

De digitale erfenis

Onderzoek naar het beheer van digitale archieven

Case study 1: De Arbeidsvoorziening

Rijksarchiefinspectie, 23 oktober 2003

RIJKSARCHIEFINSPECTIE

De hoofdinspecteur van de Rijksarchiefinspectie oefent ingevolge de Archiefwet 1995 toezicht uit op de naleving van deze wet door het Rijk, de zelfstandige bestuursorganen en de organen van publiekrechtelijke bedrijfsorganisatie.

De Rijksarchiefinspectie controleert de naleving van de archiefwetgeving en de kwaliteit van het archiefbeheer. Hiertoe voert de Rijksarchiefinspectie onderzoeken en inspecties uit. Daarnaast rapporteren de zorgdragers jaarlijks aan de Rijksarchiefinspectie over hun archiefbeheer.

De hoofdinspecteur rapporteert aan de betrokken overheidsorganen en, indien hiertoe aanleiding is, aan de voor het archiefwezen verantwoordelijke minister van Onderwijs, Cultuur en Wetenschappen. Jaarlijks brengt de hoofdinspecteur schriftelijk verslag uit over het toezicht op het archiefbeheer aan de minister, die het verslag vervolgens aan de Tweede Kamer overlegt.

De Rijksarchiefinspectie is in Den Haag gevestigd en als volgt bereikbaar:

Prins Willem-Alexanderhof 30

Postbus 11583

2502 AN DEN HAAG

Algemeen telefoonnummer 070-333 9133

Fax 070-333 9130

E-mail info@rijksarchiefinspectie.nl

Website www.rijksarchiefinspectie.nl

Inhoudsopgave

	Managementsamenvatting	5
1	Inleiding	7
1.1	Achtergrond van deze eerste case study	8
1.2	Werkwijze Rijksarchiefinspectie	9
1.3	Leeswijzer	9
2	Juridische context	11
3	Het beheer van digitale archieven in het verleden 15	
3.1	Organisatie	15
3.2	Afspraken, richtlijnen en procedures	15
3.3	Beheer in de praktijk	16
3.4	Samenvatting	18
4	De casus: de reddingsoperatie 2002-2003	19
4.1	Voorgeschiedenis	19
4.2	Het computermuseum	19
4.3	Het project	21
4.3.1	Analyse bestanden	21
4.3.2	Selectie bestanden	22
4.3.3	Toegankelijkheid bestanden	22
4.3.4	Te vernietigen bestanden	23
4.4	Samenvatting	24
5	De regels en de praktijk 25	
5.1	Inhoudelijke toegankelijkheid	25
5.2	Technische toegankelijkheid	26
5.3	Authenticiteit	26
5.4	Voorgeschreven bestandsformaten	26
5.5	Duurzaamheid 27	
5.6	Samenvatting	27
6	Geleerde lessen	29
	Bronnen	31
	Afkortingen	32

Managementsamenvatting

Wat betreft het beheer van digitale bestanden bij de centrale overheid is de situatie op dit moment als volgt:

- er is nog weinig ervaring met de archivering van digitale bestanden;
- er is in 2002 een nieuwe regeling in werking getreden, die speciaal van toepassing is op het beheer van blijvend te bewaren digitale bestanden, te weten de *Regeling geordende en toegankelijke staat archiefbescheiden*;
- veel overheidsorganisaties hebben behoefte aan praktische handreikingen voor het beheer van digitale bestanden.

De Rijksarchiefinspectie wilde, gelet op deze omstandigheden, een goed beeld krijgen van de problematiek rond het beheer van digitale bestanden in de praktijk. Die gelegenheid deed zich eind 2002 voor, toen de Arbeidsvoorziening in liquidatie ons vroeg mee te lopen in het traject om hun digitale erfenis veilig te stellen. De Arbeidsvoorziening stelde de Rijksarchiefinspectie in de gelegenheid om parallel aan het project voor de veiligstelling van de digitale erfenis een eigen onderzoek uit te voeren en daarover te rapporteren.

De problematiek bij de Arbeidsvoorziening bestond uit het feit dat de organisatie beschikte over omvangrijke, ongestructureerde digitale bestanden. De organisatie kon zich niet goed verantwoorden, mede omdat de hiertoe benodigde informatie niet beschikbaar was. Onduidelijk was, of de informatie wel of niet aanwezig was. De digitale bestanden zouden de benodigde verantwoordingsinformatie kunnen bevatten, maar men had geen idee hoe men dit terug kon vinden. We vinden het belangrijk te benadrukken dat deze situatie niet uniek is voor de Arbeidsvoorziening; ook tijdens inspecties bij andere overheidsorganisaties hebben we meermalen soortgelijke omstandigheden aangetroffen.

De belangrijkste boodschap uit dit rapport is dat het achteraf ordenen en selecteren van uw digitale erfenis een zeer kostbare, ingewikkelde en langdurige aangelegenheid is, die grote verantwoordingsrisico's met zich meebrengt. Zolang uw digitale informatievoorziening niet op orde is, kost het uw medewerkers bovendien dagelijks onnodig veel tijd om de juiste informatie boven tafel te krijgen. We vinden het juist daarom essentieel dat overheidsorganen nu hun digitale informatievoorziening goed regelen voor de toekomst, en niet op de ingeslagen weg blijven voortgaan door aan medewerkers over te laten wat ze met hun documenten doen. Onze volgende deelrapporten zullen hiervoor verdere handreikingen bieden. Maar ook in dit rapport komen al twee belangrijke risicofactoren naar voren:

- de manier waarop de digitale informatievoorziening georganiseerd is. Er moet een goede beheersorganisatie voor digitale bestanden zijn en het is noodzakelijk de verantwoordelijkheden op dit gebied goed te regelen;
- een wildgroei aan verschillende applicaties ter ondersteuning van het primaire proces en voor kantoorautomatisering staat een zorgvuldig beheer in de weg.

Het traject voor het veiligstellen van de digitale erfenis van de Arbeidsvoorziening bestond onder andere uit de voorbereidingen om een deel van de archieven over te brengen naar het Nationaal Archief. De Arbeidsvoorziening is één van de eerste organisaties die een omvangrijk digitaal bestand daadwerkelijk overbrengen. We verwachten dat de problemen die bij de Arbeidsvoorziening bestonden, zich ook kunnen voordoen bij de meeste andere

centrale overheidsorganen. Uit het project is enerzijds gebleken dat we zullen moeten accepteren dat een deel van de digitale erfenis van de afgelopen decennia verloren is gegaan. Anderzijds is het noodzakelijk om zoveel mogelijk bestanden te redden. Daarvoor kan dit rapport handreikingen bieden. Zo beschrijven we de wijze waarop de blijvend te bewaren bestanden van de Arbeidsvoorziening zijn afgescheiden van het te vernietigen materiaal. Bij het selecteren en bewaren van deze documenten is een werkwijze gehanteerd, waarmee alle betrokken partijen (Arbeidsvoorziening, Nationaal Archief en Rijksarchiefinspectie) in dit specifieke geval hebben kunnen instemmen. Wel denkt de Rijksarchiefinspectie dat deze werkwijze niet zonder meer kan worden overgenomen door andere overheidsorganisaties. Een aanpassing is onder andere nodig om de feitelijke selectie van de bestanden beter te laten aansluiten op de wettelijk bepaalde bewaar- en vernietigingstermijnen. In hoofdstuk 4 doen we hiervoor een aantal suggesties.

1 Inleiding

‘Onder welke naam heb ik dat document ook alweer opgeslagen, en in welke map?’

‘Wat is er met de e-mails gebeurd die onze vertrokken medewerker over dit onderwerp heeft verstuurd?’

‘Waarom is die database niet meer te raadplegen?’

Bovenstaande vragen zijn aan de orde van de dag in onze informatiemaatschappij. Het zal ook u geregeld overkomen, dat u op zoek bent naar een document dat u pas vorige week gemaakt hebt. Daarom zult u zich ook vast een voorstelling kunnen maken van de tijd die ambtenaren verliezen met het zoeken, herstellen of opnieuw maken van documenten. Een optimale digitale informatievoorziening is in onze optiek dan ook van essentieel belang voor een goed functionerende overheid.

Niet alleen de effectiviteit van overheidsorganen is in het geding bij een onzorgvuldig beheer van digitale informatie. Overheidsorganen zullen hun handelen ook minder goed kunnen verantwoorden, als zij hierbij geen gebruik kunnen maken van de e-mails die hun beleidsambtenaren over een bepaald onderwerp hebben verstuurd.

