

# Is een DMS bruikbaar als digitaal depot?

FILIP BOUDREZ  
EXPERTISECENTRUM DAVID VZW  
ANTWERPEN, 2013

## 0. INHOUD

---

1. Inleiding.....	1
2. Overzicht.....	2
3. De verschillen.....	3
3.1 Finaliteit.....	3
3.2 Documenttypes.....	3
3.3 Informatiemodel.....	4
3.4 Metadata.....	5
3.4.1 Hiërarchie.....	5
3.4.2 Technische metadata.....	5
3.4.3 Procesmetadata.....	5
3.4.4 Opslag.....	6
3.5 Systeemafhankelijkheid.....	6
3.6 Preservering.....	7
3.7 Zoeken.....	7
3.8 Volume.....	8
3.9 Operaties.....	9
3.10 Rechtenbeheer.....	9
4. Besluit.....	9
5. Meer informatie.....	10

## 1. INLEIDING

---

Voor het beheren en het archiveren van digitale documenten zijn diverse soorten digitale systemen beschikbaar. Een Document Management Systeem (DMS), een Records Management Applicatie (RMA) en een digitaal depot zijn wellicht de bekendste voorbeelden. Al deze systemen bieden functionaliteiten aan om digitale archiefdocumenten te beheren. Bij het zoeken naar archiveringsoplossingen voor digitale documenten stellen veel organisaties zich de vraag of één van deze systemen als archiveringsoplossing kan volstaan en in het bijzonder of een DMS-toepassing niet (her)bruikbaar is als archiveringsplatform voor documenten met een middellange of permanente bewaartermijn.

Deze bijdrage wil een antwoord formuleren op deze vragen door tien verschillen tussen een DMS en een digitaal depot op te sommen en toe te lichten. Deze verschillen hebben de typische functionaliteiten en kenmerken van een digitaal depot als uitgangspunt. De ISO-norm OAIS (Open Archival Information System) wordt hierbij als referentiekader gebruikt<sup>1</sup>. Deze bijdrage gaat ook bewust uit van de vergelijking DMS versus digitaal depot, en niet van de vergelijking RMA versus digitaal depot. Door de convergenties tussen DMS-toepassingen en Records Management Applicaties (RMA) is er de facto amper nog sprake van RMA's als afzonderlijke systemen. De meeste DMS-leveranciers bieden immers optionele modules voor records management<sup>2</sup> binnen hun DMS-platformen aan.

---

<sup>1</sup> OAIS is een referentiemodel voor de langetermijnarchivering van (digitale) informatie. Kernelementen van dit referentiemodel zijn de verantwoordelijkheden waaraan een OAIS moet voldoen, het informatiemodel voor de gearchiveerde informatie en de functies.

<sup>2</sup> Kernfunctionaliteiten van deze modules voor records management zijn het vastleggen van onwijzigbare documenten en het toepassen van bewaartermijnen en bestemmingen (bewaren of vernietigen). Ook digitale depots kunnen deze records management services aanbieden en kunnen documenten opnemen die op (middellange) termijn te vernietigen zijn.

In vergelijking met digitale depots bestaan DMS-toepassingen al veel langer en zijn ze ook wijder verspreid. Er is een brede waaier van zowel commerciële als open source DMS-pakketten op de markt beschikbaar. Voorbeelden van DMS-pakketten zijn Documentum, Corsa, Decos, OpenText, FileNet, enzovoort. Ook samenwerkingsplatformen zoals MS SharePoint en Alfresco bieden steeds meer DMS-functionaliteiten aan. Voor de vergelijking wordt uitgegaan van de doorsnee 'out of the box' functionaliteiten van een typisch DMS. Hierbij is geen rekening gehouden met pakkeetspecifieke functionaliteiten of de mogelijkheid tot extra maatwerk om ontbrekende functionaliteiten toe te voegen

## 2. OVERZICHT

De onderstaande tabel biedt een overzicht van de tien verschillen tussen een DMS en een digitaal depot.

