

TOEKOMST VOOR ONS **DIGITAAL GEHEUGEN**

duurzame toegang tot informatie in Nederland



Redactie: Inge Angevaare

Mede gebaseerd op inputnotities van Petra Helwig (Nationaal Archief, sector overheid/archieven), René van Horik (Data Archiving and Networked Services, sector wetenschap), Annelies van Nispen (Digitaal Erfgoed Nederland, sector cultuur/erfgoed), en Thaeis (bedrijvenonderzoek) (gepubliceerd op <http://www.ncdd.nl/activiteiten-natverkenning.php>).

Met veel dank aan Ingrid Dillo voor het kritisch meelezen.

De interimrapportage onder dezelfde titel is besproken tijdens de nationale werkconferentie *Toekomst voor ons digitaal geheugen* die op 18 september 2009 in Den Haag is gehouden en de deelnemers hebben volledig ingestemd met de resultaten van het onderzoek. Enkele kleine correcties zijn doorgevoerd in dit definitieve document.

november 2009

NCDD – Nationale Coalitie Digitale Duurzaamheid
Postbus 90407
2509 LK Den Haag
<http://www.ncdd.nl>
info@ncdd.nl

Het rapport werd gefinancierd door het ministerie van Onderwijs, Cultuur en Wetenschap, directie Onderzoek en Wetenschapsbeleid.



TOEKOMST VOOR ONS **DIGITAAL GEHEUGEN**

duurzame toegang tot informatie in Nederland

(definitief rapport)

'Er komt zoveel data beschikbaar dat een onderzoeker die hoeveelheid soms niet meer kan bevatten' (een wetenschapper)

'We moeten veel scherper selecteren. ... Uiteindelijk gaat het erom te bepalen wat het de samenleving waard is om de informatiehuishouding van de overheid te bewaren.'
(een archiefinspecteur)

Inhoudsopgave

0.	Managementsamenvatting	7
1.	Inleiding	11
	1.1. Duurzame toegang tot digitale informatie: de uitdaging	11
	1.2. De situatie in Nederland anno 2009	11
	1.3. De Nationale Verkenning Digitale Duurzaamheid	12
2.	De publieke sector en het bewaren van informatie	15
	2.1. De publieke sector: inleiding	15
	2.2. Redenen om informatie te bewaren	16
	2.3. Wetten en beleidskaders	17
	2.4. Rollen en verantwoordelijkheden ten aanzien van informatie	19
3.	Duurzame toegang tot digitale objecten: risico's en beheersmaatregelen	23
	3.1. De risico's	23
	3.2. Technische beheersmaatregelen	24
	3.3. Beleidsmatige en organisatorische beheersmaatregelen	25
	3.4. In Nederland gebruikte richtlijnen	26
	3.5. Impact van de richtlijnen	30
4.	De publieke sector en digitale informatie	31
	4.1. Rol en belang van digitale objecten	31
	4.2. De noodzaak van selectie	34
	4.3. Duurzame toegang tot digitale objecten in de praktijk	36
5.	Kosten en financiering	51
	5.1. Inleiding	51
	5.2. Onderzoek naar de kosten van duurzame toegang	47
	5.3. Factoren die de kosten beïnvloeden	52
	5.4. Ervaringscijfers in de Nederlandse context	53
	5.5. Financiering	55
	5.6. De kosten-batenanalyse	56
6.	(Intersectorale) samenwerking	59
	6.1. Intersectorale samenwerkingsverbanden elders	59
	6.3. Perspectieven voor Nederland	60
7.	Conclusies en aanbevelingen	63
Bijlagen:		
	1. Bronnenlijst	69
	2. Lijst van geïnterviewden	77
	3. Lijst van gebruikte termen	80

'Ja, we hebben een keer een crash gehad. Geen back-up, al onze data foetsie. Dat is nog steeds een trauma in de organisatie. Daar willen we het liever niet meer over hebben.' (NN)

0. Managementsamenvatting

Digitale informatie is een essentieel bestanddeel van de hedendaagse kenniseconomie. De publieke sector investeert veel in de productie en verwerving van digitale informatie en wil een deel van die informatie langdurig toegankelijk houden voor gebruik in de toekomst.

Maar digitale informatie vereist specifieke zorg die niet overal kan worden gegeven; gebrek aan zorg kan leiden tot verlies van belangrijke bestanden. Hoofdstuk 1 beschrijft hoe de Nationale Coalitie Digitale Duurzaamheid het initiatief heeft genomen om te laten onderzoeken hoe de publieke sector omgaat met zijn digitale gegevens: waar wordt duurzaam beheer voortvarend aangepakt? Waar loopt digitale informatie het risico om verloren te gaan? En hoe kunnen de risico's beheerst worden? De *Nationale Verkenning Digitale Duurzaamheid* werd gefinancierd door de directie OWB van het ministerie van OCW.

Omdat het te onderzoeken veld groot en divers is en de materie relatief onbekend, kozen de onderzoekers voor semigestructureerde interviews met belanghebbenden in de sectoren wetenschap, overheid/archieven en cultuur/erfgoed, en vulden deze informatie aan met literatuuronderzoek.

Hoofdstuk 2 introduceert de sectoren en onderzoekt welke redenen zij hebben om informatie langdurig te bewaren. De *overheid* moet verantwoording kunnen afleggen aan de samenleving voor haar handelen en wil het cultuurhistorisch belang dienen. In de Archiefwet heeft zij daarom een wettelijk kader vastgesteld voor langdurige bewaring van informatie, op papier of digitaal. Goede praktijk in de *wetenschap* vereist dat onderzoek geverifieerd moet kunnen worden; ook kan nieuw onderzoek gebaseerd worden op vastgelegde gegevens. Voor het cultureel erfgoed is langetermijntoegang altijd al een kerntaak. De private sector blijkt informatie niet langer te bewaren dan strikt noodzakelijk om procesrechtelijke redenen.

Wie precies verantwoordelijk is voor de langetermijntoegang is door de overheid goed vastgelegd in de Archiefwet en bij cultureel erfgoedinstellingen vrijwel vanzelfsprekend, maar in de wetenschap is vaak onduidelijk wie eigenaar van gegevens is en/of wie verantwoordelijk is voor langetermijntoegang. Niemand heeft nog verantwoordelijkheid genomen voor het hele web 2.0, waar enorme hoeveelheden audiovisueel erfgoed ontstaan.

Hoofdstuk 3 beantwoordt de vraag waarom digitale informatie zo bijzonder is en welke beheersmaatregelen nodig zijn om digitale informatie langdurig toegankelijk te houden. Digitale informatie is niets anders dan een reeks nullen en enen, die alleen betekenis krijgt door de specifieke combinatie van hardware en software waarmee de gegevens zijn geproduceerd. Omdat hardware en software voortdurend veranderen, breekt al gauw het moment aan waarop een (nieuwe) computer de (oude) informatie niet meer kan lezen. Dit risico wordt beperkt door de informatie onder een voortdurend bewakingsregime te plaatsen en steeds opnieuw aan te passen aan nieuwe hardware en software. Daarvoor is een technisch geavanceerde, betrouwbare beheersomgeving nodig alsmede structurele inspanning op het gebied van research en development.

Maar ook in organisatorische zin zijn maatregelen nodig. Digitale informatie vereist zorg van wieg tot graf. Producenten, beheerders en gebruikers van informatie moeten hun handelingen op elkaar afstemmen om breuken in de continuïteit te voorkomen. Eisen voor duurzaamheidsbeleid zijn onlangs vastgelegd in een aantal richtlijnen: de *Baseline informatiehuishouding* voor de overheid, het *Data Seal of Approval* voor wetenschappelijke onderzoeksdata en *De Basis voor duurzaamheid* voor het cultureel erfgoed.

In hoofdstuk 4 kijken we naar de praktijk. In de *wetenschap* is het gebruik van digitale informatie het diepst doorgedrongen. De betawetenschappen hebben zich inmiddels georganiseerd in wereldwijde datanetwerken van een enorme omvang en verscheidenheid. De sociale wetenschappen en de geesteswetenschappen volgden. Op een enkele discipline na is de wetenschap anno 2009 volstrekt ondenkbaar geworden zonder digitale data.

De sectoren *overheid* en *cultuur/erfgoed* zitten nog middenin de overgang van fysiek en analoog naar digitaal. Hoewel de overheid werkt aan elektronische dienstverlening en ook computers gemeengoed zijn, is het te bewaren *archiefstuk* nog vaak een printje, omdat nog alleen op papier de benodigde authenticiteit kan worden gewaarborgd. Ook in het *cultureel erfgoed* ligt de meeste aandacht nog bij het fysieke object; digitale kopieën zijn vaak gebruikskopieën voor online informatie over collecties. Niettemin is de verwachting dat de rol van digitale informatie in de komende jaren sterk zal groeien.

De vraag doet zich voor of *alle* geproduceerde informatie ook bewaard moet worden. De respondenten in het onderzoek beantwoorden deze vraag ontkennend. Alles bewaren is te kostbaar en onwenselijk; we moeten een gerichte selectie maken van informatie die ook in de toekomst van belang is voor de samenleving. De *overheid* heeft uit het analoge tijdperk het meest uitgebreide selectiemechanisme geërfd (Archiefwet), maar dit behoeft aanpassing aan het digitale tijdperk. De *wetenschap* heeft twee gezichten: ten aanzien van *publicaties* wordt de praktijk uit het analoge tijdperk voortgezet: het *peer review* systeem selecteert materiaal op basis van kwaliteit. Voor ruwe en bewerkte onderzoeksdata ontbreken vooralsnog selectiecriteria.

Voor het beheren van (geselecteerde) digitale objecten kent Nederland inmiddels een negental gespecialiseerde instellingen die voldoen aan de richtlijnen en die in staat zijn om de *bitstream* goed te bewaren. Een aantal andere digitale depots is in (diverse stadia van) ontwikkeling.

Maar de capaciteit van deze instellingen is beperkt; er wordt heel veel meer informatie geproduceerd dan zij kunnen verwerken.

Bovendien blijkt de *productie* van digitale informatie vaak niet te gebeuren in overeenstemming met de richtlijnen voor duurzaam beheer. Belemmeringen die daarbij een rol spelen zijn:

- gebrek aan bewustzijn;
- gebrek aan informatie en kennis over duurzame toegang;
- staat van de informatiehuishouding zelf;

- gebrek aan duurzame financiering;
- het (individuele) belang van de producent strookt niet met het algemene belang;
- gebrek aan betrouwbare opslagcapaciteit;
- gebrek aan menskracht, IT-capaciteit;
- gebrek aan concrete, praktische gereedschappen;
- gebrek aan diensten;
- gebrek aan organisatie die is aangepast aan het digitale tijdperk.

Hoofdstuk 5 behandelt de kosten van duurzame toegankelijkheid en concludeert dat die niet zomaar berekend kunnen worden, omdat de variabelen per instelling en per collectie te veel verschillen. Ook hanteren organisaties verschillende methodes voor het toerekenen van kosten aan duurzame toegankelijkheid. Onderzoek heeft wel een aantal algemene principes blootgelegd:

- De kosten per object nemen af naarmate het archief groter is.
- Mettertijd vlakken de kosten per object af.
- Niet opslag maar personeel voor het beheer is de grootste kostenfactor.
- Timing is belangrijk: als niet bij creatie goede beheersmaatregelen worden genomen is reparatie achteraf onmogelijk of exorbitant duur.
- De duurste fase lijkt acquisitie en opname in het systeem (ingest).

Ook de financiering komt in hoofdstuk 5 ter sprake. In vrijwel alle gevallen treedt de overheid op als financier van voorzieningen voor langetermijntoegang, maar zij maakt hierbij vaak gebruik van incidentele middelen (voor digitaliseringsprojecten, voor wetenschappelijk onderzoek). Het gebruik van incidentele middelen leidt tot aanzienlijke risico's: na afloop van de projecten kan de langetermijnzorg niet gegarandeerd worden.

De kosten-batenanalyse voor langetermijntoegang tot digitale informatie kan nauwelijks worden gemaakt. De voordelen liggen immers in de toekomst en zijn niet te kwantificeren.

Hoofdstuk 6 kijkt naar de perspectieven voor (intersectorale) samenwerking. Buitenlandse initiatieven als het National Digital Information Infrastructure and Preservation Program (NDIIPP, USA), de Digital Preservation Coalition (DPC, UK) en nestor (Duitsland) passeren de revue. Wat Nederland betreft zijn de respondenten vrijwel unaniem in hun oordeel dat samenwerking duurzame toegankelijkheid bevordert – al lopen de meningen over de manier waarop die samenwerking gestalte moet krijgen uiteen. Samenwerking in het kader van duurzame toegankelijkheid *binnen* sectoren krijgt momenteel vooral vorm in de archiefsector, o.a. in het consortium van het Nationaal Archief en de Regionale Historische Centra.

Hoofdstuk 7, ten slotte, identificeert de elementen die onderdeel moeten zijn van een plan voor een landelijke infrastructuur voor duurzame toegankelijkheid, en benoemt organisaties en initiatieven die daaraan hun bijdrage moeten leveren:

1. **Probleembewustzijn:** er moet nog veel werk worden verzet om alle belanghebbenden te doordringen van de risico's en ook het langetermijnbelang steeds weer onder de aandacht te brengen.
2. **Rollen en verantwoordelijkheden:** digitale bewaarplaatsen zelf moeten bruggen slaan naar de producenten van digitale informatie om zorg van wieg tot graf te bevorderen; de informatieketen is gebaat bij centrale regie. In de wetenschap ligt hier een belangrijke rol voor financiers die subsidies afhankelijk kunnen maken van duurzaam databeheer.
3. **Faciliteiten voor opslag en duurzaam beheer:** er is behoefte aan een landelijk dekkend netwerk van opslagfaciliteiten, waarbij schaalgrootte wordt nagestreefd voor zover die het primaire proces niet belemmert; regionaal, nationaal, en/of sectoraal, dat moet nader onderzocht worden.
4. **Diensten en gereedschappen:** vooral kleinere organisaties die digitale informatie beheren zijn gebaat bij een flexibel aanbod van diensten en gereedschappen dat hen in staat stelt hun data goed te beheren. Hierbij moet vooral aandacht zijn voor laagdrempelige, praktische gereedschappen. Tegelijkertijd moeten kleine organisaties niet wachten tot een dergelijk aanbod is gerealiseerd: ook met weinig middelen kunnen de risico's van verlies al flink beperkt worden.
5. **Structurele financiering:** duurzame toegankelijkheid vraagt om (aanvullende) structurele financiering voor bewarende instellingen; een fonds met bijdragen uit digitaliserings- en onderzoekprojecten zou een aanvullende bijdrage kunnen leveren aan die financiering.
6. **Selectie:** om duurzame toegankelijkheid betaalbaar te houden moeten alle betrokken organisaties nog meer dan voorheen selecteren welke informatie echt voor de toekomst bewaard moet worden; bestaande selectiemechanismes moeten digitaal 'hertaald' worden.
7. **Kennis en expertise:** een landelijk kenniscentrum moet ervoor zorgen dat kennis en expertise professioneel en efficiënt worden verspreid in het netwerk.

1. Inleiding

Deze inleiding beschrijft in het kort de problematiek van duurzame toegang tot digitale informatie en schetst de achtergrond en uitvoering van de *Nationale Verkenning Digitale Duurzaamheid*.

1.1. Duurzame toegang tot digitale informatie: de uitdaging

Wanneer informatie wordt toevertrouwd aan papier, ontstaat er een combinatie van drager/informatie die onlosmakelijk met elkaar verbonden is en die door ieder ander mens begrepen wordt zonder dat daar extra hulpmiddelen voor nodig zijn. Als zich geen rampen van buiten voordoen, zoals brand, waterschade of schimmels, kan de informatie eeuwenlang bruikbaar blijven.

Digitale gegevens, daarentegen, bestaan uit een onzichtbare brij van enen en nullen (de *bitstream*) die alleen betekenis krijgt als die wordt afgespeeld op *dezelfde* combinatie van hardware en software als waarmee de bitstream is geproduceerd. Omdat er voortdurend nieuwe hardware en software wordt geïntroduceerd, dreigt digitale informatie al binnen een paar jaar onbruikbaar te worden. Daarbij komt dat de dragers van de gegevens (floppy disks, cd's, dvd's, magneetbanden) zelf ook een beperkte levensduur hebben.

Eind vorige eeuw is een aantal archieven, bibliotheken en wetenschappelijke instellingen zich bewust geworden van deze risico's en is men gaan werken aan mogelijke oplossingen.

1.2. De situatie in Nederland anno 2009

Anno 2009 is er technisch flinke vooruitgang geboekt, maar zijn experts er ook van overtuigd geraakt dat we hier te maken hebben met een blijvende uitdaging. Bovendien is het besef gegroeid dat technische oplossingen alleen niet genoeg zijn: er spelen beleidsmatige, organisatorische en financiële factoren die tot nu toe te weinig aandacht hebben gekregen.

Dit besef leidde in 2007 tot de oprichting van de *Nationale Coalitie Digitale Duurzaamheid* (NCDD), een uniek intersectoraal samenwerkingsverband waarin elf landelijke organisaties participeren die te maken hebben met het beheren van digitale informatie:

- 3TU.Datacentrum (universiteitsbibliotheken van Delft, Eindhoven, Twente)
- Data Archiving and Networked Services (DANS)
- Nederlands Instituut voor Beeld en Geluid (NIBG)
- Koninklijke Bibliotheek (KB)
- Koninklijke Nederlandse Akademie voor Wetenschappen (KNAW)
- Ministerie van Binnenlandse Zaken en Koninkrijksrelaties (ministerie BZK)
- Nationaal Archief (NA)
- Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO)
- SURFfoundation

- Centraal Bureau voor de Statistiek (CBS, geassocieerd lid)
- Digitaal Erfgoed Nederland (DEN, geassocieerd lid)

De Coalitie heeft zich ten doel gesteld om te bevorderen dat er binnen vijf jaar een organisatorische en technische infrastructuur tot stand komt voor duurzame toegang tot digitale informatie in de publieke sector in Nederland. Die infrastructuur moet het mogelijk maken om gericht die informatie te *selecteren* die voor de toekomst van belang is en die zodanig *duurzaam op te slaan en toegankelijk te houden* dat toekomstige generaties de informatie kunnen (*her*)gebruiken en erop kunnen vertrouwen dat de informatie *authentiek* en *betrouwbaar* is.

Om plannen voor een landelijke infrastructuur te onderbouwen was er behoefte aan een inventarisatie van de huidige stand van zaken ten aanzien van duurzame toegankelijkheid. Dit leidde tot de *Nationale Verkenning Digitale Duurzaamheid*.

1.3. De Nationale Verkenning Digitale Duurzaamheid

1.3.1. Doelstelling

In 2006 publiceerde de Engelse Digital Preservation Coalition het rapport *Mind the Gap*, waarin de problematiek van digitale duurzaamheid voor het eerst in een (Engels) organisatorisch kader werd gezet [Mind the Gap, 2006]. Een dergelijke verkenning, zo oordeelde de NCDD, zou ook voor Nederland moeten worden uitgevoerd om een gedegen fundament te leggen onder een door de NCDD te ontwikkelen strategie.

De directie Onderwijs en Wetenschapsbeleid van het ministerie van Onderwijs, Cultuur en Wetenschap stelde middelen ter beschikking om het onderzoek uit te voeren. Het ministerie stelde de Koninklijke Bibliotheek aan als projectleider en beheerder van de middelen.

Doel van het project was om in kaart te brengen hoe belangrijk digitale informatie is voor het functioneren van de publieke sector, en hoe die informatie wordt beheerd. Waar gebeurt dat duurzaam? Waar lopen categorieën informatie gevaar om verloren te gaan? Welke belemmeringen spelen daarbij een rol? Hoe en door wie kunnen die weggenomen worden?

Het onderzoek spitste zich toe op de publieke sector. Om mogelijke *best practices* te identificeren werden vijf grote Nederlandse bedrijven uit de private sector geïnterviewd.

De publieke sector werd tijdens het onderzoek opgesplitst in drie subsectoren die een dermate eigen dynamiek hebben dat onderscheid zinvol leek:

- sector wetenschap
- sector overheid/archieven
- sector cultuur – toegespitst op het cultureel erfgoed.

Hoewel een dergelijke indeling de werkelijkheid nooit volledig kan dekken en er sprake is van aanzienlijke overlap tussen de sectoren, is de driedeling overwegend nuttig gebleken voor het onderzoek en wordt hij daarom ook in dit rapport gebruikt.

1.3.2. Opzet van het onderzoek

In 2008 voerde KPMG een ‘voorverkenning’ uit om te onderzoeken hoe het diverse en complexe werkveld het beste bevraagd zou kunnen worden. Daaruit bleek dat er voor het verzamelen van de juiste informatie nogal wat kennis van de sectoren én van de problematiek

van duurzame toegang nodig was, bijvoorbeeld om dóór te vragen wanneer respondenten zelf niet goed weten hoe ze hun situatie moeten beschrijven. Daarom werd besloten het onderzoek uit te voeren onder regie van de NCDD zelf, die immers de instellingen verenigt met de meeste kennis op dit gebied.

1.3.2.1. Projectorganisatie

Uit de gelederen van de NCDD werden drie onderzoekers aangetrokken die voor een deel van hun tijd gedetacheerd werden bij het project: Petra Helwig van het Nationaal Archief (NA) voor de sector overheid/archieven, René van Horik van Data Archiving and Networked Services (DANS) voor de sector wetenschap, en Annelies van Nispen van Digitaal Erfgoed Nederland (DEN) voor de sector cultureel erfgoed. Het project stond onder leiding van Inge Angevaere, coördinator van de NCDD, die ook verantwoordelijk is voor de tekst van dit rapport. Namens projectleider Koninklijke Bibliotheek begeleidde Ingrid Dillo het project.

De Verkenning werd begeleid door een stuurgroep met deskundigen uit de diverse sectoren, onder voorzitterschap van Ingrid Dillo: Marco de Niet, Directeur van Digitaal Erfgoed Nederland, Roel Gathier, plaatsvervangend Algemeen Directeur van SRON – Netherlands Institute for Space Research, Kato Vierbergen-Schuit, Senior coördinerend beleidsmedewerker bij het ministerie van Binnenlandse Zaken en Koninkrijksrelaties en projectleider van 'Informatie op Orde', en Titia van der Werf, Adjunct-directeur van het Internationaal Instituut voor Sociale Geschiedenis (IISG).

Het deelonderzoek dat erop was gericht de praktijk in de private sector te onderzoeken en mogelijke *best practices* te identificeren ten behoeve van de publieke sector, werd uitbesteed aan strategisch adviesbureau Thaesis.

1.3.2.2. De looptijd

Het oriëntatietraject liep van 1 januari 2008 tot september 2008. Het onderzoek zelf werd uitgevoerd tussen 1 januari 2009 en 30 juni 2009.

1.3.2.3. Afbakening van het onderzoek

Dit onderzoek betreft alle digitale gegevens waarin de overheid heeft geïnvesteerd, die in principe openbaar zijn, en die op de lange termijn toegankelijk moeten blijven. Met andere woorden:

- In de sectoren cultuur en wetenschap zijn de *collecties* onderwerp van onderzoek, niet de interne bedrijfsvoeringsinformatie;
- In de sector overheid is de bedrijfsvoering wel onderwerp van onderzoek, maar worden categorieën vertrouwelijke informatie (bijv. patiëntendossiers, interne personeelsdossiers) uitgesloten.

1.3.2.4. Terminologie

Voortschrijdend inzicht heeft de NCDD ertoe gebracht liever te spreken van *duurzame toegang/duurzame toegankelijkheid* dan van *digitale duurzaamheid*, omdat duurzame bewaring zonder toegankelijkheid geen zin heeft. Waar de oude term in namen verankerd is, hebben we die in dit rapport gehandhaafd als synoniem voor duurzame toegankelijkheid.

1.3.3. Onderzoeksmethodiek

Tijdens het oriëntatietraject door KPMG was een referentiekader opgesteld voor het ondervragen van respondenten, online of door middel van telefonische enquêtes. Dit kader bevatte weliswaar alle vragen die beantwoord zouden moeten worden, maar er bleken nauwelijks respondenten te vinden die al die antwoorden zouden kunnen verstrekken. De vragen veronderstelden namelijk kennis van de problematiek van duurzame toegang én overzicht bij de respondenten over de situatie binnen hun eigen organisatie. En juist daar bleek het in de praktijk vaak aan te ontbreken.

Daarom werd besloten het onderzoek uit te voeren door middel van:

- literatuuronderzoek
- semigestructureerde interviews.

Over (deelaspecten) van de problematiek van duurzame toegang is zowel nationaal als internationaal veel onderzoek gedaan (zie de bronnenlijst). Om de resultaten van die onderzoeken in een Nederlands kader te plaatsen en de verbinding te leggen met de praktijk, werden semigestructureerde interviews gehouden. Uit het netwerk van de NCDD werden in een eerste fase mensen geselecteerd met een brede sectorkennis. In de tweede fase werden concrete organisaties nader uitgediept. Bij het selecteren van te interviewen personen en organisaties hebben de onderzoekers gestreefd naar een zo breed mogelijke spreiding over soorten organisaties, bestuurslagen (overheid) en disciplines (wetenschap). Een lijst van geïnterviewden is te vinden in bijlage 2.

Het raamwerk voor de interviews bestond steeds uit de onderzoeksvragen zoals die hierboven genoemd zijn, maar de concrete vragen werden steeds aangepast aan de rol van de geïnterviewde (producent, digitaal archief, inspecteur, werkzaam in het onderwijs, etc.).

Waardevolle informatie werd tevens verzameld tijdens diverse congressen en expert meetings die gedurende het onderzoek werden gehouden. We noemen hier de DEN-conferentie van december 2008, workshops van SURFfoundation rond onderzoeksdata, de expert meeting van het Gemeentearchief Rotterdam over het uitbesteden van technisch beheer, de werkconferentie rond edepots van het ICTU Kennislab, de internationale KB/LIBER/NCDD conferentie voor wetenschappelijke bibliotheken 'Curating Research', en diverse internationale congressen (zie bronnenlijst). Niet alleen de lezingen droegen bij aan het overzicht, maar ook de vele informele momenten met vakgenoten.

Ten slotte deelden de onderzoekers informatie met twee simultaan lopende projecten: het onderzoek van de Taskforce Archieven dat leidde tot het visiedocument *e-Archief in ontwikkeling* [e-Archief in ontwikkeling, 2009] en het onderzoek naar de effecten van 'Informatie op Orde' door de Algemene Rekenkamer (deze rapportage wordt eind 2009 verwacht).

2. De publieke sector en het bewaren van informatie

Na een korte algemene inleiding behandelt dit hoofdstuk de *redenen* die de verschillende sectoren hebben om informatie (analoog of digitaal) te willen bewaren, de *wettelijke en beleidskaders* die daarvoor ontwikkeld zijn en de *rollen en verantwoordelijkheden* die daarin gedefinieerd worden.

2.1. De publieke sector: inleiding

De publieke sector wordt gekenmerkt door een enorme diversiteit aan organisaties en collecties, op allerlei niveaus:

De *overheid* bestaat uit meer dan zestienhonderd organisaties en instanties en omvat drie bestuurslagen: de centrale of rijksoverheid waaronder de dertien ministeries vallen, de twaalf provincies, en 441 gemeenten. Daarnaast zijn er 27 waterschappen en een aantal zelfstandige bestuursorganen die niet onder het rijk vallen, zoals de kamers van koophandel en openbare lichamen voor bedrijf en beroep [Overheid, website].

Binnen de sector *cultuur en media* onderscheidt het ministerie van Onderwijs, Cultuur en Wetenschap drie hoofdgroepen: cultureel erfgoed; media, letteren en bibliotheken; en kunsten. Het onderhavige onderzoek heeft zich toegespitst op *collectiehoudende* organisaties. Die vinden we met name in de groep cultureel erfgoed. Vanuit de media zijn de publieke en regionale omroepen van belang, alsook het Nederlands Instituut voor Beeld en Geluid (NIBG).

Het cultureel erfgoed wordt onderverdeeld in vier beleidsterreinen [Cultuurbeleid in Nederland, 2007, p. 65]:

- musea,
- monumenten,
- archeologie,
- archieven.