En ten slotte maakt een deel van alle e-mails, databases, Worddocumenten, et cetera, deel uit van ons cultuurhistorisch erfgoed. Voor digitale bestanden van de overheid is dat niet anders voor papieren documenten. De daarvoor geselecteerde bestanden moeten uiteindelijk worden overgebracht naar een archiefbewaarplaats, zoals het Nationaal Archief in Den Haag. Nu al lopen we het risico dat latere generaties niet meer kunnen reconstrueren op welke wijze de Nederlandse overheid in de laatste decennia van de twintigste eeuw heeft gefunctioneerd, om de eenvoudige reden dat te veel documenten uit die periode zijn weg- of zoekgeraakt.

Deze belangen alleen al zijn reden genoeg voor de Rijksarchiefinspectie om een onderzoek uit te voeren naar het beheer van digitale archieven bij de centrale overheid. Toch was er nóg een reden om dit onderzoek te starten: in 2002 is de *Regeling geordende en toegankelijke staat archiefbescheiden* in werking getreden, waarin speciale eisen zijn gesteld aan het beheer van blijvend te bewaren digitale archiefbescheiden. De Rijksarchiefinspectie wilde graag zicht krijgen op de toepassing en de toepasbaarheid van deze regeling.

Ons onderzoek valt uiteen in drie delen. Dit rapport behandelt het eerste deelonderzoek. Dit eerste onderzoek vond in 2003 plaats bij de Arbeidsvoorziening, een organisatie die op dat moment al in liquidatie was. Het zal daarom duidelijk zijn dat we bij de Arbeidsvoorziening alleen hebben kunnen kijken naar de vraag in hoeverre de in het verleden gevormde digitale bestanden achteraf nog te redden waren. De andere twee onderzoeken besteden vooral aandacht aan bestanden die in het heden en in de toekomst worden aangemaakt, en de wijze waarop vooraf voldoende waarborgen kunnen worden ingebouwd voor een duurzaam informatiebeheer.

Uitgangspunt bij de drie deelonderzoeken is, dat de beschreven casussen herkenbaar zijn voor andere overheidsorganen. Voor dit rapport geldt dan ook dat het weliswaar de situatie bij de Arbeidsvoorziening beschrijft, maar dat deze beschrijving naar verwachting een goede afspiegeling geeft van de omstandigheden bij veel overheidsorganen. Om die reden is het rapport in brede zin een hulpmiddel voor het verbeteren van de digitale informatievoorziening. Ook uw organisatie zal – net als de Arbeidsvoorziening – in meer of

mindere mate te maken hebben met ongestructureerde digitale bestanden. We geven aan wat u op dit moment nog met die bestanden kunt doen en in hoeverre u achteraf aan de wetgeving op dit gebied kunt voldoen.

1.1 Achtergrond van deze eerste case study

Het publiekrechtelijk zelfstandig bestuursorgaan (zbo) de Arbeidsvoorziening heeft dertien jaar bestaan, van 1991 tot en met 2003.¹ De taken van de Arbeidsvoorziening bestonden in grote lijnen uit de bemiddeling, scholing en reïntegratie van werklozen. Voor de uitvoering van deze taken was tal van applicaties in gebruik. Een lid van de Raad van Bestuur van de CWI gaf tijdens het symposium 'Een duurzame nalatenschap' aan dat de centra alleen al 3000 applicaties voor kantoorautomatisering (KA) erfden van de Arbeidsvoorziening. Daarnaast waren er zogenaamde corporate systemen in gebruik, dit zijn systemen die speciaal zijn ontworpen ter ondersteuning van de werkzaamheden van de Arbeidsvoorziening. Het belangrijkste corporate systeem was PGI (Primair en Gemeenschappelijk Informatiesysteem), voor de registratie van werklozen.

In het najaar van 2002 benaderde de Arbeidsvoorziening in liquidatie de Rijksarchiefinspectie. De Arbeidsvoorziening was al geruime tijd bezig met een project om zowel haar papieren als digitale documenten veilig te stellen. De organisatie had te maken met een enorme hoeveelheid ongestructureerde digitale bestanden (in totaal zo'n zestien terabyte). Het veiligstellen van deze bestanden was nodig om de volgende redenen.

- Een deel van deze bestanden moet ook in de toekomst gebruikt kunnen worden door de rechtsopvolgers van de Arbeidsvoorziening: de Centra voor Werk en Inkomen (CWI) en het ministerie van Sociale Zaken en Werkgelegenheid (SZW), deze laatste in verband met de toekenning van gelden van het Europees Sociaal Fonds (ESF).
- Een deel van de archieven komt voor blijvende bewaring in aanmerking. Voor dat deel is de overbrenging naar het Nationaal Archief in voorbereiding.

De Arbeidsvoorziening heeft de Rijksarchiefinspectie verzocht om inbreng van archivistische expertise. We waren van oordeel dat de casus bij de Arbeidsvoorziening een goede gelegenheid was om van dichtbij waar te nemen welke problemen en knelpunten zich bij de archivering van digitale bestanden voordoen. De geleerde lessen zouden we vervolgens kunnen verspreiden bij de andere organen van de centrale overheid. Dit laatste was mede mogelijk als gevolg van de grote bereidheid van de Arbeidsvoorziening om de opgedane kennis te delen. Van belang voor de beslissing van de Rijksarchiefinspectie om een rol te spelen in het project van de Arbeidsvoorziening was ook dat de situatie bij de centrale overheid er in 2002 als volgt uit zag:

- er was nog weinig ervaring met de archivering van digitale bestanden;
- er was net een nieuwe regeling in werking getreden, die speciaal van toepassing is op het beheer van blijvend te bewaren digitale bestanden, te weten de *Regeling geordende en toegankelijke staat archiefbescheiden*;
- veel overheidsorganisaties hadden behoefte aan praktische handreikingen voor het beheer van digitale bestanden.

1 Al ruim voor de definitieve opheffing van de Arbeidsvoorziening zijn delen van de organisatie afgesplitst, als gevolg van het kabinetsbesluit de Arbeidsvoorziening op te delen in verschillende organisaties. De Arbeidsvoorziening heeft zich bij deze grootscheepse reorganisatie mede gericht naar de voorschriften voor archiefbeheer bij organisatieveranderingen uit de archiefwetgeving. Op de gevolgen van de privatisering voor het archiefbeheer heeft de Rijksarchiefinspectie zich in dit onderzoek niet gericht, daar zullen we dan ook niet ingaan op dit rapport.

Deze redenen leidden tot ons besluit om een deel van het traject bij de Arbeidsvoorziening mee te lopen. Tegelijkertijd volgden we een eigen onderzoekstraject. Hierin stonden de volgende drie vragen centraal:

- Op welke wijze werd het digitaal archiefbeheer in het verleden vormgegeven en op welke wijze werd de regelgeving geïmplementeerd en gevolgd?
- Zijn de eisen die de regelgeving stelt aan het beheer van digitale bestanden achteraf toe te passen?
- Wat moeten overheidsorganen doen om de digitale problematiek in de toekomst op te lossen of te voorkomen?

Dit rapport beantwoordt deze drie vragen. Het rapport moet niet worden gezien als een inspectierapport, maar als een case study waar alle partijen uit kunnen leren: de Arbeidsvoorziening en haar rechtsopvolgers, andere organen van de centrale overheid, en de Rijksarchiefinspectie zelf.

1.2 Werkwijze Rijksarchiefinspectie

De Rijksarchiefinspectie heeft tijdens dit onderzoek gebruik gemaakt van een eigen vragenlijst over het beheer van digitale informatie. Het is de bedoeling dat deze vragenlijst gedurende de drie deelonderzoeken verder wordt uitgewerkt en verbeterd, zodat overheidsorganen deze lijst uiteindelijk kunnen gebruiken als een checklist voor het beheer van digitale archieven.

De lijst bevat vragen over zowel de organisatorische als de technische aspecten van dit beheer. Het gaat om vragen als:

- was er in het verleden iemand belast met de vernietiging van digitale bestanden, en waren er afspraken over de wijze waarop deze vernietiging moest plaatsvinden?
- welke kwaliteitscontroles vonden plaats bij conversie en migratie van bestanden?
- hoe werden Worddocumenten opgeborgen?

De volledige vragenlijst treft u aan als bijlage 1 bij dit rapport. Deze is voorzien van de antwoorden die wij gedurende dit onderzoek gekregen hebben.

We zijn in de vragenlijst ten minste uitgegaan van de eisen die de archiefwetgeving heeft gesteld. Voornaamste regeling in dit kader is de eerder genoemde *Regeling geordende en toegankelijke staat archiefbescheiden*. Op de archiefwetgeving gaan we in het volgende hoofdstuk nader in. Aanvullend hebben we gebruik gemaakt van Nederlandse en internationale standaarden voor documentenbeheer. Voor een verantwoording van de vragenlijst verwijzen we u naar bijlage 1.