Verschil	DMS	Digitaal Depot
FINALITEIT	aanmaken, beheren en delen van documenten	(middel)langetermijnarchivering van digitaal archief
DOCUMENTTYPES	kantoordocumenten	kantoordocumenten, websites, databases, video, audio, masterscans, film, tekeningen, enz.
INFORMATIEMODEL	digitale documenten en digitale content als informatie-eenheden	<i>information packages</i> als informatie-eenheden
METADATA	metadata over documenten/mappen	metadata over documenten/mappen, werkprocessen, <i>representation information, fixity information</i> , preserveringsacties
SYSTEEM-AFHANKELIJKHEID	stelselafhankelijke opslag van metadata, documenten en content	stelselafhankelijke opslag van <i>information packages</i>
PRESERVERING	passief	actief
ZOEKEN	op inhoud en metadata van documenten en content	op de context en metadata van de archiefdocumenten
VOLUME	max. een tiental miljoen objecten, objecten van kleine tot middelgrote omvang	enkele tientallen miljoenen objecten, objecten variërend van kleine tot heel grote omvang
OPERATIES	individueel document/map	bulkverwerkingen
RECHTENBEHEER	lees- en schrijfrechten voor interne gebruikers, doorgaans rol- en groepsgebaseerd	onderscheid tussen openbaar en niet-openbaar voor interne en externe gebruikers, auteursrechten

## 3. DE VERSCHILLEN

---

### 3.1 Finaliteit

Een DMS is gericht op het aanmaken, het beheren, het zoeken en het delen van digitale documenten. Typische DMS-functionaliteiten zijn dan ook het registreren van metadata, het (geautomatiseerd) toepassen van versiebeheer, het reserveren van documenten voor bewerking (check in – check out), het fijnmazig beveiligen van documenten(versies), het auditen van raadpleging en editeren, enzovoort. Een DMS ondersteunt het digitaal werken en digitaal documentbeheer. Een DMS bevat zowel documentatie als archief en zowel te vernietigen als te bewaren documenten.

De voorbije jaren legden DMS-producten zich ook meer toe op workflow, documentgeneratie en content management. DMS-producenten noemen hun product dan ook steeds meer Enterprise Content Management (ECM)- systemen. De workflow stuurt de werkprocessen aan en ondersteunt het geautomatiseerd aanmaken van documenten. Op basis van de aanwezige content en procesinformatie wordt 'on the fly' een document gegenereerd. Diverse content items zijn de bouwstenen van een document. Het wijzigen van één content item resulteert in een nieuw document of in een nieuwe versie, zonder dat hierbij noodzakelijk de vorige documentversie werd vastgelegd of reconstrueerbaar blijft. De reconstructie van een welbepaalde documentversie is enkel mogelijk wanneer alle bouwstenen (incl. sjablonen of formulieren) achteraf nog beschikbaar zijn.

Een digitaal depot heeft de reconstructie van digitale of gedigitaliseerde archiefdocumenten op (middel)lange termijn als finaliteit. De archiefdocumenten hoeven niet meer aangemaakt, gegenereerd of geëditteerd te worden. De documenten zijn vastgelegd: de dossiers zijn gevormd en afgesloten. Het bewaren van de mogelijkheid tot reconstructie en interpretatie houdt meer in dan enkel de status van een document te wijzigen en het een 'archief' of 'read-only' kenmerk te geven. Een digitaal depot heeft de langetermijnarchivering van digitale archiefdocumenten ten behoeve van één of meerdere doelgroepen als finaliteit<sup>3</sup>. Dit omvat niet alleen het bewaren van de leesbaarheid van digitale documenten, maar heeft ook betrekking op het bewaren van de leesbaarheid van de digitale metadata en van de relatie tussen documenten en hun metadata, zodat alle doelgroepen de bewaarde digitale archiefdocumenten autonoom kunnen begrijpen. De archiefdocumenten moeten te allen tijde reconstrueerbaar en begrijpbaar blijven. Om dat mogelijk te maken moeten alle componenten die samen een *information package* vormen systeemonafhankelijk reconstrueerbaar zijn.

### 3.2 Documenttypes

Een DMS is hoofdzakelijk gericht op het aanmaken en het beheren van de gangbare digitale kantoordocumenten zoals tekstdocumenten, e-mails, presentaties, spreadsheets, foto's, ... en hun metadata. Daarnaast bevat een DMS ook nog veel andere types objecten die het digitaal werken en digitaal dossierbeheer ondersteunen (bijv. agenda's, taaklijstjes, favorietenlijstjes, notificaties, ...). Een DMS ondersteunt hiervoor meerdere contenttypes. Bij de meeste DMS-toepassingen wordt een contenttype voor een digitaal document doorgaans gelijk gesteld aan een bestandsformaat, wat dan weer niet hetzelfde is als een documenttype of redactionele vorm.