In 2005 waren er 775 musea, waarvan het merendeel door lokale overheden (of particulier) wordt gefinancierd [CBS StatLine]. In 2008 waren er 122 archieven (1 nationaal, 112 gemeentelijk of regionaal, 9 polder- en waterschapsarchieven) [Almanak voor het archiefwezen, 2008-2009].

Binnen de *wetenschap* telt de Nederlandse Onderzoek Databank (NOD), een openbare online databank met informatie over wetenschappelijk onderzoek in Nederland, vijftien algemene universiteiten, vier bijzondere universiteiten, ca. 7.600 hoogleraren en universitair hoofddocenten, bijna 40.000 onderzoekers, 750 universitaire en niet-universitaire onderzoeksinstituten en 120 onderzoeksscholen. De NOD bevat 20.000 beschrijvingen van lopende projecten en 18.000 beschrijvingen van afgesloten onderzoek [NOD, website].

Drie grote koepelorganisaties zijn actief op het gebied van de wetenschap: de Koninklijke Nederlandse Akademie van Wetenschappen (KNAW), de Organisatie voor Wetenschappelijk Onderzoek (NWO) en de Nederlandse Organisatie voor Toegepast Natuurwetenschappelijk Onderzoek (TNO).

Toepassingsgericht wetenschappelijk onderzoek vindt plaats bij het bedrijfsleven, de overheid en hogescholen. Vooral grote bedrijven in de industriesector verrichten veel onderzoek. Daarnaast hebben ministeries instellingen die onderzoek uitvoeren, zoals het Koninklijk Nederlands Meteorologisch Instituut (KNMI) dat verbonden is aan het ministerie van Verkeer en Waterstaat. Ook het Centraal Bureau voor de Statistiek (CBS), dat onder het ministerie van Economische Zaken valt, levert belangrijke gegevens voor wetenschappelijk onderzoek.

2.2. Redenen om informatie te bewaren

Op hun website noemen de gezamenlijke Archiefinspecties vier redenen om zorgvuldig om te gaan met de archieven van de *overheid* [Archiefinspecties, website]:

- Archieven vormen een belangrijk hulpmiddel om de eigen activiteiten goed te kunnen uitvoeren. Rekeningen moeten worden betaald, gebouwen moeten worden onderhouden.
- Overheidsorganen moeten verantwoording af kunnen leggen tegenover politiek en samenleving.
- Elke burger heeft recht op informatie van overheidsorganen, voor zover die in documenten is vastgelegd (Wet Openbaarheid van Bestuur)
- Archieven kunnen grote culturele en historische waarde hebben.

Het element verantwoording zien we terugkomen in de sector *wetenschap*: de resultaten van wetenschappelijk onderzoek moeten verifieerbaar en controleerbaar zijn; om uitkomsten te toetsen moeten experimenten herhaald kunnen worden. Dat vereist zorgvuldige documentatie. Maar de wetenschap heeft nog meer redenen om informatie te bewaren: nieuw onderzoek bouwt altijd voort op eerder gedaan onderzoek. Bovendien zijn de ruwe gegevens die onderzoek mogelijk maken soms uniek, zoals bijvoorbeeld in de meteorologie, of alleen tegen hoge kosten opnieuw bijeen te brengen.

Het belang van de *cultuurhistorie* werd in 2007 aldus onderstreept door de Raad voor Cultuur:

‘Een samenleving die vooruit wil, doet er verstandig aan óók achterom te kijken. Het verleden moet niet worden vergeten, maar opgepoetst en gebruikt. Zonder continuïteit raken we op drift en koersen we van het ene incident naar het volgende. Ons collectieve geheugen, waarvan belangrijke delen zijn opgeslagen in archieven, musea en bibliotheken, schept een onmisbare basis voor die continuïteit.’ [Innoveren, participeren!, 2007, p. 12]

‘... het advies vertrekt vanuit een oriëntatie op “cultureel burgerschap”, waarmee de Raad recht wil doen aan het toenemende belang van cultuur voor het functioneren van de samenleving. Burgerschap en maatschappelijke participatie zijn in een tijdperk van migratie, globalisering en heterogenisering van de samenleving sterk onder druk komen te staan. Het advies biedt aanknopingspunten om de gesignaleerde problemen in kansrijke beleidsopties te vertalen. Zo stelt de Raad dat burgerschap staat of valt met goed geïnformeerde burgers, en in het verlengde daarvan met instellingen die onbelemmerd en bemiddelend toegang bieden tot bronnen van cultuur en informatie. De toegankelijkheid van het publieke (informatie)domein moet worden gewaarborgd en dat is een taak die de overheid zeer serieus moet nemen en in beleid hecht moet vastleggen.’ [Innoveren, participeren!, 2007, p. 8]

Alle organisaties, publiek en privaat, delen het belang van informatie voor de directe bedrijfsvoering, maar na een jaar of tien ontstaat een belangrijk onderscheid: de private sector bewaart de meeste informatie niet langer dan strikt noodzakelijk, om fiscale of procesrechtelijke redenen [Bedrijvenonderzoek, p. 16]. Voor de meeste bedrijven leidt dit al snel tot vernietiging van gegevens; slechts enkele sectoren zijn gebonden aan langere bewaartermijnen (bijv. de farmaceutische industrie). Bij de overheid gaat het onmiddellijke belang over in een verantwoordingsbelang en een cultuurhistorisch belang: na selectie wordt een deel van de geproduceerde informatie bewaard 'voor de eeuwigheid'.

2.3. Wetten en beleidskaders

De noodzaak om (selecties uit) de geproduceerde informatie korter of langer te bewaren, komt tot uitdrukking in een aantal wetten en al dan niet vastgelegde praktijken binnen de diverse sectoren. Daarin worden regels voor goed informatiebeheer gegeven en verantwoordelijken benoemd. In deze paragraaf beschrijven we die regels; in hoofdstuk 4 zal de uitvoeringspraktijk aan de orde komen.

2.3.1. De overheid en de Archiefwet

Het bekendste en belangrijkste kader voor het bewaren van overheidsinformatie is de Archiefwet 1995. Centraal in de Archiefwet staat de verplichting dat overheidsorganen hun archiefbescheiden in 'goede, geordende en toegankelijke staat' moeten brengen en bewaren. Een aantal onderdelen van de wet is nader uitgewerkt in het Archiefbesluit 1995 en in zogenaamde ministeriële regelingen. Alle organisaties die archieven produceren ('archiefvormers') zijn verplicht om hun eigen voorzieningen te treffen voor goed archiefbeheer [Prins & Matthijssen, 2000, pp. 30-46].

De Archiefwet bepaalt dat de verantwoordelijkheid voor de 'goede, geordende en toegankelijke staat' van rijksarchieven berust bij de overheidsorganisatie zelf, de archiefvormer (met het ministerie van Binnenlandse Zaken als coördinator), totdat de stukken na uiterlijk twintig jaar worden overgebracht naar het Nationaal Archief – dat onder het ministerie van Onderwijs, Cultuur en Wetenschap valt.

De 'overbrenging' naar een archief markeert een belangrijke wijziging in het openbaarheidsregime, van de Wet Openbaarheid van Bestuur naar de Archiefwet. Zoals een respondent het samenvatte: 'Bij een beroep op de WOB moet je je verzoek heel specifiek indienen, bij een beroep op de Archiefwet mag je bladeren.' Lokale overheden dragen hun archieven over aan gemeentearchieven of regionale historische centra – en soms blijft de gemeentesecretaris zelf verantwoordelijk.

De respondenten in het onderzoek zijn het erover eens dat de Archiefwet aangepast moet worden aan het digitale tijdperk; sommigen vragen om heel specifieke richtlijnen voor duurzame bewaring, anderen zijn van mening dat zulke richtlijnen per definitie door de praktijk worden ingehaald.

2.3.2. De Auteurswet, het Databankenrecht en de Wet Bescherming persoonsgegevens

Bij het digitaliseren van fysieke objecten en het omgaan met digitale gegevens spelen ook andere wettelijke kaders als de Auteurswet, het Databankenrecht en de Wet bescherming

persoonsgegevens [Beunen & Schiphof, 2006; *International study on the impact of copyright law on digital preservation*, 2008, met name het hoofdstuk over Nederland, p. 46 ff.] – maar strikt genomen zijn deze niet van belang voor de thematiek van duurzame bewaring en toegang. Zij worden in dit rapport dan ook verder niet behandeld.

2.3.3. Het verdrag van Malta

In 1992 ondertekende Nederland het Verdrag van Malta (Valletta) dat de omgang met het Europees archeologisch erfgoed regelt [Rijksdienst voor het Cultureel Erfgoed, website]. Archeologisch bodemonderzoek leidt er meestal toe dat de originele onderzoeksomgeving verloren gaat. Het is daarom wettelijk vastgelegd dat na voltooiing van het archeologisch onderzoek de vondsten en alle bijbehorende documentatie bij een depot in bewaring moeten worden gegeven. Het deponeren van digitale documentatie vindt plaats binnen de kaders van de Kwaliteitsnorm Nederlandse Archeologie (KNA) en/of het wetenschappelijk onderzoek en onderwijs [Wansleeben & van den Dries, 2008, p. 21].

2.3.4. Andere beleidskaders

Het kabinet heeft in 2006 in de nota *Informatie op Orde* zijn visie op de informatiehuishouding van de rijksoverheid gepresenteerd. Daarin worden acht zogenaamde actielijnen uitgezet, die binnen de overheid moeten leiden tot duidelijke afspraken over het beheer van overheidsinformatie, uitgaande van een integrale benadering: voor de gebruiker van de informatie mag het niet uitmaken welke organisatie de informatiehuishouding beheert. Ook werd aangekondigd dat er meer samenhang gebracht zou gaan worden in de wettelijke kaders voor de informatiehuishouding en het beheer [Informatie op Orde, 2006].

In 2007 stelde het Kabinet een beleidslijn vast die het gebruik van open standaarden bij de overheid stimuleert onder het motto: ‘comply-or-explain and commit’ [Nederland Open in Verbinding, 2007, p. 9]. Ook het gebruik van open source software wordt in de beleidslijn aanbevolen.

Het digitaal erfgoed staat centraal in Unesco’s *Charter on the preservation of the digital heritage*. Dit charter bevordert momenteel vooral het bewustzijn van de problematiek. In 2009 wordt het charter herzien [Preserving the digital heritage, 2007]. De Raad van de Europese Unie heeft zogenaamde ‘Conclusies van de Raad over digitalisering en online toegankelijkheid van cultureel materiaal en digitale bewaring’ vastgesteld [Conclusies van de Raad over digitalisering, 2006]. De Europese Unie implementeert deze in haar subsidieprogramma’s.

De eisen die aan wetenschappelijk onderzoek en aan het borgen van de resultaten worden gesteld, liggen niet wettelijk vast. Niettemin zijn er gezaghebbende kaders ontstaan die de wetenschapspraktijk bepalen:

De OECD (Organisation for Economic Co-operation and Development) werkt al vanaf 2000 aan het opstellen van algemene principes voor toegang tot onderzoeksresultaten die met publieke middelen gefinancierd zijn. Nederland speelde een voortrekkersrol bij de totstandkoming van de *Declaration on Access to Research Data from Public Funding* (Parijs, 2004) die in 2007 werd uitgewerkt in de *OECD Principles and Guidelines for Access to Research Data From Public Funding* [OECD, 2007]. Hierin worden dertien richtinggevende principes genoemd. Eén van

deze principes, relevant voor deze verkenning, is *sustainability*, nader omschreven als 'taking measures to guarantee long term access to data'. Ondanks het feit dat de OECD Principles niets zeggen over de implementatie, hebben ze veel invloed, omdat beleidsorganisaties en financiers er vaak aan refereren en op basis van de *Principles* verantwoordelijkheden benoemen.

In Nederland stelde de Algemene Rekenkamer al in 1991 vast dat wetenschappelijk onderzoek gebaat is bij een effectief beleid op het gebied van het beheer van computerbestanden [Machineleesbare gegevensbestanden, 1991]. Het rapport bepleit secundair gebruik van bestaande onderzoeksdata en beveelt aan dat eigenaren en beheerders van de bestanden schriftelijke afspraken maken over het beheer en de archivering van de bestanden.

De Gedragscode Wetenschapsbeoefening van de Vereniging van Universiteiten (VSNU) bepaalt dat onderzoek gerepliceerd moet kunnen worden om de juistheid te kunnen testen [Nederlandse Gedragscode Wetenschapsbeoefening, 2004]. Anno 2009 betekent dit dat digitale onderzoeksgegevens voor minimaal vijf jaar bewaard moeten blijven.

Zowel de KNAW als NWO hebben in hun strategische plannen benadrukt dat open toegang tot wetenschappelijke informatie essentieel is [Duurzame wetenschap, 2006, p. 22; Wetenschap gewaardeerd, 2006]. Dit heeft er o.a. toe geleid dat NWO binnen het programma investeringssubsidies middelgroot of bij subsidie-instrumenten voor dataverzamelingen verplicht stelt dat voor de borging van de onderzoeksdata een zogenaamd *datacontract* moet worden afgesloten met DANS.

2.4. Rollen en verantwoordelijkheden ten aanzien van informatie

In zijn algemeenheid kunnen we stellen dat in de sectoren overheid/archieven en cultureel erfgoed *in principe* goed benoemd is wie waarvoor verantwoordelijk is [Sorgdrager, 2006, pp. 11-14]:

- Archiefvormers/overheidsorganisaties: Elke minister en iedere provincie of gemeentebestuur et cetera is verantwoordelijk voor de kwaliteit van de eigen informatiehuishouding, inclusief een adequaat archiefbeheer.
- Minister van Binnenlandse Zaken en Koninkrijksrelaties (BZK): Het ministerie van Binnenlandse Zaken en Koninkrijksrelaties heeft tot taak om de organisatie en kwaliteit van de rijksdienst te bevorderen. BZK is eigenaar en opdrachtgever van de baten-lastendienst Centrale Archief Selectiedienst.
- Centrale Archief Selectiedienst (CAS): Bewerkt en structureert voor veel organisaties in de rijksoverheid het archiefmateriaal dat uiteindelijk wordt overgebracht naar het Nationaal Archief. De CAS vormt daarmee een schakel tussen de rijksoverheidsorganisaties en het Nationaal Archief.
- Minister van Onderwijs, Cultuur en Wetenschap (OCW): Is verantwoordelijk voor cultureel erfgoed en daarmee voor een zorgvuldig beheer van archieven in rijksarchiefbewaarplaatsen. Is tevens verantwoordelijk voor de Archiefwet.
- Nationaal Archief: Is de centrale bewaarplaats voor de overgebrachte archieven uit de rijksoverheid. De directeur van het Nationaal Archief is tegelijk Algemeen

- Regionale Historische Centra: Bewaren o.a. de archieven van organisaties van de rijksoverheid die op decentraal niveau actief zijn, en van de provincies.
- Raad voor Cultuur: Wettelijk adviesorgaan van de regering voor het cultuurbeleid.
- Erfgoedinspectie: Houdt toezicht op de naleving van de Archiefwet bij de rijksoverheid.
- Provinciale archiefinspecteurs, gemeentearchivarissen: Houden toezicht op de naleving van de Archiefwet bij provinciale en lokale overheden.
- SamenwerkingsProjecten Archieforganisaties Rijksdienst (SPAR): In 2007 opgericht samenwerkingsverband tussen de dertien departementen op het gebied van archiefwerkzaamheden

De genoemde taakafbakening geldt voor erfgoed dat in beheer is bij de publieke sector; voor grote hoeveelheden cultureel audiovisueel erfgoed die momenteel ontstaan op web 2.0 is nog geen verantwoordelijkheid toebedeeld.

In de sector wetenschap kunnen rollen en verantwoordelijkheden ten aanzien van de informatie die voortkomt uit wetenschappelijk onderzoek niet afgeleid worden uit een wettelijk kader. Onderzoekers werken vaak als individuen of in spontaan georganiseerde groepen. In het rapport *Dealing with data* [Dealing with data, 2007] heeft Liz Lyon een kader geschetst waarbinnen de verantwoordelijkheden benoemd zouden kunnen worden, maar dit is nog geen gevestigde praktijk (tabel 2.1) [Dealing with data, 2007].

De meeste uitgevers zien voor zichzelf niet echt een rol ten aanzien van langetermijntoegang tot wetenschappelijke publicaties, zo wees een recente studie door de Association of Learned and Professional Society Publishers (ALPSP) uit [Long-term Preservation, 2008]. Zij leggen de verantwoordelijkheid vooral bij bibliotheken, die immers ook in het analoge tijdperk de langetermijnbewaring voor hun rekening namen. Nationale bibliotheken als de KB pakken die uitdaging op; de rol van wetenschappelijke bibliotheken lijkt te verschuiven naar het bieden van *toegang*, tenzij de bibliotheek zich ontwikkelt tot een echt datacentrum, een rol die de Nederlandse *institutional repositories* (die vaak bij de bibliotheek zijn ondergebracht) in toenemende mate op zich lijken te nemen [Angevaere, 2009].

Tabel 2.1. *Mogelijke verdeling van rollen en verantwoordelijkheden ten aanzien van onderzoeksdata in de wetenschap* [Dealing with data, 2007; vert. René van Horik]

Rol	Rechten	Verantwoordelijkheden
<i>Wetenschapper:</i> creatie en gebruik van data	Primair gebruik van de data Wetenschappelijke erkenning Eigendomsrecht Ontvangen van training en advies	Beheer van de data gedurende het onderzoeksproject In acht nemen van relevante richtlijnen en regelgeving Data bewerken zodat anderen deze ook kunnen gebruiken
<i>Instelling:</i> bewaren van en toegang verschaffen tot data	Verkrijgen van kopie van de data	Instellen interne richtlijnen voor databeheer Databeheer op korte termijn Toepassen van standaarden Geven van training en advies aan wetenschappers Promoten van 'repository' dienst
<i>Datacentrum:</i> datacuratie van en toegang verschaffen tot data	Verkrijgen van kopie van de data Selectie van data van waarde voor langetermijn-bewaring	Beheer van data voor de lange termijn Toepassen van standaarden Aanbieden van training m.b.t. datadeponering Promoten van dienstverlening Beschermen van relevante rechten Aanbieden van <i>tools</i> voor hergebruik van de data
<i>Gebruiker:</i> (her)gebruiken van bestaande data	Hergebruiken van data (<i>non-exclusive</i>) Toegang tot metadata om bruikbaarheid te kunnen bepalen	Erkennen van licentievoorwaarden Erkennen van dataproducenten Effectief beheer van (afgeleide) data
<i>Financier:</i> opstellen van, aansluiten bij, reageren op regelgeving	Implementeren van regelgeving Eisen dat uitvoerders zich aan regelgeving houden	Breed perspectief m.b.t. belanghebbenden in acht nemen Participeren in coördinatie van de data strategie Ontwikkelen van beleid met belanghebbenden Coördinatie van beleidsontwikkeling Monitoren en opleggen van regelgeving op het gebied van data Zorgdragen voor langetermijndatabeheer na afloop van het project Stimuleren van datacuratie en financieren van deskundige dienstverlening Steunen van ontwikkeling van expertise van datacuratoren
<i>Uitgever:</i> integriteit van de wetenschappelijke output garanderen	Verwachten dat data beschikbaar is om de publicatie te ondersteunen Verzoeken om data op te slaan in langetermijnopslagplaats	Betrekken van stakeholders bij ontwikkeling van publicatiestandaarden Verwijzen naar data Ontwikkelen en monitoren van standaarden

'Ik schrik er nog steeds van als ik zie hoe slecht de overheid met zijn informatie omgaat. Soms wordt de meest elementaire anti-virussoftware niet eens bijgehouden.'
(een archivaris)

'Als ik hoor dat Archief X zijn duurzame opslag gaat uitbesteden, maak ik me zorgen ... maar ja, als ik hoor dat ze het zelf gaan doen, maak ik me ook zorgen ...' (ICT manager)

3. Duurzame toegang tot digitale objecten: risico's en beheersmaatregelen

Om de vraag te kunnen beantwoorden waar in de publieke sector digitale gegevens goed worden beheerd en waar informatie het risico loopt verloren te gaan, moeten we eerst bepalen wat de specifieke risico's zijn die digitale objecten lopen en wat we verstaan onder 'goed' beheer.

3.1. De risico's

In de jaren tachtig van de vorige eeuw begon langzaam maar zeker het besef te ontstaan dat digitale technieken niet alleen ongekende mogelijkheden boden, maar dat er ook risico's aan kleven waarvoor oplossingen gezocht moesten worden. In 1995 vatte Jeff Rothenberg de technische problemen samen [Rothenberg, 1995]:

- de dragers (cd's, dvd's, floppy disks) hebben een beperkte levensduur, in de praktijk soms niet meer dan vijf jaar;
- de informatie bestaat uit een op zichzelf betekenisloze 'bitstream' van nullen en enen, die alleen met behulp van exact de juiste hardware/software-combinatie tot leven kan worden gewekt;
- en zelfs als aan die randvoorwaarde is voldaan, kan de betekenis alleen achterhaald worden als aan de bitstream zelf allerhande contextinformatie is toegevoegd ('metadata').

Het is tekenend dat Rothenberg het woord 'documents' in de latere versie van zijn artikel uit 1999 verving door 'information', omdat er geen sprake meer is van 'documenten' in traditionele zin. Om aan te geven hoe anders digitale informatie zich gedraagt dan papieren objecten, ontwikkelde de National Archives van Australië het begrip 'performance' om het woord document te vervangen: een bitstream leidt in samenspel met de juiste hard- en software tot een 'performance' [National Archives of Australia, 2002]. Hoewel dit concept buiten Australië niet echt is overgenomen, geeft het wel aan dat we hier te maken hebben met een totaal ander soort 'materiaal' dan papier: waar voor papieren informatie onveranderlijke (geklimatiseerde) omgevingen kunnen worden gebouwd waarin het papier eeuwenlang bruikbaar kan worden gehouden, is dat voor digitale informatie onmogelijk. Het zou betekenen dat we de vooruitgang van de techniek stil moeten zetten, en dat is net zo onwenselijk als onhaalbaar. Met iedere nieuwe ontwikkeling in hardware en software zullen we dus moeten bezien of de *bitstreams* die we beheren nog wel bruikbaar zijn, en zo niet, wat we moeten aanpassen om toegankelijkheid te waarborgen. Dat vereist een regime van continue zorg, van wieg tot graf.

De continue zorg voor digitale informatie moet beginnen bij *creatie*. Daar worden al keuzes gemaakt (over het bestandsformaat, over de metadata) die van grote invloed zijn op de mate waarin een digitaal object langdurig toegankelijk gehouden kan worden. Maar de producent heeft vaak een kortetermijnbelang: het eigen onderzoek, het afgeven van een bouwvergunning, beschikbaarstelling op de website. Het extra werk dat nodig is om ook langetermijntoegang mogelijk te maken is bij de producent onbekend of onbemind. In het

analoge tijdperk kon die langetermijnzorg wachten totdat het document – vaak jaren later – werd overgedragen aan een archief, maar dat kan nu niet meer. Als we nu twintig jaar wachten met het verlenen van langetermijnzorg is de kans groot dat de informatie al onbruikbaar is geworden. Experts spreken daarom van een *records continuum* of een *data curation continuum*: korte en langetermijnbelangen ten aanzien van digitale gegevens spelen tegelijkertijd [Jonker, 2009; Treloar, Groenewegen en Harboe-Ree, 2007]. Maar onze samenleving is zo niet georganiseerd, er is vaak geen sprake van centrale regie over de hele levenscyclus van informatie. Gebrek aan afstemming kan dan gemakkelijk leiden tot informatieverlies.

3.2. Technische beheersmaatregelen

Anno 2009 zijn de nodige beheersmaatregelen ontwikkeld om de risico's van digitale informatie te beheersen, maar dé oplossing is nog niet beschikbaar. De meeste experts zijn het er inmiddels over eens dat de techniek zich zal blijven ontwikkelen met nieuwe mogelijkheden, nieuwe risico's en nieuwe beheersmaatregelen.

Zonder al te diep op de techniek in te gaan, noemen we hier enkele technische aspecten die bijdragen aan het beheersen van de risico's die digitale objecten lopen [voor een overzicht van de beschikbare literatuur ten aanzien van de techniek zie digitaleduurzaamheid.nl, website]:

3.2.1. De juiste opslagmedia

Er zijn dragers met een beperkte levensduur (cd, dvd, floppy disk) en dragers met een langere levensduur [zie o.a. DAVID vademecum, website, C2].

3.2.2. De juiste bestandsformaten

Iedere softwaretoepassing creëert weer een ander bestandsformaat; er zijn er duizenden in omloop, het één toekomstvaster dan het ander. Van MS Office documenten is bekend dat ze zeer afhankelijk zijn van specifieke software(versies); dat maakt ze moeilijk duurzaam te bewaren. PDF-A is momenteel zeer gangbaar voor langetermijnbewaring van tekstdocumenten, omdat de leesbaarheid veel minder afhankelijk is van specifieke software. Obscure bestandsformaten gelden weer als minder duurzaam dan veel gebruikte bestandsformaten, omdat men mag verwachten dat voor veel gebruikte formaten altijd wel iemand een vertaling naar het volgende bestandsformaat zal ontwikkelen. Een platformafhankelijk formaat als .xml zal het weer beter doen dan een platformafhankelijk formaat.

3.2.3. Migratie, emulatie en normalisatie

Migratie, emulatie en normalisatie zijn de bekendste strategieën ('preservation strategies') die momenteel worden genoemd om digitale informatie als het ware mee te laten groeien met de techniek:

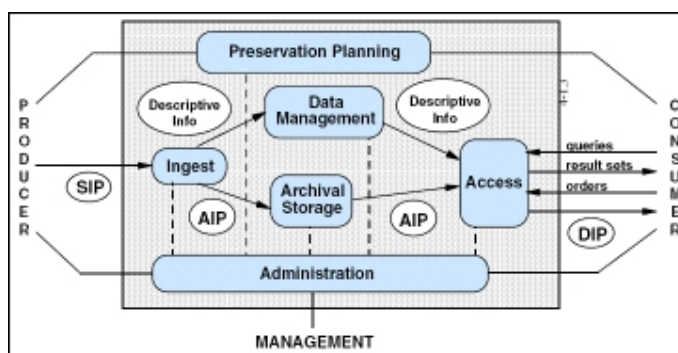
- Migratie is een techniek waarbij een bitstream zo wordt aangepast dat de informatie geraadpleegd kan worden op een andere software-hardwarecombinatie dan die waarop de informatie is geproduceerd. Deze techniek is niet zonder gevaren, want de bitstream moet veranderd worden; hoe kan men controleren of het eindresultaat voldoende correspondeert met het origineel?

- Bij emulatie wordt niet de bitstream veranderd, maar wordt software gebouwd die een nieuwe computer kan laten functioneren als een ouder model. Vooralsnog lijkt deze techniek duurder dan migratie, maar op de lange termijn en voor bijzondere digitale objecten als bijvoorbeeld websites zou deze techniek wellicht de enige oplossing kunnen zijn. In Nederland ontwikkelden de Koninklijke Bibliotheek en het Nationaal Archief de emulatiesoftware Dioscuri [Dioscuri, website].
- Normalisatie is eigenlijk een bijzondere vorm van migratie. Een organisatie kan besluiten om in het digitale archief niet alle mogelijke aangeboden bestandsformaten op te nemen, maar alle informatie bij binnenkomst te converteren naar één robuust bestandsformaat.

3.3. Beleidsmatige en organisatorische beheersmaatregelen

In de loop der jaren groeide het besef dat technische oplossingen alleen niet voldoende zijn om de digitale risico's te beheersen. Om technische oplossingen te kunnen ontwikkelen en implementeren zijn beleidsmatige en organisatorische randvoorwaarden nodig.

Een belangrijke mijlpaal was de totstandkoming van het *Open Archival Information System*, oftewel OAIS-model, dat in 2000 werd ontwikkeld door de organisatie die de data van de NASA ruimtemissies beheert, het Consultative Committee for Space Data Systems (CCSDS), en dat later werd verheven tot ISO-standaard 14721 [Lavoie, 2004]. Het model zorgde voor een begrippenkader waarmee de techniek in een beheersorganisatie kan worden ingebed en de relaties met de buitenwereld worden gedefinieerd. Dit model is zo richtinggevend gebleken in de manier waarop over de materie wordt gedacht, dat iedere organisatie die duurzaam met zijn data om wil gaan gericht beleid moet maken ten aanzien van de functies die het OAIS-model benoemt: de digitale opslag op zich ('archival storage'), 'administration', 'data management' en 'preservation planning', en de relaties met de producenten ('ingest') en die met de gebruikers ('access').