1.3 Leeswijzer

Dit rapport start met een korte omschrijving van de archiefwetgeving en enkele cruciale begrippen daaruit in hoofdstuk 2. Vervolgens bespreken we in hoofdstuk 3 het beheer van digitale archieven bij de Arbeidsvoorziening in het verleden. Hiermee schetsen we de situatie zoals deze was bij aanvang van het project om de digitale gegevens van de Arbeidsvoorziening veilig te stellen. Hoofdstuk 4 beschrijft de eigenlijke reddingsoperatie. De *Regeling geordende en toegankelijke staat archiefbescheiden* schrijft een aantal zaken voor in het verleden gevormde bestanden voor. Hoofdstuk 5 gaat in op de vraag in hoeverre deze eisen achteraf nog toe te passen zijn. Hoofdstuk 6 geeft een overzicht van

geleerde lessen, waarmee u uw voordeel kunt doen bij de inrichting van uw digitale informatievoorziening en bij het veiligstellen van uw digitale erfenis.

2 Juridische context

De Arbeidsvoorziening werd bij het project om de digitale gegevens veilig te stellen geconfronteerd met de wetgeving op dit terrein. Ook de Rijksarchiefinspectie heeft, zoals in de inleiding al naar voren kwam, tijdens haar onderzoek de wetgeving als leidraad gebruikt. Daarom beschrijft dit hoofdstuk kort de geldende wetgeving op het gebied van digitaal archiefbeheer, en leggen we enkele cruciale begrippen uit de wetgeving uit. Het is belangrijk om vast te stellen dat een deel van de wetgeving waarover we in dit hoofdstuk praten, nog niet was vastgesteld op het moment dat de (digitale) bestanden van de Arbeidsvoorziening gevormd werden. In het kader van dit onderzoek gaat het dan met name om de *Regeling duurzaamheid archiefbescheiden* (vastgesteld in 2001) en de *Regeling geordende en toegankelijke staat archiefbescheiden* (vastgesteld in 2002). Deze regelingen zijn wél (gedeeltelijk) van toepassing op de in het verleden gevormde bestanden. Op deze kwestie gaan we in hoofdstuk 5 nader in.

2.1 Voorschriften voor de digitale informatievoorziening

De Arbeidsvoorziening valt vanaf het begin van haar bestaan onder de werking van de Archiefwet.² Dat betekent dat alle organisatieonderdelen onder de werking van deze wet vallen, of het nu om de centrale en regionale besturen gaat of om de arbeidsbureaus en de ondersteunende diensten. Het betekent overigens niet dat alle documenten onder de werking van de Archiefwet vallen. Privé-documenten (zoals e-mails naar familieleden en vrienden), maken geen deel uit van het archief van de Arbeidsvoorziening. Ook documenten die het persoonlijk domein van de opsteller niet verlaten hebben (zoals concepten die niet aan collega's, leidinggevendenden of derden zijn voorgelegd), maken geen deel uit van het archief van de organisatie en vallen daarom niet onder de werking van de Archiefwet. Het digitale archief van de Arbeidsvoorziening bevatte grote aantallen privé-documenten en niet-uitgewisselde documenten. Hierop komen we nog terug in hoofdstuk 4.

Voor wat betreft de documenten die wél als archiefbescheiden moeten worden aangemerkt, gelden allereerst de algemene regels die de archiefwetgeving bevat. In grote lijnen komt het erop neer dat de Arbeidsvoorziening haar (digitale) archiefbescheiden in goede, geordende en toegankelijke staat moest brengen en houden, gedurende de periode dat deze bewaard moesten blijven. Deze termijnen moeten zijn gesteld in een - conform de archiefwettelijke procedure - vastgestelde selectielijst. De Arbeidsvoorziening beschikte over een dergelijke lijst.

Verder diende de Arbeidsvoorziening te beschikken over een regeling voor het beheer van de archiefbescheiden (archiefbeheersregels) en over voldoende en deskundig personeel.

Een deel van de digitale documenten komt op grond van de selectielijst voor blijvende bewaring in aanmerking. Dat wil zeggen dat deze moeten worden overgebracht naar het Nationaal Archief. Voor deze documenten is sinds maart 2002 de eerder genoemde *Regeling geordende en toegankelijke staat archiefbescheiden* gaan gelden. Deze regeling bevat specifieke eisen over digitale documenten, bijvoorbeeld over toegankelijkheid, het

2 Eerst de *Archiefwet 1962* en het *Archiefbesluit 1968*, sinds 1 januari 1996 onder de *Archiefwet* en het *Archiefbesluit 1995*. Vanaf september 2001 zijn ook de regelingen *Duurzaamheid archiefbescheiden* en *Bouw en inrichting archiefruimten en archiefbewaarplaatsen* van toepassing. Sinds maart 2002 is de *Regeling geordende en toegankelijke staat van toepassing*.

beheer van metadata, conversie en migratie. Op deze begrippen gaan we in de volgende paragraaf kort in. Verder is in dit verband ook de *Regeling duurzaamheid archiefbescheiden* relevant. Deze stelt eisen aan de duurzaamheid van gegevensdragers.

Behalve de archiefwetgeving is tijdens dit onderzoek weer eens gebleken hoe belangrijk privacywetgeving is bij het beheren en selecteren van digitale informatie. Met name bij een organisatie zoals de Arbeidsvoorziening, die zeer veel privacygevoelige informatie beheert. Hierop zullen we in hoofdstuk 4 nader ingaan.

2.2 Cruciale begrippen

Het is onze bedoeling om in deze paragraaf kort enkele kernbegrippen te introduceren. We denken dat deze begrippen van belang zijn voor een beter inzicht in de problematiek van het beheer van digitale archieven.

Inhoudelijke toegankelijkheid – Onder inhoudelijke toegankelijkheid verstaan we dat documenten terug te vinden zijn via de mappenstructuur, via trefwoorden, onderwerpsomschrijvingen, afzenders, registratienummers, en dergelijke. De *Regeling geordende en toegankelijke staat archiefbescheiden* stelt verschillende eisen aan de inhoudelijke toegankelijkheid van digitale documenten, waaronder de eis dat een overheidsorgaan moet beschikken over een overzicht van de beheerde bestanden en (zij het impliciet) een documentair structuurplan.

Technische toegankelijkheid – Met technische toegankelijkheid bedoelen we dat documenten raadpleegbaar en leesbaar kunnen worden gemaakt door middel van de juiste apparatuur en programmatuur. De *Regeling geordende en toegankelijke staat archiefbescheiden* schrijft dan ook voor dat die programmatuur, apparatuur en documentatie hierover beschikbaar moeten blijven. Als dit niet mogelijk is, is conversie of migratie van de gegevens naar nieuwe programmatuur en / of apparatuur vereist.

Authenticiteit – Van digitale archiefbescheiden moeten we zeker weten dat ze niet veranderen. We moeten kunnen vaststellen dat we met het authentieke document te maken hebben. Authenticiteit wil zeggen dat de inhoud, structuur en vorm van de archiefbescheiden hetzelfde blijven als op het moment van ontstaan. Het is daarom ook niet verwonderlijk dat in de *Regeling geordende en toegankelijke staat archiefbescheiden* speciale eisen zijn gesteld om voldoende authenticiteit te waarborgen.

Conversie en migratie - Het overgaan op andere hardware of software kan gevolgen hebben voor de authenticiteit van documenten en dus voor de betrouwbaarheid van uw informatievoorziening. De *Regeling geordende en toegankelijke staat archiefbescheiden* heeft daarom extra eisen gesteld aan conversie en migratie. Want dat zijn de begrippen die bij het overgaan op nieuwe hardware en software horen. De regeling definieert conversie als volgt: 'het omzetten in of het overzetten van gegevens in een ander opslagformaat'. En migratie: 'het overzetten van gegevens en applicatiesoftware naar een ander platform'.

Metadata – Metadata zijn gegevens over de organisatorische en de technische omgeving van een document. Bijvoorbeeld: wie heeft het document gemaakt, wanneer, in welk werkproces, met behulp van welke software, om welke versie van het stuk gaat het? Deze gegevens moeten bij een document worden opgeslagen. Metadata waarborgen enerzijds dat het stuk later terug te vinden is en anderzijds kan met de metadata worden

aangetoond dat het document het originele, authentieke document is. De *Regeling geordende en toegankelijke staat archiefbescheiden* geeft aan welke gegevens er in elk geval bij een document moeten worden opgeslagen.

