

Digitaal archiefbeheer en het ADVN

Het is bekend dat de samenleving in hoog tempo en exponentieel digitaliseert. Een steeds groter deel van onze culturele erfenis bestaat uit niet-tastbare gegevens, uit bits en bytes. Digitaal betekent niet alleen efficiënt, economisch en gebruiksvriendelijk, maar ook zeer kwetsbaar voor de tand des tijds. Pessimistische historici en archivariissen vrezen nu al voor de digitale middeleeuwen die zouden ontstaan door de onherroepelijke verdwijning van digitaal erfgoed. Een oplossing in de vorm van een doordacht digitaal archiefbeheer is dan ook een absolute noodzaak.

Het begrip 'digitaal archiefbeheer' kent verschillende invullingen die telkens een stadium in de evolutie ervan weergeven. De oudste en eenvoudigste vorm bestaat uit de digitalisering of automatisering van de werkprocessen die bij archiefontsluiting komen kijken. Een voorbeeld daarvan is de transformatie van gegevens uit een steekkaartenbak naar een catalogus die online beschikbaar is. Een volgend stadium in de evolutie van het digitale archiefbeheer is de digitalisering van archief. De omzetting van archief naar digitaal formaat wordt niet alleen gebruikt om kwetsbare documenten te beschermen, maar ook om de raadpleging te optimaliseren. In deze bijdrage gaan we vooral in op de meest actuele vorm van digitaal archiefbeheer, namelijk de bewaring en ontsluiting van zogenaamd *digital born* archief. Dit zijn archiefstukken die van bij hun ontstaan een digitaal karakter hebben. Het gaat onder meer om digitale tekstbestanden, rekenbladen, afbeeldingen, databanken, websites en e-mail. Archivariissen vragen zich af wat daar binnen enkele decennia

nog van overblijft. Kunnen we ons digitale erfgoed behoeden voor een historische vergeetput?

Problemen bij het beheer

De grootste bedreiging voor het digitale erfgoed bestaat uit de technologische vooruitgang zelf. Denk maar aan de evolutie die hardware en software in de voorbije kwarteeuw hebben ondergaan. Opslagmedia die twintig of zelfs tien jaar terug nog courant waren, zijn vandaag onleesbaar geworden. Bestandsformaten van toen zijn niet meer compatibel met de huidige computerprogramma's – als het programma zélf nog bestaat. De historische of juridische waarde van een digitaal document is dan ook bijna altijd langer dan de levensduur van de digitale omgeving. Nemen we ter verduidelijking een Vlaamse vereniging als fictief voorbeeld. De secretaris en penningmeester hielden in de jaren 1980 en 1990 de boekhouding en de ledenlijsten bij aan de hand van computerprogramma XYZ. De gegevens registreerden ze op 5,25" diskettes en later op cd-r. Het bestuur van de vereniging besluit anno 2008 zijn archief over te dragen aan een archiefinstelling. Behalve een dozijn archiefdozen met bestuursverslagen en activiteitendossiers bestaat het verenigingsarchief ook uit een doos met allerhande diskettes en cd's. Ondanks de voorspelde eeuwige duurzaamheid is een aantal cd's onbruikbaar geworden. De weinige bestanden die de archivaris wel nog kan recupereren, zijn wegens de afwezigheid van een compatibel computerprogramma onleesbaar.

Technologische veroudering is echter niet het enige probleem bij het beheer van digitaal archief. Zo brengt de exponentieel groeiende hoeveelheid digitale documenten een veel complexere selectie ter bewaring of vernietiging met zich mee. De onoverzichtelijke hoeveelheid aan tekstbestanden en e-mails verhoogt namelijk het risico op het verlies van waardevolle documenten. Verder vereisen de verscheidene types van digitale objecten, van rekenbladen tot websites, ook specifieke archiveringsoplossingen. En de problemen houden niet op bij de bewaring van het bestand, ook de authenticiteit en de ontstaanscontext moeten bewaard blijven. Ten slotte is er nog de blijvende toegankelijkheid voor derden na bewaring. Digitaal archiefbeheer vereist dus een doordachte bewaarstrategie.