In 2002 publiceerde OCLC het rapport *Trusted digital repositories: attributes and responsibilities* [Trusted Digital Repositories, 2002], dat gebaseerd was op het OAIS-model en waarin wordt beschreven welke kenmerken en verantwoordelijkheden een rol spelen bij het

inrichten van een duurzame digitale bewaarplaats voor onderzoeksdata. Deze bewaarplaats wordt een 'Trusted Digital Repository' (TDR) genoemd.

Ook ontstonden toetsingsinstrumenten als TRAC (Trustworthy Repositories Audit & Certification Criteria and Checklist), DRAMBORA (Digital Repository Audit Method Based on Risk Assessment), DAF (Data Audit Framework) en de *Kriterienkatalog vertrauenswürdige digitale Langzeitarchive* van het Duitse Kompetenznetzwerk 'nestor' [zie de verschillende websites en, o.a., Sierman, 2009]. Het Landelijk Overleg Provinciale Archiefinspecteurs (LOPAI) werkte TRAC uit tot een concepttoetsingskader voor digitale depots in de Nederlandse archiefsector, *ED3, Eisen Duurzaam Digitaal Depot* [ED3, 2008].

Deze toetsingsinstrumenten worden nog niet in de praktijk toegepast, omdat zij zich in verschillende ontwikkelingsstadia bevinden. Niettemin hebben zij de inspiratie geleverd voor een aantal normenkaders die in Nederland gehanteerd worden.

3.4. In Nederland gebruikte richtlijnen

Hieronder beschrijven wij enkele normenkaders voor omgang met digitale informatie die in Nederland worden gebruikt. Naast een algemene norm (NEN-ISO 15489-1) besteden we aandacht aan drie recentelijk in Nederland tot ontwikkeling gekomen instrumenten: de Baseline bij de rijksoverheid [De Baseline, 2008], het door DANS ontwikkelde Data Seal of Approval voor de wetenschap [Data Seal of Approval, website]; en DEN's De Basis voor duurzaamheid [De Basis voor duurzaamheid, 2009].

3.4.1. NEN-ISO 15489-1

De eerste norm die genoemd moet worden is NEN-ISO 15489-1 (nl) die op een tamelijk algemeen en abstract niveau beschrijft wat er allemaal gedaan moet worden om tot een adequate inrichting van een informatiehuishouding te komen – digitaal of analoog [NEN-ISO 15489-1, 2001]. Elementen zijn beleid, toedeling van verantwoordelijkheden, kwaliteitsbewaking, auditing en opleidingen. Met zijn nadruk op archiefbescheiden, verantwoording en bedrijfskapitaal zien we dat NEN-ISO 15489-1 vooral bij de overheid wordt toegepast en verwerkt in specifieke normen, zoals de Baseline. Ook het bedrijfsleven noemt de norm als handvat [Bedrijfsonderzoek, 2009, p. 11]. De norm is niet geschikt voor het beheer van collecties.

3.4.2. De Baseline informatiehuishouding rijksoverheid

In het kader van het programma 'Informatie op Orde' is in 2008 de *Baseline informatiehuishouding rijksoverheid: normenkader voor duurzaam toegankelijke en betrouwbare overheidsinformatie* vastgesteld. De Baseline geldt zowel voor analoge als digitale informatie, maar de 'zeven normen' ademen in hun formuleringen de digitale praktijk anno 2009 [de Baseline, 2008, deel III].

Tabel 3.1. De Baseline: de zeven normen

BL1	<i>Verantwoordelijkheden.</i> De verantwoordelijkheden die het lijnmanagement namens de minister draagt voor de duurzame toegankelijkheid en betrouwbaarheid van informatie zijn belegd en beschreven. (De toelichting verwijst naar NEN-ISO 15489)
BL2	<i>Inrichting organisatie.</i> De inrichting van organisatie, processen, personeel en hulpmiddelen is kwantitatief en kwalitatief toereikend voor de borging van duurzame toegankelijkheid en betrouwbaarheid. (Uit de toelichting: 'In de digitale wereld dienen al vanaf het moment van creatie van informatie de juiste maatregelen ten behoeve van duurzame toegankelijkheid en betrouwbaarheid te worden genomen.')
BL3	<i>Informatiebeleid.</i> Voor verschillende aspecten van het informatiebeheer zijn binnen het ministerie of bestuursorgaan standaarden gedefinieerd en in gebruik. (Uit de toelichting: 'Standaarden betreffen: begrippenkaders, architecturen, metadatasets, etc.').
BL4	<i>Informatieontwerp.</i> Er is een – geprioriteerde – classificatie gemaakt van producten, processen, informatie en verantwoordelijkheden, waarbij rekening is gehouden met wet- en regelgeving.
BL5	<i>Informatiesysteem.</i> Een – al of niet geautomatiseerd – systeem waarmee overheidsinformatie wordt beheerd ondersteunt aantoonbaar de eisen van duurzame toegankelijkheid en betrouwbaarheid, op het niveau van het geldende beheerregime.
BL6	<i>Creatie en gebruik van informatie.</i> Bij het creëren en gebruiken van overheidsinformatie worden de kwaliteitseisen voor duurzame toegankelijkheid en betrouwbaarheid in acht genomen.
BL7	<i>Beheer van informatie.</i> De duurzame toegankelijkheid en betrouwbaarheid van overheidsinformatie is tot het moment van verwijdering gewaarborgd, en de verwijdering verloopt conform vastgelegde procedures.

3.4.3. Data Seal of Approval (DSA)

Eerder genoemde toetsingskaders als TRAC en DRAMBORA (zie par. 3.3) zijn vaak omvangrijke documenten die een intensieve inspanning van de betrokken organisaties vereisen – zo kan ook de KB beamen die in 2006 een proef deed met een voorloper van het TRAC-instrument. Dat bracht DANS ertoe om een laagdrempelig *Data Seal of Approval* (DSA) te ontwikkelen. Het DSA wil ertoe bijdragen dat onderzoeksdata op de lange termijn vindbaar, toegankelijk, bruikbaar en betrouwbaar blijven, en dat er in de literatuur altijd naar verwezen kan blijven worden. Inmiddels heeft DANS het DSA overgedragen aan een internationale ‘editorial board’ bestaande uit vertegenwoordigers van verschillende Europese data-archieven. Het bestuur zal het DSA periodiek beoordelen en indien nodig bijstellen. Daarnaast zal het bestuur het beoordelingstraject bewaken dat ingesteld gaat worden om organisaties te laten voldoen aan de richtlijnen die onderdeel zijn van het DSA.

Tabel 3.2. *Data Seal of Approval Guidelines* (versie juni 2009)

DSA01	The data producer deposits the research data in a data repository qualified according to the DSA guidelines
DSA02	The data producer provides the research data in formats recommended by the data repository
DSA03	The data producer provides the research data together with the metadata requested by the data repository
DSA04	The data repository has an explicit mission in the area of digital archiving and promulgates it
DSA05	The data repository uses due diligence to ensure compliance with legal regulations and contracts
DSA06	The data repository applies documented processes and procedures for managing data storage
DSA07	The data repository has a plan for long-term preservation of its digital assets
DSA08	Archiving takes place according to explicit workflows across the data life cycle
DSA09	The data repository assumes responsibility from the data producers for access and availability of the digital objects
DSA10	The data repository enables the users to utilize the research data and refer to them
DSA11	The data repository ensures the integrity of the digital objects and the metadata
DSA12	The data repository ensures the authenticity of the digital objects and the metadata
DSA13	The technical infrastructure explicitly supports the tasks and functions described in internationally accepted archival standards like OAIS
DSA14	The data consumer complies with access regulations set by the data repository
DSA15	The data consumer conforms to and agrees with any codes of conduct that are generally accepted in higher education and scientific research for the exchange and proper use of knowledge and information
DSA16	The data consumer respects the applicable licences of the data repository regarding the use of the research data

3.3.4. De Basis voor duurzaamheid

Ten behoeve van het cultureel erfgoed ontwikkelt DEN in samenwerking met het veld ‘De Basis’ – en set van minimale eisen voor digitalisering van cultureel erfgoed [DEN De Basis, website]. De Basis duurzaamheid is vastgesteld voor twee jaar en zal in 2011 geëvalueerd en mogelijk herzien worden. De richtlijn voor duurzaamheid bevat veel elementen die we ook in de eerder genoemde richtlijnen terugvinden ten aanzien van beleid en organisatie. Op enkele punten is De Basis specifiek dan De Baseline en het DSA [DEN De Basis voor duurzaamheid, website]:

Tabel 3.3. De Basis voor duurzaamheid – een selectie

[NB: de nummering is niet origineel, maar in dit rapport toegevoegd om verwijzing te vergemakkelijken]

[DB0]	<i>Digitale duurzaamheid vanaf het begin.</i> Keuzes die nu gemaakt worden bij de vervaardiging van digitale data, zoals het bestandsformaat, de kwaliteit en de configuratie, zijn van belang voor de toegankelijkheid ervan op langere termijn. Raadpleging is alleen mogelijk als de instelling een strategie bepaalt voor het langdurig bewaren van digitale informatie en ervoor zorgt dat toekomstige ICT-omgevingen om kunnen blijven gaan met de bestanden die nu vervaardigd worden. Om de consequenties voor de lange termijn te kunnen overzien, is een voortdurend besef van digitale duurzaamheid nodig.
[DB1]	Fase 1. Beleid Iedere erfgoedinstelling produceert digitale data voor gebruik door derden. Om deze reden dient iedere erfgoedinstelling een duurzaamheidsstrategie te kiezen en daar beleid op te maken. Juist ook vanwege het beheer en behoud van (kennis over) digitaal erfgoed kan geen enkele erfgoedinstelling het zich permitteren het onderwerp duurzaamheid te negeren. Bij het opstellen van duurzaamheidsbeleid dient de instelling aandacht te besteden aan de volgende vier aspecten:
[DB1.1]	<i>Documentatie.</i> Het inzicht van de instelling in de problematiek wordt in de vorm van een beleidsdocument (bij voorkeur een informatieplan) expliciet gemaakt en vastgelegd.
[DB1.2]	<i>Organisatie.</i> De instelling neemt maatregelen om interne kennisdeling over digitale duurzaamheid te ondersteunen. De verantwoordelijkheden worden beredeneerd belegd binnen de organisatie.
[DB1.3]	<i>Middelen.</i> Om duurzame toegankelijkheid te ondersteunen wordt er structureel een budget voor gereserveerd.
[DB1.4]	<i>Technische oplossing.</i> Iedere erfgoedinstelling die digitaliseert dient te bepalen welke systemen in eigen beheer genomen worden en welke elders worden ingezet om duurzame toegankelijkheid te garanderen. Gaat de instelling zelf investeren in een digitaal archief, of wordt gebruik gemaakt van een 'repository' van een andere instelling?
[DB2]	Fase 2. Plannen maken. Nadat het beleid is vastgesteld, zal de erfgoedinstelling plannen maken voor het digitaliseren van de collecties en het beheren van de aldus geproduceerde digitale data. In deze fase worden de volgende eisen gesteld aan alle erfgoedinstellingen:
[DB2.1]	<i>Wat wordt bewaard?</i> Erfgoedinstellingen dienen criteria op te stellen waarmee bepaald kan worden welke data langdurig bewaard dienen te worden. Deze criteria hangen nauw samen met het collectieprofiel, het acquisitie- of het presentatiebeleid.
[DB2.2]	<i>Criteria voor bestandsformaten.</i> Erfgoedinstellingen dienen criteria te formuleren voor het accepteren dan wel verwerpen van bestandsformaten die in gebruik zijn in de digitale dienstverlening.
[DB2.3]	<i>Uitvoerbare plannen.</i> Erfgoedinstellingen dienen op basis van de vastgestelde criteria en nog vóór de vervaardiging van digitale data heldere en uitvoerbare plannen klaar te hebben voor duurzame opslag en toegankelijkheid.
[DB2.4]	<i>Kennis van externe voorzieningen.</i> Erfgoedinstellingen dienen kennis te nemen van criteria die andere partijen stellen voor nationale voorzieningen die duurzame opslag en toegankelijkheid regelen.
[DB3]	Fase 3. Uitvoering Instellingen die concrete oplossingen toepassen voor duurzame toegankelijkheid van digitale data, dienen zich aan de volgende basiseisen te houden:
[DB3.1]	<i>Kennisuitwisseling.</i> Iedere erfgoedinstelling stelt in de organisatie één aanspreekpunt aan voor de duurzaamheidsactiviteiten ter bevordering van kennisuitwisseling met andere uitvoerende partijen.
[DB3.2]	<i>Referentiemodel.</i> Erfgoedinstellingen die bouwen aan een eigen repository, gebruiken het referentiemodel OAIS als uitgangspunt.
[DB4]	Fase 4. Evaluatie Erfgoedinstellingen die systemen beheren die zorg dragen voor duurzame toegankelijkheid van digitale data, worden gehouden aan de volgende basiseis voor evaluatie van de eigen duurzame oplossing:
[DB4.1]	<i>Toetsing.</i> Erfgoedinstellingen die een eigen repository beheren hanteren TRAC voor de toetsing van de repository.

3.5. Impact van de richtlijnen

Voordat we in hoofdstuk 4 naar de praktijk in de publieke sector gaan kijken is het belangrijk te benadrukken dat bovengenoemde richtlijnen, met uitzondering van NEN-ISO 15489, *erg jong* zijn. Dat betekent dat we de praktijk niet zomaar kunnen 'afrekenen' op de specifieke bewoordingen van deze normen/richtlijnen. Wel geven de richtlijnen een globale indruk van de huidige consensus ten aanzien van de vraag wat 'goed' beheer inhoudt:

- er is sprake van vastgesteld beleid gericht op duurzame toegang (BL3, DSA04-08, DB1.1);
- er is een organisatie ingericht met het oog op duurzame toegang (BL2, DSA08, DB1.2);
- de inrichting van een bewarende organisatie is in overeenstemming met het OAIS referentiemodel (DSA13, DB3.2);
- al bij creatie van digitale gegevens worden kwaliteitseisen voor duurzame toegankelijkheid toegepast, en dit blijft zo gedurende de hele levenscyclus van de informatie (BL6, DSA01, DB0).

Ten slotte moet worden opgemerkt dat op bovengenoemde richtlijnen in de praktijk (nog) niet of nauwelijks wordt getoetst. Het toetsingskader van de Erfgoedinspectie, waaronder vrijwel alle erfgoedinstellingen vallen, is vooral gericht op fysiek beheer en de kwaliteit van registratie.

De Baseline is inmiddels door de overheid als norm vastgesteld, en De Basis duurzaamheid wordt als voorwaarde genoemd in een aantal subsidieregelingen.

4. De publieke sector en digitale informatie

Dit hoofdstuk benoemt de rol en het belang van digitale objecten voor de drie sectoren om vervolgens stil te staan bij de vraag of alle geproduceerde informatie ook bewaard moet worden. Ten slotte kijken we naar de praktijk: waar wordt digitale informatie goed beheerd en welke belemmeringen spelen een rol wanneer er geen sprake is van goed beheer?

4.1. Rol en belang van digitale objecten

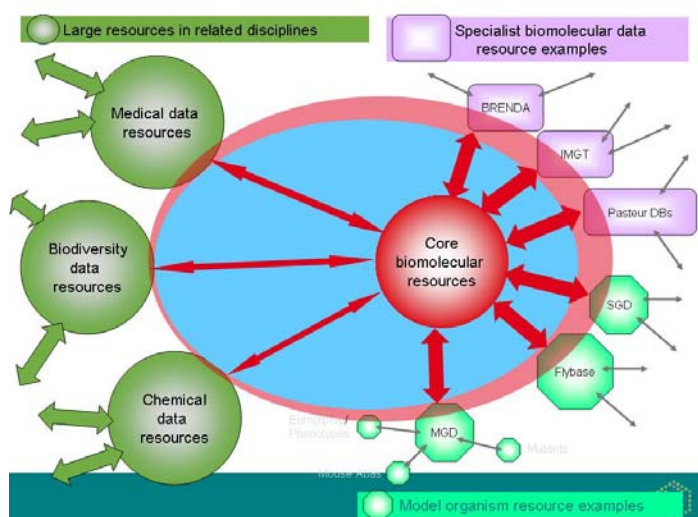
Van de drie onderzochte sectoren is de *wetenschap* de sector die de mogelijkheden van het digitale tijdperk het snelst heeft onderkend en toegepast. De huidige wetenschap is volstrekt ondenkbaar zonder grootschalig en gebruik van digitale gegevens.

De exacte wetenschappen namen het voortouw in deze ontwikkeling. Ruimtevaart, natuurkunde en chemie ontwikkelden grootschalige onderzoeksfaciliteiten gebaseerd op digitale onderzoeksdata. Maar ook de sociale wetenschappen ontdekten al in de jaren zestig het nut van de computer bij het produceren en analyseren van kwantitatieve onderzoeksdata. Binnen de geesteswetenschappen zijn taalkunde (tekstcorpora), archeologie en sociale en economische geschiedenis sterk geautomatiseerd; literatuurwetenschap, kunstgeschiedenis, filosofie en theologie maken nog maar weinig gebruik van digitale gegevens.

De toepassing van ICT in het wetenschappelijk onderzoek heeft geleid tot de ontwikkeling van wat 'e-science' wordt genoemd. Onder die term worden samenwerkingsverbanden gevat die gebruik maken van gedistribueerde rekenkracht en databronnen overal over de wereld, wat leidt tot grootschalige virtuele onderzoeksfaciliteiten. De natuurkunde en chemie lopen voorop in deze ontwikkelingen.

Figuur 1 illustreert de grootschaligheid en complexiteit van dergelijke infrastructuren.

Fig. 1. Overzicht van een netwerkinfrastructuur in de bioinformatica



[bron: Cameron, 2008]

Nederland beschikt over een goede netwerkinfrastructuur (SURF6/Gigaport) en vernieuwde supercomputers (Huygens/Blue Gene). Daarin worden o.a. gegevens van de Large Hadron Collider (LHC) van CERN opgeslagen. In het kader van het BIG GRID project werken het Nationaal Instituut voor Subatomaire Fysica (NIKHEF), de NWO National Computing Facilities Foundation (NCF) en het Netherlands Bioinformatics Centre (NBIC) aan een grootschalige e-science infrastructuur voor Nederland.

Naar de *hoeveelheid* digitale objecten die omgaan in de wetenschap kunnen we alleen maar raden. We noemen enkele willekeurige getallen: het LOFAR (Low Frequency Array) project van het Nederlandse ASTRON, dat 25.000 radioantennes plaatst, verwacht maandelijks ten minste 1000 terabyte aan gegevens te genereren; de Large Hadron Collider in Genève zal naar verwachting jaarlijks 15000 terabyte aan data voortbrengen. Ter vergelijking: de totale collectie van 10 miljoen wetenschappelijke artikelen in het e-Depot van de KB nam in 2007 10 terabyte aan ruimte in.

De *verscheidenheid* aan digitale objecten is in de wetenschap momenteel het grootst van alle sectoren. We onderscheiden:

- wetenschappelijke publicaties
- octrooien en patenten
- onderzoeksdata, onder te verdelen in:
 - ♦ gegevens die het resultaat zijn van wetenschappelijke experimenten die in principe herhaald kunnen worden (al dan niet tegen hoge kosten);
 - ♦ modellen en simulaties, waarbij het model en de metadata van hogere waarde zijn dan de gegevens die uit de modellen voortkomen;
 - ♦ observatiegegevens van specifieke fenomenen op een bepaald tijdstip of een bepaalde plaats; deze data zijn uniek en kunnen niet opnieuw gegenereerd worden;
 - ♦ afgeleide data: van 'ruwe' data afgeleide gegevens;
 - ♦ referentiedata, bijvoorbeeld teksten uit de literatuur, genendatabank, chemische structuren.
- een aparte categorie vormen de *verrijkte publicaties*: publicaties die rechtstreeks gekoppeld zijn aan verzamelingen onderzoeksdata.

Binnen de *overheid* en *cultuur/erfgoed* zien we dat digitale objecten vooralsnog een veel kleinere rol spelen. De overheid voert een actief beleid om de voordelen van elektronische informatie-uitwisseling te gebruiken ten behoeve van een goede dienstverlening aan de burger [o.a. Een dementerende overheid, 2005, p. 7], maar de te bewaren *archiefbescheiden* zijn momenteel nog vaak van papier. Veel instanties maken gebruik van elektronische document management systemen, maar aan het eind van de rit wordt een uitdraai gemaakt die wordt voorzien van een zogenaamde 'natte' handtekening; dát document is het bewijsstuk in de zin van de Archiefwet. In interviews lieten diverse archiefinspecteurs weten dat ze deze praktijk ook aanmoedigen, omdat het digitale proces in veel gevallen nog niet betrouwbaar genoeg is om de voor archiefbescheiden zo belangrijke authenticiteit te garanderen. Die authenticiteit ontleent een bestand aan het *proces* dat het heeft doorgemaakt. Zoals een archiefinspecteur

opmerkte: 'Een archivaris heeft niets aan losse flodders informatie'; het gaat om de combinatie tussen de objecten en de rol in het proces.

Wanneer archiefvormers of archieven papieren archiefbescheiden willen vervangen door gedigitaliseerde versies, is daarvoor toestemming nodig van het Nationaal Archief (het rijk) of Provinciale Staten (lokale overheden). In de praktijk wordt die toestemming verleend door de Archiefinspectie, die strenge eisen stelt. Zo'n *machtiging tot substitutie* is ook nodig wanneer instanties hun hele werkproces willen digitaliseren. Momenteel hebben slechts twee gemeentes in Nederland daarvoor toestemming (Oss en Deventer), en dan nog alleen voor enkele met name genoemde trajecten.

Dit alles leidt ertoe dat de respondenten uit ons onderzoek van mening verschillen over de mate waarin er nu al sprake is van digitale archiefbescheiden. Maar niemand twijfelt eraan dat de toekomst aan het digitale archief is en dat de sector zich dus moet voorbereiden.

Nu al krijgen archieven in toenemende mate te maken met digitale objecten uit andere bronnen dan de archiefvormers in de overheid.

Digitaal cultureel erfgoed kent drie verschijningsvormen [ABC-DE, 2008, p. 59]:

- oorspronkelijk digitaal (*born digital*) erfgoed: erfgoedmateriaal dat van origine digitaal is, zoals elektronische archieven, digitale kunst of foto's die met een digitale camera zijn gemaakt;
- gedigitaliseerd erfgoed: erfgoed dat van origine niet digitaal is, maar waarvan een digitale reproductie is gemaakt;
- digitale informatie over erfgoed, bijvoorbeeld beschrijvingen, detailfoto's of digitale reconstructies van het erfgoedobject.

Oorspronkelijk digitaal materiaal (*born digital*) is voornamelijk te vinden in de sector media: de lopende productie van de publieke omroepen stroomt in digitale vorm de archieven van het Nederlands Instituut voor Beeld en Geluid binnen. Ook nemen de collecties digitale foto's snel in omvang toe. Het Nederlands Instituut voor Mediakunst Montevideo (NIMk) meldt dat het nog slechts op zeer beperkte schaal originele digitale kunst verzamelt [interview].

In vele grote en kleine, landelijke en lokale projecten wordt analoog erfgoed gedigitaliseerd. Soms gebeurt dat om te voorkomen dat analoge informatie definitief verloren gaat (bijvoorbeeld *Beelden voor de Toekomst*), maar vaak is het doel een groter publieksbereik via internet (*Het Geheugen van Nederland*, *De Leeuwarder Courant digitaal*, *Fotocollectie Zeeuws Documentatiecentrum*, en vele andere).

In 2008 voerde DEN het onderzoek 'De digitale feiten' uit in de cultureel erfgoedsector. Bevraagd werden erfgoedinstanties die (delen van) hun collecties gedigitaliseerd hebben (instellingen zonder digitale collecties werden niet meegerekend). De respondenten gaven gemiddeld aan dat 26% van hun collecties is gedigitaliseerd, 42% nog gedigitaliseerd moet worden en 32% niet hoeft te worden gedigitaliseerd. Dat al 26% gedigitaliseerd zou zijn komt overeen met de berekeningen in *De Collectiebalans* die een digitaliseringsgraad van museale

collecties rapporteert die ligt tussen de 17% en 37% [De Collectiebalans, 2008]. Maar er zijn grote verschillen binnen de sector: van de schilderijencollecties is inmiddels 65% digitaal gereproduceerd, maar van audio-opnamen maar 6% en van oude drukken slechts 0,3% [De digitale feiten, 2009, resp. pp. 28-29, p. 7].

Fig. 4.1. Omvang gedigitaliseerd erfgoed [De digitale feiten, 2008]

Omvang gedigitaliseerd erfgoed uit het rapport De Digitale Feiten					
Erfgoedmateriaal	eenheid	materiële collectie	waarvan digitaal beschreven	waarvan digitaal geproduceerd	aantal getelde collecties
Archiefmateriaal					
Collectie (macro niveau)	meter	223.871	152.385 (68%)	786 (0,4%)	39
Objecten (micro niveau)	aantal		1.390.225	43.859	14
Tekst					
Oude drukken (tot 1850)	band	1.051.219		2.554 (0,3%)	51
Overige boeken	band	6.489.020		1.458 (0,02%)	63
Tijdschriften	jaargang	1.329.401	304.498 (23%)	3.219 (0,2%)	37
Microfiches/microfilm	aantal	958.349	373.497 (43%)	29.021 (3%)	24
Bladmuziek	aantal	100.072	90.734 (90%)	3	10
Overig gedrukt materiaal	aantal	6.698.867		320.720 (5%)	24
Beelden					
Kaarten en plattegronden	aantal	607.688		21.657 (4%)	47
Foto's zonder negatieven	aantal	6.730.317		1.155.030 (17%)	63
Negatieven	aantal	10.174.666	1.267.466 (12%)	740.585 (7%)	36
Prenten	aantal	892.994	599.559 (67%)	185.044 (21%)	49
Tekeningen	aantal	378.262	268.155 (71%)	120.030 (32%)	40
Posters	aantal	310.537	234.644 (76%)	194.840 (63%)	38
Briefkaarten	aantal	203.946	110.323 (54%)	85.366 (42%)	29
Schilderijen	aantal	48.904	47.118 (96%)	31.035 (65%)	46
Overig 2-dimensionaal materiaal	aantal	1.011.041	79.987 (8%)	56.120 (6%)	30
Objecten					
Natuurhistorische objecten	aantal	15.063.787		9.800 (0,06)	9
3-dimensionale objecten van beeldende kunst	aantal	272.179	265.401 (97%)	134.552 (50%)	35
Overige objecten in museale collecties	aantal	1.899.141	1.435.279 (75%)	689.982 (36%)	47
Multimedia					
Film	uur	103.879	36.342 (35%)	17.784 (17%)	35
Video-opnamen	uur	221.776	32.015 (14%)	28.989 (13%)	35
Muziek en andere geluidsopnames	uur	561.451	53.310 (9%)	32.658 (6%)	32
Overige Multimedia items	aantal	2.219	606 (27%)	410 (18%)	15

Uit dit overzicht blijkt dat de instellingen veel digitale informatie over hun fysieke collecties beheren (metadata, kolom 4), maar dat de hoeveelheid echt gedigitaliseerd materiaal een stuk kleiner is. Toch praten we nu al over vele miljoenen objecten en de hoeveelheden nemen met de dag toe – *Beelden voor de Toekomst* heeft begin 2009 al 1 petabyte (=1024 terabyte) aan gedigitaliseerd audiovisueel materiaal opgeleverd en de KB verwacht de komende drie jaar 21 miljoen objecten oftewel 275 terabyte aan bestanden uit de door de KB geleide landelijke digitaliseringsprojecten (*Metamorfoze*).

