie, veroudering van dragers, kwaliteitscontrole, rechtenbeheer, selectiecriteria, enzovoort, centraal staan. Aan culturele zijde participeren VTI, FARO, Muziekcentrum Vlaanderen, BAM en de Boekentoren (Universiteitsbibliotheek Gent). Uit de mediasector participeren de openbare, commerciële en regionale omroepen. De IBBT-onderzoeksgroepen en het VRT Medialab zorgen voor de wetenschappelijke omkadering.

Het project wil niet alleen onderzoek verrichten, kennis en kunde verzamelen en de resultaten ervan vrij beschikbaar maken. Het wil tevens op een gemeenschappelijk innovatieplatform concreet demonstreren hoe een gemeenschappelijke inspanning tot tastbare resultaten en efficiëntiewinst kan leiden. Het project zal tevens initiatieven nemen om de verworven kennis actief te verspreiden bij alle relevante instanties, binnen en buiten het project. Om de verschillende factoren met betrekking tot de preservatie van het digitaal erfgoed enigszins te structureren binnen het project, werd voor een onderverdeling in zes werkpakketten geopteerd. In het werkpakket 'gebruikersnoden', WP1, wordt onderzocht wat de verlangens of eisen zijn van de respectieve gebruikers. WP2, 'archivering en selectie', gaat in op de criteria en overwegingen die meespelen om documenten al dan niet digitaal te bewaren. Vervolgens staat in WP3 de opstelling van een gelaagd metadata-model centraal, waarin de digitale objecten met hun metadata zullen worden ondergebracht. WP4 houdt zich bezig met het 'beheer van rechten op digitale objecten' en WP5, 'ar-

chitectuur voor digitale bewaring en ontsluiting', onderzoekt mogelijke ontsluitingsmodellen voor specifieke doelgroepen. In WP6 ten slotte wordt een demonstrator opgesteld als opslagprototype en dit met behulp van de bevindingen en onderzoeksresultaten uit de vorige werkpakketten.

Uiteraard kunnen beeldbanken als 'WESTHOEK verbeeldt' heel wat voordeel halen uit dergelijke projecten. Zo heeft 'WESTHOEK verbeeldt' doorheen het project keuzes moeten maken. Er werd bijvoorbeeld voor gekozen om vrijwilligers metadata te laten toekennen, wat een enorme verrijking van de beeldbank betekent. Het is wel aangeraden bij te houden welke metadata door wie zijn ingevoerd en of ze gecontroleerd werden. Voor het toekomstige gebruik van de metadata is dat belangrijke informatie. Wat de ontsluiting betreft werd ervoor gekozen om met een commerciële partner in zee te gaan, Memorix, en een kant-en-klare oplossing te gebruiken. Een pragmatische beslissing met een mooi resultaat. De enthousiaste reacties van de gebruikers onderstrepen dit nog eens.

Hoe zit het echter met de technische openheid van het platform en de opslag op lange(re) termijn? Daar bieden deze oplossingen niet steeds een antwoord op, en dat wordt ook niet onmiddellijk verwacht. Er zijn wel steeds nieuwe inzichten, die een verandering van *workflow* tot gevolg hebben of een andere aanpak van de digitalisering en de metadatering. Een initiatief als BOM-VI kan dan heel wat ondersteuning en expertise bieden aan digitaliseringsprojecten, maar moet daarnaast ook kunnen inspelen op keuzes die in vroegere projecten werden gemaakt. Heel belangrijk is dat een initiatief als BOM-VI aan deze projecten een langetermijnoplossing kan bieden waar dat bij lokale instellingen niet steeds het geval is. Hoe dit na het beëindigen van het project zal worden verder gezet, zal dan ook met argusogen worden gevolgd.

1 Inge van Nieuwerburgh is coördinator digitale bibliotheek aan de Universiteitsbibliotheek UGent.

2 Zie: <http://lib.ugent.be>.

3 Zie: www.recollectinglandscapes.be.

4 Zie: www.architectuurfocus.ugent.be.

5 Technical Guidelines for Digitizing Archival Materials for Electronic Access: Creation of Production Master Files – Raster Images; NARA; juni 2004.

6 Richtlijnen en procedures voor de uitvoering van projecten in het kader van 'Het Geheugen van Nederland'; Projectbureau Het Geheugen van Nederland, maart 2006

7 TASI, Still images, moving images and sound advice, a JISC Advisory Service; zie: www.tasi.ac.uk.

8 Zie: www.bom-vl.be. Zie ook de eerdere bijdrage in dit tijdschrift: Brecht DECLERCQ, Hélène VERREYKE & Jeroen WALTERUS, 'Audiovisueel erfgoed. Het bewaren en ontsluiten van multimediale data in Vlaanderen', *faro | tijdschrift over cultureel erfgoed*, 1 (2008) 4, p.16-23