3 Het beheer van digitale archieven in het verleden

In het vorige hoofdstuk hebben we een korte beschrijving gegeven van de wet- en regelgeving die voor de Arbeidsvoorziening van toepassing was. In dit hoofdstuk beschrijven we in hoeverre de Arbeidsvoorziening in het verleden aan deze eisen voldeed. Op die manier geven we een schets van de situatie zoals deze was bij aanvang van het project om de digitale gegevens veilig te stellen. We moeten hierbij vooral uitgaan van de regelgeving die vóór de inwerkingtreding van de *Regeling geordende en toegankelijke staat archiefbescheiden* bestond, omdat deze regeling pas in werking trad in 2002. Op dat moment was de Arbeidsvoorziening al in liquidatie. We gaan in dit hoofdstuk eerst in op de vraag of de Arbeidsvoorziening beschikte over voldoende en opgeleid personeel voor het beheer van digitale archieven. Vervolgens bespreken we de beschikbaarheid van (schriftelijke) afspraken voor het beheer, waaronder archiefbeheersregels. Ten slotte behandelen we het beheer in de praktijk, waarbij aandacht wordt besteed aan de waarborgen voor de goede, geordende en toegankelijke staat van archiefbescheiden.

Overigens zal blijken dat de eisen die vanuit de archiefwetgeving aan de informatievoorziening worden gesteld, goed aansluiten op de eisen die een organisatie aan haar eigen informatievoorziening stelt. Dat is ook niet verwonderlijk, als men zich bedenkt dat in beide gevallen het basisprincipe is dat de bestanden goed toegankelijk, betrouwbaar en raadpleegbaar zijn.

3.1 Organisatie

Een eerste voorwaarde voor een goed toegankelijke, betrouwbare en raadpleegbare digitale informatievoorziening, is een professionele beheersorganisatie. Een professionele beheersorganisatie garandeert drie zaken:

- een zorgvuldig technisch beheer van de applicaties;
- een zorgvuldig beheer van de gegevens en documenten in de applicaties;
- een gewaarborgde inbreng en kwaliteitsbewaking vanuit de invalshoek van de documentaire informatievoorziening.

We constateerden bij de Arbeidsvoorziening met betrekking tot deze drie zaken het volgende:

- voor de Arbeidsvoorziening bestond een professionele beheersorganisatie voor het technisch beheer van de applicaties. Dit beheer werd verzorgd door het Facilitair Bedrijf van de Arbeidsvoorziening, Werkling;
- noch de verantwoordelijkheid voor het beheer van gegevens in de corporate systemen, noch die voor het beheer van de gegevens in de KA-omgeving was duidelijk belegd. In de praktijk lag de verantwoordelijkheid bij de gebruikers;
- de informatiespecialisten van Docuraat (het organisatieonderdeel belast met het beheer van de papieren archieven) hadden geen bemoeienis met het beheer van digitale bestanden.

3.2 Afspraken, richtlijnen en procedures

Uit bovenstaande observaties kunnen we concluderen dat (de regie voor) het beheer van digitale bestanden gedeeltelijk bij Werkling lag, gedeeltelijk bij de gebruikers en niet bij de

documentair informatiespecialisten. Juist in een dergelijke situatie, waarbij geen informatiespecialisten betrokken zijn bij het beheer, moeten goede afspraken en richtlijnen beschikbaar zijn voor het beheer van de documenten. We doelen op afspraken over onderwerpen als:

- het opslaan van documenten (moet dat op de persoonlijke schijf of op de gezamenlijke schijf, in welke map);
- het aanmaken van gezamenlijke mappen;
- het bewaren van e-mails;
- wie verantwoordelijk is voor tijdige vernietiging;
- welk type documenten moet blijven bewaard.

Zonder afspraken is op de langere termijn voor niemand meer duidelijk waar bepaalde documenten zijn opgeslagen, raken er documenten kwijt en worden bestanden te vroeg (of te laat!) vernietigd. Daarom behandelt deze paragraaf de aanwezigheid van afspraken, richtlijnen en procedures.

De verantwoordelijkheden en bevoegdheden voor de digitale informatiehuishouding, zoals beschreven in de vorige paragraaf, waren niet geregeld in archiefbeheersregels ex artikel 14 van het *Archiefbesluit 1995*, noch in een ander document. De Arbeidsvoorziening beschikte evenmin over een e-mailprotocol³. In het algemeen ontbraken richtlijnen voor de digitale informatievoorziening. De gebruikers konden dus in principe zelf bepalen wat ze met hun documenten deden. Dit geldt in mindere mate voor een aantal regio's, dat voor de gemeenschappelijke schijven een vaste structuur ontworpen had en afspraken had over het type documenten dat op de gemeenschappelijke schijven werd geplaatst. Verder werden er ook afspraken gemaakt in projectplannen bij grote conversieprojecten van de corporate systemen en in back-upprotocollen.

3.3 Beheer in de praktijk

In de vorige paragrafen hebben we moeten vaststellen dat met name het beheer van KA-documenten plaatsvond zonder de deskundige inbreng van informatieprofessionals en dat voor dit beheer geen algemene afspraken waren gemaakt. Al in de inleiding noemden we de gevolgen hiervan voor de dagelijkse praktijk: na twaalf jaar beschikte de Arbeidsvoorziening over een enorme hoeveelheid ongestructureerde informatie, die bovendien verspreid was over een wildgroei aan applicaties. Alleen al voor de kantoorautomatisering waren er 3000 applicaties in gebruik.

We moeten ervoor oppassen te concluderen dat deze ongestructureerde digitale erfenis uniek is voor de Arbeidsvoorziening. We denken dat deze situatie in de afgelopen decennia bij zeer veel overheidsorganen is ontstaan. Dit als gevolg van organisatorische en functionele scheidingen tussen de organisatieonderdelen, belast met de documentaire informatievoorziening en archiefbeheer, en de organisatieonderdelen, belast met ICT. Ook tijdens onze inspecties bij andere organisaties zien we dat er geen of nauwelijks afstemming plaatsvindt tussen de technische beheerders, de gegevensbeheerders en de informatiespecialisten. Verder hebben maar weinig gebruikers van applicaties in het verleden geleerd hoe zij om moesten gaan met het beheer van digitale documenten die zij zelf opgemaakt of ontvangen hadden.

Als we meer gedetailleerd naar de dagelijkse praktijk kijken, dan moeten we de volgende drie zaken van elkaar onderscheiden:

³ *Veel ministeries hebben tegenwoordig wel een dergelijk protocol, waarin onder andere beschreven staat op welke wijze medewerkers hun zakelijke e-mails moeten archiveren of aanbieden ter archivering.*

- het technisch beheer van de applicaties;
- het functioneel gegevensbeheer in de corporate systemen;
- het functioneel gegevensbeheer in de KA-omgeving.

Het technisch beheer van de applicaties

In de praktijk was dit onderdeel van de digitale informatievoorziening tamelijk goed op orde. Het Facilitair Bedrijf van de Arbeidsvoorziening, Werklinq, zorgde voor het ontwerp, de implementatie en het onderhoud van de applicaties. In het algemeen verliep het onderhoud, bijvoorbeeld conversie van systemen, zorgvuldig. Voorafgaand aan elke grootscheepse conversie of migratie stelde een hiertoe in het leven geroepen projectgroep een projectplan op. Helaas hebben we moeten vaststellen, dat de vereiste inbreng vanuit de documentaire informatievoorziening ontbrak bij het ontwerp, de implementatie en het onderhoud van de systemen. Gevolg hiervan was dat bij conversie alleen rekening werd gehouden met de belangen van de contemporaine gebruikers. Specifieke eisen aan de technische en inhoudelijke toegankelijkheid en authenticiteit van de bestanden waren niet in de projectplannen vastgelegd. Men controleerde bijvoorbeeld niet of alle metagegevens bij een conversie werden meegenomen. Deze metagegevens zijn van belang om achteraf de authenticiteit van gegevens te kunnen vaststellen en de toegankelijkheid te kunnen waarborgen.

Het functioneel gegevensbeheer in de corporate systemen

Het (dagelijks) beheer en de mutatie van gegevens in de corporate systemen voerden de betrokken medewerkers van de Arbeidsvoorziening zelf uit. De verantwoordelijkheid voor het functioneel applicatiebeheer van deze corporate systemen, waaronder het tijdig schonen van de gegevens op grond van de selectielijst, was niet belegd. Dit functioneel beheer in het algemeen en het tijdig schonen in het bijzonder, was in dit specifieke geval ook van belang om de eisen uit de *Wet bescherming persoonsgegevens* na te kunnen komen. Deze wet schrijft namelijk tijdige vernietiging van persoonsgegevens voor.

Het functioneel gegevensbeheer in de KA-omgeving

Worddocumenten, Powerpointpresentaties, Excel-bestanden, et cetera werden zowel opgeslagen op gemeenschappelijke als op persoonlijke schijven. De wijze waarop deze bestanden beheerd werden verschilde per regio. Niet overal was er sprake van beheer van de gemeenschappelijke schijven. Voor zover bij onze gesprekspartners bekend, lag er bij geen van de regio's een documentair structuurplan ten grondslag aan de ordening van de gemeenschappelijke schijven.