4.2. De noodzaak van selectie

Enkele jaren geleden werd er nog gespeculeerd dat het in het digitale tijdperk mogelijk zou worden om *zo veel mogelijk* informatie (automatisch) te bewaren en dat intelligente zoekmachines zouden zorgen voor goede doorzoekbaarheid. Dit denken was mede ingegeven door de dalende kosten van opslagmedia. Inmiddels is het besef doorgedrongen dat vooral het *beheer* van de informatie over langere termijn bepalend is voor de kosten. Het thema *selectie* is daarmee weer hoger op de agenda komen te staan: welke informatie is zo waardevol voor de samenleving van de toekomst dat langetermijnbewaring de moeite waard is?

Experts becijferden al in 2007 dat de mensheid meer digitale informatie produceert dan zij opslagmogelijkheden heeft [The digital universe, 2007]; alles opslaan is dus sowieso een

onmogelijkheid. Toch is er een stroming vakgenoten die vindt dat we ten minste moeten proberen om zoveel mogelijk te bewaren omdat wij niet kunnen beoordelen wat de toekomst nodig zal hebben (o.a. Steve Knight, LIFE2 conferentie 2008). De Raad voor Cultuur en de Raad voor het openbaar bestuur zijn een andere mening toegedaan:

‘Selectie moet, juist ook bij een digitale informatievoorziening. Digitale informatie vraagt zorg van de wieg tot het graf en een strenge voorselectie is daarbij noodzakelijk want álle informatie bewaren achten de Raden onbetaalbaar. Selectie moet ook leiden tot duurzame toegankelijkheid van overheidsinformatie voor iedereen’ [Informatie: grondstof met toekomstwaarde, 2008, aanbiedingsbrief]

De sector overheid/archieven heeft al in het analoge tijdperk mechanismen ontwikkeld om materiaal te selecteren voor archivering; naar schatting belandt momenteel zo’n vijf tot tien procent van de geproduceerde informatie in een archief. ‘Archiveren is selecteren’ is het motto van veel archivariissen. Niettemin behoeven de regels en processen voor selectie aanpassing aan het digitale tijdperk, daarover zijn alle betrokkenen het eens [Gewaardeerd verleden, 2007; Archiveren is vooruitzien, 2009].

Het digitale tijdperk stelt niet alleen nieuwe eisen aan *wat* er geselecteerd wordt, maar ook aan *wanneer* er geselecteerd wordt. In hun visienota concluderen de brancheverenigingen van archivariissen BRAIN en KVAN:

‘Selectie van *digital born* archieven kan niet anders dan aan de bron gebeuren, bij het inrichten van processen en het ontstaan van procesgebonden neerslag. Daar moeten archiefinstellingen en archiefvormers dan ook eendrachtig samenwerken.’ [Archiveren is vooruitzien, 2009, p. 11]

De verschuiving van het tijdstip van selecteren naar het moment van creatie heeft grote gevolgen voor de selectiemechanismes: niet langer heeft men tien of twintig jaar de tijd om te bepalen welke objecten echt belangrijk blijken te zijn.

Voor wetenschappelijke *publicaties* is een effectief kwaliteitssysteem ontwikkeld dat in de praktijk leidt tot selectie: het *peer review* systeem. Onderzoekers die hun resultaten willen publiceren sturen artikelen naar gezaghebbende tijdschriften waar collega-onderzoekers het kaf van het koren scheiden en ervoor zorgen dat alleen de beste artikelen uiteindelijk gepubliceerd worden. Dat was al de praktijk in het analoge tijdperk, en zo gebeurt het nog.

Anders is het gesteld met informatiestromen die geen wortels hebben in de analoge wereld maar hun bestaan danken aan het digitale tijdperk. Daarbij denken we vooral aan de enorme hoeveelheden ruwe en bewerkte onderzoeksdata die in de wetenschap worden geproduceerd. Bij DANS beoordeelt een wetenschappelijke raad in grote lijnen welke soorten onderzoeksdata in aanmerking komen voor bewaring en DANS zelf stelt technische kwaliteitseisen, maar op de wetenschappelijke waarde van de informatie wordt niet geselecteerd. Een DANS medewerker zei daarover: ‘Wie zijn wij om te bepalen wat wetenschappelijk relevant is en wat niet?’ Het DSA bevat dan ook geen selectierichtlijn.

Het gebrek aan selectiecriteria kan ertoe leiden dat de wal het schip keert: bij de Europese Synchrotron Facility loopt de beschikbare capaciteit zo snel vol dat een selectieregime in werking treedt dat puur op de beschikbare capaciteit is gebaseerd: na maximaal zes maanden wordt alle informatie gewist en is het aan individuele onderzoekers om informatie die zij willen bewaren zelf in veiligheid te brengen [Keeping the records of science accessible, 2008, p. 6].

In de sfeer van cultuur/erfgoed pleit De Basis voor gerichte selectie (DB2.1). Binnen digitaliseringsprojecten wordt al een keuze gemaakt uit fysieke objecten die gedigitaliseerd moeten worden, maar daarna moet opnieuw bekeken worden welke bestanden echt met alle zorg omringd moeten worden omdat zij de originelen vervangen, en welke bestanden gebruikskopieën bevatten die, zo nodig, opnieuw gemaakt kunnen worden. In zulke gevallen kan opnieuw digitaliseren goedkoper zijn dan duurzame bewaring.

Het NIBG geeft aan dat men momenteel alle uitzendingen van de publieke omroepen integraal opslaat – maar dat men terug zou kunnen vallen op de selectiemechanismen uit het analoge tijdperk als de noodzaak daartoe zou ontstaan.

In het bedrijvenonderzoek bleek dat het bedrijfsleven wel sterk selecteert, maar dat heeft te maken met de andersoortige motieven in de private sector om informatie te bewaren (zie hoofdstuk 2). [Bedrijvenonderzoek, 2009].

4.3. Duurzame toegang tot digitale objecten in de praktijk

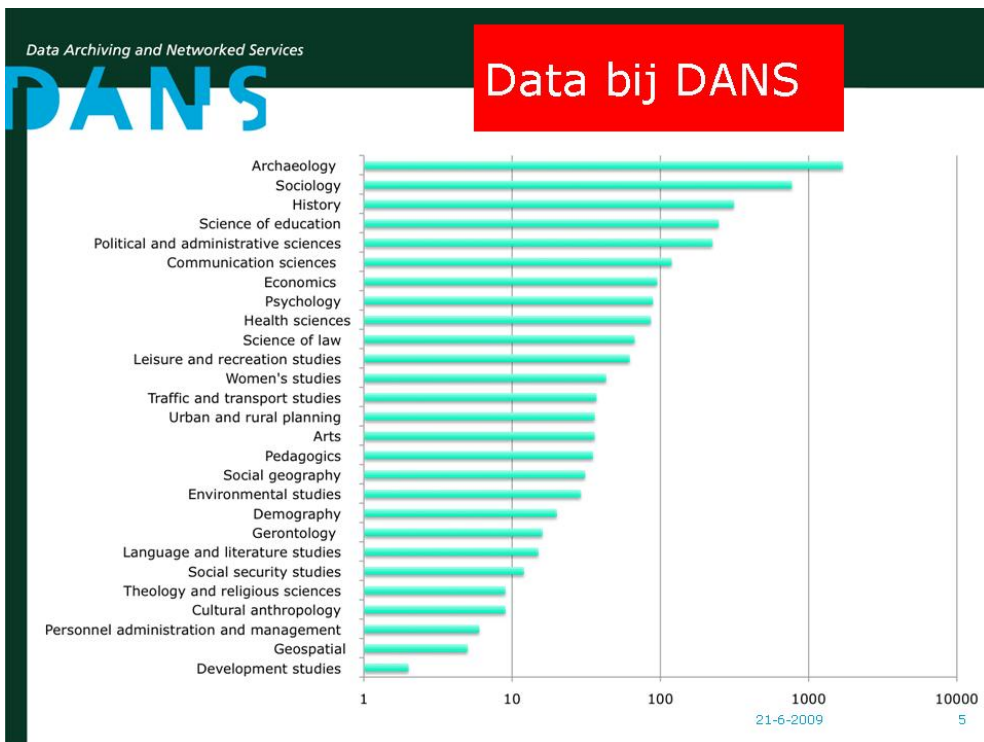
We willen nu onderzoeken hoe men in de praktijk omgaat met digitale objecten die op langere termijn beschikbaar moeten blijven voor raadpleging en (her)gebruik. In theorie zouden we willen beginnen met een kwantitatieve inventarisatie van alle digitale informatie waarin de overheid heeft geïnvesteerd om vervolgens te bezien welke categorieën wel en niet duurzaam worden beheerd. Helaas is een dergelijke inventarisatie in het kader van dit onderzoek een brug te ver gebleken, enerzijds omdat het aantal te onderzoeken instellingen en individuen (1600 overheidsinstellingen, 800 musea, 750 onderzoekinstellingen, 40.000 onderzoekers, 1 miljoen ambtenaren) de mogelijkheden van het onderzoek ver te buiten gingen, anderzijds omdat instellingen en individuen vaak zelf niet weten wat ze in huis hebben aan digitale informatie en welke risico's die informatie loopt. Het is juist tekenend voor de dynamiek van het digitale tijdperk dat ieder kwantitatief overzicht een momentopname is die niets zegt over de nabije of verre toekomst. Ook in geografische zin kan nauwelijks bepaald worden welke informatie zich in Nederland bevindt: in volume is de (beta) wetenschap momenteel de grootste sector en die heeft zich georganiseerd in internationale gedistribueerde informatienetwerken.

We kiezen daarom voor een wat pragmatischer ingang en beginnen dicht bij huis, op een aantal plaatsen waarvan het onderzoek uitwijst dat de uitdaging van toegang tot digitale informatie 'voortvarend wordt opgepakt', zoals de onderzoeksopdracht het beschrijft.

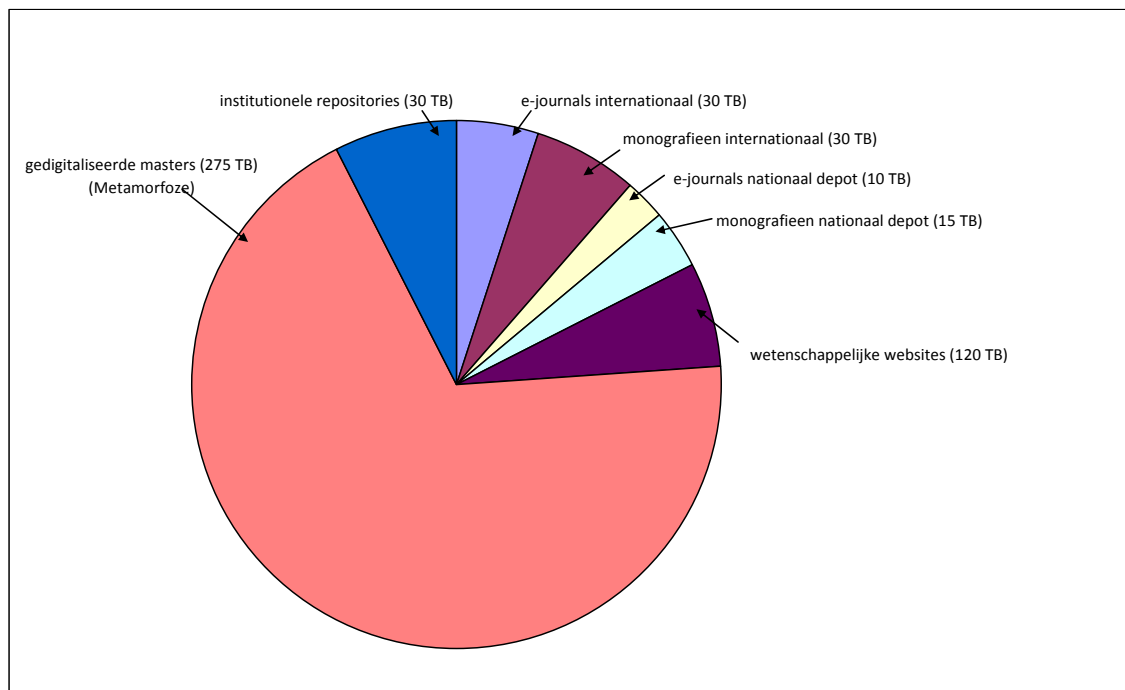
4.3.1. Digitale bewaarplaatsen in Nederland

Een viertal organisaties heeft in Nederland een nationale taak ten aanzien van het bewaren van informatie, analoog of digitaal. Het lag in de lijn van hun missie om ook faciliteiten te ontwikkelen voor de opslag van digitaal materiaal:

- de Koninklijke Bibliotheek (KB) opende in 2003 het e-Depot als digitale pendant van het Depot voor Nederlandse Publicaties;
- het Nederlands Instituut voor Beeld en Geluid (NIBG) ontwikkelde samen met Technicolor (het oude NOB) Het Digitale Archief voor de uitzendingen van de publieke omroep;
- Data Archiving and Networked Services (DANS) werd in 2005 opgericht door de KNAW en NWO als nationaal archief voor onderzoeksdata in de alfa- en gammawetenschappen; DANS is de opvolger van een aantal data-archieven dat al sinds 1964 (Steinmetzarchief) was ontstaan;
- het Nationaal Archief (NA) heeft in november 2009 zijn digitaal depot geopend.



Koninklijke Bibliotheek: het e-Depot tegen 2013



KB, NBIG, DANS en NA voldoen in hoge mate aan de richtlijnen: beleid en organisatie zijn helemaal afgestemd op *duurzame* toegang tot informatie in hun eigen sector. Dit kan ook gezegd worden van een aantal andere instellingen:

- De Justitiële Informatiedienst (JustID) ontwikkelde het CDD+ systeem waarin de dossiers uit de strafrechtssketen digitaal worden opgeslagen en centraal toegankelijk gemaakt voor alle betrokkenen;
- Het Gemeentearchief Rotterdam ontwikkelt simultaan met het Nationaal Archief een digitaal depot;
- Het KNMI heeft het KODAC-systeem ontwikkeld voor de beschikbaarstelling van klimaatgegevens; klimaatgegevens worden daarin duurzaam opgeslagen;
- Het Max-Planck-Institute for Psycholinguistics in Nijmegen bewaart uniek audiomateriaal van bedreigde en uitgestorven talen voor de lange termijn in het DoBeS archief [DoBeS, website];
- Het Instituut Nederlandse Lexicologie (INL) onderhoudt sinds 2004 de Centrale voor Taal- en Spraaktechnologie (TST-Centrale) van de Taalunie [TST, website].

Voor al deze organisaties geldt overigens niet dat zij de problemen voor de eeuwigheid hebben opgelost. In zijn algemeenheid kan gesteld worden dat zij de *bitstreams* momenteel hebben veiliggesteld, maar dat concrete duurzaamheidsstrategieën (zie par. 3.2.3) nog in ontwikkeling zijn, niet in de laatste plaats omdat het bewaarde materiaal nog relatief jong is. In interviews geven diverse organisaties aan dat het moment waarop ze daadwerkelijk zullen moeten gaan migreren/emuleren omdat de bestandsformaten verouderen wel steeds dichterbij komt.

Duurzame digitale bewaarplaatsen komen ook elders tot ontwikkeling:

- Het IISG bouwt een digitaal depot (TDR) voor vele goeddeels vertrouwelijke bestanden over sociale geschiedenis en de arbeidersbeweging; het project is gebaseerd op het OAIS referentiekader en loopt tot 2010;
- Het Stadsarchief Amsterdam werkt aan betrouwbare digitale opslag, maar doet dat in kleine stappen [interview];
- De 3TU.Federatie ontwikkelt een datacentrum voor de opslag van onderzoeksdata uit de technische wetenschappen (momenteel nog een tweejarig project);
- De kring van archivariissen Limburg (KVAL) heeft besloten een project op te zetten voor de vorming van een Limburgs e-depot; dit project is momenteel nog in een zeer pril stadium [Kaderstellende notitie, 2009];
- Het Nederlands Architectuur Instituut (NAi) heeft beleid vastgelegd voor de ontwikkeling van duurzame toegankelijkheid van het materiaal van architecten en architectenbureaus, en heeft de eerste stappen gezet om die te realiseren.

Bij al deze organisaties is er ten minste sprake van gericht beleid om digitale informatie voor de lange termijn beschikbaar te houden. Dit betekent nog niet dat de organisatie ook al volledig gerealiseerd is, maar dat die wel in aanbouw is conform de bestaande richtlijnen.

Naast de aangewezen *duurzame* bewaarplaatsen kunnen we vooral in de wetenschap een breed palet aan opslagmogelijkheden van min of meer tijdelijke aard identificeren – maar deze

voldoen geen van alle aan de richtlijnen die een expliciete missie en organisatie voor de lange termijn dicteren. We noemen hier de *institutional repositories* die door de universiteiten zijn ingericht om digitale objecten op te slaan en de vele faciliteiten bij onderzoeksorganisaties zelf – zoals bijvoorbeeld in 2008 door de KNAW geïnventariseerd in een intern document [Rapport inventarisatie digitale duurzaamheid, 2008]. De *institutional repositories* geven soms wel blijk van een langetermijnintentie (zie o.a. de websites van Igitur Utrecht en de UB van Wageningen UR), maar hebben die alleen geïmplementeerd voor de publicaties, en wel door een archiveringscontract te sluiten met de KB. De risico's van het overige materiaal worden wel beperkt door bijvoorbeeld goede back-upprocedures, maar de organisaties hebben nog geen OAIS-organisatie zoals vereist door de richtlijnen (BL2, DSA08, DB1.2).

Onder de tijdelijke bewaarplaatsen kunnen we in de sfeer van cultuur/erfgoed ook het Nederlands Instituut voor Mediakunst /Montevideo NIMk scharen, dat collecties video- en mediakunst beheert van diverse musea en dat momenteel een bewaartermijn hanteert van tien jaar [interview]. Duurzaamheidsbeleid is daar nog niet tot ontwikkeling gekomen. NIMk meldt in een interview dat *net-based art* nog buiten de scope van het instituut valt, en dat digitale kunst (*web-based art*) van vijf jaar geleden nu soms al niet meer toegankelijk is.

Concluderend mogen we stellen dat alle informatie die in de negen bestaande duurzame bewaarplaatsen tot stand komt of wordt opgenomen in principe wordt bewaard conform de richtlijnen uit hoofdstuk 3 en daarmee een goede kans maakt om duurzaam toegankelijk te blijven – onder voorwaarde dat de organisaties bestendig blijken te zijn en voldoende financiering beschikbaar blijft om steeds opnieuw die maatregelen te treffen die de voortschrijdende techniek vereist (zie hoofdstuk 5).

Maar de capaciteit die momenteel beschikbaar is om duurzame toegang mogelijk te maken tot digitale informatie is verre van toereikend. Er wordt heel veel meer geproduceerd – zoals we in de volgende paragraaf zullen zien.

4.3.2. De productie van digitale gegevens

We verbreden ons onderzoeksveld nu naar de *productie* van digitale gegevens. Wie zijn daarbij betrokken? Hoe is de productie georganiseerd? Wordt het langetermijnbelang bij creatie van digitale informatie meegenomen? En bereikt de geproduceerde informatie de duurzame bewaarplaatsen?

4.3.2.1. De overheid

In 1988 publiceerde de Algemene Rekenkamer een rapport waarin geconcludeerd werd dat de centrale overheid geen greep had op het archiveringsproces – de aanwas, de selectie en de toegankelijkheid van archieven. Dit gold voor de archieven op papier, maar ook voor 'machineleesbare gegevens' [Archiefbeheer en -behoud bij het rijk, 1988]. In 2005 maakte een rapport van de Rijksarchiefinspectie (de huidige Erfgoedinspectie) duidelijk dat er nog weinig was verbeterd:

'Een belangrijk probleem is dat overheidsorganisaties vaak geen goed overzicht hebben van de plaatsen waar zij hun verantwoordings- en digitale informatie beheren. Een groot

deel van deze informatie wordt op persoonlijke en gezamenlijke schijven beheerd, buiten de beheersafspraken om.’ [Een dementerende overheid, 2005, p. 5]

In 2008 vatten de Raad voor Cultuur en de Raad voor het Openbaar Bestuur de situatie aldus samen:

‘Een goede en betrouwbare informatiehuishouding is van vitaal belang voor de overheid. Politiek en ambtelijk management gaan echter tamelijk zorgeloos met dit belang om. De aandacht voor een ordentelijke informatiehuishouding schiet structureel te kort. Dit is een oud probleem dat verregaand versterkt wordt door de voortschrijdende digitalisering. Doorgaan op dezelfde weg betekent volgens de Raden onvermijdelijk de aantasting van belangrijke rechtsstatelijke waarden zoals transparantie, zorgvuldigheid, toegankelijkheid, verantwoording en verantwoordelijkheid. Dat is tamelijk zwart-wit gesteld maar over de ordentelijkheid van de informatiehuishouding is nuanceren nu eenmaal niet mogelijk.’ [Informatie: grondstof met toekomstwaarde, 2008, inleiding]

Inmiddels had de rijksoverheid besloten tot het Programma ‘Informatie op Orde’, dat onder gezamenlijke regie van de ministeries van Binnenlandse Zaken en Onderwijs, Cultuur en Wetenschap was gestart [Informatie op Orde, 2006]. De Raden vonden het programma niet altijd ver genoeg gaan, maar de overheid zette hier wel een agenda neer voor ‘interactie en openbaarheid’, ‘integrale benadering’ en ‘juridische samenhang’. Het programma leidde tot de vaststelling van algemene richtlijnen voor de informatiehuishouding in de Baseline (zie paragraaf 3.4.2). Op het moment dat dit rapport wordt geschreven is het nog te vroeg om te oordelen of het programma leidt tot de gewenste verbeteringen. Eind 2009 zal de Algemene Rekenkamer daarover rapporteren.

Wat weten we van de praktijk bij de provinciale en lokale overheden? In de interviews schetsen de archiefinspecteurs een weinig rooskleurig beeld. Eind 2008/begin 2009 heeft de Denktank e-depot van de Kring van Archivarissen Limburg (KVAL) onderzoek gedaan naar de praktijk bij Limburgse gemeenten. De notitie van de Denktank opent met de volgende zinsnedes:

‘Door de meeste overheidsorganisaties worden digitale archiefbescheiden niet duurzaam bewaard en beheerd terwijl daar wel een wettelijke plicht toe is. ... Het huidige beheersinstrumentarium is bij veel instellingen qua organisatie, beleid, kennis, procedures, beheersmaatregelen, hardware en software verre van toereikend.’ [Kaderstellende notitie Limburgs e-depot, 2009, p. 4].

De denktank trof bij vrijwel alle onderzochte archiefdiensten gedigitaliseerd materiaal aan, soms ook van oorsprong digitaal materiaal, en ‘in enkele gevallen beschikken de gemeenten over digitale archiefbescheiden’. De denktank concludeert vervolgens onomwonden:

‘In geen enkele situatie is sprake van duurzaam digitaal beheer. De bestanden staan op cd’s, dvd’s, USB-harde schijven, servers die zowel intern staan als extern ... een overzicht van wat waar staat kunnen de archiefdiensten in de regel niet geven.’ En: ‘Geen enkele archiefinstelling heeft een beleid met een daaruit voortvloeiend actief beheer van de

aanwezige digitale bestanden of dragers waarop de bestanden staan' [Kaderstellende notitie Limburgs e-depot, 2009, p. 10].

Er is geen enkele reden om aan te nemen dat deze situatie uniek is voor Limburg. Een Brabantse archiefinspecteur die medeopsteller van de notitie was, bevestigde dat het Limburgse beeld landelijke geldigheid heeft. Een Amsterdamse archivaris meldde in een ander verband: 'Het is ontluisterend om te zien dat de meest elementaire antivirussoftware soms niet geïnstalleerd is op de computers van de ambtenaren' [interviews].

4.3.2.2. Cultuur/erfgoed

De productiefase van het digitaal erfgoed dat de instellingen momenteel beheren is redelijk overzichtelijk omdat ze vaak zélf het materiaal gedigitaliseerd hebben. In principe zou dit kunnen leiden tot een goed beheersregime, maar een respondent in het onderzoek vat de situatie als volgt samen: 'Hoezo duurzame toegang, dat is voor ons een brug te ver, we zitten nog volop in de transitiefase van analoog naar digitaal'.

Dit beeld komt terug in veel interviews en ook in publicaties van DEN dat regelmatig onderzoek doet naar de inbedding van ICT in de culturele sector. Erfgoedinstellingen hebben wel vastgelegd beleid voor het beheren van hun fysieke collecties, maar meestal nog niet voor hun digitale bestanden en/of collecties. Het beleid van de overheid is er het laatste decennium ook vooral op gericht geweest erfgoededucatie en publieksbereik te stimuleren.

De subsidieregeling Digitaliseren met Beleid vereist een Informatieplan waarin duurzaamheidsbeleid is opgenomen [Digitaliseren met beleid, 2006] en ook DEN stimuleert het opstellen van informatiebeleidsplannen, waarvan er inmiddels ca. 100 zijn gedeponneerd in het ICT-register [DEN, website]. De duurzaamheidsparagraaf in die plannen blijft echter vaak beperkt tot het regelen van voldoende opslagcapaciteit en het maken van back-ups.

Een aantal instellingen is zich wel degelijk bewust van de risico's die digitale collecties lopen. Uit ons onderzoek noemen we hier het Fries Museum/Prinsessehof, het Nederlands Architectuur Instituut en de Zeeuwse Bibliotheek. Deze en andere instellingen overwegen niet direct om zelf een digitaal archief in te richten, maar zouden hun data willen onderbrengen in een betrouwbaar depot (TDR) elders. Ze kijken daarbij naar de grote nationale instellingen die digitale depots hebben ingericht voor hun eigen digitale collecties en, in het geval van de KB en NIBG, voor door de eigen instelling geleide nationale digitaliseringsprogramma's (*Het Geheugen van Nederland, Metamorfoze, Beelden voor de Toekomst*). Zo heeft de Rijksdienst voor het Cultureel Erfgoed (vroeger: RACM) samenwerking met DANS gezocht om duurzame toegankelijkheid tot archeologische onderzoeksdata te waarborgen.

Ten slotte het internet. De KB is begonnen met het harvesten van een selectie van het Nederlandse web en heeft voor 45.000 websites in de komende drie jaar 21 terabyte capaciteit gereserveerd, maar dit kan niet anders zijn dan het topje van de ijsberg. Een volledig onontgonnen gebied is het hele interactieve web 2.0: Hyves, YouTube, blogs, Picasa, Twitter en al die andere plekken waar particulieren hun levens documenteren. Sommige experts vertrouwen erop dat het internet robuust genoeg zal blijken te zijn om althans een gedeelte van web 2.0 voor de toekomst te bewaren; anderen vrezen dat de geschiedenissen van hele

generaties zullen verdwijnen in een zwart gat. Ligt hier een taak voor de publieke sector? Zelfs die vraag is nog niet beantwoord.

4.3.2.3. De wetenschap

In de wetenschap maken we onderscheid tussen een drietal categorieën digitale gegevens:

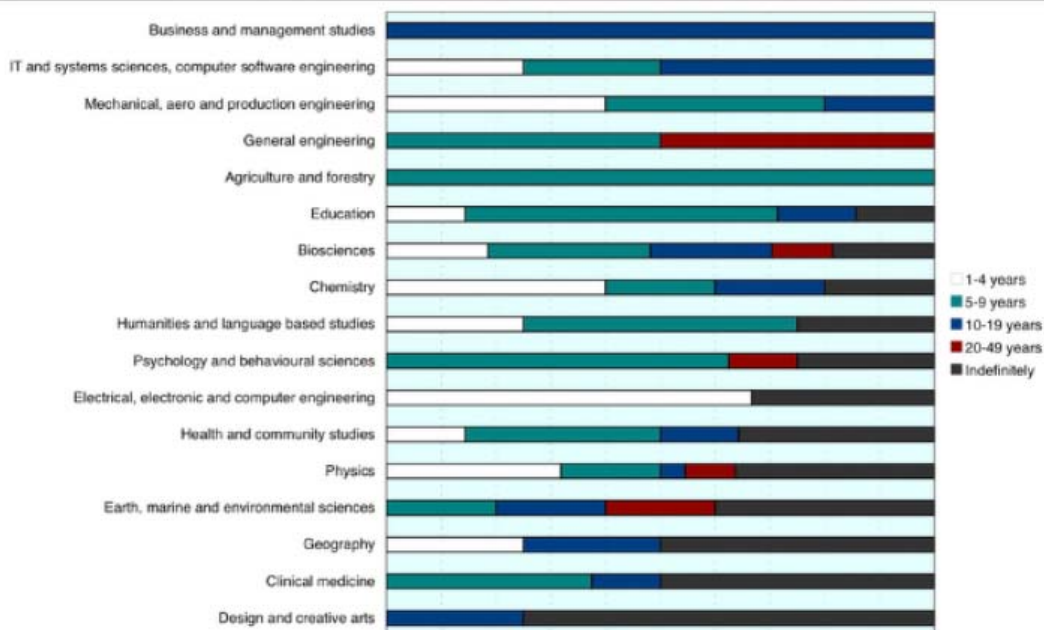
- wetenschappelijke publicaties
- octrooien en patenten
- ruwe en bewerkte onderzoeksgegevens.