Een DMS kan doorgaans minder goed overweg met onderling gelinkte bestanden die samen één document vormen of met grote bestanden. Om performantieredenen wordt in de meeste DMS-implementaties een maximaal toegelaten bestandsgrootte gehanteerd. Het opladen en het indexeren van grotere bestanden vergt (te) veel resources van de webserver en de DMS-toepassing. Een veel gehanteerde restrictie is 50 of 100 MB. Onderling gelinkte bestanden die niet als een compound bestandstype<sup>4</sup> worden opgeslagen, kunnen doorgaans niet met de standaardfunctionaliteiten van een

<sup>3</sup> Dit is één van de zes verantwoordelijken die de ISO-norm *Open Archival Information System* voorschrijft (*Reference model for an open archival information system (OAIS). Recommend practice, Magenta Book, juni 2012, p. 3-1.*)

<sup>4</sup> Een compound bestand is een computerbestand waarin meerdere bestanden en streams als één

DMS worden opgenomen en beheerd. Voorbeelden hiervan zijn gelinkte Excel-spreadsheets, Shapefiles (GIS-informatie), InDesign- of QuarkXpress-bestanden (DeskTopPublishing). Dergelijke bestanden vragen speciale connectoren tussen de clientsoftware en het DMS, of bijzondere oplossingen zoals het werken met virtuele documenten of contentless objecten. Dit laatste houdt dan weer in dat enkel de metadata en niet de bestanden zelf in het DMS worden bewaard.

Digitale archiefdocumenten kunnen om het even welke documentvorm, grootte of bestandsformaat hebben. Documentvormen en opslagformaten zijn geen criteria voor hun archiefstatus. Het documenttype van digitale archiefdocumenten kan bijgevolg heel divers zijn: elke vorm van digitale informatie met archiefstatus moet op een efficiënte wijze in het digitale depot kunnen worden opgenomen. Naast de hier boven opgesomde digitale kantoordocumenten dienen bijvoorbeeld ook hoge resolutiescans, websites, databases, audio, video, GIS-informatie, CAD, DVD's, enz. in een digitaal depot archiveerbaar te zijn. Zelfs computerprogramma's met archiefstatus moeten in het digitale depot kunnen worden opgenomen. Een concreet voorbeeld van dit laatste zijn enkele computerprogramma's die werden ontwikkeld door de jeugddienst van de stad Antwerpen binnen de context van projecten van jeugdpeilingen<sup>5</sup>.

### 3.3 Informatiemodel

Binnen het informatiemodel van DMS'-en vormen de documenten en de atomen waaruit ze bestaan de informatie-eenheden. De meeste DMS'-en zijn object georiënteerd en hanteren hierbij een model waarbij één digitaal object of één computerbestand met één document overeenkomt. Het uitgangspunt is doorgaans een één-op-één relatie tussen een digitaal object en een digitaal document, wat samenhangt met het feit dat een DMS zich hoofdzakelijk op gewone kantoordocumenten richt. In het beste geval kunnen door middel van virtuele mappen of DMS-specifieke objecttypes wel metadata worden gedeeld tussen de verschillende digitale objecten die samen één digitaal document vormen, maar zullen bijv. onderlinge links van zogenaamde 'compound objects' niet werken<sup>6</sup> tenzij alle digitale objecten één voor één buiten het DMS zijn gebracht of geëxporteerd. Dergelijke complexe digitale documenten zijn dus niet rechtstreeks raadpleegbaar in het DMS, maar moeten voor consultatie eerst geëxporteerd worden. Een aantal DMS-toepassingen kunnen deze documenten ook niet als geheel opnemen, tenzij ze eerst worden verpakt als ZIP-bestanden. Een voorbeeld van dit laatste zijn Shapefiles: dit is een set van minimaal drie digitale objecten die samen één document met geografische informatie bevatten en die voor raadpleging in dezelfde map moeten staan. Een ander voorbeeld zijn websites of onderling gelinkte DeskTopPublishingbestanden (bijv. InDesign-bestand + de afzonderlijk opgeslagen afbeeldingen).

Het uitgangspunt van een digitaal depot is niet de opslag van documenten, maar de opslag van *information packages*. De *information packages* voor de digitale archiefdocumenten kunnen heel diverse en complexe vormen aannemen. Ze bestaan niet alleen uit meer dan één informatieobject, maar in het informatiemodel van een digitaal depot moet er ook plaats zijn voor de opslag van meerdere representaties of manifestaties van hetzelfde document. Eén van de vereisten van een *trusted digital repository* is de duurzame opslag van documenten in de vorm waarin ze initieel werden opgenomen<sup>7</sup>. Aangezien dit formaat niet altijd een geschikt archiveringsformaat is of blijft, moet de mogelijkheid voorzien worden om documenten ook in een tweede formaat of in het archiveringsformaat te bewaren. Elke representatie deelt dezelfde contextuele en beschrijvende metadata, maar heeft zijn eigen metadata met technische kenmerken en preserveringsacties.