De informatie op persoonlijke schijven, waaronder e-mail, werd beheerd door de eigenaars van de persoonlijke schijven. Gezien het feit dat er geen afspraken waren over de ordening op de persoonlijke schijven, de selectie en vernietiging van de bestanden, en er een

e-mailprotocol ontbrak, betekent dit dat de gebruikers zelf bepaalden welke documenten zij voor kortere of langere termijn bewaarden. Uit onze steekproeven in de KA-omgeving blijkt dat dit per gebruiker sterk verschilt. De één bewaarde alleen werkgerelateerde e-mails, de ander alle e-mails, een derde bewaarde in het geheel geen e-mails. Ook verschilden de termijnen waarop documenten vernietigd werden van persoon tot persoon.

Noch Werklinq, noch Docuraat, was betrokken bij het inhoudelijk beheer van de documenten die in de KA-omgeving werden opgemaakt en ontvangen.

3.4 Samenvatting

Samenvattend kunnen we stellen dat de Arbeidsvoorziening onvoldoende waarborgen had geschapen om ervoor te zorgen dat de digitale bestanden in een goede, geordende en toegankelijke staat bleven gedurende de termijn waarbinnen ze bewaard moesten blijven.

De inrichting van de beheersorganisatie komt hier naar voren als een belangrijke risicofactor voor goed beheer van digitale bestanden. Met name het niet helder beleggen van verantwoordelijkheden en een onvoldoende samenwerking hebben een inadequaar archiefbeheer tot gevolg.

Belangrijk is te vermelden dat als gevolg hiervan één van de basisregels uit de archiefwetgeving niet was uitgevoerd: in de digitale omgeving was geen onderscheid gemaakt tussen te bewaren en te vernietigen documenten. Dit had grote gevolgen voor het project voor de veiligstelling van digitale gegevens, zoals we in het volgende hoofdstuk zullen zien.

4 De casus: de reddingsoperatie 2002-2003

In het voorgaande hoofdstuk hebben we kunnen lezen als gevolg van welke factoren de Arbeidsvoorziening eind 2001 beschikte over een grote, ongestructureerde digitale erfenis.

In dit hoofdstuk bespreken we het project uit 2003, waarvan het de bedoeling was om van deze erfenis de bestanden te redden die voor hetzij de rechtsopvolgers van de Arbeidsvoorziening, hetzij het Nationaal Archief van belang waren. We bespreken allereerst de redenen waarom de organisatie besloot tot een dergelijk project. Daarna besteden we aandacht aan het zogeheten 'computermuseum'. Dit is de omgeving waarin de eind 2002 aanwezige gegevens en bestanden werden veiliggesteld, met gebruikmaking van de door de Arbeidsvoorziening gebruikte apparatuur en programmatuur. Vervolgens behandelen we op thematische wijze het hele project.

4.1 Voorgeschiedenis

De Arbeidsvoorziening had voor aanvang van het project om de digitale bestanden veilig te stellen ook zelf al vastgesteld dat de kwaliteit van de informatievoorziening ernstig tekortschoot. De organisatie kon zich niet goed verantwoorden, mede omdat de hiertoe benodigde informatie niet beschikbaar was. Onduidelijk was, of de informatie wel of niet nog aanwezig was. Verder wist men niet of de benodigde informatie in het papieren archief of in de digitale omgeving terug te vinden moest zijn. De Arbeidsvoorziening beschikte, kortom, over een omvangrijk - beperkt toegankelijk - papieren archief, en over een enorm ongestructureerd digitaal archief. Deze bestanden zouden de benodigde verantwoordingsinformatie kunnen bevatten, maar men had geen idee hoe men dit terug kon vinden.

Dit besef heeft ertoe bijgedragen dat besloten is tot de inrichting van een computermuseum, kort nadat het voornemen de Arbeidsvoorziening op te heffen werd uitgesproken. In december 2002 werden de gegevens uit de KA-omgeving (inclusief e-mail) en uit de corporate systemen in 'bevoren' vorm in het computermuseum gezet. Dat was de stand van zaken op het moment dat de Rijksarchiefinspectie werd betrokken bij het project voor het veiligstellen van de digitale gegevens.

4.2 Het computermuseum

Waar hebben we het eigenlijk over, als we over de bestanden in het computermuseum spreken? Om die vraag te beantwoorden, moeten we drie zaken aan de orde stellen:

- de keuze welke systemen in het computermuseum werden opgenomen;
- het feit dat het om een 'bevriezing' gaat van de gegevens die op een bepaald moment beschikbaar waren;
- welk soort documenten in het computermuseum zijn opgeslagen.

Welke systemen in het computermuseum

De keuze welke systemen in het computermuseum zouden worden opgenomen en welke niet, is al gemaakt in een fase waarin de Rijksarchiefinspectie nog niet bij het project betrokken was (juli 2002). De keuze is gemaakt ten gunste van de systemen die de meeste 'risico-informatie' bevatten, bijvoorbeeld over de besteding van ESF-gelden. In feite is op dat moment al een selectie gemaakt tussen 'bewaren' en 'vernietigen'.

Bevriezing van gegevens

De documenten in het computermuseum vormen een momentopname: het zijn niet alle documenten die in het verleden zijn opgemaakt of ontvangen, maar alle documenten die er in december 2002 nog waren. Het feit dat het computermuseum alleen de documenten en gegevens bevat die op een bepaald moment bevroren waren, beperkt de volledigheid. Vooral bij e-mail is het in het algemeen gebruikelijk dat men alleen de meest recente stukken op de persoonlijke schijf bewaart. Uit onderzoek is gebleken, dat een deel van de medewerkers hun e-mail helemaal gewist had. Afgesloten documenten in Word, Excel, et cetera blijven gewoonlijk wel langer op persoonlijke of gemeenschappelijke schijven staan.

Maar ook de historische gegevens uit de corporate systemen zijn bijvoorbeeld niet bewaard gebleven. Dit heeft niet alleen te maken met het feit dat het computermuseum een momentopname is, maar ook met het feit dat de systemen geen archiveringsfunctie voor historische gegevens hebben.

Het soort documenten in het computermuseum

De documenten en gegevens die in de corporate systemen zijn opgeslagen, maken allemaal deel uit van het archief van de Arbeidsvoorziening in de zin van de Archiefwet. Dit wil niet zeggen dat alle systemen en de daarin opgeslagen gegevens voor eeuwig bewaard moeten blijven. Aan de hand van de selectielijst moest worden vastgesteld welke (delen van) systemen enige tijd bewaard moesten blijven en welke onmiddellijk weg konden.

De KA-bestanden en de e-mail bestonden niet alleen uit archief van de Arbeidsvoorziening. Uit een aantal beperkte steekproeven van de Rijksarchiefinspectie bleek, dat de KA- en Exchange-omgeving ook privé-documenten bevatten en documenten die nooit waren uitgewisseld. Zoals we in paragraaf 2.1 al zagen, behoren deze documenten niet tot het archief van de Arbeidsvoorziening. Deze bestanden zouden dus noch de rechtsopvolgers van de Arbeidsvoorziening, noch het Nationaal Archief, moeten bewaren. Daar komt nog bij dat met name de privé-documenten privacygevoelig materiaal bevatten, dus niet gedurende langere tijd bewaard mogen blijven en al helemaal niet in de openbaarheid mogen komen.

Overigens geldt ook voor documenten die wél tot het archief behoren dat zij vaak privacygevoelige informatie bevatten. Het gaat dan met name om gegevens over werkzoekenden. Deze documenten moeten op het hiervoor vastgestelde moment vernietigd worden, tenzij ze op basis van de selectielijst als te bewaren zijn aangemerkt. Samenvattend waren er dus verschillende typen documenten opgeslagen in het computermuseum:

- blijvend te bewaren archiefbescheiden, niet privacygevoelig en privacygevoelig;
- op termijn te vernietigen archiefbescheiden, niet privacygevoelig en privacygevoelig;
- niet uitgewisselde documenten, niet behorend tot het archief van de Arbeidsvoorziening, niet privacygevoelig en privacygevoelig;
- privédocumenten, niet privacygevoelig en privacygevoelig.

Tot slot moeten we nog opmerken dat het om documenten gaat van alle onderdelen van de Arbeidsvoorziening, dat wil zeggen het Centraal Bestuur Arbeidsbureaus, de Regionale Besturen Arbeidsbureaus, de Arbeidsbureaus en om een restcategorie

(documenten op servers voor ESF, voor de bedrijfsvoering, servers voor de Centra voor vakopleidingen, et cetera).

4.3 Het project

Doel van het project voor de veiligstelling van digitale gegevens was drieledig.