Eerst de wetenschappelijke *publicaties*. In grote lijnen zijn hier de praktijken uit het analoge tijdperk in stand gebleven: een wetenschapper doet onderzoek en publiceert de resultaten in een artikel of proefschrift. In het digitale tijdperk worden de publicaties online beschikbaar gesteld op een uitgeversplatform; duurzame bewaring geschiedt bij (nationale) bibliotheken of gespecialiseerde organisaties. De KB heeft een voortrekkersrol gespeeld bij het waarborgen van duurzame toegang tot wetenschappelijke publicaties door de consequentie te trekken uit de internationalisering van de wetenschap en het e-Depot open te stellen voor alle publicaties van grote internationale uitgevers. Ook neemt de KB alle publicaties op uit de (tijdelijke) digitale depots (*institutional repositories*) van de universiteiten. De KB heeft de aanvoer van publicaties door vaste leveranciers zo veel mogelijk geautomatiseerd.

Voor de langetermijntoegang tot *octrooien en patenten* heeft het European Patent Office de 'Esp@cenet' portal opgezet [Espacenet, website]. Dit informatiesysteem maakt 60 miljoen octrooipublicaties uit 80 landen toegankelijk.

Rest de categorie *ruwe en bewerkte onderzoeksdata*. Zoals we in hoofdstuk 2 hebben gezien, gaat het hier om enorme hoeveelheden digitale gegevens van allerlei soorten en maten die ook nog eens over de hele wereld verspreid zijn. Ze worden geproduceerd door (grootschalige) onderzoeksfaciliteiten als radiotelescopen of de befaamde LHC deeltjesversneller, door overheidsorganisaties (statistische gegevens, meteorologische gegevens), door onderzoeksinstellingen, maar ook door (alleen al in Nederland 40.000) individuele onderzoekers.

De enorme diversiteit in de sector wetenschap maakt het lastig om algemene uitspraken te doen over 'de' productiefase van onderzoeksdata. Daarvoor zijn de verschillen tussen de disciplines maar soms ook binnen disciplines te groot. Waar onderzoekers geheel afhankelijk zijn van longitudinale gegevens, zoals de meteorologie en de sociale wetenschappen, zijn al vroeg initiatieven ontwikkeld om de productiefase en duurzame archivering zo goed mogelijk aan elkaar te koppelen. In andere disciplines is dat belang veel minder groot en staat de productie van data veel losser van (enige vorm van) duurzame opslag. In 2008 onderzocht de *UK Research Data Service Feasibility Study* hoe onderzoekers zelf aankijken tegen de gewenste bewaartermijnen van hun data:



Perceived longer term value of current research data holdings by discipline (Q13)

Source: Charles Beagrie survey 2008

[bron: UKRDS, 2008]

In de studie *To share or not to share: publication and quality assurance of research data outputs* [2008] inventariseerde het Engelse Research Information Network bij meer dan honderd onderzoekers in acht verschillende disciplines hun motivatie om hun onderzoeksdata al dan niet beschikbaar te stellen voor duurzame opslag en hergebruik. De studie concludeert:

‘Some researchers are motivated to publish their data by factors such as altruism, encouragement from peers, or hope of opening up opportunities for collaboration. But the lack of explicit career rewards, and in particular the perceived failure of the Research Assessment Exercise (RAE) explicitly to recognise and reward the creating and sharing of datasets – as distinct from the publication of papers – are major disincentives. Many researchers wish to retain exclusive use of the data they have created until they have extracted all the publication value they can. When combined with the perceived lack of career rewards for data creation and sharing, this constitutes a major constraint on the publishing of data. Other disincentives include lack of time and resources; lack of experience and expertise in data management and in matters such as the provision of good metadata; legal and ethical constraints; lack of an appropriate archive service; and

fear of exploitation or inappropriate use of the data' [To share or not to share, 2008, Executive summary, p. 3]

Ten slotte mag niet onvermeld blijven dat er uiteraard ook disciplines zijn waar duurzame opslag minder relevant is omdat het goedkoper is om een experiment of observatie overnieuw te doen dan om informatie langdurig betrouwbaar te bewaren. Zoals een Nederlandse astronoom het zei: 'De kosmos is ons archief' [interview].

4.3.3. Geïdentificeerde belemmeringen en inspiratie voor oplossingen

In deze paragraaf identificeren we de belangrijkste belemmeringen voor duurzame toegang zoals die uit de literatuur bekend zijn of door de respondenten in het onderzoek zijn genoemd. We verbinden die waar mogelijk met initiatieven die inspiratie kunnen bieden voor mogelijke oplossingen (*).

4.3.3.1. Gebrek aan bewustzijn

Over de hele linie geven respondenten aan dat de bewustwording bij beleidsmakers én bij de ambtenaren/onderzoekers op de werkvloer vaak nog te wensen overlaat, waardoor het onderwerp niet de prioriteit krijgt die het verdient.

* Alle in dit rapport genoemde richtlijnen (De Baseline, het Data Seal of Approval en De Basis duurzaamheid) spelen een rol in het bevorderen van het bewustzijn van de risico's.

* Op initiatief van DANS werkt een aantal wetenschappelijke instellingen samen bij de productie van het blad '*e-data en research*', dat het belang van datadeponering en de beschikbaarheid van datasets onder de aandacht brengt van alfa- en gammawetenschappers.

4.3.3.2. Gebrek aan informatie en kennis over duurzame toegang

De informatiebronnen die in de loop van de tijd zijn ontstaan, zoals de website digitaalduurzaamheid.nl, de duurzaamheidspagina van de Taskforce Archiveren, de website van het KB e-Depot, en het geplande kenniscentrum van NBIG komen moeizaam van de grond of worden na een enthousiaste start slecht bijgehouden. Hier is niet zozeer sprake van onwil van de betrokken organisaties als wel van een – in de praktijk – lage prioriteit. Het onderhouden van een kenniscentrum vereist een gerichte inspanning waarvoor meestal geen formatieruimte wordt vrijgemaakt.

Incidenteel worden bijeenkomsten georganiseerd om kennis te delen, maar tot nu toe bekijken ze nog niet, wederom vaak vanwege tijdgebrek. Professionals missen een kader voor de in dit vak broodnodige *éducation permanente*.

* In dit kader is het een goede ontwikkeling dat SURFfoundation in juni 2009 heeft besloten tot de oprichting van een Onderzoeksdata Forum om kennis uit te wisselen tussen bewarende instellingen (o.a. de beheerders van de institutional repositories, DANS, NIOO, etc.) [SURFfoundation, 2009].

* Het Gemeentearchief Rotterdam mag genoemd worden als een partij die steeds actief kennis deelt met collega-instellingen (zie verslagen op de NCDD weblog).

* De NCDD zelf brengt informatie over alle betrokken organisaties bij elkaar op zijn website, maar de schaal is vooralsnog beperkt.

4.3.3.3. Staat van de informatiehuishouding zelf

De respondenten in het onderzoek geven aan dat de gebrekkige staat van de informatiehuishouding van de overheid momenteel het grootste risico vormt voor duurzame toegang tot archiefbescheiden. De Raad voor Cultuur en de Raad voor het openbaar bestuur missen een integrale 'visie op de rol en betekenis van informatie' [Informatie: grondstof met toekomstwaarde, 2008, p. 7]. Ook in de wetenschap komen databestanden die niet geschikt zijn voor duurzame archivering regelmatig voor.

* De overheid heeft 'Informatie op Orde' ontwikkeld om de kwaliteit van de informatiehuishouding te verbeteren, en de Vereniging van Nederlandse Gemeenten heeft aangegeven zich daarbij te willen aansluiten.

* De rijksoverheid heeft recentelijk besloten Chief Information Officers aan te stellen om de informatiehuishouding te verbeteren, als eerste bij de Belastingdienst en het ministerie van BZK.

* In het project 'Waardevolle data en diensten' onderzocht het 3TU.Datacentrum op welke wijze experts van het datacentrum onderzoekers actief kunnen ondersteunen bij het organiseren van hun informatie *tijdens* de onderzoeksfase [Waardevolle data en diensten, 2008]

4.3.3.4. Gebrek aan duurzame financiering

Het onderwerp financiering wordt uitgebreid besproken in hoofdstuk 5.

4.3.3.5. Het (individuele) belang van de producent strookt niet met het algemene belang

Over de hele linie zien we dat de belangen van de producenten van digitale informatie vooral in het particuliere domein en bij de korte termijn liggen: het verlenen van een bouwvergunning aan een burger, het publiceren van het eigen wetenschappelijk onderzoek. Het langetermijnbelang wordt niet acuut ervaren en krijgt geen prioriteit.

* Ook hier kan het project 'Waardevolle data en diensten' tot inspiratie dienen: het datacentrum probeert diensten te leveren die de kwaliteit van het onderzoek voor de onderzoeker verhogen *tijdens* de onderzoeksfase en er bijna ongemerkt voor zorgen dat de data na afloop duurzaam toegankelijk kunnen worden gehouden in het datacentrum. NWO stelt in toenemende mate verplicht dat onderzoekers die door NWO gefinancierd worden hun data na afloop deponeren bij DANS.

* DANS zelf heeft het deponeringssysteem EASY ontwikkeld, dat het wetenschappers zo gemakkelijk mogelijk maakt om data online te deponeren.

* Het *Standard Evaluation Protocol 2003-2009 for Public Research* dat door KNAW, NWO en VSNU is vastgesteld, opent de mogelijkheid om te onderzoeken of andere vormen van wetenschappelijke output dan wetenschappelijke publicaties gebruikt zouden kunnen worden om de kwaliteit van een instelling te evalueren [Standard Evaluation Protocol, p. 10]. Uitwerking van deze mogelijkheid kan bijdragen tot het aantrekkelijker maken van deponering van onderzoeksdata.

* In het programma investeringssubsidies middelgroot stelt NWO een datacontract voor deponering bij DANS verplicht wanneer dataverzamelingen worden opgebouwd.

4.3.3.6. Gebrek aan betrouwbare opslagcapaciteit

Organisaties die niet de schaalgrootte hebben om een eigen digitaal depot in te richten bepleiten de oprichting van gedeelde bewaarplaatsen voor digitale informatie; een respondent sprak van een 'datahotel'. De meningen verschillen over de manier waarop die voorzieningen georganiseerd moeten worden: getrapt geografisch (lokaal, regionaal, provinciaal, landelijk, internationaal), per discipline, per aandachtsgebied/cultuursector, of gespecialiseerd naar bestandsformaat.

* In juni 2009 zijn het Nationaal Archief en de Regionale Historische Centra overeengekomen om binnen vijf jaar een landelijke e-depotvoorziening te ontwikkelen voor algemeen gebruik binnen de archiefsector [e-Archief in ontwikkeling, 2009].

* De gemeentelijke archivariissen in Limburg overwegen momenteel een regionaal e-depot in te richten [Kaderstellende notitie Limburgs e-depot, 2009].

* In het onderzoek gaven diverse respondenten aan dat er tevens gekeken zou moeten worden naar een regionale intersectorale bundeling. Daar liggen immers de beste kansen om ook aan de kant van *toegang* faciliteiten te delen. Bovendien zouden de cultuurverschillen kleiner zijn dan tussen de landelijke en de lokale of regionale organisaties. Met name genoemd werden een mogelijk Zeeuws e-depot en de Friese samenwerking tussen Tresoar, Fries Museum/Prinsessehof, het Fries Museum en de Fryske Akademy, die via het glasvezelnet back-upfaciliteiten voor elkaar regelen.

* JustID ontwikkelde in het CDD+ systeem een manier om dossiers uit vele verschillende organisaties gezamenlijk op te slaan zonder dat de betrokken instellingen de regie over hun dossiers verliezen.

* Diverse meer technische projecten onderzoeken mogelijkheden om een diversiteit aan bestanden gedistribueerd op te slaan in een zogenaamde GRID, bijvoorbeeld iRODS, TIPR, etc. [websites].

4.3.3.7. Gebrek aan menskracht, IT-capaciteit

Zelfs als een aangewezen bewaarplaats bestaat, ontbreekt het bij organisaties aan de benodigde menskracht en IT-capaciteit om de digitale collecties zodanig te bewerken dat ze in het digitaal depot kunnen worden opgenomen; dit meldt bijvoorbeeld het CBS ten aanzien van

gedigitaliseerde publicaties en ten aanzien van de database achter StatLine waarvoor een wettelijke bewaarplicht bestaat waaraan momenteel nog niet wordt voldaan.

4.3.3.8. Gebrek aan concrete, praktische gereedschappen

Voor het duurzaam toegankelijk houden van digitale gegevens zijn technische gereedschappen nodig zoals migratietools en emulatietools. Een respondent in het onderzoek sprak in dit verband van een 'digitaal instrumentenbord' en een andere benadrukte dat de gereedschappen vooral laagdrempelig en praktisch moeten zijn. 'We need survival practices rather than best practices', zei Clifford Lynch hierover tijdens een recente conferentie [eigen aantekening DigCCurr 2009].

- * Technisch onderzoek naar concrete gereedschappen vindt vooral in Europees verband plaats, in grote projecten als PLANETS (bibliotheken en archieven), CASPAR Preserves (wetenschappelijke datacollecties) en PrestoPRIME (audiovisueel materiaal) [websites].

4.3.3.9. Gebrek aan diensten

Een aantal bewarende organisaties zou niet alle benodigde functies en expertise altijd zelf in huis willen hebben, maar waar nodig een beroep kunnen doen op een nationaal of internationaal dienstenaanbod. Men kijkt daarbij al gauw naar de vier nationale instellingen (DANS, KB, NA en NIBG), maar die kunnen niet altijd voldoen aan de vraag. Een voorbeeld: JustID is momenteel actief op zoek naar een organisatie die op verzoek databestanden kan migreren.

- * Het NA experimenteert in 2009 met diverse vormen van dienstverlening aan kleinere instellingen. Het NA onderscheidt daarbij een aantal scenario's: vervroegd overbrengen (pilot met het ministerie van BZK); trusted repository, ook voor te vernietigen materiaal (idem); trusted repository voor RHC's en andere archiefdiensten (pilot met Utrechts Archief); en een service oriented architecture (pilot met het Kadaster) [Slats, 2008]. Commerciële bedrijven bieden momenteel diverse conversiediensten aan, maar het is de vraag of die altijd kunnen worden toegepast in de publieke sector, zeker waar de *look and feel* van een document bewaard moet blijven, een eis die in de private sector niet wordt gesteld [Bedrijvenonderzoek, 2009]

- * NIBG heeft de dienst ProArchive opgezet: kleinere instellingen kunnen hun audiovisueel materiaal tegen betaling onderbrengen bij NIBG en een palet aan diensten afnemen. Momenteel loopt een pilot met het Zeeuws Archief. ProArchive meldt dat er wel offertes worden aangevraagd, maar dat er nog weinig concrete contracten worden afgesloten voor het beheer van audiovisuele data. In het onderzoek bleek dat cultuurhistorische instellingen schrikken van de prijzen voor duurzaam beheer van digitale AV-collecties; misschien is ook een meer gedifferentieerd aanbod nodig (bijv. wel opslag maar geen metadatering).

4.3.3.10. Gebrek aan organisatie die is aangepast aan het digitale tijdperk

Het beschikbaar komen van onderzoeksdata naast de wetenschappelijke publicatie is een relatief nieuw fenomeen van internationale en grootschalige proporties waarvoor nog geen

adequate organisatie is ontwikkeld. Ook de combinatie van een traditionele publicatie met onderzoeksdata, de zogenaamde *verrijkte publicatie*, levert organisatorische problemen op omdat de publicatie bijvoorbeeld bij de KB wordt gearhiveerd en de data in onderzoekscentra als DANS en 3TU, of in het buitenland.

- * Aangezien het hier een internationaal fenomeen betreft, zal de inspiratie vooral in Europees verband gevonden worden:

In 2006 publiceerde het European Strategy Forum on Research Infrastructures (ESFRI) de eerste *European Roadmap for Research Infrastructures* die in 2008 een Nederlandse vertaling kreeg [Nationale roadmap grootschalige onderzoeksfaciliteiten, 2008]. Binnen de ESFRI roadmap zijn een aantal initiatieven ontstaan die expliciet langetermijntoegang tot onderzoeksdata nastreven. We noemen hier DARIAH voor de geestes- en maatschappijwetenschappen, CESSDA voor de sociale wetenschappen, CLARIN voor taalkundige bronnen en BBMRI voor biologische databronnen voor gezondheidsonderzoek [zie de respectievelijke websites].

- * Een roadmap voor *duurzame toegang* tot wetenschappelijke bronnen wordt momenteel ontwikkeld in het PARSE.insight project [PARSE.insight, 2009].

- * In Nederland noemen we het SURFshare programma, waarin DANS, de KB, 3TU en andere organisaties samenwerken aan de ontwikkeling van een infrastructuur voor verrijkte publicaties [SURFshare, website].

- * De sector overheid hinkt momenteel nog op twee gedachtes: enerzijds is er de theorievorming van het *records continuüm*, anderzijds hakt de Archiefwet dit continuüm nog steeds in de fase *vóór* en de fase *na* overbrenging. Voor goed beheer van digitale informatie is zo'n scherpe grens eigenlijk onwenselijk. De brancheorganisatie van archieven BRAIN, de organisatie van archivariissen KVAN, het Nationaal Archief en de Regionale Historische Centra pleiten in hun visiedocumenten van 2009 voor geleidelijke aanpassing van de Archiefwet. Strikt genomen, echter, is een aanpassing van het wettelijk regime niet noodzakelijk om in praktische zin het beheer van archieven in de dynamische fase te ondersteunen: het Gemeentearchief Rotterdam ontwikkelt momenteel één systeem waarin zowel de concernadministratie als het archief is ondergebracht, met onderscheiden juridische regimes. Ook het NA experimenteert met verschillende scenario's (zie paragraaf 4.3.3.9).

- * In lijn met het advies van De Raden voor Cultuur en voor het openbaar bestuur, pleiten BRAIN en KVAN in hun visienota 'Archiveren is vooruitzien' voor het onder één bewindspersoon brengen van de hele informatiehuishouding van de overheid [Informatie: grondstof met toekomstwaarde, 2008; Archiveren is vooruitzien, 2009, p. 9].

- * Een ander voorbeeld is het CDD+ systeem van de Justitiële Informatiedienst dat informatie uit de hele strafrechtsketen bij elkaar brengt en de digitale dossiers gezamenlijk beheert terwijl de producenten van de informatie ieder hun eigen regie houden over hun eigen aandeel in de dossiers.

'De risico's die digitale data lopen kun je altijd indammen als je er maar genoeg geld in steekt - geld voor nog meer back-ups, voor nog meer personeel, voor nog meer onderzoek, en ga zo maar door.' (Richard Wright, BBC, iPRES 2008)

'Eigenlijk zouden we voor elk archiefstuk een soort duurzaamheidstoeslag moeten vragen, naar analogie van de verwijderingsbijdrage.' (een archivaris)

5. Kosten en financiering

Dit hoofdstuk inventariseert wat we weten over de kosten van duurzame toegankelijkheid en over de financiering.

5.1. Inleiding

Toen Jeff Rothenberg zijn 'Longevity' artikel schreef [Rothenberg, 1995], zag hij vooral technische obstakels voor duurzame toegang tot digitale informatie. Dertien jaar later zei Richard Wright van de BBC tijdens de iPRES conferentie in Londen: 'De risico's die digitale informatie lopen kun je altijd indammen als je er maar genoeg geld in steekt, geld voor nog meer back-ups, voor nog meer personeel, voor nog meer onderzoek, en ga zo maar door.' [Wright, 2008] Dit citaat geeft aan dat de vraag wat duurzame toegang kost niet eenduidig beantwoord kan worden, zeker niet over de breedte van een zeer diverse nationale infrastructuur. Er zijn vele variabelen in het spel die per sector, per organisatie en per collectie sterk kunnen verschillen. Hoe goed is het materiaal bij opname in het archief? Hoe zeker wil men zijn dat er echt nooit iets fout zal gaan? Hoe snel moet het materiaal beschikbaar zijn? Gaat het om grote gelijkvormige collecties of losse objecten uit diverse bronnen? Wat voor zoekmogelijkheden wil men de gebruiker bieden? En ook: welke kosten rekent men toe aan 'duurzame toegang'? Een Rotterdamse archivaris zei daarover: 'Ontsluiten moeten we toch doen, dus dat rekenen we niet toe aan het digitaal depot,' terwijl bijvoorbeeld de KB die kosten wel meeneemt in de 'total cost of ownership' van digitaal materiaal.

De kosten voor research & development vormen een andere variabele die moeilijk door te rekenen is. Momenteel staat de techniek in de kinderschoenen en vormt onderzoek een belangrijke kostenpost. Het is ook de algemene verwachting dat R&D nodig zal blijven om technische ontwikkelingen te volgen, maar de intensiteit van de inspanning is niet te voorspellen.

5.2. Onderzoek naar de kosten van duurzame toegang

Met name in het Angelsaksische taalgebied is veel onderzoek gedaan naar de kosten van duurzame toegang. Eind 2008 is de huidige kennis op dit gebied samengevat in het interim-rapport van de Amerikaans/Engelse Blue Ribbon Task Force on Sustainable Digital Preservation and Access [Sustaining the digital investment, 2008, vooral hoofdstuk 3 en tabel 3.1]. Dit rapport bevat de resultaten van het meest gedetailleerde onderzoek naar ervaringscijfers tot nu toe, de studie *Keeping research data safe* uit 2008 [Beagrie, Chruszcz & Lavoie, 2008] dat op zijn beurt weer verwijst naar Nederlands onderzoek door het Nationaal Archief [Testbed digitale bewaring, 2005].

We zullen alle cijfers die het rapport van de Blue Ribbon Task Force bevat hier niet herhalen, want uiteindelijk geven ze weinig richting. De auteurs concluderen:

'Even for those projects that explicitly build on earlier work, it is clear that within any given project the costs captured are generally focused upon only a small subset of activities within the digital preservation lifecycle (for example, storage costs). In short,

the structure for previous studies rarely supports direct comparisons.’ [Sustaining the digital investment, p. 37]

Als ervaringscijfers niet genoeg informatie opleveren, kunnen we dan wellicht modellen ontwikkelen die ons in staat stellen om ramingen voor de toekomst te geven? Die opdracht werd ter hand genomen door het LIFE (Lifecycle Information for e-Literature)-project in Engeland, dat in 2008 zijn tweede rapport opleverde [LIFE, website; Ayris et al., 2008]. Net als het Testbed digitale bewaring eerder deed, leverde het rapport waardevolle informatie op ten aanzien van het soort kosten dat bij duurzame toegang komt kijken, en in die zin zal de informatie van belang zijn voor individuele organisaties die overwegen duurzame archivering ter hand te nemen. Maar er is ook kritiek op het LIFE2 model: juist de kosten voor de infrastructuur worden er niet in meegenomen.

In het eerste half jaar van 2009 liet DANS een econoom in het kader van een mastersopleiding onderzoek doen naar de kosten van duurzame toegang bij DANS. Zij paste technieken als het *activity-based costing model* en de *balanced score card* toe op de werkprocessen bij DANS en inventariseerde hoeveel tijd en geld DANS besteedt aan diverse producten en processen. Ook zij concludeerde dat die kosten alleen afgemeten kunnen worden aan de doelstellingen die een organisatie zich stelt: wat zijn de producten die de organisatie wil leveren? Wat moet de kwaliteit van die producten zijn? Pas als die goed benoemd zijn kunnen organisaties hun producten en processen SMART benoemen en ervaringscijfers verzamelen die het mogelijk maken om *op den duur* steeds betere schattingen te maken die zullen leiden tot betere begrotingen. (De resultaten van het onderzoek zijn bij het ter perse gaan van dit rapport nog niet beschikbaar) [interview].

5.3. Factoren die de kosten beïnvloeden

Hoewel exacte cijfers dus niet voorhanden zijn, kunnen uit de literatuur wel een aantal factoren gedestilleerd worden die van grote invloed zijn op de kosten:

5.3.1. Schaalgrootte

Alle genoemde rapporten zijn het erover eens dat schaalgrootte een belangrijke factor is [o.a. Testbed digitale archivering, 2005, p. 18; Beagrie et al., 2008, p. 5-6]. Naarmate de hoeveelheid gearchiveerd materiaal toeneemt, dalen de kosten per eenheid behoorlijk. Deze winst gaat geheel of gedeeltelijk verloren als het aantal te verwerken bestandsformaten toeneemt of de complexiteit ervan. Ook het aantal toeleveranciers waarmee een digitaal archief contacten moet onderhouden verhoogt de kosten per eenheid – net als de al dan niet aanwezige motivatie/kennis van wetenschappers om mee te werken.

5.3.2. Mettertijd vlakken de kosten af

Het opstarten van een duurzame digitale bewaarplaats is een kostbare zaak. Zo berekende het Gemeentearchief Rotterdam dat het inrichten van een serverruimte met een minimale beheersorganisatie M€1,6 kost. Maar ook het inrichten van de beheersorganisatie vereist veel investeringen. Naarmate die organisatie volwassener wordt en verwerkingsprocessen efficiënter worden, vlakken de kosten af [o.a. Beagrie et al., op basis van de cijfers van de Archaeological Data Service, 2008, p. 6].

5.3.3. Personeel is de grootste kostenfactor

Er wordt nogal eens geschermd met het feit dat opslag steeds goedkoper wordt. Dat feit op zich willen we hier niet bestrijden, maar we hebben ook steeds meer opslagcapaciteit nodig. Bovendien blijken de kale opslagkosten maar deel te zijn van de totale kosten van duurzame toegang. In een interview geeft Technicolor aan dat het een kaal opslagtariaf voor relatief langzame bewaring op tapes hanteert van €80 per TB per jaar; Jim Gray van Microsoft noemde al in 2003 een bedrag van K\$300 voor het totale beheer, inclusief ontsluiting, voor een TB aan informatie [Informatie: Grondstof met toekomstwaarde, 2008, p. 15]. Wat die getallen waard zijn kan hier moeilijk worden ingeschat. Wel zijn alle bronnen het erover eens dat de kosten vooral in het beheer, en dan vooral in personeel gaan zitten. Beagrie et al. noemen een getal van 70% [Beagrie et al., 2008, p. 28]. De KB geeft de volgende verdeling: personeel voor het e-Depot productieproces: 30%; projectmedewerkers en research & development: 25%; onderhoud, hardware en software licenties: 25%; opslag: 20% [Ras, 2009].

5.3.4. Timing

Een goede aanpak *bij creatie* van de digitale informatie, met name ten aanzien van metadatering, levert later in het proces grote besparingen op, zo concludeerde het Nationaal Archief al in 2005 [Testbed digitale bewaring, 2005, p. 15]. Achteraf reconstrueren is duur, en wordt steeds duurder naarmate er meer tijd verstreken is.

5.3.5. De duurste fase lijkt acquisitie en opname (ingest)

Gezien het feit dat er nog zo weinig langetermijnervaring is met duurzame archivering, is enige voorzichtigheid geboden, maar de UK Data Archive schat dat 42% van de kosten worden gemaakt voor acquisitie en opname in het systeem (ingest), 23% voor opslag en duurzaamheid, en 35% voor toegangsregimes [Beagrie et al., 2008, p. 5].