---

digitaal object worden opgeslagen.

<sup>5</sup> Bijv. [http://zoeken.felixarchief.be/zHome/Home.aspx?id\\_isad=136649](http://zoeken.felixarchief.be/zHome/Home.aspx?id_isad=136649) en [http://zoeken.felixarchief.be/zHome/Home.aspx?id\\_isad=136650](http://zoeken.felixarchief.be/zHome/Home.aspx?id_isad=136650)

<sup>6</sup> Het eigen DMS-specifieke linkmechanisme is hier de oorzaak van. DMS-links werken op basis van ID's en niet met de paden van mappen en/of documenten waar naar wordt verwezen.

<sup>7</sup> *Trustworthy Repositories Audit & Certification (TRAC)*, 2007, B2, p. 25.

## 3.4 Metadata

Op het vlak van metadata zijn er meerdere verschillen tussen een DMS en een digitaal depot:

- hiërarchie
- technische metadata
- procesmetadata
- opslag

### 3.4.1 Hiërarchie

Een belangrijk uitgangspunt in een digitaal depot is de ondersteuning van hiërarchie in het metadatamodel. Dat is noodzakelijk vanwege de diverse aggregatieniveaus die in een archief worden onderscheiden en omdat archieven hiërarchisch worden beschreven. Niet elke DMS-toepassing ondersteunt echter hiërarchie in metadata of het overerven van metadata. In een DMS worden metadata voornamelijk op documentniveau vastgelegd en gehanteerd.

### 3.4.2 Technische metadata

Een DMS-toepassing houdt standaard maar een beperkte set aan technische metadata over de digitale objecten bij. In de regel blijft dit beperkt tot een identificatie van het formaat. Meer gedetailleerde technische metadata zoals formaatversie, formaatprofiel, creërende applicatie, .... ontbreken en worden niet expliciet geregistreerd. En aangezien de meeste DMS-toepassingen de formaatidentificatie baseren op de extensie en/of het mime-type van het bestand is de identificatie van het bestandsformaat niet altijd even accuraat. Zo wordt bijvoorbeeld voorbijgegaan aan het gebruik van dezelfde extensie voor verschillende versies van een bestandsformaat. Bekende voorbeelden hiervan zijn de extensies .doc, .xls, .ppt voor de formaten van respectievelijk MS Word, MS Excel en MS PowerPoint. Nochtans verschilt het MS Word 2.0-formaat fundamenteel van het MS Word 97-2003-formaat. Het MS Excel 5.0-formaat of MS PowerPoint 4-formaat zijn ook in geen enkel opzicht te vergelijken hun 97-2003 opvolgers. Ook voor de verschillende versies van het PDF-formaat wordt altijd dezelfde extensie .pdf gebruikt.

In een digitaal depot is een grotere behoefte aan (accurate) technische metadata ter ondersteuning van de langetermijnarchivering van de digitale archiefdocumenten. Elke bytestream in een digitaal depot moet voorzien worden van zijn karakteriserende metadata. Deze metadata zijn te beschouwen als een onderdeel van de *representation information*<sup>8</sup> van de *information packages*. Naast de technische informatie over bestandsformaat, versie, profiel en creërende applicatie omvat deze *representation information* ook de documentatie en de semantiek van datasets. Onderlinge relaties tussen documenten, representaties, objecten, ... moeten op een expliciet en duurzaam worden vastgelegd.

### 3.4.3 Procesmetadata

De procesinformatie is een essentieel onderdeel van de metadata van digitale archiefdocumenten. Deze metadata moeten samen met documenten worden gearhiveerd. Een DMS bevat echter lang niet altijd de informatie over het proces waarbinnen de documenten zijn gebruikt of gecreëerd. Een

---

<sup>8</sup> Representation information is de OAIS-verzamelterm voor alle metadata die nodig zijn om digitale informatie op een semantische en bruikbare wijze te archiveren.