Digitale bewaring

De afgelopen vijftien jaar werden tal van bewaarstrategieën gelanceerd. De meest eenvoudige aanpak is de omgekeerde digitalisering of het overzetten van digitale documenten naar microfilm of afdruk op papier. Elk te bewaren document of e-mailbericht afdrukken is



echter niet alleen onuitvoerbaar, het druist ook in tegen een belangrijk archiefprincipe, namelijk dat archiefstukken in hun primaire vorm worden bewaard. Een andere voorbijgestreefde oplossing is de ontwikkeling van een computermuseum waar de originele hard- en software bewaard blijven. Hoewel dit idee een optie kan zijn voor een korte termijnbewaring, is de instandhouding van oude computerconfiguraties alleen al vanwege het ontbreken van wisselstukken en productondersteuning onbegonnen werk. Een derde bewaarstrategie bestaat uit de continue conversie van bestanden van een lagere naar een hogere versie. Ook deze aanpak kan op korte termijn nuttig zijn. Het grootste nadeel is echter de blijvende afhankelijkheid van het programma waarbij documenten op langere termijn alsnog onleesbaar kunnen worden door hun specifieke formaat.

De meest gangbare manieren om digitale archiefdocumenten te bewaren zijn emulatie en migratie. Bij de emulatiestrategie kunnen de documenten in hun oorspronkelijk formaat worden geraadpleegd. De originele computeromgeving wordt dan virtueel nagebootst op een actueel computersysteem. Bij migratie worden de archiefdocumenten omgezet naar een gestandaardiseerd archiveringsformaat zoals XML of PDF/A. Beide strategieën hebben hun voor- en nadelen. De migratiestrategie maakt het document platformafhankelijk, maar is wel sterk afhankelijk van het als standaard erkende archiveringsformaat en kent een risico op verlies van authenticiteit. Bij emulatie kan de authenticiteit gegarandeerd worden, omdat de documenten in hun oorspronkelijk formaat gearchiveerd worden. Het grote nadeel van emulatie is een nieuwe platformafhankelijkheid en de onzekerheid rond de continuïteit ervan.

In een onderzoeksrapport van het David-project (Digitaal Archiveren in Vlaamse Instellingen en Diensten)¹ wordt een middenweg tussen emulatie en migratie voorgesteld. Deze combinatiestrategie houdt in dat er twee representaties van het document worden bewaard: het oorspronkelijke digitale archiefdocument en een naar een archiveringsformaat geëmigreerde versie. Door de bewaring van het oorspronkelijke document blijven vele opties open, terwijl tegelijkertijd door de migratie in een lange termijnoplossing wordt voorzien. De huidige strategie van het ADVN volgt die gulden middenweg. Om de raadpleging van de digitale archiefdocumenten te vereenvoudigen wordt buitendien nog een derde representatie van het bestand



aangemaakt. In de praktijk wordt een bestand dus in drie vormen bewaard: het originele bestand (verscheidene formaten), het archiveringsbestand (bv. in PDF/A of XML voor kantoordocumenten en e-mail, TIFF voor afbeeldingen) en het consulteringsbestand (bv. PDF voor kantoordocumenten en JPEG voor afbeeldingen). Het ADVN hanteert deze strategie in een eenvoudig bestandssysteem, maar volgt aandachtig de actuele evolutie in de ontwikkeling van digitale depots.

Actieve verwerving

In het digitale depot wordt niet alleen overgedragen digitaal archiefmateriaal opgeslagen. Het ADVN onderzoekt ook hoe een aantal digitale bronnen, zoals digitale periodieken en websites, actief kunnen worden verworven. Sinds geruime

tijd neemt het ADVN digitale nieuwsbrieven en tijdschriften in haar collectie op. In een voorbereidend stadium werd daarbij bijzondere aandacht besteed aan het in kaart brengen van de verscheidene technische formaten van het aanbod. In een volgende fase werden diverse obstakels uit de weg geruimd die zich voordoen bij het verzamelen, bewaren en ontsluiten van dit type periodieken.

De ontsluiting van digitale archiefdocumenten gebeurt via databestanden die op het Filemakerplatform werden ontwikkeld. Bij het ontwerp van die bestanden is rekening gehouden met de gangbare internationale standaarden. Zo is het bestand 'Archievenoverzicht' conform ISAD(G) en vormen de richtlijnen van het expertisecentrum David een leidraad voor de beschrijving van de digitale bestanden. Hoewel het eigen automatiseringssysteem nog steeds significante voordelen biedt zoals een flexibele architectuur en een performante ontsluiting van de collecties in de leeszaal, onderzoekt het ADVN ook de mogelijkheden van andere geïntegreerde ontsluitingsystemen. De eventuele implementering ervan dient de ontsluiting compatibeler te maken met andere gegevensbestanden en in te schakelen in (inter)nationale netwerken.

Digitaal archiefbeheer is een zeer complexe materie die heel wat bijkomende inspanningen vergt, zowel inzake personeelsinzet, technische omgeving als financiële middelen. Permanente opvolging is een must. [rc]

(1) Alle rapporten, adviezen en richtlijnen van het David- en cDavid-project zijn na te lezen op de website van het expertisecentrum David (<http://www.edavid.be>).