5.3.6. Duurzaamheidsstrategieën: emulatie versus migratie

Over het algemeen staat emulatie (zie hoofdstuk 3) te boek als de duurste duurzaamheidsstrategie voor digitale informatie, omdat het ontwikkelen van emulatoren kostbaar is. Toch heeft de KB in 2005 berekend dat emulatie op de *lange* termijn goedkoper zou kunnen zijn [Oltmans en Kol, 2005]. Het probleem is uiteraard wel dat voor emulatie op de korte termijn grote investeringen moeten worden gedaan; de kosten voor opeenvolgende migratierondes zullen zich geleidelijker doen voelen.


5.4. Ervaringscijfers in de Nederlandse context

Er zijn nog niet veel ervaringscijfers bekend uit de Nederlandse context; respondenten die in diverse ontwikkelingsstadia verkeren van hun digitaal depot gaven aan dat ze met hun schattingen nog niet naar buiten willen treden.

5.4.1. De Koninklijke Bibliotheek

In 2007 schatte de KB dat het e-Depot, inclusief research & development, jaarlijks M€3,655 kostte. Die kosten moesten worden omgeslagen over ca. 10 miljoen digitale objecten, voornamelijk PDF's van uitgevers [Addressing the future, 2007, p. 45]. Twee jaar later heeft de KB besloten tot een forse capaciteitsuitbreiding om met name de masters van digitaliseringsprojecten en het Nederlands Webarchief duurzaam te kunnen opslaan. De

kosten gaan omhoog naar M€6, de capaciteit van 15 terabyte naar meer dan 300 terabyte (inmiddels houdt men al rekening met 650 terabyte).



What does it cost?

costs		2010	2011	2012	2013
staffing	operations (8 fte)	450	500	500	500
	research (6 fte)	1500	1500	1500	1500
	maintenance (2 fte)	200	250	250	250
	kC	2.150	2.250	2.250	2.250
materiel	Hosting & backup	1500	1500	1500	1500
	LTP-system licence	1000	1000	1000	1000
	Application development	1000	1000	1000	1000
	Preservation actions	250	250	250	250
	kC	3.750	3.750	3.750	3.750
Total amount kC		5.900	6.000	6.000	6.000

[bron: Ras, 2009; kosten gebaseerd op 300 TB. De materiële kosten in dit overzicht blijven constant omdat de KB meerjarige contracten afsluit met leveranciers om kosten gelijkmatiger te kunnen budgetteren.]

5.4.2. NIBG

ProArchive van NIBG berekent voor de opslag kosten door *per uur materiaal, per jaar*. Hierbij wordt onderscheid gemaakt tussen opslag in het formaat MXF D10-30 en MXF D10-50. Managed storage (opslag in diverse vormen, servers, 24x7 beheer, archiefmanagement, security, browse files en keyframes) in MXF D10-30 kost per uur, per jaar €22,74, MXF D10-50 kost €27,22. Daarnaast wordt de kosten opslag op SALT-tapes voor digitale opslag en back-up doorberekend. Deze tapes kosten per twee (40 uur MXF-10-30 en 27 uur MXFD10-50) €110. Importeren van materiaal en andere diensten worden ook eenmalig per uur doorberekend.

5.4.3. NIMk/Montevideo

Binnen het nog niet afgeronde project *Play-out* heeft NIMk/Montevideo de kosten berekend voor interne opslag op LTO-tapes, inclusief arbeidskosten, materiële kosten, en de niet onaanzienlijke elektriciteitskosten. NIMk komt op €350 per terabyte per jaar [Play-out, 2009].

5.4.4. Pictura

Momenteel hanteert het bedrijf Pictura een basistarief voor opslag van €800 per terabyte per jaar. Nadere specificaties ontbreken.

5.4.5. Rekenmodel

Bij de presentatie van het rapport *De digitale feiten* vertelde Peter Wouters, hoofd Publieksbereik van het Gelders Archief, dat het archief een rekenmodel in Excel had ontwikkeld om de kosten van digitaliseringprojecten in beeld te krijgen. Dit kostenmodel is vervolgens ook gebruikt bij de Gelderland Bibliotheek. Op basis van dit model berekende het archief dat vijf jaar opslag van gedigitaliseerd materiaal even veel kost als de initiële investering voor digitalisering. En daarbij was nog geen sprake van een echt 'trustworthy'

digitaal depot, maar van een gemengd pakket waarbij alleen waardevolle bronnen echt duurzaam werden opgeslagen.

5.5. Financiering

5.5.1. De overheid als voornaamste financier

In Nederland komt de financiering van duurzame toegang tot publieke informatie voor het overgrote deel van de overheid zelf – landelijk, provinciaal of lokaal. Het rapport van de Blue Ribbon Task Force beveelt aan om vooral ook private bronnen te mobiliseren voor de financiering, maar deze aanbeveling lijkt wel zeer bepaald door de Amerikaanse context waarin publieke zaken veel vaker door private fondsen worden gefinancierd [Sustaining the digital investment, 2008].

Waar financiering van de archieven door de overheid voor de hand ligt, is in de wetenschap niet duidelijk welke partij de kosten op zich moet nemen. Onderzoekers verwachten dat ze gratis toegang hebben tot wetenschappelijke data en dat zij niet hoeven op te draaien voor de kosten van opslag en beheer.

De financiering van de ICT infrastructuur komt voornamelijk van specifieke programma's (NWO, BIG GRID) en stimuleringsprogramma's als het Besluit Subsidies Investeren Kennisinfrastructuur (BSIK). De organisatie van de ICT-infrastructuur is versnipperd tussen organisaties als NWO, de Stichting Nationale Computer Faciliteiten (NCF), SURF en SARA Reken- en Netwerkdiensten, alsmede BSIK en NWO-programma's. De economische levensduur van een ICT systeem is slechts vijf tot zeven jaar, maar de ICT-infrastructuur voor wetenschappelijk onderzoek bevat elementen die langer relevant zijn dan de levensduur van het project. Ook is de verantwoordelijkheid gedistribueerd over verschillende partijen. ICTRegie adviseert daarom om alle ontwikkel- en uitvoeringswerk voor de ICT-infrastructuur voor het wetenschappelijk onderzoek onder te brengen onder de paraplu van SURF. Tevens pleit ICTRegie voor het oprichten van een e-science research center dat multidisciplinair onderzoek verricht [Towards a competitive ICT-infrastructure, 2008].

De KB, inclusief het e-Depot, wordt gefinancierd door het ministerie van OCW, Directie Onderzoek- en Wetenschapsbeleid.

Kunst en cultuur kennen een lange traditie van gedecentraliseerde overheidszorg [Cultuurbeleid in Nederland, p. 40]. Voor een deel gebeurt overheidsfinanciering direct en voor een deel door middel van fondsen als de Mondriaan Stichting en subsidieregelingen.

Incidenteel leveren ook in Nederland private fondsen een bijdrage aan de financiering van publiek goed. In de erfgoedsector, bijvoorbeeld, hebben de Prins Bernhard Cultuurfonds en het VSB-fonds bijgedragen aan digitaliseringsprojecten. Maar de bijdragen zijn niet structureel en komen snel onder druk te staan als het economisch minder gaat.

5.5.2. Projectmatige financiering: een structureel probleem

De diverse richtlijnen (hoofdstuk 3) benadrukken dat het langetermijnbelang overal moet worden ingebed: in het beleid, in de organisatie, en dus ook in de financiering. Hoe verhoudt de praktijk zich tot die richtlijn?

- Wetenschappelijke onderzoeksprojecten worden projectmatig gefinancierd. Na afloop van het project ontbreken vaak de financiële middelen om de geproduceerde data te bewaren.
- In de culturele sector worden grote bedragen besteed aan het digitaliseren van fysiek erfgoed: kranten, geluidsdragers, films, boeken. Maar voor het langdurig bewaren van de digitale bestanden wordt geen geld gereserveerd. Zo wordt in het kader van het grootschalige project 'Beelden voor de Toekomst' voor M€154 aan digitale bestanden gecreëerd. De opslag daarvan is meegefinancierd tot aan het eind van het project, 2014. Voor de periode daarna zijn geen maatregelen getroffen.
- Ook R&D-inspanningen ten aanzien van dataopslag worden overwegend tijdelijk gefinancierd. Na afloop vallen de projectorganisaties uit elkaar en is borging van de resultaten onzeker.

Het gebrek aan structurele financiering is een ernstige belemmering voor het toegankelijk houden van ons digitaal geheugen. De organisaties die langetermijnarchivering op zich nemen zullen op de één of andere manier structureel gefinancierd moeten worden. Uit het analoge tijdperk zijn met name bibliotheken en musea bekend als voorbeelden van bewaarplaatsen die duurzaam worden gefinancierd. Dergelijke organisaties moeten ook voor ons digitale erfgoed worden ingericht – bijvoorbeeld door een deel van de gelden die beschikbaar komen voor onderzoek en digitaliseringsprojecten te reserveren voor langetermijnbewaring.

5.6. De kosten-batenanalyse

De vraag of duurzame toegang 'duur' is, kan op allerlei manieren beantwoord worden. Het is waar dat de samenleving nu in een fase verkeert waarin forse investeringen nodig zijn om oplossingen te vinden voor nieuw ontstane technische problemen. Maar fysieke opslag en beschikbaarstelling is ook duur. Al in 2006 berekende Adrian Wilson van de National Archives of Australia dat het beschikbaarstellen van een papieren document ongeveer GBP6 kostte; beschikbaarstellen via internet kostte 13 pence [Rusbridge, 2006].

Misschien moeten we de vraag herformuleren: welke eisen stelt de samenleving aan het bewaarde collectief geheugen? Of: wat heeft de samenleving ervoor over om (een deel van) het collectief geheugen voor de toekomst veilig te stellen?

De berekeningen zouden zeer gebaat zijn bij een inschatting van de baten die de bewaarde informatie de samenleving in de toekomst gaat brengen, maar die inschatting is niet te maken. Incidenteel kunnen wij misschien laten zien dat twee uur achter de computer vijf dagen laboratoriumwerk kan besparen, maar dat zijn de uitzonderingen. Wie kwantificeert de baten van een medicijn dat wellicht vijf maanden eerder beschikbaar komt? Wie hangt een prijskaartje aan de laatste geluidsbanden van uitgestorven talen?

Tijdens de DigCCurr conferentie 2009 zei Clifford Lynch van de Coalition for Networked Information dat het een van zijn grootste zorgen is dat we nog niet in staat zijn gebleken om de zaak voor duurzame toegang tot informatie echt hard te maken, om te laten zien wat de consequenties zullen zijn van zorgeloze omgang met ons digitaal geheugen [Lynch, 2009]. De Amerikaanse econoom Paul Courant vergeleek de zorg voor digitale informatie met de zorg voor onze natuur: ook daar is sprake van een publiek langetermijnbelang dat niet in geld is uit te drukken [Courant, 2008].

'Je moet het nog wel uitleggen. Maar als de partners eenmaal zien wat een digitale informatiestroom hen kan opleveren, dan zijn ze om, dan willen ze fors investeren. Want ze betalen gewoon te veel aan de papieren informatiehuishouding' (programmaleider Justitiële Informatiedienst)

6. (Intersectorale) samenwerking

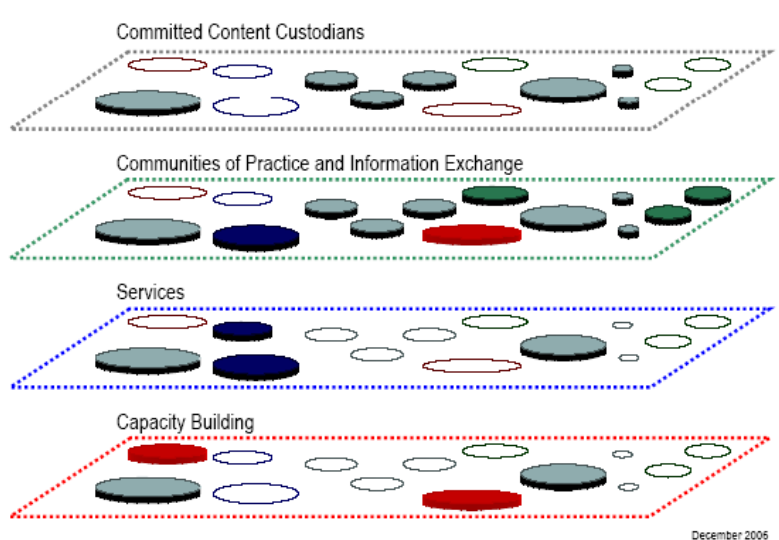
Voordat we de mogelijkheden voor intersectorale samenwerking in Nederland onder de loep nemen, kijken we naar een aantal ervaringen in het buitenland.

6.1. Intersectorale samenwerkingsverbanden elders

De Engelse collega van de NCDD, de Digital Preservation Coalition, werd in 2003 opgericht, net als de NCDD door een aantal belanghebbende organisaties, maar zonder de infrastructurele missie die de NCDD heeft opgesteld. De DPC publiceerde in 2006 als eerste een Nationale Verkenning onder de titel *Mind the Gap* waarin vooral het probleem duidelijk op de kaart werd gezet [Mind the Gap, 2006]. Gevraagd naar de ontwikkelingen na *Mind the Gap* vertelde een bestuurslid van de DPC dat er minder vooruitgang is geboekt dan gehoopt. De cultuurverschillen tussen de sectoren bleken hardnekkig, en samenwerking wordt ook niet bevorderd door de Engelse financieringsstructuren die veel gecompliceerder zijn dan in Nederland. Niettemin groeit het ledenaantal van de DPC gestaag en is de DPC met name als kenniscentrum alom tegenwoordig met o.a. het *Handbook Digital Preservation* en het kwartaaloverzicht 'What's New in Digital Preservation' [Digital Preservation Coalition, 2008]. De leden van de DPC vragen nu vooral om *praktische* gereedschappen en kennis. Ook vragen de leden in toenemende mate om informatie en diensten exclusief voor leden [interviews].

In de Verenigde Staten lanceerde het Congres in 2000 het 'National Digital Information Infrastructure & Preservation Program' (NDIIPP), met een budget van M\$100 (waarvan uiteindelijk maar M\$40 werd uitgegeven) en de opdracht om een nationale strategie voor digitale duurzaamheid te formuleren. Anders dan in Nederland of Engeland kreeg de Library of Congress duidelijk de leiderschapsrol toebedeeld, en ook anders dan in Nederland bevordert NDIIPP actief publiekprivate samenwerkingsverbanden. Martha Anderson, de projectleider geeft grif toe dat het beschikbare geld in het begin een sterke motivatie was om te participeren in NDIIPP, maar de noodzaak om 50% te matchen was voor veel instellingen een belemmering. NDIIPP heeft nu een structureel budget van M\$7 en een bureau van vier medewerkers. NDIIPP heeft 150 partnerorganisaties overal in het land en steunt projecten die op de één of andere manier samenwerking bevorderen. Martha Anderson vertelde dat het formuleren van een nationale strategie lange tijd nog een brug te ver was. Momenteel werkt het NDIIPP bureau aan een allereerste versie van een dergelijk stuk [NDIIPP, website].

Ter voorbereiding op die notitie propageert NDIIPP een 'network of networks', waarin organisaties rollen en functies op zich nemen binnen een landelijk dekkend netwerk (figuur 6.1).



Figuur 6.1. Het National Content Stewardship Network zoals voorgesteld door NDIIPP [bron: Anderson, 2008]

NDIIPP maakt het liefst gebruik van natuurlijke netwerken die zijn ontstaan rond clusters *content* en probeert vervolgens de samenwerking als een olievlek uit te breiden. ‘Learn by doing’ is het motto [Potter, 2009].

In Duitsland werd door een aantal publieke organisaties in 2003 het project ‘nestor’ opgericht – een ‘Kompetenznetzwerk Langzeitarchivierung’. Zoals de naam al zegt is kennisuitwisseling en gezamenlijke opbouw van expertise hier het enige doel. Nestor werkt met zeer actieve ‘Arbeitsgruppen’, waarvoor individuele leden de verantwoording nemen [nestor, website]. In 2009 loopt het zesjarige project af; een doorstart als ‘Kooperationsverbund’ per 1 juli 2009 is inmiddels voorzien.

De archieven en bibliotheken van Australië en Nieuw-Zeeland werken traditiegetrouw nauw samen, ook ten aanzien van duurzame toegang. In Denemarken zijn pogingen gedaan een intersectorale coalitie tot stand te brengen, maar daar is het niet gelukt.

6.2. Perspectieven voor Nederland

De respondenten uit dit onderzoek zijn vrijwel unaniem in hun inschatting dat samenwerking de duurzame toegankelijkheid van ons digitaal geheugen zal bevorderen – maar over de vraag hoe die vorm moet krijgen lopen de meningen uiteen, van zeer informele informatie-uitwisseling tot één landelijke opslagfaciliteit.

Het digitale tijdperk vraagt in elk geval nauwere samenwerking tussen de partners *binnen* de sectoren: tussen de producenten, de archieven en de gebruikers.

In de archeologie zien we hoe een groot gemeenschappelijk belang kan leiden tot goede samenwerking tussen een veelheid aan partners: de Rijksdienst voor Cultureel Erfgoed, DANS, provinciale depots en commerciële archeologische bedrijven bouwen gezamenlijk aan EDNA, een geïntegreerde informatiestructuur voor zowel onderzoeksdata als publicaties in de archeologie [interview].

Ook het CDD+ systeem van de Justitiële Informatiedienst is een voorbeeld van een situatie waarin een gedeeld probleem (de traagheid van informatieuitwisseling op papier in de strafrechtsketen) kan leiden tot nauwe samenwerking.

Gevraagd naar de perspectieven voor intersectorale samenwerking benadrukken een aantal respondenten dat ze momenteel vooral versnippering zien: 'DEN, DANS, NCDD ... ze doen allemaal iets, maar wat doen ze nu precies?' [interviews] Niettemin worden uitwisseling van kennis en expertise genoemd als het terrein waar op korte termijn de grootste winst behaald kan worden.

Sommige respondenten benadrukken dat de dynamiek in hun sector zo *anders* is dan in de andere sectoren, dat delen alleen incidenteel nut heeft. Anderen signaleren in toenemende mate dezelfde problemen, zij het dat de sectoren er verschillende namen aan geven. De kwestie van authenticiteit van gegevens, bijvoorbeeld, speelde in het analoge tijdperk wel voor archieven, maar niet of nauwelijks voor bibliotheken. Nu krijgen bibliotheken er ook mee te maken. Theorieën als het records continuüm [Jonker, 2009], die door archieven zijn ontwikkeld, kunnen een inspiratiebron zijn voor organisatiemodellen binnen wetenschap en cultuur. De wetenschap is weer het verst gevorderd in het bouwen van gedistribueerde netwerken en het bevorderen van interoperabiliteit door middel van *persistent identifiers* die voorkomen dat internetbronnen onvindbaar worden.

Onderzoek heeft aangetoond dat gebruikers een zo geïntegreerd mogelijk informatieaanbod willen. De herkomst van de bronnen (archief, bibliotheek, museum) vinden ze nauwelijks relevant [Perceptions of library and information resources, 2005]. In de *front office* is de ambitie daarmee duidelijk. Maar juist in het digitale tijdperk zijn er in de *back office* allerlei mogelijkheden om schaalvoordelen te realiseren zonder dat de gebruiker daar last van heeft.

Het Nationaal Archief en de Regionale Historische Centra hebben vlak voor de afronding van dit onderzoek een visienota gepresenteerd waarin zij aangeven als consortium te willen werken aan een algemene e-depotvoorziening voor de archiefwereld die binnen vijf jaar gerealiseerd moet zijn [e-Archief in ontwikkeling, 2009]. De vraag doet zich voor of zo'n voorziening niet ook een oplossing zou kunnen bieden voor andere bewarende instellingen in het cultureel erfgoed. De collecties liggen immers dicht bij elkaar. Ook lokale archieven zouden wellicht willen aansluiten.

Bibliotheken met digitale collecties kijken momenteel vooral naar de KB, die echter haar prioriteit legt bij landelijke projecten en programma's. De Zeeuwse Bibliotheek zoekt een veilig heenkomen voor haar digitale collecties; de Universiteit van Amsterdam heeft een aantal digitale collecties inmiddels kunnen onderbrengen bij DANS.

Of het tot concrete intersectorale samenwerkingsverbanden kan en moet komen, hangt af van vele factoren. Voor echte interoperabiliteit zouden bijvoorbeeld de metadatasystemen in archieven en bibliotheken op elkaar moeten worden aangepast, en de terminologie (archieven spreken van 'actoren', bibliotheken van 'auteurs'). Maar ook moeten culturele verschillen tussen de sectoren slijten; het kost de tijd om het noodzakelijke vertrouwen te kweken [zie

o.a. Zorich, Waibel & Erway, 2008, en Walters & MacDonald, 2008]. Een aantal respondenten schat in dat de cultuurverschillen op lokaal niveau kleiner zijn dan op landelijk niveau en dat juist daar kansen liggen om tot intersectorale samenwerking te komen tussen archieven, bibliotheken en musea.

Eén respondent benadrukt dat de grote bewarende instellingen elkaar nodig zullen hebben om hun betrouwbaarheid en de continuïteit van de zorg voor de digitale bestanden goed te regelen.

7. Conclusies en aanbevelingen

Digitale informatie is een essentieel bestanddeel van de hedendaagse kenniseconomie. De publieke sector investeert veel in de productie en verwerving van digitale informatie en wil een deel van die informatie langdurig toegankelijk houden voor gebruik door toekomstige generaties.

Het borgen van duurzame toegang tot informatie vergt in het digitale tijdperk heel andere inspanningen dan in het analoge tijdperk:

- Digitale informatie vereist zorg van wieg tot graf; al bij creatie moeten maatregelen genomen worden om langetermijntoegang mogelijk te maken. Zorgvuldig beheer en voortdurende research & development blijven nodig gedurende de hele levenscyclus van de informatie.
- Er is een technisch geavanceerde beheersomgeving nodig die steeds vernieuwd moet worden, die kostbaar is en veel expertise vraagt.

De Nationale Verkenning Digitale Duurzaamheid wijst uit dat aan deze beide randvoorwaarden nog lang niet overal wordt voldaan:

- Het aantal bewaarplaatsen met een duurzaam regime is te beperkt en de capaciteit te klein.
- Grote categorieën digitale gegevens die geproduceerd of verworven worden blijven buiten bereik van de duurzame bewaarplaatsen, bijvoorbeeld omdat het bewustzijn van de risico's onvoldoende is of omdat de (versnipperde) taak- en rolverdeling uit het analoge tijdperk de zorg van wieg tot graf in de weg staat.

Een onveranderde situatie zal zonder twijfel leiden tot verlies aan informatie, tot kapitaalvernietiging en een verzwakking van de kenniseconomie.

Een landelijk infrastructuur voor langetermijntoegang

De partijen die lid zijn van de Nationale Coalitie Digitale Duurzaamheid willen hun krachten bundelen om te komen tot een infrastructuur van faciliteiten, diensten en gereedschappen, waarbinnen:

1. het bewustzijn van het langetermijnbelang van digitale informatie tijdens de hele levenscyclus van digitale informatie aanwezig is;
2. rollen en verantwoordelijkheden duidelijk zijn en een goede aansluiting mogelijk maken tussen producenten van digitale informatie en bewaarplaatsen;
3. voldoende efficiënt ingerichte faciliteiten voor opslag en duurzaam beheer beschikbaar zijn;
4. de partijen diensten en gereedschappen zo flexibel mogelijk delen;
5. er *structurele* financiering is voor de activiteiten die nodig zijn om duurzame toegankelijkheid mogelijk te maken;
6. goede selectiemethodes beschikbaar zijn, aangepast aan het digitale tijdperk;
7. belanghebbenden kennis en expertise (nationaal en internationaal) delen.

Ad 1. Probleembewustzijn

Vrijwel alle respondenten in het onderzoek melden dat het *bewustzijn* van de problematiek van digitale informatie nog lang niet overal voldoende is doorgedrongen.

- Blijf voortdurend aandacht vragen voor het langetermijnbelang: lokaal, regionaal, landelijk en in de diverse disciplines. Investeer voortdurend in informatie, communicatie en een politieke lobby.
→ NCDD-bureau, NCDD-leden

Ad 2. Rollen en verantwoordelijkheden

Veel producenten van digitale informatie ervaren het langetermijnbelang niet als het hunne: de onmiddellijke dienst aan de burger staat voorop, of het publiceren van het eigen onderzoek. Een dergelijke tegenstelling tussen kortetermijnbelang en langetermijnbelang zal nooit helemaal uit te bannen zijn, maar een aantal kansen kan worden benut om de tegenstelling kleiner te maken:

- Geef de onderzoekers in de wetenschap zelf een belang bij het deponeren van hun onderzoeksgegevens – door hen formeel verantwoordelijkheid te geven dan wel materieel of immaterieel te belonen voor deponering. Hier ligt een rol voor de *financiers* van wetenschappelijk onderzoek, die subsidies afhankelijk kunnen maken van deponering van de data. Ook moet in internationaal verband geijverd worden voor een systeem waarbij goede datasets wetenschappelijk prestige opleveren.
→ financiers van wetenschappelijk onderzoek, universiteiten, onderzoekorganisaties, koepelorganisaties
- Maak expliciete afspraken met onderzoekers over wie eigenaar is van de onderzoeksgegevens en wie verantwoordelijk is voor welke aspecten van langetermijnzorg (metadatering, opslag, etc.).
→ universiteiten, onderzoekorganisaties, koepelorganisaties
- Stel verplicht dat ieder onderzoek- of digitaliseringsprojectplan ten minste expliciet ingaat op de noodzaak om (selecties van) de geproduceerde digitale objecten langdurig te bewaren en benoemt wie daarvoor verantwoordelijk is.
→ NCDD-leden, ministerie van OCW, NWO, KNAW, universiteiten
- Werk aan gerichte selectiemechanismes voor onderzoeksgegevens.
→ universiteiten, onderzoekorganisaties, koepelorganisaties, onderzoekers
- Sla vanuit de instellingen die langetermijnbewaring als hun kerntaak hebben de brug naar producenten van digitale informatie door al in de productiefase het primaire proces te ondersteunen en het langetermijnbelang in te bedden in het kortetermijnbelang.
→ digitale archieven, bibliotheken, datacentra

- Verwijs in de nieuwe versie van het *Standard Evaluation Protocol* expliciet naar de noodzaak om wetenschappelijke output duurzaam te bewaren.
→ VSNU, KNAW, NWO
- Breng de verantwoordelijkheid voor de informatiehuishouding van de overheid onder één bewindspersoon.
→ kabinet
- Breid het programma 'Informatie op Orde' van de rijksoverheid uit naar lokale en provinciale overheden.
→ lokale overheden, ministerie van BZK, ministerie van OCW
- Beleg een conferentie over de verantwoordelijkheid voor langetermijnbewaring van erfgoed dat op web 2.0 ontstaat.
→ NCDD-bureau, NCDD-leden, ministerie van OCW

Ad 3. Faciliteiten voor opslag en duurzaam beheer

Opslagfaciliteiten zijn duur in ontwikkeling en beheer, en onderzoek heeft aangetoond dat schaalgrootte aanzienlijke invloed heeft op de prijs per eenheid. Dit pleit ervoor om opslagfaciliteiten zoveel mogelijk te delen. De efficiency-eis kan echter in conflict komen met specifieke eisen die producenten en/of gebruikers stellen aan hun informatie. Daarom zal aanvullend onderzoek (in NCDD-verband) nodig zijn waarin de belanghebbenden de informatiearchitectuur voor duurzame opslag analyseren en bepalen waar schaalgrootte wel en niet tot voordelen kan leiden; enkele respondenten vermoeden dat de beste kansen voor gezamenlijke faciliteiten in de onderste laag van de informatiearchitectuur liggen, in de *storage layer* – dit kan ook leiden tot uitbesteden van de dataopslag bij commerciële bedrijven.

Fig. 7.1. Lagen in de informatiearchitectuur

Toegang	<i>Kansen voor generieke faciliteiten het kleinst</i>
Applicaties	
Bitstream opslag	<i>Kansen voor generieke faciliteiten het grootst</i>

De meningen verschillen over de manier waarop centrale voorzieningen georganiseerd moeten worden: getrapt geografisch (lokaal, regionaal, provinciaal, landelijk, internationaal), per discipline, per aandachtsgebied/cultuursector, of gespecialiseerd naar bestandsformaat.