DMS zal de procesinformatie doorgaans enkel bevatten wanneer de workflow binnen het DMS zelf wordt gebouwd. Is dit laatste het geval, dan wordt de procesinformatie niet in de metadata van de dossiers of de documenten geregistreerd, maar afzonderlijk in het DMS opgeslagen. Deze procesmetadata kunnen ook in de programmatie of business-logica zijn gebouwd zodat het nodig is om deze informatie verder te expliciteren bij archivering<sup>9</sup>. Functioneert het DMS echter louter als 'docstore' en beheert het enkel de documenten die vanuit een andere proces- of bedrijfsapplicatie worden aangemaakt, dan zal het DMS in de regel niet de procesinformatie bevatten. In dit laatste geval zal de procesinformatie binnen de proces- of bedrijfsapplicatie zijn opgeslagen.

### 3.4.4 Opslag

Een DMS bewaart zijn metadata doorgaans in de database van de toepassing of in afzonderlijke XML-bestanden. De metadata worden hierdoor gescheiden van de documenten opgeslagen. Dat geldt overigens ook voor de audit-trails die een DMS bijhoudt. Opgeslagen dossier-, document-, procesmetadata en audit-trails kunnen wel worden gearhiveerd of geëxporteerd in de vorm van XML of CSV-bestanden, maar dan nog is de samenhang met de dossiers of de documenten waarop ze betrekking hebben nog niet vastgelegd.

Al deze metadata zijn een essentieel onderdeel van het *information package* dat een digitaal depot duurzaam in tijd moet overbrengen. Voor de duurzame archivering van die metadata zijn diverse oplossingen beschikbaar<sup>10</sup>. Ze hebben met elkaar gemeen dat de metadata niet alleen in de database worden opgeslagen, maar dat die ook bij de digitale documenten mee worden bewaard in het opslagsysteem van het digitale depot.

## 3.5 Systeemafhankelijkheid

Documenten en hun metadata die in een DMS worden beheerd, worden op een systeemafhankelijke wijze opgeslagen. Voor de raadpleging van de documenten en hun metadata is de DMS-toepassing vereist. Dit geldt voor zowel de metadata als de documenten. Zelfs al worden de documenten op een rechtstreeks te benaderen fileserver bewaard, dan nog hebben ze in de meeste gevallen niet de oorspronkelijke bestandsnaam meer. De meeste DMS-toepassingen die documenten op een fileserver bewaren, hanteren intern andere bestandsnamen. Deze bestandsnamen zijn doorgaans korter en bevatten een systeemnaam.

De problematiek van de systeemafhankelijkheid blijkt overduidelijk uit de complexiteit van een export of migratie van de documenten en al hun metadata. Beide operaties zijn zonder betekenisvol informatieverlies niet mogelijk zonder een grondige kennis van de DMS-toepassing en meer in het bijzonder de informatiearchitectuur en het datamodel. Deze informatie is doorgaans enkel beschikbaar bij de leverancier en/of de implementator.

DMS-toepassingen zijn bovendien ook maar softwaretoepassingen als een ander. Net als elke andere toepassing zijn ze het onderwerp van versiebeheer en een ontwikkelingscyclus. In de ICT-wereld is er doorgaans elke drie jaar een nieuwe major release van een toepassing. Achterwaartse ondersteuning van oudere versies is dan weer in de tijd beperkt. Met andere woorden, de inhoud van een DMS moet regelmatig naar een nieuwe versie worden overgezet. Ook de databasemanagementsoftware is regelmatig aan updates en nieuwe versies toe. Tijdens al deze operaties zijn voornamelijk de metadata het onderwerp van conversie (aanpassingen in het datamodel, omzettingen naar andere datatypes).

Een essentieel kenmerk van een digitaal depot is dat de digitale archiefdocumenten en hun metadata op een duurzame wijze zijn opgeslagen. De duurzaamheid van een digitaal depot is niet afhankelijk

<sup>9</sup> Een primaire bron voor deze informatie zijn dan de functionele en de technische analyses.

<sup>10</sup> Zie F. BOUDREZ, *Digitaal archiveren en digitale duurzaamheid*, Antwerpen, 2009.

van de digitaaldepotsoftware – want die is per definitie niet duurzaam – maar van de duurzaamheid van de *information packages* die het digitale depot bevat. Met andere woorden, versiebeheer of migraties naar andere depotsoftware mogen geen impact hebben op de *information packages*. Deze *information packages* moeten zonder wijziging kunnen worden overgezet naar andere (toekomstige) systemen die als digitaal depot functioneren.