1. Een deel van de documenten zou over moeten gaan naar de CWI ('dynamisch archief').
2. Een deel van de documenten moest nog gedurende enige tijd in geordende en toegankelijke staat bewaard blijven omdat op basis van deze documenten mogelijk verantwoording zou moeten worden afgelegd. Daarnaast moesten ook de systemen die nodig waren voor de afronding van de bedrijfsvoering van Arbeidsvoorziening nog enige tijd bewaard blijven.
3. Een deel van de documenten kwam voor blijvende bewaring in aanmerking en zou moeten worden overgebracht naar het Nationaal Archief.

In feite kunnen we vaststellen dat het project alleen is doorgegaan met het derde doel. De documenten die de CWI nodig hadden voor de uitvoering van hun taken, waren al overgegaan naar de CWI. Tot het in goede, geordende en toegankelijke staat brengen en bewaren van de documenten die vanwege de bewijsvoering (en op basis van de selectielijst) nog enige tijd bewaard moesten blijven, is niet overgegaan. Voor deze documenten is gekozen voor een minimumvariant, te weten het tegen zo laag mogelijke kosten in stand houden van het computermuseum. Dit houdt in dat de documenten die nu nog aanwezig zijn slechts voor beperkte tijd bewaard zullen worden, namelijk totdat de functionaliteit van het computermuseum uitdooft. De Arbeidsvoorziening en CWI vonden de kosten voor alternatieve oplossingen te hoog.

4.3.1 Analyse bestanden

In paragraaf 4.1 hebben we beschreven welk type bestanden in het computermuseum aanwezig was. Een dergelijke inhoudelijke analyse had ook aan de start van het project moeten plaatsvinden. In werkelijkheid is binnen het project eerst gekeken naar de manier waarop de bestanden geschikt gemaakt konden worden voor overbrenging aan het Nationaal Archief.

In 2002 werd bijvoorbeeld al een onderzoek gedaan naar de mogelijkheden om de corporate systemen over te zetten naar het opslagformaat XML (Extensible Markup Language). Dit is een formaat dat is voorgeschreven in de *Regeling geordende en toegankelijke staat archiefbescheiden* voor bestanden die voor altijd te bewaren zijn. Later is echter vast komen te staan dat de corporate systemen, met uitzondering van het Intranetsysteem BRAINS, geen blijvend te bewaren gegevens bevatten. Het onderzoek had achterwege kunnen blijven als eerst was vastgesteld voor welk type bestanden een duurzame oplossing gevonden moest worden. Pas daarna had het vinden van (technische) oplossingen aan de orde moeten zijn. Het is overigens wel zo dat bij het onderzoek naar de omzetting naar het opslagformaat XML belangrijke en waardevolle ervaringsgegevens zijn verkregen over de complexiteit en aanpak van een dergelijke operatie.

Het feit dat bij aanvang van het project niet inhoudelijk is onderzocht welk soort bestanden in het computermuseum was opgeslagen, heeft gedurende het gehele project belemmerend gewerkt. Ook voor de KA-omgeving en de e-mail bleek het belangrijk om vast te stellen in hoeverre we werkelijk over archief spraken, en in hoeverre het blijvend te bewaren materiaal betrof. Pas in de laatste fase is vastgesteld welke type documenten nu precies voor preservatie in aanmerking kwam, dat wil zeggen: welke documenten nu

moeten worden aangemerkt als archief en welk deel daarvan voor blijvende bewaring in aanmerking komt.

4.3.2 Selectie bestanden

Het is overigens niet eenvoudig gebleken om vast te stellen welke documenten precies voor preservatie in aanmerking komen. Vooral niet als het om reeds gevormde, ongestructureerde bestanden gaat, hetgeen bij de Arbeidsvoorziening het geval was (wat betreft de Word-, Excel- en Exchange-bestanden). Het zou noodzakelijk zijn om elk document afzonderlijk te bestuderen en vervolgens aan een handeling uit een selectielijst te koppelen, om de selectie conform de selectielijst uit te voeren. Het zal duidelijk zijn dat deze weg niet bewandeld kon worden, vooral gezien de enorme omvang van de digitale archieven van de Arbeidsvoorziening.

Daarom is gezocht naar een methode van geautomatiseerde selectie. Men heeft daarbij wel geprobeerd om zo dicht mogelijk bij de selectie op basis van de selectielijst te blijven. In eerste instantie was het idee om de documenten te selecteren van personen, die een rol hadden gespeeld in 'belangrijke' werkprocessen, dat wil zeggen: die een rol hebben gespeeld in handelingen waarvan de neerslag op basis van de selectielijst blijvend te bewaren is. Dit bleek niet de juiste weg. Veel documenten van de betreffende personen bleken gewist of niet meer te reconstrueren. Bovendien waren veel blijvend te bewaren documenten op andere plekken in de organisatie te vinden dan bij de geselecteerde personen.

In tweede instantie is een lijst van trefwoorden opgesteld, waarmee de te bewaren documenten met behulp van een zoekmachine gefilterd zouden kunnen worden uit de totale KA- en Exchange-omgeving. Naar ons oordeel zou het wenselijk zijn geweest om de trefwoorden te baseren op onderzoek in de papieren dossiers (zie ook de volgende paragraaf). Vanwege de tijdsdruk die inmiddels op het project was komen te staan, is hier van afgezien en heeft het Nationaal Archief een lijst opgesteld met algemene trefwoorden. Daaraan toegevoegd is een 'negatieve trefwoordenlijst', dat wil zeggen: een lijst met trefwoorden die per se niet in de geselecteerde bestanden mochten voorkomen. Speciale aandacht ging daarbij uit naar privacygevoelige bestanden. Deze trefwoordenlijsten zijn toegepast en verbeterd in een pilot met de bestanden van de Regio Rijnmond van de Arbeidsvoorziening. Op de resultaten van deze pilot heeft de stuurgroep positief gereageerd. De overige bestanden zullen dus op deze wijze geselecteerd worden.

4.3.3 Toegankelijkheid bestanden

De trefwoorden die waren gekozen voor de selectie van de bestanden, waren zo algemeen dat ze niet te koppelen aan werkprocessen, en dus niet aan handelingen uit de selectielijst.

We zijn van oordeel dat nader onderzoek (bij een ander overheidsorgaan) zou moeten uitwijzen of er ook mogelijkheden zijn om de reeds gevormde bestanden wel (geautomatiseerd) achteraf aan een werkproces te koppelen. We denken daarbij aan het opstellen van een specifieke trefwoordenlijst per in de selectielijst omschreven handeling. Deze trefwoordenlijst zou tot stand moeten komen door intensief dossieronderzoek. Een dergelijke wijze van selecteren zou direct de basis voor de ordening van geselecteerde bestanden leggen. Hiervan kon bij de Arbeidsvoorziening geen sprake zijn, omdat de trefwoorden te algemeen waren en niet gekoppeld waren aan werkprocessen. Dat betekent dat aan de geselecteerde documenten geen metadata konden worden toegevoegd over het werkproces. De bestanden van de Arbeidsvoorziening krijgen de volgende metadata mee:

- archiefbloknnummer (aangeleverd door het Nationaal Archief);

- naam archief: 'Digitaal archief Arbeidsvoorziening, periode 1994-2001';
- omgeving: Kantoorautomatisering;
- naam van de regio / concern;
- servernaam;
- mappenstructuur waarin het bestand is opgeslagen;
- oorspronkelijk en nieuw formaat van het gearchiveerde bestand;
- oorspronkelijke bestandsnaam;
- datum laatste wijziging;
- gebruikte conversieprogrammatuur;
- conversie uitgevoerd door;
- conversie uitgevoerd op datum / tijd.

We kunnen op grond van bovenstaande lijst vaststellen dat de metadata bij de verduurzaamde documenten nauwelijks inhoudelijke informatie bevatten. De metadata die ontbreken zijn in feite cruciaal voor een optimale toegankelijkheid van het geselecteerde digitale archief. Om een goede toegankelijkheid te realiseren zouden in elk geval de volgende gegevens aan de documenten moeten worden toegevoegd (eventueel op een hoger aggregatieniveau):

- behandelaar dossier;
- classificatienummer;⁴
- handeling;
- werkproces.

Deze en andere metadata moeten ook op grond van de *Regeling geordende en toegankelijke staat* aan de bestanden zijn toegevoegd. Tenzij een goede methode zou worden ontwikkeld om documenten te koppelen aan classificatienummer, handeling en werkproces, is dit alleen te realiseren door genoemde metadata handmatig aan documenten toe te kennen, of de documenten handmatig te ordenen in mappen. Bij de Arbeidsvoorziening is niet voor deze weg gekozen. Gezien de omvang van het totale bestand zou dit ook alleen tegen zeer hoge kosten gerealiseerd kunnen worden. De bestanden zullen nu nader toegankelijk worden gemaakt met een zoekmachine, hetgeen niet binnen de kaders van de archiefwetgeving valt. Op grond van de Archiefwet moeten de bestanden wel de juiste metadata meekrijgen en goed gestructureerd worden. Dit is nodig om documenten in de juiste context te kunnen plaatsen, wat bij toegankelijkheid door middel van een zoekmachine niet mogelijk is.