- Peil de behoeftes aan opslagcapaciteit meer in detail om het landelijk aanbod daarop af te kunnen stemmen.
→ NCDD-leden via de diverse brancheorganisaties; coördinatie NCDD-bureau

- Streef waar mogelijk schaalgrootte na voor opslagcapaciteit om de kosten te beheersen.
→ allen
- Zoek naar aansluiting tussen diverse lopende initiatieven zoals het plan van het consortium van het Nationaal Archief en Regionale Historische Centra [e-Archief in ontwikkeling, 2009] en het Limburgs e-depot [Kaderstellende notitie Limburgs e-depot, 2009]. Bezie of een dergelijke infrastructuur ook soelaas kan bieden voor kleinere erfgoedinstellingen en bibliotheken, en/of daarnaast regionale intersectorale faciliteiten ontwikkeld moeten worden.
→ alle betrokken instellingen; NCDD als coördinator van het debat
- Zoek internationale fora op om voor de wetenschappelijke disciplines die internationaal georiënteerd zijn duurzame bewaarplaatsen te ontwikkelen, bijvoorbeeld de Europese Alliance for Permanent Access, DARIAH, CESSDA, en CLARIN.
→ KNAW, NWO, VSNU, universiteiten, onderzoekinstellingen

Ad 4. Diensten en gereedschappen

Met name de kleinere organisaties vragen om praktische, laagdrempelige diensten en gereedschappen om hun digitale gegevens goed te kunnen beheren. De richtlijnen voor duurzame toegankelijkheid zoals die inmiddels zijn ontwikkeld (zie hoofdstuk 3) zijn tamelijk *high-level*; ze gaan uit van de ontwikkeling van een organisatie die geheel is gericht op langetermijnbewaring en sluiten daarom niet altijd aan bij de praktijk van instellingen die een deeltaak hebben ten aanzien van bewaring van digitale gegevens.

- Ontwikkel een landelijk, sectoroverschrijdend pakket aan praktische, laagdrempelige diensten en gereedschappen dat kleinere organisaties kan helpen om hun data goed te beheren ('digitaal instrumentenbord'); overweeg het dienstenpakket onder te brengen bij een organisatie die zich specialiseert in diensten aan derden in plaats van bij één van de grote bewarende instellingen.
→ NCDD-leden; coördinatie NCDD-bureau
- Ontwerp een praktisch stappenplan dat kleinere organisaties kan helpen om de risico's van verlies van data ten minste te verkleinen, zowel in de productiefase als op de middellange termijn.
→ NCDD-leden, coördinatie NCDD-bureau
- Wacht niet op initiatieven van de grote instellingen, maar neem in elk geval praktische maatregelen die de risico's van verlies al flink verkleinen (inventarisatie maken, selecteren voor langetermijnbewaring, keuze bestandsformaten, back-up maken, metadateren, etc.)
→ kleinere archieven, erfgoedinstellingen, bibliotheken, producenten van digitale objecten

- Wees pragmatisch bij het ontwikkelen van technische faciliteiten, diensten en gereedschappen – een tijdshorizon van vijf jaar is al betrekkelijk lang in een werkveld dat zich zo snel ontwikkelt als duurzame toegankelijkheid.
→ alle instellingen met digitale depots.

Ad 5. Structurele financiering

Het is van essentieel belang dat de financierende organisaties de ontwikkeling van de infrastructuur voor digitale informatie *structureel* ondersteunen om te voorkomen dat breuken in de continuïteit ontstaan.

- Voorzie instellingen die al een eind op weg zijn in het ontwikkelen van duurzame toegankelijkheid, zoals de KB, het NIBG, DANS, NA, GAR, IISG, 3TU.Datacentrum van (uitzicht op) (aanvullende) *structurele* financiering voor het onderhouden van de infrastructuur voor duurzame toegankelijkheid.
→ ministerie van OCW, ministerie van BZK, universiteiten
- Overweeg additionele financiering uit een fonds dat wordt gevormd door een deel van de middelen voor onderzoek- en digitaliseringsprojecten voor duurzame bewaring te reserveren.
→ NCDD-leden, ministerie van OCW, NWO, KNAW, universiteiten

Ad 6. Selectie

Om de beheerskosten te beperken moet er nog meer dan in het analoge tijdperk geselecteerd worden welke informatie inderdaad voor de lange termijn toegankelijk moet blijven.

- Maak actief beleid om de selectiemechanismen uit het analoge verleden digitaal te hertalen en nieuwe mechanismes te ontwikkelen voor oorspronkelijke digitale objecten (web 2.0, onderzoeksdata).
→ NCDD-leden, alle sectoren
- Streef bij oorspronkelijk digitaal materiaal naar selectie bij creatie en automatiseer de selectie zoveel mogelijk.
→ erfgoedinstellingen, bibliotheken, archieven
- Weeg bij digitalisering goed af welke digitale objecten inderdaad voor de lange termijn bewaard moeten worden. Digitale kopieën van fysiek erfgoed die alleen zijn bedoeld voor het verlenen van online informatie behoeven beduidend minder zorg dan zogenaamde substitutie-kopieën die het origineel vervangen.
→ alle sectoren

Ad 7. Kennis en expertise

Buiten de grote bewarende instellingen geven alle respondenten aan dat kennis en expertise in Nederland momenteel versnipperd zijn en moeilijk toegankelijk. Het overzicht ontbreekt. Websites worden enthousiast opgezet, maar worden daarna slecht bijgehouden of zijn te specifiek; kennissessies worden incidenteel georganiseerd, maar de uitwisseling verwatert vaak weer snel door tijdgebrek.

- Geef kennisuitwisseling de prioriteit die het verdient om innovatie en ontwikkeling te bevorderen.
→ alle instellingen
- Maak een eind aan de versnippering van kennis en informatie over duurzame toegankelijkheid in Nederland. Zet een landelijk kenniscentrum op waar kennisdisseminatie de eerste prioriteit heeft.
→ alle NCDD-leden; coördinatie NCDD-bureau; samenwerken met o.a. ICTU, DEN
- Zet (al dan niet binnen een kenniscentrum) een landelijk gecoördineerde *technology watch* op.
→ alle bewarende instellingen; coördinatie NCDD.

De rol van de NCDD

De diverse partijen zijn in eerste instantie verantwoordelijk voor het nemen van passende maatregelen in hun eigen sector. De NCDD biedt een forum waar partijen uit diverse sectoren met elkaar om de tafel zitten en afstemming bereiken. Daarnaast ligt er een rol voor de NCDD bij de ontwikkeling van landelijke kennisfaciliteiten en een lobby bij de overheid voor de benodigde middelen.

In de discussies ten aanzien van de infrastructuur moeten ook andere belanghebbenden betrokken worden:

- brancheorganisaties als BRAIN, UKB of VSNU, de Nederlandse Museumvereniging, het AVA_Net (de opvolger van de audiovisuele sectie van KVAN en DIVA), de Taskforce Archieven, enz.
- belangrijke bewarende instellingen als JustID, KNMI, Max-Planck-Institute for Psycholinguistics, en andere.

Bijlagen

Bijlage 1. Bronnenlijst

- *ABC-DE* (2008), *Woordenboek voor het digitaal erfgoed*, uitgegeven door Digitaal Erfgoed Nederland (DEN) met steun van de Stichting GAAN. Den Haag: Digitaal Erfgoed Nederland, december 2008.
- *Addressing the uncertain future of preserving the past: towards a robust strategy for digital archiving and preservation* (2007), RAND technical report in opdracht van de Koninklijke Bibliotheek, door Stijn Hoorens, Jeff Rothenberg, Constantijn van Oranje, Martijn van der Mandele en Ruth Levitt, <http://www.kb.nl/hrd/dd/dd_links_en_publicaties/publicaties/rand_report_e-depot_TR510_3c_Cover.pdf>, geraadpleegd 9 februari 2009.
- *Almanak van het Nederlands Archiefwezen* 2008/2009. 's-Gravenhage: Koninklijke Vereniging van Archivarissen in Nederland.
- Altenhoener, Reinhard (2008): 'Implementing a Cooperative Long-term Preservation Infrastructure Solution for Heterogeneous Institutions - Report on Activities in Progress in Germany', 74th International IFLA Conference and Council, 10-14 August 2008, Québec, Canada, <<http://www.ifla.org/IV/ifla74/papers/084-Altenhoener-en.pdf>>, geraadpleegd 21 januari 2009.
- Anderson, Martha (2008): 'Evolving a Network of Networks: The Experience of Partnerships in the National Digital Information Infrastructure and Preservation Program', *International Journal of Curation* vol. 3 nr. 1, pp. 4-14, <<http://www.ijdc.net/index.php/ijdc/article/viewFile/59/38>>, geraadpleegd 14 januari 2009.
- Angevaere, Inge (2009): 'Taking care of digital collections and data: "curation" and organisational choices for research libraries', *LIBER Quarterly* vol. 19, nr. 1, pp. 1-12, <<http://liber.library.uu.nl/publish/articles/000278/article.pdf>>, geraadpleegd 17 mei 2009.
- Ayris, P., R. Davies, R., R. McLeod, R. Miao, H. Shenton en P. Wheatley (2008), *The LIFE2 final project report. Research report*. LIFE Project, London, UK, <<http://eprints.ucl.ac.uk/11758/>>, geraadpleegd 21 mei 2009.
- *Archiefbeheer en -behoud bij het Rijk* (1988), Algemene Rekenkamer.
- *Archiveren = vooruitzien* (2009), Visie van BRAIN en KVAN op de koers en inrichting van het archiefwezen, juni, <http://www.archiefbrain.nl/actueel.php?nieuws=1&nieuws_id=36>, geraadpleegd 20 juni 2009.
- Baseline Informatiehuishouding Rijksoverheid (2008), Normenkader voor duurzaam toegankelijke en betrouwbare overheidsinformatie, Programma Informatie op Orde, <http://www.e-overheid.nl/sites/kennislabs/de_baseline/de_baseline.html>, geraadpleegd 4 juni 2009.
- *De Basis voor duurzaamheid* (2009), Stichting Digitaal Erfgoed Nederland, versie 1, <<http://www.den.nl/docs/20090323032201>>, geraadpleegd 23 mei 2009.
- Beagrie, Neil; Julia Chruszcz and Brian Lavoie (2008): *Keeping research data safe: a cost model and guidance for UK universities*, <<http://www.jisc.ac.uk/media/documents/publications/keepingresearchdatasafe0408.pdf>>, geraadpleegd 12 januari 2009.
- Beagrie, Neil (2009): 'Google pulls its research datasets service', weblog 22 januari, <<http://blog.beagrie.com/archives/2009/01/22/google-pulls-its-research-datasets-service/>>, geraadpleegd 23 januari 2009.
- *Bedrijvenonderzoek Nationale Verkenning Digitale Duurzaamheid* (2009), link via <<http://www.ncdd.nl/activiteiten-natverkenning.php>> .
- *Beelden voor de toekomst* (2006), projectplan, <http://www.beeldenvoordetoekomst.nl/assets/documents/beeldenvoordetoekomst_2006.pdf>, geraadpleegd 5 juni 2009.
- Berentsen, R.H. (2007), *Digitale duurzaamheid van archieven in het CDD+*, afstudeerscriptie Technische Bedrijfskunde, Universiteit Twente, <http://essay.utwente.nl/58355/1/scriptie_R_Berentsen.pdf>, geraadpleegd 24 februari 2009.
- Beunen, Annemarie en Tjeerd Schiphof (2006), *Juridische wegwijzer archieven en musea online*, Taskforce Archieven, <<http://www.taskforce-archieven.nl/projects/juridischewegwijzer>>, geraadpleegd 17 mei 2009.

- Bossenbroek, Martin en Hans Jansen (2008), 'Auteurswet obstakel voor digitale bibliotheek', *NRC Handelsblad*, 17 april, <<http://www.kb.nl/nieuws/2008/auteursrecht.html>>.
- Cameron, Graham (2008), 'EBI and Information Infrastructure', presentatie tijdens de conferentie van de Alliance for Permanent Access, 4 november, Boedapest, <<http://www.alliancepermanentaccess.eu/documenten/cameron.ppt>>, geraadpleegd 26 mei 2009.
- CBS StatLine, <<http://statline.cbs.nl/StatWeb/publication/?DM=SLNL&PA=70692NED&D1=a&D2=16>>, geraadpleegd 3 juni 2009.
- Courant, Paul (2008), 'Some economics of digital preservation', presentation at the LIFE2 Conference, London, 23 June 2008, powerpoint presentatie <<http://www.life.ac.uk/2/conference.shtml>>, geraadpleegd 6 maart 2009.
- *De collectiebalans: een onderzoek naar het wel en wee van museumcollecties in Nederland* (2008), Lucas Veeger, in opdracht van ICN, <<http://project.collectiewijzer.nl/>>, geraadpleegd 27 juni 2009.
- *Conclusies van de Raad over digitalisering en onlinetoegekankelijkheid van cultureel materiaal en digitale bewaring* (2006), Raad van de Europese Unie, 2006/C 297/01, <<http://eur-lex.europa.eu/Notice.do?mode=dbl&lang=en&ihmlang=en&lng1=en,nl&lng2=cs,da,de,el,en,es,et,fi,fr,hu,it,lt,lv,nl,pl,pt,sk,sl,sv,&val=437034:cs&page=>>>, geraadpleegd 22 mei 2009.
- *Cultuurbeleid in Nederland* (2007), red. Cas Smithuijsen; Den Haag/Amsterdam : Ministerie van OCW/Boekmanstudies.
- *Dealing with data: roles, rights, responsibilities and relationships* [2007], Consultancy report by Liz Lyon, UKOLN, <http://www.jisc.ac.uk/media/documents/programmes/digitalrepositories/dealing_with_data_report-final.pdf>, geraadpleegd 4 juni 2009.
- Digitaliseren met beleid (2006), subsidieregeling van het ministerie van OCW, <<http://www.senternovem.nl/digitaliserenmetbeleid/index.asp>>
- *Duurzame wetenschap* [2006]. Strategisch plan 2007-2010. KNAW, <<http://www.knaw.nl/publicaties/pdf/20061041.pdf>>, geraadpleegd 3 juni 2009.
- *e-Archief in ontwikkeling: duurzame digitale toegang tot informatie* (2009), Visiedocument 2009-2014, Consortium van Regionale Historische Centra en het Nationaal Archief, *nog geen url beschikbaar bij publicatie van dit rapport; zie website Nationaal Archief*.
- *ED3, Eisen Duurzaam Digitaal Depot* (2008), toetsingskader voor de beheersomgeving van blijvend te bewaren digitale informatie, versie 1, Landelijk Overleg Archief Inspecteurs, <http://www.lopai.nl/pdf/ED3_v1.pdf>.
- *Een dementerende overheid? De risico's van digitaal beheer van verantwoordingsinformatie bij de centrale overheid* (2005), Rijksarchiefinspectie, <<http://www.erfgoedinspectie.nl/media/publications/Een%20dementerende%20overheid%20versie%20DEF%20aerial.pdf>>, geraadpleegd 12 januari 2009.
- Digitaal Erfgoed Nederland (2008), Gespreksverslagen Analyseronde 2 over 'De Basis' met 14 culturele instellingen, september/december 2008, <<http://wiki.den.nl/Analyseperiode2>>, geraadpleegd 23 januari 2009.
- *De digitale feiten: Onderzoek naar de omvang en kosten van gedigitaliseerd cultureel erfgoed* (2009), Stichting DEN, Den Haag, januari, <<http://www.den.nl/docs/20080319152256/>>
- Digital Preservation Coalition (DPC) (2008): *Promoting Digital Preservation. Annual Report 2007-2008*, <<http://www.dpconline.org/graphics/reports/index.html#annual>>, geraadpleegd 21 januari 2009.
- *The expanding digital universe: a forecast of worldwide information growth through 2010* (2007), an IDC White Paper by John F. Gantz et al., <<http://www.emc.com/collateral/analyst-reports/expanding-digital-idc-white-paper.pdf>>
- *Gewaardeerd verleden: bouwstenen voor een nieuwe waarderingsmethodiek voor archieven* (2007), Rapport van de Commissie waardering en selectie [cie. Jeurgens], <http://www.nationaalarchief.nl/images/waardering_en_selectie.pdf>, geraadpleegd 10 februari 2009.
- Gladney, H.M. (2007): 'Digital Preservation in a National Context: Questions and Views of an Outsider', *D-Lib Magazine* vol. 13 nr. 1/2, <<http://www.dlib.org/dlib/january07/gladney/01gladney.html>>, geraadpleegd 21 januari 2009.

- *De grenzen voorbij: wetenschappelijke informatievoorziening in een nieuw perspectief* [2007], beleidsplan 2007-2010 van UKB, het samenwerkingsverband van universiteitsbibliotheken en de Koninklijke Bibliotheek, <<http://www.ukb.nl/organisatie/algemenestukken/ukb07-001.pdf>>, geraadpleegd 6 juni 2009.
- Grijpink, J.H.A.M. (2005), 'Onze informatiesamenleving in wording: de uitdaging van grootschalige informatie-uitwisseling in de rechtstaat', Rede uitgesproken bij de openbare aanvaarding van het ambt van bijzonder hoogleraar . . . Universiteit Utrecht, 19 januari 2005, <<http://www.keteninformatisering.nl/2005%20oratie%20keteninformatisering%20in%20de%20rechtstaat.pdf>>, geraadpleegd 2 maart 2009.
- Grijpink, J.H.A.M. (2006), *Keteninformatisering in kort bestek: Theorie en praktijk van grootschalige informatie-uitwisseling*. Den Haag: LEMMA, 2006.
- *Informatie: grondstof met toekomstwaarde. Contouren van een visie op de rol en betekenis van informatie* (2008), Advies van de Raad voor Cultuur en de Raad voor het openbaar bestuur, <http://www.cultuur.nl/adviezen_vervolg.php?id=4&advies=6333>, geraadpleegd 21 januari 2009.
- *Informatie op orde: Kabinetsvisie op vindbare en toegankelijke overheidsinformatie* (2006), gezamenlijke uitgave van het Ministerie van BZK en het Ministerie van OCW, <<http://www.minocw.nl/documenten/Nota-Informatie-op-orde.pdf>>, geraadpleegd 17 mei 2009.
- *Infrastructure planning and data curation: A comparative study of international approaches to enabling the sharing of research data* (2008), prepared by Kaivo Ruusalepp for DCC and JISC, november 2008, <http://www.dcc.ac.uk/docs/publications/reports/Data_Sharing_Report.pdf>, geraadpleegd 23 maart 2009.
- *Innoveren, Participeren! Adviesagenda cultuurbeleid en culturele basisinfrastructuur* (2007), Raad voor Cultuur, <<http://www.agendacultuurbeleid.nl/indexAdvies.html>>, geraadpleegd 17 mei 2009.
- *International Study on the Impact of Copyright Law on Digital Preservation* (2008), A joint report of The Library of Congress National Digital Information Infrastructure and Preservation Program, the Joint Information Systems Committee, the Open Access to Knowledge (OAK) Law Project, the SURFfoundation, <http://www.digitalpreservation.gov/library/resources/pubs/docs/digital_preservation_final_report2008.pdf>, geraadpleegd 12 januari 2009.
- JISC (2003), JISC Circular 6/03 (Revised), *An invitation for expressions of interest to establish a new Digital Curation Centre for research into and support of the curation and preservation of digital data and publications*, <<http://www.dcc.ac.uk/docs/6-03Circular.pdf>>, geraadpleegd 15 februari 2009.
- Jonker, Agnes E.M. (2009): 'Een scherpere kijk op de informatiehuishouding', *Informatieprofessional* nr. 1, p. 24-29.
- *notitie Limburgs e-depot* (2009), opgesteld door KVAL Denktank @-depot, versie 1.0, 10 maart 2009. http://www.rijckheydt.nl/upload/100/pdf/Archiefinspectie%20-%20dossier%20digitaal%20archiveren/Notitie_KVAL_edepot_pdf.pdf, geraadpleegd 16 mei 2009.
- *Keeping the records of science accessible: Can we afford it? Report on the 2008 Annual Conference of the Alliance for Permanent Access* (2008), reported by Inge Angevaere, <http://www.alliancepermanentaccess.eu/documenten/Alliance2008conference_report.pdf>, geraadpleegd 6 maart 2009.
- 'Landschap wetenschappelijk onderzoek Nederland', Nederlandse Onderzoek Databank (NOD), <<http://www.onderzoekinformatie.nl/nl/oi/landschap/>>, geraadpleegd 18 mei 2009.
- Lavoie, Brian F. (2004), *The Open Archival Information System Reference Model: introductory guide*, DPC Technology Watch Series Report 04-01, January, <http://www.dpconline.org/docs/lavoie_OAIS.pdf>, geraadpleegd 16 mei 2009.
- *The LIFE2 Final Project Report*, <<http://eprints.ucl.ac.uk/11758/>>, geraadpleegd 4 juni 2009.
- *Long-term preservation: Results from a survey investigating preservation strategies amongst ALPSP publisher members* (2008), prepared by Sarah Durrant, <http://www.alpdp.org/ngen_public/article.asp?id=&did=47&aid=27202&st=&oaid=-1>, geraadpleegd 26 juni 2009.
- Lynch, Clifford (2009), niet gepubliceerde presentatie op de DigCCurr conferentie 2009, zie <<http://digitaalduurzaam.blogspot.com/2009/04/digccurr-4-de-visies.html>>, geraadpleegd 24 mei 2009.

- *Machineleesbare gegevensbestanden: archivering en beheer bij het Rijk* (1991). Rapport van de Algemene Rekenkamer, 's-Gravenhage.
- *Mind the Gap: assessing digital preservation needs in the UK* (2006), prepared for the Digital Preservation Coalition by Martin Waller, Robert Sharp (Tessella), <<http://www.dpconline.org/graphics/reports/mindthegap.html>>.
- National Archives of Australia (2002), *An approach to the preservation of digital records*, <http://www.naa.gov.au/Images/An-approach-Green-Paper_tcm2-888.pdf>
- *Nederland Open in Verbinding* (2007), Een actieplan voor het gebruik van open standaarden en open source software bij de (semi-)publieke sector, Den Haag: Ministerie van Economische Zaken, <<http://www.ososs.nl/files/Actieplan%20Nederland%20Open%20in%20Verbinding.pdf>>, geraadpleegd 19 juni 2009.
- Nederlandse Roadmap Grootchalige Onderzoeksfaciliteiten (2008), Nederlandse Commissie voor Grootchalige Onderzoeksfaciliteiten, <<http://www.minocw.nl/documenten/Roadmapdefinitief.pdf>>, geraadpleegd 6 juni 2009.
- 'Nederlandse Gedragscode Wetenschapsbeoefening' (2004), VSNU, Vereniging van Universiteiten, te downloaden vanaf <<http://www.vsnul.nl/universiteiten/publicaties.htm>>, geraadpleegd 18 mei 2009.
- *NEN-ISO 15489-1 (nl), Informatie en documentatie – Informatie en archiefmanagement*, ICS 01.140.20, Nederlands Normalisatie-instituut, november 2001.
- OECD Principles and guidelines for access to research data from public funding (2007), <<http://www.oecd.org/dataoecd/9/61/38500813.pdf>>, geraadpleegd 23 mei 2009.
- Oltmans, Erik en Nanda Kol (2005), 'Emulation and migration: a comparison in terms of costs', DCC/DPC Seminar on Costs Models, British Library, July 26, 2005, <<http://www.dpconline.org/docs/events/050726oltmans.pdf>>, geraadpleegd 21 mei 2009.
- PARSE.Insight (2009), *Draft Road Map*, Deliverable D2.1, <http://www.parse-insight.eu/downloads/Parseinsight_draft_roadmap_20090327.pdf>, geraadpleegd 24 mei 2009.
- *Perceptions of library and information resources* (2005), OCLC, <<http://www.oclc.org/reports/2005perceptions.htm>>, geraadpleegd 6 juni 2009.
- *Play-out: Nieuwe technieken voor toegankelijkheid en conservering van de videokunstcollectie in Nederland* (2009), nog ongepubliceerd rapport NIMk/Montevideo.
- Potter, Abbey (2009), 'NDIIPP update', presentatie voor de Digital Preservation Interest Group, ALA, 24 januari, <http://www.digitalpreservation.gov/library/resources/pubs/docs/ALAMid2009_Potter.pdf>, geraadpleegd 26 mei 2009.
- *Preserving our digital heritage: Plan for the National Digital Information Infrastructure and Preservation Program* (2002). A collaborative initiative of the Library of Congress. <http://www.digitalpreservation.gov/library/resources/pubs/docs/ndiipp_plan.pdf>, geraadpleegd 11 februari 2009.
- *Preserving the digital heritage* (2007), edited by Yola de Lusenet en Vincent Wintermans, Netherlands National Commission for Unesco, <<http://www.unesco.nl/documents/Preserving%20the%20digital%20heritage.pdf/view>>, geraadpleegd 22 mei 2009.
- Prins, J.E.J. en L.J. Matthijssen (2000): *De digitale overheid en de wet: de juridische kaders voor het gebruik van digitale documenten bij overheden*. Den Haag: Programma Digitale Duurzaamheid.
- *Rapport e-Depot, Stand van zaken 2008* (2009), Koninklijke Bibliotheek, intern document.
- *Rapport inventarisatie digitale duurzaamheid* (2008), KNAW, intern document.
- Ras, Marcel (2009), 'Long-term preservation of scientific publications in practice: The KB e-Depot', onderdeel van de presentatie Policies, preconditions and costs' op de conferentie *Curating Research*, Den Haag, 17 april 2009, <<http://www.kb.nl/hrd/congressen/curatingresearch2009/presentations/beagrie-ras.pdf>>, geraadpleegd 5 juni 2009.
- 'Reactie van Nederlandse erfgoedinstellingen op het Groenboek *Auteursrecht in de kennis-economie*' (2008), 28 november, <<http://forum.archieven.org/index.php?topic=19119.0>>, geraadpleegd 28 januari 2009.