### 3.6 Preservering

In een DMS wordt de leesbaarheid van digitale documenten op een passieve wijze bewaard. In een aantal DMS'en zijn wel viewers voor een aantal formaten ingebouwd of is PDF-renditie voorzien, maar verder dan dit gaat het verzekeren van de leesbaarheid niet. Bovendien richt deze functionaliteit zich enkel op de documenten zelf, en niet op hun metadata. De leesbaarheid van de metadata en hun duurzame relatie tot de documenten waarop ze betrekking hebben is afhankelijk van het DMS zelf. In een DMS worden digitale documenten op een passieve wijze beheerd.

Een aantal DMS-toepassingen hanteren ook hun eigen productafhankelijke opslagwijze voor documenttypes zoals e-mail<sup>11</sup>. Deze functionaliteit is er hoofdzakelijk op gericht om e-mails in een HTML-weergave in een webbrowser te kunnen presenteren. Andere DMS-toepassingen importeren e-mails standaard in HTML-formaat, zonder dat hierbij alle essentiële eigenschappen van de e-mail worden omgezet. Net zoals bij de standaardrendities naar PDF gaat men daarbij voorbij aan het bewaren van de integriteit van digitale documenten.

Diverse DMS-toepassingen voorzien ook in informatiebeveiliging door de mogelijkheid aan te bieden digitale documenten in geëncrypteerde vorm te bewaren. Het toepassen van encryptie in de opslag van digitale documenten is een extra reconstructieschakel die haaks staat op het vermijden van extra reconstructieschakels bij de langetermijnarchivering van digitale documenten.

De focus van een digitaal depot ligt op de langetermijnbewaring van digitale archiefdocumenten. Hierbij wordt uitgegaan van een actief preserveringsbeleid dat zich richt op alle facetten van het functioneel proces digitaal archiveren: de gegevensdragers, de bitintegriteit van de digitale objecten, de leesbaarheid van digitale documenten, de begrijpbaarheid en de betrouwbaarheid van digitale archiefdocumenten (metadata)<sup>12</sup>. Een actief preserveringsbeleid is een essentiële functie binnen een digitaal depot en richt zich op het bewaren van de authenticiteit, de integriteit, de betrouwbaarheid en de bruikbaarheid van de digitale documenten en hun metadata.

Vreemd genoeg zijn in een DMS de metadata regelmatig het onderwerp van conversie of migratie. Bij het upgraden van de DMS-toepassing of het databasemanagementsysteem worden de metadata geconverteerd of gemigreerd. Dit wordt binnen de meeste organisaties beschouwd als een pure ICT-aangelegenheid zonder dat hierbij de bewaring van de integriteit van de metadata wordt bewaakt.

### 3.7 Zoeken

Eén van de basisfunctionaliteiten van een DMS-toepassing zijn geavanceerde zoekmogelijkheden op de inhoud en de metadata van documenten. Elk DMS houdt bijvoorbeeld een full-text index op de inhoud van de documenten en hun metadata bij. Zoeken in een DMS is bijgevolg sterk afhankelijk van de gebruikte zoekwoorden en de inhoud van de indexen. Deze manier van zoeken is dus in het bijzonder geschikt voor gebruikers die de inhoud van documenten kennen en weten welke zoekwoorden moeten worden gehanteerd.

Bij archieven zijn de gebruikers echter ook externen die niet betrokken zijn bij de werkprocessen en de archiefvorming (bijv. burgers, onderzoekers, ...). Bij een digitaal depot verloopt de bevraging van de

<sup>11</sup> Een voorbeeld hiervan is het EMCMPF-formaat voor e-mails van Documentum. E-mails in bijv. .msg-formaat worden bij import automatisch omgezet naar het EMCMPF-formaat. Dit EMCMPF-formaat is versieafhankelijk (*EMC Documentum Webtop Version 6.5 SP3. Release notes*, maart 2010, p.197)

<sup>12</sup> F. BOUDREZ, *Digitaal archiveren en digitale duurzaamheid*, Antwerpen, 2009.



opgeslagen informatie dan ook veeleer vanuit de context van de archiefdocumenten. Beschrijvingen van archiefvormers en de werkprocessen, de metadata van de reeksen en de archiefbestanddelen zijn de primaire zoekbronnen, en niet zozeer de woorden die in documenten voorkomen.