4.3.4 Te vernietigen bestanden

Op het moment dat we dit rapport schrijven, is nog niet duidelijk op welke wijze de overdracht aan de CWI en SZW van documenten die op termijn te vernietigen zijn, zal plaatsvinden. Wel is duidelijk dat voor het gehele computermuseum een onderhoudscontract wordt afgesloten. Het gaat om een minimumcontract, waarbij applicaties die niet meer door de leveranciers onderhouden worden 'uitdoven'. Andere beheershandelingen worden niet verricht. Naar ons oordeel is er onvoldoende aandacht besteed aan deze bestanden. Mocht één van de rechtsopvolgers van de Arbeidsvoorziening documenten uit het computermuseum nodig hebben om zich te kunnen verantwoorden, staan zij nog altijd voor dezelfde vraag als bij aanvang van het project: Deze bestanden zouden de benodigde verantwoordingsinformatie kunnen

⁴ Het classificatienummer maakt onderdeel uit van een classificatieschema. Binnen een classificatieschema worden de onderlinge relaties gedefinieerd tussen digitale archiefdocumenten en de daarbij behorende aggregaten.
NEN-ISO 15489-1

bevatten, maar hoe kunnen we die informatie terugvinden? Ook blijft het computermuseum bestanden bevatten die op basis van de selectielijst en de privacywetgeving vernietigd hadden moeten worden.

4.4 Samenvatting

Uit dit hoofdstuk blijkt duidelijk dat de toegankelijkheid en selectie van digitale bestanden al moet worden geregeld bij het ontstaan en ontvangen van documenten. Het realiseren van toegankelijkheid en selectie achteraf is niet of alleen tegen zeer hoge kosten mogelijk. Verder hebben we gezien dat veel informatie verloren gaat, als de digitale informatievoorziening vooraf niet goed wordt ingericht. Als gevolg hiervan kan verantwoording in een latere fase lastig of zelfs onmogelijk worden.

We kunnen verder stellen dat de Arbeidsvoorziening zich veel moeite heeft getroost om in elk geval de documenten veilig te stellen die uit cultuurhistorisch perspectief van belang zijn. Het is niet gelukt om deze documenten op basis van de selectielijst te selecteren. Ook op andere punten heeft men de archiefwetgeving niet na kunnen leven. Eén oorzaak is naar ons oordeel dat een dergelijk project nooit eerder is uitgevoerd, waardoor geëxperimenteerd moest worden met mogelijke oplossingen. Een andere oorzaak is dat een alternatief traject langer zou hebben geduurd en meer kosten met zich mee zouden hebben gebracht. Hier is niet voor gekozen. Tenslotte hebben we moeten vaststellen dat de archiefwetgeving niet zonder meer op een digitale erfenis is toe te passen. Hierop gaan we in het volgende hoofdstuk nader in.

5 De regels en de praktijk

Eén van de centrale vragen van dit onderzoek is of de eisen uit de archiefwetgeving aan het beheer van digitale bestanden zijn toe te passen op bestanden die reeds gevormd zijn. Bij het stellen van deze vraag gingen we er eigenlijk vanuit dat vooral toepassing van de *Regeling geordende en toegankelijke staat* problemen zou opleveren. Bij de afronding van dit onderzoek moeten we echter vaststellen dat de problemen al beginnen bij de hoofdeisen die de archiefwetgeving én een verantwoorde bedrijfsvoering stellen: namelijk dat de documenten in een goede, geordende en toegankelijke staat beschikbaar zijn. Met name voor wat betreft de KA- en de Exchangeomgeving hebben we al eerder in dit rapport vast moeten stellen dat deze niet zodanig waren ingericht, dat stukken binnen een redelijke termijn zouden kunnen worden geraadpleegd. Mede als gevolg hiervan is selectie een enorm knelpunt gebleken. Eigenlijk was selectie van de ongestructureerde bestanden niet goed meer mogelijk. De uiteindelijk gekozen selectiemethode was geen methode op basis van de selectielijst, de wettelijk verplichte grondslag.

In deze paragraaf gaan we met name in op de toepasbaarheid van de *Regeling geordende en toegankelijke staat archiefbescheiden* en de *Regeling duurzaamheid archiefbescheiden*.

De *Regeling geordende en toegankelijke staat archiefbescheiden* kent een overgangsregime voor oudere bestanden. We zijn er tijdens dit onderzoek echter vanuit gegaan dat alleen bestanden, die sinds 1996 niet meer zijn gewijzigd, onder een lichter regime zouden vallen. Voor een verantwoording van deze keuze verwijzen we naar bijlage 2. In de praktijk bleek overigens ook dit onderscheid niet opportuun of verlichtend te werken, als er geen duidelijk beheersregime voor alle bestanden bestaat.

5.1 Inhoudelijke toegankelijkheid

Overzicht en structuurplan

Niet voor niets stelt de *Regeling geordende en toegankelijke staat* dat een overzicht van archiefbescheiden moet zijn geordend overeenkomstig het ten tijde van de vorming van het archief geldende structuurplan. In een papieren omgeving is het al niet eenvoudig om achteraf een overzicht te maken van alle documenten, als de documenten ongestructureerd op stapels liggen. In dat geval geldt dat elk document afzonderlijk moet worden bekeken, en dat er achteraf een structuur moet worden vastgelegd. In het vorige hoofdstuk zagen we dat dat voor digitale documenten niet anders is. We merkten op dat de Arbeidsvoorziening de documenten eigenlijk alleen conform de eisen zou kunnen ordenen, als elk document afzonderlijk werd geopend en geanalyseerd. Het achteraf (handmatig) opzetten van een structuur van digitale bestanden en het maken van een overzicht is naar ons oordeel alleen te realiseren als de omvang van de digitale bestanden beperkt is. Organisaties die in het verleden hun digitale archief al hadden ingericht conform een structuurplan kunnen natuurlijk wel aan de eis voldoen.

Samenvattend kunnen we dus stellen dat het erg lastig en kostbaar is om in het verleden gevormde, ongestructureerde bestanden alsnog te ordenen aan de hand van een structuurplan.

Metadata

In paragraaf 4.3.4 hebben we al moeten vaststellen dat de set metadata die aan elk document moet zijn gekoppeld, bij reeds gevormde documenten alleen handmatig kan

worden gecompleteerd. Hierbij gelden dus dezelfde problemen als bij het achteraf invoeren van een structuurplan: men kan alleen aan de eisen uit de regeling voldoen als de organisatie in het verleden al de juiste metadata toekende, als het slechts om een beperkt digitaal archiefbestand gaat of als er veel (financiële) middelen beschikbaar worden gesteld.

5.2 Technische toegankelijkheid

De technische toegankelijkheid van met name de corporate systemen kan worden gewaarborgd door hetzij de oude systemen te onderhouden, hetzij door conversie of migratie. In beide gevallen is het noodzakelijk dat men zowel de beschikking heeft over de juiste programmatuur als apparatuur, als over documentatie die aangeeft hoe de programmatuur heeft gewerkt. Dergelijke eisen zijn dan ook gesteld in de *Regeling geordende en toegankelijke staat archiefbescheiden*.

Zulke informatie kan staan in een functioneel ontwerp. Bij de Arbeidsvoorziening was duidelijk dat de benodigde documentatie gedeeltelijk ontbrak. Binnen het project heeft dit niet direct tot problemen geleid, omdat de corporate systemen niet waren geselecteerd voor overbrenging naar het Nationaal Archief. Vanzelfsprekend kan in andere gevallen het ontbreken van de documentatie wel tot problemen leiden. Als documentatie in het verleden niet is bewaard, kan deze fout in het heden nog maar moeilijk worden hersteld.

5.3 Authenticiteit

Als (digitale) documenten veranderen, bijvoorbeeld als gevolg van een onzorgvuldig voorbereide conversie, dan is de authenticiteit (en dus de bewijskracht) van deze documenten aangetast. Daarom is het belangrijk om – voorafgaand aan een conversie – vast te stellen welke aspecten van een document absoluut niet mogen veranderen. De *Regeling geordende en toegankelijke staat archiefbescheiden* schrijft dan ook voor dat elke organisatie functionele eisen opstelt ten aanzien van de inhoud, structuur en vorm voor de documenten die men beheert.