- Rijksdienst voor het Cultureel Erfgoed, website, <<http://www.cultureelerfgoed.nl/werken/wetten-en-regels>>, geraadpleegd 3 juni 2009.
- 'Roles and Responsibilities for Data Curation: The Data Librarian' (2008), powerpoint presentation by Robin Rice at the 2nd DCC/RIN Data Management Forum, 28 November 2008, <<http://www.dcc.ac.uk/events/data-forum-2008-november/presentations/05.ppt#256,1,Roles>> & Responsibilities for Data Curation: the Data Librarian, geraadpleegd 27 maart 2009.
- Rosenthal, David (2009), Presentation for the opening plenary of the CNI Spring Task Force Meeting, <<http://blog.dshr.org/2009/04/spring-cni-plenary-remix.html>>, geraadpleegd 16 mei 2009.
- Rothenburg, Jeff (1995): 'Ensuring the Longevity of Digital Information' [eerste versie gepubliceerd in *Scientific American*, vol. 272, nr. 1, pp. 24-29; herziene versie 1999, <<http://www.clir.org/pubs/archives/ensuring.pdf>>, geraadpleegd 16 mei 2009.
- Rusbridge, Chris (2006): 'Excuse me . . . Some Digital Preservation Fallacies', *Ariadne* nr. 46, <<http://www.ariadne.ac.uk/issue46/rusbridge/>>, geraadpleegd 28 januari 2009.
- Rusbridge, Chris (2009), 'A national research data infrastructure?', Digital Curation weblog, 5 februari, <<http://digitalcuration.blogspot.com/2009/02/national-research-data-infrastructure.html>>, geraadpleegd 16 februari 2009.
- Sawyer, Donald (2009), Presentatie op de DigCCurr-conferentie, 1-3 april 2009, Chapel Hill, North Carolina. Helaas is de presentatie niet gepubliceerd. De informatie komt uit eigen aantekeningen van IA.
- *Schwerpunktinitiative Digitale Information der Allianz-Partnerorganisationen* (2008), <http://www.wissenschaftsrat.de/texte/Allianz-digitale%20Info.pdf> / *Priority Initiative 'Digital Information'* (2008) by the Alliance of German Science Organisations, http://www.dfg.de/forschungsfoerderung/wissenschaftliche_infrastruktur/lis/download/allianz_initiative_digital_information_en.pdf, geraadpleegd 2 februari 2009.
- *To Share or Not to Share: Publication and Quality Assurance of Research Data Outputs* (2008), Report commissioned by the Research Information Network (RIN), <<http://www.rin.ac.uk/data-publication>>, geraadpleegd 12 januari 2009.
- Sierman, Barbara (2009), 'The jigsaw puzzle of digital preservation', *Liber Quarterly* vol. 19 nr. 1, pp. 13-21, <<http://liber.library.uu.nl/publish/articles/000262/article.pdf>>, geraadpleegd 17 mei 2009; een Nederlandstalige versie verscheen in druk in Bart de Nil & Jeroen Walterus (red.), *Erfgoed 2.0: nieuwe perspectieven voor digitaal erfgoed*, Brussel: FARO, Vlaams steunpunt voor cultureel erfgoed, pp. 123-140.
- Slats, Jacqueline (2008), 'Wat is een e-depot?', presentatie op de Studiedag Digitale Depots in Antwerpen, <http://www.gemeentearchief.rotterdam.nl/content/images/stories/content/vakgenoten/edepot/151008_wat_is_een_edepot.pdf>, geraadpleegd 24 juni 2009.
- Sorgdrager, W., en J.K.T. Postma, *Meters maken! Rapport Commissie CAS en archiefketen*, Den Haag 2006, pp. 11-14, <<http://www.handelingenbank.info/rss/rapport.pdf>>, geraadpleegd 20 juni 2009.
- Standard Evaluation Protocol 2003-2009 for Public Research (2003), VSNU, KNAW en NWO, <http://www.qanu.nl/comasy/uploadedfiles/sep2003-2009.pdf>, geraadpleegd 26 juni 2009.
- *Stewardship of Research Data in Canada: A Gap Analysis*, Research Data Strategy Working Group, October 2008, <<http://data-donnees.gc.ca/docs/GapAnalysis.pdf>>, geraadpleegd 4 februari 2009.
- SURFfoundation (2009), 'Samenwerken aan verbeterde toegang onderzoeksdata', nieuwbericht 25 juni 2009, <<http://www.surffoundation.nl/nl/actueel/Pages/Samenwerkenaanbeteretogangtotonderzoeksdata.aspx>>, geraadpleegd 27 juni 2009.
- SURFshare programma 2007-2010, SURFfoundation, <http://www.surffoundation.nl/SiteCollectionDocuments/SURFshare-programma_20072010.pdf>, geraadpleegd 6 juni 2009.
- *Sustaining the Digital Investment: Issues and Challenges of Economically Sustainable Digital Preservation* (2008), Interim Report of the Blue Ribbon Task Force on Sustainable Digital Preservation and Access, <http://brtf.sdsc.edu/biblio/BRTF_Interim_Report.pdf>, geraadpleegd 12 januari 2009.

- Testbed digitale bewaring (2005), *Costs of digital preservation*. The Hague: Nationaal Archief of the Netherlands, <<http://www.digitaalduurzaamheid.nl/bibliotheek/docs/CoDPv1.pdf>>, geraadpleegd 21 mei 2009.
- *Towards a competitive ICT infrastructure for scientific research in the Netherlands* (2008), ICTRegie, <http://www.ICTRegie.nl/publicaties/nl_08-NROI-258_Advies_ICT_infrastructuur_vdef.pdf>, geraadpleegd 12 februari 2009.
- Treloar, Andrew, David Groenewegen en Cathrine Harboe-Ree (2007), 'The data curation continuum', *D-Lib Magazine* vol. 13, no. 9/10, <<http://www.dlib.org/dlib/september07/treloar/09treloar.html>>, geraadpleegd 20 juni 2009.
- *Trusted digital repositories: attributes and responsibilities* (2002), An RLG-OCLC Report, <<http://www.oclc.org/programs/ourwork/past/trustedrep/repositories.pdf>>, geraadpleegd 16 mei 2009.
- *The UK Research Data Service Feasibility Study Interim Report* (2008a), 7 July 2008, <[http://www.ukrds.ac.uk/UKRDS%20SC%2010%20July%2008%20Item%205%20\(2\).doc](http://www.ukrds.ac.uk/UKRDS%20SC%2010%20July%2008%20Item%205%20(2).doc)>, geraadpleegd 3 februari 2009.
- *The UK Research Data Service Feasibility Study, Final Report* (2008b), 8 December 2008, <<http://www.ukrds.ac.uk/HEFCE%20UKRDS%20Final%20Report%20V%201.1.doc>>, geraadpleegd 26 mei 2009.
- *Waardevolle data en diensten: motiveren door ondersteunen van dataproducenten en onderzoeken van voordelen van data publicatie* (2008), Project in het kader van SURFshare door 3TU.Federatie; penvoerende instelling TU Delft Library. Projectbeschrijving versie 1.0, 15 december 2008, niet gepubliceerd intern document.
- 'Wat digitaal is past in het edepot' [2009], *Wageningen UR nieuwsbrief 2009-1*, <http://library.wur.nl/WebQuery/nieuws?nieuws/@isn=816&wg_sfx=nieuwsbrief>, geraadpleegd 6 juni 2009.
- Wansleeben, Milco en Marjolein van den Dries (2008), *Wegwijzer digitaal deponeren archeologie*, DANS Data Guide 3, Amsterdam: Aksant, <http://www.dans.knaw.nl/nl/dans_publicaties/pdfarcheologie/>, geraadpleegd 22 mei 2009.
- Walters, Tyler O. and Robert H. McDonald (2008): 'Creating Trust Relationships for Distributed Digital Preservation Federations', *Proceedings of the Fifth International Conference on Preservation of Digital Objects*, The British Library, 29-30 September, pp. 197-204, <http://www.bl.uk/ipres2008/presentations_day2/31_Walters.pdf>, geraadpleegd 27 mei 2009.
- 'Werkagenda Taskforce Archieven 2009: van pionieren naar verankeren', intern document Taskforce Archieven, versie 19 november 2008. Website taskforce: <<http://www.taskforce-archieven.nl/>>.
- *Wetenschap gewaardeerd!* [2006], NWO strategienota 2007-2010, <http://www.nwo.nl/nwohome.nsf/pages/NWOP_5SME25>, geraadpleegd 3 juni 2009.
- Wever, Tom (2008), presentatie tijdens de workshop digitale duurzaamheid op de DEN-conferentie 2008, uitspraak uit persoonlijke notities, zie <<http://digitaalduurzaam.blogspot.com/2008/12/duurzaamheid-ook-een-kwestie-van.html>>, geraadpleegd 27 mei 2009.
- Wright, Richard, Matthew Addis en Ant Miller (2008), 'Significance of storage in the "cost of risk" of digital preservation', presentatie op de iPRES conferentie 2008, <http://www.bl.uk/ipres2008/presentations_day1/21_Wright.pdf>, geraadpleegd 21 mei 2009. Het citaat staat niet in de tekst, maar is ter plekke opgetekend door IA; zie <<http://www.digitaalduurzaam.blogspot.com>>, linkerkolom met citaten, geraadpleegd 21 mei 2009.
- Zorich, Diane M., Gunter Waibel and Ricky Erway (2008): *Beyond the Silos of the LAMs: Collaboration among Libraries, Archives and Museums*, Report produced by OCLC Programs and Research, <<http://www.oclc.org/programs/reports/2008-05.pdf>>, geraadpleegd 28 januari 2009.

Bijeenkomsten en congressen

- 'Curating Research', internationale conferentie voor wetenschappelijke bibliotheken bij de KB op 17 april 2009, <http://www.kb.nl/curatingresearch>, en verslag op <http://digitaalduurzaam.blogspot.com/2009/04/curating-research-1-wrap-up.html>, geraadpleegd 17 mei 2009.
- DEN-conferentie 'Naar buiten', 9-10 december 2008, http://www.den.nl/innovatie/de_conferentie2008, en verslag op <http://digitaalduurzaam.blogspot.com/2008/12/duurzaamheid-ook-een-kwestie-van.html>, geraadpleegd 17 mei 2009.
- Gemeentearchief Rotterdam, expert meeting over de vraag of het (technisch) beheer van een digitaal archief mag en kan worden uitbesteed aan derden, 21 april 2009, verslag op <http://digitaalduurzaam.blogspot.com/2009/04/de-bits-en-de-bytes-van-het-archief.html>, geraadpleegd 17 mei 2009.
- KennisLAB, werkconferentie 'Het e-depot in!', 17 juni 2009, http://www.e-overheid.nl/sites/kennislab/evenementen/werkconferentie_het_edepot_in-1.xml, geraadpleegd 24 juni 2009.
- SURFfoundation workshop onderzoeksdata, verslag op <http://digitaalduurzaam.blogspot.com/2009/03/hoe-houden-we-onderzoeksdata.html>, geraadpleegd 17 mei 2009.

Websites

- 3TU Datacentrum, http://www.3tu.nl/nl/nieuws/nieuws_nederlands/?tx_ttnews%5Btt_news%5D=60&tx_ttnews%5Byear%5D=2008&tx_ttnews%5Bmonth%5D=02&cHash=284f9a9ce9, geraadpleegd 24 juni 2009.
- Archiefinspecties, <http://www.archiefinspecties.nl>, geraadpleegd 17 mei 2009.
- Blog 'duurzame toegang', <http://www.digitaalduurzaam.blogspot.com/>, geraadpleegd 24 juni 2009.
- CBS, <http://www.cbs.nl/nl-NL/menu/home/default.htm>, geraadpleegd 24 juni 2009.
- CESSDA, <http://www.cessda.org/project>, geraadpleegd 24 juni 2009.
- CLARIN, <http://www.clarin.eu>, geraadpleegd 24 juni 2009.
- Cultureel erfgoed, <http://www.cultureelerfgoed.nl/werken/wetten-en-regels>, geraadpleegd 24 juni 2009.
- DAF, Data Audit Framework, <http://www.data-audit.eu/>, geraadpleegd 24 juni 2009.
- DANS, Data Archiving and Networked Services, <http://www.dans.knaw.nl>, geraadpleegd 24 juni 2009.
- DARIAH, <http://www.dariah.eu/>, geraadpleegd 24 juni 2009.
- Data Seal of Approval, <http://www.datasealofapproval.org/>, geraadpleegd 24 juni 2009.
- DAVID Vademecum, <http://www.edavid.be/davidhandboek/index.html>, geraadpleegd 24 juni 2009.
- DEN De Basis, <http://www.den.nl/debasis>; De Basis voor duurzaamheid <http://www.den.nl/docs/20090323032201>, geraadpleegd 4 juni 2009.
- Digitaleduurzaamheid.nl, overzicht van literatuur over technische aspecten: <http://www.digitaleduurzaamheid.nl/index.cfm?paginakeuze=62&categorie=2>, geraadpleegd 5 juni 2009.
- Dioscuri, open source emulatiesoftware <http://dioscuri.sourceforge.net/>, geraadpleegd 18 juni 2009.
- DRAMBORA, <http://www.repositoryaudit.eu>, geraadpleegd 24 juni 2009.
- DoBeS, Dokumentation Bedrohter Sprachen, <http://www.mpi.nl/>, geraadpleegd 24 juni 2009.
- e-data en research, <http://www.edata.nl/>, geraadpleegd 6 juni 2009.
- EDNA, <http://edna.itor.org/nl/>, geraadpleegd 6 juni 2009.
- Esp@cenet, portal van the European Patent Office, www.espacenet.com, geraadpleegd 1 juli 2009.

- BBMRI, <<http://www.bbmri.eu>>, geraadpleegd 6 juni 2009.
- ICT-Register voor het digitaal erfgoed, Digitaal Erfgoed Nederland, <<http://www.den.nl/register>>, geraadpleegd 23 mei 2009.
- ICTU, Stichting ICTU, <http://www.ictu.nl/index.php?option=com_frontpage&Itemid=1>, geraadpleegd 24 juni 2009.
- iRODS, <http://www.diceresearch.org/DICE_Site/About.html>, geraadpleegd 6 juni 2009.
- KB e-Depot, <<http://www.kb.nl/dnp/e-depot/e-depot.html>>, geraadpleegd 24 juni 2009.
- KNAW, Koninklijke Nederlandse Akademie voor Wetenschappen, <<http://www.knaw.nl>>
- Kriterienkatalog vertrauenswürdige digitale Langzeitarchive, <http://www.langzeitarchivierung.de/modules.php?op=modload&name=PagEd&file=index&page_id=2#section10>, geraadpleegd 24 juni 2009.
- LIFE2, Lifecycle Information for e-Literature, <<http://www.life.ac.uk/2/>>, geraadpleegd 21 mei 2009.
- LOFAR, <<http://www.lofar.nl/>>, geraadpleegd 20 juni 2009.
- LOCKSS, <<http://www.lockss.org/lockss/Home>>, geraadpleegd 1 juni 2009.
- Max-Planck-Institute for Psycholinguistics, <<http://www.mpi.nl/>>, geraadpleegd 24 juni 2009.
- MIXED, Migration to intermediate XML for electronic data, <<http://www.dans.knaw.nl/nl/projectenpagina/mixed/>>, geraadpleegd 6 juni 2009.
- Mondriaanstichting, <<http://www.mondriaanfoundation.nl>>, geraadpleegd 24 juni 2009.
- Nationaal Archief, <<http://www.nationaalarchief.nl>>, geraadpleegd 24 juni 2009.
- Nederlands Instituut voor Mediakunst/Montevideo, <<http://www.nimk.nl/nl/>>, geraadpleegd 26 juni 2009.
- Nederlandse Museumvereniging, <<http://www.museumvereniging.nl>>, geraadpleegd 24 juni 2009.
- NOD, Nederlandse Onderzoek Databank, <<http://www.onderzoekinformatie.nl>>, geraadpleegd 18 mei 2009.
- NWO, <<http://www.nwo.nl/nwohome.nsf/pages/index>>, geraadpleegd 24 juni 2009.
- Overheid, <<http://www.overheid.nl>>, geraadpleegd 17 mei 2009.
- PARSE.insight, <<http://www.parse-insight.eu>>, geraadpleegd 24 mei 2009.
- Pictura, <<http://www.pictura-im.nl/nl/e-depot>>, geraadpleegd 21 juni 2009.
- PLANETS, <<http://www.planets-project.eu/about>>, geraadpleegd 24 juni 2009.
- PrestoPRIME, <www.prestoprime.org>, geraadpleegd 24 juni 2009.
- ProArchive, <<http://instituut.beeldengeluid.nl/index.aspx?ChapterID=8633>>, geraadpleegd 24 juni 2009.
- SURFfoundation, <<http://www.surfoundation.nl>>, geraadpleegd 24 juni 2009.
- SURFshare, verrijkte publicaties, <<http://www.surfoundation.nl/nl/themas/openonderzoek/verrijktepublicaties/Pages/Default.aspx>>, geraadpleegd 26 juni 2009.
- TIPR, Towards Interoperable Preservation Repositories, <<http://wiki.fcla.edu:8000/TIPR>>, geraadpleegd 24 juni 2009.
- TRAC, Trustworthy Repositories Audit and Certification, <<http://www.crl.edu/content.asp?l1=13&l2=58&l3=162&l4=91>>, geraadpleegd 24 juni 2009.
- TST, Centrale voor Spraak en Taaltechnologie, <http://www.inl.nl/index.php?option=com_content&task=view&id=448&Itemid=552>, geraadpleegd 24 juni 2009.
- VSB-fonds, <<http://www.vsbfonds.nl/>>, geraadpleegd 24 juni 2009.

Bijlage 2. Lijst van geïnterviewde personen en instanties

Alkoven		Patricia	Hoofd collectie	Nederlands Architectuur Instituut
Anderson		Martha	Director	NDIPP USA
Beek	van der	Hanneke	Information officer	Instituut Collectie Nederland
Beer	de	Hylle	Gemeentearchivaris Deventer	Saxion
Berendse		Martin	Directeur en Rijksarchivaris	Nationaal Archief
Bongenaar		Herman	Archiefinspecteur	Erfgoedcentrum DIEP
Bussel	van	Geert-Jan	Docent /directeur	UVA/HVA & Van Bussel documentservices
Ceton		Nick	Business development manager	Technicolor
Dekker		Ron	Directeur Financien, instituten en infrastructuur	NWO
Dieren	van	Arjen	Manager digitaal werken	Rabobank
Doorn		Peter	Directeur	DANS
Duyvelaar		Kees	Senior medewerker informatiebeleid	Vereniging Nederlandse Gemeenten (VNG)
Fleurbaay		Ellen	Hoofd afdeling publieksdiensten	Stadsarchief Amsterdam
Gathier		Roel	Acting General Director	SRON
Goedewaagen		Aart	Adviseur digitale bewaring	Justitiële Informatiedienst
Gordebeke		Erik	ERM architect	Achmea
Heijden	van der	Marianne	Informatiespecialist	NIOO-KNAW
Heijne		Maria	Directeur	3TU.Datacentrum
Helmus		Wilbert	Hoofd collectie	Fries Museum/Prinsessehof /SIMIN
Hoekstra		Rik	Projectmanager	Min. BZK, P-Direkt
Horsman		Peter	Wetenschappelijk medewerker	Universiteit van Amsterdam
Houbraken		Chido	Archiefinspecteur	Provincie Flevoland
Jansen		Hans	Directeur E-strategie	Koninklijke Bibliotheek
Jansen		Martin	I.D.M. adviseur	Gemeente Deventer
Janssen		Rik	Senior Beleidsmedewerker	KNAW

Jongmans		Lydia	Beleidsmedewerker cultuur	Vereniging Nederlandse Gemeenten (VNG)
Ketellapper		Wieger	Directeur algemene zaken	Buma/Stemra
Kilbride		William	Executive Director	DPC - UK
Kleijn		Erik	Hoofd gebouwd erfgoed	Rijksdienst voor het cultureel erfgoed
Koch		Ingmar	Beleidsmedewerker archiefinspectie Noord-Brabant en Limburg	Provincie Noord-Brabant
Kramer		Rob	Hoofdinspecteur collecties en archieven	Min. OCW, Erfgoedinspectie
Krutzen		Ben	Global programme manager records management	Shell
Lazorchak		Butch	digital archivist	Library of Congress
Lensink		Hester	Conservator	Instituut Collectie Nederland
Linde		Frank	Directeur	NIKHEF
Manders		Hans	Coordinator R&D informatie- en waarneeminfrastructuur	KNMI
Meijers		Enno	Hoofd ICT	Zeeuwse Bibliotheek
Mossink		Wilma		SURFfoundation
Niet	de	Marco	Directeur	Stichting DEN
Noordhoek		Irmgard	Sectormanager mediadocumentatie	Beeld&Geluid
Ras		Marcel	Hoofd E-depot	KB
Rombouts		Jeroen	Productonderzoeker	3TU Datacentrum
Rongen		Rob	IT Demand manager	Philips
Schumann		Natascha	Coordinator	nestor
Schuurman Hess		Jacqueline	Archiefsinspecteur	Gemeentearchief Rotterdam
Shenton		Helen	Board of Directors	DPC - UK
Slats		Jacqueline	Programmamanager Digitaal Nationaal Archief	Nationaal Archief
Sligman		Henk	Teammanager document management	Gemeente Deventer
Spek		Wouter	Executive Director	Alliance for Permanent Access
Steenhuis		Jantje	Gemeentearchivaris / directeur	Gemeentearchief Rotterdam
Tindemans		Peter	Acting director	Alliance for permanent access
Valentijn		Edwin	Hoogleraar "astronomische informatietechnologie"	Kapteyn instituut, Rijksuniversiteit Groningen

Valk	van der	Aad	Senior beleidsadviseur	Beeld&Geluid
Leeuwen	van	Alex	Archiefinspecteur	Erfgoedcentrum DiEP
Oss	van	Jeroen	Gemeentearchivaris	Gemeentearchief Utrecht
Velthuijs		Paula	Inspecteur collecties en archieven	Min. OCW, Erfgoedinspectie
Velzen	van	Ernst	Sectormanager ICT	Beeld&Geluid
Venema		Dany	Hoofd expertise- en dienstencentrum	Justitiële Informatiedienst
Vernooy		Marjan	Programmamanager ICT en onderzoek	SURFfoundation
Vries	de	Bert	Directeur	Historisch centrum overijssel
Wansleebe		Milco	Projectleider EDNA	DANS
Weijers		Gaby	Collectie en Conservering	NIMk/Montevideo
Weijers		Don	Hoofd bibliotheek	CBS
Werf	van der	Titia	Adjunct-directeur	IISG
Westenbrink		Renee	Beleidsmedewerker	VSNU
Westerhof		Hans	Programmadirecteur Beelden voor de toekomst	Beeld&Geluid
Westrienen	van	Gerard		SURFfoundation
Wit	de	Gerard	Hoofd cultuurhistorische informatie	Rijksdienst voor het cultureel erfgoed
Woutersen-Windhouer		Saskia		NIOO-KNAW

Bijlage 3. Lijst van gebruikte termen

Archiefbescheiden – informatie opgemaakt, ontvangen en onderhouden als bewijs en informatie door een organisatie of persoon bij het vervullen van wettelijke verplichtingen of bij zakelijke transacties [NEN-ISO 15489-1, p. 6]

Archiefvormer (ook wel zorgdrager genoemd) – dit is de term binnen de overheid voor de organisatie (landelijk, regionaal of lokaal) die bij het uitoefenen van overheidstaken archiefbescheiden produceert en daarvoor verantwoordelijkheid draagt totdat overbrenging naar een archief plaatsvindt.

Authentiek – een record is authentiek als kan worden bewezen dat het record is wat het beweert te zijn, dat het opgemaakt of verzonden is door de persoon die beweert het te hebben opgemaakt of verzonden en dat het opgemaakt of verzonden is op het tijdstip als aangegeven bij het record [ISO 15489].

Bestandsformaten – een bestandsformaat of file format is een code om digitale teksten, illustraties (grafische bestanden, gescande beelden, video e.d.) of geluiden op te slaan, te bewaren en te verzenden. Er bestaan vele verschillende bestandsformaten voor tekst (Office, PDF, enz.), illustraties (TIFF, PICT, EPS, BMP enz.), video (MPEG, AVI, Quicktime, enz.), en geluiden (WAVE, mp3, AIFF, enz.).

Bitstream – de ononderbroken rij van enen en nullen van een digitaal object, zonder metadata.

Born digital – van oorsprong digitaal materiaal; in tegenstelling tot gedigitaliseerd materiaal is er geen fysiek object aanwezig op basis waarvan het digitale object opnieuw gereconstrueerd zou kunnen worden.

Cloud computing – term voor gedistribueerde opslag van data op het internet; in tegenstelling tot *grid* computing is er geen sprake van centrale regie.

Datacuratie (*data curation*) – Term die vooral in de wetenschap wordt gebruikt voor het geheel aan handelingen dat nodig is om data toegankelijk en bruikbaar te houden [JISC, 2003].

Digitaal object – verzamelnaam voor alle soorten digitale bestanden: tekst, audiovisueel, simulaties, enz. enz.

Digitaal depot (ook wel: e-depot, elektronisch depot, duurzame bewaarplaats) – het geheel van organisatie, beleid, processen en procedures, financieel beheer, personeel, databeheer, databeveiliging en aanwezige hard- en software, dat het duurzaam beheren van te bewaren digitale archiefbescheiden mogelijk maakt' [ED3, Eisen Duurzaam Digitaal Depot, 2008]; zie ook OAIS.

Digitale duurzaamheid – zie *duurzame toegang*

Digitalisering – proces waarbij van een fysiek object (document, film, etc.) een digitale kopie wordt gemaakt.

Duurzame toegang (ook: digitale duurzaamheid) – algemene term voor het beheersregime waarbinnen digitale objecten duurzaam worden opgeslagen en toegankelijk gehouden opdat toekomstige generaties de objecten kunnen (her)gebruiken en erop kunnen vertrouwen dat de informatie authentiek is. Over het algemeen is de tijdshorizon langer dan 20 jaar.

Duurzaamheidstrategie – strategie om objecten duurzaam toegankelijk te houden, emulatie en migratie.

e-depot – zie digitaal depot.

Emulatie – duurzaamheidsstrategie gericht op het behoud van digitale bronnen zoals ze oorspronkelijk zijn vervaardigd en gebruikt. Door emulatietechnieken toe te passen wordt het mogelijk het gedrag van een verouderde computer na te bootsen op een andere (nieuwere) computer [ABC-DE, 2008].

Grid computing – de techniek waarbij meerdere systemen worden verenigd tot een virtueel systeem ongeacht het onderliggende hardware en/of software platform [<http://labyrinth.rienkjonker.nl/glossary>]. Het verschil met *cloud* computing is dat de betrokken organisaties zelf regie houden over het grid.

Informatieplan – plan waarin een instelling het ICT-beleid vastlegt. Hiermee is het een onmisbare schakel in de voorbereiding op digitaliseringsactiviteiten. Het informatieplan is opgebouwd uit een informatiebeleidsplan, een digitaliseringsplan en een duurzaamheidsplan [ABC-DE, 2008].

Infrastructuur – het geheel aan voorzieningen dat nodig is om duurzame toegang tot informatie te waarborgen; de NCDD onderscheid een technische infrastructuur en een organisatorische.

Institutional repository – een digitaal archief van wetenschappelijke publicaties en onderzoeksdata van een bepaalde instelling. Dergelijke archieven worden momenteel door alle Nederlandse universiteiten en door een aantal onderzoeksinstituten aangelegd. Hoewel er vaak een intentie is om informatie voor de lange termijn te bewaren, wordt die momenteel nog slechts gedeeltelijk gerealiseerd, vooral door middel van archiveringscontracten ten aanzien van publicaties met de KB.

Metadata – gestructureerde gegevens over fysieke of digitale objecten. Soorten metadata zijn: beschrijvende metadata (nodig voor het identificeren en vinden van objecten), structurele metadata (deze leggen de relatie vast tussen individuele objecten die gezamenlijk een eenheid vormen) en administratieve metadata (deze richten zich op beheer en management van objecten) [ABC-DE, 2008].

Migratie – duurzaamheidsstrategie om digitale informatie voor de lange termijn bruikbaar te houden, door het geschikt te maken voor nieuwe (versies van) software of computers. Op basis van migratietechnieken worden bestanden overgezet van een verouderd systeem naar een nieuwer, met zo veel mogelijk behoud van authenticiteit [ABC-DE, 2008].

OAIS – Reference model for an Open archival information system (2003) is een raamwerk voor termen en concepten die relevant zijn voor langetermijntoegang. Een OAIS wordt gedefinieerd als ‘an archive, consisting of an organisation of people and systems that has accepted the responsibility to preserve information and make it available for a designated community’. Belangrijk in het model zijn de zes functionele entiteiten *ingest*, *archival storage*, *data management*, *administration*, *preservation planning* en *access* [Lavoie, 2004]. Voor een Nederlandstalige beschrijving zie [ED3, 2008].

Open standaard – Een standaard waarover in een open procedure is beslist en waarvan het beheer bij een non-profit organisatie met een vrij toetredingsbeleid is ondergebracht. Open standaarden moeten zijn gepubliceerd en de kosten voor het gebruik moeten laag zijn. Ze hebben geen beperkende voorwaarden over het hergebruik [ABC-DE, 2008].

Opslagmedium – drager om data op te slaan. Bijvoorbeeld een harde schijf, CD-ROM of tape [ABC-DE, 2008]

Records continuüm – het concept dat archiefdocumenten in de tijd verschillende belangen voor verschillende belangengroepen kunnen dienen - al dan niet gelijktijdig - en daartoe voortdurend betrouwbaar archiefbeheer behoeven, ongeacht wie dat uitvoert [<http://labyrinth.rienkjonker.nl/glossary>]

Substitutie – het vervangen van archiefbescheiden door reproducties. De originele bescheiden worden vervolgens vernietigd. De reproducties nemen dus volledig de plaats in van de oorspronkelijke bescheiden. zij worden daarmee archiefbescheiden in de zin van de Archiefwet. De nieuwe informatiedragers, bijvoorbeeld microfilms en optische schijven, moeten dan ook voldoen aan de eisen

die worden gesteld in de regeling duurzaamheid archiefbescheiden
[<http://www.nationaalarchief.nl/archiefbeheer/archiefzorg/substitutie/>]

Technology watch – het volgen van technische ontwikkelingen in de wereld om tijdig te kunnen ingrijpen wanneer hardware, software en/of bestandsformaten verouderen.

Zorgdrager – zie archiefvormer.