Nog meer dan een DMS moet een digitaal depot deel uitmaken van een geïntegreerd archiefbeheer. Binnen dit geïntegreerd archiefbeheer moeten zowel analoge, gedigitaliseerde als digital born archieven beheerd worden en opzoekbaar zijn. Het realiseren van een integratie met het beschrijvingssysteem van papieren archief is één mogelijkheid. Het hanteren van een overkoepelend beschrijvingssysteem voor zowel papieren als digitaal archief is een andere. DMS-toepassingen zijn hiervoor niet zo geschikt. Een DMS kan via een federated search of door de opname van contentless objecten wel zoeken in andere gegevensbronnen, maar dit is geen doorsnee out of the box functionaliteit.

### 3.8 Volume

Het verschil tussen een DMS en een digitaal depot inzake volume situeert zich op twee vlakken: het aantal digitale objecten en de omvang van de digitale objecten.

Een aantal DMS-toepassingen hebben (hard-coded) een maximum limiet voor het totale aantal digitale objecten die ze kunnen bevatten<sup>13</sup>. Limieten zijn dikwijls ook van toepassing op het maximum aantal documenten in een site, een map, een collectie of een partitie, al naargelang het architectuurmodel van het DMS. Van andere DMS-toepassingen is dan weer bekend dat ze te kampen hebben met performantieproblemen van zodra ze heel veel (major versions) digitale documenten bevatten<sup>14</sup>. Vooral het uitvoeren van zoekopdrachten vraagt dan veel tijd, zodat bijzondere maatregelen voor het behouden van de performantie nodig zijn.

Daarnaast kunnen DMS-toepassingen niet altijd even vlot overweg met digitale objecten van een grote omvang. Sommige DMS-toepassingen bewaren de digitale objecten als BLOB's<sup>15</sup> in een database, zodat de databaser restricties de maximale omvang van de bestanden bepalen<sup>16</sup>. Dit soort van beperkingen zijn doorgaans als (technische) limiet ingebouwd zodat hier geen workarounds of tweaks voor zijn. Veel DMS-toepassingen werken ook volledig webgebaseerd en kunnen de up- en download van heel grote objecten niet aan. Om die reden wordt in veel DMS-implementaties standaard een maximale omvang per digitaal object gehanteerd.

Een digitaal depot moet in staat zijn om zowel een heel groot volume aan digitale objecten te beheren als om heel grote digitale objecten (DVD, gedigitaliseerde geluids- en beeldopnamen, GIS-informatie, ...) op te nemen en te beheren. Het aantal digitale objecten dat wordt opgenomen in een digitaal depot kan oplopen tot tientallen miljoenen. Digitale objecten groter dan enkele gigabytes zijn geen uitzondering. In enkele uitzonderlijke gevallen kunnen sommige digitale objecten die thuis horen in een digitaal depot zelfs groter zijn dan enkele honderden gigabytes. Het grootste digitale object in

<sup>13</sup> Zo zijn er o.a. beperkingen inzake een maximum aantal documenten die binnen één map kunnen worden gepresenteerd. Andere beperkingen hebben dan weer te maken met het databasemanagementsysteem dat door de DMS-toepassing wordt gebruikt (bijv. maximum aantal records/tabel, maximale omvang per object).

<sup>14</sup> Dit leidt dikwijls tot enkele vreemde ingrepen zoals het afzetten van geautomatiseerd versiebeheer, het periodiek verwijderen van minor versions of het uitsluiten van grote bestanden voor opname of indexing.

<sup>15</sup> BLOB: Binary Large Object. Dit is een datatype voor velden in een (Oracle-)databank. Gelijkaardige datatypes in MS SQL-omgeving zijn images of varbinary. In deze velden kunnen binaire gegevens en zelfs hele computerbestanden worden opgeslagen (Bijv. een MS Word-of PDF-document dat in één veld wordt opgeslagen).

<sup>16</sup> Sharepoint 2010 hanteert bijvoorbeeld de volgende restricties (<http://technet.microsoft.com/en-us/library/cc262787.aspx>):

- een maximale bestandsomvang van 2 GB per object ( = max. BLOB-omvang in een SQL-database)
- een maximale omvang van 200 GB per contentdatabase



het digitale depot van de stad Antwerpen is 400 GB klein.

### 3.9 Operaties

De basisfunctionaliteiten van DMS-toepassingen zijn doorgaans enkel toepasbaar op een beperkte selectie van documenten en/of mappen. Vanwege de performantie hanteren de meeste DMS-toepassingen en implementaties een beperking op het maximum aantal documenten en/of mappen dat gezamenlijk kan worden bewerkt.