Natuurlijk kan elke organisatie op elk gewenst moment deze functionele eisen vastleggen. Vanaf dat moment kan bij elke conversie en migratie gecontroleerd worden, of de inhoud, structuur en vorm van de documenten – gemeten aan de eisen die de organisatie daar zelf aan heeft gesteld – ongeoorloofd gewijzigd wordt. Echter: als deze eisen nu worden opgesteld, garandeert dat niet dat de authenticiteit van documenten in het verleden ongewijzigd is gebleven. Kortom: op dit moment is niet meer vast te stellen of de authenticiteit is aangetast van documenten die in het verleden zijn gemaakt.

Dit neemt niet weg dat het noodzakelijk is om functionele eisen op te stellen, zodat in elk geval bij toekomstige conversies en migraties zorgvuldig wordt nagegaan of de documenten niet ongeoorloofd wijzigen. We vinden het dan ook een tekortkoming dat de Arbeidsvoorziening aan de vooravond van de verduurzaming naar PDF- en XML-formaat dergelijke eisen niet heeft opgesteld. Binnen dit project hebben we niet kunnen vaststellen in hoeverre de oorspronkelijke documenten en de verduurzaamde documenten van elkaar afwijken.

5.4 Voorgeschreven bestandsformaten

Ondanks het feit dat voorafgaand aan de conversie naar nieuwe bestandsformaten geen functionele eisen zijn opgesteld, waardoor de authenticiteit van de verduurzaamde bestanden niet kan worden getoetst, merken we toch op dat deze conversie op het eerste gezicht geen grote problemen leek op te leveren.

5.5 Duurzaamheid

De Arbeidsvoorziening heeft een poging ondernomen om optische schijven aan te schaffen die voldoen aan de eisen uit de *Regeling duurzaamheid archiefbescheiden*. In de praktijk bleek dit een probleem, omdat er vrijwel geen leveranciers zijn die de voorgeschreven producten aanbieden. De RAI zal dit knelpunt, en overigens ook de andere in dit hoofdstuk genoemde knelpunten en ervaringen, doorgeven aan de evaluatiecommissie voor de drie ministeriële regelingen (waaronder de *Regeling duurzaamheid archiefbescheiden*).

5.6 Samenvatting

De *Regeling geordende en toegankelijke staat archiefbescheiden* is slechts in beperkte mate en tegen zeer hoge kosten toepasbaar op omvangrijke, in het verleden gevormde, ongestructureerde bestanden. Verder hebben we gezien dat het overgangsregime uit de *Regeling geordende en toegankelijke staat archiefbescheiden* in de praktijk niet verlichtend werkt, als er geen duidelijk beheersregime voor alle bestanden geldt.

6 Geleerde lessen

Algemeen

1. Uit het project is gebleken hoe belangrijk het samenspel tussen ICT'ers en documentair informatiespecialisten is. In hoofdstuk 3 beschreven we dat er tussen Werkling en Docuraat van Arbeidsvoorziening nauwelijks interactie was. Binnen het project 'veiligstelling digitale gegevens' werden de samenwerking en het begrip tussen de verschillende professies steeds beter. De archiefexperts kregen steeds meer inzicht in de technische (on-)mogelijkheden; de ICT'ers kregen meer oog voor het belang van een inhoudelijke benadering. Hoewel de Rijksarchiefinspectie al geruime tijd intensieve samenwerking tussen beide professies propageert, hebben we binnen dit project aan den lijve kunnen ondervinden welke verschillen in benadering van de digitale informatievoorziening tussen beide beroepsgroepen bestaan. Tevens werd duidelijk welke winst het oplevert als zij met elkaar in overleg treden.

Geleerde lessen over het omgaan met de digitale erfenis

2. Het is belangrijk dat overheidsorganen eerst vaststellen voor welk type bestanden een oplossing gevonden moet worden, alvorens over te gaan tot het vinden van (technische) oplossingen voor verduurzaming. Dat betekent dat eerst moet worden vastgesteld of het om archiefbescheiden gaat of ook om andere documenten; en of het om te bewaren of te vernietigen bestanden gaat.
3. De resultaten van de selectie op basis van trefwoorden achteraf waren redelijk positief. We denken dan ook dat selectie op basis van trefwoorden in combinatie met een zoekmachine goed kan werken voor de selectie van in het verleden gevormde, grote, ongestructureerde bestanden. Wel denken we dat voor het formuleren en bijstellen van de trefwoorden ruim de tijd moet worden genomen, en dat dit zodanig moet worden uitgevoerd dat de trefwoorden terug te voeren zijn op handelingen uit de selectielijst. Voor kleinere ongestructureerde bestanden denken we dat het de voorkeur verdient om de documenten eerst te ordenen, en vervolgens te selecteren.
4. De *Regeling geordende en toegankelijke staat archiefbescheiden* is slechts in beperkte mate toepasbaar op omvangrijke, in het verleden gevormde, ongestructureerde bestanden. De toepasbaarheid op in het verleden gevormde kleine bestanden en op in het verleden gevormde gestructureerde bestanden verdient nader onderzoek. De Rijksarchiefinspectie richt zich in de volgende twee deelonderzoeken naar het beheer van digitale archieven op in het heden en de toekomst te vormen bestanden.
5. Bij het opruimen moet niet alleen aandacht worden besteed aan te bewaren bestanden, maar ook aan te vernietigen bestanden. Deze kunnen nog van belang zijn vanwege de bedrijfsvoering of verantwoording. Overheidsorganen moeten zich ook bewust zijn van het feit dat voor deze documenten een vastgestelde vernietigingstermijn geldt.
6. Overheidsorganen moeten rekening houden met het feit dat het zorgvuldig opruimen van de digitale erfenis een kostbare aangelegenheid is, zowel in tijd als in geld. Het niet goed regelen van de toegankelijkheid, selectie en duurzaamheid van digitale bestanden vooraf, leidt tot zeer hoge kosten achteraf.

Geleerde lessen over toekomstig beheer van digitale archieven

7. Het is belangrijk dat overheidsinstellingen prioriteit geven aan het inrichten van een goede beheersorganisatie voor digitale bestanden en aan het regelen van verantwoordelijkheden op dit gebied.
8. We benadrukken het belang van de invoering van een structuurplan voor de documenten die in de toekomst gevormd worden. Voor organisaties die gaan werken met een record management systeem is het vanzelfsprekend dat men vóór de implementatie nadenkt over de (logische) structuur. Maar ook voor organisaties die hun (KA-)bestanden niet met behulp van een record management systeem beheren, is het noodzakelijk een structuurplan op te stellen waarin onder meer de mappenstructuur is weergegeven. In zo'n situatie is het eveneens van belang om afspraken vast te leggen over het type stukken dat op de gemeenschappelijke schijf moet komen te staan en op welke plaats.
9. Selectie van digitale archiefbestanden aan de bron verdient de absolute voorkeur boven selectie achteraf. Selectie aan de bron is alleen goed mogelijk als organisaties een record management systeem invoeren.
10. Het is zeer belangrijk om met name de documentatie over systemen die specifiek voor uw bedrijfsprocessen zijn ontworpen, zorgvuldig te bewaren.

Bronnen

- *Archiefbesluit 1995*
- Archiefschool, *Softwarespecificaties voor Records Management Applicaties voor de Nederlandse Overheid* (versie 3.0, maart 2002)
- *Archiefwet 1995*
- Cap Gemini, Ernst en Young, diverse rapportages en plannen van aanpak over de preservering van digitale systemen van de Arbeidsvoorziening (2002)
- Cap Gemini, Ernst en Young, diverse voortgangsrapportages (2003)
- Het Expertise Centrum, *FTO Verduurzaming KA-omgeving van Arbeidsvoorziening i.l. Eindrapport* (augustus 2003) en eerdere versies (2003)
- Horsman, P., 'Entiteiten en attributen. Metadatamodel' (versie 2.0), in: *Functionele eisen voor archiveringssystemen* (februari 2003)
- KB en Partners, *Preservering digitaal Organisatiegeheugen. Adviesrapport* (2001)
- *Regeling duurzaamheid archiefbescheiden* (2001)
- *Regeling geordende en toegankelijke staat archiefbescheiden* (2002)
- Sanders, R., 'Werklijn in de etalage', in: *Computable Nieuws*, www.computable.nl (26 oktober 2001)
- Selectielijst van de Arbeidsvoorzieningsorganisatie 1991-1997

Afkortingen

CWI	Centra voor Werk en Inkomen
ESF	Europees Sociaal Fonds
ICT	Informatie- en Communicatie Technologie
KA	Kantoorautomatisering
PDF	Portable Document
PGI	Primair en Gemeenschappelijk Informatiesysteem
SZW	Ministerie van Sociale Zaken en Werkgelegenheid
XML	Extensible Markup Language
Zbo	Zelfstandig bestuursorgaan