Gelet op het grote volume aan digitale objecten in een digitaal depot moeten de operaties (opname, conversie, migratie, ...) in bulk kunnen worden uitgevoerd.

### 3.10 Rechtenbeheer

In een DMS kan een heel fijnmazig rechtenbeheer worden opgezet. Op het niveau van mappen, documenten en soms zelfs van documentversies kunnen op een gedetailleerde wijze rechten worden geplaatst voor zowel individuele gebruikers, groepen of rollen binnen de organisatie. Grote categorieën in de rechten zijn lees- en wijzigingsrechten, maar soms worden nog veel specifiekere rechten toegekend. Ook op het vlak van metadata worden rechten geplaatst of de rechten worden overgenomen van de documenten of mappen.

In een digitaal depot zijn wijzigingsrechten in principe niet meer van toepassing: alle archiefdocumenten zijn per definitie 'read-only'. In tegenstelling tot een DMS gaat het in een digitaal depot hoofdzakelijk om het onderscheid openbaar of niet-openbaar en wordt de openbaarheid niet op documentniveau geregeld, maar op reeks- en/of bestanddeelniveau. Een bijzondere rechtengroep is nodig voor de digitale depotbeheerders. Deze groep heeft wel wijzigingsrechten. Zowel openbare als niet-openbare archiefdocumenten moeten raadpleegbaar zijn voor zowel interne als externe gebruikers.

Aangezien gearchiveerde digitale documenten ook door externe gebruikers worden geraadpleegd, moet binnen een digitaal depot ook de nodige aandacht worden besteed aan auteursrechten. De uitzonderingen op de auteurswet houden ondermeer in dat auteursrechtelijk beschermde werken wel in het gebouw van het archief kunnen worden geraadpleegd, maar niet er buiten of op het internet.

## 4. BESLUIT

---

DMS-toepassingen en digitale depots verschillen in een aantal opzichten fundamenteel van elkaar. In deze bijdrage werden tien verschillen besproken. De belangrijkste conclusie is dat DMS-toepassingen niet zomaar 'out of the box' inzetbaar zijn als digitaal depot of zelfs als platform voor een digitaal depot.

Naast de verschillen zijn er natuurlijk ook gelijkenissen. Hierdoor rijst de vraag of mits maatwerk of extra ontwikkelingen aan het DMS niet aan een aantal ontbrekende digitaal depot functionaliteiten tegemoet kan worden gekomen. Met maatwerk of extra ontwikkelingen kunnen ongetwijfeld enkele ontbrekende functionaliteiten in het DMS worden voorzien. Vraag is echter wat de compatibiliteit van dit maatwerk met toekomstige versies van het DMS is, welke kostprijs dit maatwerk met zich meebrengt en of een pure digitaal depotoplossing niet goedkoper is.

Naast de aandachtspunten compatibiliteit en kostprijs blijft een fundamenteel element de systeemafhankelijke opslag van metadata en documenten binnen elk DMS. Zelfs wanneer het DMS hiervoor gebruik maakt van XML, is dit nog niet afdoende. Het betreft immers in veel gevallen een systeemafhankelijke XML-vorm die niet voldoet aan de eisen van *representation information*. Het uitgangspunt van een digitaal depot is en blijft immers de archivering van duurzame

*informationpackages* waarvan de verschillende componenten systeemafhankelijk reconstrueerbaar moeten zijn en blijven. Tot op heden biedt geen enkel DMS hiervoor voldoende garanties en functionaliteiten. Hierdoor rijst dan ook de vraag of het wel opportuun is om een RMA binnen een DMS te bouwen, en of het niet verstandiger is om vanuit een digitaal depot records management functies aan te bieden.

## 5. MEER INFORMATIE

---

- F. BOUDREZ, *Basisprocessen voor een digitaal archiefdepot*, Antwerpen, 2006.
- F. BOUDREZ, *Digitaal archiveren en digitale duurzaamheid*, Antwerpen, 2009.
- F. BOUDREZ, *Wat is een digitaal depot?*, Zevenbergen, 11 november 2009.
- B. LOHMAN, *Waarom is een DMS geen digitaal depot?*, Zevenbergen, 11 november 2009.
- ED3. *Eisen Duurzaam Digitaal Depot, Toetsingskader voor langetermijnbeheer van blijvend te bewaren digitale informatie*, versie 2, LOPAI, december 2012.
- *Trustworthy Repositories Audit & Certification: Criteria and Checklist*, 2007.